EV826334317US

PTO/SB/05 (07-06)
Approved for use through 01/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office. U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	PA3909US
First Inventor	Carlos Avendano
Title	System and Method for Utilizing Omni-Directional Microphones For Speech Enhancement
Express Mail Label No.	EV 826 334 317 US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. ☑ **Fee Transmittal Form** (e.g., PTO/SB/17) *(Submit an original and a duplicate for fee processing)*
2. ☑ **Applicant claims small entity status.** See 37 CFR 1.27.
3. ☑ **Specification** [*Total Pages* 26] Both the claims and abstract must start on a new page *(For information on the preferred arrangement, see MPEP 608.01(a))*
4. ☑ **Drawing(s)** (*35 U.S.C. 113*) [*Total Sheets* 9]
5. **Oath or Declaration** [*Total Sheets* 3]
   a. ☑ Newly executed (original or copy)
   b. ☐ A copy from a prior application (37 CFR 1.63(d)) *(for continuation/divisional with Box 18 completed)*
      i. ☐ DELETION OF INVENTOR(S) Signed statement attached deleting inventor(s) name in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).
6. ☐ **Application Data Sheet.** See 37 CFR 1.76
7. ☐ **CD-ROM or CD-R** in duplicate, large table or Computer Program *(Appendix)* ☐ Landscape Table on CD
8. **Nucleotide and/or Amino Acid Sequence Submission** *(if applicable, items a. – c. are required)*
   a. ☐ Computer Readable Form (CRF)
   b. Specification Sequence Listing on:
      i. ☐ CD-ROM or CD-R (2 copies); or
      ii. ☐ Paper
   c. ☐ Statements verifying identity of above copies

ADDRESS TO:

Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

ACCOMPANYING APPLICATION PARTS

9. ☑ **Assignment Papers** (cover sheet & document(s))
   Name of Assignee Audience, Inc.
10. ☐ **37 CFR 3.73(b) Statement** *(when there is an assignee)* ☑ **Power of Attorney**
11. ☐ **English Translation Document** *(if applicable)*
12. ☐ **Information Disclosure Statement** (PTO/SB/08 or PTO-1449) ☐ Copies of citations attached
13. ☐ **Preliminary Amendment**
14. ☑ **Return Receipt Postcard** (MPEP 503) *(Should be specifically itemized)*
15. ☐ **Certified Copy of Priority Document(s)** *(if foreign priority is claimed)*
16. ☑ **Nonpublication Request** under 35 U.S.C. 122(b)(2)(B)(i). Applicant must attach form PTO/SB/35 or equivalent.
17. ☑ Other: Check for $690.00

18. If a CONTINUING APPLICATION, *check appropriate box, and supply the requisite information below and in the first sentence of the specification following the title, or in an Application Data Sheet under 37 CFR 1.76:*

☐ Continuation ☐ Divisional ☒ Continuation-in-part (CIP) of prior application No.: 11/343,524

Prior application information: *Examiner* Unknown *Art Unit:* 2644

19. CORRESPONDENCE ADDRESS

☑ The address associated with Customer Number: 22830 ***OR*** ☐ Correspondence address below

Name					
Address					
City		State		Zip Code	
Country		Telephone		Email	

Signature	Susan Yee	Date	January 29, 2007
Name (Print/Type)	Susan Yee	Registration No. (Attorney/Agent)	41,388

This collection of information is required by 37 CFR 1.53(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**
If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Sony v. Jawbone
U.S. Patent No. 8,321,213
Sony Ex. 1007

012907 1769

PTO/SB/17 (01-06)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number

Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).

FEE TRANSMITTAL For FY 2006

U.S. PTO

☑ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 690.00

Complete if Known	
Application Number	Unknown
Filing Date	January 26, 2007
First Named Inventor	Carlos Avendano
Examiner Name	Unknown
Art Unit	Unknown
Attorney Docket No.	PA3909US

METHOD OF PAYMENT (check all that apply)

☑ Check ☐ Credit Card ☐ Money Order ☐ None ☐ Other (please identify): _____

☑ Deposit Account Deposit Account Number: 06-0600 Deposit Account Name: Carr & Ferrell LLP

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

☐ Charge fee(s) indicated below ☐ Charge fee(s) indicated below, **except for the filing fee**

☑ Charge any additional fee(s) or underpayments of fee(s) under 37 CFR 1.16 and 1.17 ☑ Credit any overpayments

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

FEE CALCULATION (All the fees below are due upon filing or may be subject to a surcharge.)

1. BASIC FILING, SEARCH, AND EXAMINATION FEES

	FILING FEES		SEARCH FEES		EXAMINATION FEES		
		Small Entity		Small Entity		Small Entity	
Application Type	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fees Paid ($)
Utility	300	150	500	250	200	100	500
Design	200	100	100	50	130	65	
Plant	200	100	300	150	160	80	
Reissue	300	150	500	250	600	300	
Provisional	200	100	0	0	0	0	

2. EXCESS CLAIM FEES

Fee Description	Fee ($)	Small Entity Fee ($)
Each claim over 20 (including Reissues)	50	25
Each independent claim over 3 (including Reissues)	200	100
Multiple dependent claims	360	180

Total Claims		Extra Claims	Fee ($)		Fee Paid ($)
26	- 20 or HP =	6	x 25	=	150

HP = highest number of total claims paid for, if greater than 20.

Indep. Claims		Extra Claims	Fee ($)		Fee Paid ($)
3	- 3 or HP =	0	x 100	=	0

HP = highest number of independent claims paid for, if greater than 3.

Multiple Dependent Claims	
Fee ($)	Fee Paid ($)

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets		Extra Sheets		Number of each additional 50 or fraction thereof	Fee ($)		Fee Paid ($)
	- 100 =		/ 50 =	(round **up** to a whole number)	x	=	

4. OTHER FEE(S)

	Fees Paid ($)
Non-English Specification, $130 fee (no small entity discount)	
Other (e.g., late filing surcharge): Assignment	40

SUBMITTED BY

Signature	Susan Yee	Registration No. (Attorney/Agent) 41,388	Telephone 650.812.3400
Name (Print/Type)	Susan Yee		Date January 29, 2007

This collection of information is required by 37 CFR 1.136. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT

By: Carlos Avendano

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The present application claims the priority benefit of U.S. Provisional Patent Application No. 60/850,928, filed October 10, 2006, and entitled "Array Processing Technique for Producing Long-Range ILD Cues with Omni-Directional Microphone Pair;" the present application is also a continuation-in-part of U.S. Patent Application No. 11/343,524, and entitled "System and Method for Utilizing Inter-Microphone Level Differences for Speech Enhancement," both of which are herein incorporated by reference.

BACKGROUND OF THE INVENTION

Field of Invention

[0002] The present invention relates generally to audio processing and more particularly to speech enhancement using inter-microphone level differences.

Description of Related Art

[0003] Currently, there are many methods for reducing background noise and enhancing speech in an adverse environment. One such method is to use two or more microphones on an audio device. These microphones are in prescribed positions and allow the audio device to determine a level difference between the microphone signals. For example, due to a space difference between the microphones, the difference in times of arrival of the signals from a speech source to the microphones may be utilized to localize the speech source. Once localized, the signals can be spatially filtered to suppress the noise originating from the different directions.

[0004] In order to take advantage of the level difference between two omni-directional microphones, a speech source needs to be closer to one of the microphones. That is, in order to obtain a significant level difference, a distance from the source to a first microphone needs to be shorter than a distance from the source to a second microphone. As such, a speech source must remain in relative closeness to the microphones, especially if the microphones are in close proximity as may be required by mobile telephony applications.

[0005] A solution to the distance constraint may be obtained by using directional microphones. Using directional microphones allow a user to extend an effective level difference between the two microphones over a larger range with a narrow inter-level difference (ILD) beam. This may be desirable for applications such as push-to-talk (PTT) or videophones where a speech source is not in as close a proximity to the microphones, as for example, a telephone application.

[0006] Disadvantageously, directional microphones have numerous physical drawbacks. Typically, directional microphones are large in size and do not fit well in small telephones or cellular phones. Additionally, directional microphones are difficult to mount as they required ports in order for sounds to arrive from a plurality of directions. Slight variations in manufacturing may result in a mismatch, resulting in more expensive manufacturing and production costs.

[0007] Therefore, it is desirable to utilize the characteristics of directional microphones in a speech enhancement system, without the disadvantages of using directional microphones, themselves.

SUMMARY OF THE INVENTION

[0008] Embodiments of the present invention overcome or substantially alleviate prior problems associated with noise suppression and speech enhancement. In general, systems and methods for utilizing inter-microphone level differences (ILD) to attenuate noise and enhance speech are provided. In exemplary embodiments, the ILD is based on energy level differences of a pair of omni-directional microphones.

[0009] Exemplary embodiments of the present invention use a non-linear process to combine components of the acoustic signals from the pair of omni-directional microphones in order to obtain the ILD. In exemplary embodiments, a primary acoustic signal is received by a primary microphone, and a secondary acoustic signal is received by a secondary microphone (e.g., omni-directional microphones). The primary and secondary acoustic signals are converted into primary and secondary electric signals for processing.

[0010] A differential microphone array (DMA) module processes the primary and secondary electric signals to determine a cardioid primary signal and a cardioid secondary signal. In exemplary embodiments, the primary and secondary electric signals are delayed by a delay node. The cardioid primary signal is then determined by taking a difference between the primary electric signal and the delayed secondary electric signal, while the cardioid secondary signal is determined by taking a difference between the secondary electric signal and the delayed primary electric signal. In various embodiments the delayed primary electric signal and the delayed secondary electric signal are adjusted by a gain. The gain may be a ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

[0011] The cardioid signals are filtered through a frequency analysis module which takes the signals and mimics the frequency analysis of the cochlea (i.e., cochlear domain) simulated in this embodiment by a filter bank. Alternatively, other filters such as short-time Fourier transform (STFT), sub-band filter banks, modulated complex

lapped transforms, cochlear models, wavelets, etc. can be used for the frequency analysis and synthesis. Energy levels associated with the cardioid primary signal and the cardioid secondary signals are then computed (e.g., as power estimates) and the results are processed by an ILD module using a non-linear combination to obtain the ILD. In exemplary embodiments, the non-linear combination comprises dividing the power estimate associated with the cardioid primary signal by the power estimate associated with the cardioid secondary signal. The ILD may then be used as a spatial discrimination cue in a noise reduction system to suppress unwanted sound sources and enhance the speech.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1a and FIG. 1b are diagrams of two environments in which embodiments of the present invention may be practiced.

[0013] FIG. 2 is a block diagram of an exemplary audio device implementing embodiments of the present invention.

[0014] FIG. 3 is a block diagram of an exemplary audio processing engine.

[0015] FIG. 4a illustrates an exemplary implementation of the DMA module, frequency analysis module, energy module, and the ILD module.

[0016] FIG. 4b is an exemplary implementation of the DMA module.

[0017] FIG. 5 is a block diagram of an alternative embodiment of the present invention.

[0018] FIG. 6 is a polar plot of a front-to-back cardioid directivity pattern and ILD diagram produced according to embodiments of the present invention.

[0019] FIG. 7 is a flowchart of an exemplary method for utilizing ILD of omni-directional microphones for speech enhancement.

[0020] FIG. 8 is a flowchart of an exemplary noise reduction process.

DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0021] The present invention provides exemplary systems and methods for utilizing inter-microphone level differences (ILD) of at least two microphones to identify frequency regions dominated by speech in order to enhance speech and attenuate background noise and far-field distracters. Embodiments of the present invention may be practiced on any audio device that is configured to receive sound such as, but not limited to, cellular phones, phone handsets, headsets, and conferencing systems. Advantageously, exemplary embodiments are configured to provide improved noise suppression on small devices and in applications where the main audio source is far from the device. While some embodiments of the present invention will be described in reference to operation on a cellular phone, the present invention may be practiced on any audio device.

[0022] Referring to FIG. 1a and FIG. 1b, environments in which embodiments of the present invention may be practiced are shown. A user provides an audio (speech) source 102 to an audio device 104. The exemplary audio device 104 comprises two microphones: a primary microphone 106 relative to the audio source 102 and a secondary microphone 108 located a distance, d, away from the primary microphone 106. In exemplary embodiments, the microphones 106 and 108 are omni-directional microphones.

[0023] While the microphones 106 and 108 receive sound (i.e., acoustic signals) from the audio source 102, the microphones 106 and 108 also pick up noise 110. Although the noise 110 is shown coming from a single location in FIG. 1a and FIG. 1b, the noise 110 may comprise any sounds from one or more locations different than the audio source 102, and may include reverberations and echoes.

[0024] Embodiments of the present invention exploit level differences (e.g., energy differences) between the acoustic signals received by the two microphones 106 and 108 independent of how the level differences are obtained. In FIG. 1a, because the

primary microphone 106 is much closer to the audio source 102 than the secondary microphone 108, the intensity level is higher for the primary microphone 106 resulting in a larger energy level during a speech/voice segment, for example. In FIG. 1b, because directional response of the primary microphone 106 is highest in the direction of the audio source 102 and directional response of the secondary microphone 108 is lower in the direction of the audio source 102, the level difference is highest in the direction of the audio source 102 and lower elsewhere.

[0025] The level difference may then be used to discriminate speech and noise in the time-frequency domain. Further embodiments may use a combination of energy level differences and time delays to discriminate speech. Based on binaural cue decoding, speech signal extraction, or speech enhancement may be performed.

[0026] Referring now to FIG. 2, the exemplary audio device 104 is shown in more detail. In exemplary embodiments, the audio device 104 is an audio receiving device that comprises a processor 202, the primary microphone 106, the secondary microphone 108, an audio processing engine 204, and an output device 206. The audio device 104 may comprise further components necessary for audio device 104 operations. The audio processing engine 204 will be discussed in more details in connection with FIG. 3.

[0027] As previously discussed, the primary and secondary microphones 106 and 108, respectively, are spaced a distance apart in order to allow for an energy level differences between them. Upon reception by the microphones 106 and 108, the acoustic signals are converted into electric signals (i.e., a primary electric signal and a secondary electric signal). The electric signals may themselves be converted by an analog-to-digital converter (not shown) into digital signals for processing in accordance with some embodiments. In order to differentiate the acoustic signals, the acoustic signal received by the primary microphone 106 is herein referred to as the primary acoustic signal, while the acoustic signal received by the secondary microphone 108 is herein referred to as the secondary acoustic signal.

[0028] The output device 206 is any device which provides an audio output to the user. For example, the output device 206 may be an earpiece of a headset or handset, or a speaker on a conferencing device.

[0029] FIG. 3 is a detailed block diagram of the exemplary audio processing engine 204, according to one embodiment of the present invention. In exemplary embodiments, the audio processing engine 204 is embodied within a memory device. In operation, the acoustic signals (i.e., X_1 and X_2) received from the primary and secondary microphones 106 and 108 are converted to electric signals and processed through a differential microphone array (DMA) module 302. The DMA module 302 is configured to use DMA theory to create directional patterns for the close-spaced microphones 106 and 108. The DMA module 302 may determine sounds and signals in a front and back cardioid region about the audio device 104 by delaying and subtracting the acoustic signals captured by the microphones 106 and 108. Signals (i.e., sounds) received from these cardioid regions are hereinafter referred to as cardioid signals. In one example, sounds from a sound source 102 within the cardioid region are transmitted by the primary microphone 106 as a cardioid primary signal. Sounds from the same sound source 102 are transmitted by the secondary microphone 108 as a cardioid secondary signal.

[0030] For a two-microphone system, the DMA module 302 can create two different directional patterns about the audio device 104. Each directional pattern is a region about the audio device 104 in which sounds generated by an audio source 102 within the region may be received by the microphones 106 and 108 with little attenuation. Sounds generated by audio sources 102 outside of the directional pattern may be attenuated.

[0031] In one example, one directional pattern created by the DMA module 302 allows sounds generated from an audio source 102 within a front cardioid region around the audio device 104 to be received, and a second pattern allows sounds from a

second audio source 102 within a back cardioid region around the audio device 104 to be received. Sounds from audio sources 102 beyond these regions may also be received but the sounds may be attenuated.

[0032] The cardioid signals from the DMA module 302 are then processed by a frequency analysis module 304. In one embodiment the frequency analysis module 304 takes the cardioid signals and mimics the frequency analysis of the cochlea (i.e., cochlear domain) simulated by a filter bank. In one example, the frequency analysis module 304 separates the cardioid signals into frequency bands. Alternatively, other filters such as short-time Fourier transform (STFT), sub-band filter banks, modulated complex lapped transforms, cochlear models, wavelets, etc. can be used for the frequency analysis and synthesis. Because most sounds (e.g., acoustic signals) are complex and comprise more than one frequency, a sub-band analysis on the acoustic signal determines what individual frequencies are present in the complex acoustic signal during a frame (e.g., a predetermined period of time). In one embodiment, the frame is 8 ms long.

[0033] Once the frequencies are determined, the signals are forwarded to an energy module 306 which computes energy level estimates during an interval of time (i.e., power estimates). The power estimate may be based on bandwidth of the cochlea channel and the cardioid signal. The power estimates are then used by the inter-microphone level difference (ILD) module 308 to determine the ILD.

[0034] In various embodiments, the DMA module 302 sends the cardiod signals to the energy module 306. The energy module 306 computes the power estimates prior to the analysis of the cardiod signals by the frequency analysis module 304.

[0035] Referring to FIG. 4a, one implementation of the DMA module 302, frequency analysis module 304, energy module 306, and the ILD module 308 is provided. In this implementation, the acoustic signals received by the microphones 106 and 108 are processed by the DMA module 302. The exemplary DMA module 302

delays the primary acoustic signal, X_1, via a delay node 402, $z^{-\tau 1}$. Similarly, the DMA module 302 delays the secondary acoustic signal, X_2, via a second delay node 40, $z^{-\tau 2}$.

[0036] In exemplary embodiments, a cardioid primary signal (C_f) is mathematically determined in the frequency domain (Z transform) as

$$C_f = X_1 - z^{-\tau 1} gX_2$$

while the cardioid secondary signal (C_b) is mathematically determined as

$$C_b = gX_2 - z^{-\tau 2} X_1.$$

The gain factor, g, is computed by the gain module 406 to equalize the signal levels. Prior art systems can suffer loss of performance when the microphone signals have different levels. The gain module is further discussed herein.

[0037] In various embodiments, the cardioid signals can be processed through the frequency analysis module 304. The filter coefficient may be applied to each microphone signal. As a result, the output of the frequency analysis module 304 may comprise a filtered cardioid primary signal, $\alpha C_f(t,\omega)$ and a filtered cardioid secondary signal, $\beta C_f(t,\omega)$, where t represents the time index (t=0,1,...N) and ω represents the frequency index (ω=0,1,...K).

[0038] The energy module 306 takes the signals from the frequency analysis module 304 and calculates the power estimates associated with the cardioid primary signal (C_f) and the cardioid secondary signal(C_b). In exemplary embodiments, the power estimates may be mathematically determined by squaring and integrating an absolute value of the output of the frequency analysis module 304. Power estimates of the signals from the cardioid primary signal and the cardioid secondary signal are referred to herein as components. For example, the energy level associated with the primary microphone signal may be determined by

$$E_f(t,\omega) = \int_{\text{frame}} \left| C_f(t',\omega) \right|^2 dt',$$

and the energy level associated with the secondary microphone signal may be determined by

$$E_b(t,\omega) = \int_{\text{frame}} \left| C_b(t',\omega) \right|^2 dt'.$$

[0039] Given the calculated energy levels, the ILD may be determined by the ILD module 308. In exemplary embodiments, the ILD is determined in a non-linear manner by taking a ratio of the energy levels, such as

$$\text{ILD}(t, \omega) = E_f(t,\omega) / E_b(t,\omega)$$

Applying the determined energy levels to this ILD equations results in

$$ILD(t,\omega) = \frac{\int_{\text{frame}} \left| C_f(t',\omega) \right|^2 dt'}{\int_{\text{frame}} \left| C_b(t',\omega) \right|^2 dt'}.$$

[0040] By nonlinearly combining the energy level (i.e., component) of the cardioid primary signal with the energy level (i.e., component) of the cardioid secondary signal , sounds from audio sources 102 within a front-to-back cardioid region (depicted in FIG. 6) about the audio device 104 may be effectively received. The spatial extent over which the signal can be retrieved can be specified and controlled by the ILD region selected. In contrast, if the cardioid primary signal and the cardioid secondary signal are combined linearly (e.g., the signals are subtracted,) sounds from audio sources 102 within a hypercardioid region may be effectively received. The hypercardioid region may be larger (broader) than the front-to-back cardioid ILD region selected, thus the non-linear combination via ILD can produce a narrower and more spatially selective beam.

[0041] Once the ILD is determined, the signals are processed through a noise reduction system 310. Referring back to FIG. 3, in exemplary embodiments, the noise

reduction system 310 comprises a noise estimate module 312, a filter module 314, a filter smoothing module 316, a masking module 318, and a frequency synthesis module 320.

[0042] According to an exemplary embodiment of the present invention, a Wiener filter is used to suppress noise/enhance speech. In order to derive the Wiener filter estimate, however, specific inputs are needed. These inputs comprise a power spectral density of noise and a power spectral density of the primary acoustic signal.

[0043] In exemplary embodiments, the noise estimate is based only on the acoustic signal from the primary microphone 106. The exemplary noise estimate module 312 is a component which can be approximated mathematically by

$$N(t,\omega) = \lambda_I(t,\omega)E_1(t,\omega) + (1-\lambda_I(t,\omega))\min[N(t-1,\omega), E_1(t,\omega)]$$

according to one embodiment of the present invention. As shown, the noise estimate in this embodiment is based on minimum statistics of a current energy estimate of the primary acoustic signal, $E_1(t,\omega)$ and a noise estimate of a previous time frame, $N(t\text{-}1,\omega)$. As a result, the noise estimation is performed efficiently and with low latency.

[0044] $\lambda_I(t,\omega)$ in the above equation is derived from the ILD approximated by the ILD module 308, as

$$\lambda_I(t,\omega) = \begin{cases} \approx 0 & \text{if} \quad ILD(t,\omega) < threshold \\ \approx 1 & \text{if} \quad ILD(t,\omega) > threshold \end{cases}$$

That is, when at the primary microphone 106 is smaller than a threshold value (e.g., *threshold* = 0.5) above which speech is expected to be, λ_I is small, and thus the noise estimator follows the noise closely. When ILD starts to rise (e.g., because speech is present within the large ILD region), λ_I increases. As a result, the noise estimate module 312 slows down the noise estimation process and the speech energy does not contribute significantly to the final noise estimate. Therefore, exemplary embodiments of the present invention may use a combination of minimum statistics and voice activity detection to determine the noise estimate.

[0045] A filter module 314 then derives a filter estimate based on the noise estimate. In one embodiment, the filter is a Wiener filter. Alternative embodiments may contemplate other filters. Accordingly, the Wiener filter may be approximated, according to one embodiment, as

$$W = \left(\frac{P_s}{P_s + P_n} \right)^{\varphi},$$

where P_s is a power spectral density of speech and P_n is a power spectral density of noise. According to one embodiment, P_n is the noise estimate, $N(t,\omega)$, which is calculated by the noise estimate module 312. In an exemplary embodiment, $P_s = E_1(t,\omega) - \gamma N(t,\omega)$, where $E_1(t,\omega)$ is the energy estimate associated with the primary acoustic signal (e.g., the cardioid primary signal) calculated by the energy module 306, and $N(t,\omega)$ is the noise estimate provided by the noise estimate module 312. Because the noise estimate changes with each frame, the filter estimate will also change with each frame.

[0046] γ is an over-subtraction term which is a function of the ILD. γ compensates bias of minimum statistics of the noise estimate module 312 and forms a perceptual weighting. Because time constants are different, the bias will be different between portions of pure noise and portions of noise and speech. Therefore, in some embodiments, compensation for this bias may be necessary. In exemplary embodiments, γ is determined empirically (e.g., 2-3 dB at a large ILD, and is 6-9 dB at a low ILD).

[0047] φ in the above exemplary Wiener filter equation is a factor which further limits the noise estimate. φ can be any positive value. In one embodiment, nonlinear expansion may be obtained by setting φ to 2. According to exemplary embodiments, φ is determined empirically and applied when a body of $W = \left(\frac{P_s}{P_s + P_n} \right)$ falls below a

prescribed value (e.g., 12 dB down from the maximum possible value of W, which is unity).

[0048] Because the Wiener filter estimation may change quickly (e.g., from one frame to the next frame) and noise and speech estimates can vary greatly between each frame, application of the Wiener filter estimate, as is, may result in artifacts (e.g., discontinuities, blips, transients, etc.). Therefore, an optional filter smoothing module 316 is provided to smooth the Wiener filter estimate applied to the acoustic signals as a function of time. In one embodiment, the filter smoothing module 316 may be mathematically approximated as

$$M(t,\omega) = \lambda_s(t,\omega)W(t,\omega) + (1-\lambda_s(t,\omega))M(t-1,\omega)$$

, where λ_s is a function of the Wiener filter estimate and the primary microphone energy, E_1.

[0049] As shown, the filter smoothing module 316, at time (t) will smooth the Wiener filter estimate using the values of the smoothed Wiener filter estimate from the previous frame at time (t-1). In order to allow for quick response to the acoustic signal changing quickly, the filter smoothing module 316 performs less smoothing on quick changing signals, and more smoothing on slower changing signals. This is accomplished by varying the value of λ_s according to a weighed first order derivative of E_1 with respect to time. If the first order derivative is large and the energy change is large, then λ_s is set to a large value. If the derivative is small then λ_s is set to a smaller value.

[0050] After smoothing by the filter smoothing module 316, the primary acoustic signal is multiplied by the smoothed Wiener filter estimate to estimate the speech. In the above Wiener filter embodiment, the speech estimate is approximated by $S(t,\omega) = C_f(t,\omega) * M(t,\omega)$, where $C_f(t,\omega)$ is the cardioid primary signal. In exemplary embodiments, the speech estimation occurs in the masking module 318.

[0051] Next, the speech estimate is converted back into time domain from the cochlea domain. The conversion comprises taking the speech estimate, $S(t,\omega)$, and adding together the phase shifted signals of the cochlea channels in a frequency synthesis module 320. Once conversion is completed, the signal is output to the user.

[0052] It should be noted that the system architecture of the audio processing engine 204 of FIG. 3 is exemplary. Alternative embodiments may comprise more components, less components, or equivalent components and still be within the scope of embodiments of the present invention. Various modules of the audio processing engine 204 may be combined into a single module. For example, the functionalities of the frequency analysis module 304 and energy module 306 may be combined into a single module. Furthermore, the functions of the ILD module 308 may be combined with the functions of the energy module 306 alone, or in combination with the frequency analysis module 304. As a further example, the functionality of the filter module 314 may be combined with the functionality of the filter smoothing module 316.

[0053] Referring now to FIG. 4b, a practical implementation of the DMA module 302 according to one embodiment of the present invention. In exemplary embodiments, microphone differences are compensated by using a filter 412, F(z), that equalizes the microphones 106 and 108. Since the filter 412 is a non-causal filter, in some embodiments, a delay is applied to the primary microphone signal with a delay node 414, D(z). The application of the delay node 414 results in an alignment of the two channels.

[0054] To implement a fractional delay, allpass filters 416 and 418 (e.g., $A_1(z)$ and $A_2(z)$) are applied to the signals. However, the application of the allpass filters 416 and 418 introduces a delay. As a result, two more delay nodes 420 and 422 (e.g., $D_1(z)$ and $D_2(z)$) are required.

[0055] A secondary acoustic signal magnitude may be modified to match a magnitude of the primary acoustic signal by applying a gain which is computed by the gain module 406. The gain module 406 computes the magnitude of both signals (e.g., X_1 and X_2) and derives the gain, g, as the ratio between the magnitude of the primary acoustic signal to the magnitude of the secondary acoustic signal. The gain can then be used to calculate the cardioid primary signal and the cardioid secondary signal [Notice the change I made to the figure CA].

[0056] Since the allpass filters 416 and 418 produce a desired fractional delay up to one-half the Nyquist frequency, the processing is applied at twice the system sampling rate.

[0057] As a result, a sampling rate conversion (SRC) node 424 and 426 is provided. The outputs of the SRC nodes 424 and 426 are the cardioid primary and cardioid secondary signals, C_f and C_b.

[0058] FIG. 5 is a block diagram of an alternative embodiment of the present invention. In this embodiment, the acoustic signals from the microphones 106 and 108 are processed by a frequency analysis module 304 prior to processing by a DMA module 302. According to the present embodiment, the frequency analysis module 304 takes the acoustic signals (i.e., X_1 and X_2) and mimics a cochlea implementation using a filter bank, such as a fast Fourier transform. Alternatively, other filters such as short-time Fourier transform (STFT), sub-band filter banks, modulated complex lapped transforms, cochlear models, wavelets, etc. can be used for the frequency analysis and synthesis. The output of the frequency analysis module 304 may comprise a plurality of signals (e.g., one per sub-band or tap.)

[0059] The secondary acoustic signal magnitude is modified to match the magnitude of the primary acoustic signal by computing the magnitude of both signals and deriving the gain, g, as the ratio between the magnitude of the primary acoustic

signal to the magnitude of the secondary acoustic signal. Subsequently, the signals may be processed through the DMA module 302. In the present embodiment, phase shifting of the signals (e.g., using $e^{j\omega\tau_f}$) is utilized to achieve a fractional delay of the signals.

[0060] The remainder of the process through the energy module 306 and the ILD module 308 is similar to the process described in connection with FIG. 4a, but on a per sub-band or tap basis.

[0061] FIG. 6 is a polar plot of a front-to-back cardioid directivity pattern 602 and ILD diagram produced according to exemplary embodiments of the present invention. The cardioid directivity pattern 602 illustrates a range in which the acoustic signals may be received. As shown, by using the non-linear combination process and delay lines (e.g., 420 and 422), the range of the cardioid directivity pattern 602 may be extended in the forward and backward directions (i.e., along the x-axis). The extension in the forward and backward directions allows significant ILD cues to be obtained from acoustic sources further away from the microphones 106 and 108. As a result, the omni-directional microphones 106 and 108 can achieve acoustic characteristics that mimic those of directional microphones.

[0062] Referring now to FIG. 7, a flowchart of an exemplary method for utilizing ILD of omni-direction microphones for noise suppression and speech enhancement is shown. In step 702, acoustic signals are received by the primary microphone 106 and the secondary microphone 108. In exemplary embodiments, the microphones are omni-directional microphones. In some embodiments, the acoustic signals are converted by the microphones to electronic signals (i.e., the primary electric signal and the secondary electric signal) for processing.

[0063] Differential array analysis is then performed on the acoustic signals by the DMA module 302. In exemplary embodiments, the DMA module 302 is configured to determine the cardioid primary signal and the cardioid secondary signal by delaying,

subtracting, and applying a gain factor to the acoustic signals captured by the microphones 106 and 108. Specifically, the DMA module 302 determines the cardioid primary signal by taking a difference between the primary electric signal and a delayed secondary electric signal. Similarly, the DMA module 302 determines thecardioid secondary signal by taking a difference between the secondary electric signal and a delay primary electric signal.

[0064] In step 706, the frequency analysis module 304 performs frequency analysis on the cardioid primary and secondary signals. According to one embodiment, the frequency analysis module 304 utilizes a filter bank to determine individual frequencies present in the complex cardioid primary and secondary signals.

[0065] In step 708, energy estimates for the cardioid primary and secondary signals are computed. In one embodiment, the energy estimates are determined by the energy module 306. The exemplary energy module 306 utilizes a present cardioid signal and a previously calculated energy estimate to determine the present energy estimate of the present cardioid signal.

[0066] Once the energy estimates are calculated, inter-microphone level differences (ILD) are computed in step 710. In one embodiment, the ILD is calculated based on a non-linear combination of the energy estimates of the cardioid primary and secondary signals. In exemplary embodiments, the ILD is computed by the ILD module 308.

[0067] Once the ILD is determined, the cardioid primary and secondary signals are processed through a noise reduction system in step 712. Step 712 will be discussed in more detail in connection with FIG. 8. The result of the noise reduction processing is then output to the user in step 714. In some embodiments, the electronic signals are converted to analog signals for output. The output may be via a speaker, earpieces, or other similar devices.

[0068] Referring now to FIG. 8, a flowchart of the exemplary noise reduction process (step 712) is provided. Based on the calculated ILD, noise is estimated in step 802. According to embodiments of the present invention, the noise estimate is based only on the acoustic signal received at the primary microphone 106. The noise estimate may be based on the present energy estimate of the acoustic signal from the primary microphone 106 and a previously computed noise estimate. In determining the noise estimate, the noise estimation is frozen or slowed down when the ILD increases, according to exemplary embodiments of the present invention.

[0069] In step 804, a filter estimate is computed by the filter module 314. In one embodiment, the filter used in the audio processing engine 208 is a Wiener filter. Once the filter estimate is determined, the filter estimate may be smoothed in step 806. Smoothing prevents fast fluctuations which may create audio artifacts. The smoothed filter estimate is applied to the acoustic signal from the primary microphone 106 in step 808 to generate a speech estimate.

[0070] In step 810, the speech estimate is converted back to the time domain. Exemplary conversion techniques apply an inverse frequency of the cochlea channel to the speech estimate. Once the speech estimate is converted, the audio signal may now be output to the user.

[0071] The above-described modules can be comprises of instructions that are stored on storage media. The instructions can be retrieved and executed by the processor 202. Some examples of instructions include software, program code, and firmware. Some examples of storage media comprise memory devices and integrated circuits. The instructions are operational when executed by the processor 202 to direct the processor 202 to operate in accordance with embodiments of the present invention. Those skilled in the art are familiar with instructions, processor(s), and storage media.

[0072] The present invention is described above with reference to exemplary embodiments. It will be apparent to those skilled in the art that various modifications

may be made and other embodiments can be used without departing from the broader scope of the present invention. Therefore, these and other variations upon the exemplary embodiments are intended to be covered by the present invention.

CLAIMS

1. A system for enhancing speech, comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal;

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal; and

an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

2. The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level-equalized secondary electric signal.

3. The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain.

4. The system of claim 3 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

5. The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the level-equalized secondary electric signal and a delayed primary electric signal.

6. The system of claim 1 further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal.

7. The system of claim 1 further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

8. The system of claim 1 further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference.

9. The system of claim 1 further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

10. The system of claim 9 further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

11. The system of claim 1 further comprising a masking module configured to determine a speech estimate.

12. The system of claim 11 further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output.

13. The system of claim 1, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal.

14. A method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; and

non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

15. The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between the primary electric signal and a delayed secondary electric signal.

16. The method of claim 14 wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain.

17. The method of claim 16 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

18. The method of claim 14 wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal.

19. The method of claim 14 wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal.

20. The method of claim 14 further comprising determining an energy estimate for each of the acoustic signals during a frame.

21. The method of claim 14 further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference.

22. The method of claim 21 further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

23. The method of claim 22 further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal

24. The method of claim 22 further comprising smoothing the filter estimate.

25. The method of claim 14 wherein the cardioid primary signal and the cardioid secondary signal is of a sub-band of the primary electric signal.

26. A machine readable medium having embodied thereon a program, the program providing instructions for a method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; and

non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference.

SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT

ABSTRACT

[0073] Systems and methods for utilizing inter-microphone level differences (ILD) to attenuate noise and enhance speech are provided. In exemplary embodiments, primary and secondary acoustic signals are received by omni-directional microphones, and converted into primary and secondary electric signals. A differential microphone array module processes the electric signals to determine a cardioid primary signal and a cardioid secondary signal. The cardioid signals are filtered through a frequency analysis module which takes the signals and mimics a cochlea implementation (i.e., cochlear domain). Energy levels of the signals are then computed, and the results are processed by an ILD module using a non-linear combination to obtain the ILD. In exemplary embodiments, the non-linear combination comprises dividing the energy level associated with the primary microphone by the energy level associated with the secondary microphone. The ILD is utilized by a noise reduction system to enhance the speech of the primary acoustic signal.

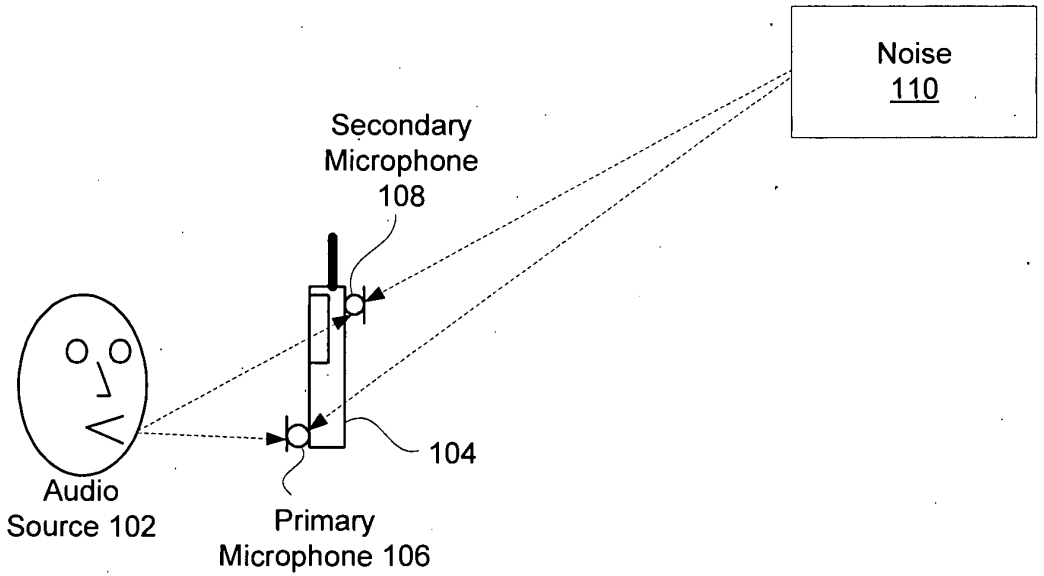

Noise
110
Secondary
Microphone
108
104
Audio
Source 102
Primary
Microphone 106

FIG. 1a

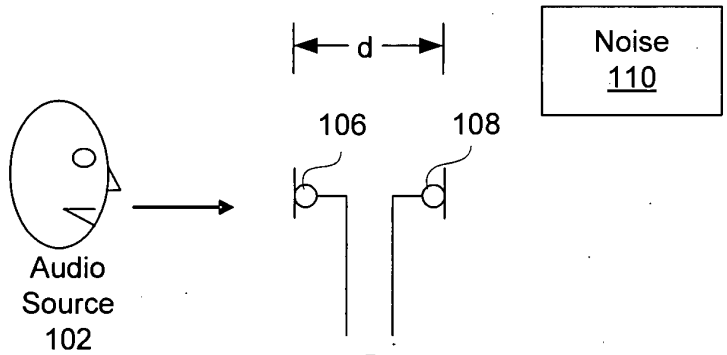

d
Noise
110
106
108
Audio
Source
102

FIG. 1b

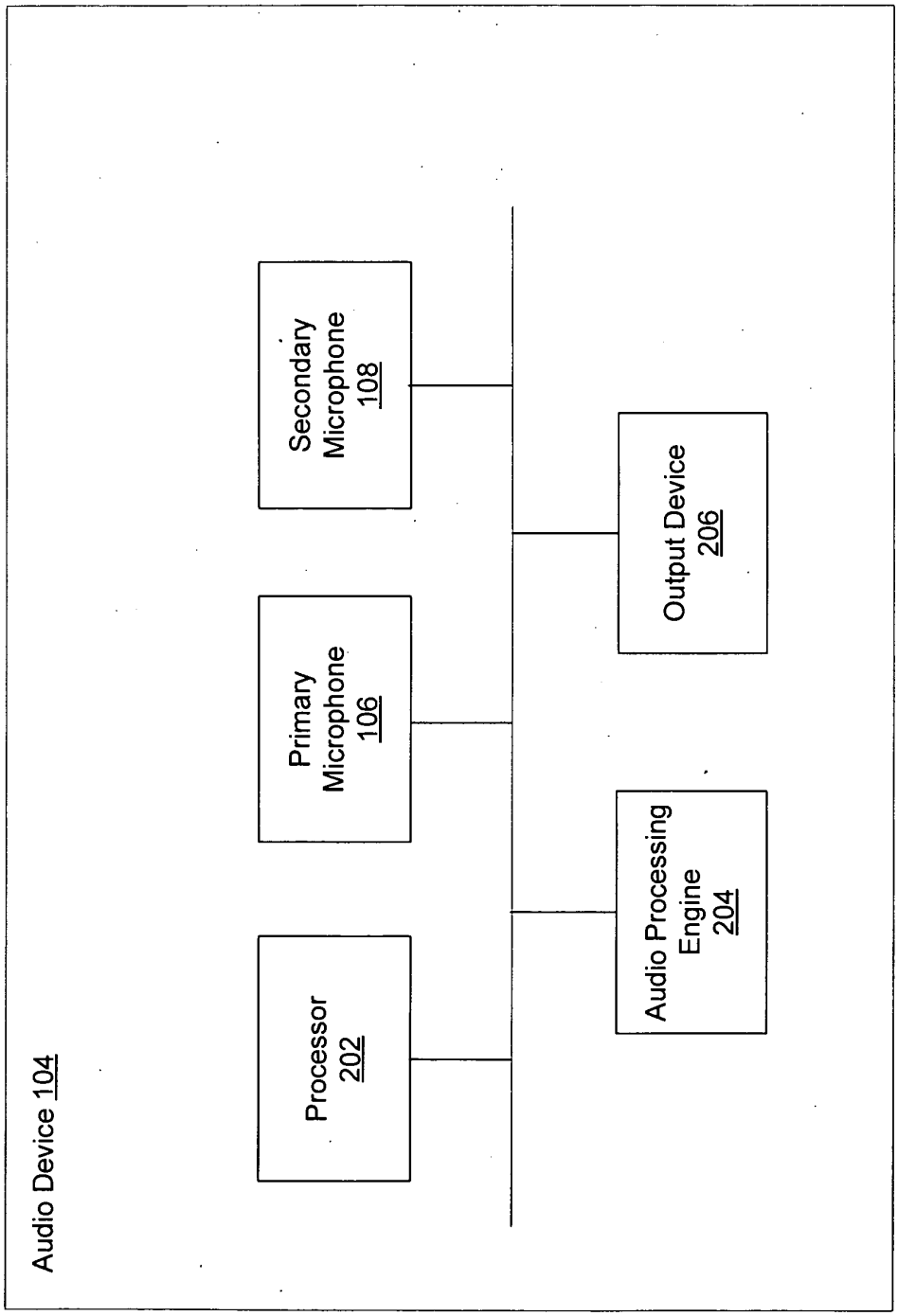

Audio Device 104
Processor 202
Primary Microphone 106
Secondary Microphone 108
Audio Processing Engine 204
Output Device 206

FIG. 2

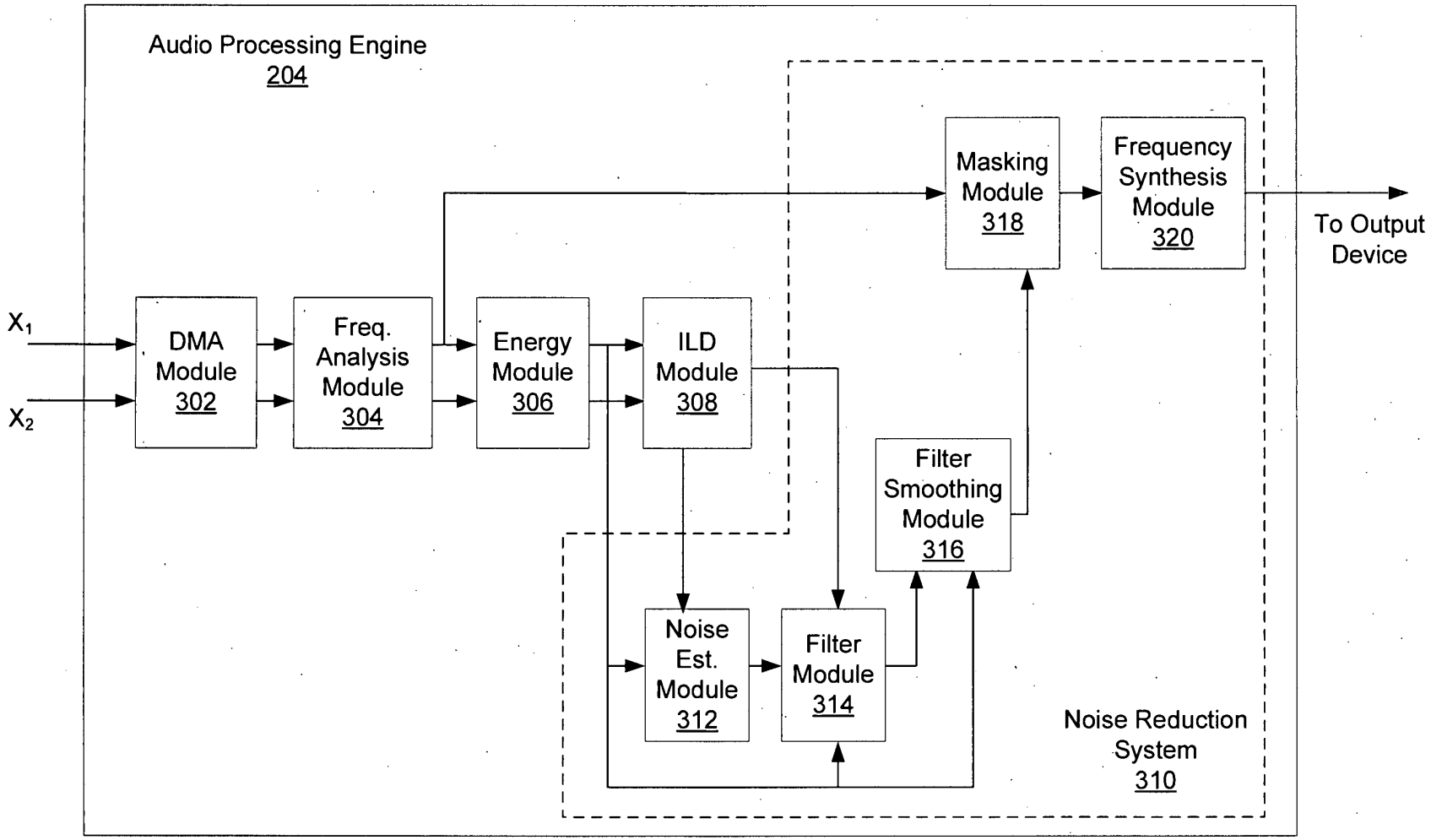

Audio Processing Engine
204
X_1
X_2
DMA Module 302
Freq. Analysis Module 304
Energy Module 306
ILD Module 308
Masking Module 318
Frequency Synthesis Module 320
To Output Device
Filter Smoothing Module 316
Noise Est. Module 312
Filter Module 314
Noise Reduction System 310

FIG. 3

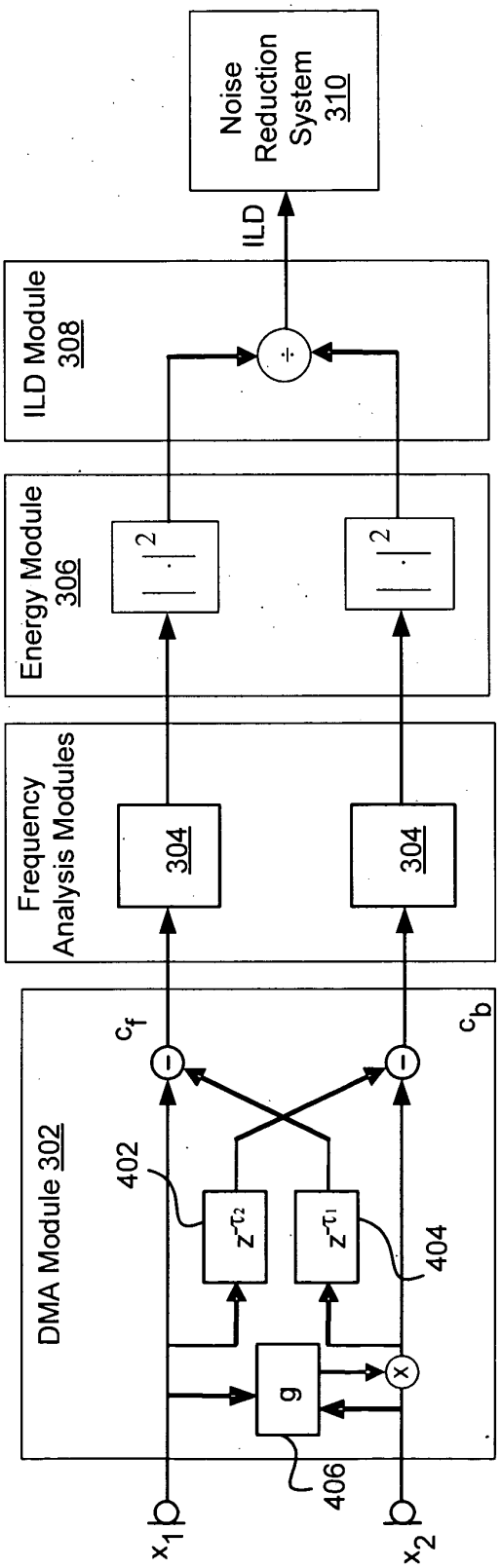

Noise Reduction System 310
ILD
ILD Module 308
Energy Module 306
Frequency Analysis Modules
304
304
DMA Module 302
c_f
c_b
402
404
406
$z^{-\tau_2}$
$z^{-\tau_1}$
g
x_1
x_2

FIG. 4a

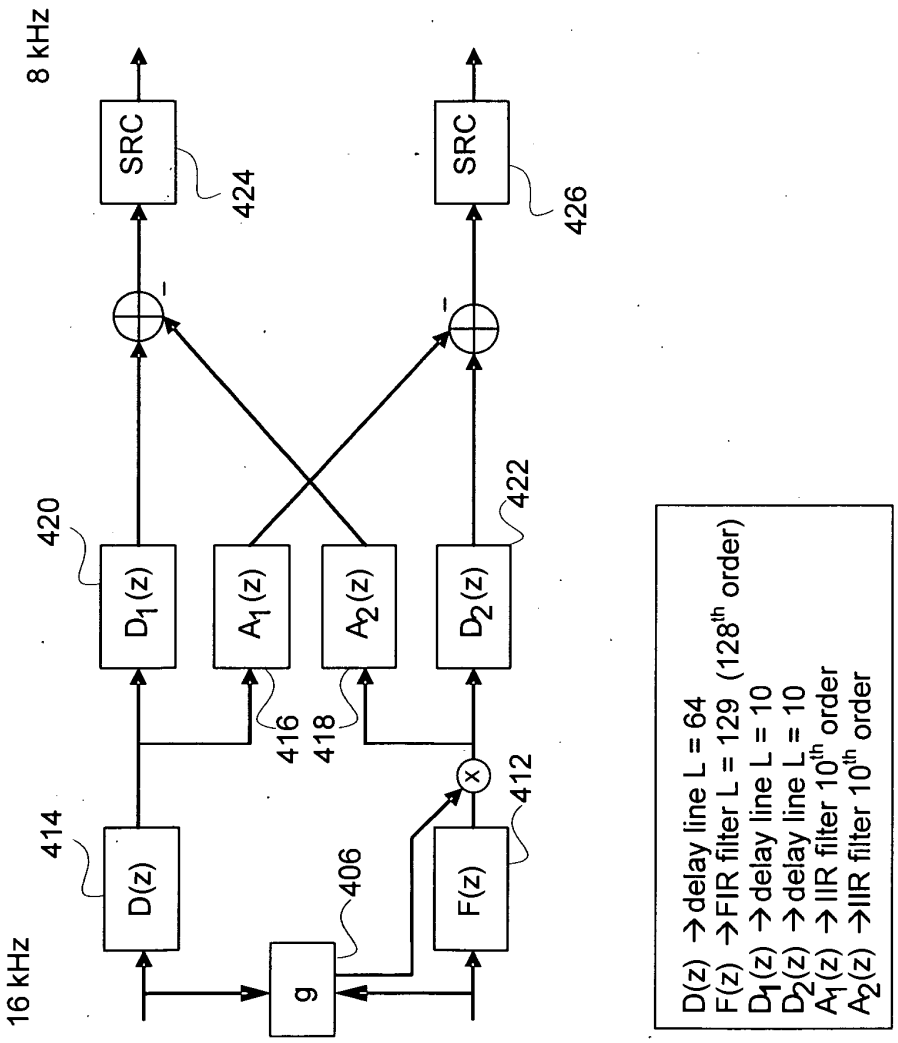

16 kHz
8 kHz
D(z)
414
406
g
F(z)
412
$D_1(z)$
420
$A_1(z)$
416
$A_2(z)$
418
$D_2(z)$
422
SRC
424
SRC
426
D(z) →delay line L = 64
F(z) →FIR filter L = 129 (128th order)
$D_1(z)$ →delay line L = 10
$D_2(z)$ →delay line L = 10
$A_1(z)$ →IIR filter 10th order
$A_2(z)$ →IIR filter 10th order

FIG. 4b

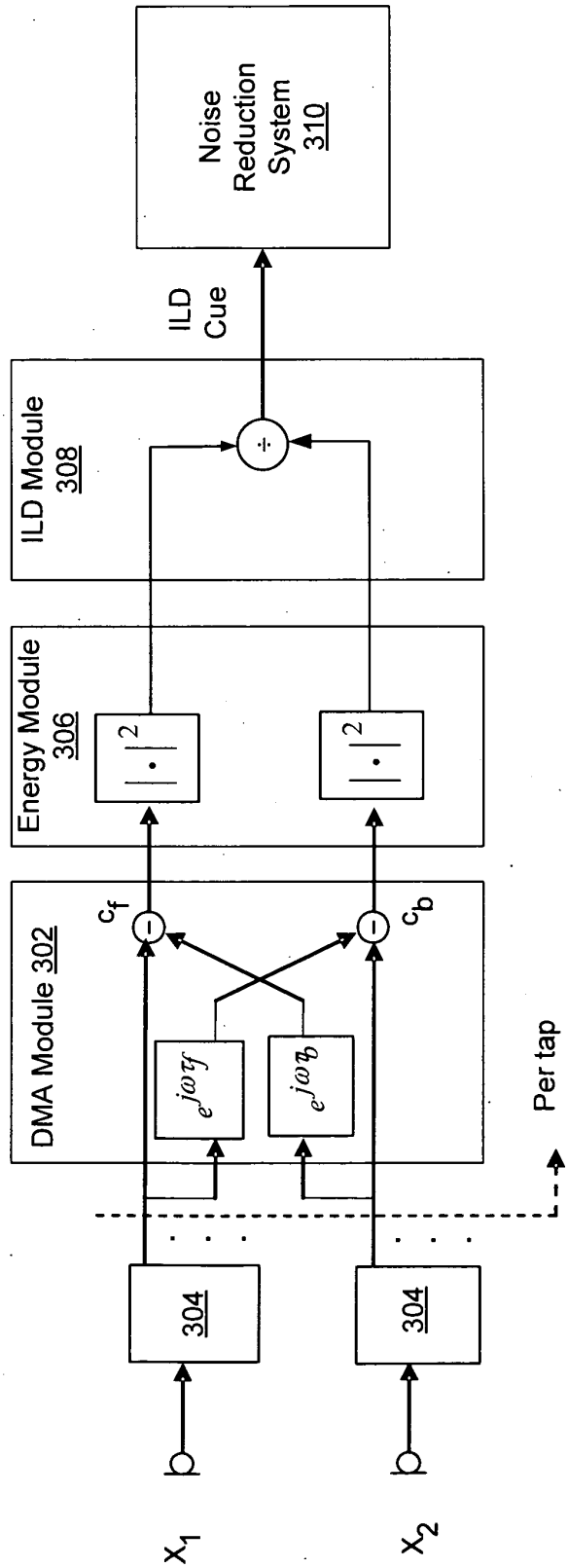

Noise Reduction System 310
ILD Cue
ILD Module 308
Energy Module 306
DMA Module 302
c_f
c_b
$e^{j\omega\tau_f}$
$e^{j\omega\tau_b}$
Per tap
304
304
x_1
x_2

FIG. 5

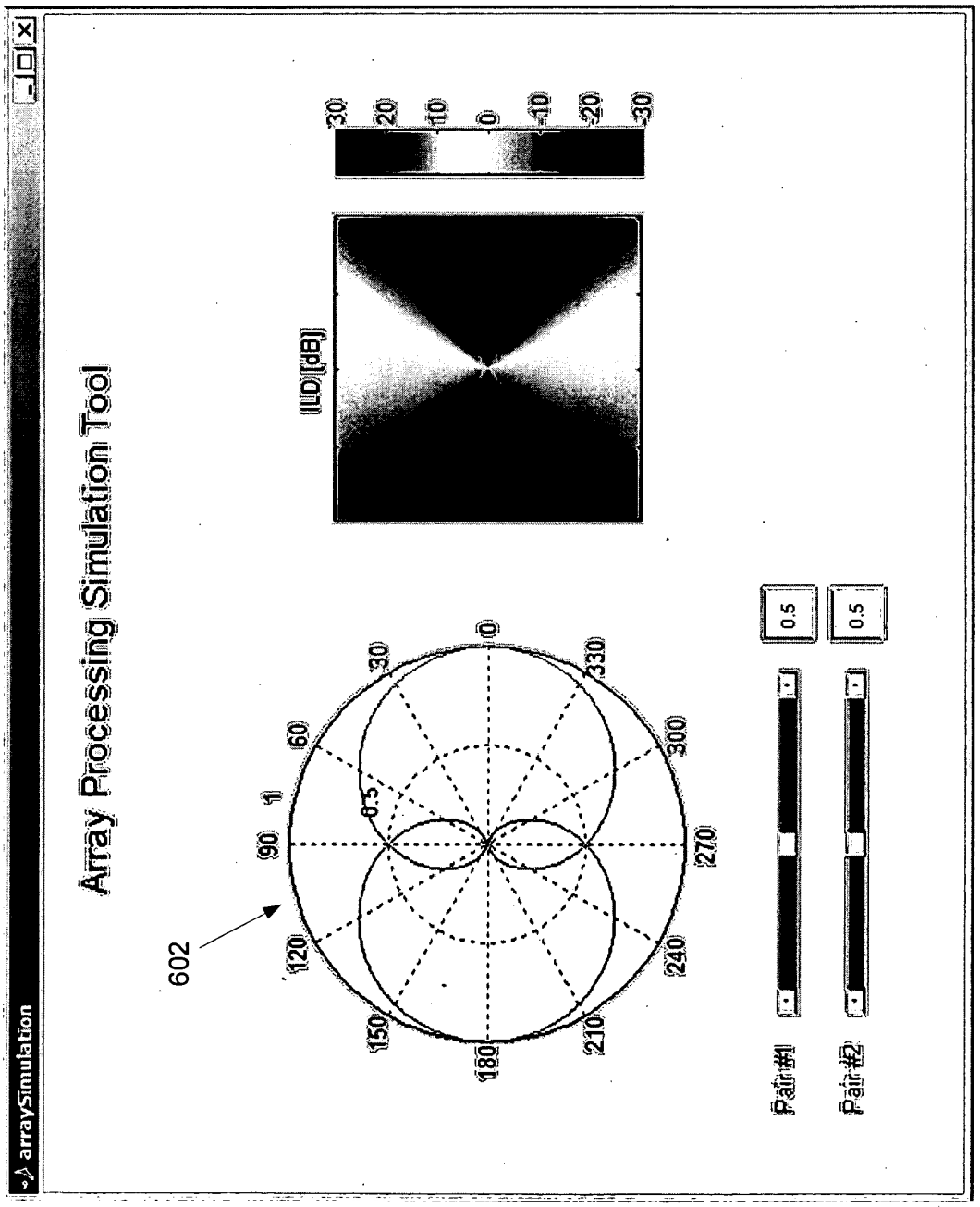

arraySimulation
Array Processing Simulation Tool
602
ILD [dB]
30
20
10
0
-10
-20
-30
0
30
60
90
120
150
180
210
240
270
300
330
0.5
1
0.5
0.5
Pair #1
Pair #2

FIG. 6

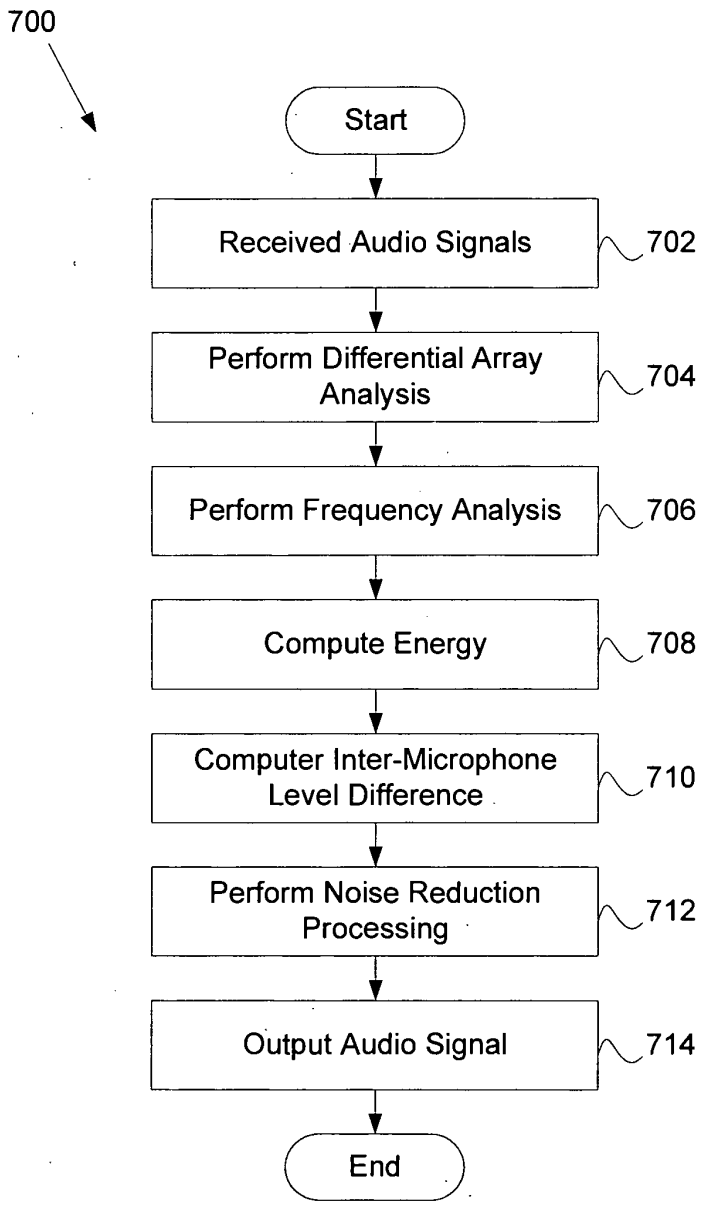

700
Start
Received Audio Signals
702
Perform Differential Array Analysis
704
Perform Frequency Analysis
706
Compute Energy
708
Computer Inter-Microphone Level Difference
710
Perform Noise Reduction Processing
712
Output Audio Signal
714
End

FIG. 7

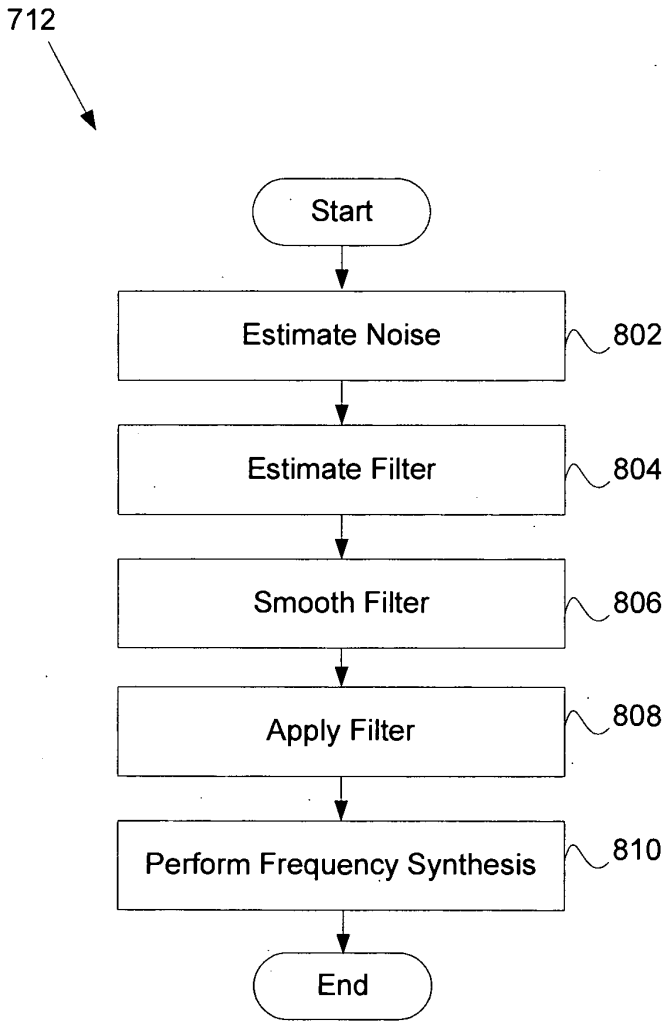

712
Start
Estimate Noise
802
Estimate Filter
804
Smooth Filter
806
Apply Filter
808
Perform Frequency Synthesis
810
End

FIG. 8

ATTORNEY'S DOCKET NO.: PA3909US

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original and first inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled

"System and Method for Utilizing Omni-Directional Microphones For Speech Enhancement"

the specification of which is attached hereto.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56, including for continuation-in-part applications, material information which became available between the filing date of the prior application and the national or PCT international filing date of the continuation-in-part application.

I hereby claim foreign priority benefits under Title 35, United States Code §119(a)-(d) or (f) or §365(b) of any foreign application(s) for patent, inventor's or plant breeder's rights certificate(s), or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent, inventor's or plant breeder's rights certificate(s), or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s):

Number	Country	Foreign Filing Date	Priority Claimed?		Certified Copy Attached?	
_____	_____	_____	☐	☐	☐	☐
(Number)	(Country)	(MM/DD/YYYY)	Yes	No	Yes	No
_____	_____	_____	☐	☐	☐	☐
(Number)	(Country)	(MM/DD/YYYY)	Yes	No	Yes	No

I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Prior Provisional Applications:

60/850,928	October 10, 2006
(Application Number)	(Filing Date)
(Application Number)	(Filing Date)

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s), or §365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application.

Prior U.S. or PCT Applications Designating the U.S.:

11/343,524	January 30, 2006	Pending
(Application Number)	(Filing Date)	(Status -- patented, pending, abandoned)
(Application Number)	(Filing Date)	(Status -- patented, pending, abandoned)

POWER OF ATTORNEY: I hereby appoint the attorney(s) and/or agent(s) associated with the customer number **22830** to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

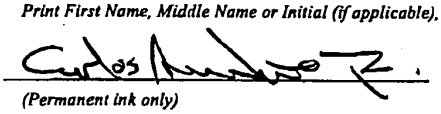
SEND ALL CORRESPONDENCE TO:

Customer Number 22830
Carr & Ferrell LLP
2200 Geng Road
Palo Alto, CA 94303
Phone: (650) 812-3400
Fax: (650) 812-3444

ATTORNEY'S DOCKET NO.: PA3909US

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first inventor: Carlos Avendano Citizenship: U.S.A.
Print First Name, Middle Name or Initial (if applicable), Last Name

Inventor's signature:  Dated: 1/29/07
(Permanent ink only) *(Date of Signature)*

Mailing Address: 1330 Villa Street, Mountain View, CA 94041
(Where the inventor customarily receives mail: Number, Street, City, State, Zip Code)

Residence Address: Campbell, CA
(if the inventor lives at a location different than his/her mailing address)

PTO/SB/06 (12-04)
Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875 Effective December 8, 2004

Application or Docket Number: 11/499732

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY FEE ($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	150.00		N/A	300.00
SEARCH FEE (37 CFR 1.16(k), (l), or (m))	N/A	N/A	N/A	$250		N/A	$500
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	$100		N/A	$200
TOTAL CLAIMS (37 CFR 1.16(i))	24 minus 20 =	* 4	X$ 25 =	150	OR	X$50 =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	3 minus 3 =	* —	X100 =			X200 =	
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))			+180=			+360=	
			TOTAL	650		TOTAL	

* If the difference in column 1 is less than zero, enter "0" in column 2.

APPLICATION AS AMENDED – PART II

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE ($)
AMENDMENT A	Total (37 CFR 1.16(i))	*	Minus	**	=	X$ 25 =		OR	X$50 =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X100 =		OR	X200 =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))					+180=		OR	+360=	
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)	OR	RATE ($)	ADDITIONAL FEE ($)
AMENDMENT B	Total (37 CFR 1.16(i))	*	Minus	**	=	X$ 25 =		OR	X$50 =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X100 =		OR	X200 =	
	Application Size Fee (37 CFR 1.16(s))									
	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))					+180=		OR	+360=	
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PATENT APPLICATION SERIAL NO. _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

02/01/2007 JBALINAN 00000014 11699732

01 FC:2011	150.00	OP
02 FC:2111	250.00	OP
03 FC:2311	100.00	OP
04 FC:2202	150.00	OP

PTO-1556
(5/87)

*U.S. Government Printing Office: 2002 — 489-267/69033

PTO/SB/35 (09-04)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U. S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

NONPUBLICATION REQUEST UNDER 35 U.S.C. 122(b)(2)(B)(i)

First Named Inventor	Carlos Avendano
Title	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
Attorney Docket Number	PA3909US

I hereby certify that the invention disclosed in the attached application **has not and will not be** the subject of an application filed in another country, or under a multilateral agreement, that requires publication at eighteen months after filing.

I hereby request that the attached application not be published under 35 U.S.C. 122(b).

Susan Yee — Signature

January 29, 2007 — Date

Susan Yee — Typed or printed name

41,388 — Registration Number, if applicable

650.812.3400 — Telephone Number

This request must be signed in compliance with 37 CFR 1.33(b) and submitted with the application **upon filing.**

Applicant may rescind this nonpublication request at any time. If applicant rescinds a request that an application not be published under 35 U.S.C. 122(b), the application will be scheduled for publication at eighteen months from the earliest claimed filing date for which a benefit is claimed.

If applicant subsequently files an application directed to the invention disclosed in the attached application in another country, or under a multilateral international agreement, that requires publication of applications eighteen months after filing, the applicant **must** notify the United States Patent and Trademark Office of such filing within forty-five (45) days after the date of the filing of such foreign or international application. **Failure to do so will result in abandonment of this application (35 U.S.C. 122(b)(2)(B)(iii)).**

This collection of information is required by 37 CFR 1.213(a). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

EV826334317US

PTO/SB/05 (07-06)
Approved for use through 01/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office. U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	PA3909US
First Inventor	Carlos Avendano
Title	System and Method for Utilizing Omni-Directional Microphones For Speech Enhancement
Express Mail Label No.	EV 826 334 317 US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. ☑ **Fee Transmittal Form** (e.g., PTO/SB/17)
   (Submit an original and a duplicate for fee processing)
2. ☑ **Applicant claims small entity status.** See 37 CFR 1.27.
3. ☑ **Specification** [*Total Pages* 26]
   Both the claims and abstract must start on a new page
   (For information on the preferred arrangement, see MPEP 608.01(a))
4. ☑ **Drawing(s)** (*35 U.S.C. 113*) [*Total Sheets* 9]
5. **Oath or Declaration** [*Total Sheets* 3]
   a. ☑ Newly executed (original or copy)
   b. ☐ A copy from a prior application (37 CFR 1.63(d))
   (for continuation/divisional with Box 18 completed)
   i. ☐ DELETION OF INVENTOR(S)
   Signed statement attached deleting inventor(s) name in the prior application, see 37 CFR 1.63(d)(2) and 1.33(b).
6. ☐ **Application Data Sheet.** See 37 CFR 1.76
7. ☐ **CD-ROM or CD-R** in duplicate, large table or Computer Program *(Appendix)*
   ☐ Landscape Table on CD
8. **Nucleotide and/or Amino Acid Sequence Submission**
   (if applicable, items a. – c. are required)
   a. ☐ Computer Readable Form (CRF)
   b. Specification Sequence Listing on:
      i. ☐ CD-ROM or CD-R (2 copies); or
      ii. ☐ Paper
   c. ☐ Statements verifying identity of above copies

ADDRESS TO:
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

ACCOMPANYING APPLICATION PARTS

9. ☑ **Assignment Papers** (cover sheet & document(s))
   Name of Assignee Audience, Inc.
10. ☐ **37 CFR 3.73(b) Statement** *(when there is an assignee)* ☑ **Power of Attorney**
11. ☐ **English Translation Document** *(if applicable)*
12. ☐ **Information Disclosure Statement** (PTO/SB/08 or PTO-1449)
    ☐ Copies of citations attached
13. ☐ **Preliminary Amendment**
14. ☑ **Return Receipt Postcard** (MPEP 503)
    (Should be specifically itemized)
15. ☐ **Certified Copy of Priority Document(s)**
    (if foreign priority is claimed)
16. ☑ **Nonpublication Request** under 35 U.S.C. 122(b)(2)(B)(i). Applicant must attach form PTO/SB/35 or equivalent.
17. ☑ Other: Check for $690.00

18. If a CONTINUING APPLICATION, *check appropriate box, and supply the requisite information below and in the first sentence of the specification following the title, or in an Application Data Sheet under 37 CFR 1.76:*

☐ Continuation ☐ Divisional ☒ Continuation-in-part (CIP) of prior application No.: 11/343,524

Prior application information: *Examiner* Unknown *Art Unit:* 2644

19. CORRESPONDENCE ADDRESS

☑ The address associated with Customer Number: 22830 ***OR*** ☐ Correspondence address below

Name					
Address					
City		State		Zip Code	
Country		Telephone		Email	

Signature		Date	January 29, 2007
Name (Print/Type)	Susan Yee	Registration No. (Attorney/Agent)	41,388

This collection of information is required by 37 CFR 1.53(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/SB/17 (01-06)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number

1769
012907

U.S. PTO

Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).

FEE TRANSMITTAL For FY 2006

☑ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 690.00

Complete if Known	
Application Number	Unknown
Filing Date	January 26, 2007
First Named Inventor	Carlos Avendano
Examiner Name	Unknown
Art Unit	Unknown
Attorney Docket No.	PA3909US

METHOD OF PAYMENT (check all that apply)

☑ Check ☐ Credit Card ☐ Money Order ☐ None ☐ Other (please identify): _____

☑ Deposit Account Deposit Account Number: 06-0600 Deposit Account Name: Carr & Ferrell LLP

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

☐ Charge fee(s) indicated below ☐ Charge fee(s) indicated below, **except for the filing fee**

☑ Charge any additional fee(s) or underpayments of fee(s) under 37 CFR 1.16 and 1.17 ☑ Credit any overpayments

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

FEE CALCULATION (All the fees below are due upon filing or may be subject to a surcharge.)

1. BASIC FILING, SEARCH, AND EXAMINATION FEES

	FILING FEES		SEARCH FEES		EXAMINATION FEES		
		Small Entity		Small Entity		Small Entity	
Application Type	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fees Paid ($)
Utility	300	150	500	250	200	100	500
Design	200	100	100	50	130	65	
Plant	200	100	300	150	160	80	
Reissue	300	150	500	250	600	300	
Provisional	200	100	0	0	0	0	

2. EXCESS CLAIM FEES

Fee Description	Fee ($)	Small Entity Fee ($)
Each claim over 20 (including Reissues)	50	25
Each independent claim over 3 (including Reissues)	200	100
Multiple dependent claims	360	180

Total Claims		Extra Claims	Fee ($)		Fee Paid ($)
26	- 20 or HP =	6	x 25	=	150

HP = highest number of total claims paid for, if greater than 20.

Indep. Claims		Extra Claims	Fee ($)		Fee Paid ($)
3	- 3 or HP =	0	x 100	=	0

HP = highest number of independent claims paid for, if greater than 3.

Multiple Dependent Claims	
Fee ($)	Fee Paid ($)

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets		Extra Sheets		Number of each additional 50 or fraction thereof	Fee ($)		Fee Paid ($)
	- 100 =		/ 50 =	(round **up** to a whole number)	x	=	

4. OTHER FEE(S)

	Fees Paid ($)
Non-English Specification, $130 fee (no small entity discount)	
Other (e.g., late filing surcharge): Assignment	40

SUBMITTED BY

Signature	Susan Yee	Registration No. (Attorney/Agent) 41,388	Telephone 650.812.3400
Name (Print/Type)	Susan Yee		Date January 29, 2007

This collection of information is required by 37 CFR 1.136. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	DRAWINGS	TOT CLAIMS	IND CLAIMS
11/699,732	01/29/2007	2615	650	PA3909US	9	26	3

CONFIRMATION NO. 7721

22830
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CA94303

FILING RECEIPT

Date Mailed: 02/26/2007

Receipt is acknowledged of this regular Patent Application. It will be considered in its order and you will be notified as to the results of the examination. Be sure to provide the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION when inquiring about this application. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please mail to the Commissioner for Patents P.O. Box 1450 Alexandria Va 22313-1450. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections (if appropriate).**

Applicant(s)

Carlos Avendano, Campbell, CA;

Assignment For Published Patent Application

Audience, Inc.

Power of Attorney: The patent practitioners associated with Customer Number 22830

Domestic Priority data as claimed by applicant

This appln claims benefit of 60/850,928 10/10/2006
and is a CIP of 11/343,524 01/30/2006

Foreign Applications

If Required, Foreign Filing License Granted: 02/26/2007

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US11/699,732**

Projected Publication Date: Request for Non-Publication Acknowledged

Non-Publication Request: Yes

Early Publication Request: No

** SMALL ENTITY **

Title

System and method for utilizing omni-directional microphones for speech enhancement

Preliminary Class

381

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4158).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the

subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

IFW

PTO/SB/21 (09-04)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

OIPE IAP
MAY 24 2007
PATENT & TRADEMARK OFFICE

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission	11

Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

ENCLOSURES *(Check all that apply)*

- ☐ Fee Transmittal Form
  - ☐ Fee Attached
- ☐ Amendment/Reply
  - ☐ After Final
  - ☐ Affidavits/declaration(s)
- ☐ Extension of Time Request
- ☐ Express Abandonment Request
- ☑ Information Disclosure Statement
- ☐ Certified Copy of Priority Document(s)
- ☐ Reply to Missing Parts/ Incomplete Application
  - ☐ Reply to Missing Parts under 37 CFR 1.52 or 1.53
- ☐ Drawing(s)
- ☐ Licensing-related Papers
- ☐ Petition
- ☐ Petition to Convert to a Provisional Application
- ☐ Power of Attorney, Revocation Change of Correspondence Address
- ☐ Terminal Disclaimer
- ☐ Request for Refund
- ☐ CD, Number of CD(s) _____
  - ☐ Landscape Table on CD
- ☐ After Allowance Communication to TC
- ☐ Appeal Communication to Board of Appeals and Interferences
- ☐ Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- ☐ Proprietary Information
- ☐ Status Letter
- ☑ Other Enclosure(s) (please Identify below):
  1) Confirmation postcard;
  2) Duplicate of IDS;
  3) PTO/SB/08 (x2);
  4) Copies of 3 cited nonpatent references

Remarks

Total page number does not include postcard and nonpatent references

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	Carr & Ferrell LLP Cust. No. 22830		
Signature	Susan Yee		
Printed name	Susan Yee		
Date	May 21, 2007	Reg. No.	41,388

CERTIFICATE OF TRANSMISSION/MAILING

I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:

Signature	Susan Yee		
Typed or printed name	Susan Yee, Reg. No. 41,388	Date	May 21, 2007

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and1.14. This collection is estimated to 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.



IN THE

UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Carlos Avendano
APPLICATION NO.:	11/699,732
FILING DATE:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
ART UNIT:	2615
EXAMINER:	Unknown
ATTY.DKT.NO.:	PA3909US

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Mail Stop Amendment, Commissioner for Patents, Alexandria, VA 22313-1450, on the date printed below:

Dated: May 21, 2007

By: Susan Yee
Susan Yee, Reg. No. 41,388

MAIL STOP AMENDMENT
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. §§ 1.56, and 1.97-1.98

SIR:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08 listing several references; copies of the non-

patent literature are enclosed. The Examiner is requested to make these references of official record in the application. The references cited may be material to examination of the application and are submitted in compliance with Applicant's duty of disclosure as defined by 37 C.F.R. § 1.56.

No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

Applicant believes no fee is due with this submission. If a fee is due, however, the Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600. A copy of this statement is submitted in duplicate for this purpose.

Respectfully submitted,
Carlos Avendano

Dated: May 21, 2007

By: _____

Susan Yee, Reg. No. 41,388
Carr & Ferrell *LLP*
2200 Geng Road
Palo Alto, CA 94303
TEL: (650) 812-3400
FAX: (650) 812-3444

PTO/SB/08A (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

OIPE
MAY 24 2007
PATENT & TRADEMARK OFFICE

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 3

Complete if Known	
Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

U. S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
		US- 4,649,505	03-10-1987	Zinser, Jr. et al.	
		US- 5,319,736	06-07-1994	Hunt	
		US- 5,381,473	01-10-1995	Andrea et al.	
		US- 5,694,474	12-02-1997	Ngo et al.	
		US- 6,002,776	12-14-1999	Bhadkamkar et al.	
		US- 6,061,456	05-09-2000	Andrea et al.	
		US- 6,072,881	06-06-2000	Linder	
		US- 6,222,927	04-24-2001	Feng et al.	
		US- 6,317,501	11-13-2001	Matsuo	
		US- 6,363,345	03-26-2002	Marash et al.	
		US- 6,430,295	08-06-2002	Handel et al.	
		US- 6,584,203	06-24-2003	Elko et al.	
		US- 6,717,991	04-06-2004	Gustafsson et al.	
		US- 6,738,482	05-18-2004	Jaber	
		US- 6,760,450	07-06-2004	Matsuo	
		US- 6,785,381	08-31-2004	Gartner et al.	
		US- 6,795,558	09-21-2004	Matsuo	
		US- 7,099,821	08-29-2006	Visser et al.	
		US- 7,171,008	01-30-2007	Elko	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PTO/SB/08A (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 2 of 3

Complete if Known	
Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

U. S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
		US- 2001/0016020	08-23-2001	Gustafsson et al.	
		US- 2001/0031053	10-18-2001	Feng et al.	
		US- 2002/0009203	01-24-2002	Erten	
		US- 2002/0041693	04-11-2002	Matsuo	
		US- 2002/0080980	06-27-2002	Matsuo	
		US- 2002/0106092	08-08-2002	Matsuo	
		US- 2002/0116187	08-22-2002	Erten	
		US- 2003/0039369	02-27-2003	Bullen	
		US- 2003/0072460	04-17-2003	Gonopolskiy et al.	
		US- 2003/0099345	05-29-2003	Gartner et al.	
		US- 2003/0138116	07-24-2003	Jones et al.	
		US- 2003/0169891	09-11-2003	Ryan et al.	
		US- 2005/0185813	08-25-2005	Sinclair et al.	
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

PTO/SB/08B (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 3 of 3

Complete if Known	
Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
		AVENDANO, CARLOS, "Frequency-Domain Source Identification and Manipulation in Stereo Mixes for Enhancement, Suppression and Re-panning Applications," 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 19-22, 2003, pp. 55-58, New Paltz, New York, USA.	
		WIDROW, B. et al., "Adaptive Atenna Systems," December 1967, pp. 2143-2159, Vol. 55 No. 12, Proceedings of the IEEE.	
		ELKO, GARY W., "Differential Microphone Arrays," Audio Signal Processing for Next-Generation Multimedia Communication Systems, 2004, pp. 12-65, Kluwer Academic Publishers, Norwell, Massachusetts, USA.	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

1 Applicant's unique citation designation number (optional). 2 Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

OIPE IAP22 OCT 11 2007 PATENT & TRADEMARK OFFICE

PTO/SB/36 (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U. S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

RESCISSION OF PREVIOUS NONPUBLICATION REQUEST
(35 U.S.C. 122(b)(2)(B)(ii))
AND, IF APPLICABLE,
NOTICE OF FOREIGN FILING
(35 U.S.C. 122(b)(2)(B)(iii))

Send completed form to:
Mail Stop PG Pub
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450
FAX: (571) 273-8300

Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Title	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
Atty Docket Number	PA3909US
Art Unit	2615
Examiner	Unknown

A request that the above-identified application not be published under 35 U.S.C. 122(b) (nonpublication request) was included with the above-identified application on filing pursuant to 35 U.S.C. 122(b)(2)(B)(i).
I hereby **rescind** the previous nonpublication request.

If a notice of foreign or international filing is or will be required by 35 U.S.C. 122(b)(2)(B)(iii) and 37 CFR 1.213(c), I hereby provide such notice. This notice is being provided no later than forty-five (**45**) days after the date of such foreign or international filing.

*If a notice of subsequent foreign or international filing required by 35 U.S.C. 122(b)(2)(B)(iii) and 37 CFR 1.213(c) was not filed within forty-five (**45**) days after the date of filing of the foreign or international application, the application is ABANDONED, and a petition to revive under 37 CFR 1.137(b) is required. See 37 CFR 1.137(f).*

Signature	October 9, 2007 Date
Susan Yee Typed or printed name	41,388 Registration Number, if applicable
650-812-3400 Telephone Number	

This request must be signed in compliance with 37 CFR 1.33(b).

If information or assistance is needed in completing this form, please contact the Pre-Grant Publication Division at (703)605-4283 or by e-mail at PGPub@USPTO.gov.

CERTIFICATE OF MAILING OR TRANSMISSION

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Mail Stop PG Pub, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, or facsimile transmitted to the U.S. Patent and Trademark Office on the date shown below.

Signature			
Name (Print/Type)	Susan Yee, Reg. No. 41,388	Date	October 9, 2007

This collection of information is required by 37 CFR 1.213(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 6 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Mail Stop PG Pub, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

OIPE IAP22 OCT 11 2007 PATENT & TRADEMARK OFFICE

TJW

PTO/SB/21 (09-04)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission	2

Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

ENCLOSURES *(Check all that apply)*

- ☐ Fee Transmittal Form
  - ☐ Fee Attached
- ☐ Amendment/Reply
  - ☐ After Final
  - ☐ Affidavits/declaration(s)
- ☐ Extension of Time Request
- ☐ Express Abandonment Request
- ☐ Information Disclosure Statement
- ☐ Certified Copy of Priority Document(s)
- ☐ Reply to Missing Parts/ Incomplete Application
  - ☐ Reply to Missing Parts under 37 CFR 1.52 or 1.53
- ☐ Drawing(s)
- ☐ Licensing-related Papers
- ☐ Petition
- ☐ Petition to Convert to a Provisional Application
- ☐ Power of Attorney, Revocation Change of Correspondence Address
- ☐ Terminal Disclaimer
- ☐ Request for Refund
- ☐ CD, Number of CD(s) _____
  - ☐ Landscape Table on CD
- ☐ After Allowance Communication to TC
- ☐ Appeal Communication to Board of Appeals and Interferences
- ☐ Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- ☐ Proprietary Information
- ☐ Status Letter
- ☑ Other Enclosure(s) (please Identify below):

1) Confirmation postcard;
2) Rescission of Previous Nonpublication Request (1 pg)

Remarks

Total page number does not include postcard.

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	Carr & Ferrell LLP Cust. No. 22830		
Signature	Susan Yee		
Printed name	Susan Yee		
Date	October 9, 2007	Reg. No.	41,388

CERTIFICATE OF TRANSMISSION/MAILING

I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:

Signature	Susan Yee		
Typed or printed name	Susan Yee, Reg. No. 41,388	Date	October 9, 2007

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and1.14. This collection is estimated to 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(c) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/699,732	01/29/2007	Carlos Avendano	PA3909US

CONFIRMATION NO. 7721

22830
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CA94303

Date Mailed. 10/18/2007

NOTICE OF NEW OR REVISED PROJECTED PUBLICATION DATE

The above-identified application has a new or revised projected publication date. The current projected publication date for this application is 01/24/2008. If this is a new projected publication date (there was no previous projected publication date), the application has been cleared by Licensing & Review or a secrecy order has been rescinded and the application is now in the publication queue.

If this is a revised projected publication date (one that is different from a previously communicated projected publication date), the publication date has been revised due to processing delays in the USPTO or the abandonment and subsequent revival of an application. The application is anticipated to be published on a date that is more than six weeks different from the originally-projected publication date.

More detailed publication information is available through the private side of Patent Application Information Retrieval (PAIR) System. The direct link to access PAIR is currently http://pair.uspto.gov. Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Questions relating to this Notice should be directed to the Office of Patent Publication at 1-888-786-0101.

PART 1 - ATTORNEY/APPLICANT COPY

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/699,732	01/29/2007	Carlos Avendano	PA3909US

CONFIRMATION NO. 7721

22830
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CA 94303

NONPUBLICATION RESCISSION LETTER

OC000000026334561

Date Mailed: 10/19/2007

Communication Regarding Rescission Of Nonpublication Request and/or Notice of Foreign Filing

Applicant's rescission of the previously-filed nonpublication request and/or notice of foreign filing is acknowledged. The paper has been reflected in the Patent and Trademark Office's (USPTO's) computer records so that the earliest possible projected publication date can be assigned.

The projected publication date is 01/24/2008.

If applicant rescinded the nonpublication request before or on the date of "foreign filing,"[1] then no notice of foreign filing is required.

If applicant foreign filed the application after filing the above application and before filing the rescission, and the rescission did not also include a notice of foreign filing, then a notice of foreign filing (not merely a rescission) is required to be filed within 45 days of the date of foreign filing. See 35 U.S.C. § 122(b)(2)(B)(iii), and Clarification of the United States Patent and Trademark Office's Interpretation of the Provisions of 35 U.S.C. § 122(b)(2)(B)(ii)-(iv), 1272 Off. Gaz. Pat. Office 22 (July 1, 2003).

If a notice of foreign filing is required and is not filed within 45 days of the date of foreign filing, then the application becomes abandoned pursuant to 35 U.S.C. § 122(b)(2)(B)(iii). In this situation, applicant should either file a petition to revive or notify the Office that the application is abandoned. See 37 CFR 1.137(f). Any such petition to revive will be forwarded to the Office of Petitions for a decision. Note that the filing of the petition will not operate to stay any period of reply that may be running against the application.

Questions regarding petitions to revive should be directed to the Office of Petitions at (571) 272-3282. Questions regarding publications of patent applications should be directed to the patent application publication hotline at (703) 605-4283 or by e-mail pgpub@uspto.gov.

[1] Note, for purpose of this notice, that "foreign filing" means "filing an application directed to the same invention in another country, or under a multilateral international agreement, that requires publication of applications 18 months after filing".

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(c) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/699,732	01/29/2007	Carlos Avendano	PA3909US

CONFIRMATION NO. 7721

22830
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CA94303

Title: System and method for utilizing omni-directional microphones for speech enhancement

Publication No. US-2008-0019548-A1
Publication Date: 01/24/2008

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently http://pair.uspto.gov/. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Pre-Grant Publication Division, 703-605-4283

PTO/SB/08a (10-08)
Approved for use through 11/30/2008. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

U.S.PATENTS						
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	4649505		1987-03-10	Zinser, Jr. et al.	
	2	4658426		1987-04-14	Chabries et al.	
	3	5208864		1993-05-04	Kaneda	
	4	5319736		1994-06-07	Hunt	
	5	5381473		1995-01-10	Andrea et al.	
	6	5402496		1995-03-28	Soli et al.	
	7	5574824		1996-11-12	Slyh et al.	
	8	5602962		1997-02-11	Kellermann	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	9	5694474		1997-12-02	Ngo et al.	
	10	5757937		1998-05-26	Itoh et al.	
	11	5796819		1998-08-18	Romesburg	
	12	6002776		1999-12-14	Bhadkamkar et al.	
	13	6061456		2000-05-09	Andrea et al.	
	14	6072881		2000-06-06	Linder	
	15	6222927	B1	2001-04-24	Feng et al.	
	16	6317501	B1	2001-11-13	Matsuo	
	17	6363345	B1	2002-03-26	Marash et al.	
	18	6430295	B1	2002-08-06	Handel et al.	
	19	6549630	B1	2003-04-15	Bobisuthi	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	20	6717991	B1	2004-04-06	Gustafsson et al.	
	21	6738482	B1	2004-05-18	Jaber	
	22	6760450	B2	2004-07-06	Matsuo	
	23	6785381	B2	2004-08-31	Gartner et al.	
	24	6795558	B2	2004-09-21	Matsuo	
	25	6882736	B2	2005-04-19	Dickel et al.	
	26	6917688	B2	2005-07-12	Yu et al.	
	27	6978159	B2	2005-12-20	Feng et al.	
	28	7031478	B2	2006-04-18	Belt et al.	
	29	7155019	B2	2006-12-26	Hou	
	30	7171008	B2	2007-01-30	Elko	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	31	7174022	B1	2007-02-06	Zhang et al.	
	32	7206418	B2	2007-04-17	Yang et al.	
	33	7246058	B2	2007-07-17	Burnett	

If you wish to add additional U.S. Patent citation information please click the Add button.

U.S.PATENT APPLICATION PUBLICATIONS

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20010016020	A1	2001-08-23	Gustafsson et al.	
	2	20010031053	A1	2001-10-18	Feng et al.	
	3	20020009203	A1	2002-01-24	Erten	
	4	20020041693	A1	2002-04-11	Matsuo	
	5	20020080980	A1	2002-06-27	Matsuo	
	6	20020106092	A1	2002-08-08	Matsuo	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	7	20020116187	A1	2002-08-22	Erten	
	8	20030026437	A1	2003-02-06	Janse et al.	
	9	20030039369	A1	2003-02-27	Bullen	
	10	20030072460	A1	2003-04-17	Gonopolskiy et al.	
	11	20030099345	A1	2003-05-29	Gartner et al.	
	12	20030138116	A1	2003-07-24	Jones et al.	
	13	20030147538	A1	2003-08-07	Elko	
	14	20030169891	A1	2003-09-11	Ryan et al.	
	15	20040133421	A1	2004-07-08	Burnett et al.	
	16	20050049864	A1	2005-03-03	Kaltenmeier et al.	
	17	20050060142	A1	2005-03-17	Visser et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

If you wish to add additional U.S. Published Application citation information please click the Add button.

FOREIGN PATENT DOCUMENTS

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐

If you wish to add additional Foreign Patent Document citation information please click the Add button

NON-PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☒ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☐ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Susan Yee/	Date (YYYY-MM-DD)	2009-02-02
Name/Print	Susan Yee	Registration Number	41388

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Carlos Avendano
APPLICATION NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Vivian C. Chin
GROUP ART UNIT:	2615
ATTY.DKT.NO.:	PA3909US

INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. § 1.56, and 1.97-1.98

SIR:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08, listing several references. The Examiner is requested to make these references of official record in the application.

The references cited may be material to examination of the application and are submitted in compliance with Applicant's duty of disclosure as defined by 37 C.F.R. § 1.56. No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

Applicant believes no fee is due with this submission. If a fee is due, however, the Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600.

Respectfully submitted,
Carlos Avendano

February 2, 2009

By: /Susan Yee/

Susan Yee, Reg. No. 41,388
Carr & Ferrell *LLP*
2200 Geng Road
Palo Alto, CA 94303
TEL: (650) 812-3400
FAX: (650) 812-3444

PTO/SB/08a (10-08)
Approved for use through 11/30/2008. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	7099821	B2	2006-08-29	Visser et al.	
	2	7146316	B2	2006-12-05	Alves	
	3	7142677	B2	2006-11-28	Gonopolskiy	
If you wish to add additional U.S. Patent citation information please click the Add button.						Add
U.S.PATENT APPLICATION PUBLICATIONS						Remove
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20050152559	A1	2005-07-14	Gierl et al.	
	2	20050185813	A1	2005-08-25	Sinclair et al.	
	3	20050213778	A1	2005-09-29	Buck et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	4	20050276423	A1	2005-12-15	Aubauer et al.	
	5	20060120537	A1	2006-06-08	Burnett et al.	
	6	20070021958	A1	2007-01-25	Visser et al.	
	7	20070027685	A1	2007-02-01	Arakawa et al.	
	8	20070033020	A1	2007-02-08	(Kelleher) Francois et al.	
	9	20070078649	A1	2007-04-05	Hetherington et al.	
	10	20070116300	A1	2007-05-24	Chen	
	11	20070150268	A1	2007-06-28	Acero et al.	
	12	20070154031	A1	2007-07-05	Avendano et al.	
	13	20070165879	A1	2007-07-19	Deng et al.	
	14	20030061032	A1	2003-03-27	Gonopolskiy	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	15	20060074646	A1	2006-04-06	Alves et al.	

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	04184400	JP		1992-07-01	Yoshio		☐
	2	06269083	JP		1994-09-22	Toru et al.		☐
	3	62110349	JP		1987-05-21	Miura		☐

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1	MARC MOONEN et al. "Multi-Microphone Signal Enhancement Techniques for Noise Suppression and Dereverberation," source(s): http://www.esat.kuleuven.ac.be/sista/yearreport97/node37.html.	☐
	2	STEVEN BOLL et al. "Suppression of Acoustic Noise in Speech Using Two Microphone Adaptive Noise Cancellation", source(s): IEEE Transactions on Acoustic, Speech, and Signal Processing. Vol. v ASSP-28, n 6, December 1980, pgs. 752-753.	☐
	3	CHEN LIU et al. "A two-microphone dual delay-line approach for extraction of a speech sound in the presence of multiple interferers", source(s): Acoustical Society of America. Vol. 110, 6, December 2001, pgs. 3218-3231.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	4	COHEN et al. "Microphone Array Post-Filtering for Non-Stationary Noise", source(s): IEEE. May 2002	☐
	5	JINGDONG CHEN et al. "New Insights into the Noise Reduction Wiener Filter", source(s): IEEE Transactions on Audio, Speech, and Langauge Processing. Vol. 14, 4, July 2006, pgs. 1218-1234.	☐
	6	RAINER MARTIN et al. "Combined Acoustic Echo Cancellation, Dereverberation and Noise Reduction: A two Microphone Approach", source(s): Annales des Telecommunications/Annals of Telecommunications. Vol. 29, 7-8, July -Aug 1994, pgs. 429-438.	☐
	7	MITSUNORI MIZUMACHI et al. "Noise Reduction by Paired-Microphones Using Spectral Subtraction", source(s): 1998 IEEE. pgs. 1001-1004.	☐
	8	LUCAS PARRA et al. "Convolutive blind Separation of Non-Stationary", source(s): IEEE Transactions on Speech and Audio Processing. Vol 8, 3, May 2008, pgs. 320-327.	☐
	9	ISREAL COHEN. "Multichannel Post-Filtering in Nonstationary Noise Environment", source(s): IEEE Transactions on Signal Processing. Vol. 52, 5, May 2004, pgs. 1149-1160.	☐
	10	R.A. GOUBRAN. "Acoustic Noise Suppression Using Regressive Adaptive Filtering", source(s): 1990 IEEE. pgs. 48-53	☐
	11	IVAN TASHEV et al. "Microphone Array of Headset with Spatial Noise Suppressor", source(s): http://research. microsoft.com/users/ivantash/Documents/Tashev_MAforHeadset_HSCMA_05.pdf. (4 pages).	☐
	12	MARTIN FUCHS et al. "Noise Suppression for Automotive Applications Based on Directional Information", source(s): 2004 IEEE. pgs. 237-240.	☐
	13	JEAN-MARC VALIN et al. "Enhanced Robot Audition Based on Microphone Array Source Separation with Post-Filter", source(s): Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, September 28- October 2, 2004, Sendai, Japan. pgs. 2123-2128.	☐
	14	JONT B. ALLEN. "Short Term Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform", IEEE Transactions on Acoustics, Speech, and Signal Processing. Vol ASSP-25, 3. June 1977. pgs. 235-238.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	15	JONT B. ALLEN et al. "A Unified Approach to Short-Time Fourier Analysis and Synthesis", Proceedings of the IEEE. Vol. 65, 11, November 1977. pgs. 1558-1564.	☐
	16	C. Avendano, "Frequency-Domain Techniques for Source Identification and Manipulation in Stereo Mixes for Enhancement, Suppression and Re-Panning Applications," in Proc. IEEE Workshop on Application of Signal Processing to Audio and Acoustics, Waspaa, 03, New Paltz, NY, 2003.	☐
	17	B. Widrow et al., "Adaptive Antenna Systems," Proceedings IEEE, Vol. 55, No. 12, pp. 2143-2159, December, 1967.	☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☒ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☐ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Susan Yee/	Date (YYYY-MM-DD)	2008-02-02
Name/Print	Susan Yee	Registration Number	41388

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Carlos Avendano
APPLICATION NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Vivian C. Chin
GROUP ART UNIT:	2615
ATTY.DKT.NO.:	PA3909US

COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

List of Co-Pending Patents or Patent Applications That May Be Directed Towards Similar Subject Matter

Examiner's Initials	Application Number	First-Named Inventor	Title	Filing Date
	11/343,524	Carlos Avendano	System and Method for Utilizing Inter-Microphone Level Differences for Speech Enhancement	01/30/2006

DERWENT-ACC-NO: 1992-272247
DERWENT-WEEK: 199233
COPYRIGHT 2007 DERWENT INFORMATION LTD
TITLE: Sound noise suppressor for voice response unit - estimates noise level from power spectrum difference between output signals of two microphones NoAbstract

PATENT-ASSIGNEE: NIPPON TELEGRAPH & TELEPHONE CORP[NITE]

PRIORITY-DATA: 1990JP-0313388 (November 19, 1990)

PATENT-FAMILY:
PUB-NO PUB-DATE LANGUAGE PAGES MAIN-IPC JP 04184400 A July 1, 1992 N/A
007
G10L 003/02

APPLICATION-DATA:
PUB-NO APPL-DESCRIPTOR APPL-NO APPL-DATE JP 04184400A N/A
1990JP-0313388
November 19, 1990

INT-CL (IPC): G10L003/02, G10L009/00 , H04R003/00

ABSTRACTED-PUB-NO:

EQUIVALENT-ABSTRACTS:

TITLE-TERMS: SOUND NOISE SUPPRESS VOICE RESPOND UNIT ESTIMATE NOISE LEVEL POWER SPECTRUM DIFFER OUTPUT SIGNAL TWO MICROPHONE NOABSTRACT

DERWENT-CLASS: P86 W04

EPI-CODES: W04-T; W04-V;

SECONDARY-ACC-NO:
Non-CPI Secondary Accession Numbers: N1992-208187

NOISE REMOVING DEVICE

Patent number: JP4184400

Publication date: 1992-07-01

Inventor: NAKADAI YOSHIO

Applicant: NIPPON TELEGRAPH & TELEPHONE

Classification:

- international: ***G10L15/20; G10L21/02; H04B1/10; H04R3/00;*** **G10L15/00; G10L21/00; H04B1/10; H04R3/00;** (IPC1-7): G10L3/02; G10L9/00; H04R3/00

- european:

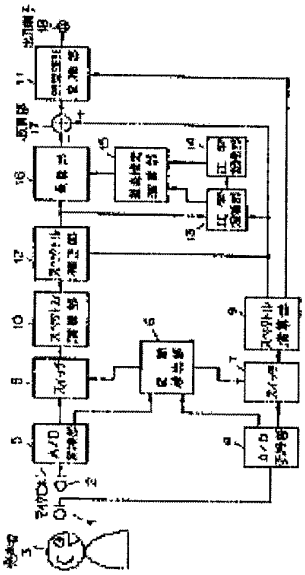
Application number: JP19900313388 19901119

Priority number(s): JP19900313388 19901119

Report a data error here

Abstract of **JP4184400**

PURPOSE:To allow the noise removal which is strong to the input of noises having time changes, hardly removes voice components and has a high noise removing capacity by decreasing noise components and removing the voice components from the signals in which the noises and voices coexist. CONSTITUTION:The noises are received in the environment where there are the ambient noises and the outputs of A/D converting sections 4 and 5 are respectively inputted to spectrum computing sections 9 and 10 where the filter coefft. of a spectrum correcting section 12 is learned by utilizing the ratio of the long time average value to the short-time power spectra. The ratio of the spectra of the voices of an uttering person 3 is calculated in the state of having no ambient noises and the state of having the noises. The waveforms of the sections decided to be the voices by a segment detecting section 6 from the digital waveforms by the A/D converting section 4 to 5 of the voices of the uttering person 3 are subjected to the correction by the spectrum correcting section 12 only for the output of the spectrum computing section 10 and are respectively sent to a ratio computing section 13. The noise component are estimated by a noise estimating computing section 15 on accordance with the results of the calculation. The results of the calculation of the estimated noise power spectra are sent to a subtraction section 17 and are subtracted, by which the voice waveforms after the noise removal are obtd.



Data supplied from the ***esp@cenet*** database - Worldwide

⑲日本国特許庁(JP)　⑪特許出願公開

⑫公開特許公報(A)　平4−184400

⑤Int.Cl.[5]		識別記号	庁内整理番号
G 10 L	3/02	301	8842−5D
	9/00	F	8622−5H
H 04 R	3/00	320	8622−5H

⑬公開　平成4年(1992)7月1日

審査請求　未請求　請求項の数 1　(全7頁)

⑭発明の名称　雑音除去装置

⑳特　願　平2−313388

㉒出　願　平2(1990)11月19日

⑫発明者　中　葺　芳　夫　東京都千代田区内幸町1丁目1番6号　日本電信電話株式会社内

⑪出願人　日本電信電話株式会社　東京都千代田区内幸町1丁目1番6号

⑭代理人　弁理士　草野　卓

明　細　書

1．発明の名称

雑音除去装置

2．特許請求の範囲

(1) 同一発声者からの音声信号をその発声者に近接した2か所の位置から同時に受信できる第1および第2の信号入力器と、

これら第1，第2の信号入力器からの出力をそれぞれパワースペクトルに変換する第1，第2のスペクトル演算手段と、

上記第2の信号入力器の出力パワースペクトルに対する上記第1の信号入力器の出力パワースペクトルの比率を求める比率演算手段と、

周囲雑音が存在しない状態での収容時の上記比率を記憶する比率記憶手段と、

上記記憶した比率と、現在計算される上記比率と、上記第2の信号入力器の出力パワースペクトルとからその出力パワースペクトルに重畳されている雑音パワースペクトルを推定する雑音推定演算手段と、

この推定雑音パワースペクトルを上記第1の信号入力器の出力パワースペクトルより除去する減算手段と、

を具備する雑音除去装置。

3．発明の詳細な説明

「産業上の利用分野」

この発明は、雑音と音声とが混在する信号から、雑音成分を低減し音声成分を取り出す雑音除去装置に関するものである。

「従来の技術」

従来より、音声認識装置などにおいて、自動車騒音下などの種々の雑音環境下でも音声の特徴を安定して得ることを目的として、音声入力段にさまざまな雑音除去手法を用いることが提案されている。

例えば、装置の信号入力器を1組とし、入力される信号について予め目的とする音声信号を含む音声区間と音声信号を含まない雑音区間とがわかっているときに、雑音区間についてその平均パワースペクトルを求め、音声区間全体にわたってこ

の平均パワースペクトルを差し引いて雑音除去を行う方法がある。

また、信号入力器を2組とし、信号入力器に2組に同時に入力される信号について、音声信号と雑音とのSN比が大きい入力信号パワースペクトルからSN比の小さい入力信号パワースペクトルを差し引いて雑音除去を行う方法がある。

「発明が解決しようとする課題」

前者の方法では、雑音パワースペクトルが時間的変動を伴う場合に、除去しようとする雑音パワースペクトルと実際の雑音パワースペクトルとの間に違いが生じ、音声成分を除去してしまうなどの不適切な雑音除去をなす場合がある。つまり雑音は一般的に、時間的に安定した確率分布をもって発生することは保証されていない。従って、音声区間外で求めた雑音パワースペクトルを用いて音声区間内の雑音パワースペクトルを除去する方法は、両者のパワースペクトルが異なった振幅値を持つ可能性があるため、雑音除去方法として適切ではない。

後者の方法では、信号入力器の配置方法により2つの入力信号のSN比が近い値を持つような場合に、除去する側の信号パワースペクトルに音声成分が混在するために、1入力による場合と同様に音声成分も除去するような不適切な除去結果を生じる。また、この方法では、除去する側の信号パワースペクトルに音声成分が混入することを防ぐ目的で、一方の信号入力器を発声者から離して配置する対処方法が取られるが、このときには2組の信号入力器に入力される雑音信号のパワースペクトルの間に時間相関性が成立しなくなり、除去しようとするパワースペクトルと実際の雑音パワースペクトルとの間に違いが生じ、不適切な雑音除去をなす場合がある。

この発明の目的は、雑音環境下で2つの音声入力器を用いて、従来より存在する1つの信号入力器による雑音除去装置および2つの信号入力器による雑音除去装置と比較して、時間変化のある雑音の入力に強く、また、音声成分の除去が生じにくく、雑音除去能力の高い、雑音除去装置を提供することにある。

「課題を解決するための手段」

この発明では、発声者の近くに位置する第1の信号入力器と、この第1の信号入力器に近接して位置する第2の信号入力器との2組の信号入力器を用い、それぞれの入力信号に対し短時間信号分析を行い信号パワースペクトルがスペクトル演算手段でそれぞれ求められ、周囲雑音のない場合に発声者によって音声が発声されたときの2つの入力信号のパワースペクトルの比率が比率記憶手段に予め記憶され、除去を目的とする周囲雑音のある場合に発声者によって音声が発声されたときの2つの入力信号のパワースペクトルの比率が比率演算手段で求められ、この比率と、先に記憶しておいたパワースペクトルの比率とを比較することにより、周囲雑音がある場合とない場合とでのパワースペクトルの比率の差分から信号パワースペクトル中の周囲雑音のパワースペクトル成分が雑音推定演算手段で推定され、この推定雑音パワースペクトル成分を発声者の近くに位置する信号入力器で受信したパワースペクトル成分から演算手段で除去されて、雑音を除去した音声パワースペクトルが得られる。

つまり、発声者の位置が固定で、周囲雑音が複数の信号入力器に対してほぼ同一に入力される場合、複数の信号入力器のうち、発声者に最も近い信号入力器に対する他の信号入力器のパワースペクトルの比率を観測すれば、雑音が多く存在する場合とあまり存在しない場合とではパワースペクトルの比率が異なるため、存在しない場合の比率と存在する場合の比率とを比較して入力信号中の雑音パワースペクトルを推定することが可能となり、この推定パワースペクトルを発声者に最も近い信号入力器のパワースペクトルから除去することにより雑音除去された音声パワースペクトルを得るものである。

「実施例」

第1図にこの発明の実施例を示す。マイクロホン1および2は、いずれも同一周波数帯域の音声信号を入力するものであり、マイクロホン1は雑

音を除去した音声信号を取り出すことを目的とした音声信号入力器であり、また、マイクロホン2は推定雑音パワースペクトルを生成し、これをマイクロホン1の信号パワースペクトルより除去することを目的とした音声信号入力器である。マイクロホン1および2の指向特性は、同一であり、例えば、単一指向性マイクロホンとする。この実施例では、マイクロホン1および2は、周囲雑音の雑音源より遠くに配置し、また、マイクロホンを発声者に近接配置したときはマイクロホンの位置が発声者からわずか遠くに移動しても音圧レベルの減少効果が大きいことを利用して雑音除去および音声区間検出を行うものであるため、マイクロホン1の音声対雑音比がマイクロホン2のそれよりも大きくなるようにする。すなわち、マイクロホン2よりもマイクロホン1の方を発声者3に近接させて配置するものとする。

発声者3と2つのマイクロホン1，2までの各距離およびマイクロホン1から2までの距離は、任意であるが、両マイクロホン1，2で受信する周波数帯域内の信号について2つのマイクロホン1，2間での位相のずれが無いように配置することが望ましい。例えば、発声者3からマイクロホン1までの距離を40cm、マイクロホン1から2までの距離を5cmとし、発声者3およびマイクロホン1，2の位置関係はこれらが一直線上に並ぶようにする。また、2つのマイクロホン1，2間およびこれらと発声者3との配置関係は発声者3の発声中には固定されているものとする。

A／D変換部4および5は、それぞれマイクロホン1および2の各出力アナログ波形信号を同一のサンプリング周波数でディジタル波形信号に変換するものである。このとき、サンプリング周波数はナイキストのサンプリング定理からマイクロホン1および2の受信する周波数帯域の上限値の2倍より大きいものとし、例えば、マイクロホン1および2の受信する周波数帯域が300Hzから4kHz とすると、A／D変換部4および5のサンプリング周波数は、例えば、10kHz とする。

区間検出部6は、A／D変換部4および5の出力波形を入力して、A／D変換部4および5の出力波形に音声を含むかどうかを検出するものである。スイッチ7および8は、区間検出部6の検出結果から音声区間と判断された場合にA／D変換部4および5の各出力をそれぞれスイッチ7および8を通じてスペクトル演算部9および10にそれぞれ供給する。

スペクトル演算部9および10は、それぞれA／D変換部4および5の出力波形から短時間パワースペクトルを計算する。パワースペクトルを計算する区間長（分析フレーム長）は、例えば、A／D変換部4および5のサンプリング周波数10kHz に対し256サンプル毎、すなわち25.6msec毎とし、パワースペクトルの計算には、例えば、高速フーリエ変換を使用する。また、スペクトル演算部9からは、短時間パワースペクトル計算時に分離したA／D変換部4の出力波形の位相情報を時間波形変換部11へ送る。

スペクトル補正部12はマイクロホン1および2とA／D変換部4および5とで生じたマイクロホン1と2との間のパワースペクトル形状の違いを補正するディジタルフィルタである。比率演算部13は、スペクトル演算部9および10で計算されたパワースペクトルの比率を計算するものであり、比率記憶部14は、比率演算部13の計算結果を記憶するものである。雑音推定演算部15は、比率演算部13の計算結果と比率記憶部14の記憶数値とから、スペクトル演算部10で計算されたマイクロホン2で受信される音声信号中に含まれる雑音パワースペクトルの割合を計算する。乗算部16は、雑音推定演算部15の計算結果を、スペクトル演算部10で計算したパワースペクトルに乗じ、推定雑音パワースペクトルを出力する。減算部17は、スペクトル演算部9で計算したパワースペクトルより乗算部16で出力される推定雑音パワースペクトルを除去する。時間波形変換部11は、減算部17の出力パワースペクトルとA／D変換部4の出力波形の位相情報とから雑音除去後の時間波形を得るものである。出力端子18は時間波形変換部11の出力が出力される出力端

である。

この実施例での動作を説明する。まず、マイクロホン2からの音声信号のパワースペクトルをマイクロホン1からの音声信号のパワースペクトルへ適合させるスペクトル補正部12のフィルタ係数は既に学習されて決定されているものとする。この学習方法は、例えば、除去を目的とする周囲雑音のある環境において、マイクロホン1および2により雑音を受信させ、スイッチ7および8によりA／D変換部4および5の出力をスペクトル演算部9および10にそれぞれ入力させ、これにより得られる短時間パワースペクトルの長時間平均値の比率を利用してスペクトル補正部12のフィルタ係数を学習させるものである。すなわち、スペクトル演算部9および10の各1フレーム当りの短時間パワースペクトルの100フレーム間の平均値の値をそれぞれ$MP_1(\omega)$、$MP_2(\omega)$とすると、スペクトル補正部12のフィルタ係数$FP(\omega)$は(1)式で求められる。

$$FP(\omega)=MP_1(\omega)/MP_2(\omega) \quad (1)$$

ここで、ωはフーリエ変換の計算ポイントであり、例えば、$0\leqq\omega\leqq5$kHzの範囲である。このようにしてスペクトル演算部10からスペクトル演算部9へのパワースペクトルの適合は、スペクトル演算部10の出力にフィルタ係数$FP(\omega)$を乗じて求める。

次に、周囲雑音がない状態で発声者3によって発声された音声のマイクロホン1，2間のパワースペクトルの比率を算出する。周囲雑音がない状態は、例えば、マイクロホン1および2を周囲雑音のない場所で使用することにより実現できる。発声者3が任意の音声を発声すると、音声はマイクロホン1および2により受信され、A／D変換部4および5に送られ、ディジタル波形信号として出力される。この波形から、区間検出部6により、音声かどうかが判定される。区間検出部6の波形検出アルゴリズムは、例えば、A／D変換部4および5の出力波形の短時間パワーをそれぞれK_1，K_2とするとき、その比率K_rを

$$K_r=K_1/K_2 \quad (2)$$

としたとき、K_rがあるしきい値K_{th}に対して、$K_r\geqq K_{th}$となる区間について音声区間であると判定するものである。この判定結果より、音声と判定された区間について、区間検出部6からの信号によりスイッチ7および8が閉じられ、A／D変換部4および5の各ディジタル波形出力は、スペクトル演算部9および10にそれぞれ送られる。スペクトル演算部9および10で計算されたパワースペクトルは、スペクトル演算部10の出力については、スペクトル補正部12による補正を受けて、それぞれ比率演算部13に送られる。比率演算部13においては、スペクトル演算部9および10から送られたパワースペクトル波形を一定時間長分、例えば、100フレーム分蓄積し、(3)式および(4)式のように比率を求める。スペクトル演算部9で計算されたパワースペクトルを$P_1(\omega)$、スペクトル演算部10で計算されスペクトル補正部12によって補正されたパワースペクトルを$P_2(\omega)$とし、$P_1(\omega)$および$P_2(\omega)$の各100フレーム間の平均値をそれぞれ$AP_1(\omega)$および$AP_2(\omega)$とすると、比率演算部13で求めるパワースペクトル比率$R(\omega)$および長時間パワースペクトル比率$AR(\omega)$は、

$$R(\omega)=P_1(\omega)/P_2(\omega) \quad (3)$$

$$AR(\omega)=AP_1(\omega)/AP_2(\omega) \quad (4)$$

となる。一般に、受信する音声信号のゆらぎにより、$R(\omega)$の値は$AR(\omega)$の値よりも安定せずにばらつく。この実施例では、パワースペクトル比率として、$R(\omega)$の代わりに$AR(\omega)$を採用する。$AR(\omega)$は比率記憶部14へ送られて記憶される。

次に、除去を目的とする周囲雑音がある状態で発声者3によって発声された音声のマイクロホン1，2間のパワースペクトルの比率を算出する。このときの2つのマイクロホン1，2および発声者3の配置関係は上記の周囲雑音がない状態で発声者3に発声してもらった場合と同一である。上記の周囲雑音がない状態での発声と同様に、発声者3が任意の音声を発声すると、音声はマイクロホン1および2により受信され、A／D変換部4

および5によりディジタル波形信号として出力される。この波形から、区間検出部6により、音声を含む区間かそうでないかが判定され、音声と判定された区間の波形は、スイッチ7および8を介して、スペクトル演算部9および10に送られる。スペクトル演算部9および10で計算されたパワースペクトルは、スペクトル演算部10の出力についてのみスペクトル補正部12による補正を受けて、それぞれ比率演算部13に送られる。

比率演算部13では、上記の場合と同様に、スペクトル演算部9の出力およびスペクトル補正部12の出力をそれぞれ$Pn_1(\omega)$、$Pn_2(\omega)$としたとき、雑音のある場合のパワースペクトル比率$Q(\omega)$を(5)式により求める。

$$Q(\omega)=Pn_1(\omega)/Pn_2(\omega) \quad (5)$$

次に、$R(\omega)$および$Q(\omega)$の計算結果をもとに、雑音推定演算部15により、$Pn_2(\omega)$における雑音成分を推定する。

マイクロホン1および2で観測される雑音のパワースペクトルをそれぞれ$N_1(\omega)$および$N_2(\omega)$、また、音声のパワースペクトルをそれぞれ$S_1(\omega)$および$S_2(\omega)$とすると、$Pn_1(\omega)$および$Pn_2(\omega)$は、

$$Pn_1(\omega)=S_1(\omega)+N_1(\omega) \quad (6)$$

$$Pn_2(\omega)=S_2(\omega)+N_2(\omega) \quad (7)$$

となる。ここで、スペクトル補正部12により、マイクロホン1，2間で雑音に対するパワースペクトル補正を行っていることから、

$$N_1(\omega)=N_2(\omega) \quad (8)$$

の関係が成立する。また、比率演算部13で求めるパワースペクトル比率$R(\omega)$と長時間パワースペクトル比率$AR(\omega)$とにおいて、パワースペクトル比率が音声信号に依存せず、発声者3とマイクロホン1および2との位置関係だけに依存して決定されると仮定すれば、

$$R(\omega)=AR(\omega) \quad (9)$$

の関係が成立する。さらに(3)式においてマイクロホン1および2に入力される信号は発声者3による音声信号が主で周囲雑音信号が無視できるほど小さいものと仮定すれば、$R(\omega)$は、

$$R(\omega)=S_1(\omega)/S_2(\omega) \quad (10)$$

と書き換えられる。そこで、(8)式，(9)式，(10)式の関係を(6)式，(7)式に代入し、また、

$$X(\omega)=S_2(\omega)/N_2(\omega) \quad (11)$$

なる$X(\omega)$を定義すると、(6)式および(7)式は次のように書き換えられる。

$$\begin{aligned}Pn_1(\omega)&=S_1(\omega)+N_1(\omega)\\&=AR(\omega)\,S_2(\omega)+N_2(\omega)\\&=\{AR(\omega)\,X(\omega)+1\}N_2(\omega)\end{aligned} \quad (12)$$

$$\begin{aligned}Pn_2(\omega)&=S_2(\omega)+N_2(\omega)\\&=\{X(\omega)+1\}N_2(\omega)\end{aligned} \quad (13)$$

従って、$Q(\omega)$は、

$$Q(\omega)=Pn_1(\omega)/Pn_2(\omega)=\frac{AR(\omega)\,X(\omega)+1}{X(\omega)+1} \quad (14)$$

となる。(14)式を変形すると、

$$X(\omega)=\frac{Q(\omega)-1}{AR(\omega)-Q(\omega)} \quad (15)$$

となり、また、(13)式から、

$$N_2(\omega)=G(\omega)\,Pn_2(\omega) \quad (16)$$

ただし、

$$G(\omega)=\frac{1}{X(\omega)+1} \quad (17)$$

が得られる。よって、(15)式，(16)式および(17)式によりマイクロホン2の入力信号のパワースペクトル$Pn_2(\omega)$から、比率記憶部14に記憶した$AR(\omega)$および現に入力されたマイクロホン1，2の出力のパワースペクトルの比率$Q(\omega)$を使用して、その時、入力された雑音信号パワースペクトル$N_2(\omega)$が推定されたことになる。ただし、マイクロホン1および2の位置関係から以下の不等式が成立することが必要である。

$$AR(\omega)\geq Q(\omega)\geq 1 \quad (18)$$

理想的には、全てのωについて(18)式は成立する。ところが、実際には雑音成分のゆらぎにより、ωによっては(18)式が成立しない場合がある。よって、これらの方程式をもとに、雑音推定演算部15では、(18)式が成立しない場合の近似推定を含めて、以下のようなアルゴリズムで(15)式および

(17)式の推定を行う。まず、高速フーリエ変換の計算ポイントω毎に(18)式が成立するかどうかを確認する。次に、(18)式が成立するものについて、(15)式の分母である、

$$M(\omega) = AR(\omega) - Q(\omega) \quad (19)$$

を計算し、$M(\omega) = 0$となるωについては、$G(\omega) = 0$を代入し、$M(\omega) > 0$となるωについては(15)式および(17)式を計算して$G(\omega)$を得る。また、雑音パワースペクトルのゆらぎによって(18)式が成立しないωについては、同一フレームのパワースペクトル上で、(18)式が成立するωについて$Q(\omega)$の値を累計し、その平均値Q_{av}を求める。ここで、

$$M_{av}(\omega) = AR(\omega) - Q_{av} \quad (20)$$

を求め、$M_{av}(\omega) > 0$となるωについては、

$$X(\omega) = \frac{Q(\omega) - 1}{M_{av}(\omega)} \quad (21)$$

を求め、近似的な$X(\omega)$を算出し、(17)式に代入して$G(\omega)$を得る。また、$M_{av}(\omega) \leq 0$となるωについては$G(\omega) = 0$とする。このようにして高速フーリエ変換の計算ポイントω毎に$G(\omega)$を算出し、最後に、事前に設定したしきい値G_{th}に対し$G(\omega) > G_{th}$となる$G(\omega)$について、$G(\omega) = G_{th}$として、雑音パワースペクトルのゆらぎにより$G(\omega)$が巨大な値をとる場合を制限して、雑音推定演算部15での$G(\omega)$の計算を完了する。

この$G(\omega)$は雑音推定演算部15から乗算部16へ送られる。乗算部16では、$G(\omega)$をもとに(16)式の計算が行われ、推定雑音パワースペクトル$N_2(\omega)$が算出される。この算出結果は減算部17へ送られ、スペクトル演算部9で計算したパワースペクトル$P_{n_1}(\omega)$より推定雑音パワースペクトル$N_2(\omega)$が減算され、つまり、

$$P_{s_1}(\omega) = P_{n_1}(\omega) - N_2(\omega) \quad (22)$$

なる演算が行われ、雑音除去された信号パワースペクトル$P_{s_1}(\omega)$を得る。この$P_{s_1}(\omega)$は、時間波形変換部11へ送られ、先にスペクトル演算部9より伝送された位相情報を利用して逆フーリエ変換されて、雑音除去後のディジタルの時間波形信号として出力端子18より得る。

音声認識装置などで、特に元の時間波形信号に戻す必要がない場合は、雑音除去後の音声パワースペクトル$P_{s_1}(\omega)$を得るようにしてもよい。

以降、音声区間についてはこれらの雑音除去操作が行われ、出力端子18より雑音除去後の音声波形を得る。

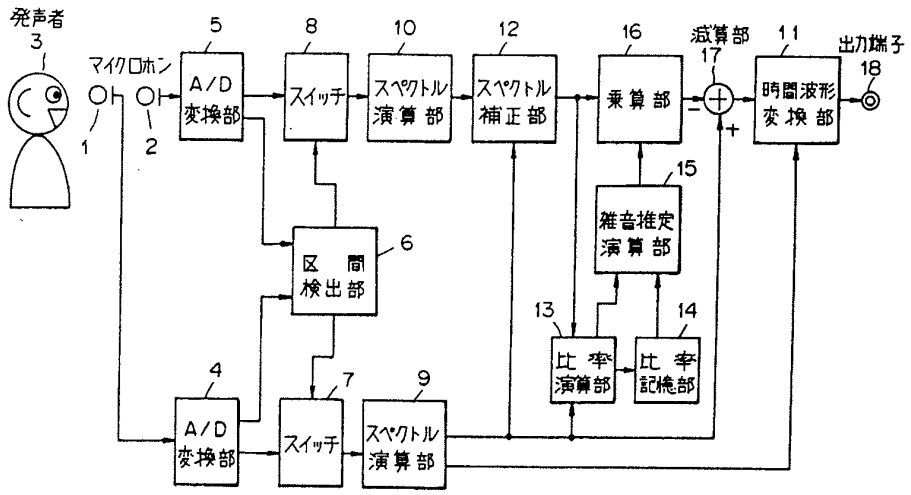
「発明の効果」

以上述べたように、この発明では、①2組の入力器によって雑音の混在する音声を入力することにより、音声に重畳する雑音パワースペクトルと同一の時間変化をする雑音パワースペクトルを推定することが容易となる、②1組の入力器のみ使用した従来の雑音除去方式よりも雑音の時間変動に対して追従性が良くなり、音声パワースペクトルの成分を誤って除去するなどの不適切な雑音除去を防ぐことができる、効果を有する。

4．図面の簡単な説明

第1図はこの発明の実施例を示すブロック図である。

第1図


発声者
3
マイクロホン
1
2
A/D
変換部
5
スイッチ
8
スペクトル
演算部
10
スペクトル
補正部
12
乗算部
16
減算部
17
時間波形
変換部
11
出力端子
18
雑音推定
演算部
15
区間
検出部
6
比率
演算部
13
比率
記憶部
14
A/D
変換部
4
スイッチ
7
スペクトル
演算部
9

DERWENT-ACC-NO: 1994-346040
DERWENT-WEEK: 200122
COPYRIGHT 2007 DERWENT INFORMATION LTD
TITLE: Noise suppression equipment for microphone - uses digital filters and adaptive noise cancellation circuit to determine noise level

PATENT-ASSIGNEE: SONY CORP[SONY]

PRIORITY-DATA: 1993JP-0076371 (March 10, 1993)

PATENT-FAMILY:
PUB-NO PUB-DATE LANGUAGE PAGES MAIN-IPC JP 06269083 A September 22, 1994 N/A
011
H04R 003/00

JP 3154151 B2
April 9, 2001
N/A
012
H04R 003/00

APPLICATION-DATA:
PUB-NO APPL-DESCRIPTOR APPL-NO APPL-DATE JP 06269083A N/A
1993JP-0076371
March 10, 1993

JP 3154151B2
N/A
1993JP-0076371
March 10, 1993

JP 3154151B2
Previous Publ.
JP 6269083
N/A

INT-CL (IPC): G10L003/02, G10L015/20 , G10L021/02 , H03H017/02 , H03H017/04 , H03H021/00 , H04R001/40 , H04R003/00

ABSTRACTED-PUB-NO: JP 06269083A

BASIC-ABSTRACT:

The equipment consists of two microphones (11, 12) whose outputs are digitised using A/D convertors (13, 14). The output of these convertors are passed through a subtraction circuit (15) to get a differential signal. The output of the first A/D convertor is passed through a band pass filter (22). The differential signal is passed through another band pass filter (23). These outputs serve as main input and reference input to an adaptive noise cancellation circuit (10).

The output of the first A/D convertor is passed through a high pass filter (21) whose cut-off frequency is fixed in advance. The outputs of this filter and the noise cancellation circuit are passed through an adder (24). The output of the adder is passed through a D/A convertor (16) and the audio signal without wind noise is obtained at the output terminal (17).

USE/ADVANTAGE - E.g for use in carried in built-in-camera VTR and in recording and reproduction of audio signal. Reduces wind noise.

CHOSEN-DRAWING: Dwg.2/8
TITLE-TERMS: NOISE SUPPRESS EQUIPMENT MICROPHONE DIGITAL FILTER ADAPT NOISE CANCEL CIRCUIT DETERMINE NOISE LEVEL

DERWENT-CLASS: P86 U22 W04

EPI-CODES: U22-G01A5; U22-G01D; W04-B10; W04-G03; W04-M01K; W04-M01X;

SECONDARY-ACC-NO:
Non-CPI Secondary Accession Numbers: N1994-271812

MICROPHONE EQUIPMENT

Patent number: JP6269083

Publication date: 1994-09-22

Inventor: SASAKI TORU; OKUBO HITOSHI; MIZUUCHI TAKAYUKI; GYOTOKU KAORU

Applicant: SONY CORP

Classification:

- international: ***G10L15/20; G10L21/02; H03H17/00; H03H17/02; H03H17/04; H03H21/00; H04B3/04; H04R1/40; H04R3/00;*** **G10L15/00; G10L21/00; H03H17/00; H03H17/02; H03H17/04; H03H21/00; H04B3/04; H04R1/40; H04R3/00;** (IPC1-7): H04R3/00; G10L3/02; H03H17/02; H03H17/04; H03H21/00; H04R1/40

- european:

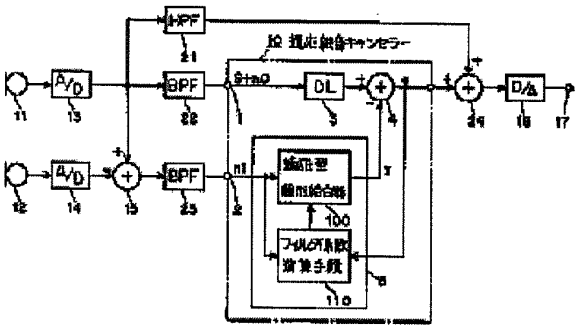
Application number: JP19930076371 19930310

Priority number(s): JP19930076371 19930310

Report a data error here

Abstract of **JP6269083**

PURPOSE:To obtain a microphone equipment capable of obtaining excellent sound collection quality while sufficiently reducing wind noise. CONSTITUTION:First and second microphone elements 11, 12 are arranged close to each other. A difference output of the 1st and 2nd microphone elements is obtained from a subtractor circuit 15. A component of a predetermined band whose frequency is a predetermined frequency or over in the output signal of the 1st microphone element 11 is extracted by a 1st filter 21. A component of a predetermined band whose frequency is below a predetermined frequency in output signals of the 1st microphone element 11 is extracted by a 2nd filter and the extracted signal component is used for a major input signal and it is fed to an adaptive noise canceller 10. A component for a predetermined band whose frequency is below a predetermined frequency in the difference output from the subtractor circuit 15 is extracted by a 3rd filter and the extracted signal component is applied to the adaptive noise canceller 10 as a reference input signal. The output signal of the 1st filter and the output signal of the adaptive noise canceller 10 are added by an adder circuit 24 and the sum is led out at an output terminal 17.



Data supplied from the ***esp@cenet*** database - Worldwide

(19)日本国特許庁（ＪＰ）　(12)公開特許公報(A)　(11)特許出願公開番号

特開平6－269083

(43)公開日　平成6年(1994)9月22日

(51)Int.Cl.[5]		識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｈ０４Ｒ	3/00	３２０	7346－5H		
Ｇ１０Ｌ	3/02	３０１　Ｅ	9379－5H		
Ｈ０３Ｈ	17/02	Ｌ	7037－5Ｊ		
	17/04	Ａ	7037－5Ｊ		
	21/00		7037－5Ｊ		

審査請求　未請求　請求項の数5　ＦＤ　(全 11 頁)　最終頁に続く

(21)出願番号　特願平5－76371

(22)出願日　平成5年(1993)3月10日

(71)出願人　000002185
ソニー株式会社
東京都品川区北品川６丁目７番35号

(72)発明者　佐々木　徹
東京都品川区北品川６丁目７番35号　ソニー株式会社内

(72)発明者　大久保　仁
東京都品川区北品川６丁目７番35号　ソニー株式会社内

(72)発明者　水内　崇行
東京都品川区北品川６丁目７番35号　ソニー株式会社内

(74)代理人　弁理士　佐藤　正美

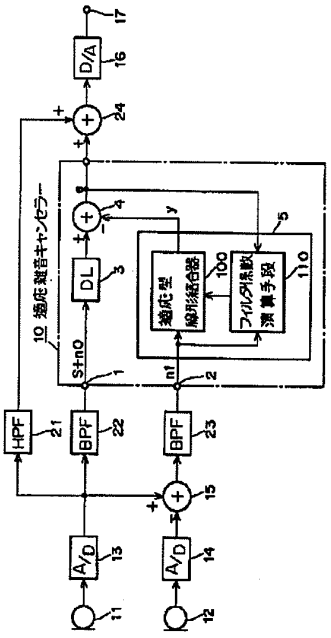
最終頁に続く

(54)【発明の名称】　マイクロホン装置

(57)【要約】

【目的】　風雑音を十分に低減しながら、良好な収音品質が得られるマイクロホン装置を実現する。

【構成】　第１及び第２のマイクロホン素子１１及び１２を近接して配置する。第１及び第２のマイクロホン素子の出力の差分出力を、減算回路１５から得る。第１のマイクロホン素子１１の出力信号の内の所定周波数以上の所定帯域の成分を第１のフィルタ２１で抽出する。第１のマイクロホン素子１１の出力信号の内の所定周波数以下の所定帯域の成分を第２のフィルタで抽出し、抽出した信号成分を主要入力信号として適応雑音キャンセラー１０に供給する。減算回路１５からの差分出力の内の所定周波数以下の所定帯域の成分を第３のフィルタで抽出し、抽出した信号成分を参照入力信号として適応雑音キャンセラー１０に供給する。第１のフィルタの出力信号と、適応雑音キャンセラー１０の出力とを加算回路２４で加算して、出力端子１７に導出する。



1

【特許請求の範囲】

【請求項1】　近接して配置した第1及び第2のマイクロホン素子と、

上記第1及び第2のマイクロホン素子の出力の差分出力を得る手段と、

主要入力信号に含まれる雑音成分を、参照入力信号に基づいて適応的に低減するようにする適応雑音キャンセラーと、

上記第1のマイクロホン素子の出力信号の内の所定周波数以上の所定帯域の成分を抽出する第1のフィルタと、

上記第1のマイクロホン素子の出力信号の内の所定周波数以下の所定帯域の成分を抽出し、抽出した信号成分を上記主要入力信号として上記適応雑音キャンセラーに供給する第2のフィルタと、

上記第1及び第2のマイクロホン素子の差分出力の内の所定周波数以下の所定帯域の成分を抽出し、抽出した信号成分を上記参照入力信号として上記適応雑音キャンセラーに供給する第3のフィルタと、

上記第1のフィルタの出力信号と、上記適応雑音キャンセラーの出力とを加算して、装置出力とする加算手段とを備えるマイクロホン装置。

【請求項2】　請求項1に記載のマイクロホン装置において、

上記第3フィルタの出力信号のレベルに応じて、上記適応雑音キャンセラーの適応動作を制御することを特徴とするマイクロホン装置。

【請求項3】　請求項1に記載のマイクロホン装置において、

上記第1のフィルタの出力信号のレベルに応じて、上記適応雑音キャンセラーの適応動作を制御することを特徴とするマイクロホン装置。

【請求項4】　請求項1に記載のマイクロホン装置において、

上記第3フィルタの出力信号のレベルと、上記第1のフィルタの出力信号のレベルとの比に応じて、上記適応雑音キャンセラーの適応動作を制御することを特徴とするマイクロホン装置。

【請求項5】　請求項1に記載のマイクロホン装置において、

上記第2のフィルタは、上記第1のマイクロホン素子の出力信号の内の所定周波数以下の帯域を複数帯域に分割して、各分割帯域毎の成分を抽出するものであり、

上記第2のフィルタにより抽出された各分割帯域の信号成分は、各分割帯域に対応して設けられた複数の適応雑音キャンセラーに、それぞれ主要入力信号として供給され、

上記第3のフィルタは、上記第1及び第2のマイクロホン素子の差分出力の内の所定周波数以下の帯域を上記複数帯域に分割して、各分割帯域毎の成分を抽出するものであり、

2

上記第3のフィルタにより抽出された各分割帯域の信号成分は、上記複数の適応雑音キャンセラーに、それぞれ参照入力信号として供給され、

上記第1のフィルタの出力信号と、上記複数の適応雑音キャンセラーの複数の出力とを上記加算手段において加算するようにしたことを特徴とするマイクロホン装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、例えば、カメラー体型ＶＴＲに搭載するマイクロホンやスピーチ／歌唱用マイクロホンなど広範囲に利用できるマイクロホン装置に関し、特に、自然風、呼気、空調などの気流にさらされるマイクロホン装置に大きな効果を発揮するものである。

【0002】

【従来の技術】マイクロホンは、音波による音圧の変化を振動板の機械的な振動に変換し、この機械的振動に基づき電気音響変換系を動作させて、電気信号を得るようにするものが多い。このため、マイクロホンで希望音声を収音する場合、希望音声以外の何等かの要因によって振動板に機械的振動が与えられると、これが希望音声に対しては、雑音となる。この場合に、上記の要因が風であれば、風による雑音（以下、これを風雑音と称する）が発生する。

【0003】マイクロホンにおいて発生する風雑音を低減する方法としては、

（1）ウィンドスクリーン（風防）の使用

（2）電気的／音響的ハイパスフィルタの使用

（3）低音域で無指向性を示す構成の採用

などが、従来、良く用いられている。

【0004】

【発明が解決しようとする課題】しかしながら、上記（1）の方法は、一般に、ウィンドスクリーンの外形寸法が大きいほど、また、マイクロホン素子とウィンドスクリーン内壁との距離が大きいほど、風雑音は小さくなるので、充分に風雑音を低減させようとすれば、ウィンドスクリーンは大きなものを用意しなくてはならないという問題がある。この問題は、機器の小型化や携帯性に大きな支障になる。

【0005】また、風雑音は低音域主体の雑音であるから、上記（2）の固定的な電気的／音響的ハイパスフィルタを使用する方法は有効であるが、風雑音を低減させようとすると、同時に希望音源の低音域をも低減させてしまい、収音品質を低下させてしまう欠点がある。

【0006】さらに、上記（3）の方法は、マイクロホンの指向性が有指向性であるよりも無指向性である方が、風雑音が低いという事実から使用されている。しかし、この方法による風雑音の低減効果はあまり大きくはなく、その上、実際にはマイクロホン装置を構成する際の筐体の影響などもあって、風雑音が充分低いレベルに

ならない。

【0007】出願人は、以上の問題点を回避して、風雑音を十分に低減することができるようにする方法を、特願平3－349274号として、先に、提案している。この方法は、適応雑音キャンセラーを応用したものである。

【0008】図8は、この先に提案したマイクロホン装置の一例の構成を示すブロック図で、同図において、10は、適応雑音キャンセラーである。先ず、この適応雑音キャンセラーについて説明する。

【0009】適応雑音キャンセラー10において、1は主要入力端子、2は参照入力端子であって、主要入力端子1を通じて入力された主要入力信号は遅延回路3を介して合成回路4に供給される。また、参照入力端子2を通じて入力された参照入力信号は、適応フィルタ回路5を介して合成回路4に供給され、遅延回路3からの信号から減算される。この合成回路4の出力は、適応フィルタ回路5に帰還されると共に、出力端子6に導出される。

【0010】この適応雑音キャンセラーにおいては、主要入力信号としては、希望信号sと、これと無相関の雑音信号n0とが加算されたものが入力される。一方、参照入力信号としては、雑音信号n1が入力される。参照＊

$$E[e^2]=E[s^2]+E[(n0-y)^2]+2E[s(n0-y)]$$
$$=E[s^2]+E[(n0-y)^2]$$

となる。適応フィルタ回路5が収束するものとすれば、適応フィルタ回路5は、$E[e^2]$が最小になるよう※

$$Emin[e^2]=E[s^2]+Emin[(n0-y)^2]$$

となる。

【0015】すなわち、$E[e^2]$が最小化されることによって$E[(n0-y)^2]$が最小化され、適応フィルタ回路5の出力yは、雑音信号n0の推定量になる。そして、合成回路4からの出力の期待値は、希望信号sのみとなる。すなわち、適応フィルタ回路5を調整して全出力パワーを最小化することは、減算出力eが、希望音声信号sの最小二乗推定値になることに等しい。

【0016】なお、適応フィルタ回路5はアナログ信号処理回路で実現する場合とデジタル信号処理回路で実現する場合の、いずれでも可能であるが、一般的には、DSP（デジタルシグナルプロセッサ）を用いたデジタル処理回路の構成とされる。

【0017】図8の例では、以上説明した適応雑音キャンセラー10を用いて風雑音を低減するマイクロホン装置を次のようにして実現している。すなわち、図8のマイクロホン装置においては、互いに近接して2個のマイクロホン素子11及び12を配置する。これらマイクロホン素子11、12は同特性のものが用いられ、例えば無指向性マイクロホンユニットが用いられる。

【0018】そして、一方のマイクロホン素子11の出力がA／Dコンバータ13によりデジタル信号に変換さ

＊入力の雑音信号n1は、希望信号sとは無相関であるが、雑音信号n0とは相関があるようにされている。

【0011】適応フィルタ回路5は、参照入力雑音信号n1をフィルタリングして、雑音信号n0に近似する信号yを出力する。この場合は、適応フィルタ回路5においては、所定の適応のアルゴリズムにより、合成回路4の減算出力（残差出力）eが最小になるように、参照入力雑音信号n1のフィルタリングのフィルタ係数を更新してゆく。

【0012】この適応フィルタ回路5の出力信号yとして、雑音信号n0と逆相、等振幅の信号を得るようにすることもできる。遅延回路3は、適応フィルタ回路5での演算処理に要する時間遅れや適応フィルタでの伝播時間その他を補償して、減算処理する信号との時間合わせをするためのものである。

【0013】以下に、適応雑音キャンセラーの原理について説明する。

【0014】今、希望信号s，雑音n0，雑音n1，出力信号yが統計的に定常であり、平均値が0であると仮定すると残差出力eは、

$$e=s+n0-y$$

となる。これを二乗したものの期待値は、希望信号sが雑音n0及び出力yと無相関であるから、

※に、適応フィルタ係数を更新するものである。このとき、$E[s^2]$は影響を受けないので、

れ、そのデジタル信号が主要入力信号として、適応雑音キャンセラー10の主要入力端子1に供給される。また、他方のマイクロホン素子12の出力がA／Dコンバータ14によりデジタル信号に変換され、そのデジタル信号が減算回路15に供給される。

【0019】この減算回路15においては、A／Dコンバータ13からのマイクロホン素子11の出力のデジタル信号と、この他方のマイクロホン素子12の出力のデジタル信号との差分が求められ、その差分出力が参照入力信号として、適応雑音キャンセラー10の参照入力端子2に供給される。

【0020】そして、適応雑音キャンセラー10の出力端子6に得られる信号は、D／Aコンバータ16によりアナログ信号に戻され、マイクロホン装置の出力信号として出力端子17に導出される。

【0021】この図8のマイクロホン装置の風雑音の低減動作について説明する。このマイクロホン装置により、風雑音が生じるような環境において、収音を行った場合、マイクロホン素子11及び12の出力は、収音音声信号に風雑音が含まれるものとなる。

【0022】上述したように、2個のマイクロホン素子11及び12は、近接して配置されているので、音声

は、これら2個のマイクロホン素子11、12に非常に相関の強い状態で収音される。一方、風によりマイクロホン素子11とマイクロホン素子12とに発生する風雑音は、それぞれのマイクロホン素子に固有となるため、両マイクロホン素子の風雑音の間に相関はない。

【0023】したがって、減算回路15において、2個のマイクロホン素子11及び12の出力の差分演算が行われると、音声信号はキャンセルされ、減算回路15からは風雑音のみの成分が得られる。図8の構成において、主要入力として適応雑音キャンセラー10に供給されるマイクロホン11の出力信号中の風雑音成分と、減算回路15からの風雑音成分とは相関があり、この減算回路15からの風雑音の成分が上述したように、適応雑音キャンセラー10の参照入力となっているので、出力端子17には、主要入力中の風雑音が適応雑音キャンセラー10により十分に低減される。

【0024】以上説明したように、図8の構成によれば、上記のようにして風雑音を低減することができるが、音声入力に及ぼす歪みの発生が問題として残る。すなわち、前述したように、適応雑音キャンセラーにおける希望入力信号は不要雑音信号と無相関であることが前提である。しかし、実際にはマイクロホンの風雑音と希望音声の低域成分とは完全に無相関とはならず、結果的に希望音声信号が無歪みで出力されることはない。

【0025】具体的には、変調歪みのような非線形歪みが音声信号出力に付加されたり、音声信号レベルが低減されてしまうなどの現象として現われる。特に音声信号レベルが相対的に大きい場合、あるいは適応フィルタを例えばデジタルＦＩＲフィルタで構成したときにそのタップ数が小さい場合に、顕著である。

【0026】風雑音の低減効果と、希望音声への歪みの影響とは、いわゆるトレードオフの関係にあり、風雑音を十分に抑圧しようとすると、適応雑音低減処理により、主要入力中の希望音声に大きな歪が加わり、良好な収音品質にするという目的に反してしまう。また、希望音声への歪みを抑えようとすると、風雑音の低減効果が低くなり、やはり良好な収音品質は得にくくなる。

【0027】この発明は、上記の問題点を解消して、風雑音を十分に低減しながら、良好な収音品質を得ることができるマイクロホン装置を提供することを目的とする。

【0028】

【課題を解決するための手段】上記課題を解決するために、この発明によるマイクロホン装置は、基本的には適応雑音キャンセラーを応用して風雑音を低減するものであるが、風雑音が低音域主体の雑音であることに鑑み、適応雑音キャンセラーを主要入力の低音域に対してのみ働かせるようにすることを特徴とする。

【0029】すなわち、この発明によるマイクロホン装置は、後述する実施例の参照符号を対応させると、近接して配置した第1及び第2のマイクロホン素子11及び12と、上記第1及び第2のマイクロホン素子の出力の差分出力を得る手段15と、主要入力信号に含まれる雑音成分を、参照入力信号に基づいて適応的に低減するようにする適応雑音キャンセラー10と、上記第1のマイクロホン素子11の出力信号の内の所定周波数以上の所定帯域の成分を抽出する第1のフィルタ21と、上記第1のマイクロホン素子11の出力信号の内の所定周波数以下の所定帯域の成分を抽出し、抽出した信号成分を上記主要入力信号として上記適応雑音キャンセラー10に供給する第2のフィルタ22と、上記第1及び第2のマイクロホン素子11及び12の差分出力の内の所定周波数以下の所定帯域の成分を抽出し、抽出した信号成分を上記参照入力信号として上記適応雑音キャンセラー10に供給する第3のフィルタ23と、上記第1のフィルタ21の出力信号と、上記適応雑音キャンセラー10の出力とを加算して、装置出力とする加算手段24とを備える。

【0030】上記第1～第3のフィルタ21～23は、適応動作を風雑音にある程度限定するためのもので、それぞれバンドパスフィルタで構成することができる。また、第1のフィルタ21は、その目的からハイパスフィルタでも良い。第2及び第3のフィルタ22及び23は、その目的からローパスフィルタで構成するようにすることができる。

【0031】

【作用】マイクロホンの風雑音は、低音域例えば数100Ｈｚ以下を主成分とする雑音である。一方、音声信号は約100Ｈｚ以上に分布している。上記の構成のこの発明によれば、第1のフィルタにより第1のマイクロホン素子11により収音された音声信号の内の例えば数100Ｈｚ以上の中高域の周波数成分が抽出される。上記の風雑音の性質から第1のフィルタの出力には、風雑音の成分は殆ど含まれない。この第1のフィルタ21の出力は、そのまま加算手段24に供給される。

【0032】一方、第2のフィルタ22からは、第1のマイクロホン素子11により収音された音声信号の低域成分と、風雑音の主成分が得られ、これが適応雑音キャンセラー10の主要入力とされる。同様にして、第3のフィルタ23からは、第1及び第2のマイクロホン素子11及び12の出力の差分の低域成分、すなわち、風雑音の主成分が得られ、これが適応雑音キャンセラー10の参照入力とされる。

【0033】この状態で、適応動作させられると、適応雑音キャンセラー10の出力信号としては、風雑音成分がキャンセルされた収音音声の低域成分が得られる。加算手段では、この低域成分と、適応処理の影響を全く受けていない収音音声の中高域成分とが加算される。この加算出力が装置出力となるので、風雑音が十分に低減され、かつ、品質の良い収音音声信号が得られる。

7

【0034】

【実施例】以下、この発明によるマイクロホン装置の一実施例を、図を参照しながら説明する。図1は、この発明によるマイクロホン装置の一実施例を示すブロック図で、前述した図8の例と対応する部分には同一符号を付してある。

【0035】ところで、前述もしたように、マイクロホンの風雑音は数100Hz以下を主成分とする雑音である。一方、音声信号は約100Hz以上に分布している。図2に、風雑音と、音声信号の平均パワースペクトラムの一例を示す。図2において、実線31は風速2m／secのときの風雑音のパワースペクトラム、実線32は風速4m／secのときの風雑音のパワースペクトラム、実線33は男性話声のパワースペクトラムである。

【0036】この図2から、風雑音の低減は数100Hz以下について行われればよく、数100Hz以上の周波数成分については何ら処理の対象とせずに出力すればよいことが分かる。

【0037】このため、図1の例においては、近接して配置された1対のマイクロホン素子11及び12の内の一方のマイクロホン素子11の出力は、A／Dコンバータ13によりデジタル信号に変換された後、ハイパスフィルタ21に供給される。このハイパスフィルタ21は、カットオフ周波数が例えば700Hzのものが使用され、マイクロホン素子11の出力信号のうち、風雑音が殆ど含まれない中高域成分を抽出する。

【0038】そして、このハイパスフィルタ21の出力信号は、適応雑音キャンセラー10の出力端子6とD／Aコンバータ16との間に設けられた加算回路24に供給される。なお、ハイパスフィルタ21の代わりに、必要音声帯域以上の周波数帯域を通過帯域としないバンドパスフィルタを用いても良い。

【0039】A／Dコンバータ13の出力信号は、また、バンドパスフィルタ22に供給される。このバンドパスフィルタ22は、例えば100Hz～700Hzを通過帯域とするもので、マイクロホン素子11の低域成分と、風雑音の主成分を通過させる。なお、バンドパスフィルタ22の通過帯域の上側の周波数は、ハイパスフィルタ21のカットオフ周波数に必ずしも一致させなくてもよい。このバンドパスフィルタ22の出力信号は、主要入力として適応雑音キャンセラー10の主要入力端子1に供給される。

【0040】また、減算回路15からのマイクロホン素子11とマイクロホン素子12の差分信号のデジタル信号は、バンドパスフィルタ22と同様な特性のバンドパスフィルタ23に供給されて、風雑音の成分のみが抽出され、これが適応雑音キャンセラー10の参照入力端子2に入力される。

【0041】適応雑音キャンセラー10の適応フィルタ

8

回路5は、図3にも示すように、FIRフィルタ型の適応型線形結合器100と、フィルタ係数演算回路110からなっている。この適応フィルタ回路5は、マイクロコンピュータを搭載するDSPにより、ソフトウエアとして構成することができる。フィルタ係数の更新のアルゴリズムは、この例では、計算量が少なく、実用的であるため多用されているLMS（最小平均自乗）法を用いる。

【0042】LMS法について、図3を参照しながら説明する。図3に示すように、適応型線形結合器100は、それぞれ単位サンプリング時間の遅延時間Z^{-1}を有する複数個の遅延回路DL1，DL2，……DLm（mは正の整数）と、入力雑音n1及び各遅延回路DL1，DL2，……DLmの出力信号と加重係数（フィルタ係数）との掛け算を行う加重回路MX0，MX1，MX2，……MXmと、加重回路MX0～MXmの出力を加算する加算回路101を備える。加算回路101の出力は、図8の例で説明した信号yである。

【0043】加重回路MX0～MXmに供給する加重係数は、フィルタ係数演算回路110で、LMSアルゴリズムにより、合成回路4からの残差信号eと、参照入力n1とに基づいて形成される。このフィルタ係数演算回路110で実行されるアルゴリズムは、次のようになる。

【0044】今、時刻kにおける入力ベクトルX_kを、図3にも示すように、

$X_k = [x_{0k}\ \ x_{1k}\ \ x_{2k}\ \ \cdots\ x_{mk}]^T$

とし、出力をy_k、加重係数をw_{jk}（j=0,1,2,…m）とすると、入出力の関係は、次の数1に示すように、

【0045】

【数1】

$$y_k = \sum_{j=0}^{m} w_{jk} x_{jk}$$

となる。

【0046】そして、時刻kにおける加重ベクトルW_kを、

$W_k = [w_{0k}\ \ w_{1k}\ \ w_{2k}\ \ \cdots\ w_{mk}]^T$

と定義すれば、入出力関係は、

$y_k = X_k^T \cdot W_k$

で与えられる。ここで、希望の応答をd_kとすれば、残差e_kは次のように表される。

$e_k = d_k - y_k$

$= d_k - X_k^T \cdot W_k$

LMS法では、加重ベクトルの更新を、

$W_{k+1} = W_k + 2\mu \cdot e_k \cdot X_k$ … (1)

なる式（1）により順次行っていく。ここで、μは適応の速度と安定性を決める利得因子（ステップゲイン）である。

【００４７】以上説明したような図１の構成においては、適応雑音キャンセラー１０では、マイクロホン素子１１で収音された音声のうち、風雑音の主成分が含まれている低音域のみについて、主要入力であるマイクロホン素子１１の出力中の風雑音をキャンセルするように適応処理を行う。その結果、適応雑音キャンセラー１０の出力端子６には、風雑音が低減除去されたマイクロホン素子１１の出力の低域成分が得られる。

【００４８】そして、加算回路２４において、この適応雑音キャンセラー１０の出力信号と、風雑音成分を殆ど含まないマイクロホン素子１１の出力の中高域成分とが加算されるので、出力端子１７には、風雑音成分が低減されたマイクロホン素子１１の収音音声信号が得られるものである。

【００４９】なお、図１の例の場合においても、音声のレベルや適応動作の設定によっては歪みが生じることもあるが、マイクロホン素子１１の信号の中高域成分は適応処理の影響を全く受けないで、ハイパスフィルタ２１を介して出力されるので、全体としての音声品質の改善は十分に達成される。

【００５０】また、図１の例においては、１００Ｈｚ以下の成分は、風雑音には多く含まれるが、音声信号には殆ど含まれないので、バンドパスフィルタ２２及び２３を用いて１００Ｈｚ以下の周波数成分は抽出しない。しかし、バンドパスフィルタ２２及び２３の代わりにローパスフィルタを用いて、１００Ｈｚ以下の周波数成分についても適応雑音キャンセラー１０により、雑音低減処理を行うようにしてももちろん良い。

【００５１】図４は、この発明によるマイクロホン装置の他の実施例を示すブロック図である。この例は、適応雑音キャンセラー１０の参照入力は、そのときの風雑音の大きさを表しているので、この参照入力信号レベルの大きさを検出することにより、適応雑音キャンセラー１０の適応処理動作を制御するものである。

【００５２】すなわち、図４の例においては、バンドパスフィルタ２３の出力信号は、参照入力として、適応雑音キャンセラー１０に供給されると共に、レベル検出回路２５に供給され、その信号レベルが検出される。そして、このレベル検出回路２５のレベル検出出力が適応雑音キャンセラー１０の適応フィルタ回路５のフィルタ係数演算回路１１０にその係数更新の制御信号として供給される。その他は、図１の例と同様である。

【００５３】フィルタ係数演算回路１１０では、レベル検出回路２５の出力を受けて、参照入力信号（つまり、風雑音）のレベルに応じて、前述したフィルタ係数の更新式（１）のステップゲインμの大きさを制御する。すなわち、風雑音が大きいときには、適応雑音キャンセラー１０の雑音低減効果を大きくすべく、そのステップゲインμを大きめに設定し、風雑音が小さいときには、適応雑音キャンセラー１０の雑音低減効果をさほど大きくする必要もないので、そのステップゲインμを小さめに設定する。こうして、音声信号が風雑音で聴き取りづらくならないように対処できる。

【００５４】図５は、この発明によるマイクロホン装置の他の実施例を示すブロック図である。図１の例において、ハイパスフィルタ２１の出力には風雑音は殆ど含まれないので、この出力は、音声信号の中高域成分が主体であると言える。したがって、このハイパスフィルタ２１の出力信号のレベルの大きさは、音声信号の大きさを間接的に表していると解釈できる。図５の例は、上記の点を考慮して、ハイパスフィルタ２１の出力信号のレベルに応じて、図４の例と同様に、適応雑音キャンセラーの動作を制御するものである。

【００５５】すなわち、図５の例においては、ハイパスフィルタ２１の出力信号は加算回路２４に供給されると共に、レベル検出回路２６に供給され、その信号レベルが検出される。そして、このレベル検出回路２６のレベル検出出力が適応雑音キャンセラー１０の適応フィルタ回路５のフィルタ係数演算回路１１０に、その係数更新の制御信号として供給される。その他は、図１の例と同様である。

【００５６】フィルタ係数演算回路１１０では、レベル検出回路２６の出力を受けて、ハイパスフィルタ２１の出力信号レベル（つまり、音声信号レベル）に応じてフィルタ係数の更新のための、前記式（１）のステップゲインμの大きさを制御する。すなわち、例えば音声レベルが大きいときには、適応雑音キャンセラー１０の雑音低減効果をさほど大きくする必要もないので、そのステップゲインμを小さめに設定し、音声レベルが小さい時には、ステップゲインμを大きく設定するように制御する。こうして、この例においても、音声信号に加わる歪みを適宜小さくでき、良好な収音品質が得られる。

【００５７】図６は、この発明によるマイクロホン装置のさらに他の実施例を示すブロック図である。この例は、適応雑音キャンセラー１０の参照入力と、ハイパスフィルタ２１の出力信号のレベルを比較することは、その時点での信号対雑音比を（間接的に）監視することになることを利用する。

【００５８】すなわち、音声信号レベルが風雑音レベルよりも相対的に十分大きければ、適応雑音キャンセラー１０の雑音低減効果をさほど大きくする必要はなく、むしろ音声品質を良好に保つことを優先したほうが好ましい。逆に風雑音レベルが相対的に大きい場合は、適応雑音キャンセラー１０による雑音低減効果を大きくし、音声品質を少々犠牲にしても、その信号対雑音比を向上させるほうがよい。そこで、図６の例では、その時点での信号対雑音比に応じて適応雑音キャンセラー１０の適応動作を制御するようにする。

【００５９】これを実現するために、図６の例においては、ハイパスフィルタ２１の出力信号と、バンドパスフ

ィルタ２３の出力信号は、レベル比検出回路２７に供給され、このレベル比検出回路２７において、両信号のレベル比が検出される。そして、このレベル比検出回路２７で検出された両信号のレベル比が、適応雑音キャンセラー１０の適応フィルタ回路５のフィルタ係数演算回路１１０にその係数更新の制御信号として供給される。その他は、図１の例と同様である。

【００６０】フィルタ係数演算回路１１０では、上記レベル比に応じてフィルタ係数更新式（１）におけるステップゲインμを制御する。例えば、その信号対雑音比が大きいときにはステップゲインμを小さめに設定し、逆に信号対雑音比が小さいときにはステップゲインμを大きくする。こうして、良好な収音品質がえられる。

【００６１】図７は、この発明のさらに他の実施例のブロック図である。この例は、風雑音が存在する低音域を複数通りに分割し、各分割帯域毎に適応雑音キャンセラーを用いた処理を行うようにした例である。この例によれば、各分割帯域毎に適応的に風雑音が低減されるので、より効果的に風雑音を低減することが可能になる。

【００６２】図７の例は、風雑音が存在する低音域を２分割した場合の例である。すなわち、図１の例のバンドパスフィルタ２２及び２３の代わりに、１００Ｈｚ～５００Ｈｚの帯域を２分割して、その上側の分割帯域を通過帯域とするバンドパスフィルタ２２１及び２３１と、その下側の分割帯域を通過帯域とするバンドパスフィルタ２２２及び２３２とを設ける。そして、Ａ／Ｄコンバータ１３の出力信号をバンドパスフィルタ２２１及び２２２にそれぞれ供給すると共に、減算回路１５の出力信号をバンドパスフィルタ２３１及び２３２に供給する。

【００６３】また、各分割帯域数に応じて適応雑音キャンセラー５１及び５２を設ける。そして、バンドパスフィルタ２２１及び２２２の出力を、それぞれ適応雑音キャンセラー５１及び５２の主要入力端子１に供給し、バンドパスフィルタ２３１及び２３２の出力を、それぞれ適応雑音キャンセラー５１及び５２の参照入力端子２に供給する。

【００６４】この例の場合、各帯域毎の適応雑音キャンセラー５１、５２の適応フィルタ回路５におけるフィルタ係数の更新におけるステップゲインμの大きさは、その帯域毎に適切な値に設定される。もっとも、前述した図４～図６の例のように、参照入力のレベルあるいは主要入力のレベル、さらには、両者のレベル比に応じて適応的にステップゲインを制御する構成とすることもできる。

【００６５】適応雑音キャンセラー５１及び５２の出力信号は、加算回路２４に供給され、ハイパスフィルタ２１からのマイクロホン素子１１の収音出力信号の中高域成分と加算され、Ｄ／Ａコンバータ１６に供給され、アナログ信号に戻され、出力端子１７に導出される。

【００６６】図７の例は、低音域を２分割にした例であるが、３分割以上に分割してももちろん良い。

【００６７】以上説明したマイクロホン装置によれば、特殊なセンサーなどを用意しなくとも、通常のマイクロホン素子を２個用意することで、風雑音を低減させることができる。この場合に、２個のマイクロホン素子は、近接して配置されるため、機器の小型化に貢献できる。

【００６８】また、２個のマイクロホン素子をそれぞれ左右のチャンネルに振り分けることもできるので、ステレオマイクロホンを構成する場合でも、マイクロホン素子を増やすことなく実現できる。

【００６９】図４～図６の例によれば、風雑音レベルや音声信号レベルの大小によって、適応雑音キャンセラーの風雑音低減効果を適宜制御できるため、良好な収音品質と風雑音抑圧効果を同時に得ることができる。

【００７０】さらに、風雑音低減動作が自動的であるため、使用者は他の作業、例えばカメラ一体型ＶＴＲで撮影中に、モニター映像の監視などに集中できる。

【００７１】なお、上述の例では、２個のマイクロホン素子１１、１２は、無指向性のマイクロホンを使用したが、これらのマイクロホン素子は、どのような指向性であってもよい。しかし、無指向性マイクロホンを使用すれば、扱いが容易で、安価なため、実用上の効果は大きい。

【００７２】また、２個のマイクロホン素子を組み合わせて、他の指向性を得ることも容易である。さらには、低域では殆ど無指向性で、中高域では単一指向性とすることもやはり容易である。もちろん、３個以上のマイクロホン素子を利用して、この発明の主旨に沿った信号を得ることも可能である。

【００７３】なお、以上の例においては、ハイパスフィルタ２１、バンドパスフィルタ２２、２３、２２１、２２２、２３１、２３２のカットオフ周波数は、固定としたが、図２に示したように、風雑音のパワースペクトラムは、風速に応じて周波数変化することに鑑み、これらフィルタをローオフ周波数可変の構成として、風速に応じて適切なカットオフ周波数に制御するように構成することができる。カットオフ周波数可変のフィルタは、例えばＩＩＲ型デジタルフィルタを用い、その加重係数を変更することにより、実現することができる。

【００７４】カットオフ周波数の変更は、マニュアル操作で行ってもよいし、風速を計測する手段を設けて、自動的に行うようにすることもできる。

【００７５】また、Ａ／Ｄコンバータ１３、１４は、マイクロホン素子１１、１２に直接接続せずに、バンドパスフィルタ２２、２３の出力側に接続してもよい。その場合には、ハイパスフィルタ２１、バンドパスフィルタ２２、２３はアナログ回路で構成される。また、Ｄ／Ａコンバータ１６は、適応雑音キャンセラー１０の出力部に挿入されることになる。

【００７６】また、適応雑音キャンセラーの出力にバン

13

ドパスフィルタ２２、２３と同様の周波数特性のバンドパスフィルタを挿入してもよい。

【００７７】

【発明の効果】以上説明したように、この発明によれば、風雑音が低音域に主として存在していることに鑑み、マイクロホンで収音した音声を低域成分と、風雑音の成分が殆ど存在しない中高域成分とに分け、低域成分についてのみ、適応雑音キャンセラーを応用して風雑音を低減除去し、その風雑音を除去した低域成分と高域成分とを加算して出力信号を得るようにしたので、収音した音声の中高域成分は、適応雑音キャンセラーの影響は全く受けず、十分に風雑音が抑圧された良好な品質の収音出力音声信号が得られる。

【００７８】また、この発明においては、風雑音レベルや音声信号レベルの大小によって、適応雑音キャンセラーの風雑音低減効果を適宜制御するようにしたので、良好な収音品質と十分な風雑音抑圧効果を同時に得ることができる。

【図面の簡単な説明】

【図１】この発明によるマイクロホン装置の一実施例を示すブロック図である。

【図２】風雑音と音声信号の平均パワースペクトラムの一例を示す図である。

【図３】適応フィルタ回路の構成例を示すブロック図である。

【図４】この発明によるマイクロホン装置の他の一実施

14

例を示すブロック図である。

【図５】この発明によるマイクロホン装置の他の一実施例を示すブロック図である。

【図６】この発明によるマイクロホン装置の他の一実施例を示すブロック図である。

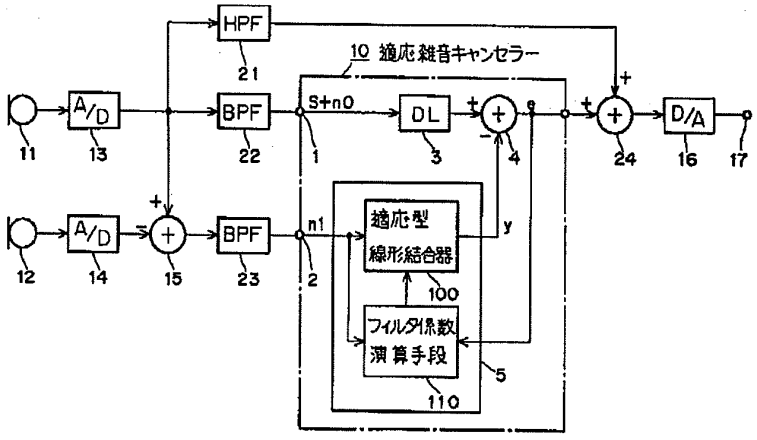
【図７】この発明によるマイクロホン装置の他の一実施例を示すブロック図である。

【図８】先願のマイクロホン装置の一実施例を示すブロック図である。

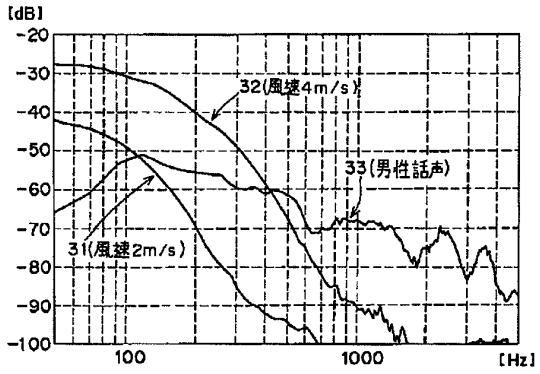
【符号の説明】

１０	適応雑音キャンセラー
１	主要入力端子
２	参照入力端子
４	合成回路
５	適応フィルタ回路
１１、１２	マイクロホン素子
１３、１４	Ａ／Ｄコンバーター
１５	減算回路
１６	Ｄ／Ａコンバータ
１７	装置出力端子
２１	ハイパスフィルタ
２２、２３	バンドパスフィルタ
２４	加算回路
２５、２６	レベル検出回路
２７	レベル比検出回路

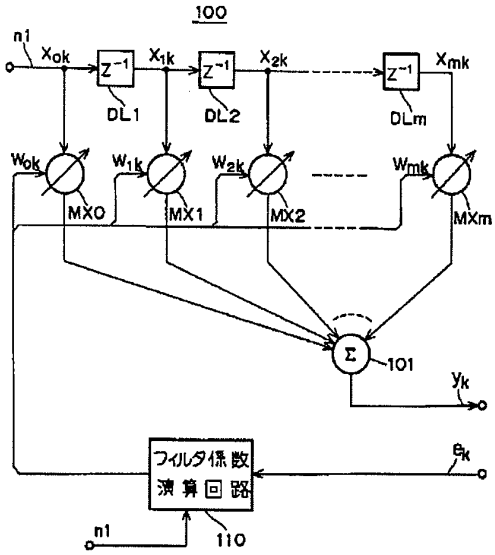
【図１】



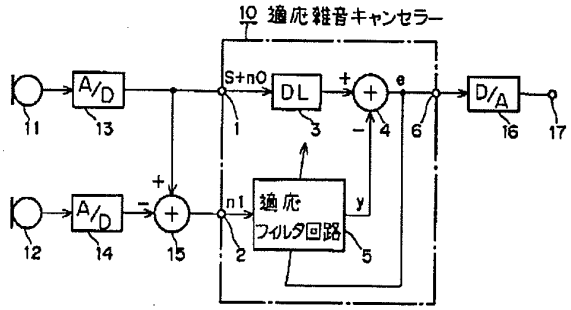
【図2】


[dB]
-20
-30
-40
-50
-60
-70
-80
-90
-100
100
1000
[Hz]
32(風速4m/s)
33(男性話声)
31(風速2m/s)

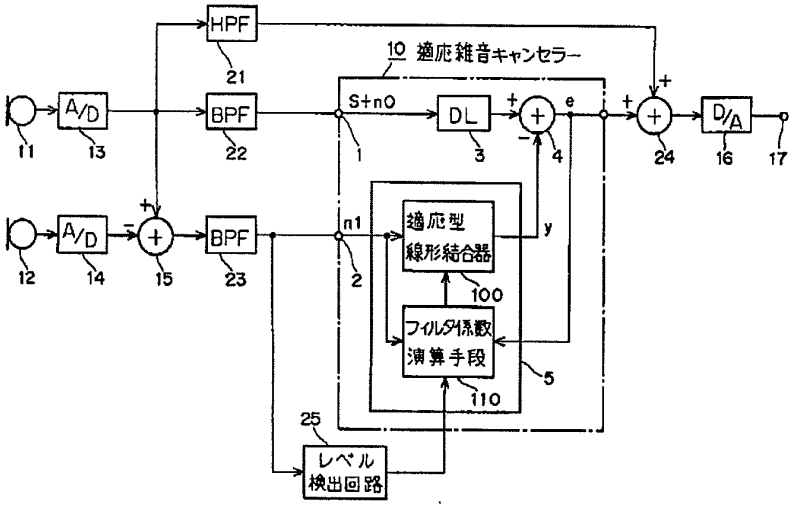
【図3】


100
n1
Xok
X1k
X2k
Xmk
Z-1
DL1
DL2
DLm
Wok
W1k
W2k
Wmk
MX0
MX1
MX2
MXm
Σ
101
yk
ek
フィルタ係数
演算回路
n1
110

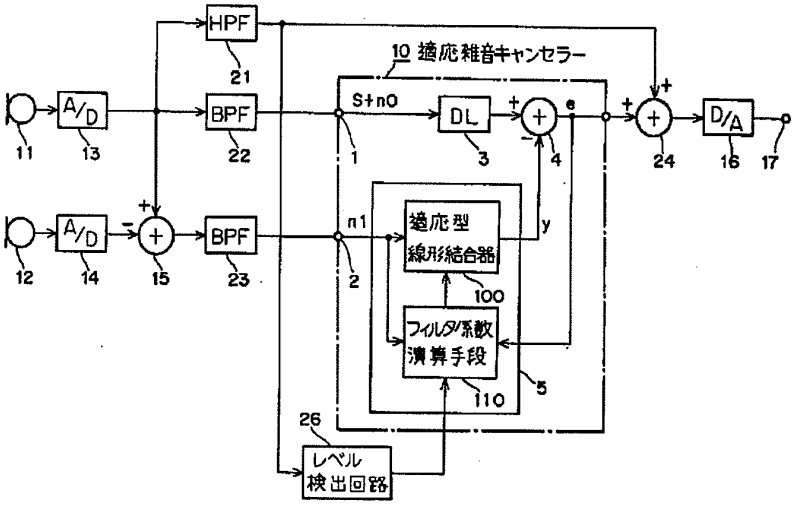
【図8】


10 適応雑音キャンセラー
A/D
A/D
S+n0
DL
D/A
適応
フィルタ回路
y
e
11
12
13
14
15
1
2
3
4
5
6
16
17

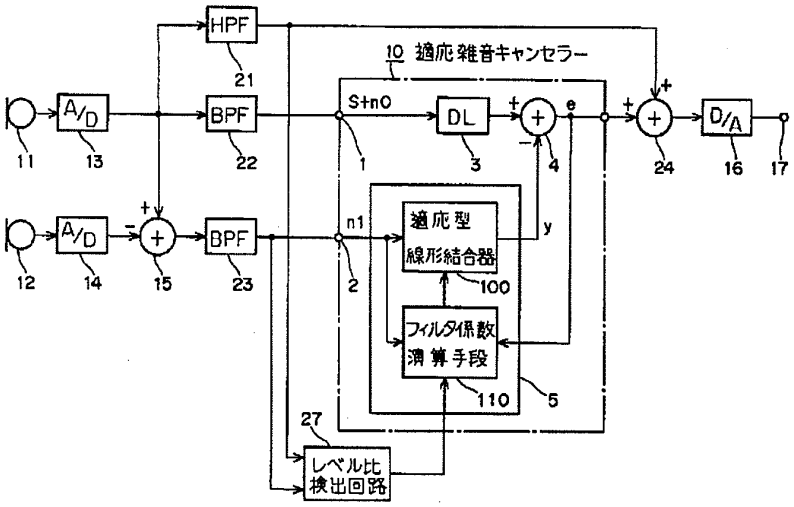
【図4】


HPF
21
10 適応雑音キャンセラー
A/D
BPF
22
S+n0
DL
3
e
D/A
24
16
17
11
13
1
4
12
14
15
23
n1
2
適応型
線形結合器
100
y
フィルタ係数
演算手段
110
5
25
レベル
検出回路

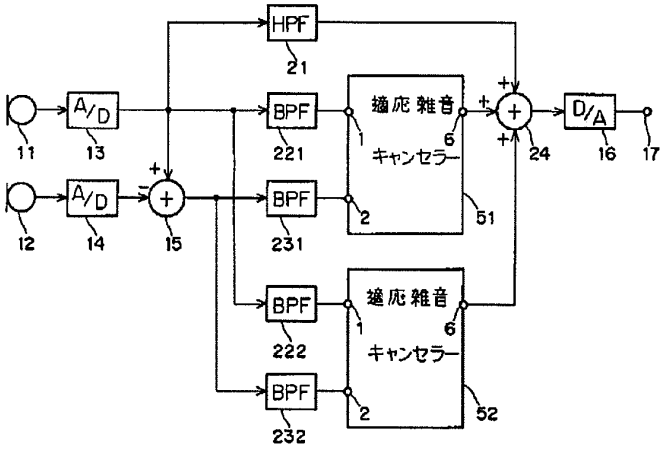
【図5】


HPF
21
10 適応雑音キャンセラー
A/D
BPF
22
S+n0
DL
3
e
D/A
24
16
17
11
13
1
4
12
14
15
23
n1
2
適応型
線形結合器
100
y
フィルタ係数
演算手段
110
5
26
レベル
検出回路

【図6】



【図7】



フロントページの続き

(51)Int.Cl.[5]	識別記号	庁内整理番号	F I	技術表示箇所
H04R 1/40		320 Z		

(72)発明者　行徳　薫
東京都品川区北品川6丁目7番35号　ソニー株式会社内

TRANSMITTER

Patent number: JP62110349
Publication date: 1987-05-21

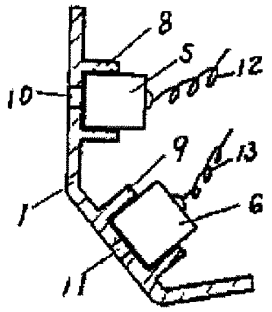
Inventor: MIURA KENZO
Applicant: MATSUSHITA ELECTRIC IND CO LTD

Classification:
- international: ***H04M1/03; H04R3/00;*** **H04M1/03; H04R3/00;** (IPC1-7): H04M1/03; H04R3/00
- european:
Application number: JP19850251321 19851108
Priority number(s): JP19850251321 19851108

Report a data error here

Abstract of **JP62110349**

PURPOSE:To effectively suppress ambient noise etc., by using two microphones, arranging them to make the mouth of a speaker come near the extension of a line connecting the microphones and taking out the difference of output of these microphones.
CONSTITUTION:Faces of a box 1 to which microphones 5, 6 are attached are formed to chevron-shape, and microphones 5, 6 are attached to each face to face to each other, and arranged to make a voice producing source such as the mouth of a speaker come close to the extension of a line connecting the microphones 5, 6. When microphones 5, 6 are arranged thus, and the difference of their output is taken out, their directional characteristics show 8-shape to the direction of voice producing source. Accordingly, ambient noise, that is, sound coming from the direction other than the voice producing source is suppressed remarkably, and only voice form the voice producing source can be transmitted effectively.



Data supplied from the ***esp@cenet*** database - Worldwide

⑲日本国特許庁(JP)　⑪特許出願公開

⑫公開特許公報(A)　昭62-110349

⑤Int.Cl.⁴	識別記号	庁内整理番号	㊸公開　昭和62年(1987)5月21日
H 04 M 1/03		7608-5K	
H 04 R 3/00	320	8524-5D	

審査請求　未請求　発明の数 1　(全3頁)

㊹発明の名称　送話器

㉑特　願　昭60-251321

㉒出　願　昭60(1985)11月8日

㉒発明者　三浦　研造　横浜市港北区綱島東4丁目3番1号　松下通信工業株式会社内

㉑出願人　松下電器産業株式会社　門真市大字門真1006番地

㉔代理人　弁理士　中尾　敏男　外1名

明　細　書

1. 発明の名称

送話器

2. 特許請求の範囲

2個のマイクロホンを結ぶ線の延長線上に近い位置に発音源が来るように上記2個のマイクロホンを配置し、これらのマイクロホンの出力差を取出して送話するように構成した送話器。

3. 発明の詳細な説明

産業上の利用分野

本発明は電話機、インターホン、無線機その他の機器に使用する送話器に関する。

従来の技術

一般に電話機、インターホン、その他の機器に使用する送話器は発音源が極めて近接することから炭素型のマイクロホンが使用されることが多い。しかしながら、この種のものでは省電力化、小型化が困難であり、最近では小型のエレクトレット型コンデンサマイクロホンが使用されることも多い。

発明が解決しようとする問題点

ところで、エレクトレット型のコンデンサマイクロホンは入力音圧に対する出力電圧の直線性が良好で感度も非常に高いため、送話者からの音声の外に周囲の騒音もそのまま拾われ伝送されることになり、受話者にとってはそれがかえって邪魔になり、必要な送話者からの音声が識別できなくなるという問題があった。また、周囲の騒音がそのまま伝送されると、その中には企業秘密に属すること、個人のプライバシーに関することなどが含まれていることがあることから副次的に企業秘密や、プライバシー等が漏洩することがあるという問題があった。

本発明は以上のような従来の欠点を除去するものであり、簡単な構成で周囲の騒音を抑圧し、送話者からの音声のみ明瞭に伝送できる優れた送話器を提供することを目的とするものである。

問題点を解決するための手段

本発明は上記目的を達成するために2個のマイクロホンを使用し、これらのマイクロホンを結ぶ線

の延長線の近くに送話者の口（発音源）が来るように上記2個のマイクロホンを配置し、これらのマイクロホンの出力の差を取出してそれを伝送するように構成したものである。

作　　用

したがって、本発明によれば、周囲の騒音等は、2つのマイクロホンで互に相殺され、送話者からの音声のみが明瞭に伝送されることになるという作用を有する。

実　施　例

第1図～第3図は本発明の送話器の一実施例を示すものであり、図中1は電話機等のハンドセットを構成する筐体、2は筐体1の一方に内蔵された受話器、3は受話器2を筐体1の他方に内蔵されたコネクタ4に接続するリード線、5，6は筐体1の他方にそれぞれ間隔をおいて配置された無指向性のエレクトレット型コンデンサマイクロホン、7は筐体1に内蔵されマイクロホン5，6の出力を入力する差動増幅回路、8，9は筐体1に一体に形成されそれぞれマイクロホン5，6を支持する支持体、10，11はマイクロホン5，6に対向して筐体1に形成された音孔、12，13，14はそれぞれマイクロホン5，6と差動増幅回路7、差動増幅回路7とコネクタ4を互に接続するリード線である。

尚、筐体1のマイクロホン5，6を取付ける面はく字状に形成され、そのそれぞれの面にマイクロホン5，6が対向するように取付けられている。そして、これらのマイクロホン5，6は、上記マイクロホン5，6を結ぶ線の延長線の近くに送話者の口等の発音源が来るように配置されている。

第4図はマイクロホン5，6と発音源の関係を示したものであり、マイクロホン5，6はそれを結ぶ線の延長線上に発音体が位置するように配置されている。第4図に示すようにマイクロホン5，6を配置し、その出力の差を取出すようにすると、その指向特性は第5図に示すように発音源の方向に対し8の字型の指向特性を示す。したがって、このようにすると周囲の騒音すなわち発音源以外の方向から到来する音に対してはそれが著しく抑圧され、発音源からの音声のみ効果的に伝送できるようになる。そして、周波数特性をとると第6図に示すようにマイクロホン5と発音源の距離がたとえば1cmと極めて小さく、発音源が指向特性の0度の位置にあるときに最も平坦で良好な特性を示す。すなわち、発音源が指向特性の90度の位置にあるときには低減が著しく低減され、また発音源が指向特性の0度の位置にあるときでも発音源とマイクロホン5の間の距離がたとえば50cmと比較的大きい場合には低減が比較的大きく低減される。しかしながら、発音源が指向特性の0度の位置にあり、かつ発音源とマイクロホン5の間の距離が比較的小さい場合には第6図実線で示すようにほゞ全周波数範囲に亘って平坦で、かつ高感度な周波数特性を得ることができ、したがって、この点でも発音源からの音声のみ効果的に伝送できることになる。

このように、上記実施例によれば2つのマイクロホン5，6を使用し、それを結ぶ線の延長線上又はそれに近い位置に発音源が来るように上記2つのマイクロホン5，6を配置し、これらのマイクロホン5，6の出力の差を取出し伝送するように構成しており、したがって、発音源からの音声のみ効果的に伝送し、発音源以外の方向から来る周囲の騒音に対してはそれを大きく抑圧することができる。そして、上記実施例によればマイクロホン5，6として無指向性のマイクロホンを使用し、差動増幅回路7によって上記マイクロホン5，6の出力の差を電気的に取出すように構成しており、したがって、筐体1に形成する音孔10，11も比較的小さくすることができるという利点を有する。また、上記実施例では、発音源とマイクロホン5の間の距離をいくらにするか明確にしていないが、電話機やインターホン、無線機等に使用する送話器では通常その距離が10cm以下と極めて小さく、したがって、先述したようにほゞ全周波数範囲に亘って平坦な周波数特性が得られるという利点を有する。

発明の効果

本発明は上記実施例より明らかなように2つの

マイクロホンを使用し、それを結ぶ線の延長線上又はそれに近い位置に送話者の口等の発音源が来るように上記2つのマイクロホンを配置し、これらのマイクロホンの出力の差を取出して伝送するように構成したものであり、したがって、周囲の騒音等は効果的に抑圧され、発音源からの音声が明瞭に伝送されるという利点を有する。

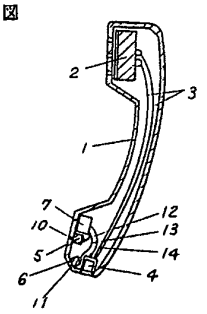
4. 図面の簡単な説明

第1図は本発明の送話器における一実施例の断側面図、第2図は同要部の拡大断面図、第3図は同要部の斜面図、第4図は本発明の送話器を構成する2つのマイクロホンと発音源の関係を示す説明図、第5図は同送話器の指向特性図、第6図は同送話器の周波数特性図である。

1…匡体、2…受話器、3，12，13，14…リード線、4…コネクタ、5，6…マイクロホン、7…差動増幅回路、8，9…支持体、10，11…音孔。

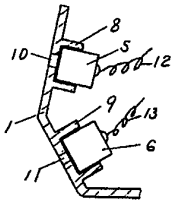
代理人の氏名　弁理士　中　尾　敏　男　ほか1名

第 1 図

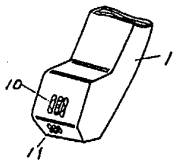


1…匡体
2…受話器
4…コネクタ
5.6…マイクロホン
7…差動増幅回路
10.11…音孔

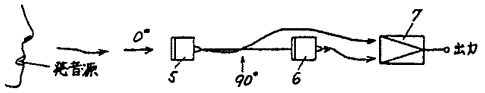
第 2 図



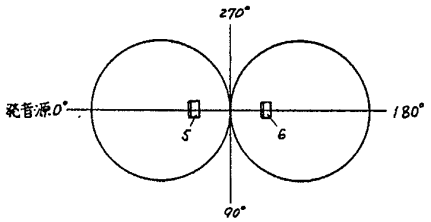
第 3 図



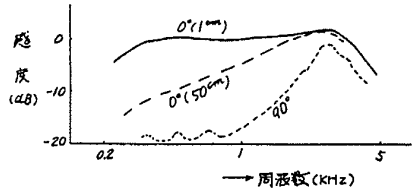
第 4 図



第 5 図



第 6 図



Electronic Acknowledgement Receipt	
EFS ID:	4721374
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	Susan Yee
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	02-FEB-2009
Filing Date:	29-JAN-2007
Time Stamp:	19:55:09
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Filed (SB/08)	PA3909US_PTO_SB_08_1.pdf	39991 821f85034c0c666efad02f1e4aa0c411172e78a7	no	8
Warnings:					
Information:					

This is not an USPTO supplied IDS fillable form					
2	Information Disclosure Statement (IDS) Filed (SB/08)	PA3909US_PTO_SB_08_2.pdf	610265 f0138579e2f90f6df4eaecc2c067e5b39e81e143	no	7
Warnings:					
Information:					
3	Information Disclosure Statement Letter	PA3909US_PTO_Letter.pdf	100537 d3d4f201dc91b38d291d1fe9d408fbdcc12a9af0	no	2
Warnings:					
Information:					
4	Information Disclosure Statement Letter	PA3909US_PTO_Copending_List.pdf	86365 ce03f2408b8ae56b40cee8f06f2ec3cc079e27ce	no	1
Warnings:					
Information:					
5	NPL Documents	PA3909US_IDS_Allen1.pdf	686876 31c4198a4f65ec9809d78671fc128ee00ce8f0f9	no	7
Warnings:					
Information:					
6	NPL Documents	PA3909US_IDS_Allen2.pdf	372213 b0f8badb20567c28ccaaa438317dcbed5af62dec	no	4
Warnings:					
Information:					
7	NPL Documents	PA3909US_IDS_Avendano.pdf	377374 6d637a3df3932692a07e61272a7647d987c317f0	no	4
Warnings:					
Information:					
8	NPL Documents	PA3909US_IDS_Boll1.pdf	253736 720347afcee90b8fe1fd8d4c12d7800fa5960ebe	no	2
Warnings:					
Information:					
9	NPL Documents	PA3909US_IDS_Chen.pdf	1564275 da172015838b3c2ac802575fbe4555404ea29326	no	17
Warnings:					
Information:					
10	NPL Documents	PA3909US_IDS_Cohen2.pdf	387172 76c315caaa0ac1c7c2de6fd7646bfb552dc75e4b	no	4

Warnings:					
Information:					
11	NPL Documents	PA3909US_IDS_Cohen.pdf	1085762 3f9eda64a93a1d3cad39b1a16bca01c89fc3b143	no	12
Warnings:					
Information:					
12	NPL Documents	PA3909US_IDS_Fuchs.pdf	391667 fa505a6223dac0f4d956f1da0fc39569eecb8488	no	4
Warnings:					
Information:					
13	NPL Documents	PA3909US_IDS_Goubran.pdf	340895 aa1b46ac076e87266fd1bb37b78071c9be261df4	no	6
Warnings:					
Information:					
14	Foreign Reference	PA3909US_IDS_JP04184400_Yoshio.pdf	388389 e2e46463275 7b3ce2bc9ebec96fbd4297d12f6ae	no	9
Warnings:					
Information:					
15	Foreign Reference	PA3909US_IDS_JP06269083_Toru.pdf	848950 1efa98dcd81bae9bf83e0aa862f92a7e855368be	no	14
Warnings:					
Information:					
16	Foreign Reference	PA3909US_IDS_JP62110349_Miura.pdf	184461 b03b8c45ae5ab3e3ddf81f6d1b996c0073dc2763	no	4
Warnings:					
Information:					
17	NPL Documents	PA3909US_IDS_Liu.pdf	1220193 f7dc93a72741d90ec111d20a41a4151765c90ce2	no	14
Warnings:					
Information:					
18	NPL Documents	PA3909US_IDS_Martin.pdf	813636 d96dced25204e8de4c26877940a694c4e8771de9	no	11
Warnings:					
Information:					
19	NPL Documents	PA3909US_IDS_Mizumachi.pdf	343011 704e776619c944a42a6f730145faf233b25e9ef3	no	4

Warnings:					
Information:					
20	NPL Documents	PA3909US_IDS_Moonen.pdf	83003 7a49e303ccf047b75ce1670f563a4c9df34a0e44	no	2
Warnings:					
Information:					
21	NPL Documents	PA3909US_IDS_Parra.pdf	852566 6ac4e4b05a80477dc5f62e88d00c7f1c3f64dbfe	no	8
Warnings:					
Information:					
22	NPL Documents	PA3909US_IDS_Tashev.pdf	326809 9488d61607d85b1e871a9c24b07025836eb2b814	no	4
Warnings:					
Information:					
23	NPL Documents	PA3909US_IDS_Valin.pdf	536749 72c79464ba7dd7b7fbc7308f344e065db5d01ca6	no	6
Warnings:					
Information:					
24	NPL Documents	PA3909US_IDS_Widrow1.pdf	1355580 d3c65facc9302cea847bd89bd461340c84560414	no	17
Warnings:					
Information:					
			Total Files Size (in bytes):	13250475	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Carlos Avendano
APPLICATION NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Vivian C. Chin
GROUP ART UNIT:	2615
ATTY.DKT.NO.:	PA3909US

INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. § 1.56, and 1.97-1.98

SIR:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08, listing several references. The Examiner is requested to make these references of official record in the application.

The references cited may be material to examination of the application and are submitted in compliance with Applicant's duty of disclosure as defined by 37 C.F.R. § 1.56. No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

Applicant believes no fee is due with this submission. If a fee is due, however, the Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600.

Respectfully submitted,
Carlos Avendano

February 18, 2009

By: /Susan Yee/

Susan Yee, Reg. No. 41,388
Carr & Ferrell *LLP*
2200 Geng Road
Palo Alto, CA 94303
TEL: (650) 812-3400
FAX: (650) 812-3444

Doc code :IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (10-08)
Approved for use through 11/30/2008. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

U.S.PATENTS						
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Patent citation information please click the Add button.						

U.S.PATENT APPLICATION PUBLICATIONS						
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20030072382	A1	2003-04-17	Raleigh et al	
If you wish to add additional U.S. Published Application citation information please click the Add button.						

FOREIGN PATENT DOCUMENTS								
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button								

NON-PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☒ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☐ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Susan Yee/	Date (YYYY-MM-DD)	2009-02-18
Name/Print	Susan Yee	Registration Number	41388

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Electronic Acknowledgement Receipt	
EFS ID:	4816839
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	Susan Yee
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	18-FEB-2009
Filing Date:	29-JAN-2007
Time Stamp:	18:33:18
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement Letter	PA3909US_Supplemental_IDS_Letter.pdf	100540 e1fee7edb0b822dd3afcb342ba35e6ed9f7ca75e	no	2

Warnings:
Information:

2	Information Disclosure Statement (IDS) Filed (SB/08)	PA3909US_Supplemental_IDS_SB08.pdf	25167 2bab6b6097bb46b129d26af34744cd45c1fbdcb9	no	3

Warnings:	
Information:	
This is not an USPTO supplied IDS fillable form	
Total Files Size (in bytes):	125707

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 04/07/2011

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	
ART UNIT	PAPER NUMBER
2614	
MAIL DATE	DELIVERY MODE
04/07/2011	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	11/699,732	AVENDANO, CARLOS	
	Examiner	Art Unit	
	PAUL KIM	2614	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *29 January 2007*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-26* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-26* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *29 January 2007* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *2/18/2009, 2/2/2009, 2/2/2009, 5/24/2007*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 101

1. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

2. Claim 26 is rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

3. The claim recites, *inter alia*, "a machine readable medium having embodied thereon a program, the program providing instructions for a method for enhancing speech" After close inspection, the Examiner respectfully notes that the disclosure, as a whole, does not specifically identify what may be included as a computer readable storage medium and what is not to be included as a computer readable storage medium.

4. An Examiner is obliged to give claims their broadest reasonable interpretation consistent with the specification during examination. The broadest reasonable interpretation of a claim drawn to a computer readable medium (also called machine readable medium and other such variations) typically covers forms of non-transitory tangible media and transitory propagating signals *per se* in view of the ordinary and customary meaning of computer readable media, particularly when the specification is silent. See MPEP 2111.01. When the broadest reasonable interpretation of a claim

covers a signal, *per se*, the claim must be rejected under 35 U.S.C. § 101 as covering non-statutory subject matter.

5. Therefore, given the silence of the disclosure and the broadest reasonable interpretation, the computer readable storage medium of the claim may include transitory propagating signals. As a result, the claim pertains to non-statutory subject matter.

6. However, the Examiner respectfully submits a claim drawn to such a computer readable medium that covers both transitory and non-transitory embodiments may be amended to narrow the claim to cover only statutory embodiments to avoid a rejection under 35 U.S.C. § 101 by adding the limitation "a non-transitory machine readable medium having embodied thereon a program executable by a processor" to the claim. Such an amendment would typically not raise the issue of new matter, even when the specification is silent because the broadest reasonable interpretation relies on the ordinary and customary meaning that includes signals *per se*. For additional information, please see the Patents' Official Gazette notice published February 23, 2010 (1351 OG 212).

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 1-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Elko (US Pub. 20030147538) in view of Ryan (US Pub. 20030169891).

Regarding claim 1, Elko discloses a system for enhancing speech (see abstract), comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal (see paragraphs 8 and 95);

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal (see paragraph 26).

Elko fails to disclose an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

However, Ryan discloses an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference (see paragraphs 38 and 40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's inter-microphone level difference module with a

system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 2, Elko in view of Ryan discloses a system of claim 1, as discussed above.

Elko fails to disclose wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level-equalized secondary electric signal.

However, Ryan discloses wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level-equalized secondary electric signal (see paragraph 70).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's DMA module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 3, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain (see Elko, paragraphs 99 and 102).

Regarding claim 4, Elko in view of Ryan discloses a system of claim 3, as discussed above.

Elko fails to disclose wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

However, Ryan discloses wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal (see paragraph 42).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's gain with a system for enhancing speech of Elko. The motivation would be to provide a well balanced value for the gain.

Regarding claim 5, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the level-equalized secondary electric signal and a delayed primary electric signal (see Elko, paragraphs 99 and 102).

Regarding claim 6, Elko in view of Ryan discloses a system of claim 1, as discussed above, further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal (see Elko, paragraph 63).

Regarding claim 7, Elko in view of Ryan discloses a system of claim 1, as discussed above.

Elko fails to disclose further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

However, Ryan discloses further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal (see paragraph 39).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's energy module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 8, Elko in view of Ryan discloses a system of claim 1, as discussed above, further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference (see Elko, paragraph 51).

Regarding claim 9, Elko in view of Ryan discloses a system of claim 1, as discussed above.

Elko fails to disclose further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

However, Ryan discloses further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal (see paragraph 28).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's filter module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 10, Elko in view of Ryan discloses a system of claim 9, as discussed above.

Elko fails to disclose further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

However, Ryan discloses further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal (see paragraph 28).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's filter smoothing module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 11, Elko in view of Ryan discloses a system of claim 1, as discussed above, further comprising a masking module configured to determine a speech estimate (see Elko, paragraph 52).

Regarding claim 12, Elko in view of Ryan discloses a system of claim 11, as discussed above, further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output (see Elko, paragraph 52).

Regarding claim 13, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal (see Elko, paragraph 105).

Regarding claim 14, Elko discloses a method for enhancing speech (see abstract), comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone (see paragraphs 8 and 95);

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal (see paragraph 26).

Elko fails to disclose non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

However, Ryan discloses non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference (see paragraphs 38 and 40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's non-linearly combining components with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 15, Elko in view of Ryan discloses a method of claim 14, as discussed above.

Elko fails to disclose wherein determining the cardioid primary signal comprises taking a difference between the primary electric signal and a delayed secondary electric signal.

However, Ryan discloses wherein determining the cardioid primary signal comprises taking a difference between the primary electric signal and a delayed secondary electric signal (see paragraph 70).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's method of determining the cardioid primary signal with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 16, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain (see Elko, paragraphs 99 and 102).

Regarding claim 17, Elko in view of Ryan discloses a method of claim 16, as discussed above.

Elko fails to disclose wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

However, Ryan discloses wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal (see paragraph 42).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's gain with a method for enhancing speech of Elko. The motivation would be to provide a well balanced value for the gain.

Regarding claim 18, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal (see Elko, paragraphs 99 and 102).

Regarding claim 19, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein non-linearly combining comprises dividing the component of

the cardioid primary signal by the component of the cardioid secondary signal (see Elko, paragraph 29).

Regarding claim 20, Elko in view of Ryan discloses a method of claim 14, as discussed above.

Elko fails to disclose further comprising determining an energy estimate for each of the acoustic signals during a frame.

However, Ryan discloses further comprising determining an energy estimate for each of the acoustic signals during a frame (see paragraph 39).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's energy module with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 21, Elko in view of Ryan discloses a method of claim 14, as discussed above, further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference (see Elko, paragraph 51).

Regarding claim 22, Elko in view of Ryan discloses a method of claim 21, as discussed above.

Elko fails to disclose further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

However, Ryan discloses further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference (see paragraphs 39-40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's determining a filter estimate with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 23, Elko in view of Ryan discloses a method of claim 22, as discussed above, further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal (see Elko, paragraph 52).

Regarding claim 24, Elko in view of Ryan discloses a method of claim 22, as discussed above.

Elko fails to disclose further comprising smoothing the filter estimate.

However, Ryan discloses further comprising smoothing the filter estimate (see paragraph 28).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's smoothing the filter estimate with a method for

enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 25, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein the cardioid primary signal and the cardioid secondary signal is of a sub-band of the primary electric signal (see Elko, paragraph 105).

Regarding claim 26, Elko discloses a machine readable medium having embodied thereon a program, the program providing instructions for a method for enhancing speech (see abstract; paragraph 119), comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone (see paragraphs 8 and 95);

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal (see paragraph 26).

Elko fails to disclose non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference.

However, Ryan discloses non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference (see paragraphs 38 and 40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's non-linearly combining components with a method

for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PAUL KIM whose telephone number is (571) 270-7697. The examiner can normally be reached on Monday - Thursday, 8 AM - 4 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Vivian Chin can be reached on (571) 272-7848. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/P. K./
Examiner, Art Unit 2614

/VIVIAN CHIN/
Supervisory Patent Examiner, Art Unit 2614

Notice of References Cited	Application/Control No. 11/699,732	Applicant(s)/Patent Under Reexamination AVENDANO, CARLOS	
	Examiner PAUL KIM	Art Unit 2614	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-2003/0147538	08-2003	Elko, Gary W.	381/92
*	B	US-2003/0169891	09-2003	Ryan et al.	381/92
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 7721

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
11/699,732	01/29/2007	381	2614	PA3909US

APPLICANTS
Carlos Avendano, Campbell, CA;

**** CONTINUING DATA ***************************
This appln claims benefit of 60/850,928 10/10/2006
and is a CIP of 11/343,524 01/30/2006

**** FOREIGN APPLICATIONS ***************************

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED **** ** SMALL ENTITY **
02/26/2007

Foreign Priority claimed ☐ Yes ☑ No 35 USC 119(a-d) conditions met ☐ Yes ☑ No Verified and Acknowledged /PAUL KIM/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
		CA	9	26	3

ADDRESS

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025
UNITED STATES

TITLE

System and method for utilizing omni-directional microphones for speech enhancement

FILING FEE RECEIVED		
650	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

BIB (Rev. 05/07).

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L4	2615	(381/313,312,91,92,122).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/01/27 15:40
S1	1	(11/699732).APP.	US-PGPUB; USPAT	OR	OFF	2011/01/27 05:56
S2	20	((CARLOS) near2 (AVENDANO)).INV.	US-PGPUB; USPAT	ADJ	ON	2011/01/27 06:07
S3	53	("20010016020" | "20010031053" | "20020009203" | "20020080980" | "20020106092" | "20020116187" | "20030039369" | "20030072460" | "20030099345" | "20030138116" | "20030169891" | "20050185813" | "4649505" | "4658426" | "5208864" | "5319736" | "5381473" | "5402496" | "5574824" | "5602962" | "5694474" | "5757937" | "5796819" | "6002776" | "6061456" | "6072881" | "6222927" | "6317501" | "6363345" | "6381176" | "6430295" | "6549630" | "6584203" | "6717991" | "6738482" | "6760450" | "6760805" | "6785381" | "6795558" | "6831865" | "6882736" | "6917688" | "6978159" | "7031478" | "7089349" | "7099821" | "7142677" | "7146316" | "7155019" | "7171008" | "7174022" | "7206418" | "7246058").PN.	US-PGPUB; USPAT	ADJ	ON	2011/01/27 06:09
S5	1	("20030147538").PN.	US-PGPUB; USPAT	OR	OFF	2011/01/27 07:30
S6	1	(11/343524).APP.	US-PGPUB; USPAT	OR	OFF	2011/01/27 07:33

1/27/2011 3:47:57 PM
C:\Documents and Settings\pkim2\My Documents\EAST\Workspaces\11699732.wsp

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11699732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

SEARCHED			
Class	Subclass	Date	Examiner
381	313, 312, 91, 92, 122	1/27/2011	PK

SEARCH NOTES		
Search Notes	Date	Examiner
EAST - see EAST search notes	1/26/2011	PK
Inventor search - Carlos Avendano	1/26/2011	PK

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner

/P. K./ Examiner.Art Unit 2614	

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11699732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	01/27/2011								
	1	✓								
	2	✓								
	3	✓								
	4	✓								
	5	✓								
	6	✓								
	7	✓								
	8	✓								
	9	✓								
	10	✓								
	11	✓								
	12	✓								
	13	✓								
	14	✓								
	15	✓								
	16	✓								
	17	✓								
	18	✓								
	19	✓								
	20	✓								
	21	✓								
	22	✓								
	23	✓								
	24	✓								
	25	✓								
	26	✓								

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (10-08)
Approved for use through 11/30/2008. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

U.S.PATENTS						
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	4649505		1987-03-10	Zinser, Jr. et al.	
	2	4658426		1987-04-14	Chabries et al.	
	3	5208864		1993-05-04	Kaneda	
	4	5319736		1994-06-07	Hunt	
	5	5381473		1995-01-10	Andrea et al.	
	6	5402496		1995-03-28	Soli et al.	
	7	5574824		1996-11-12	Slyh et al.	
	8	5602962		1997-02-11	Kellermann	

EFS Web 2.1.7

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

	9	5694474		1997-12-02	Ngo et al.	
	10	5757937		1998-05-26	Itoh et al.	
	11	5796819		1998-08-18	Romesburg	
	12	6002776		1999-12-14	Bhadkamkar et al.	
	13	6061456		2000-05-09	Andrea et al.	
	14	6072881		2000-06-06	Linder	
	15	6222927	B1	2001-04-24	Feng et al.	
	16	6317501	B1	2001-11-13	Matsuo	
	17	6363345	B1	2002-03-26	Marash et al.	
	18	6430295	B1	2002-08-06	Handel et al.	
	19	6549630	B1	2003-04-15	Bobisuthi	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

	20	6717991	B1	2004-04-06	Gustafsson et al.	
	21	6738482	B1	2004-05-18	Jaber	
	22	6760450	B2	2004-07-06	Matsuo	
	23	6785381	B2	2004-08-31	Gartner et al.	
	24	6795558	B2	2004-09-21	Matsuo	
	25	6882736	B2	2005-04-19	Dickel et al.	
	26	6917688	B2	2005-07-12	Yu et al.	
	27	6978159	B2	2005-12-20	Feng et al.	
	28	7031478	B2	2006-04-18	Belt et al.	
	29	7155019	B2	2006-12-26	Hou	
	30	7171008	B2	2007-01-30	Elko	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

	31	7174022	B1	2007-02-06	Zhang et al.	
	32	7206418	B2	2007-04-17	Yang et al.	
	33	7246058	B2	2007-07-17	Burnett	

If you wish to add additional U.S. Patent citation information please click the Add button.

U.S.PATENT APPLICATION PUBLICATIONS

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20010016020	A1	2001-08-23	Gustafsson et al.	
	2	20010031053	A1	2001-10-18	Feng et al.	
	3	20020009203	A1	2002-01-24	Erten	
	4	20020041693	A1	2002-04-11	Matsuo	
	5	20020080980	A1	2002-06-27	Matsuo	
	6	20020106092	A1	2002-08-08	Matsuo	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

	7	20020116187	A1	2002-08-22	Erten	
	8	20030026437	A1	2003-02-06	Janse et al.	
	9	20030039369	A1	2003-02-27	Bullen	
	10	20030072460	A1	2003-04-17	Gonopolskiy et al.	
	11	20030099345	A1	2003-05-29	Gartner et al.	
	12	20030138116	A1	2003-07-24	Jones et al.	
	13	20030147538	A1	2003-08-07	Elko	
	14	20030169891	A1	2003-09-11	Ryan et al.	
	15	20040133421	A1	2004-07-08	Burnett et al.	
	16	20050049864	A1	2005-03-03	Kaltenmeier et al.	
	17	20050060142	A1	2005-03-17	Visser et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

If you wish to add additional U.S. Published Application citation information please click the Add button.

FOREIGN PATENT DOCUMENTS

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐

If you wish to add additional Foreign Patent Document citation information please click the Add button

NON-PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	01/27/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (10-08)
Approved for use through 11/30/2008. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT
(Not for submission under 37 CFR 1.99)

Application Number	11699732
Filing Date	2007-01-29
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Vivian C. Chin
Attorney Docket Number	PA3909US

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	7099821	B2	2006-08-29	Visser et al.	
	2	7146316	B2	2006-12-05	Alves	
	3	7142677	B2	2006-11-28	Gonopolskiy	

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20050152559	A1	2005-07-14	Gierl et al.	
	2	20050185813	A1	2005-08-25	Sinclair et al.	
	3	20050213778	A1	2005-09-29	Buck et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

	4	20050276423	A1	2005-12-15	Aubauer et al.	
	5	20060120537	A1	2006-06-08	Burnett et al.	
	6	20070021958	A1	2007-01-25	Visser et al.	
	7	20070027685	A1	2007-02-01	Arakawa et al.	
	8	20070033020	A1	2007-02-08	(Kelleher) Francois et al.	
	9	20070078649	A1	2007-04-05	Hetherington et al.	
	10	20070116300	A1	2007-05-24	Chen	
	11	20070150268	A1	2007-06-28	Acero et al.	
	12	20070154031	A1	2007-07-05	Avendano et al.	
	13	20070165879	A1	2007-07-19	Deng et al.	
	14	20030061032	A1	2003-03-27	Gonopolskiy	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29	
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlos Avendano	
	Art Unit	2615	
	Examiner Name	Vivian C. Chin	
	Attorney Docket Number	PA3909US	

	15	20060074646	A1	2006-04-06	Alves et al.	

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	04184400	JP		1992-07-01	Yoshio		☐
	2	06269083	JP		1994-09-22	Toru et al.		☐
	3	62110349	JP		1987-05-21	Miura		☐

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1	MARC MOONEN et al. "Multi-Microphone Signal Enhancement Techniques for Noise Suppression and Dereverberation," source(s): http://www.esat.kuleuven.ac.be/sista/yearreport97/node37.html.	☐
	2	STEVEN BOLL et al. "Suppression of Acoustic Noise in Speech Using Two Microphone Adaptive Noise Cancellation", source(s): IEEE Transactions on Acoustic, Speech, and Signal Processing. Vol. v ASSP-28, n 6, December 1980, pgs. 752-753.	☐
	3	CHEN LIU et al. "A two-microphone dual delay-line approach for extraction of a speech sound in the presence of multiple interferers", source(s): Acoustical Society of America. Vol. 110, 6, December 2001, pgs. 3218-3231.	☐

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	4	COHEN et al. "Microphone Array Post-Filtering for Non-Stationary Noise", source(s): IEEE. May 2002	☐
	5	JINGDONG CHEN et al. "New Insights into the Noise Reduction Wiener Filter", source(s): IEEE Transactions on Audio, Speech, and Langauge Processing. Vol. 14, 4, July 2006, pgs. 1218-1234.	☐
	6	RAINER MARTIN et al. "Combined Acoustic Echo Cancellation, Dereverberation and Noise Reduction: A two Microphone Approach", source(s): Annales des Telecommunications/Annals of Telecommunications. Vol. 29, 7-8, July -Aug 1994, pgs. 429-438.	☐
	7	MITSUNORI MIZUMACHI et al. "Noise Reduction by Paired-Microphones Using Spectral Subtraction", source(s): 1998 IEEE. pgs. 1001-1004.	☐
	8	LUCAS PARRA et al. "Convolutive blind Separation of Non-Stationary", source(s): IEEE Transactions on Speech and Audio Processing. Vol 8, 3, May 2008, pgs. 320-327.	☐
	9	ISREAL COHEN. "Multichannel Post-Filtering in Nonstationary Noise Environment", source(s): IEEE Transactions on Signal Processing. Vol. 52, 5, May 2004, pgs. 1149-1160.	☐
	10	R.A. GOUBRAN. "Acoustic Noise Suppression Using Regressive Adaptive Filtering", source(s): 1990 IEEE. pgs. 48-53	☐
	11	IVAN TASHEV et al. "Microphone Array of Headset with Spatial Noise Suppressor", source(s): http://research. microsoft.com/users/ivantash/Documents/Tashev_MAforHeadset_HSCMA_05.pdf. (4 pages).	☐
	12	MARTIN FUCHS et al. "Noise Suppression for Automotive Applications Based on Directional Information", source(s): 2004 IEEE. pgs. 237-240.	☐
	13	JEAN-MARC VALIN et al. "Enhanced Robot Audition Based on Microphone Array Source Separation with Post-Filter", source(s): Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, September 28- October 2, 2004, Sendai, Japan. pgs. 2123-2128.	☐
	14	JONT B. ALLEN. "Short Term Spectral Analysis, Synthesis, and Modification by Discrete Fourier Transform", IEEE Transactions on Acoustics, Speech, and Signal Processing. Vol ASSP-25, 3. June 1977. pgs. 235-238.	☐

EFS Web 2.1.6

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/02/2009	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	15	JONT B. ALLEN et al. "A Unified Approach to Short-Time Fourier Analysis and Synthesis", Proceedings of the IEEE. Vol. 65, 11, November 1977. pgs. 1558-1564.	☐
	16	C. Avendano, "Frequency-Domain Techniques for Source Identification and Manipulation in Stereo Mixes for Enhancement, Suppression and Re-Panning Applications," in Proc. IEEE Workshop on Application of Signal Processing to Audio and Acoustics, Waspaa, 03, New Paltz, NY, 2003.	☐
	17	B. Widrow et al., "Adaptive Antenna Systems," Proceedings IEEE, Vol. 55, No. 12, pp. 2143-2159, December, 1967.	☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	01/27/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

U.S.PATENTS						
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Patent citation information please click the Add button.						

U.S.PATENT APPLICATION PUBLICATIONS						
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20030072382	A1	2003-04-17	Raleigh et al	
If you wish to add additional U.S. Published Application citation information please click the Add button.						

FOREIGN PATENT DOCUMENTS								
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button								

NON-PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

EFS Web 2.1.7

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 02/18/2009	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2615
	Examiner Name	Vivian C. Chin
	Attorney Docket Number	PA3909US

	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	01/27/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

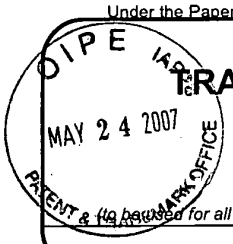
[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 05/24/2007

11699732 - GAU: 2614

PTO/SB/21 (09-04)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.



TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission	11

Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

ENCLOSURES *(Check all that apply)*

- ☐ Fee Transmittal Form
  - ☐ Fee Attached
- ☐ Amendment/Reply
  - ☐ After Final
  - ☐ Affidavits/declaration(s)
- ☐ Extension of Time Request
- ☐ Express Abandonment Request
- ☑ Information Disclosure Statement
- ☐ Certified Copy of Priority Document(s)
- ☐ Reply to Missing Parts/ Incomplete Application
  - ☐ Reply to Missing Parts under 37 CFR 1.52 or 1.53
- ☐ Drawing(s)
- ☐ Licensing-related Papers
- ☐ Petition
- ☐ Petition to Convert to a Provisional Application
- ☐ Power of Attorney, Revocation Change of Correspondence Address
- ☐ Terminal Disclaimer
- ☐ Request for Refund
- ☐ CD, Number of CD(s) _____
  - ☐ Landscape Table on CD
- ☐ After Allowance Communication to TC
- ☐ Appeal Communication to Board of Appeals and Interferences
- ☐ Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- ☐ Proprietary Information
- ☐ Status Letter
- ☑ Other Enclosure(s) (please Identify below):
  1) Confirmation postcard;
  2) Duplicate of IDS;
  3) PTO/SB/08 (x2);
  4) Copies of 3 cited nonpatent references

Remarks

Total page number does not include postcard and nonpatent references

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	Carr & Ferrell LLP Cust. No. 22830		
Signature	Susan Yee		
Printed name	Susan Yee		
Date	May 21, 2007	Reg. No.	41,388

CERTIFICATE OF TRANSMISSION/MAILING

I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:

Signature	Susan Yee		
Typed or printed name	Susan Yee, Reg. No. 41,388	Date	May 21, 2007

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and1.14. This collection is estimated to 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

OIPE IAP28 MAY 24 2007 PATENT & TRADEMARK OFFICE

IN THE

UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT:	Carlos Avendano
APPLICATION NO.:	11/699,732
FILING DATE:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
ART UNIT:	2615
EXAMINER:	Unknown
ATTY.DKT.NO.:	PA3909US

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Mail Stop Amendment, Commissioner for Patents, Alexandria, VA 22313-1450, on the date printed below:

Dated: May 21, 2007 By: Susan Yee

Susan Yee, Reg. No. 41,388

MAIL STOP AMENDMENT
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. §§ 1.56, and 1.97-1.98

SIR:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08 listing several references; copies of the non-

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

patent literature are enclosed. The Examiner is requested to make these references of official record in the application. The references cited may be material to examination of the application and are submitted in compliance with Applicant's duty of disclosure as defined by 37 C.F.R. § 1.56.

No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

Applicant believes no fee is due with this submission. If a fee is due, however, the Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600. A copy of this statement is submitted in duplicate for this purpose.

Respectfully submitted,
Carlos Avendano

Dated: May 21, 2007 By: _____

Susan Yee, Reg. No. 41,388
Carr & Ferrell *LLP*
2200 Geng Road
Palo Alto, CA 94303
TEL: (650) 812-3400
FAX: (650) 812-3444

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

PTO/SB/08A (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

OIPE MAY 24 2007 PATENT & TRADEMARK OFFICE

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 3

Complete if Known	
Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

U. S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
		US- 4,649,505	03-10-1987	Zinser, Jr. et al.	
		US- 5,319,736	06-07-1994	Hunt	
		US- 5,381,473	01-10-1995	Andrea et al.	
		US- 5,694,474	12-02-1997	Ngo et al.	
		US- 6,002,776	12-14-1999	Bhadkamkar et al.	
		US- 6,061,456	05-09-2000	Andrea et al.	
		US- 6,072,881	06-06-2000	Linder	
		US- 6,222,927	04-24-2001	Feng et al.	
		US- 6,317,501	11-13-2001	Matsuo	
		US- 6,363,345	03-26-2002	Marash et al.	
		US- 6,430,295	08-06-2002	Handel et al.	
		US- 6,584,203	06-24-2003	Elko et al.	
		US- 6,717,991	04-06-2004	Gustafsson et al.	
		US- 6,738,482	05-18-2004	Jaber	
		US- 6,760,450	07-06-2004	Matsuo	
		US- 6,785,381	08-31-2004	Gartner et al.	
		US- 6,795,558	09-21-2004	Matsuo	
		US- 7,099,821	08-29-2006	Visser et al.	
		US- 7,171,008	01-30-2007	Elko	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

Examiner Signature	/Paul Kim/	Date Considered	01/27/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

PTO/SB/08A (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 2 of 3

Complete if Known	
Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

U. S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
		US- 2001/0016020	08-23-2001	Gustafsson et al.	
		US- 2001/0031053	10-18-2001	Feng et al.	
		US- 2002/0009203	01-24-2002	Erten	
		US- 2002/0041693	04-11-2002	Matsuo	
		US- 2002/0080980	06-27-2002	Matsuo	
		US- 2002/0106092	08-08-2002	Matsuo	
		US- 2002/0116187	08-22-2002	Erten	
		US- 2003/0039369	02-27-2003	Bullen	
		US- 2003/0072460	04-17-2003	Gonopolskiy et al.	
		US- 2003/0099345	05-29-2003	Gartner et al.	
		US- 2003/0138116	07-24-2003	Jones et al.	
		US- 2003/0169891	09-11-2003	Ryan et al.	
		US- 2005/0185813	08-25-2005	Sinclair et al.	
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]

Examiner Signature	/Paul Kim/	Date Considered	01/27/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

PTO/SB/08B (07-05)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 3 of 3

Complete if Known	
Application Number	11/699,732
Filing Date	January 29, 2007
First Named Inventor	Carlos Avendano
Art Unit	2615
Examiner Name	Unknown
Attorney Docket Number	PA3909US

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
		AVENDANO, CARLOS, "Frequency-Domain Source Identification and Manipulation in Stereo Mixes for Enhancement, Suppression and Re-panning Applications," 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 19-22, 2003, pp. 55-58, New Paltz, New York, USA.	
		WIDROW, B. et al., "Adaptive Atenna Systems," December 1967, pp. 2143-2159, Vol. 55 No. 12, Proceedings of the IEEE.	
		ELKO, GARY W., "Differential Microphone Arrays," Audio Signal Processing for Next-Generation Multimedia Communication Systems, 2004, pp. 12-65, Kluwer Academic Publishers, Norwell, Massachusetts, USA.	

Examiner Signature	/Paul Kim/	Date Considered	01/27/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

1 Applicant's unique citation designation number (optional). 2 Applicant is to place a check mark here if English language Translation is attached.

This collection of information is required by 37 CFR 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 (1-800-786-9199) and select option 2.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

U.S.PATENTS						
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Patent citation information please click the Add button.						

U.S.PATENT APPLICATION PUBLICATIONS						
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20080260175	A1	2008-10-23	Elko	
If you wish to add additional U.S. Published Application citation information please click the Add button.						

FOREIGN PATENT DOCUMENTS								
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button								

NON-PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☒ See attached certification statement.

☐ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☐ A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/ James Drapinski /	Date (YYYY-MM-DD)	2011-06-30
Name/Print	James Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Electronic Acknowledgement Receipt	
EFS ID:	10434105
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	James William Drapinski
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	30-JUN-2011
Filing Date:	29-JAN-2007
Time Stamp:	22:50:33
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	3909US_IDS_Statement_to_PTO.pdf	78040 1d2a35496b8d0c1c186ef061f876c47e46b21d03	no	2
Warnings:					
Information:					

2	Information Disclosure Statement (IDS) Form (SB08)	3909US_IDS_SB08.pdf	30257 47bd5eea72d6709a9b39d46dd439d59573ae6af2	no	3

Warnings:	
Information:	
This is not an USPTO supplied IDS fillable form	
Total Files Size (in bytes):	108297

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano
APPLICATION NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and method for utilizing omni-directional microphones for speech enhancement
EXAMINER:	Kim, Paul
GROUP ART UNIT:	2614
CONFIRMATION NO.:	7721
ATTY.DKT.NO.:	PA3909US

MAIL STOP AMENDMENT
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. § 1.56, and 1.97-1.98
& Statement Under 37 C.F.R. § 1.97(e)(2)

EXAMINER:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08, listing a reference. The Examiner is requested to make this reference of official record in the application.

The reference cited may be material to examination of the application and is submitted in compliance with the Applicants' duty of disclosure as defined by 37 C.F.R. § 1.56. Additionally, the Applicants wish to make the Examiner aware of, and invite the Examiner to consider, as is appropriate, the following office action from the following U.S. patent application:

- Office Action mailed March 30, 2011 in U.S. patent application number 11/343,524, filed January 30, 2006.

- Office Action mailed April 28, 2011 in U.S. patent application number 11/825,563, filed July 6, 2007.

No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

Applicants hereby declare, pursuant to 37 C.F.R. § 1.97(e)(2), that no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application and, to the knowledge of the person signing the certification after having made a reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 C.F.R. § 1.56(c) more than three months prior to the filing of the information disclosure statement.

The Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600.

Respectfully submitted,
Carlos Avendano

June 30, 2011

By: / James Drapinski /
James Drapinski, Reg. No. 46,242
Carr & Ferrell *LLP*
120 Constitution Drive
Menlo Park, CA 94025
T: 650.812.3400
F: 650.812.3444

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano
APPLICATION NO.:	11/699,732
FILING DATE:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
ART UNIT:	2614
CONF. NO.:	7721
ATTY.DKT. NO.:	PA3909US

MAIL STOP AMENDMENT
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

RESPONSE A

Sir:

In response to the non-final Office Action (*Office Action*) mailed April 7, 2011, the Applicant submits the following response. **Amendments**, **remarks**, and a **conclusion** begin on pages **two**, **nine**, and **thirteen**, respectively. The requisite fee for an extension of time up to the current date is included herewith.

IN THE CLAIMS

This listing of claims will replace all prior versions and listing of claims in the application:

1. (original) A system for enhancing speech, comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal;

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal; and

an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

2. (currently amended) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between [[the]] a delayed primary electric signal and a delayed and level-equalized secondary electric signal.

3. (original) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain.

4. (original) The system of claim 3 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

5. (currently amended) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the ~~level-equalized~~ secondary electric signal and a delayed primary electric signal.

6. (original) The system of claim 1 further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal.

7. (original) The system of claim 1 further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

8. (original) The system of claim 1 further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference.

9. (original) The system of claim 1 further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

10. (original) The system of claim 9 further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

11. (original) The system of claim 1 further comprising a masking module configured to determine a speech estimate.

12. (original) The system of claim 11 further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output.

13. (original) The system of claim 1, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal.

14. (original) A method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; and

non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

15. (currently amended) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between [[the]] a delayed primary electric signal and a delayed secondary electric signal.

16. (original) The method of claim 14 wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain.

17. (original) The method of claim 16 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

18. (original) The method of claim 14 wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal.

19. (original) The method of claim 14 wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal.

20. (original) The method of claim 14 further comprising determining an energy estimate for each of the acoustic signals during a frame.

21. (original) The method of claim 14 further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference.

22. (original) The method of claim 21 further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

23. (original) The method of claim 22 further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal

24. (original) The method of claim 22 further comprising smoothing the filter estimate.

25. (currently amended) The method of claim 14 wherein the cardioid primary signal and the cardioid secondary signal are each ~~is~~ of a sub-band of the primary electric signal.

26. (currently amended) A non-transitory ~~machine~~ computer readable storage medium having embodied thereon a program, the program ~~providing instructions for~~ being executable by a processor to perform a method for enhancing speech, the method comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; and

non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference.

27. (new) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between a level-equalized secondary electric signal and a delayed primary electric signal.

28. (new) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a level-equalized secondary electric signal.

AMENDMENTS TO THE SPECIFICATION

Please replace paragraph [0055] of the specification as originally filed with the following amended paragraph:

A secondary acoustic signal magnitude may be modified to match a magnitude of the primary acoustic signal by applying a gain which is computed by the gain module 406. The gain module 406 computes the magnitude of both signals (e.g., X_1 and X_2) and derives the gain, g, as the ratio between the magnitude of the primary acoustic signal to the magnitude of the secondary acoustic signal. The gain can then be used to calculate the cardioid primary signal and the cardioid secondary signal ~~[Notice the change I made to the figure CA]~~.

REMARKS

In the *Office Action*, the Examiner rejected claims 1-26. No claims are canceled by this response. Claims 2, 5, 15, 25, and 26 are amended. Claims 27 and 28 are added. No new matter is introduced. Support for the amended claims may be found, for example, at paragraphs [0021], [0035], and [0036] of the pending application (*Instant Application*) as originally filed. Based on the following remarks, the Applicant respectfully requests reconsideration of the *Instant Application*.

Rejection under § 101

Claim 26 was rejected under 35 U.S.C. § 101 for being directed to non-statutory subject matter. Although the Applicant does not necessarily agree with the Examiner, claim 26 is amended to recite a "non-transitory computer readable storage medium" in order to advance prosecution of the *Instant Application*. The Applicant respectful requests that the rejection under § 101 be withdrawn.

Rejection under § 103(a)

The Examiner rejected claims 1-26 under 35 U.S.C. § 103(a) as obvious over U.S. Patent Application Publication Number 2003/0147538 (*Elko*) in view of U.S. Patent Application Publication Number 2003/0169891 (*Ryan*). Because the combination of *Elko* and *Ryan* fails to render obvious each limitation of claims 1-26, the Applicant asserts that these claims are patentable over the art.

The cited art does not teach or suggest "an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference."

Claim 1 as amended recites, among other features: "an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference." Components of the cardioid primary signal and the cardioid secondary signal may be non-linearly combined, for example, by taking a ratio. *Instant Application*, paragraph [0039]. The Examiner acknowledged that *Elko* does not teach this claimed feature: "Elko fails to disclose an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference." *Office Action*, 4. Instead, the Examiner relied on *Ryan* to teach this feature: "(see paragraphs 38 and 40)."

The Applicant respectfully disagrees with the Examiner's interpretation of *Ryan*. *Ryan* merely teaches splitting "the single input signal into four distinct frequency bands," channel processing the four distinct frequency bands individually with a plurality of channel processing/twin detector blocks, and then "the four channel signals are summed by a summer." *Ryan*, Figure 3B, and paragraphs [0035]-[0038] and [0040]. As known in the art, a linear combination is an expression constructed from a set of terms by multiplying each term by a constant and adding the results. Clearly, the processing and subsequent summation of the four channel signals is a linear combination. In contrast, the claimed invention includes a non-linear combination. Thus *Elko* and *Ryan*, either individually or in theoretical combination, do not teach the claimed "an inter-microphone level difference module configured to non-linearly

combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference."

The cited art does not teach or suggest "the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal."

Claim 2 as amended recites, among other features: "the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal." The Examiner acknowledged that *Elko* does not teach this feature: "Elko fails to disclose wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level equalized secondary electric signal." *Office Action*, 5. The Examiner suggested that *Ryan* discloses this feature: "(see paragraph 70)." *Office Action*, 5.

The Applicant respectfully disagrees with the Examiner's interpretation of *Ryan*. *Ryan* teaches that only the one output of the summation circuit 118 is equalized. *Ryan*, Figure 5 and paragraph [0074]. That is, neither the front or rear microphone signal is equalized. In contrast, the claimed secondary electric signal is equalized. Hence *Elko* and *Ryan*, either individually or in theoretical combination, do not disclose the claimed "DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal."

To support a conclusion that the claim would have been obvious requires that **all** the claimed elements were known in the prior art and that one skilled in the art could have combined those elements. See *KSR International Co. v. Teleflex Inc.*, 127 S.Ct. 1727, 1739 (2007) (emphasis added); see also MPEP § 2143. Since *Elko* in view of *Ryan* fails to

disclose all the claimed elements of independent claim 1, the rejection of claim 1 under Section 103 is overcome and thus claim 1 is patentable. Additionally, as independent claims 14 and 26 include similar elements to those of independent claim 1, claims 14 and 26 are likewise patentable for at least the same reasons. Similarly, *Elko* in view of *Ryan* fails to disclose all the claimed elements of dependent claim 2, and the rejection of claim 2 under Section 103 is overcome and thus claim 2 is patentable. In addition, dependent claim 15 includes similar elements to those of dependent claim 2, so claim 15 is likewise patentable for at least similar reasons.

Furthermore, as a dependent claim incorporates by reference all the limitations of the claim from which it depends (see 35 U.S.C. § 112 ¶ 4), claims 2-13, 16-25, and 27-28 are patentable for at least the same reasons as the independent claim from which they depend in addition to their separate distinguishing characteristics. The Applicant has, for brevity, chosen not to individually discuss the additional distinguishing characteristics of the dependent claims, aside from claims 2 and 15. The Applicant's silence should not be interpreted as belief that the dependent claims are patentable only by virtue of their dependence on a patentable independent claim.

CONCLUSION

It is respectfully submitted that the *Instant Application* is now in condition for allowance, and such action is requested. If the Examiner has any questions regarding the case, the Examiner is invited to contact the Applicant's undersigned representative.

The requisite fee for an extension of time up to the current date is included herewith. The Commissioner is hereby authorized to charge any underpayment or credit any overpayment to Deposit Account No. 06-0600 for any matter in connection with this response, including any fee for extension of time up to the current date, which may be required.

Respectfully submitted,
Carlos Avendano

August 4, 2011

By: /Bruce C. Shem/

Bruce C. Shem, Reg. no. 64,914
CARR & FERRELL *LLP*
120 Constitution Drive
Menlo Park, CA 94025
TEL: (650) 812-3400
FAX: (650) 812-3444

Electronic Patent Application Fee Transmittal	
Application Number:	11699732
Filing Date:	29-Jan-2007
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Filer:	Bruce Clovis Shem/Carol Tran
Attorney Docket Number:	PA3909US
Filed as Small Entity	

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Claims in excess of 20	2202	2	26	52
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension - 1 month with $0 paid	2251	1	65	65
Miscellaneous:				
		Total in USD ($)		**117**

Electronic Acknowledgement Receipt	
EFS ID:	10666439
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	Bruce Clovis Shem
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	04-AUG-2011
Filing Date:	29-JAN-2007
Time Stamp:	14:51:53
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$117
RAM confirmation Number	1127
Deposit Account	060600
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)	
Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)	
Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		Audience_3909US_Amendment_Response_A.pdf	131459 14791be0136339dc9c9a1f00d187dcf0cc97ce20	yes	13

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Specification	1	1
Claims	2	8
Applicant Arguments/Remarks Made in an Amendment	9	13

Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	31502 6ece0617f8c903841a3dde256b39c4e59f1faca3	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	162961		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☐ To be Mailed
11/699,732	01/29/2007	

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY ☒ RATE ($)	SMALL ENTITY ☒ FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
	☐ Application Size Fee (37 CFR 1.16(s))									
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

AMENDMENT	08/04/2011	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)	OR	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 28	Minus	** 26	= 2	X $26 =	52	OR	X $ =	
	Independent (37 CFR 1.16(h))	* 3	Minus	*** 3	= 0	X $110 =	0	OR	X $ =	
	☐ Application Size Fee (37 CFR 1.16(s))									
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE	**52**	OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/Chantae Dessau/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 12/06/2011

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	
ART UNIT	**PAPER NUMBER**
2614	
MAIL DATE	**DELIVERY MODE**
12/06/2011	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	11/699,732	AVENDANO, CARLOS	
	Examiner	Art Unit	
	PAUL KIM	2614	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *04 August 2011*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-28* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-28* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *6/30/2011*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 1-28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Elko (US Pub. 20030147538) in view of Ryan (US Pub. 20030169891).

Regarding claim 1, Elko discloses a system for enhancing speech (see abstract), comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal (see paragraphs 8 and 95);

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal (see paragraph 26).

Elko fails to disclose an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

However, Ryan discloses an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference (see paragraphs 38 and 40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's inter-microphone level difference module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 2, Elko in view of Ryan discloses a system of claim 1, as discussed above.

Elko fails to disclose wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal.

However, Ryan discloses wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal (see paragraph 70).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's DMA module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 3, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain (see Elko, paragraphs 99 and 102).

Regarding claim 4, Elko in view of Ryan discloses a system of claim 3, as discussed above.

Elko fails to disclose wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

However, Ryan discloses wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal (see paragraph 42).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's gain with a system for enhancing speech of Elko. The motivation would be to provide a well balanced value for the gain.

Regarding claim 5, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the secondary electric signal and a delayed primary electric signal (see Elko, paragraphs 99 and 102).

Regarding claim 6, Elko in view of Ryan discloses a system of claim 1, as discussed above, further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal (see Elko, paragraph 63).

Regarding claim 7, Elko in view of Ryan discloses a system of claim 1, as discussed above.

Elko fails to disclose further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

However, Ryan discloses further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal (see paragraph 39).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's energy module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 8, Elko in view of Ryan discloses a system of claim 1, as discussed above, further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference (see Elko, paragraph 51).

Regarding claim 9, Elko in view of Ryan discloses a system of claim 1, as discussed above.

Elko fails to disclose further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

However, Ryan discloses further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal (see paragraph 28).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's filter module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 10, Elko in view of Ryan discloses a system of claim 9, as discussed above.

Elko fails to disclose further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

However, Ryan discloses further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal (see paragraph 28).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's filter smoothing module with a system for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 11, Elko in view of Ryan discloses a system of claim 1, as discussed above, further comprising a masking module configured to determine a speech estimate (see Elko, paragraph 52).

Regarding claim 12, Elko in view of Ryan discloses a system of claim 11, as discussed above, further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output (see Elko, paragraph 52).

Regarding claim 13, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal (see Elko, paragraph 105).

Regarding claim 14, Elko discloses a method for enhancing speech (see abstract), comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone (see paragraphs 8 and 95);

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal (see paragraph 26).

Elko fails to disclose non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

However, Ryan discloses non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference (see paragraphs 38 and 40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's non-linearly combining components with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 15, Elko in view of Ryan discloses a method of claim 14, as discussed above.

Elko fails to disclose wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a delayed secondary electric signal.

However, Ryan discloses wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a delayed secondary electric signal (see paragraph 70).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's method of determining the cardioid primary signal with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 16, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain (see Elko, paragraphs 99 and 102).

Regarding claim 17, Elko in view of Ryan discloses a method of claim 16, as discussed above.

Elko fails to disclose wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

However, Ryan discloses wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal (see paragraph 42).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's gain with a method for enhancing speech of Elko. The motivation would be to provide a well balanced value for the gain.

Regarding claim 18, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal (see Elko, paragraphs 99 and 102).

Regarding claim 19, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal (see Elko, paragraph 29).

Regarding claim 20, Elko in view of Ryan discloses a method of claim 14, as discussed above.

Elko fails to disclose further comprising determining an energy estimate for each of the acoustic signals during a frame.

However, Ryan discloses further comprising determining an energy estimate for each of the acoustic signals during a frame (see paragraph 39).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's energy module with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 21, Elko in view of Ryan discloses a method of claim 14, as discussed above, further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference (see Elko, paragraph 51).

Regarding claim 22, Elko in view of Ryan discloses a method of claim 21, as discussed above.

Elko fails to disclose further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

However, Ryan discloses further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference (see paragraphs 39-40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's determining a filter estimate with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 23, Elko in view of Ryan discloses a method of claim 22, as discussed above, further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal (see Elko, paragraph 52).

Regarding claim 24, Elko in view of Ryan discloses a method of claim 22, as discussed above.

Elko fails to disclose further comprising smoothing the filter estimate.

However, Ryan discloses further comprising smoothing the filter estimate (see paragraph 28).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's smoothing the filter estimate with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 25, Elko in view of Ryan discloses a method of claim 14, as discussed above, wherein the cardioid primary signal and the cardioid secondary signal are each of a sub-band of the primary electric signal (see Elko, paragraph 105).

Regarding claim 26, Elko discloses a non-transitory computer readable storage medium having embodied thereon a program, the program being executable by a processor to perform a method for enhancing speech (see abstract; paragraph 119), the method comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone (see paragraphs 8 and 95);

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal (see paragraph 26).

Elko fails to disclose non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference.

However, Ryan discloses non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference (see paragraphs 38 and 40).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's non-linearly combining components with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Regarding claim 27, Elko in view of Ryan discloses a system of claim 1, as discussed above, wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between a level-equalized secondary electric signal and a delayed primary electric signal (see Elko, paragraphs 99 and 102).

Regarding claim 28, Elko in view of Ryan discloses a method of claim 14, as discussed above.

Elko fails to disclose wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a level-equalized secondary electric signal.

However, Ryan discloses wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a level-equalized secondary electric signal (see paragraph 70).

Therefore, it would've been obvious to a person of ordinary skill in the art, at the time of invention, to include Ryan's signal determining method with a method for enhancing speech of Elko. The motivation would be to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument.

Response to Arguments

3. Applicant's arguments filed 8/4/2011 have been fully considered but they are not persuasive. Applicant is reminded that the specifications are not the measure of invention. Therefore, limitations contained therein cannot be read into the claims for the purpose of avoiding the prior art. In re Sporck, 55 CCPA 743, 386 F.2d 924, 155 USPQ 687 (1968). It is not necessary that the references actually suggest, expressly or in so many words, the changes or improvements that applicant has made. The test for combining references is what the references as a whole would have suggested to one of ordinary skill in the art. In re Sheckler, 168 USPQ 716 (CCPA 1971); In re McLaughlin 170 USPQ 209 (CCPA 1971); In re Young 159 USPQ 725 (CCPA 1968). In response to applicant's arguments that the prior art references fail to disclose the claimed limitations "an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference" and "wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal," the limitations have been interpreted reasonably broad to read on the prior art references, as discussed above.

Conclusion

4. THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PAUL KIM whose telephone number is (571)270-7697. The examiner can normally be reached on Monday - Thursday, 8 AM - 4 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Vivian Chin can be reached on (571) 272-7848. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/P. K./
Examiner, Art Unit 2614

/VIVIAN CHIN/
Supervisory Patent Examiner, Art Unit 2614

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11699732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

SEARCHED			
Class	Subclass	Date	Examiner
381	313, 312, 91, 92, 122	1/27/2011	PK
381	95, 110, 122, 94.1, 94.2, 94.3, 94.7	12/2/2011	PK
704	226, 227, 233, 275	12/2/2011	PK

SEARCH NOTES		
Search Notes	Date	Examiner
EAST - see EAST search notes	1/26/2011	PK
Inventor search - Carlos Avendano	1/26/2011	PK
Search consultation - Xu Mei (Primary, AU 2614)	8/10/2011	PK
Updated search - see EAST search notes	12/2/2011	PK

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner
USPGPUBS Text Search		12/2/2011	PK

/P. K./ Examiner.Art Unit 2614	

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

U.S.PATENTS						
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Patent citation information please click the Add button.						

U.S.PATENT APPLICATION PUBLICATIONS						
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20080260175	A1	2008-10-23	Elko	
If you wish to add additional U.S. Published Application citation information please click the Add button.						

FOREIGN PATENT DOCUMENTS								
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button								

NON-PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

EFS Web 2.1.17

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 06/30/2011	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	12/02/2011

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L2	6224	((381/91,92,95,110,122) or (704/226,233,275)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 13:54
L3	1835	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer))) and (distance or separat$4)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:54
L4	164	L3 and L2	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:54
L5	1508	((381/94.1,94.2,94.3,94.7) or (704/227)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 13:55
L6	67	L3 and L5	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
L7	21458	(inter with (level or microphone))	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
L8	80	L7 and l2	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
L9	33	L7 and l5	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
S1	1	(11/699732).APP.	US-PGPUB; USPAT	OR	OFF	2011/01/27 05:56
S2	20	((CARLOS) near2 (AVENDANO)).INV.	US-PGPUB; USPAT	ADJ	ON	2011/01/27 06:07
S3	53	("20010016020" | "20010031053" | "20020009203" | "20020080980" | "20020106092" | "20020116187" | "20030039369" | "20030072460" | "20030099345" | "20030138116" | "20030169891" | "20050185813" | "4649505" | "4658426" | "5208864" | "5319736" | "5381473" | "5402496" | "5574824" | "5602962" | "5694474" | "5757937" | "5796819" | "6002776" | "6061456" | "6072881" | "6222927" | "6317501" | "6363345" | "6381176" | "6430295" | "6549630" | "6584203" | "6717991" | "6738482" | "6760450" | "6760805" | "6785381" | "6795558" | "6831865" | "6882736" | "6917688" | "6978159" |	US-PGPUB; USPAT	ADJ	ON	2011/01/27 06:09

		"7031478" | "7089349" | "7099821" | "7142677" | "7146316" | "7155019" | "7171008" | "7174022" | "7206418" | "7246058").PN.				
S5	1	("20030147538").PN.	US-PGPUB; USPAT	OR	OFF	2011/01/27 07:30
S6	1	(11/343524).APP.	US-PGPUB; USPAT	OR	OFF	2011/01/27 07:33
S7	2615	(381/313,312,91,92,122).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/01/27 15:40
S8	55	("6381176" | "20020116187" | "20050185813" | "5319736" | "5402496" | "7206418" | "20030169891" | "5796819" | "6549630" | "7146316" | "6831865" | "5381473" | "6795558" | "7099821" | "7155019" | "7246058" | "20020106092" | "20030072460" | "6317501" | "6760450" | "5208864" | "5574824" | "6882736" | "20030147538" | "6760805" | "20020009203" | "20030039369" | "6430295" | "6584203" | "6717991" | "5602962" | "6917688" | "20010016020" | "20030099345" | "6002776" | "6061456" | "6363345" | "7171008" | "6978159" | "20020080980" | "5694474" | "4658426" | "5757937" | "7142677" | "7089349" | "20010031053" | "20030138116" | "4649505" | "6072881" | "6222927" | "6738482" | "6785381" | "7031478" | "7174022" | "20080260175").PN.	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 08:23
S9	1	("20060133621").PN.	US-PGPUB; USPAT; USOCR	OR	OFF	2011/12/02 08:26
S10	7369	speech with enhanc$6	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:46
S11	1	(speech with enhanc$6) same (primary with secondary with microphone$2) same (different$5 with microphone$2 with array)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:48
S12	12	(speech with enhanc$6) same (different$5 with microphone$2 with array)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:48
S13	111	(speech with enhanc$6) and (different$5 with microphone$2 with array)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:49
S14	1	(speech with enhanc$6) same (cardioid with primary with secondary with signal)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:52
S15	3	(speech with enhanc$6) same (microphone$2 with array) same (primary with secondary with signal)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:52
S16	6224	((381/91,92,95,110,122) or (704/226,233,275)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 12:57
S17	1508	((381/94.1,94.2,94.3,94.7) or (704/227)).CCLS.	US-PGPUB;	OR	OFF	2011/12/02 12:59

			USPAT			
S18	1835	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer))) and (distance or separat$4)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:03
S19	2321	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer)))	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:04
S20	9	(inter-microphone with level with differenc$4)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:06
S21	21458	(inter with (level or microphone))	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:10
S22	58	S19 and S21	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:11

12/2/2011 1:56:57 PM
C:\Users\pkim2\Documents\EAST\Workspaces\11699732.wsp

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11699732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	01/27/2011	12/02/2011							
	1	✓	✓							
	2	✓	✓							
	3	✓	✓							
	4	✓	✓							
	5	✓	✓							
	6	✓	✓							
	7	✓	✓							
	8	✓	✓							
	9	✓	✓							
	10	✓	✓							
	11	✓	✓							
	12	✓	✓							
	13	✓	✓							
	14	✓	✓							
	15	✓	✓							
	16	✓	✓							
	17	✓	✓							
	18	✓	✓							
	19	✓	✓							
	20	✓	✓							
	21	✓	✓							
	22	✓	✓							
	23	✓	✓							
	24	✓	✓							
	25	✓	✓							
	26	✓	✓							
	27		✓							
	28		✓							

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano
APPLICATION NO.:	11/699,732
FILING DATE:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
ART UNIT:	2614
CONF. NO.:	7721
ATTY.DKT. NO.:	PA3909US

MAIL STOP AF
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

RESPONSE B

Examiner Kim:

In response to the Final Office Action (*Final Office Action*) mailed December 6, 2011, the Applicant submits the following response. **Listing of Claims**, **Remarks**, and **Conclusion** begin on pages **two**, **seven**, and **thirteen**, respectively.

IN THE CLAIMS

This listing of claims will replace all prior versions and listing of claims in the application:

1. (original) A system for enhancing speech, comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal;

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal; and

an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

2. (previously presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal.

3. (original) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain.

4. (original) The system of claim 3 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

5. (previously presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the secondary electric signal and a delayed primary electric signal.

6. (original) The system of claim 1 further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal.

7. (original) The system of claim 1 further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

8. (original) The system of claim 1 further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference.

9. (original) The system of claim 1 further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

10. (original) The system of claim 9 further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

11. (original) The system of claim 1 further comprising a masking module configured to determine a speech estimate.

12. (original) The system of claim 11 further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output.

13. (original) The system of claim 1, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal.

14. (original) A method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; and

non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

15. (previously presented) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a delayed secondary electric signal.

16. (original) The method of claim 14 wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain.

17. (original) The method of claim 16 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

18. (original) The method of claim 14 wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal.

19. (original) The method of claim 14 wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal.

20. (original) The method of claim 14 further comprising determining an energy estimate for each of the acoustic signals during a frame.

21. (original) The method of claim 14 further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference.

22. (original) The method of claim 21 further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

23. (original) The method of claim 22 further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal

24. (original) The method of claim 22 further comprising smoothing the filter estimate.

25. (previously presented) The method of claim 14 wherein the cardioid primary signal and the cardioid secondary signal are each of a sub-band of the primary electric signal.

26. (previously presented) A non-transitory computer readable storage medium having embodied thereon a program, the program being executable by a processor to perform a method for enhancing speech, the method comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; and

non-linearly combining components of the cardioid primary signal and the cardioid primary signal to obtain an inter-microphone level difference.

27. (previously presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between a level-equalized secondary electric signal and a delayed primary electric signal.

28. (previously presented) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a level-equalized secondary electric signal.

REMARKS

In the *Final Office Action*, the Examiner rejected claims 1-28. No claims are canceled, amended, or added by this response. Claims 1-28 remain pending. Based on the following remarks, the Applicant respectfully requests reconsideration of the pending application (*Instant Application*).

Rejection under § 101

In the Office Action mailed April 7, 2011 (*Previous Office Action*), Claim 26 was rejected under 35 U.S.C. § 101 for being directed to non-statutory subject matter. In the Response filed on August 4, 2011 (*Response A*), claim 26 was amended to recite a "non-transitory computer readable storage medium" in order to advance prosecution of the *Instant Application*. Since the Examiner did not explicitly address the § 101 rejection in the *Final Office Action*, the Applicant believes that the § 101 rejection is withdrawn. The Applicant respectfully requests confirmation.

Rejection under § 103(a)

The Examiner rejected claims 1-28 under 35 U.S.C. § 103(a) as obvious over U.S. Patent Application Publication Number 2003/0147538 (*Elko*) in view of U.S. Patent Application Publication Number 2003/0169891 (*Ryan*). Because the theoretical combination of *Elko* and *Ryan* fails to render obvious each limitation of claims 1-26, the Applicant asserts that these claims are patentable over the art.

The cited art does not teach or suggest (a) "an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference" and (b) "the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal."

In *Response A*, the Applicant explained that: (1) the Examiner acknowledged that *Elko* does not teach these claimed features (*Previous Office Action*, 4-5) and (2) *Ryan* did not teach the claimed features, either. *Response A*, 10-12. For brevity, the explanation is not repeated herein.

In the *Final Office Action*, the Examiner appears to agree with the Applicant that *Elko* and *Ryan* do not teach the claimed features, because as understood, the Examiner states that such teaching or suggestion is unnecessary:

> It is not necessary that the references actually suggest, expressly or in so many words, the changes or improvements that the applicant has made.
> *Final Office Action*, 15.

The Applicant respectfully disagrees.

The Applicant maintains that to support a conclusion that the claim would have been obvious **requires that all the claimed elements were known in the prior art** and that one skilled in the art could have combined those elements. See *KSR International Co. v. Teleflex Inc.*, 127 S.Ct. 1727, 1739 (2007) (emphasis added); see also MPEP § 2143. The Applicant further maintains that neither *Elko* nor *Ryan* disclose the claimed **non-linear** combination and a delayed and **level-equalized** secondary electric signal.

The Examiner suggested that the Applicant is improperly imputing limitations from the *Instant Application* into the claims:

> Applicant is reminded that the specifications are not the measure of invention. Therefore, limitations contained therein cannot be read into the claims for the purpose of avoiding prior art. In re Sporck, 55 CCPA 743, 386 F.2d 924, 155 USPQ 687 (1968).
> *Final Office Action*, 15.

However, the limitations which the Applicant explained distinguish the claims over the prior art are clearly contained within the claims. For example, claim 1 recites in part "configured to **non-linearly** combine components." By way of further example, claim 2 in part recites "a delay and **level-equalized secondary electric signal**."

The Examiner contends that the claim limitations given their **broadest reasonable interpretation** are taught by *Elko* and *Ryan*:

> In response to applicant's arguments that the prior art references fail to disclose the claimed limitations "an inter-microphone level difference module configured to nonlinearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference" and "wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal," the limitations have been **interpreted reasonably broad to read on the prior art references**, as discussed above.

Final Office Action, 15 (emphasis added).

The Applicant respectfully disagrees with the Examiner's interpretation of the claims.

The broadest reasonable interpretation must be consistent with the specification:

> The Patent and Trademark Office ("PTO") determines the scope of claims in patent applications not solely on the basis of the claim language, but upon giving claims their broadest reasonable construction "**in light of the specification as it would be interpreted by one of ordinary skill in the art**." *In re Am. Acad. of Sci. Tech. Ctr.*, 367 F.3d 1359, 1364[, 70 USPQ2d 1827] (Fed. Cir. 2004).

MPEP, § 2111 (quoting *Phillips v. AWH Corp.*, 415 F.3d 1303, 75 USPQ2d 1321 (Fed. Cir. 2005). Emphasis added).

Applicant respectfully submits that a reasonably broad interpretation of the claims in light of the specification does **not** read on *Elko* and *Ryan*.

For example, claim 1 recites in part "configured to **non-linearly** combine components." The Examiner contended that ¶¶ [0038] and [0040] of *Ryan* disclose the claimed feature (*Final Office Action*, 3):

> The channel processing blocks 58A-58D can be configured to employ a twin detector average detection scheme while compressing the input signals. This twin detection scheme includes both slow and fast attack/release tracking modules that allow for fast response to transients (in the fast tracking module), while preventing annoying pumping of the input signal (in the slow tracking module) that only a fast time constant would produce. The outputs of the fast and slow tracking modules are compared, and the compression slope is then adjusted accordingly. The compression ratio, channel gain, lower and upper thresholds (return to linear point), and the fast and slow time constants (of the fast and slow tracking modules) can be independently programmed and saved in memory 44 for each of the plurality of channel processing blocks 58A-58D.
>
> *Ryan*, ¶ [0038].

> After channel processing is complete, the four channel signals are **summed by summer** 60 to form a composite signal. This composite signal is then coupled to the post-filter 62, which may apply a post-processing filter function as discussed above. Following post-processing, the composite signal is then applied to a notch-filter 64, that attenuates a narrow band of frequencies that is adjustable in the frequency range where hearing aids tend to oscillate. This notch filter 64 is used to reduce feedback and prevent unwanted "whistling" of the device. Preferably, the notch filter 64 may include a dynamic transfer function that changes the depth of the notch based upon the magnitude of the input signal.
>
> *Ryan*, ¶ [0040] (emphasis added).

Clearly, at best *Ryan* teaches a **linear** combination (i.e., summing) and does not teach or suggest a **non-linear** combination. *Response A*, 10-11. In contrast, at least ¶ [0040] of the *Instant Application* as filed teaches "**non-linearly** combining." At the time of invention, one of ordinary skill in the art would not substitute a linear combination with a non-linear combination, or otherwise use one in the place of the other. The broadest reasonable interpretation of the claims and specification of the *Instant Application* does **not** read on the **linear** combination of *Ryan*.

By way of further example, claim 2 in part recites "a delay and **level-equalized secondary electric signal**." The Examiner contended that ¶ [0070] of *Ryan* discloses the claimed feature (*Final Office Action*, 3):

> The front and rear microphones 110, 112 are preferably omnidirectional microphones that receive an acoustical waveform and generate a front and rear microphone signal, respectively. The front microphone signal is coupled to the front allpass filter 114, and the rear microphone signal is coupled to the time delay circuit 115. The time delay circuit 115 implements a time-of-flight delay that compensates for the distance between the front and rear microphones 110, 112 and determines the specific nature of the directional microphone pattern (i.e., cardioid, hyper-cardioid, bi-directional, etc.).
>
> *Ryan*, ¶ [0070].

Clearly, *Ryan* does not teach or suggest that either the front or the rear microphone signal is equalized. *Response A*, 11. In contrast, at least ¶ [0052] of the *Instant Application* teaches "a filter 402, F(z), that equalizes the microphones." At the time of invention, one of ordinary skill in the art would not replace an equalized signal with a non-equalized signal, or otherwise use one in the place of the other. The broadest reasonable interpretation of the claims and specification of the *Instant Application* does **not** read on the **non-equalized** microphones of *Ryan*.

The Examiner's statement that "the limitations have been interpreted reasonably broad to read on the prior art references" is conclusory. A clear articulation of the rationale or reason(s) behind the Examiner's broad interpretation is **not** present in the *Final Office Action* (*Final Office Action*, 15), as required by MPEP § 2142:

> [R]ejections on obviousness cannot be sustained with mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness
>
> MPEP, § 2142 (quoting *In re Kahn*, 441 F.3d 977, 988, 78 USPQ2d 1329, 1336 (Fed. Cir. 2006)).

Accordingly, in the event that the Examiner maintains the § 103 rejection of the claims, the Applicant respectfully requests the finality of the *Final Office Action* be withdrawn and the reasoning provided.

As explained above, the rejection of claim 1 under § 103 is overcome and thus claim 1 is patentable. Additionally, as independent claims 14 and 26 include similar elements to those of independent claim 1, claims 14 and 26 are likewise patentable for at least the same reasons. Similarly, *Elko* in view of *Ryan* fails to disclose all the claimed elements of dependent claim 2, and the rejection of claim 2 under § 103 is overcome and thus claim 2 is patentable. In addition, dependent claim 15 includes similar elements to those of dependent claim 2, so claim 15 is likewise patentable for at least similar reasons.

Furthermore, as a dependent claim incorporates by reference all the limitations of the claim from which it depends (see 35 U.S.C. § 112 ¶ 4), claims 2-13, 16-25, and 27-28 are patentable for at least the same reasons as the independent claim from which they depend in addition to their separate distinguishing characteristics. The Applicant has, for brevity, chosen not to individually discuss the additional distinguishing characteristics of the dependent claims, aside from claims 2 and 15. The Applicant's silence should not be interpreted as belief that the dependent claims are patentable only by virtue of their dependence on a patentable independent claim.

CONCLUSION

It is respectfully submitted that the *Instant Application* is now in condition for allowance, and such action is requested. If the *Instant Application* is not allowed, then the Applicant requests that the finality of the *Final Office Action* be withdrawn, because the final rejection is premature as explained above. Should the Examiner have any questions regarding the case, the Examiner is invited to contact the Applicant's undersigned representative.

The Commissioner is hereby authorized to charge any underpayment or credit any overpayment to Deposit Account No. 06-0600 for any matter in connection with this response, including any fee for extension of time up to the current date, which may be required.

Respectfully submitted,
Carlos Avendano

February 6, 2012

By: /Bruce C. Shem/

Bruce C. Shem, Reg. no. 64,914
CARR & FERRELL *LLP*
120 Constitution Drive
Menlo Park, CA 94025
TEL: (650) 812-3400
FAX: (650) 812-3444

Electronic Acknowledgement Receipt	
EFS ID:	12003542
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	Bruce Clovis Shem
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	06-FEB-2012
Filing Date:	29-JAN-2007
Time Stamp:	17:19:35
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		3909US_Response.pdf	145926 75bf38cd1bd9a02f4116b20c513b8c4c54939e5d	yes	13

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Amendment After Final	1	1
	Claims	2	6
	Applicant Arguments/Remarks Made in an Amendment	7	13
Warnings:			
Information:			
	Total Files Size (in bytes):	145926	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☒ To be Mailed
11/699,732	01/29/2007	

APPLICATION AS FILED – PART I

	(Column 1)	(Column 2)	SMALL ENTITY ☒		OR	OTHER THAN SMALL ENTITY	
FOR	NUMBER FILED	NUMBER EXTRA	RATE ($)	FEE ($)		RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT 02/06/2012	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 28	Minus	** 28	= 0	X $30 =	0	OR	X $ =	
Independent (37 CFR 1.16(h))	* 3	Minus	***3	= 0	X $125 =	0	OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE	0	OR	TOTAL ADD'L FEE	

AMENDMENT	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/NICHELE PETERSON/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 02/14/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	
ART UNIT	**PAPER NUMBER**
2614	
MAIL DATE	**DELIVERY MODE**
02/14/2012	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Advisory Action **Before the Filing of an Appeal Brief**	**Application No.** *11/699,732*	**Applicant(s)** AVENDANO, CARLOS
	Examiner PAUL KIM	**Art Unit** 2614

--The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

THE REPLY FILED 06 February 2012 FAILS TO PLACE THIS APPLICATION IN CONDITION FOR ALLOWANCE.

1. ☒ The reply was filed after a final rejection, but prior to or on the same day as filing a Notice of Appeal. To avoid abandonment of this application, applicant must timely file one of the following replies: (1) an amendment, affidavit, or other evidence, which places the application in condition for allowance; (2) a Notice of Appeal (with appeal fee) in compliance with 37 CFR 41.31; or (3) a Request for Continued Examination (RCE) in compliance with 37 CFR 1.114. The reply must be filed within one of the following time periods:
   a) ☐ The period for reply expires _____months from the mailing date of the final rejection.
   b) ☒ The period for reply expires on: (1) the mailing date of this Advisory Action, or (2) the date set forth in the final rejection, whichever is later. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of the final rejection.
   Examiner Note: If box 1 is checked, check either box (a) or (b). ONLY CHECK BOX (b) WHEN THE FIRST REPLY WAS FILED WITHIN TWO MONTHS OF THE FINAL REJECTION. See MPEP 706.07(f).

Extensions of time may be obtained under 37 CFR 1.136(a). The date on which the petition under 37 CFR 1.136(a) and the appropriate extension fee have been filed is the date for purposes of determining the period of extension and the corresponding amount of the fee. The appropriate extension fee under 37 CFR 1.17(a) is calculated from: (1) the expiration date of the shortened statutory period for reply originally set in the final Office action; or (2) as set forth in (b) above, if checked. Any reply received by the Office later than three months after the mailing date of the final rejection, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

NOTICE OF APPEAL

2. ☐ The Notice of Appeal was filed on _____. A brief in compliance with 37 CFR 41.37 must be filed within two months of the date of filing the Notice of Appeal (37 CFR 41.37(a)), or any extension thereof (37 CFR 41.37(e)), to avoid dismissal of the appeal. Since a Notice of Appeal has been filed, any reply must be filed within the time period set forth in 37 CFR 41.37(a).

AMENDMENTS

3. ☐ The proposed amendment(s) filed after a final rejection, but prior to the date of filing a brief, will not be entered because
   (a) ☐ They raise new issues that would require further consideration and/or search (see NOTE below);
   (b) ☐ They raise the issue of new matter (see NOTE below);
   (c) ☐ They are not deemed to place the application in better form for appeal by materially reducing or simplifying the issues for appeal; and/or
   (d) ☐ They present additional claims without canceling a corresponding number of finally rejected claims.
   NOTE: _____. (See 37 CFR 1.116 and 41.33(a)).
4. ☐ The amendments are not in compliance with 37 CFR 1.121. See attached Notice of Non-Compliant Amendment (PTOL-324).
5. ☐ Applicant's reply has overcome the following rejection(s): _____.
6. ☐ Newly proposed or amended claim(s) _____ would be allowable if submitted in a separate, timely filed amendment canceling the non-allowable claim(s).
7. ☐ For purposes of appeal, the proposed amendment(s): a) ☐ will not be entered, or b) ☐ will be entered and an explanation of how the new or amended claims would be rejected is provided below or appended.
   The status of the claim(s) is (or will be) as follows:
   Claim(s) allowed: _____.
   Claim(s) objected to: _____.
   Claim(s) rejected: _____.
   Claim(s) withdrawn from consideration: _____.

AFFIDAVIT OR OTHER EVIDENCE

8. ☐ The affidavit or other evidence filed after a final action, but before or on the date of filing a Notice of Appeal will not be entered because applicant failed to provide a showing of good and sufficient reasons why the affidavit or other evidence is necessary and was not earlier presented. See 37 CFR 1.116(e).
9. ☐ The affidavit or other evidence filed after the date of filing a Notice of Appeal, but prior to the date of filing a brief, will not be entered because the affidavit or other evidence failed to overcome all rejections under appeal and/or appellant fails to provide a showing a good and sufficient reasons why it is necessary and was not earlier presented. See 37 CFR 41.33(d)(1).
10. ☐ The affidavit or other evidence is entered. An explanation of the status of the claims after entry is below or attached.

REQUEST FOR RECONSIDERATION/OTHER

11. ☒ The request for reconsideration has been considered but does NOT place the application in condition for allowance because: Applicant is reminded that the test for obviousness is not whether the features of the reference may be bodily incorporated into the other to produce the claimed subject matter but simply what the references make obvious to one of ordinary skill in the art. The claimed limitations have been interpreted reasonably broad to read on the prior art references, as discussed in the final rejection.
12. ☐ Note the attached Information *Disclosure Statement*(s). (PTO/SB/08) Paper No(s). _____
13. ☐ Other: _____.

/VIVIAN CHIN/ Supervisory Patent Examiner, Art Unit 2614	/P. K./ Examiner, Art Unit 2614

U.S. Patent and Trademark Office

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano
APPLICATION NO.:	11/699,732
FILING DATE:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
ART UNIT:	2614
CONF. NO.:	7721
ATTY.DKT. NO.:	PA3909US

MAIL STOP AF
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

RESPONSE B

Examiner Kim:

In response to the Final Office Action (*Final Office Action*) mailed December 6, 2011, the Applicant submits the following response. **Listing of Claims**, **Remarks**, and **Conclusion** begin on pages **two**, **seven**, and **thirteen**, respectively.

OK TO ENTER UPON FILING OF APPEALS BRIEF: /P.K./

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 02/28/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	
ART UNIT	PAPER NUMBER
2614	
MAIL DATE	DELIVERY MODE
02/28/2012	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Applicant-Initiated Interview Summary	**Application No.** 11/699,732	**Applicant(s)** AVENDANO, CARLOS	
	Examiner PAUL KIM	**Art Unit** 2614	

All participants (applicant, applicant's representative, PTO personnel):

(1) *PAUL KIM*.

(2) *BRUCE SHEM*.

(3)*JAMES DRAPINSKI*.

(4)_____.

Date of Interview: *23 February 2012*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☒ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☐112 ☐102 ☐103 ☒Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *ALL*.

Identification of prior art discussed: *ELKO AND RYAN*.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

Applicant's attorney faxed a copy of proposed amendment. The proposed amendment was discussed in the interview and the examiner indicated that the proposed amendment overcomes the prior art rejections; however, that it would require further search/consideration upon submission of the amendment.

Applicant recordation instructions: The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

Examiner recordation instructions: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☐ Attachment

/P. K./ Examiner, Art Unit 2614	/VIVIAN CHIN/ Supervisory Patent Examiner, Art Unit 2614

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

Doc code: RCEX
Doc description: Request for Continued Examination (RCE)

PTO/SB/30EFS (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

REQUEST FOR CONTINUED EXAMINATION(RCE)TRANSMITTAL
(Submitted Only via EFS-Web)

Application Number	11/699,732	Filing Date	2007-01-29	Docket Number (if applicable)	PA3909US	Art Unit	2614
First Named Inventor	Carlos Avendano			Examiner Name	Paul Kim		

This is a Request for Continued Examination (RCE) under 37 CFR 1.114 of the above-identified application.
Request for Continued Examination (RCE) practice under 37 CFR 1.114 does not apply to any utility or plant application filed prior to June 8, 1995, or to any design application. The Instruction Sheet for this form is located at WWW.USPTO.GOV

SUBMISSION REQUIRED UNDER 37 CFR 1.114

Note: If the RCE is proper, any previously filed unentered amendments and amendments enclosed with the RCE will be entered in the order in which they were filed unless applicant instructs otherwise. If applicant does not wish to have any previously filed unentered amendment(s) entered, applicant must request non-entry of such amendment(s).

☐ Previously submitted. If a final Office action is outstanding, any amendments filed after the final Office action may be considered as a submission even if this box is not checked.

- ☐ Consider the arguments in the Appeal Brief or Reply Brief previously filed on _____
- ☐ Other _____

☒ Enclosed

- ☒ Amendment/Reply
- ☐ Information Disclosure Statement (IDS)
- ☐ Affidavit(s)/ Declaration(s)
- ☐ Other _____

MISCELLANEOUS

☐ Suspension of action on the above-identified application is requested under 37 CFR 1.103(c) for a period of months _____
(Period of suspension shall not exceed 3 months; Fee under 37 CFR 1.17(i) required)

☐ Other _____

FEES

☒ **The RCE fee under 37 CFR 1.17(e) is required by 37 CFR 1.114 when the RCE is filed.**
The Director is hereby authorized to charge any underpayment of fees, or credit any overpayments, to Deposit Account No 060600

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

☒ Patent Practitioner Signature

☐ Applicant Signature

PTO/SB/30EFS (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Signature of Registered U.S. Patent Practitioner			
Signature	/Bruce C. Shem/	Date (YYYY-MM-DD)	2012-03-01
Name	Bruce C. Shem	Registration Number	64914

This collection of information is required by 37 CFR 1.114. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Electronic Patent Application Fee Transmittal					
Application Number:	11699732				
Filing Date:	29-Jan-2007				
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement				
First Named Inventor/Applicant Name:	Carlos Avendano				
Filer:	Bruce Clovis Shem/Jillian Sweetnam				
Attorney Docket Number:	PA3909US				
Filed as Small Entity					
Utility under 35 USC 111(a) Filing Fees					
Description		**Fee Code**	**Quantity**	**Amount**	**Sub-Total in USD($)**
Basic Filing:					
Pages:					
Claims:					
Miscellaneous-Filing:					
Petition:					
Patent-Appeals-and-Interference:					
Post-Allowance-and-Post-Issuance:					
Extension-of-Time:					

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Request for continued examination	2801	1	465	465
		Total in USD ($)		**465**

Electronic Acknowledgement Receipt	
EFS ID:	12205401
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	System and method for utilizing omni-directional microphones for speech enhancement
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	Bruce Clovis Shem
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	01-MAR-2012
Filing Date:	29-JAN-2007
Time Stamp:	18:29:56
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$465
RAM confirmation Number	6580
Deposit Account	060600
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees) Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees) Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)					
File Listing:					
Document Number	**Document Description**	**File Name**	**File Size(Bytes)/ Message Digest**	**Multi Part /.zip**	**Pages (if appl.)**
1		3909US_Response.pdf	146451 84a9df7d9ccb0cc2bc3d94398e95f6f9237c61d0	yes	14
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Amendment After Final		1	1	
	Claims		2	8	
	Applicant Arguments/Remarks Made in an Amendment		9	14	
Warnings:					
Information:					
2	Request for Continued Examination (RCE)	3909US_Response_RCE_SB_30.pdf	30963 2668322acaf0ff20bb33eb9c998210727655ce89	no	2
Warnings:					
This is not a USPTO supplied RCE SB30 form.					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	30222 fae2dc81b5abbe11a50d07efdee3845679926bc9	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	207636		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano
APPLICATION NO.:	11/699,732
FILING DATE:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
ART UNIT:	2614
CONF. NO.:	7721
ATTY.DKT. NO.:	PA3909US

MAIL STOP AF
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

RESPONSE C AND REQUEST FOR CONTINUED EXAMINATION

Examiner Kim:

In response to the Final Office Action (*Final Office Action*) mailed December 6, 2011 and the Advisory Action mailed on February 14, 2012 (*Advisory Action*), the Applicant submits the following response. A **Listing of Claims**, **Remarks**, and **Conclusion** begin on pages **two**, **nine**, and **fourteen**, respectively.

The present response is submitted concurrently with a Request for Continued Examination.

IN THE CLAIMS

This listing of claims will replace all prior versions and listing of claims in the application:

1. (Currently Amended) A system for enhancing speech, comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal;

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal, the differential microphone array module being further configured to determine the cardioid primary signal based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and

an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

2. (Previously Presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal.

3. (Currently Amended) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary electric signal adjusted by the gain.

4. (Original) The system of claim 3 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

5. (Previously Presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the secondary electric signal and a delayed primary electric signal.

6. (Original) The system of claim 1 further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal.

7. (Original) The system of claim 1 further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

8. (Original) The system of claim 1 further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference.

9. (Original) The system of claim 1 further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

10. (Original) The system of claim 9 further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

11. (Original) The system of claim 1 further comprising a masking module configured to determine a speech estimate.

12. (Original) The system of claim 11 further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output.

13. (Original) The system of claim 1, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal.

14. (Currently Amended) A method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal;

determining the cardioid primary signal further based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and

non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

15. (Previously Presented) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a delayed secondary electric signal.

16. (Currently Amended) The method of claim 14 wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary electric signal adjusted by the gain.

17. (Original) The method of claim 16 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

18. (Original) The method of claim 14 wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal.

19. (Original) The method of claim 14 wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal.

20. (Original) The method of claim 14 further comprising determining an energy estimate for each of the acoustic signals during a frame.

21. (Original) The method of claim 14 further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference.

22. (Original) The method of claim 21 further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

23. (Original) The method of claim 22 further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal

24. (Original) The method of claim 22 further comprising smoothing the filter estimate.

25. (Previously Presented) The method of claim 14 wherein the cardioid primary signal and the cardioid secondary signal are each of a sub-band of the primary electric signal.

26. (Currently Amended) A non-transitory computer readable storage medium having embodied thereon a program, the program being executable by a processor to perform a method for enhancing speech, the method comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal;

determining the cardioid primary signal further based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and

non-linearly combining components of the cardioid primary signal and the cardioid ~~primary~~ secondary signal to obtain an inter-microphone level difference.

27. (previously presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between a level-equalized secondary electric signal and a delayed primary electric signal.

28. (previously presented) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a level-equalized secondary electric signal.

REMARKS

In the *Final Office Action*, the Examiner rejected claims 1-28. No claims are canceled or added by this response. Claims 1, 3, 14, 16, and 26 are amended. Claims 3, 16, and 26 are amended to correct typographical errors. No new matter is introduced. Support for the claim amendments may be found at least at ¶ [0029] of the present application (*Pending Application*) as originally filed. Claims 1-28 remain pending. Based on the following remarks, the Applicant respectfully requests reconsideration of the *Pending Application*.

Interview Summary

The undersigned representative and James W. Drapinski (Reg. No. 46,242) engaged in a telephonic interview with Examiner Paul Kim on February 23, 2012 (*Telephonic Interview*). Applicants and their representatives extend their sincere thanks to the Examiner for his time and courtesy during the aforementioned interview. During the course of the telephonic interview, the Examiner stated and confirmed that the proposed claim amendments overcome the cited art. These claim amendments are present in the amended claims of the present response. The Examiner also indicated that further search may be needed.

Rejection under § 101

In the Office Action mailed April 7, 2011 (*Previous Office Action*), Claim 26 was rejected under 35 U.S.C. § 101 for being directed to non-statutory subject matter. In the Response filed on August 4, 2011 (*Response A*), claim 26 was amended to recite a "non-transitory computer readable storage medium" in order to advance prosecution of the *Instant Application*. Since the Examiner did not explicitly address the § 101 rejection in the *Final Office Action* or the *Advisory Action*, the Applicant believes that the § 101 rejection is withdrawn. The Applicant respectfully requests confirmation.

Rejection under § 103(a)

Claims 1-28 stand rejected under 35 U.S.C. § 103(a) as obvious over U.S. Patent Application Publication Number 2003/0147538 (*Elko*) in view of U.S. Patent Application Publication Number 2003/0169891 (*Ryan*). The Applicant respectfully traverses.

The cited art does not teach or suggest "the differential microphone array module being further configured to determine the cardioid primary signal based at least in part on delaying at least one of the primary electric signal and the secondary electric signal."

Independent Claim 1 as amended recites in part: "the differential microphone array module being further configured to determine the cardioid primary signal based at least in part on delaying at least one of the primary electric signal and the secondary electric signal." During the *Telephonic Interview*, the above claim amendment was discussed relative to the cited references. The Examiner stated that the claim amendment overcomes the prior art rejections.

The cited art does not teach or suggest "an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference."

In addition, Claim 1 also in part recites: "an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference." In the *Final Office Action*, the Examiner determined that *Elko* does teach the claimed feature. The Examiner also contended that Ryan teaches the claimed feature. *Final Office Action*, page 3, lines 1-7.

The Applicant maintains that to support a conclusion that the claim would have been obvious requires that *all* the claimed elements were known in the prior art and that one skilled in the art could have combined those elements. See *KSR International Co. v. Teleflex Inc.*, 127 S.Ct. 1727, 1739 (2007) (emphasis added); see also MPEP § 2143. The Applicant further maintains that neither *Elko* nor *Ryan*, singly or in any theoretical combination, disclose the claimed non-linear combination.

The Examiner suggested in the *Final Office Action* that the Applicant is improperly imputing limitations from the *Instant Application* into the claims:

> Applicant is reminded that the specifications are not the measure of invention. Therefore, limitations contained therein cannot be read into the claims for the purpose of avoiding prior art. In re Sporck, 55 CCPA 743, 386 F.2d 924, 155 USPQ 687 (1968).
>
> *Final Office Action*, 15.

However, claim 1 clearly recites the claimed feature described: "configured to non-linearly combine components."

The Examiner contends that the claim limitations given their broadest reasonable interpretation are taught by *Elko* and *Ryan*:

> In response to applicant's arguments that the prior art references fail to disclose the claimed limitations "an inter-microphone level difference module configured to nonlinearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference" and "wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal," the limitations have been interpreted reasonably broad to read on the prior art references, as discussed above.
>
> *Final Office Action*, 15.

> The claimed limitations have been interpreted reasonably broad to read on the prior art references, as discussed in the final rejection.
>
> *Advisory Action*, 1.

However, the broadest reasonable interpretation must be consistent with the specification:

> The Patent and Trademark Office ("PTO") determines the scope of claims in patent applications not solely on the basis of the claim language, but upon giving claims their broadest reasonable construction "in light of the specification as it would be interpreted by one of ordinary skill in the art." *In re Am. Acad. of Sci. Tech. Ctr.*, 367 F.3d 1359, 1364[, 70 USPQ2d 1827] (Fed. Cir. 2004).

MPEP, § 2111 (quoting *Phillips v. AWH Corp.*, 415 F.3d 1303, 75 USPQ2d 1321 (Fed. Cir. 2005)).

The Applicant respectfully submits that a reasonably broad interpretation of the claims in light of the specification does not read on *Elko* and *Ryan*.

For example, claim 1 recites in part "configured to non-linearly combine components." The Examiner contended that ¶¶ [0038] and [0040] of *Ryan* disclose the claimed feature (*Final Office Action*, 3):

> The channel processing blocks 58A-58D can be configured to employ a twin detector average detection scheme while compressing the input signals. This twin detection scheme includes both slow and fast attack/release tracking modules that allow for fast response to transients (in the fast tracking module), while preventing annoying pumping of the input signal (in the slow tracking module) that only a fast time constant would produce. The outputs of the fast and slow tracking modules are compared, and the compression slope is then adjusted accordingly. The compression ratio, channel gain, lower and upper thresholds (return to linear point), and the fast and slow time constants (of the fast and slow tracking modules) can be independently programmed and saved in memory 44 for each of the plurality of channel processing blocks 58A-58D.

Ryan, ¶ [0038].

> After channel processing is complete, the four channel signals are **summed by summer** 60 to form a composite signal. This composite signal is then coupled to the post-filter 62, which may apply a post-processing filter function as discussed above. Following post-processing, the composite signal is then applied to a notch-filter 64, that attenuates a narrow band of frequencies that is adjustable in the frequency range

> where hearing aids tend to oscillate. This notch filter 64 is used to reduce feedback and prevent unwanted "whistling" of the device. Preferably, the notch filter 64 may include a dynamic transfer function that changes the depth of the notch based upon the magnitude of the input signal.
>
> *Ryan*, ¶ [0040] (emphasis added).

Clearly, at best *Ryan* teaches a linear combination (i.e., summing) and does not teach or suggest a non-linear combination. In contrast, claim 1 recites in part "configured to non-linearly combine components." Clearly a linear combination is not the same as a non-linear combination.

For all the above reasons, the rejection of claim 1 under § 103 is overcome and thus claim 1 is patentable. Additionally, as independent claims 14 and 26 include similar elements to those of independent claim 1, claims 14 and 26 are likewise patentable for at least the same reasons.

Furthermore, as a dependent claim incorporates by reference all the limitations of the claim from which it depends (see 35 U.S.C. § 112 ¶ 4), claims 2-13, 16-25, and 27-28 are patentable for at least the same reasons as the independent claim from which they depend in addition to their separate distinguishing characteristics. The Applicant has, for brevity, chosen not to individually discuss the additional distinguishing characteristics of the dependent claims, aside from claims 2 and 15. The Applicant's silence should not be interpreted as belief that the dependent claims are patentable only by virtue of their dependence on a patentable independent claim.

CONCLUSION

Based on the foregoing amendments and remarks, the Applicant believes the rejections have been overcome, and that the *Pending Application* is in condition for allowance. Should the Examiner have any questions regarding the case, the Examiner is invited to contact the Applicant's undersigned representative.

Although amendments have been made, no acquiescence or estoppel is or should be implied thereby. Rather the amendments are made only to expedite prosecution of the present application, and without prejudice to presentation or assertion in the future of claims on the subject matter affected thereby. The Applicant reserves the right to pursue the original claims and/or similar claims in subsequent amendments, continuations, or other descendants of the present application.

The present response is submitted concurrently with a Request for Continued Examination. The Commissioner is hereby authorized to charge any underpayment or credit any overpayment to Deposit Account No. 06-0600 for any matter in connection with this response, including any fee for extension of time up to the current date, which may be required.

Respectfully submitted,
Carlos Avendano

March 1, 2012

By: /Bruce C. Shem/

Bruce C. Shem, Reg. no. 64,914
CARR & FERRELL *LLP*
120 Constitution Drive
Menlo Park, CA 94025
TEL: (650) 812-3400
FAX: (650) 812-3444

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☐ To be Mailed
11/699,732	01/29/2007	

APPLICATION AS FILED – PART I

	(Column 1)	(Column 2)	SMALL ENTITY ☒		OR	OTHER THAN SMALL ENTITY	
FOR	NUMBER FILED	NUMBER EXTRA	RATE ($)	FEE ($)		RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT 03/01/2012

	(Column 1)		(Column 2)	(Column 3)	SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 28	Minus	** 28	= 0	X $30 =	0	OR	X $ =	
Independent (37 CFR 1.16(h))	* 3	Minus	***3	= 0	X $125 =	0	OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE	0	OR	TOTAL ADD'L FEE	

AMENDMENT

	(Column 1)		(Column 2)	(Column 3)					
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)		RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/MARCIA GORDON/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	3976863		1976-08-24	Engel	
	2	3978287		1976-08-31	Fletcher, et al.	
	3	4137510		1979-01-30	Iwahara	
	4	4433604		1984-02-28	Ott	
	5	4516259		1985-05-07	Yato, et al.	
	6	4536844		1985-08-20	Lyon	
	7	4581758		1986-04-08	Coker, et al.	
	8	4628529		1986-12-09	Borth, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	9	4630304		1986-12-16	Borth, et al.	
	10	4674125		1987-06-16	Carlson, et al.	
	11	4718104		1988-01-05	Anderson	
	12	4811404		1989-03-07	Vilmur, et al.	
	13	4812996		1989-03-14	Stubbs	
	14	4864620		1989-09-05	Bialick	
	15	4920508		1990-04-24	Yassaie, et al.	
	16	5027410		1991-06-25	Williamson, et al.	
	17	5054085		1991-10-01	Meisel, et al.	
	18	5058419		1991-10-22	Nordstrom, et al.	
	19	5099738		1992-03-31	Hotz	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	20	5119711		1992-06-09	Bell, et al.	
	21	5142961		1992-09-01	Paroutaud	
	22	5150413		1992-09-22	Nakatani, et al.	
	23	5175769		1992-12-29	Hejna, Jr., et al.	
	24	5187776		1993-02-16	Yanker	
	25	5210366		1993-05-11	Sykes, Jr.	
	26	5224170		1993-06-29	Waite, Jr.	
	27	5230022		1993-07-20	Sakata	
	28	5323459		1994-06-21	Hirano	
	29	5341432		1994-08-23	Suzuki, et al.	
	30	5381512		1995-01-10	Holton, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	31	5400409		1995-03-21	Linhard	
	32	5402493		1995-03-28	Goldstein	
	33	5471195		1995-11-28	Rickman	
	34	5473702		1995-12-05	Yoshida, et al.	
	35	5473759		1995-12-05	Slaney, et al.	
	36	5479564		1995-12-26	Vogten, et al.	
	37	5502663		1996-03-26	Lyon	
	38	5544250		1996-08-06	Urbanski	
	39	5583784		1996-12-10	Kapust, et al.	
	40	5587998		1996-12-24	Velardo, Jr., et al.	
	41	5590241		1996-12-31	Park, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	42	5675778		1997-10-07	Jones	
	43	5682463		1997-10-28	Allen, et al.	
	44	5706395		1998-01-06	Arslan, et al.	
	45	5717829		1998-02-10	Takagi	
	46	5729612		1998-03-17	Abel, et al.	
	47	5732189		1998-03-24	Johnston, et al.	
	48	5749064		1998-05-05	Pawate, et al.	
	49	5792971		1998-08-11	Timis, et al.	
	50	5806025		1998-09-08	Vis, et al.	
	51	5809463		1998-09-15	Gupta, et al.	
	52	5825320		1998-10-20	Miyamori, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	53	5839101		1998-11-17	Vahatalo, et al.	
	54	5920840		1999-07-06	Satyamurti, et al.	
	55	5933495		1999-08-03	Oh	
	56	5943429		1999-08-24	Handel	
	57	5956674		1999-09-21	Smyth, et al.	
	58	5974380		1999-10-26	Smyth, et al.	
	59	5978824		1999-11-02	Ikeda	
	60	5983139		1999-11-09	Zierhofer	
	61	5990405		1999-11-23	Auten, et al.	
	62	6097820		2000-08-01	Turner	
	63	6108626		2000-08-22	Cellario, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	64	6122610		2000-09-19	Isabelle	
	65	6134524		2000-10-17	Peters, et al.	
	66	6137349		2000-10-24	Menkhoff, et al.	
	67	6140809		2000-10-31	Doi	
	68	6173255		2001-01-09	Wilson, et al.	
	69	6180273		2001-01-30	Okamoto	
	70	6216103		2001-04-10	Wu, et al.	
	71	6223090		2001-04-24	Brungart	
	72	6226616		2001-05-01	You, et al.	
	73	6263307		2001-07-17	Arslan, et al.	
	74	6266633		2001-07-24	Higgins, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	75	6339758		2002-01-15	Kanazawa, et al.	
	76	6355869		2002-03-12	Mitton	
	77	6381570		2002-04-30	Li, et al.	
	78	6434417		2002-08-13	Lovett	
	79	6449586		2002-09-10	Hoshuyama	
	80	6469732		2002-10-22	Chang, et al.	
	81	6487257		2002-11-26	Gustafsson, et al.	
	82	6496795		2002-12-17	Malvar	
	83	6513004		2003-01-28	Rigazio, et al.	
	84	6516066		2003-02-04	Hayashi	
	85	6529606		2003-03-04	Jackson, Jr. II, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	86	6622030		2003-09-16	Romesburg, et al.	
	87	6718309		2004-04-06	Selly	
	88	6792118		2004-09-14	Watts	
	89	6798886		2004-09-28	Smith, et al.	
	90	6810273		2004-10-26	Mattila, et al.	
	91	6915264		2005-07-05	Baumgarte	
	92	6944510		2005-09-13	Ballesty, et al.	
	93	6982377		2006-01-03	Sakurai, et al.	
	94	6999582		2006-02-14	Popovic, et al.	
	95	7016507		2006-03-21	Brennan	
	96	7020605		2006-03-28	Gao	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	97	7054452		2006-05-30	Ukita	
	98	7065485		2006-06-20	Chong-White, et al.	
	99	7076315		2006-07-11	Watts	
	100	7092529		2006-08-15	Yu, et al.	
	101	7092882		2006-08-15	Arrowood, et al.	
	102	7164620		2007-01-16	Hoshuyama	
	103	7171246		2007-01-30	Mattila, et al.	
	104	7209567		2007-04-24	Kozel, et al.	
	105	7225001		2007-05-29	Eriksson, et al.	
	106	7242762		2007-07-10	He, et al.	
	107	7254242		2007-08-07	Ise, et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	108	7359520		2008-04-15	Brennan, et al.	
	109	7412379		2008-08-12	Taori, et al.	
	110	7433907		2008-10-07	Nagai, et al.	
	111	7555434		2009-06-30	Nomura, et al.	

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20020002455		2002-01-03	Accardi, Anthony J. ; et al.	
	2	20020133334		2002-09-19	Coorman, Geert ; et al.	
	3	20020147595		2002-10-10	Baumgarte, Frank	
	4	20020184013		2002-12-05	Walker, Michael	
	5	20030014248		2003-01-16	Vetter, Rolf	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	6	20030033140		2003-02-13	Taori, Rakesh ; et al.	
	7	20030040908		2003-02-27	Yang, Feng ; et al.	
	8	20030063759		2003-04-03	Brennan, Robert L. ; et al.	
	9	20030095667		2003-05-22	Watts, Lloyd	
	10	20030101048		2003-05-29	Liu, Chia-Horng	
	11	20030103632		2003-06-05	Goubran, Rafik ; et al.	
	12	20030128851		2003-07-10	Furuta, Satoru	
	13	20030228023		2003-12-11	Burnett, Gregory C. ; et al.	
	14	20040013276		2004-01-22	Ellis, Richard Thompson ; et al.	
	15	20040047464		2004-03-11	Yu, Zhuliang ; et al.	
	16	20040057574		2004-03-25	Faller, Christof	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	17	20040078199		2004-04-22	Kremer, Hanoh ; et al.	
	18	20040131178		2004-07-08	Shahaf, Mark ; et al.	
	19	20040165736		2004-08-26	Hetherington, Phil ; et al.	
	20	20040196989		2004-10-07	Friedman, Sol ; et al.	
	21	20040263636		2004-12-30	Cutler, Ross ; et al.	
	22	20050025263		2005-02-03	Wu, Gin-Der	
	23	20050027520		2005-02-03	Mattila, Ville-Veikko ; et al.	
	24	20050216259		2005-09-29	Watts, Lloyd	
	25	20050228518		2005-10-13	Watts, Lloyd	
	26	20050288923		2005-12-29	Kok, Chi-Wah	
	27	20060072768		2006-04-06	Schwartz; Stephen R. ; et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	28	20060098809		2006-05-11	Nongpiur; Rajeev ; et al.	
	29	20060133621		2006-06-22	Chen; Juin-Hwey ; et al.	
	30	20060149535		2006-07-06	Choi; Woo Young ; et al.	
	31	20060184363		2006-08-17	McCree; Alan ; et al.	
	32	20060198542		2006-09-07	Benjelloun Touimi; Abdellatif ; et al.	
	33	20060222184		2006-10-05	Buck; Markus ; et al.	
	34	20070067166		2007-03-22	Pan; Xingde ; et al.	
	35	20070094031		2007-04-26	Chen; Juin-Hwey	
	36	20070100612		2007-05-03	Ekstrand; Per ; et al.	
	37	20070195968		2007-08-23	Jaber; Marwan	
	38	20070230712		2007-10-04	Belt; Harm Jan Willem ; et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	39	20070276656		2007-11-29	Solbach; Ludger ; et al.	
	40	20080033723		2008-02-07	Jang; Giljin ; et al.	
	41	20080140391		2008-06-12	Yen; Ming Hsiang ; et al.	
	42	20080201138		2008-08-21	Visser; Erik ; et al.	
	43	20080228478		2008-09-18	Hetherington; Phillip A. ; et al.	
	44	20090012783		2009-01-08	Klein; David	
	45	20090012786		2009-01-08	Zhang; Xianxian ; et al.	
	46	20090129610		2009-05-21	Kim; Kyu-hong ; et al.	
	47	20090220107		2009-09-03	Every; Mark ; et al.	
	48	20090238373		2009-09-24	Klein; David	
	49	20090253418		2009-10-08	Makinen; Jorma	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	50	20090271187		2009-10-29	Yen; Kuan-Chieh ; et al.		
	51	20090323982		2009-12-31	Solbach; Ludger ; et al.		
	52	20100094643		2010-04-15	Avendano; Carlos ; et al.		
	53	20100278352		2010-11-04	Petit; Nicolas ; et al.		
	54	20110178800		2011-07-21	Watts; Lloyd		

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	5053587	JP		1993-03-05	Nakada		☐
	2	2003/069499	WO		2003-08-21	Watts		☐
	3	2010/005493	WO		2010-01-14	Solbach et al.		☐
	4	2007/140003	WO		2007-12-06	Solbach et al.		☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	5	2003/043374	WO		2003-05-22	Watts		☐
	6	2001/74118	WO		2001-10-04	Watts		☐
	7	2007/081916	WO		2007-07-19	Avendano et al.		☐
	8	11-249693	JP		1999-09-17	Takano et al.		☐
	9	10-313497	JP		1998-11-24	Aoki et al.		☐
	10	2005195955	JP		2005-07-21	Amada et al.		☐
	11	2005110127	JP		2005-04-21	Takahashi		☐
	12	2004/010415	WO		2004-01-29	Nomura et al.		☐
	13	2004053895	JP		2004-02-19	Osamu et al.		☐
	14	2005518118	JP		2005-06-16	Watts		☐
	15	2003069499	WO		2003-08-21	Watts		☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	16	2004531767	JP		2004-10-14	Brandman		☐
	17	2002103676	WO		2002-12-27	Brandman		☐
	18	2005148274	JP		2005-06-09	Nagai et al.		☐
	19	2002080362	WO		2002-10-10	Ekstrand		☐
	20	2004533155	JP		2004-10-28	Ekstrand		☐
	21	05-172865	JP		1993-07-13	Waite		☐

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T5
	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/James W. Drapinski/	Date (YYYY-MM-DD)	2012-03-14
Name/Print	James W. Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

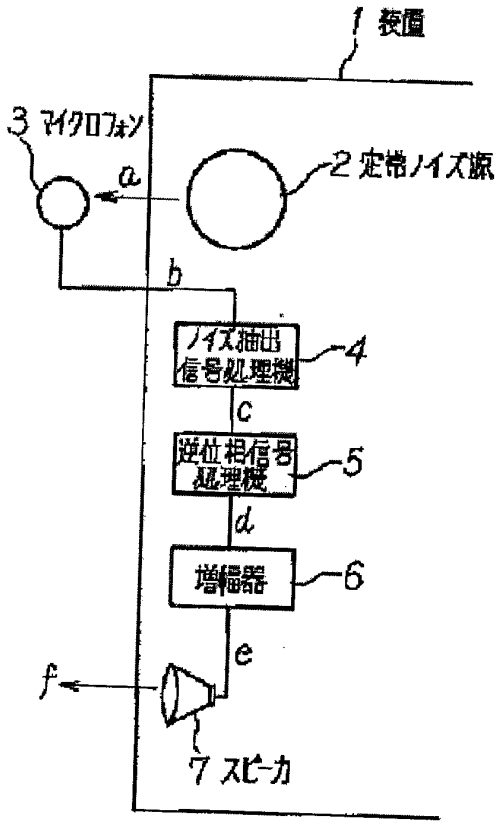
DEVICE NOISE CANCELLATION SYSTEM

Patent number: JP5053587
Publication date: 1993-03-05
Inventor: NAKADA AKIHIKO
Applicant: NIPPON ELECTRIC CO
Classification:
- international: ***G10K11/178; G10K11/16;*** **G10K11/00**; (IPC1-7): G10K11/16
- european:
Application number: JP19910213939 19910826
Priority number(s): JP19910213939 19910826

Report a data error here

Abstract of **JP5053587**

PURPOSE:To suppress the sound pressure in a sound field, where a stationary acoustic noise is generated by a device, following variation of a device noise by generating a signal which is opposite in phase from the stationary acoustic noise and feeding it back. CONSTITUTION:A stationary noise source 2 in the device 1 consists generally of a rotary mechanism such as a motor. A microphone 3 converts a stationary random noise (a) that the noise source 2 generates into an electric signal (b). A noise extraction signal processing machine 4 extracts the period, amplitude, and phase of the signal (b). An opposite-phase signal processing machine 5 outputs the signal (d) which is opposite in phase from the signal (b) with the extracted information signal (c). The opposite-phase signal (d) is amplified by an amplifier 6 up to the same amplitude with the noise (a). The amplified electric signal (e) is converted by a speaker 7 into a sound signal (f), which is outputted to the outside. Consequently, the sound pressure of the sound filed composite wave composed of the noise (a) and speaker output (f) in a measurement space approximates zero and the noise is suppressed.



Data supplied from the ***esp@cenet*** database - Worldwide

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（A）　(11)特許出願公開番号

特開平5－53587

(43)公開日　平成5年(1993)3月5日

(51)Int.Cl.[5]	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ１０Ｋ　11/16		Ｈ　7350－5Ｈ		

審査請求　未請求　請求項の数3(全　4　頁)

(21)出願番号　特願平3－213939

(22)出願日　平成3年(1991)8月26日

(71)出願人　000004237
日本電気株式会社
東京都港区芝五丁目7番1号

(72)発明者　中田　明彦
東京都港区芝五丁目7番1号　日本電気株式会社内

(74)代理人　弁理士　後藤　洋介　(外2名)

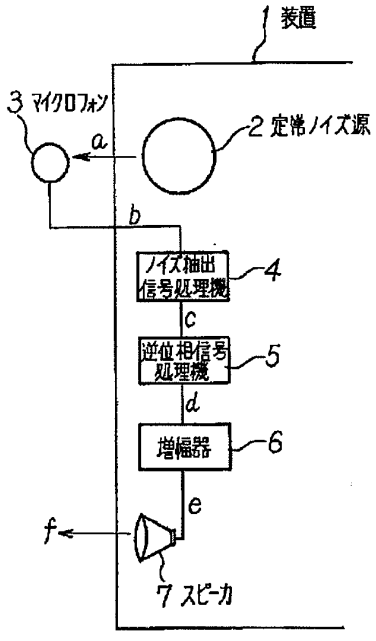
(54)【発明の名称】　装置ノイズキヤンセル方式

(57)【要約】

【目的】　装置内に存在する定常ランダムノイズの抑制。

【構成】　発明の構成は、装置1から発生した音響ノイズを、マイクロフォン3で電気信号に変換し、この信号から入力信号と逆位相の信号を生成し、スピーカにより音響信号ｆとして音場へ放射して装置からのノイズと合わせてその音場の音圧を抑圧すものである。

【効果】　装置内の静粛が得られる。



1

【特許請求の範囲】

【請求項1】　装置が発生する定常音響ノイズと逆位相の信号を発生しフィードバックして該定常ノイズの発生する音場の音圧を抑制することを特徴とする装置ノイズキャンセル方式。

【請求項2】　装置が発生する音響ノイズを抑制したい音響空間でその音響ノイズを捉えるマイクロフォンと、そのマイクロフォンの発生する信号からこの信号と逆位相の信号を発生させる逆位相信号発生手段と、該発生した逆位相信号を増幅し音場へ出力する手段とを備えている装置ノイズキャンセラー。

【請求項3】　装置が発生する音響ノイズを抑制したい音響空間でその音響ノイズを捉えるマイクロフォンと、そのマイクロフォンの発生する信号の位相、周期、および振幅を含む定常信号を抽出するノイズ抽出信号処理機と、抽出された情報を基に前記マイクロフォンの発生する信号と１８０°位相を変化させた信号を発生させる逆位相信号処理機と、発生した逆位相信号を原信号と同じ高さまで増幅する増幅器と、該増幅した音響信号を空間へ出力するスピーカとを備えている装置ノイズキャンセラー。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、音響ノイズを制御する方式、特に装置（例えばコンピュータシステム）から出る定常音響ノイズを制御する方式に関する。

【０００２】

【従来の技術】従来のコンピュータ装置は、装置から出る熱気を排気する為のファンやその他必要な機能を実現する為のモータを実装している。

【０００３】

【発明が解決しようとする課題】従来のコンピュータ装置では、排気ファン及びモータなどにより定常ノイズを発生しており、このノイズは吸音材及び装置筐体の構造強化により抑制している。しかしこの方法はパッシブ型抑制方式であるため、ノイズ特性により抑制材料が異なるので多種の材料を用意しなければならならず、また装置ノイズ発生源の経年変化によるノイズ特性変化に追従不可であり、更に製造装置毎に方式が異なる為設計コストが掛かると言う問題点があった。

【０００４】本発明は、コンピュータなどの場合、そのノイズ源は一般にモータ等の回転機構で構成される為、そのノイズ特性がランダム性ノイズではなく周期性を有する定常ランダム性ノイズが主に含まれていることに注目して、この定常ノイズ特性を利用しノイズを抑制する手法を実現しようとするものである。

【０００５】

【課題を解決するための手段】本発明によれば、装置が発生する定常音響ノイズと逆位相の信号を発生しフィードバックして該定常ノイズの発生する音場の音圧を抑制することを特徴とする装置ノイズキャンセル方式が得られる。

2

【０００６】また本発明によれば、装置が発生する音響ノイズを抑制したい音響空間でその音響ノイズを捉えるマイクロフォンと、そのマイクロフォンの発生する信号からこの信号と逆位相の信号を発生させる逆位相信号発生手段と、該発生した逆位相信号を増幅し音場へ出力する手段とを備えている装置ノイズキャンセラーが得られる。

【０００７】更に本発明によれば、装置が発生する音響ノイズを抑制したい音響空間でその音響ノイズを捉えるマイクロフォンと、そのマイクロフォンの発生する信号の位相周期、および振幅を含む定常信号を抽出するノイズ抽出信号処理機と、抽出された情報を基に前記マイクロフォンの発生する信号と１８０°位相を変化させた信号を発生させる逆位相信号処理機と、発生した逆位相信号を原信号と同じ高さまで増幅する増幅器と、該増幅した音響信号を空間へ出力するスピーカとを備えている装置ノイズキャンセラーが得られる。

【０００８】

【実施例】次に本発明について、図面を参照して説明する。

【０００９】図１は本発明の一実施例のブロック図である。装置１内にノイズ発生源２が存在する。このノイズ源は一般にモータ等の回転機構で構成される為、ノイズ特性としてランダム性ノイズではなく、周期性を有する定常ランダム性ノイズが主に含まれている。マイクロフォン３はこの定常ランダム性ノイズである音響信号を電気信号に変換する。ノイズ抽出信号処理機４はこの変換された電気信号のノイズの周期、振幅、および位相を実時間で抽出する。逆位相信号処理機５はこの抽出された情報によりマイクロホンの発生した信号と１８０°位相を変化させた逆位相信号に変換する。

【００１０】この逆位相信号は増幅器６によりマイクロフォン入力時の信号振幅と同等の振幅まで増幅される。そして増幅された電気信号はスピーカ７により音響信号に変換され外部へ放出される。

【００１１】これにより装置１で発生する定常ノイズ源２の発する音響信号と、各処理機により生成された逆位相信号の音響信号が打ち消され、測定点でのノイズを最大限に抑制する事になる。

【００１２】図２は、図１で説明した信号の流れを示したものである。

【００１３】ノイズ源２のノイズａはマイクロフォン３によりマイクロフォン出力信号ｂとして忠実に取込まれる。すなわち位相ひずみは発生しない。ノイズ抽出信号処理機４はマイクロフォン出力信号ｂから位相、周期、振幅の情報信号ｃを抽出する。この情報信号ｃを数式で表せば下記の数式１で表せる。

【００１４】

3

【数1】

$$S_{X(f)} = \int_{-\infty}^{\infty} X_{(t)}\ e^{-j2\pi ft}\,dt$$

【0015】上記のようにして得られた情報信号cを逆位相信号処理機5へ入れると逆位相信号dが出力される。この逆位相信号dは、振幅がノイズ源と異なる為、増幅器6により振幅し出力信号eとして得る。この信号eを電気信号から音響信号としてスピーカ7から出力fとして音場へ放射する。

【0016】これによりノイズaとスピーカ出力fを合成すると、得られる測定空間の音場合成波gは、位相が互いに逆と成っているので音圧は0に近づき、ノイズの抑制が実現される。

【0017】

【発明の効果】以上説明したように、本発明の装置ノイズキャンセル方式は、ノイズ発生源から出る音響ノイズを電気的にリアルタイムでフィードバックするアクティブ型ノイズキャンセル方式である為、ノイズ特性に左右されず、装置ノイズの変化に追従出来、且つ、ノイズが

4

定常ノイズであるかぎり装置固有の手法は必要なく一定方法で良い為、開発設計コストは低下する。

【図面の簡単な説明】

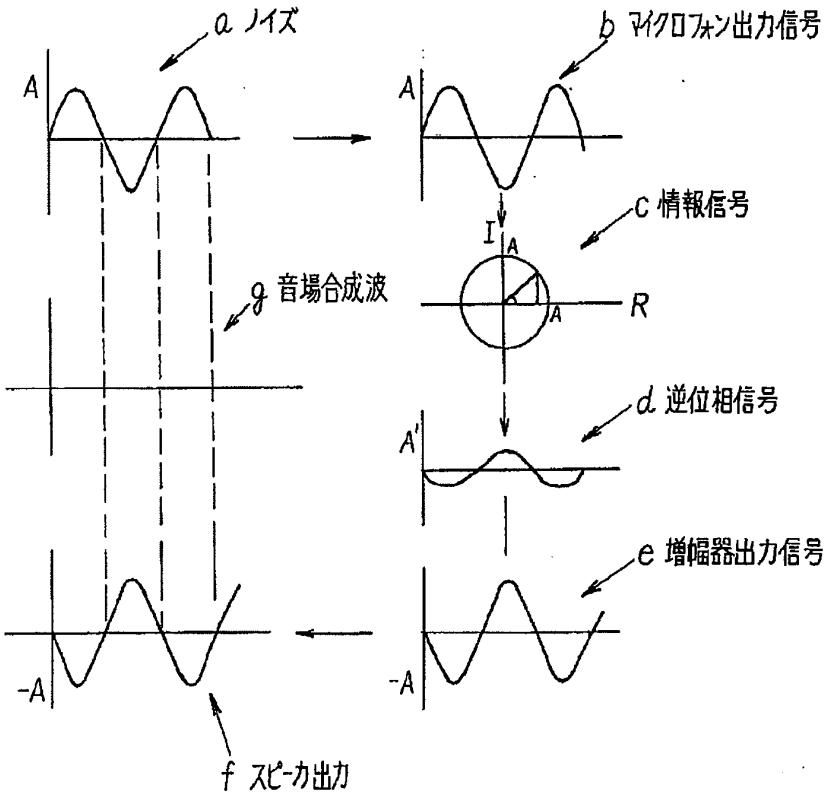
【図1】本発明の一実施例のブロック図である。

【図2】図1の信号の流れと特性を示した図である。

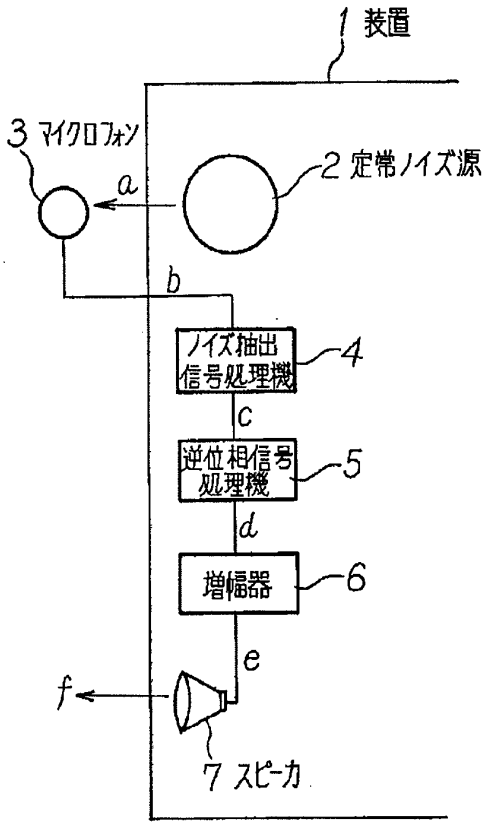
【符号の説明】

1	装置
2	定常ノイズ源
3	マイクロフォン
4	ノイズ抽出信号処理機
5	逆位相信号処理機
6	増幅器
7	スピーカ
a	ノイズ
b	マイクロフォン出力信号
c	情報信号
d	逆位相信号
e	増幅器出力信号
f	スピーカ出力
g	音場合成波

【図2】



【図1】


1 装置
3 マイクロフォン
a
2 定常ノイズ源
b
ノイズ抽出信号処理機
4
c
逆位相信号処理機
5
d
増幅器
6
e
f
7 スピーカ

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 August 2003 (21.08.2003)

PCT

(10) International Publication Number
WO 2003/069499 A1

(51) International Patent Classification[7]: G06F 17/00, H03G 5/00, G10L 19/00

(21) International Application Number: PCT/US2003/004124

(22) International Filing Date: 11 February 2003 (11.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/074,991 13 February 2002 (13.02.2002) US

(71) Applicant: AUDIENCE, INC. [US/US]; 4906 El Camino Real, Suite 205, Los Altos, CA 94022 (US).

(72) Inventor: WATTS, Lloyd; 500 Chiquita Avenue #2, Mountain View, CA 94041 (US).

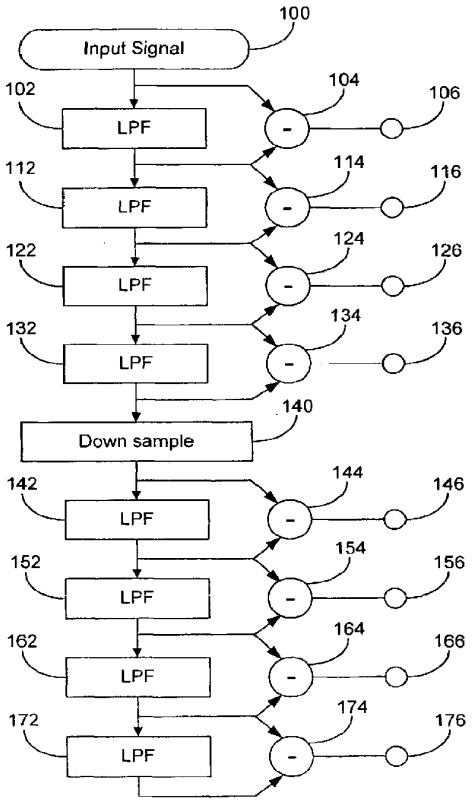
(74) Agent: VAN PELT, Lee; Van Pelt & Yi LLP, 10050 N. Foothill Blvd., Suite 200, Cupertino, CA 95014 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI,

[Continued on next page]

(54) Title: FILTER SET FOR FREQUENCY ANALYSIS



(57) Abstract: A system and method are disclosed for analyzing an input signal (100) into a plurality of frequency components. In one embodiment, the input signal (100) is processed with a first set of low pass filters (102-132) to derive a first set of frequency components wherein the first set of low pass filters (102-132) are arranged serially in a chain having a first low pass filter (102) and a last low pass filter (132), the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter (132). The output of the last low pass filter (132) is downsampled (140) to produce a downsampled signal. The downsampled signal is processed with a second set of low pass filters (142-172) to derive a second set of frequency components.

WO 2003/069499 A1

WO 2003/069499 A1

SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(48) Date of publication of this corrected version:
3 June 2004

(15) Information about Correction:
see PCT Gazette No. 23/2004 of 3 June 2004, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

FILTER SET FOR FREQUENCY ANALYSIS

Cross reference to related applications

This application claims priority to co-pending U.S. Patent Application No. 09/534,682 (Attorney Docket No. ANSCP001) entitled EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE filed March 24, 2000, which is incorporated herein by reference for all purposes.

Field of the Invention

The present invention relates generally to signal processing. A system and method for analyzing a signal into frequency components is disclosed.

Background Of The Invention

A useful step in analyzing a signal is the separation of the signal into frequency components. For some time, the fast Fourier transform or FFT algorithm has been used to analyze a time domain signal into its frequency components. For various types of processing, and in particular for processing audio signals, it would be desirable to analyze a signal into its frequency components with improved temporal resolution at high frequencies and better spectral resolution at low frequencies. Numerous techniques have been proposed for accomplishing this. Included among such techniques are systems that use a set of filters to separate the signal being analyzed into different channels or frequency components. Such filter sets operate roughly in a manner that is analogous to a biological cochlea, which includes a series of filtered output signals that correspond to different frequency channels.

Filter sets may be implemented with analog or digital filters. Previous instantiations of filter sets have been limited by practical considerations in designing filters. For example, high order bandpass filters to separate each channel output are expensive to implement. Various approaches have been implemented using combinations of high pass and low pass filters; however, more efficient techniques are needed to allow real time processing of signals for various important applications including speech recognition, source separation of audio signals and stream separation of audio signals.

Brief Description Of The Drawings

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

Figure 1 is a block diagram illustrating a filter network used in one embodiment for analyzing an input signal into a plurality of frequency components.

Figure 2 is a diagram illustrating an alternative embodiment wherein the low pass filters are not chained together at their inputs and outputs.

Figure 3 is a signal flow graph of a filter equation.

Figure 4 is a block diagram illustrating the arrangement of the filters.

Figure 5 is a diagram illustrating an example of the filter response of a second-order section with poles only.

Figure 6 is a diagram illustrating a typical filter response where Q_p is the Q of the pole, Q_z is the Q of the zero, f_{cp} is the center frequency of the pole (also referred to as f_p), and f_{cz} is the center frequency of the zero (also referred to as f_z).

Figure 7 is a diagram illustrating filter responses for filters designed according to the critical band.

Figure 8 is a diagram illustrating the phase characteristics for filters designed according to the critical band.

Figure 9A is a diagram illustrating how a filter set as described herein is used in a voice recognition system.

Figure 9B is a diagram illustrating how a filter set as described herein is used in an audio stream separation system.

Figure 9C is a diagram illustrating how a filter set as described herein is used in a spatial correlator or sound localization system.

Detailed Description

A detailed description of a preferred embodiment of the invention is provided below. While the invention is described in conjunction with that preferred embodiment, it should be understood that the invention is not limited to any one embodiment. On the contrary, the scope of the invention is limited only by the appended claims and the invention encompasses numerous alternatives, modifications and equivalents. For the purpose of example, numerous specific details are set forth in the following description in order to provide a thorough understanding of the present invention. The present invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the present invention is not unnecessarily obscured.

A filter cascade for frequency analysis is disclosed that includes a number of features. In various embodiments, the features are implemented either separately or together. For example, in some embodiments, each frequency component is computed by subtracting the output of a low pass filter from the input to the filter. In this manner a bandpass signal is derived. In some embodiments, low pass filters are chained or cascaded with each filter output being fed to the next filter input in a filter set. The output of the last filter in the set is downsampled, with the filter set itself collectively acting as a high order antialiasing filter. The downsampled filter set output comprised of lower frequency components may then be more efficiently processed. Filters in the cascade may be designed so that the Q of the filters varies with frequency.

U.S. Patent Application No. 09/534,682 which was previously incorporated by reference (hereinafter, "the 682 application") discloses a digital filter cascade for frequency analysis. The filters in the cascade are chained together and sets of filters are separated into octaves with downsampling between octaves. Filter parameters are shared among corresponding filters in different octaves. As described herein, advantages may be realized if filter parameters are varied among octaves in a manner that varies the Q, or sharpness of the filters among octaves. In one embodiment, the Q is varied substantially according to critical bandwidth.

Figure 1 is a block diagram illustrating a filter network used in one embodiment for analyzing an input signal into a plurality of frequency components. An input signal 100 is fed to a low pass filter (LPF) 102. The output of LPF 102 is subtracted from input signal 100 by a subtractor 104. The output at node 106 thus represents the difference between the signal before and after LPF 102. It emphasizes a band or channel of frequencies above the cutoff frequency of LPF 102 and whatever

the upper frequency cutoff of the input signal happens to be. The output of LPF 102 is similarly directed to the input of LPF 112 and the difference between the input and the output of LFP 112 is computed by a subtractor 114 and output at node 116. The output at node 116 represents another frequency channel that emphasizes frequencies between the cutoff frequencies of LPF 102 and LPF 112. In a similar manner, LPF 122 and LPF 132 and subtractor 124 and 134 output other frequency channels at nodes 126 and 136. The output of the nodes may be further processed as is appropriate. For example, in some embodiments, the outputs are half wave rectified and in some embodiments, the gain of the outputs is adjusted to compress or expand the dynamic range.

In different embodiments, second order or higher digital or analog filters may be used. The nature of the filters, of course determines the exact nature of each channel output that generally emphasizes a given frequency band and thus has a general bandpass character. Collectively, the channel outputs represent the frequency components of the signal. Because of the subtraction of each LPF input and output, each channel output represents a band or slice of frequencies and the sum of all the outputs represents the entire input signal.

Because the output of each LPF is fed to the input of the next LPF, forming a chain of low pass filters, the output of the last LPF in the chain has characteristics of a much higher order filter than the order of the last filter. This higher order filtering effect may be exploited when the output of the last filter in the chain is downsampled. Essentially, the chain of low pass filters used to separate out frequency channels collectively act as a high order filter that performs the function of an anti aliasing filter when the signal is downsampled.

An example of this is depicted in Figure 1 where downsampler 140 downsamples the output from LPF 132. It should be noted that only four filters are shown in the chain for the purpose of illustration. In most embodiments, more than four filters would be used to process a frequency range before downsampling. The downsampled signal output from downsampler 140 is then processed by another chain of low pass filters that includes LPF 142, LPF 142, LPF 142, LPF 142 and frequency channel outputs are derived by subtractors 144, 154, 164 and 174 at nodes 146, 156, 166 and 176.

In one embodiment, second order individual filters are used and a chain of 60 filters process one octave of the signal before downsampling. Downsampling may be implemented by simply discarding every other sample or any other appropriate technique. The amount of downsampling is determined by the Nyquist criterion. A suitable amount of oversampling may be done as desired. The combined effect of the chain of filters is that of a very high order anti aliasing filter. Thus, downsampling the signal may be done to speed the processing of lower frequency octaves without requiring an expensive high order anti aliasing filter.

It should be noted that the benefit of chaining the low pass filters is realized in certain embodiments without implementing the subtractors to calculate the frequency bands. The output of each low pass filter may be used directly to represent the energy in each frequency channel. The output of the last filter in each chain is downsampled with the filter chain itself performing the function of an antialiasing filter.

Figure 2 is a diagram illustrating an alternative embodiment wherein the low pass filters are not chained together at their inputs and outputs. Input signal 200 is fed into low pass filters 202, 204, 206, and 208. The difference between the input and the output of each low pass filter is calculated by subtractors 212, 214, 216 and 218.

Again, the differences calculated represent an analysis of the frequency bands or channels of the input signal. However, because the output of each filter is not fed to the input of the next filter, the higher order filter effect in the output of the last filter in the chain described above is not realized.

The filter cascade may be implemented using either analog or digital filters. In one embodiment, the filters are implemented as digital filters with cutoff frequencies designed to produce the desired channel resolution. Each filter has a set of coefficients (a_0, a_1, a_2, b_1, b_2) associated with it. The output of each filter is calculated according to the following function:

Equation 1. $$y_n = a_0x_n + a_1x_{n-1} + a_2x_{n-2} - b_1y_{n-1} - b_2y_{n-2}$$

where the filter output y_n is a function of the input data x_n at time n, previous inputs x_{n-1} and x_{n-2}, and previous outputs y_{n-1} and y_{n-2}. Figure 3 is a block diagram illustrating this signal flow. The output of the filter y_n is passed to the input x_n of the next filter in the cascade.

The filter response **H(z)** is given by the following:

Equation 2. $$H(z) = \frac{a_0 + a_1z^{-1} + a_2z^{-2}}{1 + b_1z^{-1} + b_2z^{-2}}$$

and $z = e^{i*(\omega/\omega s)}$, $\omega = 2\pi f$, $\omega_s = 2\pi f_s$

where f_s is the sampling frequency.

Substitution of the above into the transfer function of Equation 2 produces a filter response **H(f)**, which is a function of the filter coefficients a_0, a_1, a_2, b_1, b_2 and the sampling rate f_s.

As described in the 682 application, the filter coefficients may be reused between sets of filters with the response of the filters being altered as a result of downsampling between the sets of filters. In the embodiment shown, the filters are evenly distributed over the octaves, resulting in 60 filters per octave. 60 objects are created in a computer. Each object has a set of coefficients as described above, and additionally has ten sets of state variables, corresponding to ten filters running at frequencies that are whole octaves apart. The 60 objects using their first sets of state variables correspond to the first octave group of filters, while the 60 objects using their second sets of state variables (and sampling at a lower frequency) correspond to the second octave group of filters, and so on. In another embodiment, each object contains a set of coefficients, but only one set of state variables, and is run at a single frequency. In this case, 600 objects are required to represent 600 filters.

The filters in the first octave are tuned to the frequencies in the highest octave, 20 kHz to 10 kHz, and are sampled at 44.1 kHz, which satisfies the Nyquist sampling criterion. The filters in the second octave are tuned to half of the frequencies of the corresponding filters in the first octave, and range from 10 kHz to 5 kHz. These filters in the second octave are sampled at 22.05 kHz, half of the first sampling frequency. Coefficients for each filter are stored in memory and applied in the computations for the filters. The cascade response is the sum of responses of individual filters (which are all weak responses by themselves, but when summed, produce a much stronger response). The coefficients of the filters are determined by the desired response.

As the audio signal is passed through each filter, the signal is sampled and filtered before being passed to the next filter. Figure 4 is a block diagram illustrating the arrangement of the filters. At the end of the first octave, the signal is passed into

the first filter in the next octave, which comprises filters sampling at half the sampling rate of the first octave, as stated above. Successive octaves are downsampled in a similar manner, using the same factor of two. In this configuration, each stage acts as an anti-aliasing filter for later stages, removing the high frequencies sufficiently to allow downsampling without aliasing. No extra anti-aliasing filters are required.

Downsampling each successive octave significantly decreases the computational complexity of the system. In addition, the required precision for filter coefficients is lower, and thus, fewer bits are required to represent each coefficient. Digital low-pass filters have the property that the numerical precision required to represent the filter coefficients depends on the ratio between the cutoff frequency and the sampling

frequency. For a given sampling frequency, a filter with a low cutoff frequency will require higher-precision coefficients than a filter with a higher cutoff frequency. Without the successive downsampling technique, very high-precision filter coefficients (on the order of 23 bits) are required to represent the lowest-cutoff-frequency filters (30 Hz) at the 44 kHz sampling rate. With the successive downsampling technique, lower-precision coefficients (on the order of 12 bits) can be used to represent the 30-Hz cutoff filters, since the sampling rate is much lower in the lowest octave after many downsampling steps. This reduced precision results in lower hardware complexity (less memory, smaller registers, lower-precision arithmetic operators) and thus lower overall cost in a custom hardware implementation.

In the embodiment described in the 682 application, each filter shares filter parameters with filters that are one, two, or more octaves higher or lower, resulting in reduced storage requirements. For example, the highest frequency filter 40 in the first

octave shares filter coefficients with the highest frequency filter 50 in the second octave, the highest frequency filter 60 in the third octave, and so on. The second-highest frequency filter 42 in the first octave shares filter coefficients with the second-highest frequency filters 52 and 62 in the second and third octaves, and with all other corresponding filters (tuned to frequencies that are one, two, or more octaves lower).

Alternatively, it has been determined that the delay at low frequencies can be improved by changing the filter parameters within each octave as described below. For many systems, this is preferable to sharing filter parameters between corresponding filters in different octaves because the benefit from improved delay at low frequencies offsets increased memory storage requirements.

In one embodiment, filter coefficients are tuned to produce a desired Q (quality factor, or degree of sharpness or frequency selectivity) depending on the frequency band (determined by the frequency cutoff) being processed by the filter. Reusing filter coefficients in the cascade results in a cascade with constant Q, and all the filter responses will have the same shape (Q). This "constant-Q" configuration has the advantages of conceptual simplicity and shared filter coefficients, but has significant delays at low frequencies. For example, for a constant-Q design with a phase accumulation of four cycles at all frequencies, the delay at the 20 kHz tap will be 200 μs, while the delay at the 20 Hz tap will be 200 ms. Faster performance at low frequencies is desirable to improve the response time of the cascade, which may be accomplished by changing the filter coefficients of the filters in lower octaves.

Figure 5 is a diagram illustrating an example of the filter response of a second-order section with poles only. The filter may be described in terms of the time constant Tau and quality factor Q, or in terms of filter coefficients $\mathbf{b_1}$ and $\mathbf{b_2}$

mentioned previously. Tau is the inverse of the center frequency $\mathbf{f_c}$ and describes where the peak is, while Q describes how sharp the peak is. As can be seen from Figure 5, a higher Q results in a sharper peak, while $\mathbf{f_c}$ indicates where the peak occurs. The equations for the filter are as follows:

$$Vout(z) = \frac{1}{1 - b_1 z^{-1} - b_2 z^{-2}} Vin(z)$$

and

$$Vout(s) = \frac{1}{1 + Tau\, s/Q + Tau^2 s^2} Vin(s)$$

where the relationship between Tau, Q and $\mathbf{b_1}$, $\mathbf{b_2}$ are given in the "Lyon's Cochlear Model" Apple Technical Report #13 by Malcolm Slaney ©1988 which is herein incorporated by reference. The filter coefficients for the filter can be determined from the center frequency f_c=1/Tau, and the Q of the filter.

The filters may be designed to have zeros as well as poles, and the equation for such a system is given by

$$Vout(s) = \frac{1 + Tau_z\, s/Q_z + Tau_z^2 s^2}{1 + Tau_p\, s/Q_p + Tau_p^2 s^2} * Vin(s)$$

Figure 6 is a diagram illustrating a typical filter response where Q_p is the Q of the pole, Q_z is the Q of the zero, f_{cp} is the center frequency of the pole (also referred to as f_p), and f_{cz} is the center frequency of the zero (also referred to as f_z). The zeros arrest the dropping gain, and reverse the phase back up to zero. The closer the zero is to the pole, the sooner these effects occur. If the zero is very close to the pole, the phase trajectory may not get very far (a small fraction of a cycle) before the zero

reverses it. This property is the key to controlling the total amount of phase accumulation through the cascade, and hence the delay response of the cascade.

If 600 filters are used, and implemented with a cascade of 600 poles-only sections, each one would contribute a quarter-cycle of phase accumulation at its best frequency, resulting in a large amount of delay. In one embodiment, the filter cascade is configured so that the center frequencies decrease exponentially through the cascade. The Q's decrease gradually through the cascade, to give sharp responses at high frequencies, where delay is not an issue, and to give fast responses at low frequencies, where some loss of sharpness is acceptable in return for faster response. This implementation of nonconstant Q filters is particularly useful for signal processing systems used, for example in submarine passive sonar, speech recognition, music transcription, audio stream separation and sound localization. It should be noted that this approach is not limited to downsampled filter cascades, and may be used with filter cascades with no downsampling.

Design of a filter cascade with constant-Q involves choosing the range of cutoff frequencies and the number of taps per octave, such as a frequency range of 20 Hz to 20 kHz, 600 taps, 10 octaves (60 taps/octave). This determines f_p for each tap. Fixed values are chosen for Q_p, Q_z, and $f_{ratio}=f_z/f_p$, based on the sharpness and delay desired through the cascade. In one embodiment, values used for a constant-Q design may be $Q_p = 7.0$, $Q_z = 7.5$, and $f_{ratio} = 1.03$. In another embodiment, the values may be $Q_p = 23$, $Q_z = 26$, and $f_{ratio} = 1.01$.

For a variable-Q filter cascade using 600 taps in 10 octaves, one embodiment may employ the following values: $Q_p = 7.0$, $Q_z = 7.0$, and $f_{ratio} = 1.03$, with a sampling

rate of 44.1 kHz and 2x oversampling in the highest octave. These values are used for the first 360 taps, and then varied linearly over the next 240 taps to $Q_p = 1.6$, $Q_z = 1.6$, and $f_{ratio} = 1.1$ at tap 600 (the lowest frequency tap). This results in a design with broader filter responses at low frequencies, but much faster time response.

In another embodiment, the Q_p, Q_z, and f_{ratio} parameters are selected to match the filter responses to appropriate psychophysical critical bandwidth and loudness perception curves. Critical bandwidth is the tuning width of the filter response curves, within which signal components can interact with each other. Critical bandwidth curves are given in Rossing, 1982, "The Science of Sound" (Addison-Wesley, Reading, MA), the disclosure of which is hereby incorporated by reference. The critical bandwidth varies from a little less than 100 Hz at low frequencies to between two and three musical semitones (12% to 19%) at high frequencies. Loudness perception describes how sensitive the filters are to different frequencies. For example, the threshold of audibility at 20 Hz is about 65 dB higher than at 1 kHz.

One embodiment of a variable-Q filter cascade uses the following parameters:

Tap 0: $Q_p = 7.0$, $Q_z = 7.0$, $f_{ratio} = 1.03$
Tap 300: $Q_p = 11.0$, $Q_z = 11.0$, $f_{ratio} = 1.03$
Tap 360: $Q_p = 9.0$, $Q_z = 11.0$, $f_{ratio} = 1.03$
Tap 600: $Q_p = 1.6$, $Q_z = 1.6$, $f_{ratio} = 1.01$

with linear interpolation of parameters between the specified taps. This piecewise linear variation of the parameters gives a good fit to the psychophysical critical bandwidth and loudness perception curves. Figure 7 is a diagram illustrating filter responses for filters designed according to the critical band. The filter responses are sharp at mid-range frequencies, and very broad at low frequencies, corresponding to the critical bandwidth curve. The filters are more sensitive at mid-range

frequencies, and about 65 dB less sensitive at low frequencies, so as to match the loudness perception parameters.

Figure 8 is a diagram illustrating the phase characteristics for filters designed according to the critical band. The phase characteristics of the filters are such that there are about two cycles of phase accumulation at mid-to-high frequencies, but much less at low frequencies. This results in a faster response at low frequencies, where it is needed.

A filter cascade for analyzing a signal into frequency components has been described. In various embodiments, the filter cascade utilizes different techniques to improve temporal resolution at high frequencies and spectral resolution at low frequencies. As a result, each of the disclosed filter cascade embodiments are particularly useful as a component of a voice recognition system. In addition, the filter cascade is useful for audio stream separation and sound localization.

Figure 9A is a diagram illustrating how a filter set as described herein is used in a voice recognition system. An audio signal is input to a filter set 902 and the output of the filter set is analyzed by a feature extractor 904. The features are classified by a phoneme classifier 906 that matches features with phonemes included in a phoneme database 908. Words are derived based on the phonemes by a word search block 909 that access a word database 910.

Figure 9B is a diagram illustrating how a filter set as described herein is used in an audio stream separation system such as is described in United States Patent Application No. 60/300,012 (Attorney Docket No.ANSCP003+) by Lloyd Watts (filed June 21, 2001,) entitled: ROBUST HEARING SYSTEMS FOR INTELLIGENT MACHINES which is herein incorporated by reference. An audio signal is input to a filter set 912 and the output of the filter set is analyzed by a set of

feature extractors 914 that extract features. The features are grouped by feature grouping processor 916 into separate streams of associated audio signals.

Figure 9C is a diagram illustrating how a filter set as described herein is used in a spatial correlator or sound localization system such as is described in United States Patent Application No. 10/004,141 (Attorney Docket No. ANSCP005) by Lloyd Watts (filed November 14, 2001) entitled: COMPUTATION OF MULTI-SENSOR TIME DELAYS which is herein incorporated by reference. A right channel audio signal is input to a right channel filter set 922 and a left channel audio signal is input to a left channel filter set 924. The outputs of the filter sets are correlated by a binaural processor 926 to determine the time delay between the left and right channel input signals. The direction from which a sound emanates may be determined from the time delay.

Although the foregoing invention has been described in some detail for purposes of clarity of understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims. It should be noted that there are many alternative ways of implementing both the process and apparatus of the present invention. Accordingly, the present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein, but may be modified within the scope and equivalents of the appended claims.

WHAT IS CLAIMED IS:

CLAIMS

1. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first set of low pass filters to derive a first set of frequency components wherein the first set of low pass filters are arranged serially in a chain having a first low pass filter and a last low pass filter, the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter;

downsampling the output of the last low pass filter to produce a downsampled signal;

processing the downsampled signal with a second set of low pass filters to derive a second set of frequency components.

2. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the frequency components are derived by subtracting the output of each low pass filter from the input to the low pass filter.

3. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the second set of low pass filters have a different Q than the first set of low pass filters.

4. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the second set of low pass filters have a Q that is less sharp than the first set of low pass filters.

5. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the second set of low pass filters have a Q that differs from the Q of the first set of low pass filters substantially according to human critical bandwidth.

6. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first low pass filter to produce a first low pass filtered signal;

subtracting the first low pass filtered signal from the input signal to derive a first frequency component;

processing the signal with a second low pass filter to produce a second low pass filtered signal; and

subtracting the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

7. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first low pass filter to produce a first low pass filtered signal;

subtracting the first low pass filtered signal from the input signal to derive a first frequency component;

processing the low pass filtered signal with a second low pass filter to produce a second low pass filtered signal; and

subtracting the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

8. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first filter wherein the first filter is configured to separate part of the signal into a first output frequency channel; and

processing the signal with a second filter wherein the second filter is configured to separate part of the signal into a second output frequency channel wherein the first frequency channel emphasizes higher frequencies than the second frequency channel; and wherein the second filter has a different Q than the first filter.

9. A method of analyzing an input signal into a plurality of frequency components as recited in claim 8 wherein the second filter has a Q that is less sharp than the first filter.

10. A method of analyzing an input signal into a plurality of frequency components as recited in claim 8 wherein the second filter has a Q that differs from the Q of the first filter substantially according to human critical bandwidth.

11. A method of analyzing an input signal into a plurality of frequency components as recited in claim 8 wherein the filters are low pass filters.

12. A system for analyzing an input signal into a plurality of frequency components comprising:

a first set of low pass filters configured to derive a first set of frequency components wherein the first set of low pass filters are arranged serially in a chain

having a first low pass filter and a last low pass filter, the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter;

a downsampler configured to downsample the output of the last low pass filter to produce a downsampled signal;

a second set of low pass filters configured to process the downsampled signal to derive a second set of frequency components.

13. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the frequency components are derived by subtracting the output of each low pass filter from the input to the low pass filter.

14. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the second set of low pass filters have a different Q than the first set of low pass filters.

15. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the second set of low pass filters have a Q that is less sharp than the first set of low pass filters.

16. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the second set of low pass filters have a Q that differs from the Q of the first set of low pass filters substantially according to critical band.

17. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the system is used in a voice recognition system.

18. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the system is used for audio stream separation

19. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the system is used for sound localization.

20. A system for analyzing an input signal into a plurality of frequency components comprising:

a first low pass filter that outputs a first low pass filtered signal;

a first processor configured to subtract the first low pass filtered signal from the input signal to derive a first frequency component;

a second low pass filter that outputs a second low pass filtered signal; and

a second processor configured to subtract the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

21. A system for analyzing an input signal into a plurality of frequency components comprising:

a first low pass filter that outputs a first low pass filtered signal;

a first processor configured to subtract the first low pass filtered signal from the input signal to derive a first frequency component;

a second low pass filter configured to process the low pass filtered signal to produce a second low pass filtered signal; and

a second processor configured to subtract the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

22. A system for analyzing an input signal into a plurality of frequency components comprising:

a first filter configured to process the signal wherein the first filter is configured to separate part of the signal into a first output frequency channel; and

a second filter configured to process the signal wherein the second filter is configured to separate part of the signal into a second output frequency channel wherein the first frequency channel emphasizes higher frequencies than the second frequency channel; and wherein the second filter has a different Q than the first filter.

23. A system for analyzing an input signal into a plurality of frequency components as recited in claim 22 wherein the second filter has a Q that is less sharp than the first filter.

24. A system for analyzing an input signal into a plurality of frequency components as recited in claim 22 wherein the second filter has a Q that differs from the Q of the first filter substantially according to a critical band.

25. A system for analyzing an input signal into a plurality of frequency components as recited in claim 22 wherein the filters are low pass filters.

1/9

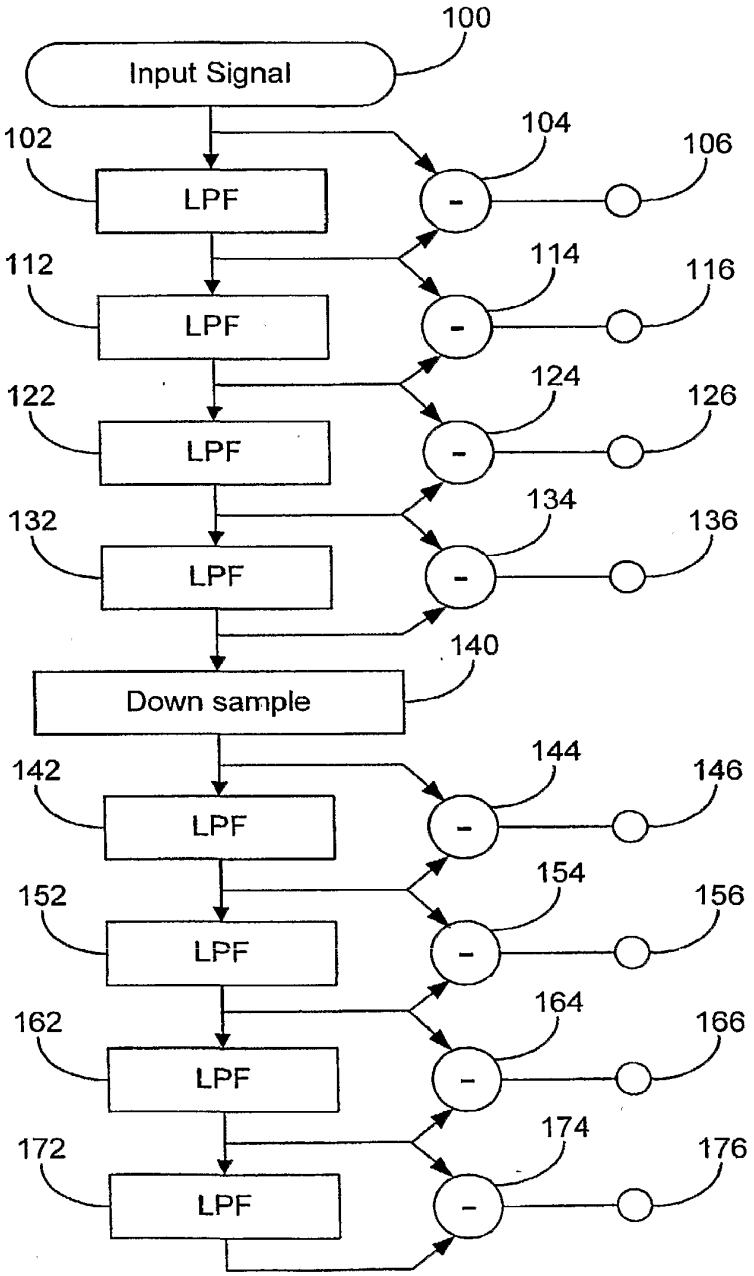
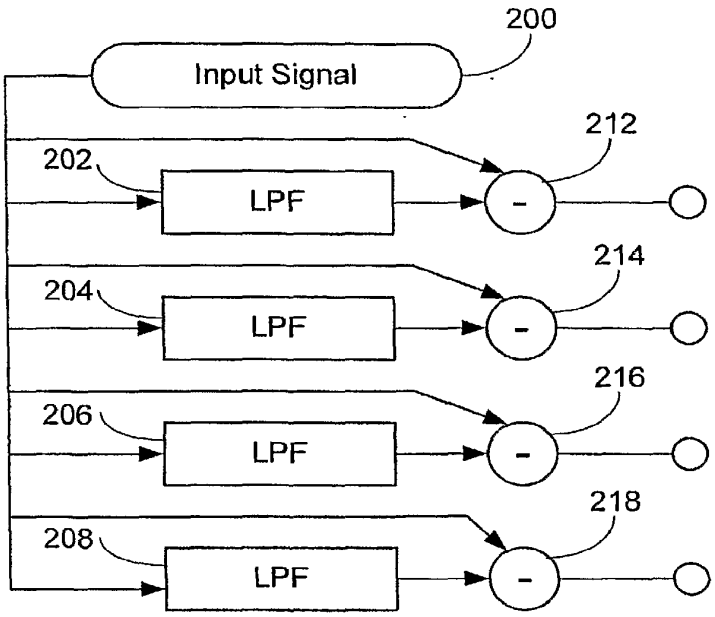


Figure 1

SUBSTITUTE SHEET (RULE 26)


200
Input Signal
212
202
LPF
214
204
LPF
216
206
LPF
218
208
LPF

3/9

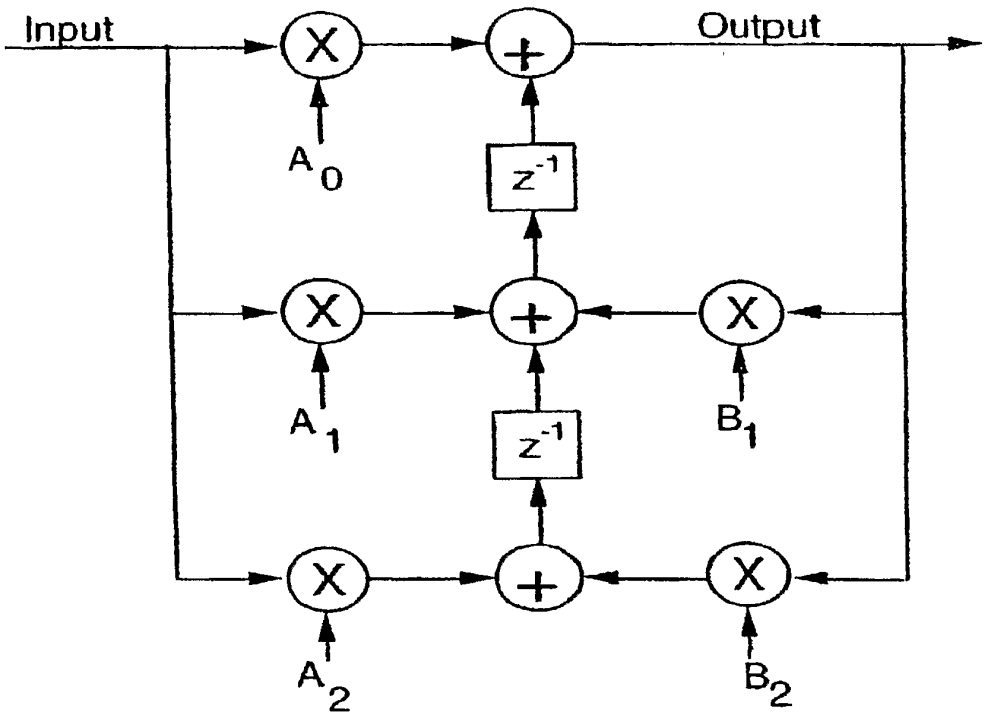


Figure 3

4/9

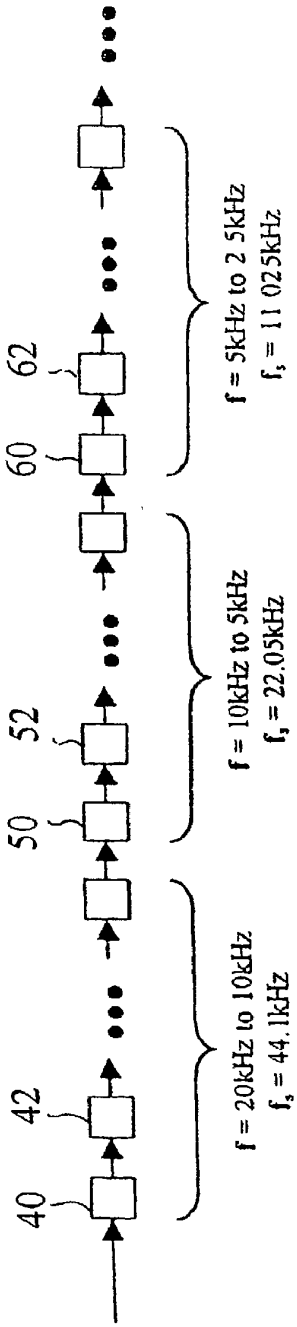

40
42
50
52
60
62
f = 20kHz to 10kHz
f_s = 44.1kHz
f = 10kHz to 5kHz
f_s = 22.05kHz
f = 5kHz to 2.5kHz
f_s = 11.025kHz

Figure 4

5/9

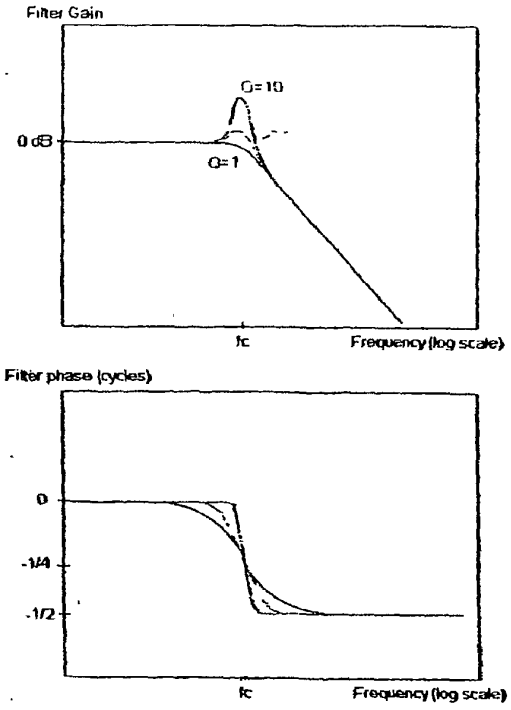


Figure 5

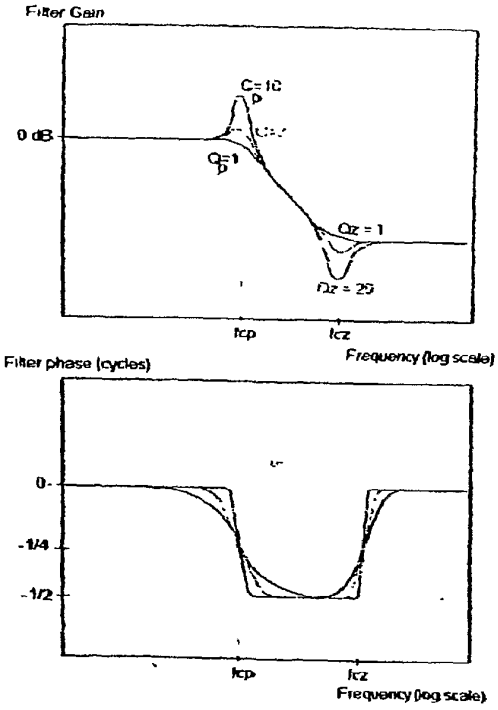


Figure 6

6/9

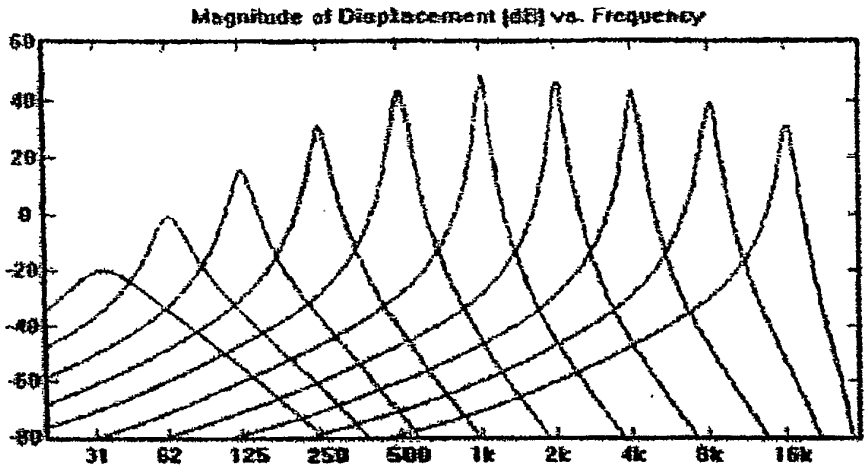


Figure 7

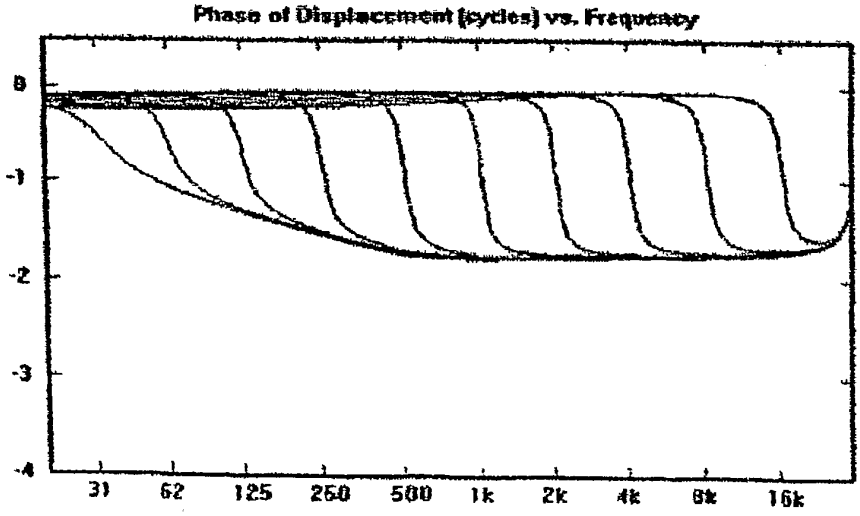


Figure 8

SUBSTITUTE SHEET (RULE 26)

7/9

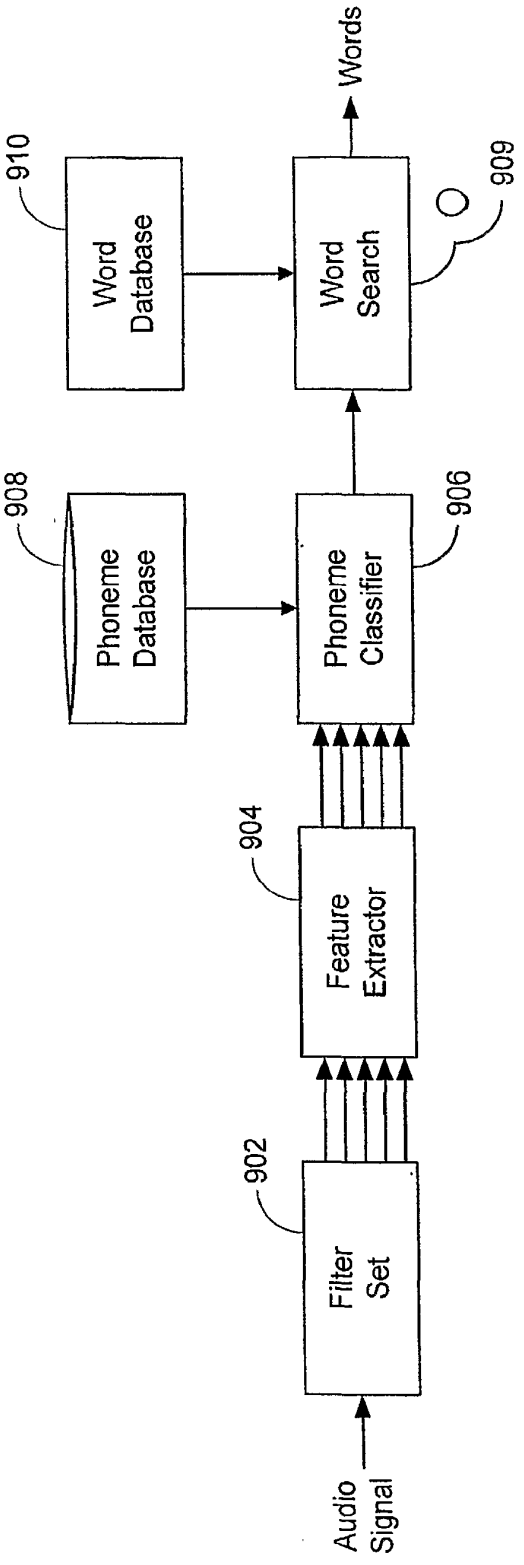

Words
910
Word Database
Word Search
909
908
Phoneme Database
Phoneme Classifier
906
904
Feature Extractor
902
Filter Set
Audio Signal

Figure 9A

8/9

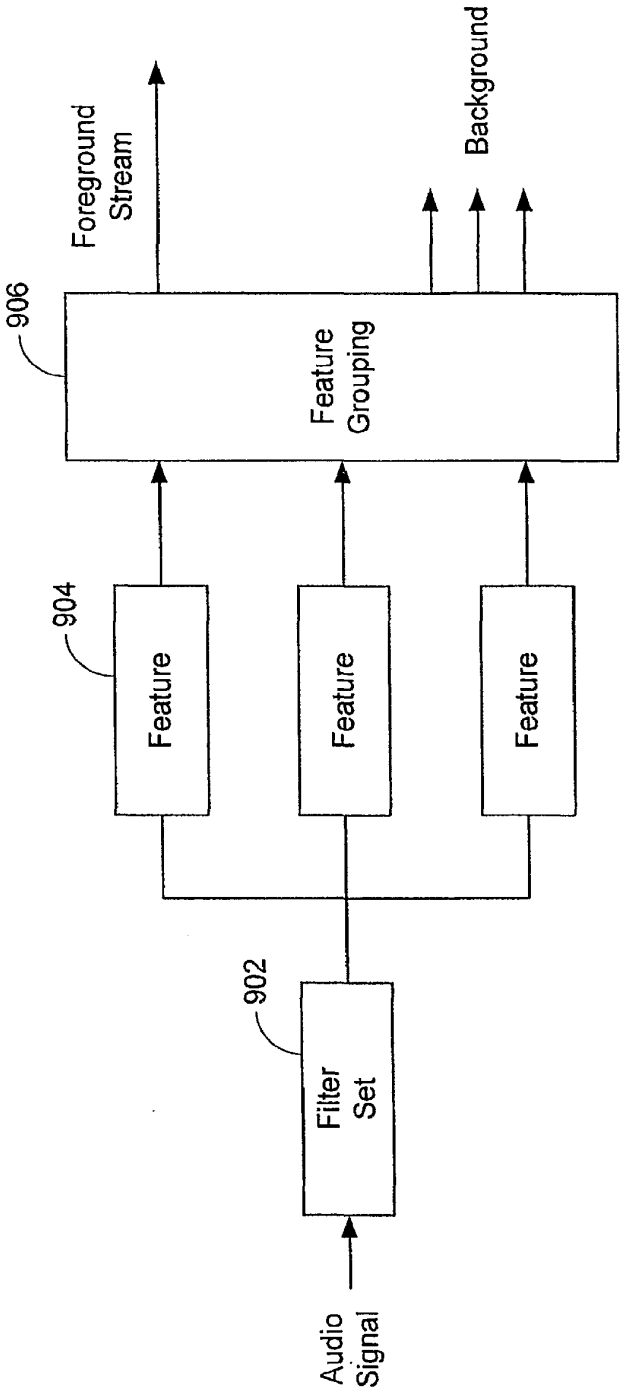

Foreground Stream
Background
906
Feature Grouping
904
Feature
Feature
Feature
902
Filter Set
Audio Signal

Figure 9B

9/9

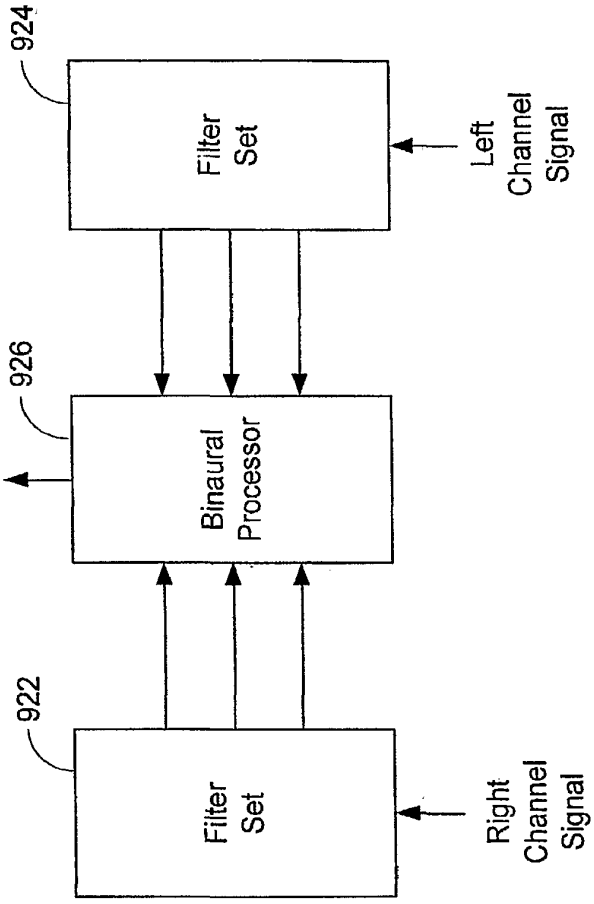

924
Filter Set
Left Channel Signal
926
Binaural Processor
922
Filter Set
Right Channel Signal

Figure 9C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/04124

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/00; H03G 5/00; G10L 19/00
US CL : 700/94; 381/98; 704/500

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 700/94; 381/98; 704/500

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	US 2002/0147595 A1 (BAUMGARTE) 10 October 2002 (10.10.2002), figures1, 2 and 7, page 3 and 5.	1,3-5, 8-12, 14-19, 22-25
A	US 6,513,004 B1 (RIGAZIO et al) 28 January 2003 (28.01.2003), entire document.	1-25
A	US 4,536,844 A (LYON) 20 August 1985 (20.08.1985), entire document.	1-25
A	US 5,473,759 A (SLANEY et al) 05 December 1995 (05.12.1995), entire document.	1-25
A	KATES, JAMES M. "A Time-Domain Digital Cochlear Model" IEEE Transactions of Siganl Processing, December 1991, Volume 39 Number 12, pages 2573-2592.	1-25

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
05 May 2003 (05.05.2003)

Date of mailing of the international search report
29 MAY 2003

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Forester W Isen
Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 January 2010 (14.01.2010)

PCT

(10) International Publication Number
WO 2010/005493 A1

(51) International Patent Classification:
A61F 11/06 (2006.01)

(21) International Application Number:
PCT/US2009/003813

(22) International Filing Date:
26 June 2009 (26.06.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/215,980 30 June 2008 (30.06.2008) US

(71) Applicant *(for all designated States except US)*: **AUDIENCE, INC.** [US/US]; 440 Clyde Avenue, Mountain View, CA 94043 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SOLBACH, Ludger** [DE/US]; C/o Audience, Inc., 440 Clyde Avenue, Mountain View, CA 94043 (US). **MURGIA, Carlo** [IT/US]; C/o Audience, Inc., 440 Clyde Avenue, Mountain View, CA 94043 (US).

(74) Agents: **BACHMANN, Stephen** et al.; Carr & Ferrell LLP, 2200 Geng Road, Palo Alto, CA 94303 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: SYSTEM AND METHOD FOR PROVIDING NOISE SUPPRESSION UTILIZING NULL PROCESSING NOISE SUBTRACTION

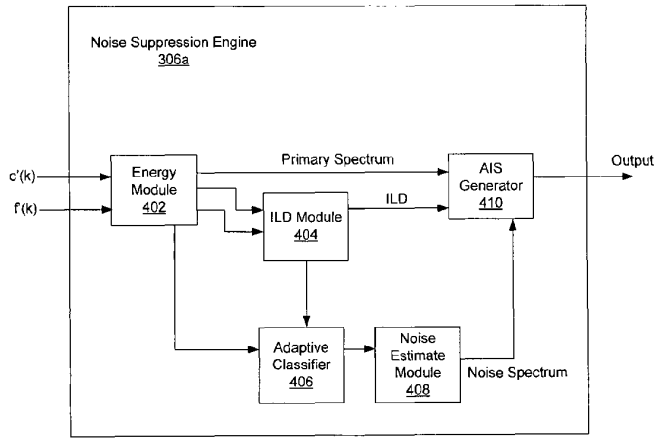


FIG. 4

(57) **Abstract**: Systems and methods for noise suppression using noise subtraction processing are provided. The noise subtraction processing comprises receiving at least a primary and a secondary acoustic signal. A desired signal component may be calculated and subtracted from the secondary acoustic signal to obtaining a noise component signal. A determination may be made of a reference energy ratio and a prediction energy ratio. A determination may be made as to whether to adjust the noise component signal based partially on the reference energy ratio and partially on the prediction energy ratio. The noise component signal may be adjusted or frozen based on the determination. The noise component signal may then be removed from the primary acoustic signal to generate a noise subtracted signal which may be outputted.

WO 2010/005493 A1

SYSTEM AND METHOD FOR PROVIDING NOISE SUPPRESSION UTILIZING NULL PROCESSING NOISE SUBTRACTION

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The present application is related to U.S. Patent Application No. 11/825,563, filed July 6, 2007 and entitled "System and Method for Adaptive Intelligent Noise Suppression," and U.S. Patent Application No. 12/080,115, filed March 31, 2008 and entitled "System and Method for Providing Close Microphone Adaptive Array Processing," both of which are herein incorporated by reference.

[0002] The present application is also related to U.S. Patent Application No. 11/343,524, filed January 30, 2006 and entitled "System and Method for Utilizing Inter-Microphone Level Differences for Speech Enhancement," and U.S. Patent Application No. 11/699,732, filed January 29, 2007 and entitled "System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement," which are incorporated by reference.

BACKGROUND OF THE INVENTION

Field of Invention

[0003] The present invention relates generally to audio processing and more particularly to adaptive noise suppression of an audio signal.

Description of Related Art

[0004] Currently, there are many methods for reducing background noise in an adverse audio environment. One such method is to use a stationary noise suppression system. The stationary noise suppression system will always provide an output noise that is a fixed amount lower than the input noise. Typically, the stationary noise suppression is in the range of 12-13 decibels (dB). The noise

suppression is fixed to this conservative level in order to avoid producing speech distortion, which will be apparent with higher noise suppression.

[0005] In order to provide higher noise suppression, dynamic noise suppression systems based on signal-to-noise ratios (SNR) have been utilized. This SNR may then be used to determine a suppression value. Unfortunately, SNR, by itself, is not a very good predictor of speech distortion due to existence of different noise types in the audio environment. SNR is a ratio of how much louder speech is than noise. However, speech may be a non-stationary signal which may constantly change and contain pauses. Typically, speech energy, over a period of time, will comprise a word, a pause, a word, a pause, and so forth. Additionally, stationary and dynamic noises may be present in the audio environment. The SNR averages all of these stationary and non-stationary speech and noise. There is no consideration as to the statistics of the noise signal; only what the overall level of noise is.

[0006] In some prior art systems, an enhancement filter may be derived based on an estimate of a noise spectrum. One common enhancement filter is the Wiener filter. Disadvantageously, the enhancement filter is typically configured to minimize certain mathematical error quantities, without taking into account a user's perception. As a result, a certain amount of speech degradation is introduced as a side effect of the noise suppression. This speech degradation will become more severe as the noise level rises and more noise suppression is applied. That is, as the SNR gets lower, lower gain is applied resulting in more noise suppression. This introduces more speech loss distortion and speech degradation.

[0007] Some prior art systems invoke a generalized side-lobe canceller. The generalized side-lobe canceller is used to identify desired signals and interfering signals comprised by a received signal. The desired signals propagate from a desired location and the interfering signals propagate from other locations. The interfering signals are subtracted from the received signal with the intention of cancelling interference.

[0008] Many noise suppression processes calculate a masking gain and apply this masking gain to an input signal. Thus, if an audio signal is mostly noise, a masking gain that is a low value may be applied (i.e., multiplied to) the audio signal. Conversely, if the audio signal is mostly desired sound, such as speech, a high value gain mask may be applied to the audio signal. This process is commonly referred to as multiplicative noise suppression.

SUMMARY OF THE INVENTION

[0009] Embodiments of the present invention overcome or substantially alleviate prior problems associated with noise suppression and speech enhancement. In exemplary embodiments, at least a primary and a secondary acoustic signal are received by a microphone array. The microphone array may comprise a close microphone array or a spread microphone array.

[0010] A noise component signal may be determined in each sub-band of signals received by the microphone by subtracting the primary acoustic signal weighted by a complex-valued coefficient σ from the secondary acoustic signal. The noise component signal, weighted by another complex-valued coefficient α, may then be subtracted from the primary acoustic signal resulting in an estimate of a target signal (i.e., a noise subtracted signal).

[0011] A determination may be made as to whether to adjust α. In exemplary embodiments, the determination may be based on a reference energy ratio (g_1) and a prediction energy ratio (g_2). The complex-valued coefficient α may be adapted when the prediction energy ratio is greater than the reference energy ratio to adjust the noise component signal. Conversely, the adaptation coefficient may be frozen when the prediction energy ratio is less than the reference energy ratio. The noise component signal may then be removed from the primary acoustic signal to generate a noise subtracted signal which may be outputted.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] FIG. 1 is an environment in which embodiments of the present invention may be practiced.

[0013] FIG. 2 is a block diagram of an exemplary audio device implementing embodiments of the present invention.

[0014] FIG. 3 is a block diagram of an exemplary audio processing system utilizing a spread microphone array.

[0015] FIG. 4 is a block diagram of an exemplary noise suppression system of the audio processing system of FIG. 3.

[0016] FIG. 5 is a block diagram of an exemplary audio processing system utilizing a close microphone array.

[0017] FIG. 6 is a block diagram of an exemplary noise suppression system of the audio processing system of FIG. 5.

[0018] FIG. 7a is a block diagram of an exemplary noise subtraction engine.

[0019] FIG. 7b is a schematic illustrating the operations of the noise subtraction engine.

[0020] FIG. 8 is a flowchart of an exemplary method for suppressing noise in an audio device.

[0021] FIG. 9 is a flowchart of an exemplary method for performing noise subtraction processing.

DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0022] The present invention provides exemplary systems and methods for adaptive suppression of noise in an audio signal. Embodiments attempt to balance noise suppression with minimal or no speech degradation (i.e., speech loss distortion). In exemplary embodiments, noise suppression is based on an audio source location and applies a subtractive noise suppression process as opposed to a purely multiplicative noise suppression process.

[0023] Embodiments of the present invention may be practiced on any audio device that is configured to receive sound such as, but not limited to, cellular phones, phone handsets, headsets, and conferencing systems. Advantageously, exemplary embodiments are configured to provide improved noise suppression while minimizing speech distortion. While some embodiments of the present invention will be described in reference to operation on a cellular phone, the present invention may be practiced on any audio device.

[0024] Referring to FIG. 1, an environment in which embodiments of the present invention may be practiced is shown. A user acts as a speech source 102 to an audio device 104. The exemplary audio device 104 may include a microphone array. The microphone array may comprise a close microphone array or a spread microphone array.

[0025] In exemplary embodiments, the microphone array may comprise a primary microphone 106 relative to the audio source 102 and a secondary microphone 108 located a distance away from the primary microphone 106. While embodiments of the present invention will be discussed with regards to having two microphones 106 and 108, alternative embodiments may contemplate any number of microphones or acoustic sensors within the microphone array. In some embodiments, the microphones 106 and 108 may comprise omni-directional microphones.

[0026] While the microphones 106 and 108 receive sound (i.e., acoustic signals) from the audio source 102, the microphones 106 and 108 also pick up noise

110. Although the noise 110 is shown coming from a single location in FIG. 1, the noise 110 may comprise any sounds from one or more locations different than the audio source 102, and may include reverberations and echoes. The noise 110 may be stationary, non-stationary, or a combination of both stationary and non-stationary noise.

[0027] Referring now to FIG. 2, the exemplary audio device 104 is shown in more detail. In exemplary embodiments, the audio device 104 is an audio receiving device that comprises a processor 202, the primary microphone 106, the secondary microphone 108, an audio processing system 204, and an output device 206. The audio device 104 may comprise further components (not shown) necessary for audio device 104 operations. The audio processing system 204 will be discussed in more details in connection with FIG. 3.

[0028] In exemplary embodiments, the primary and secondary microphones 106 and 108 are spaced a distance apart in order to allow for an energy level difference between them. Upon reception by the microphones 106 and 108, the acoustic signals may be converted into electric signals (i.e., a primary electric signal and a secondary electric signal). The electric signals may, themselves, be converted by an analog-to-digital converter (not shown) into digital signals for processing in accordance with some embodiments. In order to differentiate the acoustic signals, the acoustic signal received by the primary microphone 106 is herein referred to as the primary acoustic signal, while the acoustic signal received by the secondary microphone 108 is herein referred to as the secondary acoustic signal.

[0029] The output device 206 is any device which provides an audio output to the user. For example, the output device 206 may comprise an earpiece of a headset or handset, or a speaker on a conferencing device.

[0030] FIG. 3 is a detailed block diagram of the exemplary audio processing system 204a according to one embodiment of the present invention. In exemplary embodiments, the audio processing system 204a is embodied within a memory

device. The audio processing system 204a of FIG. 3 may be utilized in embodiments comprising a spread microphone array.

[0031] In operation, the acoustic signals received from the primary and secondary microphones 106 and 108 are converted to electric signals and processed through a frequency analysis module 302. In one embodiment, the frequency analysis module 302 takes the acoustic signals and mimics the frequency analysis of the cochlea (i.e., cochlear domain) simulated by a filter bank. In one example, the frequency analysis module 302 separates the acoustic signals into frequency sub-bands. A sub-band is the result of a filtering operation on an input signal where the bandwidth of the filter is narrower than the bandwidth of the signal received by the frequency analysis module 302. Alternatively, other filters such as short-time Fourier transform (STFT), sub-band filter banks, modulated complex lapped transforms, cochlear models, wavelets, etc., can be used for the frequency analysis and synthesis. Because most sounds (e.g., acoustic signals) are complex and comprise more than one frequency, a sub-band analysis on the acoustic signal determines what individual frequencies are present in the complex acoustic signal during a frame (e.g., a predetermined period of time). According to one embodiment, the frame is 8 ms long. Alternative embodiments may utilize other frame lengths or no frame at all. The results may comprise sub-band signals in a fast cochlea transform (FCT) domain.

[0032] Once the sub-band signals are determined, the sub-band signals are forwarded to a noise subtraction engine 304. The exemplary noise subtraction engine 304 is configured to adaptively subtract out a noise component from the primary acoustic signal for each sub-band. As such, output of the noise subtraction engine 304 is a noise subtracted signal comprised of noise subtracted sub-band signals. The noise subtraction engine 304 will be discussed in more detail in connection with FIG. 7a and FIG. 7b. It should be noted that the noise subtracted sub-band signals may comprise desired audio that is speech or non-speech (e.g., music). The results of the noise subtraction engine 304 may be output to the user or

processed through a further noise suppression system (e.g., the noise suppression engine 306). For purposes of illustration, embodiments of the present invention will discuss embodiments whereby the output of the noise subtraction engine 304 is processed through a further noise suppression system.

[0033] The noise subtracted sub-band signals along with the sub-band signals of the secondary acoustic signal are then provided to the noise suppression engine 306a. According to exemplary embodiments, the noise suppression engine 306a generates a gain mask to be applied to the noise subtracted sub-band signals in order to further reduce noise components that remain in the noise subtracted speech signal. The noise suppression engine 306a will be discussed in more detail in connection with FIG. 4 below.

[0034] The gain mask determined by the noise suppression engine 306a may then be applied to the noise subtracted signal in a masking module 308. Accordingly, each gain mask may be applied to an associated noise subtracted frequency sub-band to generate masked frequency sub-bands. As depicted in FIG. 3, a multiplicative noise suppression system 312a comprises the noise suppression engine 306a and the masking module 308.

[0035] Next, the masked frequency sub-bands are converted back into time domain from the cochlea domain. The conversion may comprise taking the masked frequency sub-bands and adding together phase shifted signals of the cochlea channels in a frequency synthesis module 310. Alternatively, the conversion may comprise taking the masked frequency sub-bands and multiplying these with an inverse frequency of the cochlea channels in the frequency synthesis module 310. Once conversion is completed, the synthesized acoustic signal may be output to the user.

[0036] Referring now to FIG. 4, the noise suppression engine 306a of FIG. 3 is illustrated. The exemplary noise suppression engine 306a comprises an energy module 402, an inter-microphone level difference (ILD) module 404, an adaptive classifier 406, a noise estimate module 408, and an adaptive intelligent suppression

(AIS) generator 410. It should be noted that the noise suppression engine 306a is exemplary and may comprise other combinations of modules such as that shown and described in U.S. Patent Application No. 11/343,524, which is incorporated by reference.

[0037] According to an exemplary embodiment of the present invention, the AIS generator 410 derives time and frequency varying gains or gain masks used by the masking module 308 to suppress noise and enhance speech in the noise subtracted signal. In order to derive the gain masks, however, specific inputs are needed for the AIS generator 410. These inputs comprise a power spectral density of noise (i.e., noise spectrum), a power spectral density of the noise subtracted signal (herein referred to as the primary spectrum), and an inter-microphone level difference (ILD).

[0038] According to exemplary embodiment, the noise subtracted signal (c'(k)) resulting from the noise subtraction engine 304 and the secondary acoustic signal (f'(k)) are forwarded to the energy module 402 which computes energy/power estimates during an interval of time for each frequency band (i.e., power estimates) of an acoustic signal. As can be seen in FIG. 7b, f'(k) may optionally be equal to f(k). As a result, the primary spectrum (i.e., the power spectral density of the noise subtracted signal) across all frequency bands may be determined by the energy module 402. This primary spectrum may be supplied to the AIS generator 410 and the ILD module 404 (discussed further herein). Similarly, the energy module 402 determines a secondary spectrum (i.e., the power spectral density of the secondary acoustic signal) across all frequency bands which is also supplied to the ILD module 404. More details regarding the calculation of power estimates and power spectrums can be found in co-pending U.S. Patent Application No. 11/343,524 and co-pending U.S. Patent Application No. 11/699,732, which are incorporated by reference.

[0039] In two microphone embodiments, the power spectrums are used by an inter-microphone level difference (ILD) module 404 to determine an energy ratio between the primary and secondary microphones 106 and 108. In exemplary

embodiments, the ILD may be a time and frequency varying ILD. Because the primary and secondary microphones 106 and 108 may be oriented in a particular way, certain level differences may occur when speech is active and other level differences may occur when noise is active. The ILD is then forwarded to the adaptive classifier 406 and the AIS generator 410. More details regarding one embodiment for calculating ILD may be can be found in co-pending U.S. Patent Application No. 11/343,524 and co-pending U.S. Patent Application No. 11/699,732. In other embodiments, other forms of ILD or energy differences between the primary and secondary microphones 106 and 108 may be utilized. For example, a ratio of the energy of the primary and secondary microphones 106 and 108 may be used. It should also be noted that alternative embodiments may use cues other then ILD for adaptive classification and noise suppression (i.e., gain mask calculation). For example, noise floor thresholds may be used. As such, references to the use of ILD may be construed to be applicable to other cues.

[0040] The exemplary adaptive classifier 406 is configured to differentiate noise and distractors (e.g., sources with a negative ILD) from speech in the acoustic signal(s) for each frequency band in each frame. The adaptive classifier 406 is considered adaptive because features (e.g., speech, noise, and distractors) change and are dependent on acoustic conditions in the environment. For example, an ILD that indicates speech in one situation may indicate noise in another situation. Therefore, the adaptive classifier 406 may adjust classification boundaries based on the ILD.

[0041] According to exemplary embodiments, the adaptive classifier 406 differentiates noise and distractors from speech and provides the results to the noise estimate module 408 which derives the noise estimate. Initially, the adaptive classifier 406 may determine a maximum energy between channels at each frequency. Local ILDs for each frequency are also determined. A global ILD may be calculated by applying the energy to the local ILDs. Based on the newly calculated global ILD, a running average global ILD and/or a running mean and variance (i.e.,

global cluster) for ILD observations may be updated. Frame types may then be classified based on a position of the global ILD with respect to the global cluster. The frame types may comprise source, background, and distractors.

[0042] Once the frame types are determined, the adaptive classifier 406 may update the global average running mean and variance (i.e., cluster) for the source, background, and distractors. In one example, if the frame is classified as source, background, or distracter, the corresponding global cluster is considered active and is moved toward the global ILD. The global source, background, and distractor global clusters that do not match the frame type are considered inactive. Source and distractor global clusters that remain inactive for a predetermined period of time may move toward the background global cluster. If the background global cluster remains inactive for a predetermined period of time, the background global cluster moves to the global average.

[0043] Once the frame types are determined, the adaptive classifier 406 may also update the local average running mean and variance (i.e., cluster) for the source, background, and distractors. The process of updating the local active and inactive clusters is similar to the process of updating the global active and inactive clusters.

[0044] Based on the position of the source and background clusters, points in the energy spectrum are classified as source or noise; this result is passed to the noise estimate module 408.

[0045] In an alternative embodiment, an example of an adaptive classifier 406 comprises one that tracks a minimum ILD in each frequency band using a minimum statistics estimator. The classification thresholds may be placed a fixed distance (e.g., 3dB) above the minimum ILD in each band. Alternatively, the thresholds may be placed a variable distance above the minimum ILD in each band, depending on the recently observed range of ILD values observed in each band. For example, if the observed range of ILDs is beyond 6dB, a threshold may be place such that it is midway between the minimum and maximum ILDs observed in each band over a certain specified period of time (e.g., 2 seconds). The adaptive classifier is further

discussed in the U.S. nonprovisional application entitled "System and Method for Adaptive Intelligent Noise Suppression," serial number 11/825,563, filed July 6, 2007, which is incorporated by reference.

[0046] In exemplary embodiments, the noise estimate is based on the acoustic signal from the primary microphone 106 and the results from the adaptive classifier 406. The exemplary noise estimate module 408 generates a noise estimate which is a component that can be approximated mathematically by

$$N(t,\omega) = \lambda_I(t,\omega)E_1(t,\omega) + (1-\lambda_I(t,\omega))\min[N(t-1,\omega), E_1(t,\omega)]$$

according to one embodiment of the present invention. As shown, the noise estimate in this embodiment is based on minimum statistics of a current energy estimate of the primary acoustic signal, $E_1(t,\omega)$ and a noise estimate of a previous time frame, $N(t-1,\omega)$. As a result, the noise estimation is performed efficiently and with low latency.

[0047] $\lambda_I(t,\omega)$ in the above equation may be derived from the ILD approximated by the ILD module 404, as

$$\lambda_I(t,\omega) = \begin{cases} \approx 0 & \text{if } ILD(t,\omega) < threshold \\ \approx 1 & \text{if } ILD(t,\omega) > threshold \end{cases}$$

That is, when the primary microphone 106 is smaller than a threshold value (e.g., *threshold* = 0.5) above which speech is expected to be, λ_I is small, and thus the noise estimate module 408 follows the noise closely. When ILD starts to rise (e.g., because speech is present within the large ILD region), λ_I increases. As a result, the noise estimate module 408 slows down the noise estimation process and the speech energy does not contribute significantly to the final noise estimate. Alternative embodiments, may contemplate other methods for determining the noise estimate or noise spectrum. The noise spectrum (i.e., noise estimates for all frequency bands of an acoustic signal) may then be forwarded to the AIS generator 410.

[0048] The AIS generator 410 receives speech energy of the primary spectrum from the energy module 402. This primary spectrum may also comprise some residual noise after processing by the noise subtraction engine 304. The AIS

generator 410 may also receive the noise spectrum from the noise estimate module 408. Based on these inputs and an optional ILD from the ILD module 404, a speech spectrum may be inferred. In one embodiment, the speech spectrum is inferred by subtracting the noise estimates of the noise spectrum from the power estimates of the primary spectrum. Subsequently, the AIS generator 410 may determine gain masks to apply to the primary acoustic signal. More detailed discussion of the AIS generator 410 may be found in U.S. Patent Application No. 11/825,563 entitled "System and Method for Adaptive Intelligent Noise Suppression," which is incorporated by reference. In exemplary embodiments, the gain mask output from the AIS generator 410, which is time and frequency dependent, will maximize noise suppression while constraining speech loss distortion.

[0049] It should be noted that the system architecture of the noise suppression engine 306a is exemplary. Alternative embodiments may comprise more components, less components, or equivalent components and still be within the scope of embodiments of the present invention. Various modules of the noise suppression engine 306a may be combined into a single module. For example, the functionalities of the ILD module 404 may be combined with the functions of the energy module 304.

[0050] Referring now to FIG. 5, a detailed block diagram of an alternative audio processing system 204b is shown. In contrast to the audio processing system 204a of FIG. 3, the audio processing system 204b of FIG. 5 may be utilized in embodiments comprising a close microphone array. The functions of the frequency analysis module 302, masking module 308, and frequency synthesis module 310 are identical to those described with respect to the audio processing system 204a of FIG. 3 and will not be discussed in detail.

[0051] The sub-band signals determined by the frequency analysis module 302 may be forwarded to the noise subtraction engine 304 and an array processing engine 502. The exemplary noise subtraction engine 304 is configured to adaptively subtract out a noise component from the primary acoustic signal for each sub-band.

As such, output of the noise subtraction engine 304 is a noise subtracted signal comprised of noise subtracted sub-band signals. In the present embodiment, the noise subtraction engine 304 also provides a null processing (NP) gain to the noise suppression engine 306a. The NP gain comprises an energy ratio indicating how much of the primary signal has been cancelled out of the noise subtracted signal. If the primary signal is dominated by noise, then NP gain will be large. In contrast, if the primary signal is dominated by speech, NP gain will be close to zero. The noise subtraction engine 304 will be discussed in more detail in connection with FIG. 7a and FIG. 7b below.

[0052] In exemplary embodiments, the array processing engine 502 is configured to adaptively process the sub-band signals of the primary and secondary signals to create directional patterns (i.e., synthetic directional microphone responses) for the close microphone array (e.g., the primary and secondary microphones 106 and 108). The directional patterns may comprise a forward-facing cardioid pattern based on the primary acoustic (sub-band) signals and a backward-facing cardioid pattern based on the secondary (sub-band) acoustic signal. In one embodiment, the sub-band signals may be adapted such that a null of the backward-facing cardioid pattern is directed towards the audio source 102. More details regarding the implementation and functions of the array processing engine 502 may be found (referred to as the adaptive array processing engine) in U.S. Patent Application No. 12/080,115 entitled "System and Method for Providing Close-Microphone Array Noise Reduction," which is incorporated by reference. The cardioid signals (i.e., a signal implementing the forward-facing cardioid pattern and a signal implementing the backward-facing cardioid pattern) are then provided to the noise suppression engine 306b by the array processing engine 502.

[0053] The noise suppression engine 306b receives the NP gain along with the cardioid signals. According to exemplary embodiments, the noise suppression engine 306b generates a gain mask to be applied to the noise subtracted sub-band signals from the noise subtraction engine 304 in order to further reduce any noise

components that may remain in the noise subtracted speech signal. The noise suppression engine 306b will be discussed in more detail in connection with FIG. 6 below.

[0054] The gain mask determined by the noise suppression engine 306b may then be applied to the noise subtracted signal in the masking module 308. Accordingly, each gain mask may be applied to an associated noise subtracted frequency sub-band to generate masked frequency sub-bands. Subsequently, the masked frequency sub-bands are converted back into time domain from the cochlea domain by the frequency synthesis module 310. Once conversion is completed, the synthesized acoustic signal may be output to the user. As depicted in FIG. 5, a multiplicative noise suppression system 312b comprises the array processing engine 502, the noise suppression engine 306b, and the masking module 308.

[0055] Referring now to FIG. 6, the exemplary noise suppression engine 306b is shown in more detail. The exemplary noise suppression engine 306b comprises the energy module 402, the inter-microphone level difference (ILD) module 404, the adaptive classifier 406, the noise estimate module 408, and the adaptive intelligent suppression (AIS) generator 410. It should be noted that the various modules of the noise suppression engine 306b functions similar to the modules in the noise suppression engine 306a.

[0056] In the present embodiment, the primary acoustic signal (c''(k)) and the secondary acoustic signal (f''(k)) are received by the energy module 402 which computes energy/power estimates during an interval of time for each frequency band (i.e., power estimates) of an acoustic signal. As a result, the primary spectrum (i.e., the power spectral density of the primary sub-band signals) across all frequency bands may be determined by the energy module 402. This primary spectrum may be supplied to the AIS generator 410 and the ILD module 404. Similarly, the energy module 402 determines a secondary spectrum (i.e., the power spectral density of the secondary sub-band signal) across all frequency bands which is also supplied to the ILD module 404. More details regarding the calculation of power estimates and

power spectrums can be found in co-pending U.S. Patent Application No. 11/343,524 and co-pending U.S. Patent Application No. 11/699,732, which are incorporated by reference.

[0057] As previously discussed, the power spectrums may be used by the ILD module 404 to determine an energy difference between the primary and secondary microphones 106 and 108. The ILD may then be forwarded to the adaptive classifier 406 and the AIS generator 410. In alternative embodiments, other forms of ILD or energy differences between the primary and secondary microphones 106 and 108 may be utilized. For example, a ratio of the energy of the primary and secondary microphones 106 and 108 may be used. It should also be noted that alternative embodiments may use cues other then ILD for adaptive classification and noise suppression (i.e., gain mask calculation). For example, noise floor thresholds may be used. As such, references to the use of ILD may be construed to be applicable to other cues.

[0058] The exemplary adaptive classifier 406 and noise estimate module 408 perform the same functions as that described in accordance with FIG. 4. That is, the adaptive classifier differentiates noise and distractors from speech and provides the results to the noise estimate module 408 which derives the noise estimate.

[0059] The AIS generator 410 receives speech energy of the primary spectrum from the energy module 402. The AIS generator 410 may also receive the noise spectrum from the noise estimate module 408. Based on these inputs and an optional ILD from the ILD module 404, a speech spectrum may be inferred. In one embodiment, the speech spectrum is inferred by subtracting the noise estimates of the noise spectrum from the power estimates of the primary spectrum. Additionally, the AIS generator 410 uses the NP gain, which indicates how much noise has already been cancelled by the time the signal reaches the noise suppression engine 306b (i.e., the multiplicative mask) to determine gain masks to apply to the primary acoustic signal. In one example, as the NP gain increases, the estimated SNR for the inputs decreases. In exemplary embodiments, the gain mask output from the AIS generator

410, which is time and frequency dependent, may maximize noise suppression while constraining speech loss distortion.

[0060] It should be noted that the system architecture of the noise suppression engine 306b is exemplary. Alternative embodiments may comprise more components, less components, or equivalent components and still be within the scope of embodiments of the present invention.

[0061] FIG. 7a is a block diagram of an exemplary noise subtraction engine 304. The exemplary noise subtraction engine 304 is configured to suppress noise using a subtractive process. The noise subtraction engine 304 may determine a noise subtracted signal by initially subtracting out a desired component (e.g., the desired speech component) from the primary signal in a first branch, thus resulting in a noise component. Adaptation may then be performed in a second branch to cancel out the noise component from the primary signal. In exemplary embodiments, the noise subtraction engine 304 comprises a gain module 702, an analysis module 704, an adaptation module 706, and at least one summing module 708 configured to perform signal subtraction. The functions of the various modules 702 - 708 will be discussed in connection with FIG. 7a and further illustrated in operation in connection with FIG. 7b.

[0062] Referring to FIG. 7a, the exemplary gain module 702 is configured to determine various gains used by the noise subtraction engine 304. For purposes of the present embodiment, these gains represent energy ratios. In the first branch, a reference energy ratio (g_1) of how much of the desired component is removed from the primary signal may be determined. In the second branch, a prediction energy ratio (g_2) of how much the energy has been reduced at the output of the noise subtraction engine 304 from the result of the first branch may be determined. Additionally, an energy ratio (i.e., NP gain) may be determined that represents the energy ratio indicating how much noise has been canceled from the primary signal by the noise subtraction engine 304. As previously discussed, NP gain may be used

by the AIS generator 410 in the close microphone embodiment to adjust the gain mask.

[0063] The exemplary analysis module 704 is configured to perform the analysis in the first branch of the noise subtraction engine 304, while the exemplary adaptation module 306 is configured to perform the adaptation in the second branch of the noise subtraction engine 304.

[0064] Referring to FIG. 7b, a schematic illustrating the operations of the noise subtraction engine 304 is shown. Sub-band signals of the primary microphone signal c(k) and secondary microphone signal f(k) are received by the noise subtraction engine 304 where k represents a discrete time or sample index. c(k) represents a superposition of a speech signal s(k) and a noise signal n(k). f(k) is modeled as a superposition of the speech signal s(k), scaled by a complex-valued coefficient σ, and the noise signal n(k), scaled by a complex-valued coefficient ν. ν represents how much of the noise in the primary signal is in the secondary signal. In exemplary embodiments, ν is unknown since a source of the noise may be dynamic.

[0065] In exemplary embodiments, σ is a fixed coefficient that represents a location of the speech (e.g., an audio source location). In accordance with exemplary embodiments, σ may be determined through calibration. Tolerances may be included in the calibration by calibrating based on more than one position. For a close microphone, a magnitude of σ may be close to one. For spread microphones, the magnitude of σ may be dependent on where the audio device 102 is positioned relative to the speaker's mouth. The magnitude and phase of the σ may represent an inter-channel cross-spectrum for a speaker's mouth position at a frequency represented by the respective sub-band (e.g., Cochlea tap). Because the noise subtraction engine 304 may have knowledge of what σ is, the analysis module 704 may apply σ to the primary signal (i.e., $\sigma(s(k)+n(k))$ and subtract the result from the secondary signal (i.e., $\sigma s(k)+\nu(k)$) in order to cancel out the speech component σ s(k) (i.e., the desired component) from the secondary signal resulting in a noise

component out of the summing module 708. In an embodiment where there is not speech, α is approximately $1/(\nu-\sigma)$, and the adaptation module 706 may freely adapt.

[0066] If the speaker's mouth position is adequately represented by σ, then $f(k)-\sigma c(k) = (\nu-\sigma)n(k)$. This equation indicates that signal at the output of the summing module 708 being fed into the adaptation module 706 (which, in turn, applies an adaptation coefficient $\alpha(k)$) may be devoid of a signal originating from a position represented by σ (e.g., the desired speech signal). In exemplary embodiments, the analysis module 704 applies σ to the secondary signal f(k) and subtracts the result from c(k). Remaining signal (referred to herein as "noise component signal") from the summing module 708 may be canceled out in the second branch.

[0067] The adaptation module 706 may adapt when the primary signal is dominated by audio sources 102 not in the speech location (represented by σ). If the primary signal is dominated by a signal originating from the speech location as represented by σ, adaptation may be frozen. In exemplary embodiments, the adaptation module 706 may adapt using one of a common least-squares method in order to cancel the noise component n(k) from the signal c(k). The coefficient may be update at a frame rate according to on embodiment.

[0068] In an embodiment where n(k) is white and a cross-correlation between s(k) and n(k) is zero within a frame, adaptation may happen every frame with the noise n(k) being perfectly cancelled and the speech s(k) being perfectly unaffected. However, it is unlikely that these conditions may be met in reality, especially if the frame size is short. As such, it is desirable to apply constraints on adaptation. In exemplary embodiments, the adaptation coefficient $\alpha(k)$ may be updated on a per-tap/per-frame basis when the reference energy ratio g_1 and the prediction energy ratio g_2 satisfy the follow condition:

$$g_2 \cdot \gamma > g_1/\gamma$$

where $\gamma > 0$. Assuming, for example, that $\hat{\sigma}(k) = \sigma$, $\alpha(k) = 1/(\nu - \sigma)$, and s(k) and n(k) are uncorrelated, the following may be obtained:

$$g_1 = \frac{E\{(s(k)+n(k))^2\}}{|\nu-\sigma|^2 \cdot E\{n^2(k)\}} = \frac{S+N}{|\nu-\sigma|^2 \cdot N}$$

and

$$g_2 = \frac{|\nu-\sigma|^2 \cdot E\{n^2(k)\}}{E\{s^2(k)\}} = |\nu-\sigma|^2 \cdot \frac{N}{S},$$

where E{...} is an expected value, S is a signal energy, and N is a noise energy. From the previous three equations, the following may be obtained:

$$SNR^2 + SNR < \gamma^2 |\nu-\sigma|^4,$$

where $SNR = S/N$. If the noise is in the same location as the target speech (i.e., $\sigma = \nu$), this condition may not be met, so regardless of the SNR, adaptation may never happen. The further away from the target location the source is, the greater $|\nu-\sigma|^4$ and the larger the SNR is allowed to be while there is still adaptation attempting to cancel the noise.

In exemplary embodiments, adaptation may occur in frames where more signal is canceled in the second branch as opposed to the first branch. Thus, energies may be calculated after the first branch by the gain module 702 and g_1 determined. An energy calculation may also be performed in order to determine g_2 which may indicate if α is allowed to adapt. If $\gamma^2 |\nu-\sigma|^4 > SNR^2 + SNR^4$ is true, then adaptation of α may be performed. However, if this equation is not true, then α is not adapted.

[0069] The coefficient γ may be chosen to define a boundary between adaptation and non-adaptation of α. In an embodiment where a far-field source at 90 degree angle relative to a straight line between the microphones 106 and 108. In this embodiment, the signal may have equal power and zero phase shift between both microphones 106 and 108 (e.g., $\nu = 1$). If the SNR = 1, then $\gamma^2 |\nu-\sigma|^4 = 2$, which is equivalent to $\gamma = \text{sqrt}(2)/|1-\sigma|^4$.

[0070] Lowering γ relative to this value may improve protection of the near-end source from cancellation at the expense of increased noise leakage; raising γ has

an opposite effect. It should be noted that in the microphones 106 and 108, ν =1 may not be a good enough approximation of the far-field/90 degrees situation and may have to substituted by a value obtained from calibration measurements.

[0071] FIG. 8 is a flowchart 800 of an exemplary method for suppressing noise in an audio device. In step 802, audio signals are received by the audio device 102. In exemplary embodiments, a plurality of microphones (e.g., primary and secondary microphones 106 and 108) receive the audio signals. The plurality of microphones may comprise a close microphone array or a spread microphone array.

[0072] In step 804, the frequency analysis on the primary and secondary acoustic signals may be performed. In one embodiment, the frequency analysis module 302 utilizes a filter bank to determine frequency sub-bands for the primary and secondary acoustic signals.

[0073] Noise subtraction processing is performed in step 806. Step 806 will be discussed in more detail in connection with FIG. 9 below.

[0074] Noise suppression processing may then be performed in step 808. In one embodiment, the noise suppression processing may first compute an energy spectrum for the primary or noise subtracted signal and the secondary signal. An energy difference between the two signals may then be determined. Subsequently, the speech and noise components may be adaptively classified according to one embodiment. A noise spectrum may then be determined. In one embodiment, the noise estimate may be based on the noise component. Based on the noise estimate, a gain mask may be adaptively determined.

[0075] The gain mask may then be applied in step 810. In one embodiment, the gain mask may be applied by the masking module 308 on a per sub-band signal basis. In some embodiments, the gain mask may be applied to the noise subtracted signal. The sub-bands signals may then be synthesized in step 812 to generate the output. In one embodiment, the sub-band signals may be converted back to the time domain from the frequency domain. Once converted, the audio signal may be

output to the user in step 814. The output may be via a speaker, earpiece, or other similar devices.

[0076] Referring now to FIG. 9, a flowchart of an exemplary method for performing noise subtraction processing (step 806) is shown. In step 902, the frequency analyzed signals (e.g., frequency sub-band signals or primary signal) are received by the noise subtraction engine 304. The primary acoustic signal may be represented as c(k) = s(k) + n(k) where s(k) represents the desired signal (e.g., speech signal) and n(k) represents the noise signal. The secondary frequency analyzed signal (e.g., secondary signal) may be represented as f(k) = σs(k) + νn(k).

[0077] In step 904, σ may be applied to the primary signal by the analysis module 704. The result of the application of σ to the primary signal may then be subtracted from the secondary signal in step 906 by the summing module 708. The result comprises a noise component signal.

[0078] In step 908, the gains may be calculated by the gain module 702. These gains represent energy ratios of the various signals. In the first branch, a reference energy ratio (g_1) of how much of the desired component is removed from the primary signal may be determined. In the second branch, a prediction energy ratio (g_2) of how much the energy has been reduce at the output of the noise subtraction engine 304 from the result of the first branch may be determined.

[0079] In step 910, a determination is made as to whether α should be adapted. In accordance with one embodiment if $SNR^2 + SNR < \gamma^2|\nu - \sigma|^4$ is true, then adaptation of α may be performed in step 912. However, if this equation is not true, then α is not adapted but frozen in step 914.

[0080] The noise component signal, whether adapted or not, is subtracted from the primary signal in step 916 by the summing module 708. The result is a noise subtracted signal. In some embodiments, the noise subtracted signal may be provided to the noise suppression engine 306 for further noise suppression processing via a multiplicative noise suppression process. In other embodiments,

the noise subtracted signal may be output to the user without further noise suppression processing. It should be noted that more than one summing module 708 may be provided (e.g., one for each branch of the noise subtraction engine 304).

[0081] In step 918, the NP gain may be calculated. The NP gain comprises an energy ratio indicating how much of the primary signal has been cancelled out of the noise subtracted signal. It should be noted that step 918 may be optional (e.g., in close microphone systems).

[0082] The above-described modules may be comprised of instructions that are stored in storage media such as a machine readable medium (e.g., a computer readable medium). The instructions may be retrieved and executed by the processor 202. Some examples of instructions include software, program code, and firmware. Some examples of storage media comprise memory devices and integrated circuits. The instructions are operational when executed by the processor 202 to direct the processor 202 to operate in accordance with embodiments of the present invention. Those skilled in the art are familiar with instructions, processors, and storage media.

[0083] The present invention is described above with reference to exemplary embodiments. It will be apparent to those skilled in the art that various modifications may be made and other embodiments may be used without departing from the broader scope of the present invention. For example, the microphone array discussed herein comprises a primary and secondary microphone 106 and 108. However, alternative embodiments may contemplate utilizing more microphones in the microphone array. Therefore, there and other variations upon the exemplary embodiments are intended to be covered by the present invention.

CLAIMS

What is claimed is:

1. A method for suppressing noise, comprising:

receiving at least a primary and a secondary acoustic signal;

subtracting a desired signal component from the secondary acoustic signal to obtain a noise component signal;

performing a first determination of at least one energy ratio related to the desired signal component and the noise component signal;

performing a second determination of whether to adjust the noise component signal based on the at least one energy ratio;

adjusting the noise component signal based on the second determination;

subtracting the noise component signal from the primary acoustic signal to generate a noise subtracted signal; and

outputting the noise subtracted signal.

2. The method of claim 1 wherein subtracting the desired signal component comprises applying a coefficient representing a source location to the primary acoustic signal to generate the desired signal component.

3. The method of claim 1 wherein the at least one energy ratio comprises a reference energy ratio and a prediction energy ratio.

4. The method of claim 3 further comprising adapting an adaptation coefficient applied to the noise component signal when the prediction energy ratio is greater than the reference energy ratio.

5. The method of claim 3 further comprising freezing an adaptation coefficient applied to the noise component signal when the prediction energy ratio is less than the reference energy ratio.

6. The method of claim 1 further comprising determining a NP gain based on the at least one energy ratio indicating how much of the primary acoustic signal has been cancelled out of the noise subtracted signal.

7. The method of claim 6 further comprising providing the NP gain to a multiplicative noise suppression system.

8. The method of claim 1 wherein the primary and secondary acoustic signals are separated into sub-band signals.

9. The method of claim 1 wherein outputting the noise subtracted signal comprises outputting the noise subtracted signal to a multiplicative noise suppression system.

10. The method of claim 9 wherein the multiplicative noise suppression system comprises generating a gain mask based at least on the noise subtracted signal.

11. The method of claim 10 further comprising applying the gain mask to the noise subtracted signal to generate an audio output signal.

12. A system for suppressing noise, comprising:

a microphone array configured to receive at least a primary and a secondary acoustic signal;

an analysis module configured to generate a desired signal component which may be subtracted from the secondary acoustic signal to obtain a noise component signal;

a gain module configured to perform a first determination of at least one energy ratio related to the desired signal component and the noise component signal;

an adaptation module configured to perform a second determination of whether to adjust the noise component signal based on the at least one energy ratio, the adaption module further configured to adjust the noise component signal based on the second determination; and

at least one summing module configured to subtract the desired signal component from the secondary acoustic signal and to subtract the noise component signal from the primary acoustic signal to generate a noise subtracted signal.

13. The system of claim 12 wherein the analysis module is configured to apply a coefficient representing a source location to the primary acoustic signal to generate the desired signal component.

14. The system of claim 12 wherein the at least one energy ratio comprises a reference energy ratio and a prediction energy ratio.

15. The system of claim 14 wherein the adaptation module is configured to adapt an adaptation coefficient applied to the noise component signal when the prediction energy ratio is greater than the reference energy ratio.

16. The system of claim 14 wherein the adaptation module is configured to freeze an adaptation coefficient applied to the noise component signal when the prediction energy ratio is less than the reference energy ratio.

17. The system of claim 12 wherein further comprising a gain module configured to determine a NP gain based on the at least one energy ratio indicating how much of the primary acoustic signal has been cancelled out of the noise subtracted signal.

18. A machine readable medium having embodied thereon a program, the program providing instructions for a method for suppressing noise using noise subtraction processing, the method comprising:

receiving at least a primary and a secondary acoustic signal;

subtracting a desired signal component from the secondary acoustic signal to obtain a noise component signal;

performing a first determination of at least one energy ratio related to the desired signal component and the noise component signal;

performing a second determination of whether to adjust the noise component signal based on the at least one energy ratio;

adjusting the noise component signal based on the second determination;

subtracting the noise component signal from the primary acoustic signal to generate a noise subtracted signal; and

outputting the noise subtracted signal.

19. The machine readable medium of claim 18 wherein the at least one energy ratio comprises a reference energy ratio and a prediction energy ratio.

20. The machine readable medium of claim 19 wherein the method further comprises adapting an adaptation coefficient applied to the noise component signal when the prediction energy ratio is greater than the reference energy ratio.

21. The machine readable medium of claim 19 wherein the method further comprises freezing an adaptation coefficient applied to the noise component signal when the prediction energy ratio is less than the reference energy ratio.

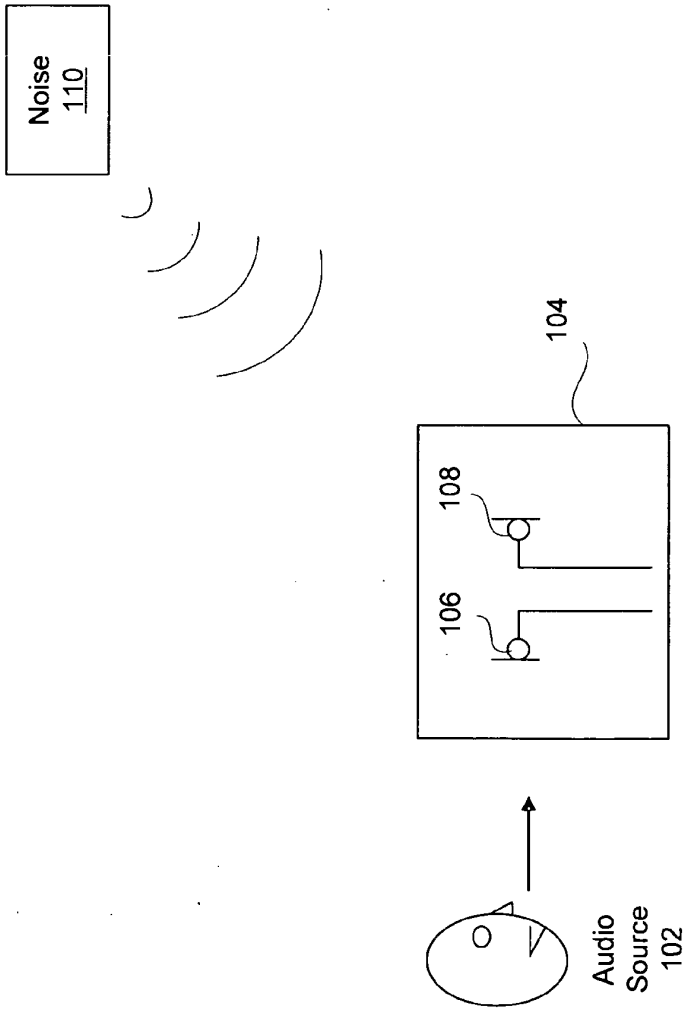


FIG. 1

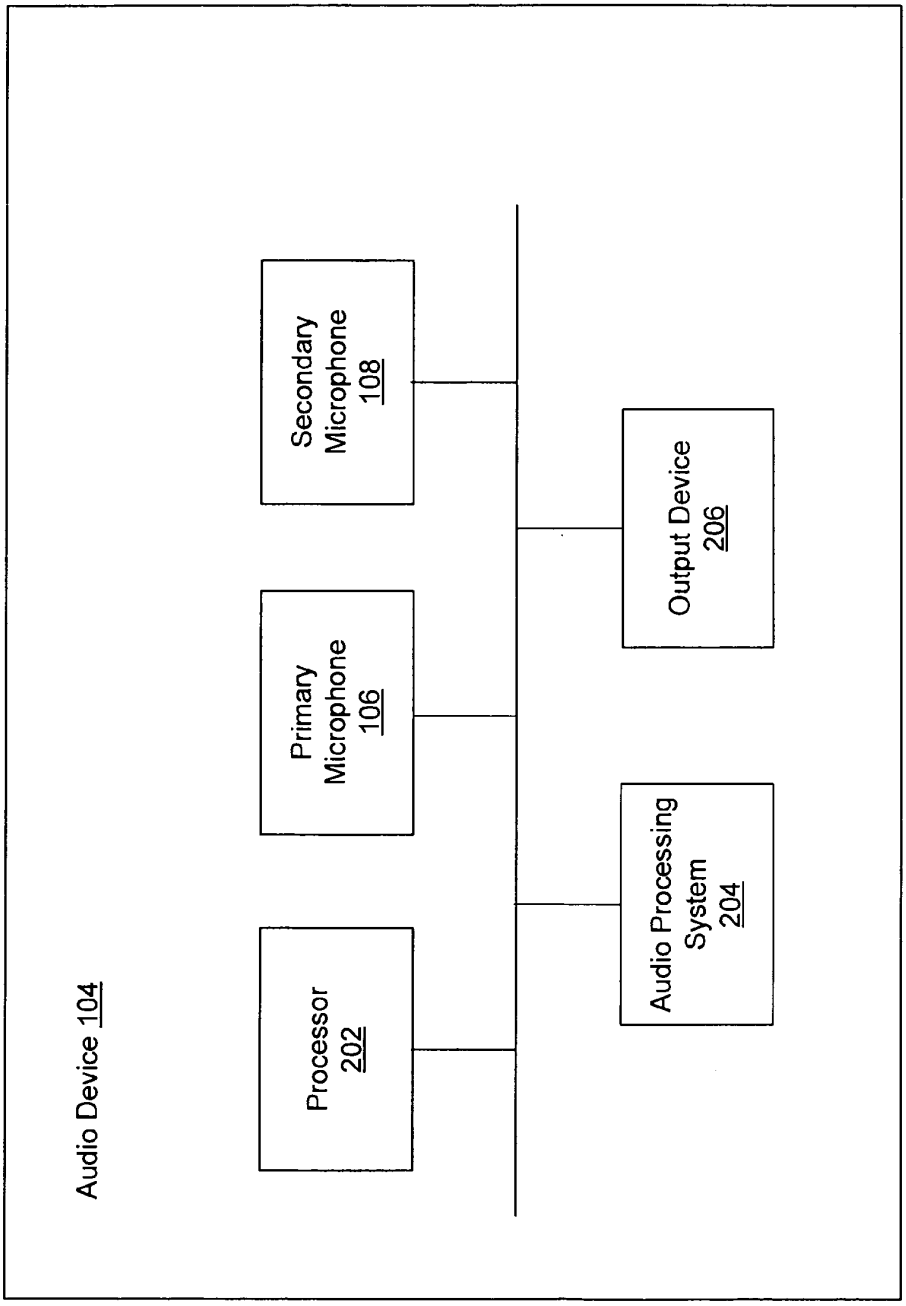

Audio Device 104
Processor 202
Primary Microphone 106
Secondary Microphone 108
Audio Processing System 204
Output Device 206

FIG. 2

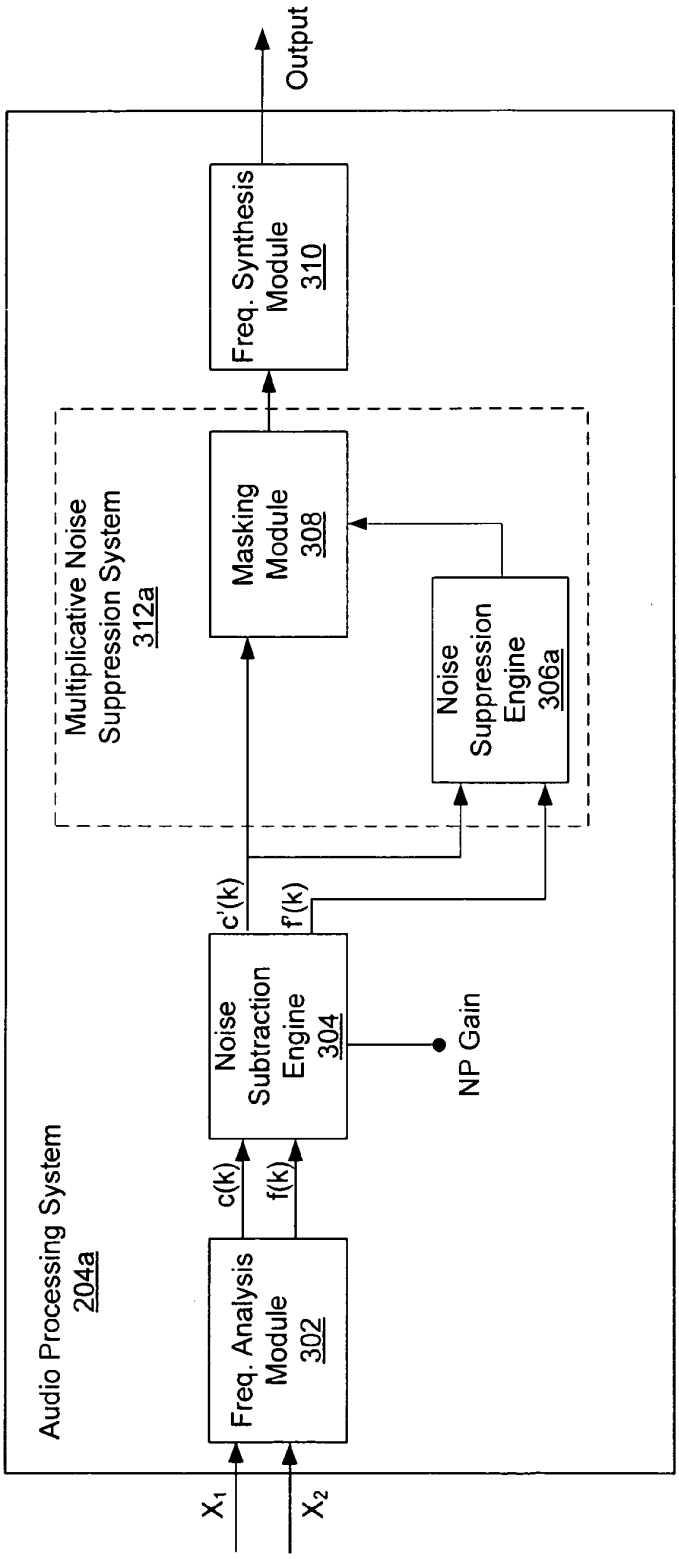

Output
Freq. Synthesis Module 310
Multiplicative Noise Suppression System 312a
Masking Module 308
Noise Suppression Engine 306a
c'(k)
f'(k)
Noise Subtraction Engine 304
NP Gain
Audio Processing System 204a
c(k)
f(k)
Freq. Analysis Module 302
X1
X2

FIG. 3

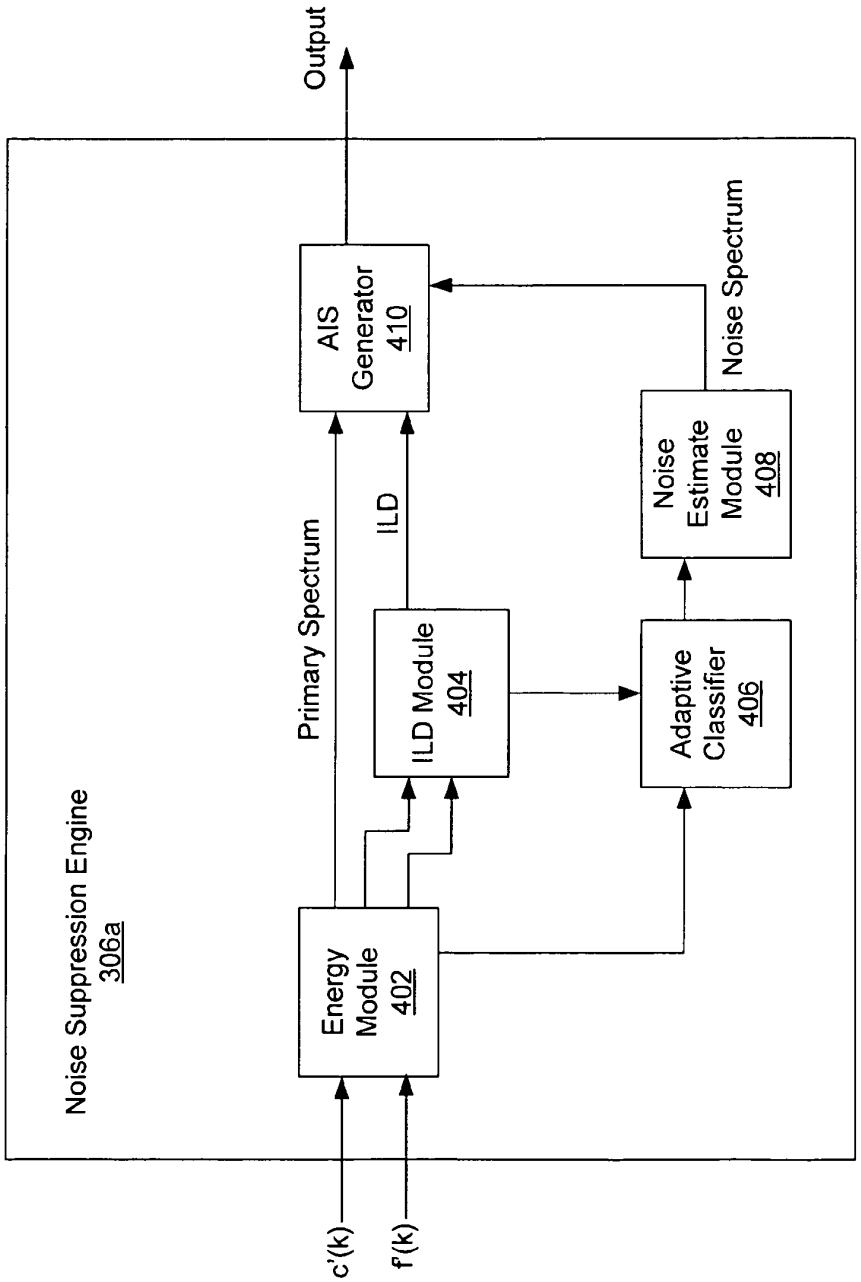

Output
Noise Suppression Engine
306a
AIS Generator 410
Noise Estimate Module 408
Noise Spectrum
Primary Spectrum
ILD
ILD Module 404
Adaptive Classifier 406
Energy Module 402
c'(k)
f'(k)

FIG. 4

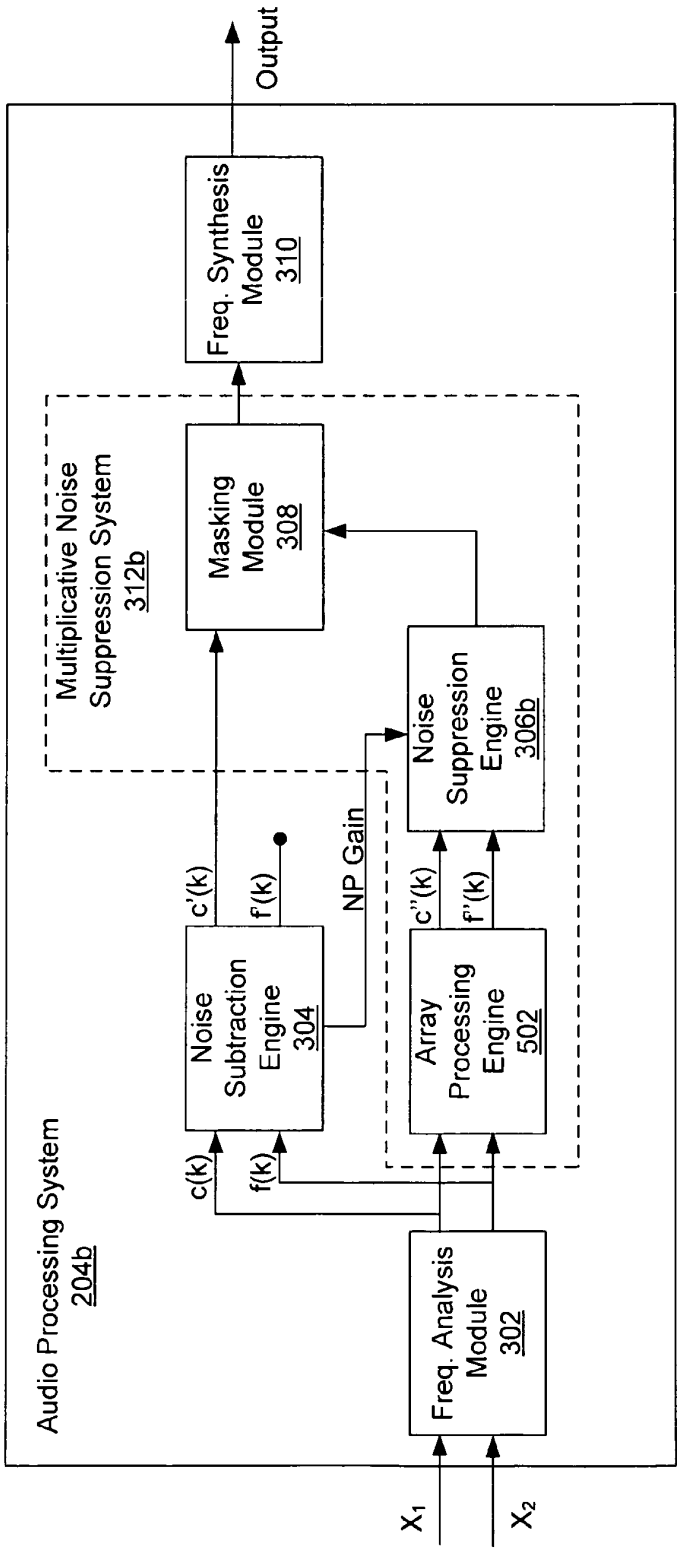

Output
Freq. Synthesis Module 310
Multiplicative Noise Suppression System 312b
Masking Module 308
Noise Suppression Engine 306b
NP Gain
c'(k)
f'(k)
c''(k)
f''(k)
Noise Subtraction Engine 304
Array Processing Engine 502
c(k)
f(k)
Audio Processing System 204b
Freq. Analysis Module 302
X1
X2

FIG. 5

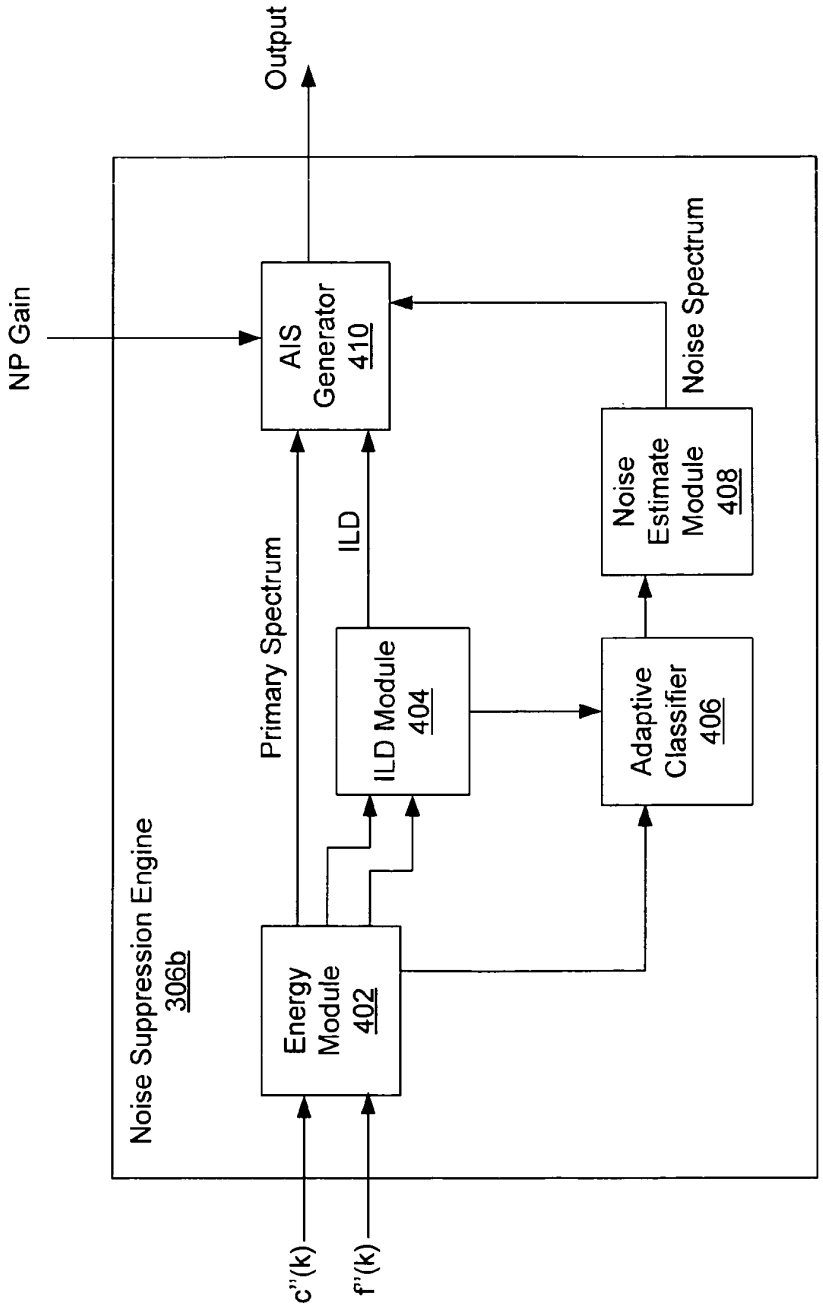

Output
NP Gain
AIS Generator 410
Noise Spectrum
Noise Estimate Module 408
Primary Spectrum
ILD
ILD Module 404
Adaptive Classifier 406
Noise Suppression Engine 306b
Energy Module 402
c''(k)
f''(k)

FIG. 6

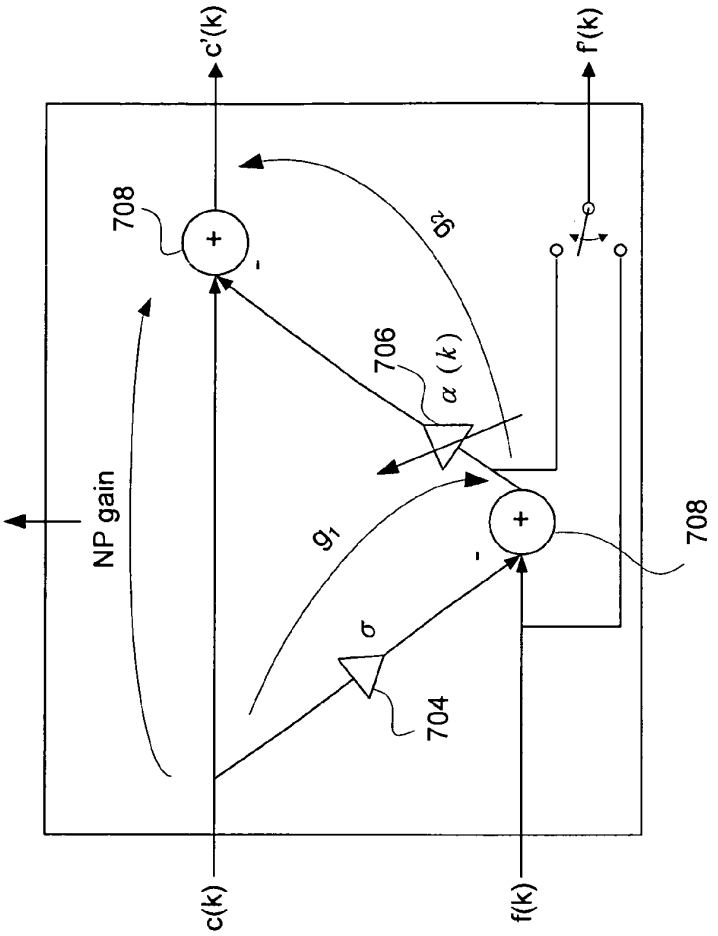

c'(k)
f'(k)
708
+
-
706
α (k)
NP gain
g2
g1
σ
704
+
-
708
c(k)
f(k)

FIG. 7b

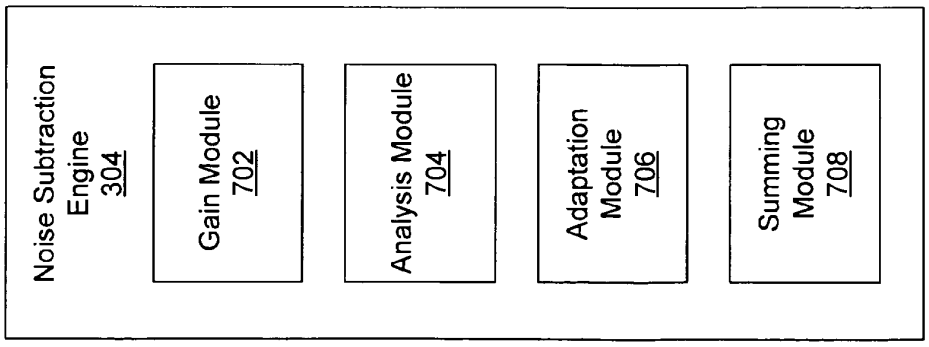

Noise Subtraction Engine 304
Gain Module 702
Analysis Module 704
Adaptation Module 706
Summing Module 708

FIG. 7a

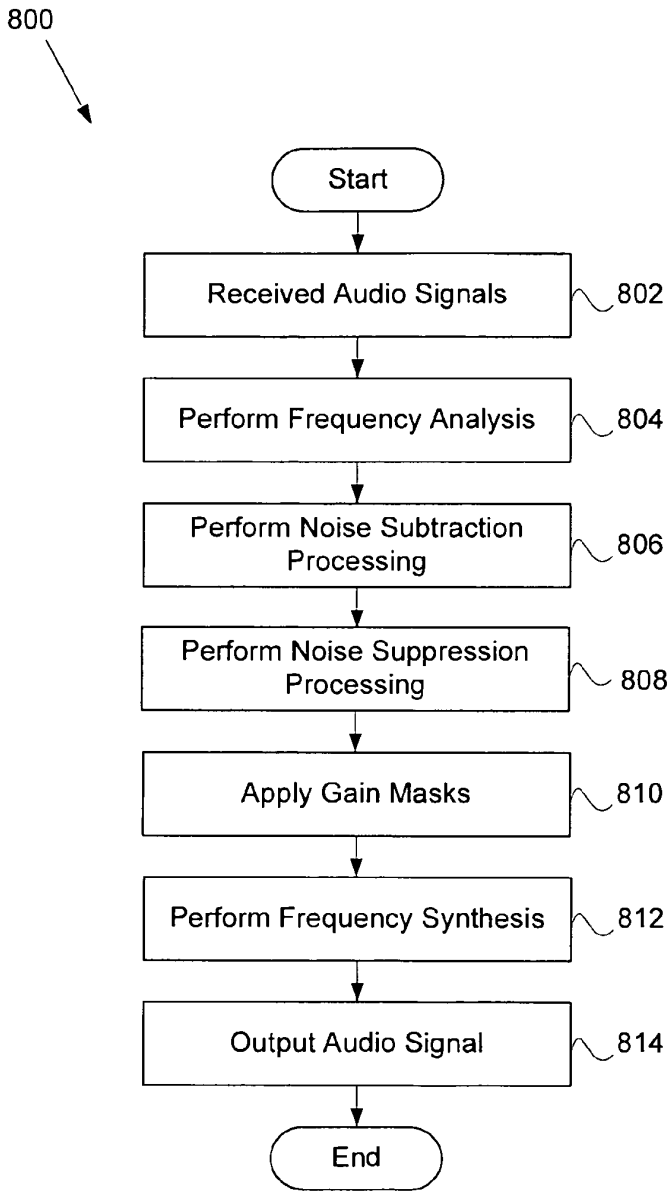

800
Start
Received Audio Signals
802
Perform Frequency Analysis
804
Perform Noise Subtraction Processing
806
Perform Noise Suppression Processing
808
Apply Gain Masks
810
Perform Frequency Synthesis
812
Output Audio Signal
814
End

FIG. 8

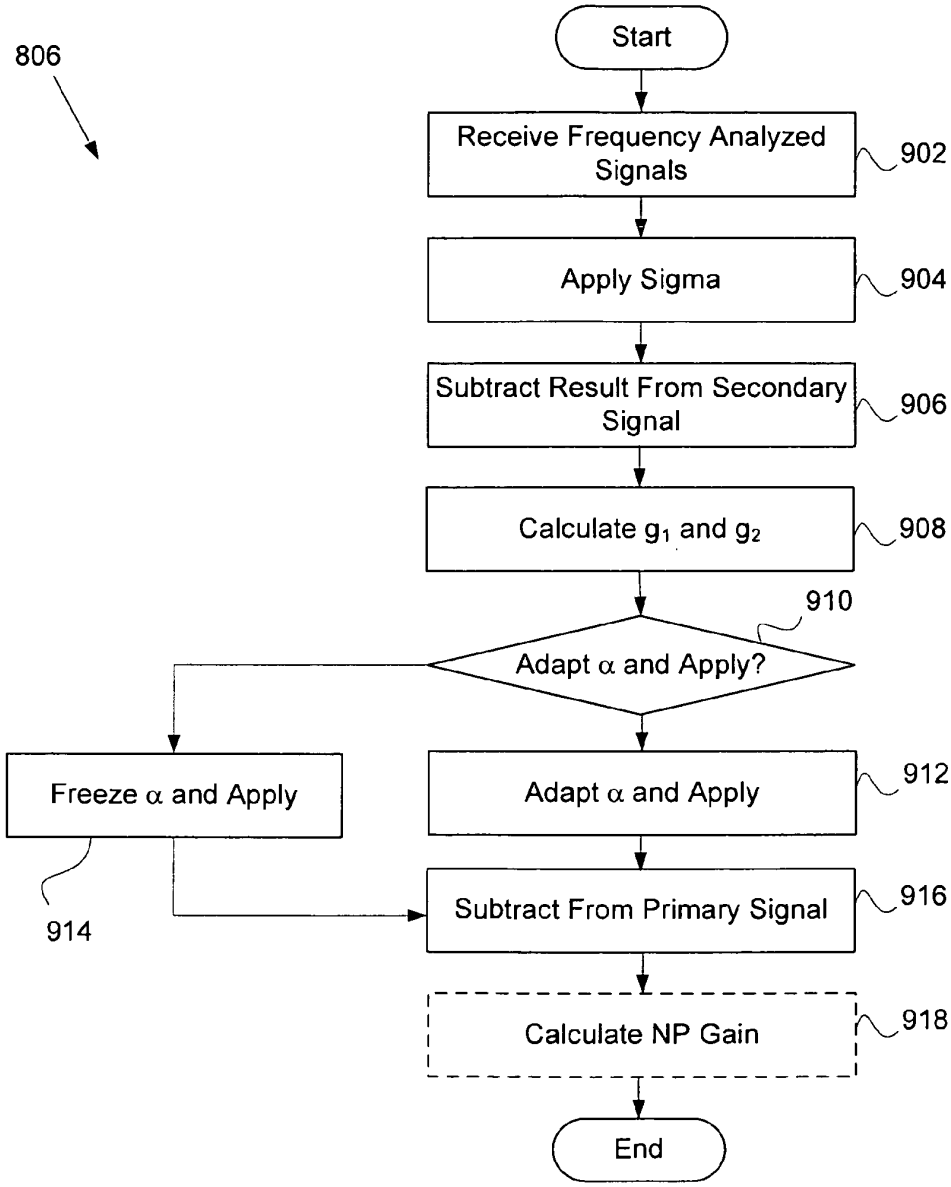

806
Start
Receive Frequency Analyzed Signals
902
Apply Sigma
904
Subtract Result From Secondary Signal
906
Calculate g_1 and g_2
908
910
Adapt α and Apply?
Freeze α and Apply
914
Adapt α and Apply
912
Subtract From Primary Signal
916
Calculate NP Gain
918
End

FIG. 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 09/03813

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A61F 11/06 (2009.01)
USPC - 381/71.11
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 381/71.11

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 381/71.1; 381/150 (keyword limited - see search terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (PGPB, USPT, USOC, EPAB, JPAB); GOOGLE
Terms: noise, suppress, acoustic, signal, subtract, reference, predict, ratio, gain, mask, filter, coefficient, primary, secondary, microphone, mobile.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2008/0019548 A1 (Avendano) 24 January 2008 (24.01.2008), entire document, especially abstract, para [0023], [0029], [0031], [0045], [0053], [0061], [0062].	1-21
Y	US 2003/0101048 A1 (Liu) 29 May 2003 (29.05.2003), entire document, especially abstract, para [0006], [0015], [0036], [0037], [0038], [0039].	1-21

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 August 2009 (10.08.2009)	27 AUG 2009
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 December 2007 (06.12.2007)

PCT

(10) International Publication Number
WO 2007/140003 A2

(51) International Patent Classification:
G10L 19/00 (2006.01)

(21) International Application Number:
PCT/US2007/012628

(22) International Filing Date: 24 May 2007 (24.05.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/441,675 25 May 2006 (25.05.2006) US

(71) Applicant *(for all designated States except US)*: **AUDIENCE, INC.** [US/US]; 1330 Villa Street, Mountain View, California 94041 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SOLBACH, Ludger** [DE/US]; c/o Audience, Inc., 1330 Villa Street, Mountain View, California 94041 (US). **WATTS, Lloyd** [CA/US]; c/o Audience, Inc., 1330 Villa Street, Mountain View, California 94041 (US).

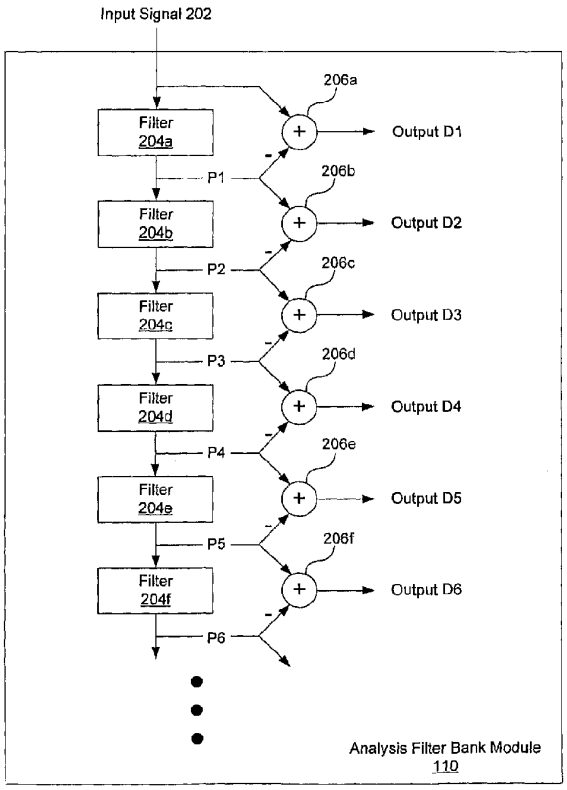
(74) Agents: **YEE, Susan** et al.; Carr & Ferrell LLP, 2200 Geng Road, Palo Alto, California 94303 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR PROCESSING AN AUDIO SIGNAL



(57) Abstract: Systems and methods for audio signal processing are provided. In exemplary embodiments, a filter cascade of complex-valued filters are used to decompose an input audio signal into a plurality of frequency components or sub-band signals. These sub-band signals may be processed for phase alignment, amplitude compensation, and time delay prior to summation of real portions of the sub-band signals to generate a reconstructed audio signal.

WO 2007/140003 A2

WO 2007/140003 A2

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SYSTEM AND METHOD FOR PROCESSING AN AUDIO SIGNAL

By: Ludger Solbach and Lloyd Watts

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] Embodiments of the present invention are related to audio processing, and more particularly to the analysis of audio signals.

Related Art

[0002] There are numerous solutions for splitting an audio signal into sub-bands and deriving frequency-dependent amplitude and phase characteristics varying over time. Examples include windowed fast Fourier transform / inverse fast Fourier transform (FFT/IFFT) systems as well as parallel banks of finite impulse response (FIR) and infinite impulse response (IIR) filter banks. These conventional solutions, however, all suffer from deficiencies.

[0003] Disadvantageously, windowed FFT systems only provide a single, fixed bandwidth for each frequency band. Typically, a bandwidth which is applied from low frequency to high frequency is chosen with a fine resolution at the bottom. For example, at 100Hz, a filter (bank) with a 50kHz bandwidth is desired. This means, however, that at 8kHz, a 50Hz bandwidth is used where a wider bandwidth such as 400Hz may be more appropriate. Therefore, flexibility to match human perception cannot be provided by these systems.

[0004] Another disadvantage of windowed FFT systems is that inadequate fine frequency resolution of sparsely sampled windowed FFT systems at high frequencies can result in objectionable artifacts (e.g., "musical noise") if modifications are applied, (e.g., for noise suppression). The number

of artifacts can be reduced to some extent by dramatically reducing the number of samples of overlap between the windowed frames size "FFT hop size" (i.e., increasing oversampling). Unfortunately, computational costs of FFT systems increase as oversampling increases. Similarly, the FIR subclass of filter banks are also computationally expensive due to the convolution of the sampled impulse responses in each sub-band which can result in high latency. For example, a system with a window of 256 samples will require 256 multiplies and a latency of 128 samples, if the window is symmetric.

[0005] The IIR subclass is computationally less expensive due to its recursive nature, but implementations employing only real-valued filter coefficients present difficulties in achieving near-perfect reconstruction, especially if the sub-band signals are modified. Further, phase and amplitude compensation as well as time-alignment for each sub-band is required in order to produce a flat frequency response at the output. The phase compensation is difficult to perform with real-valued signals, since they are missing the quadrature component for straight-forward computation of amplitude and phase with fine time-resolution. The most common way to determine amplitude and frequency is to apply a Hilbert transform on each stage output. But an extra computation step is required for calculating the Hilbert transform in real-valued filter banks, and is computationally expensive.

[0006] Therefore, there is a need for systems and methods for analyzing and reconstructing an audio signal that is computationally less expensive than existing systems, while providing low end-to-end latency, and the necessary degrees of freedom for time-frequency resolution.

SUMMARY OF THE INVENTION

[0007] Embodiments of the present invention provide systems and methods for audio signal processing. In exemplary embodiments, a filter cascade of complex-valued filters is used to decompose an input audio signal into a plurality of sub-band signals. In one embodiment, an input signal is filtered with a complex-valued filter of the filter cascade to produce a first filtered signal. The first filtered signal is subtracted from the input signal to derive a first sub-band signal. Next, the first filtered signal is processed by a next complex-valued filter of the filter cascade to produce a next filtered signal. The processes repeat until the last complex-valued filters in the cascade has been utilized. In some embodiments, the complex-valued filters are single pole, complex-valued filters.

[0008] Once the input signal is decomposed, the sub-band signals may be processed by a reconstruction module. The reconstruction module is configured to perform a phase alignment on one or more of the sub-band signals. The reconstruction module may also be configured to perform amplitude compensation on one or more of the sub-band signals. Further, a time delay may be performed on one or more of the sub-band signals by the reconstruction module. Real portions of the compensated and/or time delayed sub-band signals are summed to generate a reconstructed audio signal.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] FIG. 1 is an exemplary block diagram of a system employing embodiments of the present invention;

[0010] FIG. 2 is an exemplary block diagram of the analysis filter bank module in an exemplary embodiment of the present invention;

[0011] FIG. 3 is illustrates a filter of the analysis filter bank module, according to one embodiment;

[0012] FIG. 4 illustrates for every six (6) sub-bands a log display of magnitude and phase of the sub-band transfer function;

[0013] FIG. 5 illustrates for every six (6) stages a log display of magnitude and phase of the accumulated filter transfer functions;

[0014] FIG. 6 illustrates the operation of the exemplary reconstruction module;

[0015] FIG. 7 illustrates a graphical representation of an exemplary reconstruction of the audio signal; and

[0016] FIG. 8 is a flowchart of an exemplary method for reconstructing an audio signal.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0017] Embodiments of the present invention provide systems and methods for near perfect reconstruction of an audio signal. The exemplary system utilizes a recursive filter bank to generate quadrature outputs. In exemplary embodiments, the filter bank comprises a plurality of complex-valued filters. In further embodiments, the filter bank comprises a plurality of single pole, complex-valued filters.

[0018] Referring to FIG. 1, an exemplary system 100 in which embodiments of the present invention may be practiced is shown. The system 100 may be any device, such as, but not limited to, a cellular phone, hearing aid, speakerphone, telephone, computer, or any other device capable of processing audio signals. The system 100 may also represent an audio path of any of these devices.

[0019] The system 100 comprises an audio processing engine 102, an audio source 104, a conditioning module 106, and an audio sink 108. Further components not related to reconstruction of the audio signal may be provided in the system 100. Additionally, while the system 100 describes a logical progression of data from each component of FIG. 1 to the next, alternative embodiments may comprise the various components of the system 100 coupled via one or more buses or other elements.

[0020] The exemplary audio processing engine 102 processes the input (audio) signals inputted via the audio source 104. In one embodiment, the audio processing engine 102 comprises software stored on a device which is operated upon by a general processor. The audio processing engine 102, in various embodiments, comprises an analysis filter bank module 110, a modification module 112, and a reconstruction module 114. It should be noted that more, less, or functionally equivalent modules may be provided in the audio processing engine 102. For example, one or more the modules 110-

114 may be combined into few modules and still provide the same functionality.

[0021] The audio source 104 comprises any device which receives input (audio) signals. In some embodiments, the audio source 104 is configured to receive analog audio signals. In one example, the audio source 104 is a microphone coupled to an analog-to-digital (A/D) converter. The microphone is configured to receive analog audio signals while the A/D converter samples the analog audio signals to convert the analog audio signals into digital audio signals suitable for further processing. In other examples, the audio source 104 is configured to receive analog audio signals while the conditioning module 106 comprises the A/D converter. In alternative embodiments, the audio source 104 is configured to receive digital audio signals. For example, the audio source 104 is a disk device capable of reading audio signal data stored on a hard disk or other forms of media. Further embodiments may utilize other forms of audio signal sensing/capturing devices.

[0022] The conditioning module 106 pre-processes the input signal (i.e., any processing that does not require decomposition of the input signal). In one embodiment, the conditioning module 106 comprises an auto-gain control. The conditioning module 106 may also perform error correction and noise filtering. The conditioning module 106 may comprise other components and functions for pre-processing the audio signal.

[0023] The analysis filter bank module 110 decomposes the received input signal into a plurality of sub-band signals. In some embodiments, the outputs from the analysis filter bank module 110 can be used directly (e.g., for a visual display). The analysis filter bank module 110 will be discussed in more detail in connection with FIG. 2. In exemplary embodiments, each sub-band signal represents a frequency component.

[0024] The exemplary modification module 112 receives each of the sub-band signals over respective analysis paths from the analysis filter bank

module 110. The modification module 112 can modify/adjust the sub-band signals based on the respective analysis paths. In one example, the modification module 112 filters noise from sub-band signals received over specific analysis paths. In another example, a sub-band signal received from specific analysis paths may be attenuated, suppressed, or passed through a further filter to eliminate objectionable portions of the sub-band signal.

[0025] The reconstruction module 114 reconstructs the modified sub-band signals into a reconstructed audio signal for output. In exemplary embodiments, the reconstruction module 114 performs phase alignment on the complex sub-band signals, performs amplitude compensation, cancels the complex portion, and delays remaining real portions of the sub-band signals during reconstruction in order to improve resolution of the reconstructed audio signal. The reconstruction module 114 will be discussed in more details in connection with FIG. 6.

[0026] The audio sink 108 comprises any device for outputting the reconstructed audio signal. In some embodiments, the audio sink 108 outputs an analog reconstructed audio signal. For example, the audio sink 108 may comprise a digital-to-analog (D/A) converter and a speaker. In this example, the D/A converter is configured to receive and convert the reconstructed audio signal from the audio processing engine 102 into the analog reconstructed audio signal. The speaker can then receive and output the analog reconstructed audio signal. The audio sink 108 can comprise any analog output device including, but not limited to, headphones, ear buds, or a hearing aid . Alternately, the audio sink 108 comprises the D/A converter and an audio output port configured to be coupled to external audio devices (e.g., speakers, headphones, ear buds, hearing aid.)

[0027] In alternative embodiments, the audio sink 108 outputs a digital reconstructed audio signal. In another example, the audio sink 108 is a disk device, wherein the reconstructed audio signal may be stored onto a hard disk

or other medium. In alternate embodiments, the audio sink 108 is optional and the audio processing engine 102 produces the reconstructed audio signal for further processing (not depicted in FIG. 1).

[0028] Referring now to FIG. 2, the exemplary analysis filter bank module 110 is shown in more detail. In exemplary embodiments, the analysis filter bank module 110 receives an input signal 202, and processes the input signal 202 through a series of filters 204 to produce a plurality of sub-band signals or components (e.g., P1-P6). Any number of filters 204 may comprise the analysis filter bank module 110. In exemplary embodiments, the filters 204 are complex valued filters. In further embodiments, the filters 204 are first order filters (e.g., single pole, complex valued). The filters 204 are further discussed in FIG. 3.

[0029] In exemplary embodiments, the filters 204 are organized into a filter cascade whereby an output of one filter 204 becomes an input in a next filter 204 in the cascade. Thus, the input signal 202 is fed to a first filter 204a. An output signal, P1, of the first filter 204a is subtracted from the input signal 202 by a first computation node 206a to produce an output D1. The output D1 represents the difference signal between the signal going into the first filter 204a and the signal after the first filter 204a.

[0030] In alternative embodiments, benefits of the filter cascade may be realized without the use of the computation node 206 to determine sub-band signals. That is, the output of each filter 204 may be used directly to represent energy of the signal at the output or be displayed, for example.

[0031] Because of the cascade structure of the analysis filter bank module 110, the output signal, P1, is now an input signal into a next filter 204b in the cascade. Similar to the process associated with the first filter 204a, an output of the next filter 204b (i.e., P2) is subtracted from the input signal P1 by a next computation node 206b to obtain a next frequency band or channel (i.e., output D2). This next frequency channel emphasizes frequencies

between cutoff frequencies of the present filter 204b and the previous filter 204a. This process continues through the remainder of the filters 204 of the cascade.

[0032] In one embodiment, sets of filters in the cascade are separated into octaves. Filter parameters and coefficients may then be shared among corresponding filters (in a similar position) in different octaves. This process is described in detail in U.S. Patent Application Serial No. 09/534,682.

[0033] In some embodiments, the filters 204 are single pole, complex-valued filters. For example, the filters 204 may comprise first order digital or analog filters that operate with complex values. Collectively, the outputs of the filters 204 represent the sub-band components of the audio signal. Because of the computation node 206, each output represents a sub-band, and a sum of all outputs represents the entire input signal 202. Since the cascading filters 204 are first order, the computational expense may be much less than if the cascading filters 204 were second order or more. Further, each sub-band extracted from the audio signal can be easily modified by altering the first order filters 204. In other embodiments, the filters 204 are complex-valued filters and not necessarily single pole.

[0034] In further embodiments, the modification module 112 (FIG. 1) can process the outputs of the computation node 206 as necessary. For example, the modification module 112 may half wave rectify the filtered sub-bands. Further, the gain of the outputs can be adjusted to compress or expand a dynamic range. In some embodiments, the output of any filter 204 may be downsampled before being processed by another chain/cascade of filters 204.

[0035] In exemplary embodiments, the filters 204 are infinite impulse response (IIR) filters with cutoff frequencies designed to produce a desired channel resolution. The filters 204 may perform successive Hilbert

transformations with a variety of coefficients upon the complex audio signal in order to suppress or output signals within specific sub-bands.

[0036] FIG. 3 is a block diagram illustrating this signal flow in one exemplary embodiment of the present invention. The output of the filter 204, $y_{real}[n]$ and $y_{imag}[n]$ is passed as an input $x_{real}[n+1]$ and $x_{imag}[n+1]$, respectively, of a next filter 204 in the cascade. The term "n" identifies the sub-band to be extracted from the audio signal, where "n" is assumed to be an integer. Since the IIR filter 204 is recursive, the output of the filter can change based on previous outputs. The imaginary components of the input signal (e.g., $x_{imag}[n]$) can be summed after, before, or during the summation of the real components of the signal. In one embodiment, the filter 204 can be described by the complex first order difference equation $y(k) = g * (x(k) + b * x(k-1)) + a * y(k-1)$ where $b = r\_z * \exp(i * theta\_p)$ and $a = -r\_p * \exp(i * theta\_p)$ and "y" is a sample index.

[0037] In the present embodiment, "g" is a gain factor. It should be noted that the gain factor can be applied anywhere that does not affect the pole and zero locations. In alternative embodiments, the gain may be applied by the modification module 112 (FIG. 1) after the audio signals have been decomposed into sub-band signals.

[0038] Referring now to FIG. 4, an example log display of magnitude and phase for every six (6) sub-bands of an audio signal is shown. The magnitude and phase information is based on outputs from the analysis filter bank module 110 (FIG. 1). That is, the amplitudes shown in FIG. 4 are the outputs (i.e., output D1-D6) from the computation node 206 (FIG. 2). In the present example, the analysis filter bank module 110 is operating at a 16kHz sampling rate with 235 sub-bands for a frequency range from 80Hz to 8kHz. End-to-end latency of this analysis filter bank module 110 is 17.3 ms.

[0039] In some embodiments, it is desirable to have a wide frequency response at high frequencies and a narrow frequency response at low frequencies. Because embodiments of the present invention are adaptable to many audio sources 104 (FIG. 1), different bandwidths at different frequencies may be used. Thus, fast responses with wide bandwidths at high frequencies and slow response with a narrow, short bandwidth at low frequencies may be obtained. This results in responses that are much more adapted to the human ear with relatively low latency (e.g., 12 ms).

[0040] Referring now to FIG. 5, an example of magnitude and phase per stage of an analytic cochlea design is shown. The amplitude shown in FIG. 5 is the outputs of filters 204 of FIG. 2 (e.g., P1-P6).

[0041] FIG. 6 illustrates operation of the reconstruction module 114 according to one embodiment of the present invention. In exemplary embodiments, the phase of each sub-band signal is aligned, amplitude compensation is performed, the complex portion of each sub-band signal is removed, and then time is aligned by delaying each sub-band signal as necessary to achieve a flat reconstruction spectrum and reduce impulse response dispersion.

[0042] Because the filters use complex signals (e.g., real and imaginary parts), phase may be derived for any sample. Additionally, amplitude may also be calculated by $A = \sqrt{\left((y_{real}[n])^2 + (y_{imag}[n])^2\right)}$. Thus, the reconstruction of the audio signal is mathematically made easier. As a result of this approach, the amplitude and phase for any sample is readily available for further processing (i.e., to the modification module 112 (FIG. 1).

[0043] Since the impulse responses of the sub-band signals may have varying group delays, merely summing up the outputs of the analysis filter bank module 110 (FIG. 1) may not provide an accurate reconstruction of the audio signal. Consequently, the output of a sub-band can be delayed by the

sub-band's impulse response peak time so that all sub-band filters have their impulse response envelope maximum at a same instance in time.

[0044] In an embodiment where the impulse response waveform maximum is later in time than the desired group delay, the filter output is multiplied with a complex constant such that the real part of the impulse response has a local maximum at the desired group delay.

[0045] As shown, sub-band signals 602 (e.g., S_0, S_n, and S_m) are received by the reconstruction module 114 from the modification module 112 (FIG. 1). Coefficients 604 (e.g., a_0, a_n, and a_m) are then applied to the sub-band signal. The coefficient comprises a fixed, complex factor (i.e., comprising a real and imaginary portion). Alternately, the coefficients 604 can be applied to the sub-band signal within the analysis filter bank module 110. The application of the coefficient to each sub-band signal aligns the phases of the sub-band signal and compensates each amplitude. In exemplary embodiments, the coefficients are predetermined. After the application of the coefficient, the imaginary portion is discarded by a real value module 606 (i.e., Re{ }).

[0046] Each real portion of the sub-band signal is then delayed by a delay Z^{-1} 608. This delay allows for cross sub-band alignment. In one embodiment, the delay Z^{-1} 608 provides a one tap delay. After the delay, the respective sub-band signal is summed in a summation node 610, resulting in a value. The partially reconstructed signal is then carried into a next summation node 610 and applied to a next delayed sub-band signal. The process continues until all sub-band signals are summed resulting in a reconstructed audio signal. The reconstructed audio signal is then suitable for the audio sink 108 (FIG. 1). Although the delays Z^{-1} 608 are depicted after sub-band signals are summed, the order of operations of the reconstruction module 114 can be interchangeable.

[0047] FIG. 7 illustrates a reconstruction graph based on the example of FIG. 4 and FIG. 5. The reconstruction (i.e., reconstructed audio signal) is

obtained by combining the outputs of each filter 206 (FIG. 2) after phase alignment, amplitude compensation, and delay for cross sub-band alignment by the reconstruction module 114 (FIG. 1). As a result, the reconstruction graph is relatively flat.

[0048] Referring now to FIG. 8, a flowchart 800 of an exemplary method for audio signal processing is provided. In step 802, an audio signal is decomposed into sub-band signals. In exemplary embodiments, the audio signal is processed by the analysis filter bank module 110 (FIG. 1). The processing comprises filtering the audio signal through a cascade of filters 204 (FIG. 2), the output of each filter 204 resulting in a sub-band signal at the respective outputs 206. In one embodiment, the filters 204 are complex-valued filters. In a further embodiment, the filters 204 are single pole, complex-valued filters.

[0049] After sub-band decomposition, the sub-band signals are processed through the modification module 112 (FIG. 1) in step 804. In exemplary embodiments, the modification module 112 (FIG. 1) adjusts the gain of the outputs to compress or expand a dynamic range. In some embodiments, the modification module 112 may suppress objectionable sub-band signals.

[0050] A reconstruction module 114 (FIG. 1) then performs phase and amplitude compensation on each sub-band signal in step 806. In one embodiment, the phase and amplitude compensation occurs by applying a complex coefficient to the sub-band signal. The imaginary portion of the compensated sub-band signal is then discarded in step 808. In other embodiments, the imaginary portion of the compensated sub-band signal is retained.

[0051] Using the real portion of the compensated sub-band signal, the sub-band signal is delayed for cross-sub-band alignment in step 810. In one

embodiment, the delay is obtained by utilizing a delay line in the reconstruction module 114.

[0052] In step 812, the delayed sub-band signals are summed to obtain a reconstructed signal. In exemplary embodiments, each sub-band signal/segment represents a frequency.

[0053] Embodiments of the present invention have been described above with reference to exemplary embodiments. It will be apparent to those skilled in the art that various modifications may be made and other embodiments can be used without departing from the broader scope of the invention. Therefore, these and other variations upon the exemplary embodiments are intended to be covered by the present invention.

WHAT IS CLAIMED IS:

1. A method for processing audio signals, comprising:

filtering an input signal with a complex-valued filter of a filter cascade to produce a first filtered signal;

subtracting the first filtered signal from the input signal to derive a first sub-band signal;

filtering the first filtered signal with a next complex-valued filter of the filter cascade to produce a next filtered signal; and

subtracting the next filtered signal from the first filtered signal to derive a next sub-band signal.

2. The method of claim 1 wherein the complex-valued filter and the next complex-valued filter are single pole, complex-valued filters.

3. The method of claim 1 further comprising performing phase alignment on one or more of the sub-band signals.

4. The method of claim 3 further comprising disposing of an imaginary portion of the one or more phase aligned sub-band signals.

5. The method of claim 1 further comprising performing amplitude compensation on one or more of the sub-band signals.

6. The method of claim 1 further comprising performing a time delay on one or more of the sub-band signals for cross-sub-band alignment.

7. The method of claim 6 further comprising summing the delayed one or more sub-band signals to generate a reconstructed audio signal.

8. The method of claim 1 further comprising pre-processing the input signal prior to filtering the input signal with the complex-valued filter of the filter cascade.

9. The method of claim 1 further comprising modifying one or more of the sub-band signals based on an analysis path from the filter cascade.

10. The method of claim 1 wherein the sub-band signals are frequency components of the input signal.

11. A system for processing an audio signal, comprising:

an audio processing engine comprising a filter cascade of complex-valued filters configured to derive a plurality of sub-band signals from an input signal, the set of complex-valued filters arranged in the filter cascade whereby an output of each complex-valued filter is passed to a next complex-valued filter in the filter cascade.

12. The system of claim 11 wherein the complex-valued filters are single pole, complex-valued filters.

13. The system of claim 11 wherein the audio processing engine further comprises a reconstruction module configured to perform phase alignment on one or more of the sub-band signals.

14. The system of claim 11 wherein the audio processing engine further comprises a reconstruction module configured to perform amplitude compensation on one or more of the sub-band signals.

15. The system of claim 11 wherein the audio processing engine further comprises a reconstruction module configured to perform a time delay on one or more of the sub-band signals.

16. The system of claim 11 wherein the audio processing engine further comprises a modification module configured to modify one or more of the sub-band signals based on an analysis path from the filter cascade.

17. The system of claim 11 further comprising a conditioning module configured to pre-process the input signal prior to filtering the input signal with the filter cascade.

18. A machine-readable medium having embodied thereon a program, the program being executable by a machine to perform a method for processing an audio signal, the method comprising:

filtering an input signal with a complex-valued filter of a filter cascade to produce a first filtered signal;

subtracting the first filtered signal from the input signal to derive a first sub-band signal;

filtering the first filtered signal with a next complex-valued filter of the filter cascade to produce a next filtered signal; and

subtracting the next filtered signal from the first filtered signal to derive a next sub-band signal.

19. The machine-readable medium of claim 18 wherein the complex-valued filter and the next complex-valued filter are single pole, complex-valued filters.

20. The machine-readable medium of claim 18 wherein the method further comprises performing phase alignment on one or more of the sub-band signals.

21. The machine-readable medium of claim 18 wherein the method further comprises performing amplitude compensation on one or more of the sub-band signals.

22. The machine-readable medium of claim 18 wherein the method further comprises performing a time delay on one or more the sub-band signals.

23. The machine-readable medium of claim 18 wherein the method further comprises pre-processing the input signal prior to filtering the input signal with the filter cascade.

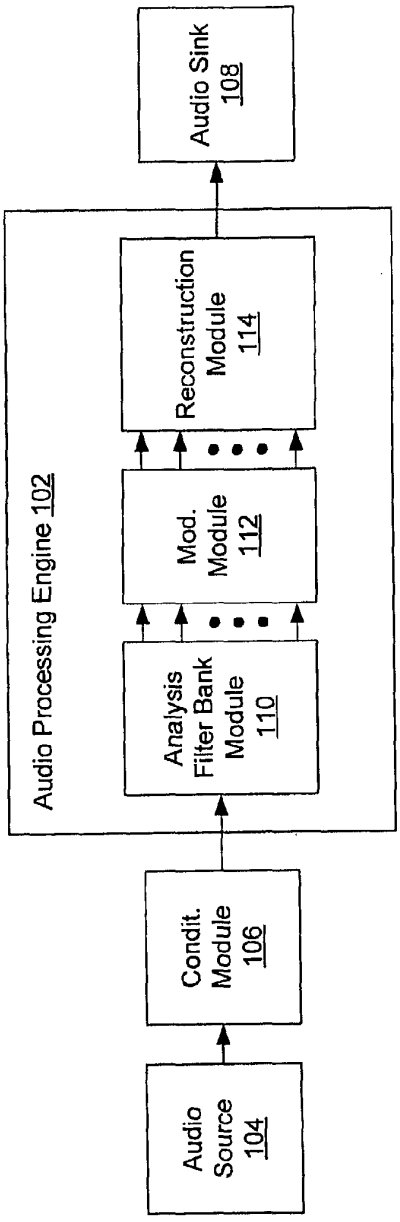

Audio Sink 108
Reconstruction Module 114
Mod. Module 112
Analysis Filter Bank Module 110
Audio Processing Engine 102
Condit. Module 106
Audio Source 104
100

FIG. 1

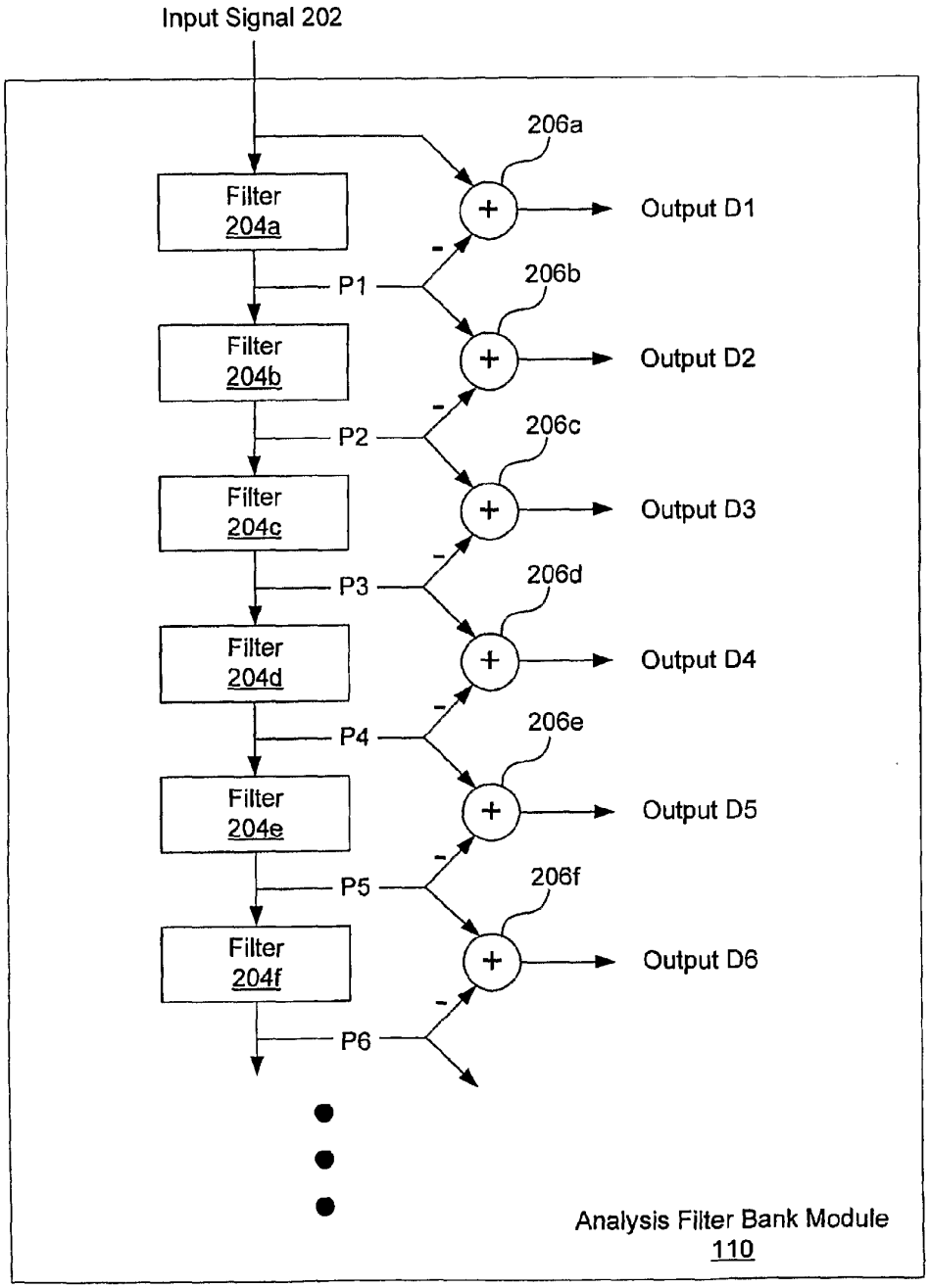

Input Signal 202
206a
Filter 204a
Output D1
P1
206b
Filter 204b
Output D2
P2
206c
Filter 204c
Output D3
P3
206d
Filter 204d
Output D4
P4
206e
Filter 204e
Output D5
P5
206f
Filter 204f
Output D6
P6
Analysis Filter Bank Module 110

FIG. 2

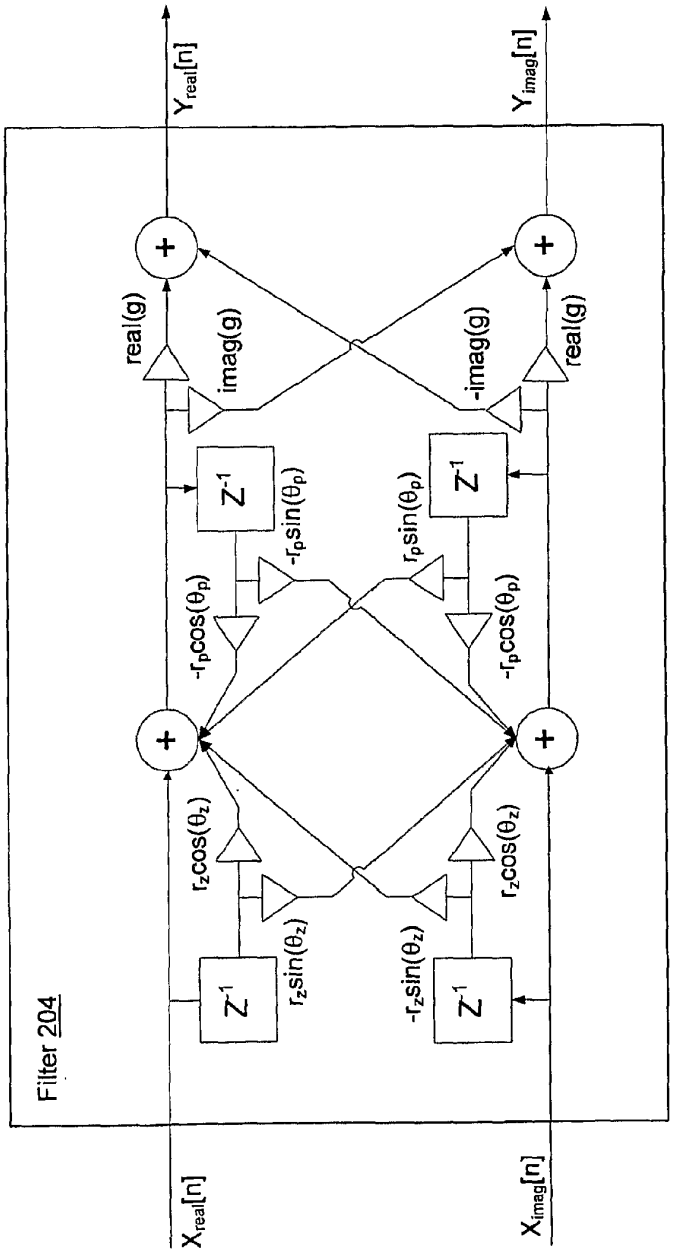

Filter 204
$Y_{real}[n]$
$Y_{imag}[n]$
$X_{real}[n]$
$X_{imag}[n]$
real(g)
imag(g)
-imag(g)
real(g)
Z^{-1}
$-r_p\sin(\theta_p)$
$r_p\sin(\theta_p)$
$-r_p\cos(\theta_p)$
$-r_p\cos(\theta_p)$
$r_z\cos(\theta_z)$
$r_z\sin(\theta_z)$
$-r_z\sin(\theta_z)$
$r_z\cos(\theta_z)$

FIG. 3

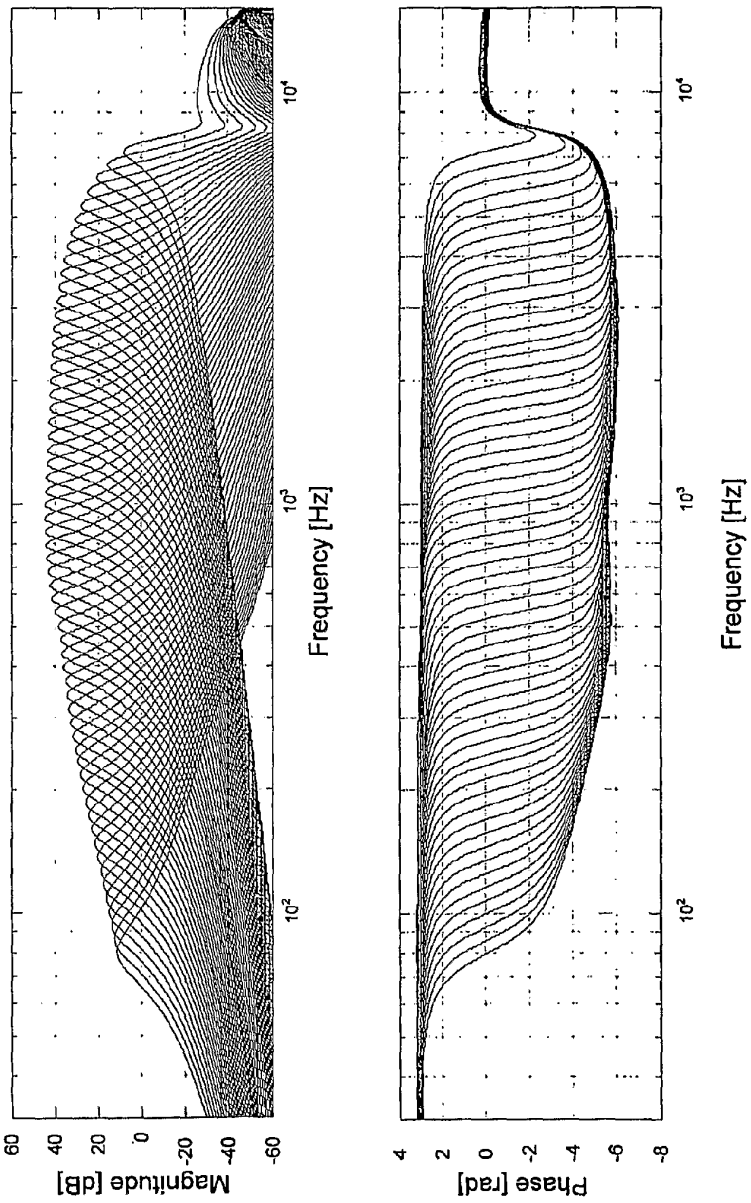

Magnitude [dB]
60
40
20
0
-20
-40
-60
10^2
10^3
10^4
Frequency [Hz]
Phase [rad]
4
2
0
-2
-4
-6
-8
10^2
10^3
10^4
Frequency [Hz]

FIG. 4

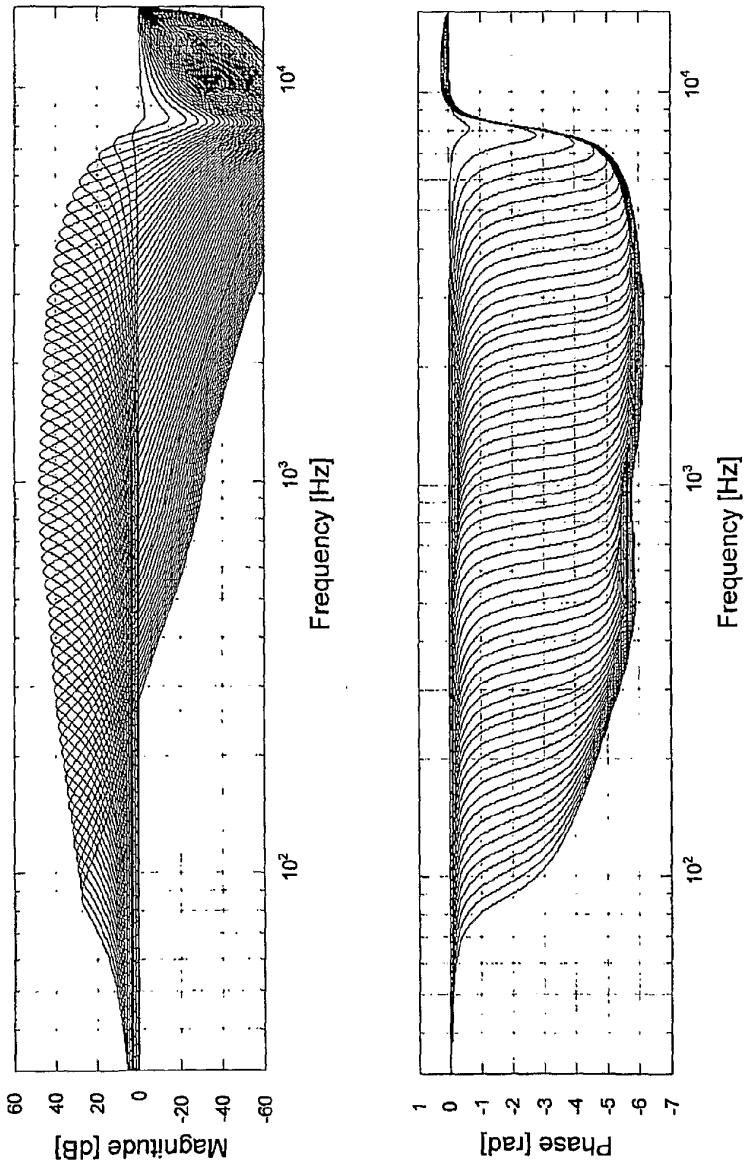

Magnitude [dB]
60
40
20
0
-20
-40
-60
10^2
10^3
10^4
Frequency [Hz]
Phase [rad]
1
0
-1
-2
-3
-4
-5
-6
-7
10^2
10^3
10^4
Frequency [Hz]

FIG. 5

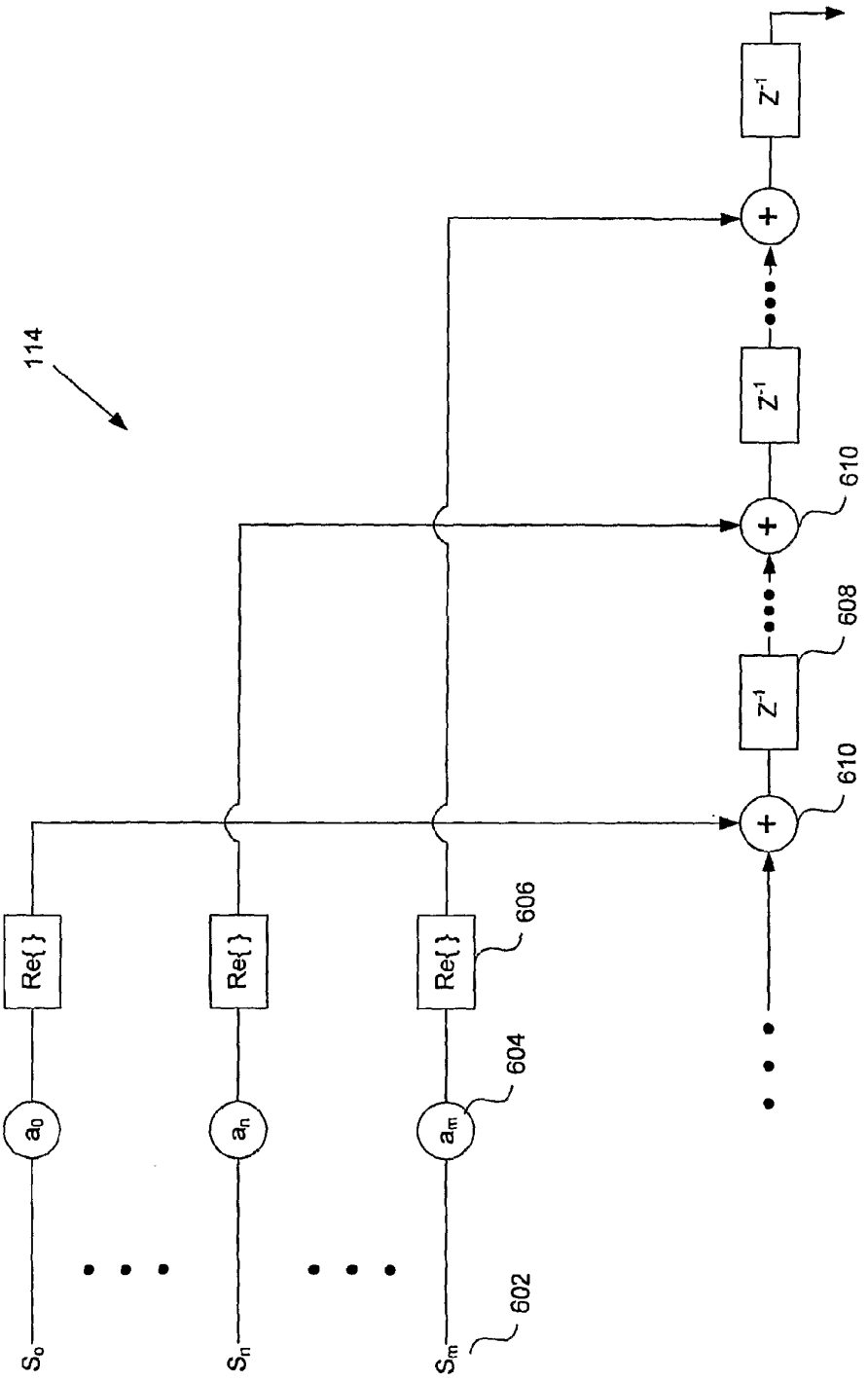


FIG. 6

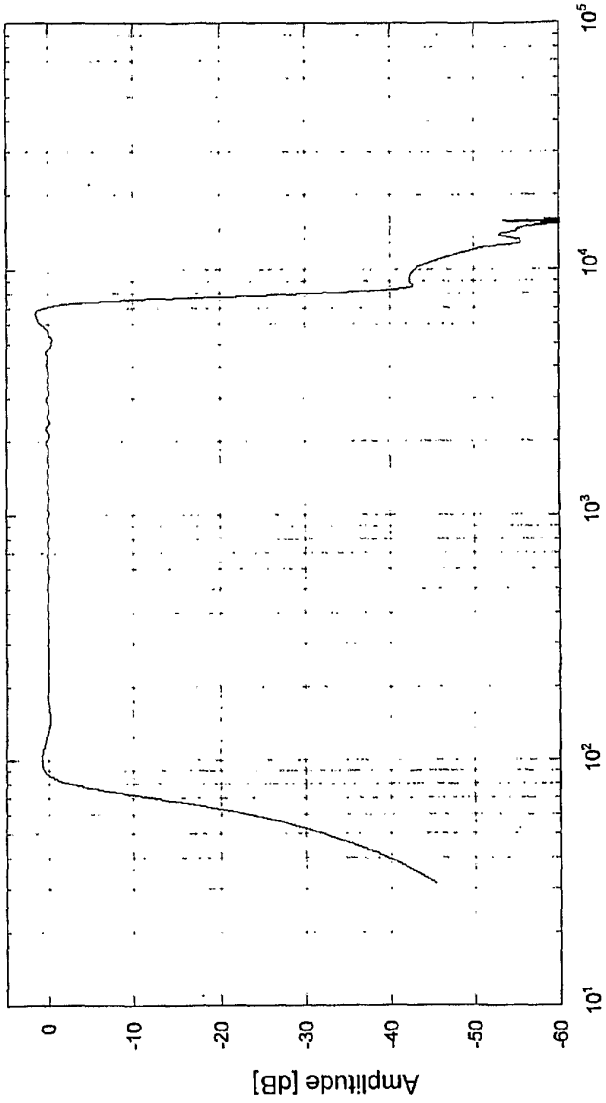

Amplitude [dB]
0
-10
-20
-30
-40
-50
-60
Frequency [Hz]
10^1
10^2
10^3
10^4
10^5

FIG. 7

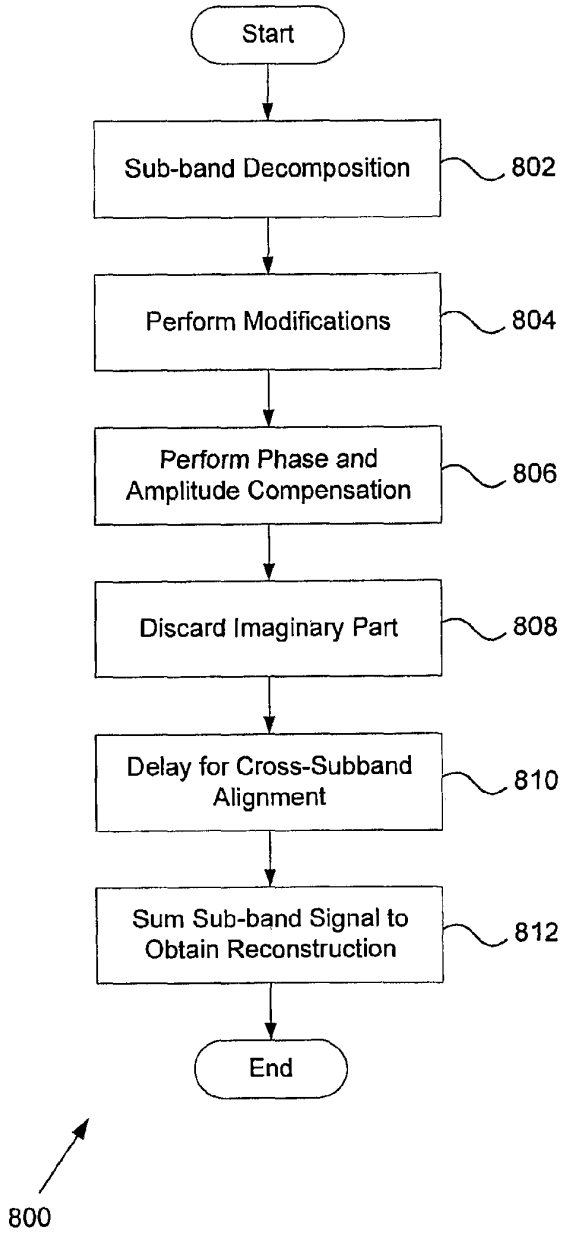

Start
Sub-band Decomposition
802
Perform Modifications
804
Perform Phase and Amplitude Compensation
806
Discard Imaginary Part
808
Delay for Cross-Subband Alignment
810
Sum Sub-band Signal to Obtain Reconstruction
812
End
800

FIG. 8

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 May 2003 (22.05.2003)

PCT

(10) International Publication Number
WO 03/043374 A1

(51) International Patent Classification[7]: H04R 5/00

(21) International Application Number: PCT/US02/36946

(22) International Filing Date:
14 November 2002 (14.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/004,141 14 November 2001 (14.11.2001) US

(71) Applicant: **AUDIENCE, INC.** [US/US]; 4906 El Camino Real, Suite 205, Los Altos, CA 94022 (US).

(72) Inventor: **WATTS, Lloyd**; 500 Chiquita Avenue #14, Mountain View, CA 94041 (US).

(74) Agent: **VAN PELT, Lee**; Van Pelt & Yi LLP, 10050 N. Foothill Blvd, Suite 200, Cupertino, CA 95014-5601 (US).

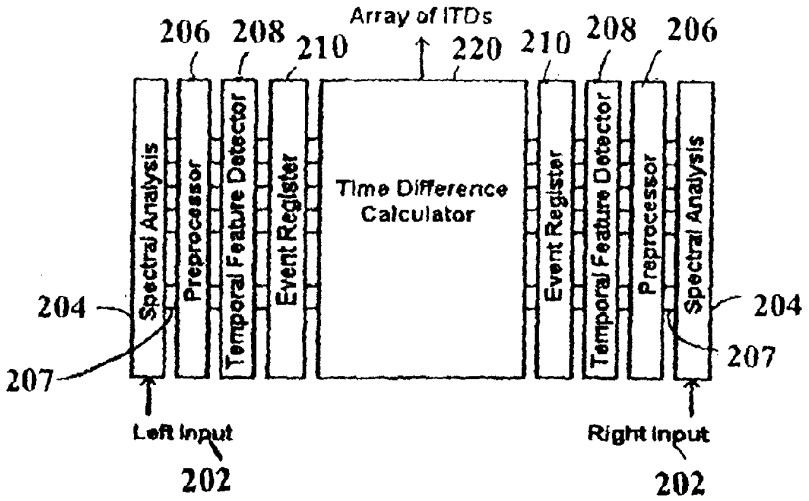
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

[Continued on next page]

(54) Title: COMPUTATION OF MULTI-SENSOR TIME DELAYS



(57) Abstract: Determining a time delay between a first signal received at a first sensor and a second signal received at a second sensor is described. The first signal is analyzed (204) to derive a plurality of first signal channels at different frequencies and the second signal is analyzed (204) to derive a plurality of second signal channels at different frequencies. A first feature detected (208) taht occurs at a first time in one of the first signal channels. A second feature is detected (208) that occurs at a second time in one of the second signal channels. The first feature is matched with the second feature and the first time is compared to the second time (220) to determine the time delay.

WO 03/043374 A1

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

COMPUTATION OF MULTI-SENSOR TIME DELAYS

CROSS REFERENCE TO RELATED APPLICATIONS

This application is related to co-pending United States Patent Application No. 09/534,682 (Attorney Docket No. ANSCP001) by Lloyd Watts filed March 24, 2000 entitled: "EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE" which is herein incorporated by reference for all purposes.

FIELD OF THE INVENTION

The present invention relates generally to sound localization. Calculation of a multisensor time delay is disclosed.

BACKGROUND OF THE INVENTION

For many audio signal processing applications, it is very useful to localize sound. Sound may be localized by precisely measuring the time delay between sound sensors that are separated in space and that both receive the sound. One of the important cues used by humans for localizing the position of a sound source is the Interaural Time Difference (ITD), that is, the difference in time of arrival of sounds at the two ears, which are sound sensors separated in space. ITD is usually computed using the algorithm proposed by Lloyd A. Jeffress in "A Place Theory of Sound Localization," J. Comp. Physiol. Psychol., Vol. 41, pp. 35-39 (1948), which is herein incorporated by reference.

Figure 1 is a diagram illustrating the Jeffress algorithm. Sound is input to a right input 102 and a left input 104. A spectral analysis 106 is performed on the two sound sources, and the outputs from the spectral analysis subsystems are input to cross correlator 108 which includes an array of delay lines and multipliers that produce a set of correlation outputs. This approach is conceptually simple, but

requires many computations to be performed in a practical application. For example, for sound sampled at 44.1 kHz, with a 600-tap spectral analysis, and a 200-tap cross-correlator size (number of delay stages), it would be necessary to perform 5.3 Billion multiplications/second.

In order for sound localization to be practically included in audio signal processing systems that would benefit from it, it is necessary for a more computationally feasible technique to be developed.

SUMMARY OF THE INVENTION

An efficient method for computing the delays between signals received from multiple sensors is described. This provides a basis for determining the positions of multiple signal sources. The disclosed computation is used in sound localization and auditory stream separation systems for audio systems such as telephones, speakerphones, teleconferencing systems, and robots or other devices that require directional hearing.

It should be appreciated that the present invention can be implemented in numerous ways, including as a process, an apparatus, a system, a device, a method, or a computer readable medium such as a computer readable storage medium or a computer network wherein program instructions are sent over optical or electronic communication links. Several inventive embodiments of the present invention are described below.

In one embodiment, determining a time delay between a first signal received at a first sensor and a second signal received at a second sensor includes analyzing the first signal to derive a plurality of first signal channels at different frequencies and analyzing the second signal to derive a plurality of second signal channels at different frequencies. A first feature detected that occurs at a first time in one of the first signal channels. A second feature is detected that occurs at a second time in one of the second signal channels. The first feature is matched with the second feature and the first time is compared to the second time to determine the time delay.

These and other features and advantages of the present invention will be presented in more detail in the following detailed description and the accompanying figures which illustrate by way of example the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

Figure 1 is a diagram illustrating the Jeffress algorithm.

Figure 2 is a diagram illustrating a system for calculating a time delay between two sound sensors that receive a signal.

Figure 3 is a diagram illustrating a sampled signal 300 corresponding to one of the output channels of the spectrum analyzer that is input to the temporal feature detector.

Figure 4A is a graph depicting the output of one frequency channel for each of two sensors.

Figure 4B is a diagram illustrating a sampled signal that contains a plurality of local peaks.

Figure 5 is a flow chart illustrating how a peak event is detected and reported to the event register in one embodiment.

Figure 6 is a diagram illustrating a sample output of the system. The time delay for events measured in each channel is plotted against the frequency of each channel.

Figure 7 is a block diagram illustrating a configuration used in one embodiment where three sensors are used with three time difference calculations being determined.

DETAILED DESCRIPTION

A detailed description of a preferred embodiment of the invention is provided below. While the invention is described in conjunction with that preferred embodiment, it should be understood that the invention is not limited to any one embodiment. On the contrary, the scope of the invention is limited only by the appended claims and the invention encompasses numerous alternatives, modifications and equivalents. For the purpose of example, numerous specific details are set forth in the following description in order to provide a thorough understanding of the present invention. The present invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the present invention is not unnecessarily obscured.

Figure 2 is a diagram illustrating a system for calculating a time delay between two sound sensors that receive a signal. A signal source produces a signal that is received by sensors 202. The signal from each sensor is input to a spectrum analyzer 204. Each spectrum analyzer processes the signal from its inputs and outputs the processed signal which includes a number of separated frequency channels to temporal feature detector. The spectrum analyzer may be implemented in a number of different ways, including Fourier Transform, Fast Fourier Transform, Gammatone filter banks, wavelets, or cochlear models. A particularly useful implementation is the cochlear model described in United States Patent Application No. 09/534,682 (Attorney Docket No. ANSCP001) by Lloyd Watts filed March 24, 2000 entitled: "EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE" which is herein incorporated by reference for all purposes. In one embodiment, a preprocessor 206 is used between the spectrum analyzer and a

temporal feature detector 208. This preprocessor is configured to perform onset emphasis to eliminate or reduce the effect of environmental reverberations. In one embodiment, onset emphasis is achieved by passing the spectrum analyzer channel output through a high-pass filter. In another embodiment, onset emphasis is achieved by providing a delayed inhibition, such that echoes arriving after a certain amount of time are suppressed.

The signal output from the spectrum analyzer is provided on a plurality of taps corresponding to different frequency bands (channels), as shown by the separate lines 207. In one embodiment, a digital filter cascade is used for each spectrum analyzer. Each filter cascade has n filters (not shown) that each respond to different frequencies. Each filter has an output connected to an input of the temporal feature detector or to an input of the preprocessor which in turn processes the signal and passes it to the temporal feature detector. Each filter output corresponds to a different channel output.

The temporal feature detector detects features in the signal and passes them to an event register 210. As the event register receives each event, the event register associates a timestamp with the event, and passes it to time difference calculator 220. When the time difference calculator receives an event from one input, it matches it to an event from the other input, and computes the time difference. This time difference may then be used to determine the position of the signal source, such as in azimuthal position determination systems and sonar (obviating the need for a sonar sweep).

The temporal feature detector greatly reduces the processing required compared to a system that correlates the signals in each of the frequency channels. Feature extraction, time stamping, and time comparison can be far less computationally intensive than correlation which requires a large number of multiplications.

Figure 3 is a diagram illustrating a sampled signal 300 corresponding to one of the output channels of the spectrum analyzer that is input to the temporal feature detector. The temporal feature detector is configured to detect features in the sampled waveform at a corresponding output tap from the spectrum analyzer. Some of the features of signal 300 that may be detected by the feature detector are signal peaks

(local maximums) 302, 304, and 306. In other embodiments, other features such as zero crossings or local minimums are detected.

In detecting a feature, it is desirable to obtain the exact time that a feature occurs, and to accurately obtain a parameter that characterizes or distinguishes the feature. In one embodiment, signal peaks are used as a feature that is detected. The amplitude of each peak may also be measured and used to characterize each peak and to distinguish it from other peaks. A number of known techniques can be used to detect peaks. In one embodiment, peaks are detected by applying the following criteria:

$$x_{n-1} > x_n \text{ and } x_{n-1} > x_{n-2}$$

Figure 4A is a graph depicting the output of one frequency channel for each of two sensors. The channel output shown for Sensor 2 includes three peaks, 402, 404, and 406, each having a different height. Likewise, the channel output for Sensor 1 includes three peaks, 412, 414, and 416 that correspond to the peaks detected by Sensor 2 with a different time delay caused by the physical separation between the two sensors. In one embodiment, each peak is detected using an appropriate method and the time and amplitude of each detected peak event is reported to the event register. Other types of events may be detected and reported as well.

Figure 5 is a flow chart illustrating how a peak event is detected and reported to the event register in one embodiment. The process starts at 600. In step 610, the system receives a sampled point from the channel and assigns a timestamp to it. In step 620, the system determines whether the amplitude of sampled point x_n is less than the amplitude of the preceding sampled point x_{n-1}. In the event that it is, then control is transferred to step 630 and the system determines whether the amplitude of the preceding sampled point x_{n-1} is less than the amplitude of the sampled point x_{n-2} preceding it. In the event that it is, control is transferred to step 640 and x_{n-1} is reported as a peak, along with its timestamp and, optionally, its amplitude .

The peak-finding method described above is appropriate when it is only necessary to obtain nearest-sample accuracy in the timing or the amplitude of the waveform peak. In some situations, however, such as when the spectral analysis step

is performed by a progressively downsampled cochlea model as described in "EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE" by Watts, which was previously incorporated by reference, it may be useful to use a more accurate peak finding method based on quadratic curve fitting. This is done by using the method described above to identify three adjacent samples containing a local maximum, and then fitting a quadratic curve to the three points to determine the temporal position (to subsample accuracy) and the amplitude of the peak.

In some embodiments, the temporal feature detector is configured to detect valleys. A number of techniques are used to detect valleys. In one embodiment, the following criteria is applied:

$$x_{n-1} < x_n \text{ and } x_{n-1} < x_{n-2}$$

The valley event x_{n-1} is reported to the event register just as a peak event is reported. As with a peak event, the amplitude of the valley event may be measured and recorded for the purpose of characterizing the valley event.

The temporal feature detector may also be configured to detect zero crossings. In one embodiment, the temporal feature detector detects positive going zero crossings by looking for the following condition:

$$x_n > 0 \text{ and } x_{n-1} \leq 0$$

The system receives a sampled point x_n and assigns a timestamp to it. If the amplitude of the sampled point x_n is greater than zero, the system further checks whether the amplitude of the preceding sampled point x_{n-1} is less than or equal to zero. If this is the case, the event x_n (or x_{n-1} if equal to zero) may be reported to the event register. Linear interpolation may be used to more exactly determine the time of zero crossing. The time of the zero crossing is reported to the event register, which uses that time to assign a timestamp to the event. If negative going zero crossings are also detected, then the fact that the zero crossing is a positive going one may also be reported to characterize and possibly distinguish the event.

The zero crossing method has the advantage of being able to use straight-line interpolation to find the timing of the zero crossing with good accuracy, because the shape of the waveform tends to be close to linear in the region near the zero. The peak-finding method allows the system to obtain a reasonably accurate estimate of recent amplitude to characterize the event.

Time difference calculator 220 retrieves events from the event registers, matches the events, and computes the multisensor time delay using the timestamps associated with the events. The time delay between individual events may be calculated or a group of events (such as periodically occurring peaks) may be collected into an event set that is compared with a similar event set detected from another sensor for the purpose of determining a time delay. Just as individual events may be identified by a parameter such as amplitude, a set of events may be identified by a parameter such as frequency or a pattern of amplitudes. Referring back to Figure 4A, there is some ambiguity as to whether Sensor 2 leads Sensor 1 by 0.5 ms or whether Sensor 1 leads Sensor 2 by 1.5ms. One method of resolving the ambiguity if the amplitude or some other property of each event is different is to match events between sensors that have similar amplitudes or amplitude patterns.

Other methods of matching events or sets of events may be used. In one embodiment, the envelope function of a group of peaks is determined and the peak of the envelope function is used as a detected event. Figure 4B is a diagram illustrating a sampled signal that contains a plurality of local peaks 452. Local peaks 452 vary in amplitude and the amplitude variation follows a peak envelope function 454. Peak envelope has a maximum at 456, which does not correspond to a local peak, but nevertheless represents a maximum for the envelope function for a collective event that comprises the group of local peaks. In one embodiment, a quadratic method is used to find the local peaks, and then another quadratic method is used to find the peak of the peak envelope function. This method of defining an acoustic event is useful for many acoustic signals where significant events tend to include several cycles. The envelope of the cycles is likely to be the most identifiable event.

In general, this technique is useful for high frequency signals where the period is small compared to the delay. There may be several events in one channel before

the first event arrives at a second channel. An ambiguity occurs as to which events in either channel correspond to each other. However, the envelope of the events varies more slowly and peaks may not occur as often, so that corresponding events occur less frequently and the ambiguity may be solved. In another embodiment, instead of detecting the envelope, the largest event in a group of events is selected. The result is similar, except that the event is detected at a local peak, instead of at the maximum of the envelope function, which does not necessarily occur at a local peak.

In some embodiments, a more computationally expensive alternative to using the peak envelope function is used. The amplitudes of the local peaks are correlated. While this is less computationally expensive than correlating the signals, the peak of the peak envelope function is preferred because it requires less computing and provides good results.

In one embodiment, the time difference calculator, as it matches the events from each sensor, also determines which sensor detects events first. For certain periodic events, determining which sensor first detects an event may be ambiguous. Referring again back to Figure 4A, if the amplitudes do not differ or are not used to distinguish events, then an ambiguity occurs as to which sensor first detects an event. In the example shown, the frequency is 500 Hz, giving a period of 2 ms. Events from the signal received by sensor #1 follow events from the signal received by sensor #2 by 0.5 ms. But events from sensor #2 can also be seen to be following events from sensor #1 by 1.5 ms. This means that either signal #1 leads signal #2 by 1.5 ms, or signal #2 leads signal #1 by 0.5 ms.

In one embodiment, the ambiguity is resolved by defining a maximum multisensor time delay that is possible. Such a maximum may be derived physically from the maximum allowed separation between the sensors. For example, if the sensors are placed to simulate a human head, the maximum multisensor time delay is approximately 0.5 ms, based on the speed of sound in air around a person's head. Thus, the 1.5 ms result is discarded, and the 0.5 ms result (signal 2 leads signal 1) is returned as the time delay. An indication may also be returned that signal 2 leads signal 1.

Figure 6 is a diagram illustrating a sample output of the system. The time delay for events measured in each channel is plotted against the frequency of each channel. Events detected on different channels are grouped and analyzed to determine whether they correspond to a single sound source. The example shown includes event groups 610, 620. 630, and 640. Event group 610 includes events that all occurred at about -0.2ms. The common multisensor time delay calculation across all frequencies indicates that it corresponds to a single sound source. In general, events at different frequency channels that arrive at approximately the same time are likely coming from a common source. The other groups likely represent ambiguities which naturally occur at high frequencies (for which the delay is longer than the period) and can be discarded.

The system groups events detected on different frequency channels and evaluates the groups to determine which groups correspond to a single sound source. In one embodiment, groups are formed by considering all events that are within a set time difference of other events already in the group. For example, all events within a 33 ms video frame are considered in one embodiment.

In one embodiment, an event group is evaluated by computing the standard deviation of the points. If the standard deviation is less than a threshold, then the group is identified as corresponding to a single source and the time delay is recorded for that source. In one embodiment, ratio of the standard deviation to the frequency is compared to a threshold. In some embodiments, the average time delay of the events included in the event group is recorded. In some embodiments, certain events, such as events with a time outside the standard deviation in the group are excluded in calculating time delay for the source.

In other embodiments, other methods are used to automatically decide whether a group of events on different channels correspond to a single source. In one embodiment, the range of the points is compared to a maximum range. In one embodiment, it is determined whether there is a consistent trend in one direction or another with change in frequency. More complex pattern recognition techniques may also be used. In one embodiment, a two dimensional filter that has a strong response when there is a vertical line (see Figure 6) of aligned events occurring at the same

time across different frequencies, with inhibition on either side of the vertical line. This results in accentuation of vertically aligned events such as event 610 and rejection of the sloped/ambiguous events such as events 620,630 and 640.

The multisensor time delay is a strong cue for the azimuthal position of sound sources located around a sensor array. The method of calculating time delay disclosed herein can be applied to numerous systems where a sound source is to be localized. In one embodiment, a sound source is identified and associated with a time delay as described above. The identification is used to identify a person that is speaking based on continuity of position. In one embodiment, a moving person or other sound source of interest is tracked by monitoring ITD measurements over time and using a tracking algorithm to track the source. In one embodiment, a continuity criteria is prescribed for a source that determines whether or not a signal is originating from the source. In one embodiment, the continuity criteria is that the ITD measurement may not change by more than a maximum threshold during a given period of time. In another embodiment where sources are tracked, continuity of the tracked path is used as a continuity criteria. In other embodiments, other continuity criteria are used. The identification may be used to separate the sound source from background sound by filtering out environmental sound that is not emanating from the person or other desired sound source. This may be used in a car (for telematic applications), a conference room, a living room, or other conversational or voice-command situation. In one application used for teleconferencing, a video camera is automatically trained on a speaker by localizing the sound from the speaker using the source identification and time delay techniques described herein.

Figure 7 is a block diagram illustrating a configuration used in one embodiment where three sensors are used with three time difference calculations being determined. The signal from each sensor is processed in a manner as describe above to detect and register the time of events. A time difference calculator is provided to determine the time difference between events detected at each sensor. The three time differences can be used to determine sound source position with greater accuracy. Three or more sensors may be arranged in a planar configuration or multiple sensors may be arranged in multiple planes.

In one embodiment, the time difference is calculated to identify the driver's position as a sound source within an automobile. Sounds not emanating from the driver's position (e.g. the radio, the vent, etc.) are filtered so that background noise can be removed when the driver is speaking voice commands to a cell phone or other voice activated system. The approximate position of the driver's voice may be predetermined and an exact position learned over time as the driver speaks.

Although the foregoing invention has been described in some detail for purposes of clarity of understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims. It should be noted that there are many alternative ways of implementing both the process and apparatus of the present invention. Accordingly, the present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein, but may be modified within the scope and equivalents of the appended claims. For example, embodiments having two and three sensors have been described in detail. In other embodiments, the disclosed techniques are used in connection with more than three sensors and with varying configurations of sensors

WHAT IS CLAIMED IS:

CLAIMS

1. A method of determining a time delay between a first signal received at a first sensor and a second signal received at a second sensor comprising:

analyzing the first signal to derive a plurality of first signal channels at different frequencies;

analyzing the second signal to derive a plurality of second signal channels at different frequencies;

detecting a first feature occurring at a first time in one of the first signal channels;

detecting a second feature occurring at a second time in one of the second signal channels;

matching the first feature with the second feature; and

comparing the first time to the second time to determine the time delay.

2. A method of determining a time delay as recited in claim 1 wherein the first feature includes a signal local maximum.

3. A method of determining a time delay as recited in claim 1 wherein the first feature includes a set of signal local maximums.

4. A method of determining a time delay as recited in claim 1 wherein the first feature includes a signal local minimum.

5. A method of determining a time delay as recited in claim 1 wherein the first feature is a signal local maximum and wherein the amplitude of signal local maximum is used to match the first feature with the second feature.

6. A method of determining a time delay as recited in claim 1 wherein the first feature is a signal local maximum and wherein the signal local maximum is interpolated from sampled points.

7. A method of determining a time delay as recited in claim 1 wherein the first feature is a zero crossing.

8. A method of determining a time delay as recited in claim 1 wherein the first feature is a zero crossing and wherein whether the zero crossing is positive going or negative going is used to match the first feature with the second feature.

9. A method of determining a time delay as recited in claim 1 wherein the first feature is a peak of an envelope function.

10. A method of determining a time delay as recited in claim 1 further including matching a plurality of features in a plurality of channels and determining a plurality of time delays and comparing the time delays to derive a single time delay that corresponds to a single source.

11. A method of determining a time delay as recited in claim 1 further including matching a plurality of features in a plurality of channels and determining a plurality of time delays and comparing the time delays to validate the event time delay.

12. A method of determining a time delay as recited in claim 1 wherein the time delay is used to localize a sound source.

13. A method of determining a time delay as recited in claim 1 wherein the time delay is used to localize a sound source for and a video camera is configured to point at the sound source.

14. A method of determining a time delay as recited in claim 1 wherein the time delay is used to localize a sound source and background sounds not emanating from the source are filtered.

15. A method of determining a time delay as recited in claim 1 wherein the time delay is used to localize a sound source that is moving and the sound source is tracked.

16. A method of determining a time delay as recited in claim 1 wherein the time delay is used to localize a sound source that is moving and the sound source is tracked using a continuity criteria.

17. A method of determining a time delay as recited in claim 1 wherein matching the first feature with the second feature includes imposing a maximum possible time delay to remove ambiguities.

18. A method of determining a time delay as recited in claim 1 wherein the first feature and the second feature are periodic and wherein comparing the first time to the second time to determine the time delay includes imposing a maximum possible time delay to determine which feature preceded the other.

19. A system for determining a time delay between a first signal and a second signal comprising:

a first sensor that receives the first signal;

a second sensor that receives the second signal;

a spectrum analyzer that analyzes the first signal to derive a plurality of first signal channels at different frequencies and that analyzes the second signal to derive a plurality of second signal channels at different frequencies;

a feature detector that detects a first feature occurring at a first time in one of the first signal channels and that detects a second feature occurring at a second time in one of the second signal channels;

an event register that records the occurrence of the first feature and the second feature; and

a time difference calculator that compares the first time to the second time to determine the time delay.

20. A computer program product for determining a time delay between a first signal received at a first sensor and a second signal received at a second sensor, the computer program product being embodied in a computer readable medium and comprising computer instructions for:

analyzing the first signal to derive a plurality of first signal channels at different frequencies;

analyzing the second signal to derive a plurality of second signal channels at different frequencies;

detecting a first feature occurring at a first time in one of the first signal channels;

detecting a second feature occurring at a second time in one of the second signal channels;

matching the first feature with the second feature; and

comparing the first time to the second time to determine the time delay.

1/8

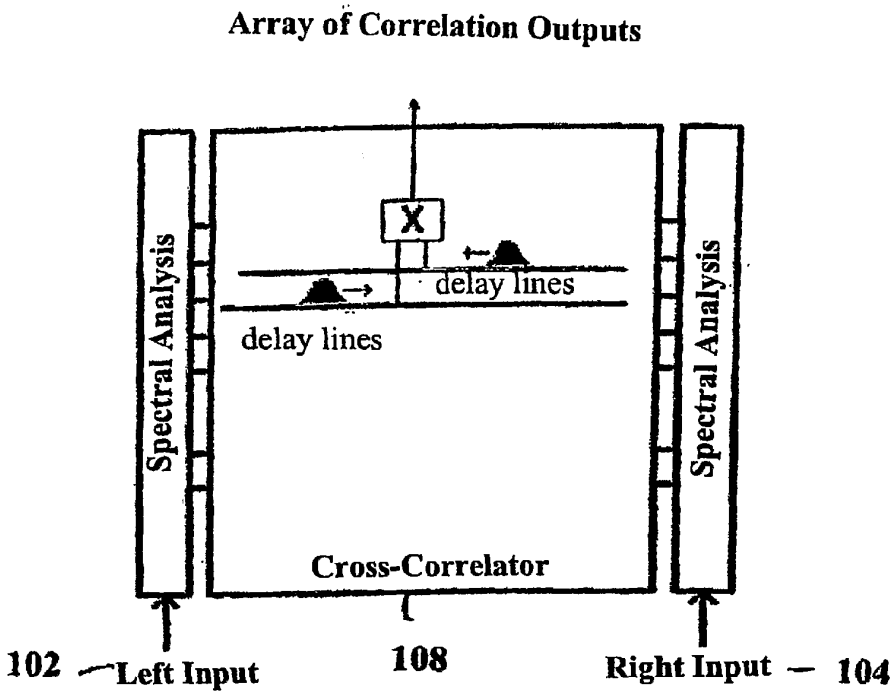

Array of Correlation Outputs
Spectral Analysis
X
delay lines
delay lines
Spectral Analysis
Cross-Correlator
102
Left Input
108
Right Input
104

Figure 1

2/8

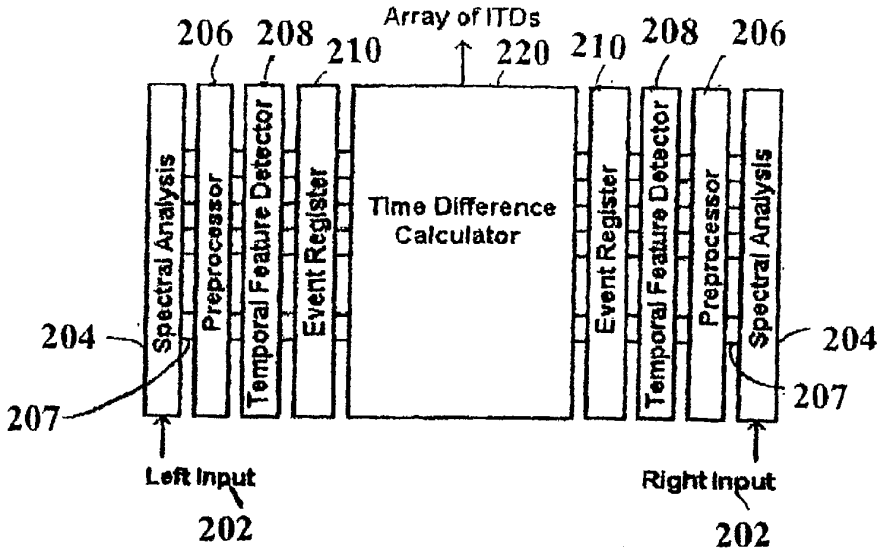

Array of ITDs
206
208
210
220
210
208
206
Spectral Analysis
Preprocessor
Temporal Feature Detector
Event Register
Time Difference Calculator
Event Register
Temporal Feature Detector
Preprocessor
Spectral Analysis
204
207
204
207
Left Input
202
Right Input
202

Figure 2

3/8

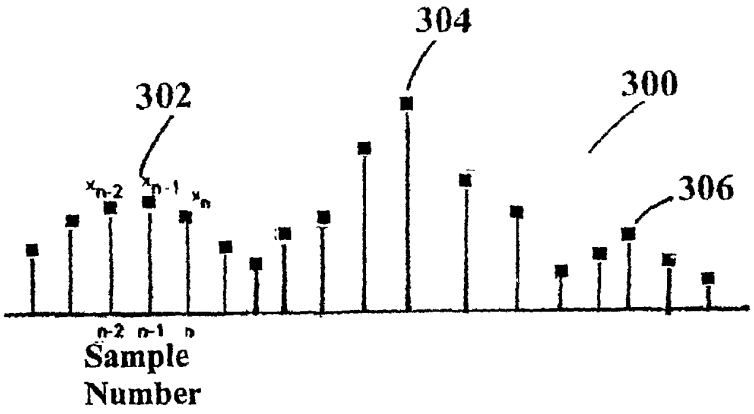


Figure 3

4/8

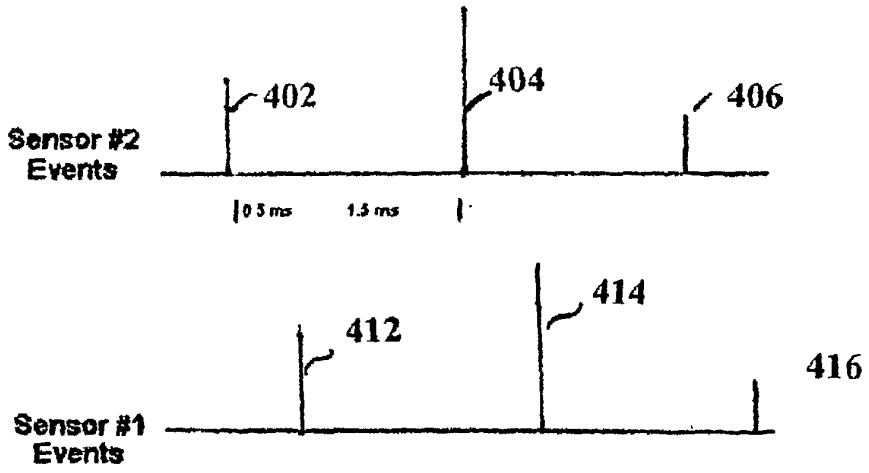

402
404
406
Sensor #2
Events
0.5 ms
1.5 ms
414
412
416
Sensor #1
Events

Figure 4A

5/8

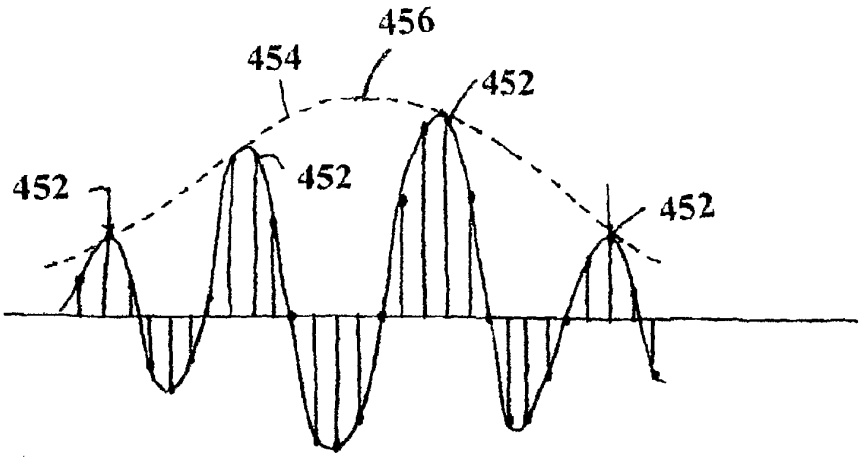


Figure 4B

6/8

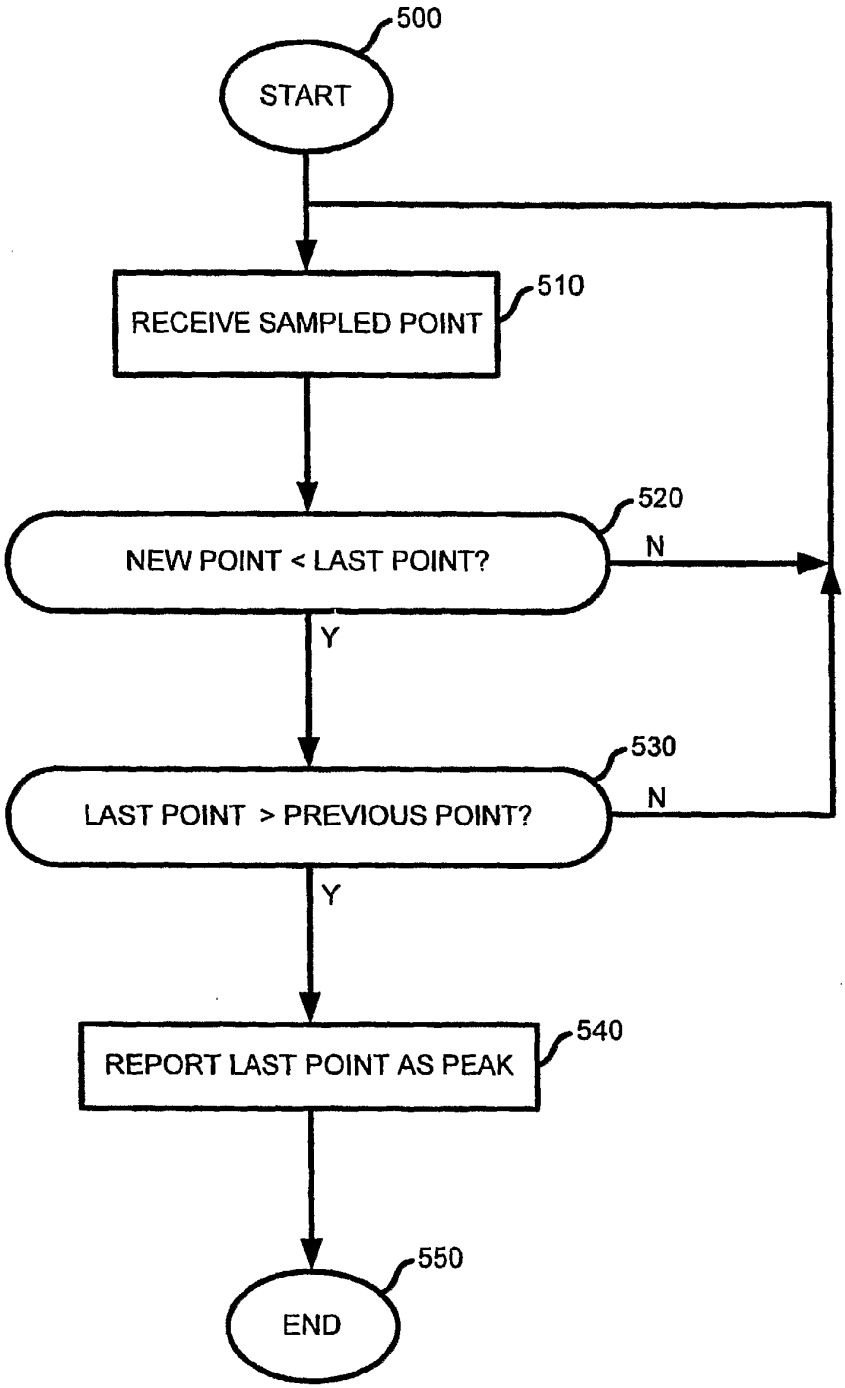

500
START
510
RECEIVE SAMPLED POINT
520
NEW POINT < LAST POINT?
N
Y
530
LAST POINT > PREVIOUS POINT?
N
Y
540
REPORT LAST POINT AS PEAK
550
END

FIG. 5

7/8

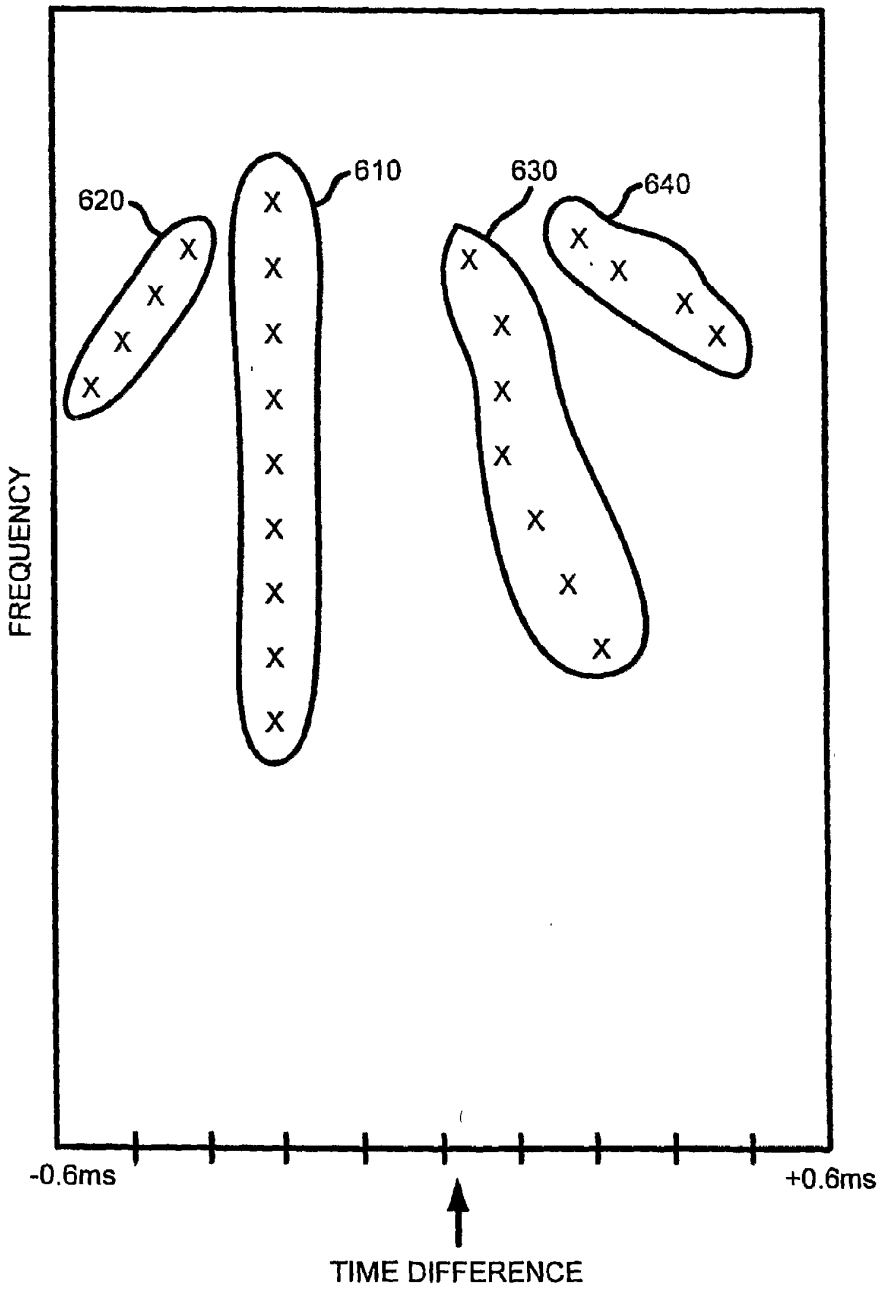


FIG. 6

8/8

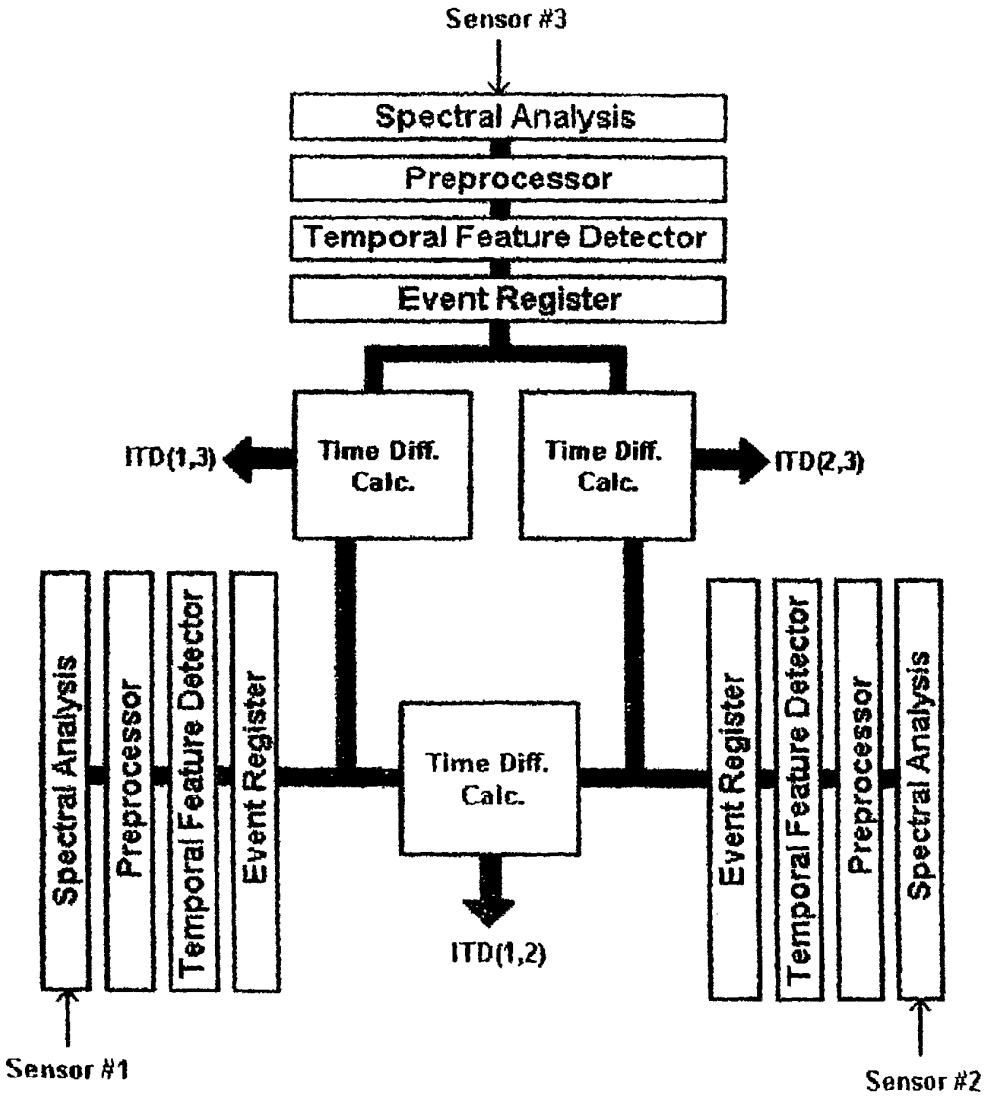

Sensor #3
Spectral Analysis
Preprocessor
Temporal Feature Detector
Event Register
ITD(1,3)
Time Diff. Calc.
Time Diff. Calc.
ITD(2,3)
Spectral Analysis
Preprocessor
Temporal Feature Detector
Event Register
Time Diff. Calc.
ITD(1,2)
Event Register
Temporal Feature Detector
Preprocessor
Spectral Analysis
Sensor #1
Sensor #2

Figure 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/36946

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : H04R 5/00
US CL : 381/17
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 381/17,1,309,310,311,300; 379/406.01

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,223,090 B1 (BRUNGART) 24 April 2001 (24.04.2001), figures 3, 6, 7	1-20
A	US 5,729,612 A (ABEL et al) 17 March 1998 (17.03.1998), see entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
12 February 2003 (12.02.2003)

Date of mailing of the international search report
03 APR 2003

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Forester W. Isen
Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 October 2001 (04.10.2001)

PCT

(10) International Publication Number
WO 01/74118 A1

(51) International Patent Classification[7]: H04R 25/00

(21) International Application Number: PCT/US01/08372

(22) International Filing Date: 15 March 2001 (15.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/534,682 24 March 2000 (24.03.2000) US

(71) Applicant: **APPLIED NEUROSYSTEMS CORPORATION** [US/US]; Suite 205, 4906 El Camino Real, Los Altos, CA 94022 (US).

(72) Inventor: **WATTS, Lloyd**; 500 Chiquita Avenue #14, Mountain View, CA 94041 (US).

(74) Agent: **VAN PELT, Lee**; Van Pelt & Yi, LLP, Suite 205, 4906 El Camino Real, Los Altos, CA 94022 (US).

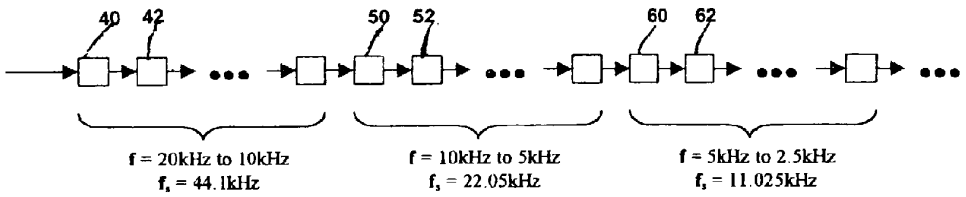
(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE



(57) Abstract: A system for processing audio signals comprises a sequence of digital filters (40, 42) each configured to process a selected frequency using a set of coefficients. A filter configured to process a certain frequency shares its coefficients with another filter that processes a frequency that is lower than the first frequency by at least one frequency interval, such as octave. The first filter samples at a certain sampling rate, and the second filter's sampling rate is determined by multiplying the first sampling rate by the ratio of the second frequency to the first frequency. The filters are evenly grouped into frequency intervals, such as octaves. Filters in an octave are sampled at a sampling frequency that is at least twice as high as the highest frequency processed in that octave.

WO 01/74118 A1

EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE

FIELD OF THE INVENTION

This invention relates generally to a method, article of manufacture, and apparatus for computing the response of a cascade of digital filters in an efficient manner that provides for high resolution while reducing computational expense and storage requirements. More particularly, this invention relates to modeling a cochlea for real-time processing of acoustic signals using an improved digital filter bank cascade.

BACKGROUND

Much effort has been devoted to modeling hearing, for applications such as automatic speech recognition, noise cancellation, hearing aids, and music. A popular approach is to model the cochlea, a coiled snail-shaped structure that is part of the inner ear as shown in FIG. 1. The cochlea is a spiraling, fluid-filled tunnel embedded in the temporal bone, and converts acoustic signals into electrical signals transmitted to the brain. Sound pressure waves strike the eardrum, causing it to move inward and moving the three small bones of the middle ear, which are the hammer, anvil, and stirrup. The movement of the bones initiates pressure waves in the cochlear fluid. These pressure waves propagate along the cochlear partition, which, as shown in FIG. 2, consists of the basilar membrane BM, tectorial membrane TM, and organ of Corti OC. The organ of Corti OC is a collection of cells, including the sensory hair cells, that sit on the basilar membrane BM. The bases (bottoms) of

these hair cells are connected to nerve fibers NF from the auditory nerve AN, and the apexes (tops) of the hair cells have hair bundles HB. There are two types of hair cells in the cochlea: inner hair cells IHC and outer hair cells OHC.

The human cochlea is believed to contain approximately 4,000 inner hair cells IHC and 12,000 outer hair cells OHC, with four cells radially abreast and spaced every 10 microns along the length of the basilar membrane BM. The tectorial membrane TM lies on top of the surface of the organ of Corti OC. A thin fluid space of about 4 to 6 microns lies between these two surfaces, which shear as the basilar membrane BM moves up and down. The hair cells are primarily transducers that convert displacement of the hair bundle HB (due to shearing between the tectorial membrane TM and the surface of the organ of Corti) into a change in the receptor current flowing through the cell, which is transmitted to the auditory nerve AN and processed by the brain.

Each point on the basilar membrane BM is tuned to a different frequency, with a spatial gradient of about 0.2 octaves/mm for a human, and about 0.32 octaves/mm for a cat. Roughly speaking, the cochlea acts like a bank of filters. The filtering allows the separation of various frequency components of the signal with a good signal-to-noise ratio. The range of audible frequencies is about 20 Hz to 16 kHz in the human cochlea and about 100 Hz to 40kHz in the cat cochlea.

Modeling the function of the cochlea has been an active area of research for many years. For example, U.S. Patent No. 4,771,196, titled "Electronically variable active analog delay line" and issued to Mead and Lyon on September 13, 1988, describes an analog filter

bank cascade for signal processing. This patent, the disclosure of which is hereby incorporated by reference, illustrates an electronically variable active analog delay line that incorporates cascaded differential transconductance amplifiers with integrating capacitors and negative feedback from the output to the input of each noninverting amplifier. "Lyon's Cochlear Model", written in 1988 by Malcolm Slaney as Apple Technical Report #13, describes a digital filter bank cascade developed by Lyon as a model of the cochlea. Further details of the Lyon model may be seen by reference to the technical report, the disclosure of which is hereby incorporated by reference.

This model uses a cascade of second-order filters, each of which requires a number of computations every time the signal is sampled. Each filter has a set of coefficients associated with it, and must also store some previous computations. If the sampling rate is increased, or the number of filters is increased in order to increase resolution, the number of computations rises proportionally. Thus, the desire for better resolution and sampling of the acoustic signal is balanced against the computations required and the storage needed for each filter. A more efficient approach, such as the approach of the present invention, would reduce the computation required for the cascade and allow for a higher quality representation of the signal.

This problem is not limited to digitized signals represented by discrete amplitude levels, nor is it limited to acoustic signals. Rather, it applies to any sampled signal (represented by discrete time values). Although the disclosure herein describes the problem and the invention in the context of audio signal processing, one skilled in the art will

recognize that the invention may be applied to any signal processing using sampling, including electrical waveform sampling and video signal processing.

SUMMARY OF THE INVENTION

It should be appreciated that the present invention can be implemented in numerous ways, including as a process, an apparatus, a system, a device, a method, or a computer readable medium such as a computer readable storage medium or a computer network wherein program instructions are sent over optical or electronic communication links. Several inventive embodiments of the present invention are described below.

Briefly, therefore, this invention provides for a method, article of manufacture, and apparatus for real-time processing of signals. In an embodiment of the invention, a system for processing audio signals comprises a sequence of digital filters each configured to process a selected frequency using a set of coefficients. A filter configured to process a certain frequency shares its coefficients with another filter that processes a frequency that is lower than the first frequency by at least one frequency interval, such as an octave. The first filter samples at a certain sampling rate, and the second filter's sampling rate is determined by multiplying the first sampling rate by the ratio of the second frequency to the first frequency. The filters are evenly grouped into frequency intervals, such as octaves. Filters in an octave are sampled at a sampling frequency that is at least twice as high as the highest frequency processed in that octave.

The advantages and further details of the present invention will become apparent to one skilled in the art from the following detailed description when taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

FIG. 1 is a sectional view of the inner and outer ear of a human ear;

FIG. 2 is a sectional view of the inner ear of a human ear;

FIG. 3 is a schematic of a signal processing system in accordance with the invention;

FIG. 4 depicts the structure of the cochlear model in a serial filter bank cascade configuration in accordance with the invention;

FIG. 5 is a signal flow graph of a filter equation in accordance with the invention; and

FIG. 6 is a schematic of the filter bank cascade showing its division into octaves and the use of downsampling.

DESCRIPTION OF THE INVENTION

Overview

A signal processing system in accordance with the invention comprises a computer configured with a cascade of digital filters arranged sequentially on a logarithmic frequency scale, through which a signal is passed. The filters are configured to process certain frequencies and are programmed with filter coefficients appropriate to the desired filter

behaviors and frequencies processed. Each successive filter in the sequence is configured to process a lower frequency than the one before it. Each filter also has a tap associated with it for extracting the filter output, and the number of filters and taps is determined by the desired resolution and frequency range. The filters are grouped into octaves, and within an octave group, a sampling rate is used that meets the Nyquist sampling criterion for the highest frequency filter in the octave. The filters in the highest octave use the same filter coefficients as filters in the lower octaves, with each successively lower octave group using a successively lower sampling rate to produce the lower frequency filters. Since the filters in each octave group remove the highest frequencies in the signal, the sampling rate can be reduced between octaves without violating the Nyquist sampling criterion.

In another embodiment of the invention, a filter can be used to process a certain frequency at a certain sampling rate and reused to process other frequencies that are one, two, or more octaves higher or lower, with a corresponding adjustment to the sampling frequency based on the highest frequency in the octave of target frequency. Another filter can be used to process another frequency in the same octave, and be reused to process other frequencies that are one, two, or more octaves higher or lower. In this manner, an array of filters covering a single octave can be used to process signals spanning multiple octaves.

In a further embodiment of the invention, the efficient digital filter bank cascade can be used as a model of a cochlea to process acoustic signals with improved accuracy and resolution, and more efficient use of computational and storage resources.

The response of this cascade of digital filters is thus computed in an efficient manner that provides for high resolution while reducing computational expense and storage requirements.

A detailed description of a preferred embodiment of the invention is provided below. While the invention is described in conjunction with that preferred embodiment, it should be understood that the invention is not limited to any one embodiment. On the contrary, the scope of the invention is limited only by the appended claims and the invention encompasses numerous alternatives, modifications, and equivalents. For the purpose of example, numerous specific details are set forth in the following description in order to provide a thorough understanding of the present invention. The present invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the present invention is not unnecessarily obscured.

Detailed Description

In accordance with the invention, a signal processing system comprises a computer configured to analyze signals, such as acoustic or audio signals. In an embodiment of the invention, the signal processing system is in the form of a software program being executed on a general-purpose computer such as an Intel Pentium-based PC running a Windows or Linux operating system, or a workstation running Unix. Other means of implementing the signal processing system may be used, such as a special-purpose hardwired system with instructions burned into a chip such as an application specific integrated circuit (ASIC) or field-programmable gate array (FPGA). As is usual in the industry, the computer (CPU) 10

may have memory 12, a display 14, a keyboard 16, a mass storage device 18, a network interface 20, and other input or output devices 22, shown in FIG. 3. Also shown in FIG. 3 is a signal input device 24 in the form of a microphone, though other types of signal input devices may be used. In accordance with common practice, the memory 12 and the mass storage device 18 can be used to store program instructions and data. The computer may further have more than one central processing unit, such as a multiprocessor Pentium-based system or Sun SPARCstation.

It will be readily apparent to one skilled in the art that more than one computer may be used, such as by using multiple computers in a parallel or load-sharing arrangement or distributing tasks across multiple computer such that, as a whole, they perform the functions of the signal processing system; i.e. they take the place of a single computer. It is intended that the disclosure cover all such configurations as if fully set forth herein.

A signal processing system in accordance with the invention comprises a computer configured with program describing a cascade of digital filters arranged sequentially on a logarithmic frequency scale, through which a signal is passed. The filters are configured to process certain frequencies and are programmed with filter coefficients appropriate to the desired filter behaviors and frequencies processed. Each successive filter in the sequence is configured to process a lower frequency than the one before it. Each filter also has a tap associated with it, and the number of filters and taps is determined by the desired resolution and frequency range. A filter is used to process a signal of a certain frequency at a certain sampling rate, and shares its filter coefficients with filters configured to process signals of frequencies that are one, two, or more octaves lower. The filter also attenuates its target

frequency and passes the signal on to the next filter in the sequence. For each successive filter in the sequence, the sampling rate may be reduced in proportion to the reduction in its target frequency. For convenience, the filters may be grouped into octaves, and each filter in an octave will be sampled at a rate that meets the Nyquist sampling criterion for the highest frequency filter in the octave. Lower octaves will be sampled at successively lower rates.

In another embodiment of the invention, a filter can be used to process a certain frequency at a certain sampling rate and reused to process other frequencies that are one, two, or more octaves higher or lower, with a corresponding adjustment to the sampling frequency based on the target frequency, in accordance with the Nyquist sampling criterion. Another filter can be used to process another frequency in the same octave, and be reused to process other frequencies that are one, two, or more octaves higher or lower. In this manner, an array of filters covering a single octave can be used to process signals spanning multiple octaves. Similarly to the above embodiment, the sampling rate can be reduced as the octave of frequencies being sampled is lowered.

The invention will be illustrated by its use in audio signal processing, utilizing a model of the cochlea. This model describes the propagation of sound in the inner ear and the conversion of acoustic signals into neural signals. It combines a series of filters that model the traveling pressure waves with half-wave rectifiers to detect the energy in the signal and several stages of automatic gain control, as shown in FIG. 4. Sound pressure waves cause displacement of the hair cells and generation of neural signals as described above. This is modeled by the filters, which, like the hair cells, are tuned to specific frequencies. The basilar membrane is attuned to high frequency sounds near the base of the cochlea, where the

sound enters, and senses progressively lower and lower frequencies as the sound pressure wave travels through the cochlea. The filters in the model are arranged similarly, with each filter attuned to a higher frequency than succeeding filters, so that the signal is gradually low-pass filtered.

In this model, the audio signal acquired from the signal input device 24 undergoes some preprocessing, and is then passed through a cascade of sequentially arranged filters 30 to model the propagation of the sound pressure waves through the cochlea, from left to right in the diagram of FIG. 4. Each filter 30 in the cascade has an output that feeds into the input of the next filter 30 in the cascade (if one is present), and a tap that allows data to be extracted from the filter 30, which in this embodiment is the data provided to the filter output. The tap has several stages of processing associated with it, such as a half-wave rectifier 32 and automatic gain control 34. Each filter is attuned to a particular frequency, and has a set of coefficients $(a_0, a_1, a_2, b_1, b_2)$ associated with it. The output of each filter is calculated according to the following function:

Equation 1. $$y_n = a_0x_n + a_1x_{n-1} + a_2x_{n-2} - b_1y_{n-1} - b_2y_{n-2}$$

where the filter output y_n is a function of the input data x_n at time n, previous inputs x_{n-1} and x_{n-2}, and previous outputs y_{n-1} and y_{n-2}. This formula is illustrated by the signal flow graph in FIG. 5. The output of the filter y_n is passed to the input x_n of the next filter in the cascade.

The filter response $H(z)$ is given by the following:

$$\textbf{Equation 2.} \qquad H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2}}$$

and

$$z = e^{i * (\omega/\omega s)}, \quad \omega = 2\pi f, \qquad \omega_s = 2\pi f_s$$

where f_s is the sampling frequency.

Substitution of the above into the transfer function of Equation 2 produces a filter response **H(f)**, which is a function of the filter coefficients a_0, a_1, a_2, b_1, b_2 and the sampling rate f_s.

In this audio signal processing embodiment, the frequency range typically used is 20 Hz to 20 kHz, since that is roughly the range of human hearing. With about 4,000 inner hair cells, a human has the equivalent of 4,000 taps spread over ten octaves, or about 400 taps per octave.

The Nyquist Theorem states that when an analog waveform is digitized, only the frequencies in the waveform below half the sampling frequency will be recorded. In order to accurately represent the original waveform, sufficient samples must be recorded to capture the peaks and troughs of the original waveform. If a waveform is sampled at less than its Nyquist frequency (which is twice the frequency of the waveform), the reconstructed

waveform will represent low frequencies not present in the original signal. This phenomenon is called "aliasing", and the high frequencies are said to be "under an alias".

Thus, since the highest frequency is 20 kHz, the Nyquist frequency is 40 kHz. The standard sampling rate for CD (compact disc) audio is slightly higher, at 44.1 kHz. A brute force approach would be to represent all 4,000 inner hair cells as 4,000 filters. Equation 1 shows that there are five multiplication operations and four addition operations per filter per sample, for a total of nine operations per filter sample. Thus, a complete representation of a human ear would require

44,100 * 4,000 * 9 = 1,587,600,000 operations per second

Such a large number of operations would make computation of the cochlear model impractical on all but the fastest computers. Digital signal processing chips typically have a multiply-accumulate instruction, and can perform one addition and multiplication as a single unit of computation. The number of computations required for a DSP to compute the formula of Equation 1 would drop to five, lowering the requirement to 882,000,000 operations per second. This is still impractical. Typically, the number of filters is reduced to a much more manageable size, on the order of 125 filters total, covering six octaves, with a sample rate of 16 kHz. Computation of this model requires

16,000 * 125 * 9 = 18,000,000 operations per second

In a DSP, this would be 10,000,000 operations per second. The number of computations required is about two orders of magnitude smaller, but significant degradation of the sampled waveform has taken place. The frequency range has been reduced by four octaves, and the filter density within the range covered is lower, which reduces the resolution. Further, with

the sampling frequency reduced to 16 kHz, the maximum frequency that can accurately be represented by the sampled waveform is now 8 kHz.

Increasing the number of filters to 600 and covering 10 octaves, as well as increasing the sampling frequency to 44.1 kHz results in significant improvement in resolution, and the frequency range covered now more closely approximates that of human hearing. This would require

44,100 * 600 * 9 = 238,140,000 operations per second

or 132,300,000 operations per second in a DSP. This is better than the original number for a complete model, but is still very computationally expensive.

In accordance with the invention, the filters are evenly distributed over the octaves, resulting in 60 filters per octave. In one embodiment, 60 objects are created in a computer. Each object has a set of coefficients as described above, and additionally has ten sets of state variables, corresponding to ten filters running at frequencies that are whole octaves apart. The 60 objects using their first sets of state variables correspond to the first octave group of filters, while the 60 objects using their second sets of state variables (and sampling at a lower frequency) correspond to the second octave group of filters, and so on. In another embodiment, each object contains a set of coefficients, but only one set of state variables, and is run at a single frequency. In this case, 600 objects are required to represent 600 filters.

The filters in the first octave are tuned to the frequencies in the highest octave, 20 kHz to 10 kHz, and are sampled at 44.1 kHz, which satisfies the Nyquist sampling criterion. The filters in the second octave are tuned to half of the frequencies of the corresponding filters in

the first octave, and range from 10 kHz to 5 kHz. These filters in the second octave are sampled at 22.05 kHz, half of the first sampling frequency. Coefficients for each filter are stored in memory and applied in the computations for the filters. As the audio signal is passed through each filter, the signal is sampled and filtered before being passed to the next filter. FIG. 6 shows the arrangement of the filters. At the end of the first octave, the signal is passed into the first filter in the next octave, which comprises filters sampling at half the sampling rate of the first octave, as stated above. Successive octaves are downsampled in a similar manner. The computational requirement for the digital filter bank of the invention would be

$$44{,}100 * 60 * 9 * (1 + {}^{1}/_{2} + {}^{1}/_{4} + {}^{1}/_{8} + {}^{1}/_{16} + {}^{1}/_{32} + {}^{1}/_{64} + {}^{1}/_{128} + {}^{1}/_{256} + {}^{1}/_{512}) \cong$$

47,581,488 operations per second, or 26,434,160 operations per second in a DSP. Thus, downsampling each successive octave has resulted in a fivefold reduction in computational requirements. This is a nontrivial improvement, particularly in the area of embedded signal processing chips, where performance, size, and cost are primary considerations.

For a given set of filter parameters $(\mathbf{a_0}, \mathbf{a_1}, \mathbf{a2}, \mathbf{b_1}, \mathbf{b_2})$ at a particular sampling rate $\mathbf{f_s}$, the second-order filter will have some resonant frequency $\mathbf{f_r}$. If the filter parameters are kept constant while the sampling rate $\mathbf{f_s}$ is divided by two, the resonant frequency $\mathbf{f_r}$ will also be divided by two, because the transfer function depends on $\mathbf{z}$, which is a normalized frequency variable; i.e. it is normalized by the sampling rate $\mathbf{f_s}$. Thus, scaling the sampling frequency scales the frequency response of the filter by the same amount. In this manner, the filter can be tuned to a frequency that is an octave lower, by sampling at half the original sampling rate without changing the filter coefficients. Downsampling again in this manner produces a filter that runs at yet another octave lower, so long as high frequencies are filtered out before

downsampling. The sampling frequency does not necessarily have to be divided by two, four, or other multiples of two, nor do the filter frequencies have to be grouped by octaves. Any scaling factor may be used, such as ten (resulting in shifts by decades rather than octaves) or other number (resulting in shifts by a corresponding interval on a logarithmic scale), which does not have to be a whole number.

Thus, in the configuration depicted in FIG. 6, any given filter shares filter parameters with filters that are one, two, or more octaves higher or lower. For example, the highest frequency filter 40 in the first octave shares filter coefficients with the highest frequency filter 50 in the second octave, the highest frequency filter 60 in the third octave, and so on. The second-highest frequency filter 42 in the first octave shares filter coefficients with the second-highest frequency filters 52 and 62 in the second and third octaves, and with all other corresponding filters (tuned to frequencies that are one, two, or more octaves lower). It will be apparent that "corresponding" refers to filters that occupy the same relative positions in their respective octaves.

In effect, the filters 40, 50, 60, and other filters in corresponding positions in other octaves are the same filter. Similarly, filters 42, 52, 62, and corresponding filters are the same filter, as are all groups of filters that differ in frequency by whole octaves. A single filter can be used to sample a target frequency, and other target frequencies that are one, two, or more octaves lower, with reduction of the sampling frequency as described above, as long as the Nyquist criterion of removing higher frequencies is observed.

This reduces storage requirements for filter coefficients, because only one set of filter coefficients (for one octave) needs to be stored. Successive octaves may reuse the filter coefficients in accordance with the invention. Another advantage of the invention is that the required precision for filter coefficients is lower, and thus, fewer bits are required to represent each coefficient. In the prior art approach, 20 bits were required for acceptable results, particularly for the low-frequency filter coefficients. The inventive digital filter bank cascade requires about 12 bits to maintain an acceptable level of stability.

The advantage of reducing precision of the filter coefficients is not limited to storage. The reduced number of bits in the operands means that the processing hardware can be made smaller. For example, the arithmetic logic unit can be made smaller, since it does not need to process as many bits, and buses can be made narrower. Further advantages of reduced precision requirements will be readily apparent to one skilled in the art, as will other advantages of the invention.

The foregoing disclosure and embodiment demonstrate the utility of the present invention in dramatically increasing the efficiency of computing digital filter bank cascades for purposes such as audio signal processing, although it will be apparent that the present invention will be beneficial for many other uses.

All references cited herein are intended to be incorporated by reference. Although the present invention has been described above in terms of specific embodiments, it is anticipated that alterations and modifications to this invention will no doubt become apparent to those skilled in the art and may be practiced within the scope and equivalents of the appended

claims. For example, one skilled in the art will recognize that the filters do not necessarily need to be evenly distributed over the octaves, or that the filters do not necessarily need to be used with an audio signal. The present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein. It is therefore intended that the following claims be interpreted as covering all such alterations and modifications as fall within the true spirit and scope of the invention.

CLAIMS

What is claimed is:

1. A system for processing audio signals, comprising a sequence of digital filters, wherein each filter is configured to process a selected frequency and at least one filter is configured to process more than one frequency.

2. The system as recited in claim 1, wherein the at least one filter includes coefficients for processing, and the coefficients are used to process more than one frequency.

3. The system as recited in claim 2, wherein the at least one filter is configured to process a first frequency and a second frequency that is at least one interval away from the first frequency.

4. The system as recited in claim 3, wherein the interval is an octave.

5. The system as recited in claim 4, wherein the at least one filter is configured to sample the first frequency at a first sampling rate and the second frequency at a second sampling rate.

6. The system as recited in claim 5, wherein the second frequency is lower than the first frequency and the second sampling rate is lower than the first sampling rate.

7. The system as recited in claim 6, wherein the second sampling rate is lower than the first sampling rate by two raised to the number of octaves spacing between the first frequency and the second frequency.

8. The system as recited in claim 7, wherein the sequence of digital filters is configured to process frequencies in a first octave at the first sampling rate.

9. The system as recited in claim 8, wherein the sequence of digital filters is further configured to process frequencies in a second octave at the second sampling rate.

10. The system as recited in claim 9, wherein each coefficient is represented by fewer than 13 bits.

11. The system as recited in claim 10,wherein each coefficient is represented by 12 bits.

12. A system for processing audio signals, comprising a sequence of digital filters each configured to process a selected frequency, wherein each filter includes coefficients for processing and a first filter configured to process a first frequency shares its coefficients with a second filter configured to process a second frequency.

13. The system as recited in claim 12, wherein the second frequency is spaced apart from the first frequency by at least one frequency interval.

14. The system as recited in claim 13, wherein the second frequency is spaced apart from the first frequency by at least one octave.

15. The system as recited in claim 14, wherein the first filter is configured to sample the first frequency at a first sampling frequency and the second filter is configured to sample a second frequency at a second sampling frequency.

16. The system as recited in claim 15, wherein the second frequency is lower than the first frequency, and the second sampling frequency is lower than the first sampling frequency by a ratio of the first frequency to the second frequency.

17. The system as recited in claim 14, wherein the filters are evenly grouped into at least a first and a second octave, the first filter being in the first octave and the second filter being in the second octave.

18. The system as recited in claim 17, wherein the filters in the first octave are sampled at a first sampling frequency that is at least twice as high as a highest frequency processed by the first octave.

19. The system as recited in claim 18, wherein the second octave is one octave lower than the first octave, and the filters in the second octave are sampled at a second sampling rate that is half as high as the first sampling frequency.

20. The system as recited in claim 19, wherein each filter in the first octave shares its coefficients with each filter in a corresponding position in the second octave.

21. A method for processing an audio signal, comprising the steps of:

(a) providing a sequence of digital filters each configured to process a selected frequency;

(b) providing each filter with coefficients for processing its selected frequency such that a first filter configured to process a first frequency shares its coefficients with a second filter configured to process a second frequency; and

(c) applying the audio signal to the sequence of digital filters.

22. A computer program product for processing an audio signal, comprising a computer usable medium having machine readable code embodied therein for performing the steps of:

(a) providing a sequence of digital filters each configured to process a selected frequency;

(b) providing each filter with coefficients for processing its selected frequency such that a first filter configured to process a first frequency shares its coefficients with a second filter configured to process a second frequency; and

(c) applying the audio signal to the sequence of digital filters.

1/4

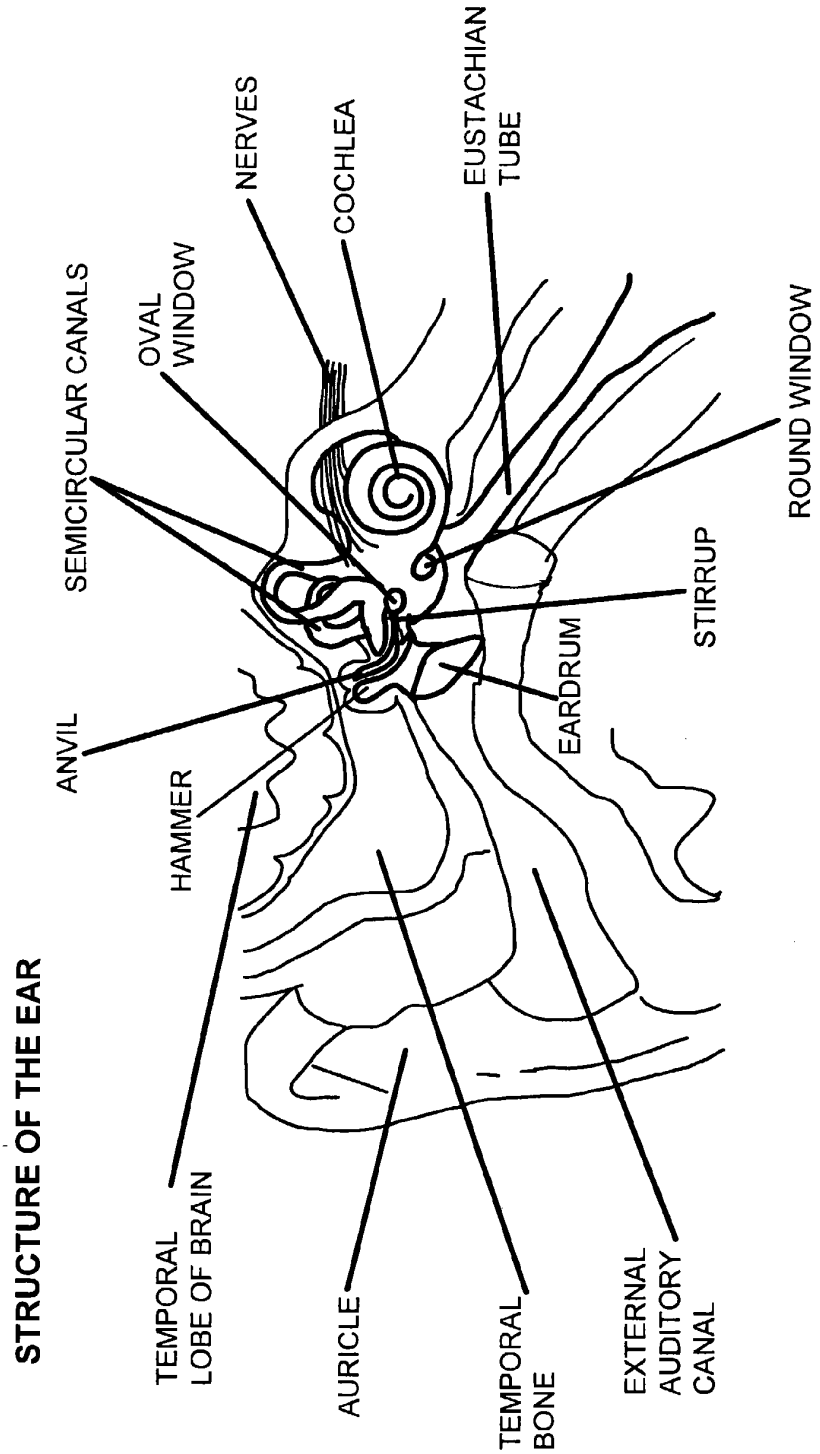

STRUCTURE OF THE EAR
TEMPORAL LOBE OF BRAIN
HAMMER
ANVIL
SEMICIRCULAR CANALS
OVAL WINDOW
NERVES
COCHLEA
EUSTACHIAN TUBE
ROUND WINDOW
STIRRUP
EARDRUM
AURICLE
TEMPORAL BONE
EXTERNAL AUDITORY CANAL

FIG. 1

2/4

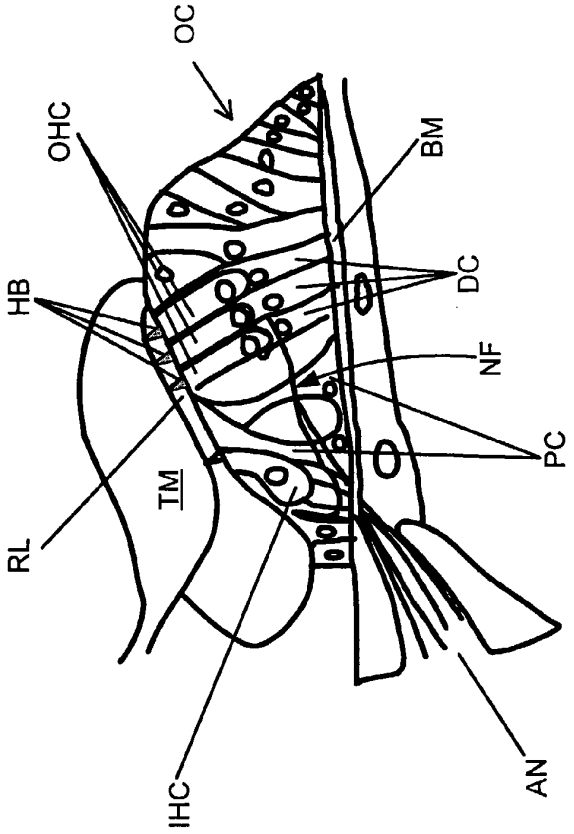


FIG. 2

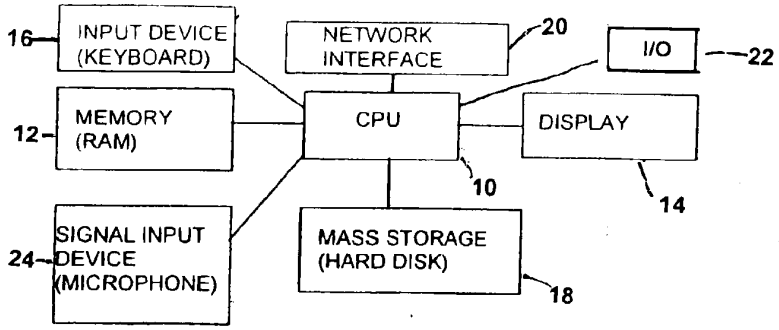


Figure 3

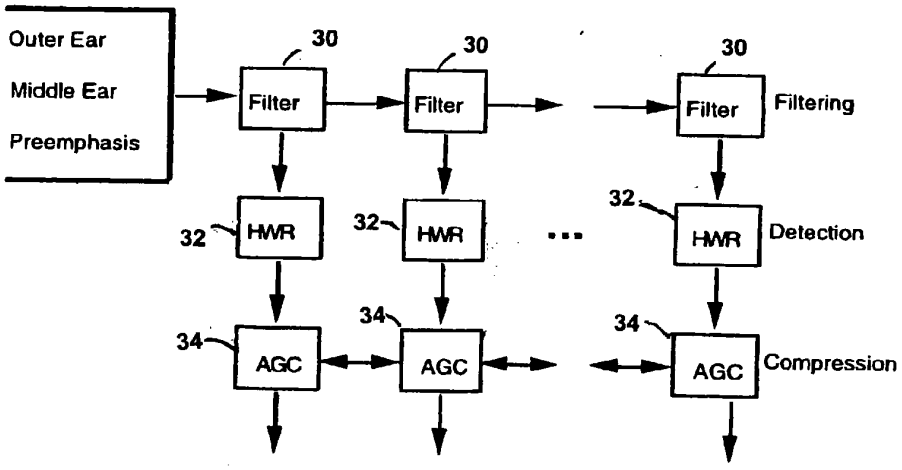


Figure 4

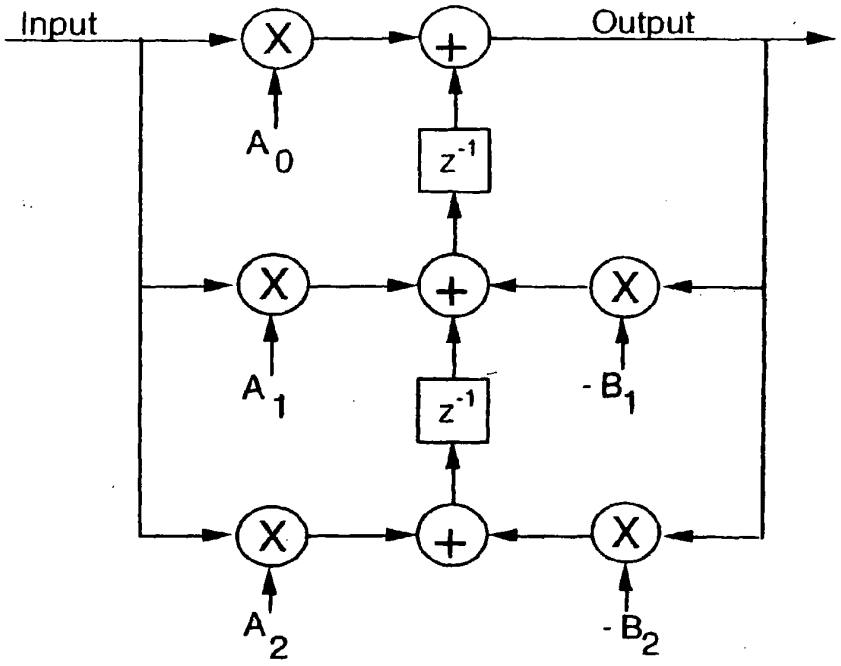


Figure 5

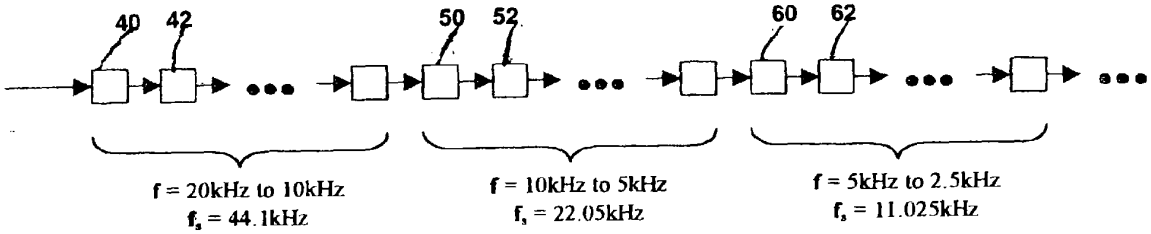


Figure 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/08372

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : H04R 25/00
US CL : 381/312, 320
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 381/312, 320

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
IEEE Explore Web site

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4,536,844 A (LYON) 20 August 1985 (20.08.1985), abstract, figure 1, column 2 lines 22-25	1-4, 12-14, 17, 21, 22
X	US 5,473,759 A (SLANEY et al) 05 December 1995 (05.12.1995), figures 2 and 4	1-4, 12-14, 17, 21, 22
X	KATES, JAMES M. "A Time-Domain Digital Cochlear Model" IEEE Transactions on Signal Processing, December 1991, Volume 39 Number 12, pages 2573-2592	1-4, 12-14, 17, 21, 22

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
18 May 2001 (18.05.2001)

Date of mailing of the international search report
08 JUN 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Forester W. Isen
Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 July 2007 (19.07.2007)

PCT

(10) International Publication Number
WO 2007/081916 A2

(51) International Patent Classification:
H04R 3/00 (2006.01) *H04R 1/02* (2006.01)

(21) International Application Number:
PCT/US2007/000463

(22) International Filing Date: 5 January 2007 (05.01.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/756,826 5 January 2006 (05.01.2006) US
11/343,524 30 January 2006 (30.01.2006) US

(71) Applicant *(for all designated States except US)*: **AUDIENCE, INC.** [US/US]; 1330 Villa Street, Mountain View, California 94041 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **AVENDANO, Carlos** [US/US]; Audience, Inc., 1330 Villa Street, Mountain View, California 94041 (US). **SANTOS, Peter** [US/US]; Audience, Inc., 1330 Villa Street, Mountain View, California 94041 (US).

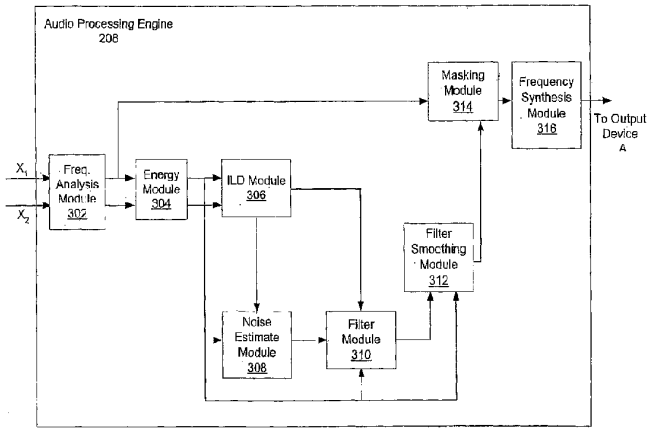
(74) Agents: **KLOKE, Daniel** et al.; Carr & Ferrell LLP, 2200 Geng Road, Palo Alto, California 94303 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR UTILIZING INTER-MICROPHONE LEVEL DIFFERENCES FOR SPEECH ENHANCEMENT



(57) Abstract: Systems and methods for utilizing inter-microphone level differences to attenuate noise and enhance speech are provided. In exemplary embodiments, energy estimates of acoustic signals received by a primary microphone and a secondary microphone are determined in order to determine an inter-microphone level difference (ILD). This ILD in combination with a noise estimate based only on a primary microphone acoustic signal allow a filter estimate to be derived. In some embodiments, the derived filter estimate may be smoothed. The filter estimate is then applied to the acoustic signal from the primary microphone to generate a speech estimate.

WO 2007/081916 A2

WO 2007/081916 A2



Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SYSTEM AND METHOD FOR UTILIZING INTER-MICROPHONE LEVEL DIFFERENCES FOR SPEECH ENHANCEMENT

By: Carlos Avendano and Peter Santos

BACKGROUND OF THE INVENTION

[0001] Presently, there are numerous methods for reducing background noise in speech recordings made in adverse environments. One such method is to use two or more microphones on an audio device. These microphones are localized and allow the device to determine a difference between the microphone signals. For example, due to a space difference between the microphones, the difference in times of arrival of the signals from a speech source to the microphones may be utilized to localize the speech source. Once localized, the signals can be spatially filtered to suppress the noise originating from different directions.

[0002] Beamforming techniques utilizing a linear array of microphones may create an "acoustic beam" in a direction of the source, and thus can be used as spatial filters. This method, however, suffers from many disadvantages. First, it is necessary to identify the direction of the speech source. The time delay, however, is difficult to estimate due to such factors as reverberation which may create ambiguous or incorrect information. Second, the number of sensors needed to achieve adequate spatial filtering is generally large (e.g., more than two). Additionally, if the microphone array is used on a small device, such as a cellular phone, beamforming is more difficult at lower frequencies because the distance between the microphones of the array is small compared to the wavelength.

[0003] Spatial separation and directivity of the microphones provides not only arrival-time differences but also inter-microphone level differences (ILD) that can be more easily identified than time differences in some applications. Therefore, there is

a need for a system and method for utilizing ILD for noise suppression and speech enhancement.

SUMMARY OF THE INVENTION

[0004] Embodiments of the present invention overcome or substantially alleviate prior problems associated with noise suppression and speech enhancement. In general, systems and methods for utilizing inter-microphone level differences (ILD) to attenuate noise and enhance speech are provided. In exemplary embodiments, the ILD is based on energy level differences.

[0005] In exemplary embodiments, energy estimates of acoustic signals received from a primary microphone and a secondary microphone are determined for each channel of a cochlea frequency analyzer for each time frame. The energy estimates may be based on a current acoustic signal and an energy estimate of a previous frame. Based on these energy estimates the ILD may be calculated.

[0006] The ILD information is used to determine time-frequency components where speech is likely to be present and to derive a noise estimate from the primary microphone acoustic signal. The energy and noise estimates allow a filter estimate to be derived. In one embodiment, a noise estimate of the acoustic signal from the primary microphone is determined based on minimum statistics of the current energy estimate of the primary microphone signal and a noise estimate of the previous frame. In some embodiments, the derived filter estimate may be smoothed to reduce acoustic artifacts.

[0007] The filter estimate is then applied to the cochlea representation of the acoustic signal from the primary microphone to generate a speech estimate. The speech estimate is then converted into time domain for output. The conversion may be performed by applying an inverse frequency transformation to the speech estimate.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1a and 1b are diagrams of two environments in which embodiments of the present invention may be practiced;

[0009] FIG. 2 is a block diagram of an exemplary communication device implementing embodiments of the present invention;

[0010] FIG. 3 is a block diagram of an exemplary audio processing engine; and

[0011] FIG. 4 is a flowchart of an exemplary method for utilizing inter-microphone level differences to enhance speech.

DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0012] The present invention provides exemplary systems and methods for recording and utilizing inter-microphone level differences to identify time frequency regions dominated by speech in order to attenuate background noise and far-field distractors. Embodiments of the present invention may be practiced on any communication device that is configured to receive sound such as, but not limited to, cellular phones, phone handsets, headsets, and conferencing systems. Advantageously, exemplary embodiments are configured to provide improved noise suppression on small devices where prior art microphone arrays will not function well. While embodiments of the present invention will be described in reference to operation on a cellular phone, the present invention may be practiced on any communication device.

[0013] Referring to FIG. 1a and 1b, environments in which embodiments of the present invention may be practiced are shown. A user provides an audio (speech) source 102 to a communication device 104. The communication device 104 comprises at least two microphones: a primary microphone 106 relative to the audio source 102 and a secondary microphone 108 located a distance away from the primary microphone 106. In exemplary embodiments, the microphones 106 and 108 are omni-directional microphones. Alternative embodiments may utilize other forms of microphones or acoustic sensors.

[0014] While the microphones 106 and 108 receive sound information from the speech source 102, the microphones 106 and 108 also pick up noise 110. While the noise 110 is shown coming from a single location, the noise may comprise any sounds from one or more locations different than the speech and may include reverberations and echoes.

[0015] Embodiments of the present invention exploit level differences (e.g., energy differences) between the two microphones 106 and 108 independent of how the level differences are obtained. In FIG. 1a because the primary microphone 106 is much closer to the speech source 102 than the secondary microphone 108, the intensity level is

higher for the primary microphone 106 resulting in a larger energy level during a speech/voice segment. In FIG. 1b, because directional response of the primary microphone 106 is highest in the direction of the speech source 102 and directional response of the secondary microphone 108 is lower in the direction of the speech source 102, the level difference is highest in the direction of the speech source 102 and lower elsewhere.

[0016] The level differences may then be used to discriminate speech and noise in the time-frequency domain. Further embodiments may use a combination of energy level difference and time delays to discriminate speech. Based on binaural cue decoding, speech signal extraction or speech enhancement may be performed.

[0017] Referring now to FIG. 2, the exemplary communication device 104 is shown in more detail. The exemplary communication device 200 is an audio receiving device that comprises a processor 202, the primary microphone 106, the secondary microphone 108, an audio processing engine 204, and an output device 206. The communication device 104 may comprise further components necessary for communication device 104 operation, but not related to noise suppression or speech enhancement. The audio processing engine 204 will be discussed in more details in connection with FIG. 3.

[0018] As previously discussed, the primary and secondary microphones 106 and 108, respectively, are spaced a distance apart in order to allow for an energy level difference between them. It should be noted that the microphones 106 and 108 may comprise any type of acoustic receiving device or sensor, and may be omni-directional, unidirectional, or have other directional characteristics or polar patters. Once received by the microphones 106 and 108, the acoustic signals are converted by an analog-to-digital converter (not shown) into digital signals for processing in accordance with some embodiments. In order to differentiate the acoustic signals, the acoustic signal received by the primary microphone 106 is herein referred to as the primary acoustic signal, while the acoustic signal received by the secondary microphone 108 is herein referred to as the secondary acoustic signal.

[0019] The output device 206 is any device which provides an audio output to the user. For example, the output device 206 may be an earpiece of a headset or handset, or a speaker on a conferencing device.

[0020] FIG. 3 is a detailed block diagram of the exemplary audio processing engine 204, according to one embodiment of the present invention. In one embodiment, the acoustic signals (i.e., X_1 and X_2) received from the primary and secondary microphones 106 and 108 (FIG. 2) are converted to digital signals and forwarded to a frequency analysis module 302. In one embodiment, the frequency analysis module 302 takes the acoustic signals and mimics a cochlea implementation (i.e., cochlea domain) using a filter bank. Alternatively, other filter banks such as short-time Fourier transform (STFT), sub-band filter banks, modulated complex lapped transforms, wavelets, etc. can be used for the frequency analysis and synthesis. Because most sounds (e.g., acoustic signal) are complex and comprise more than one frequency, a sub-band analysis on the acoustic signal determines what individual frequencies are present in the complex acoustic signal during a frame (i.e., a predetermined period of time). In one embodiment, the frame is 4ms long.

[0021] Once the frequencies are determined, the signals are forwarded to an energy module 304 which computes energy level estimates during an interval of time. The energy estimate may be based on bandwidth of the cochlea channel and the acoustic signal. The exemplary energy module 304 is a component which, in some embodiments, can be represented mathematically. Thus, the energy level of the acoustic signal received at the primary microphone 106 may be approximated, in one embodiment, by the following equation

$$E_1(t,\omega) = \lambda_E \left|X_1(t,\omega)\right|^2 + (1-\lambda_E)E_1(t-1,\omega)$$

where λ_E is a number between zero and one that determines an averaging time constant, $X_1(t,\omega)$ is the acoustic signal of the primary microphone 106 in the cochlea domain, ω represents the frequency, and t represents time. As shown, a present energy level of the primary microphone 106, $E_1(t,\omega)$, is dependent upon a previous energy level

of the primary microphone 106, $E_1(t\text{-}1,w)$. In some other embodiments, the value of λ_E can be different for different frequency channels. Given a desired time constant T (e.g., 4 ms) and the sampling frequency f_s (e.g. 16 kHz), the value of λ_E can be approximated as

$$\lambda_E = 1 - e^{-\frac{1}{Tf_s}}$$

[0022] The energy level of the acoustic signal received from the secondary microphone 108 may be approximated by a similar exemplary equation

$$E_2(t,\omega) = \lambda_E \left| X_2(t,\omega) \right|^2 + (1-\lambda_E) E_2(t-1,\omega)$$

where $X_2(t,w)$ is the acoustic signal of the secondary microphone 108 in the cochlea domain. Similar to the calculation of energy level for the primary microphone 106, energy level for the secondary microphone 108, $E_2(t,\omega)$, is dependent upon a previous energy level of the secondary microphone 108, $E_2(t\text{-}1,\omega)$.

[0023] Given the calculated energy levels, an inter-microphone level difference (ILD) may be determined by an ILD module 306. The ILD module 306 is a component which may be approximated mathematically, in one embodiment, as

$$ILD(t,\omega) = \left[1 - 2\frac{E_1(t,\omega)E_2(t,\omega)}{E_1^{\,2}(t,\omega) + E_2^{\,2}(t,\omega)} \right] * sign(E_1(t,\omega) - E_2(t,\omega))$$

where E_1 is the energy level of the primary microphone 106 and E_2 is the energy level of the secondary microphone 108, both of which are obtained from the energy module 304. This equation provides a bounded result between -1 and 1. For example, ILD goes to 1 when the E_2 goes to 0, and ILD goes to -1 when E_1 goes to 0. Thus, when the speech source is close to the primary microphone 106 and there is no noise, ILD =1, but as more noise is added, the ILD will change. Further, as more noise is picked up by both of the microphones 106 and 108, it becomes more difficult to discriminate speech from noise.

[0024] The above equation is desirable over an ILD calculated via a ratio of the energy levels, such as $ILD(t,\omega) = \frac{E_1(t,\omega)}{E_2(t,\omega)}$, where ILD is not bounded and may go to infinity as the energy level of the primary microphone gets smaller.

[0025] In an alternative embodiment, the ILD may be approximated by

$$ILD(t,\omega)=\frac{E_1(t,\omega)-E_2(t,\omega)}{E_1(t,\omega)+E_2(t,\omega)}.$$

Here, the ILD calculation is also bounded between -1 and 1. Therefore, this alternative ILD calculation may be used in one embodiment of the present invention.

[0026] According to an exemplary embodiment of the present invention, a Wiener filter is used to suppress noise/enhance speech. In order to derive a Wiener filter estimate, however, specific inputs are required. These inputs comprise a power spectral density of noise and a power spectral density of the source signal. As such, a noise estimate module 308 may be provided to determine a noise estimate for the acoustic signals.

[0027] According to exemplary embodiments, the noise estimate module 308 attempts to estimate the noise components in the microphone signals. In exemplary embodiments, the noise estimate is based only on the acoustic signal received by the primary microphone 106. The exemplary noise estimate module 308 is a component which can be approximated mathematically by

$$N(t,\omega)=\lambda_1(t,\omega)E_1(t,\omega)+(1-\lambda_1(t,\omega))\min[N(t-1,\omega),E_1(t,\omega)]$$

according to one embodiment of the present invention. As shown, the noise estimate in this embodiment is based on minimum statistics of a current energy estimate of the primary microphone 106, $E_1(t,\omega)$ and a noise estimate of a previous time frame, $N(t-1,\omega)$. Therefore the noise estimation is performed efficiently and with low latency.

[0028] $\lambda_1(t,\omega)$ in the above equation is derived from the ILD approximated by the ILD module 306, as

$$\lambda_1(t,\omega)=\begin{cases}\approx 0 & \text{if } ILD(t,\omega) < threshold\\ \approx 1 & \text{if } ILD(t,\omega) > threshold\end{cases}$$

That is, when speech at the primary microphone 106 is smaller than a threshold value (e.g., *threshold* = 0.5) above which speech is expected to be, λ_1 is small, and thus the noise estimator follows the noise closely. When ILD starts to rise (e.g., because speech is detected), however, λ_1 increases. As a result, the noise estimate module 308 slows down

the noise estimation process and the speech energy does not contribute significantly to the final noise estimate. Therefore, exemplary embodiments of the present invention may use a combination of minimum statistics and voice activity detection to determine the noise estimate.

[0029] A filter module 310 then derives a filter estimate based on the noise estimate. In one embodiment, the filter is a Weiner filter. Alternative embodiments may contemplate other filters. Accordingly, the Weiner filter approximation may be approximated, according to one embodiment, as

$$W = \left(\frac{P_s}{P_s + P_n} \right)^{\alpha}$$

, where P_s is a power spectral density of speech and P_n is a power spectral density of noise. According to one embodiment, P_n is the noise estimate, $N(t,\omega)$, which is calculated by the noise estimate module 308. In an exemplary embodiment, $P_s = E_1(t,\omega) - \beta N(t,\omega)$, where $E_1(t,\omega)$ is the energy estimate of the primary microphone 106 from the energy module 304, and $N(t,\omega)$ is the noise estimate provided by the noise estimate module 308. Because the noise estimate changes with each frame, the filter estimate will also change with each frame.

[0030] β is an over-subtraction term which is a function of the ILD. β compensates bias of minimum statistics of the noise estimate module 308 and forms a perceptual weighting. Because time constants are different, the bias will be different between portions of pure noise and portions of noise and speech. Therefore, in some embodiments, compensation for this bias may be necessary. In exemplary embodiments, β is determined empirically (e.g., 2-3 dB at a large ILD, and is 6-9 dB at a low ILD).

[0031] α in the above exemplary Weiner filter equation is a factor which further suppresses the noise estimate. α can be any positive value. In one embodiment, nonlinear expansion may be obtained by setting α to 2. According to exemplary embodiments, α is determined empirically and applied when a body of $W = \left(\frac{P_s}{P_s + P_n} \right)$

falls below a prescribed value (e.g., 12 dB down from the maximum possible value of W, which is unity).

[0032] Because the Weiner filter estimation may change quickly (e.g., from one frame to the next frame) and noise and speech estimates can vary greatly between each frame, application of the Weiner filter estimate, as is, may result in artifacts (e.g., discontinuities, blips, transients, etc.). Therefore, an optional filter smoothing module 312 is provided to smooth the Wiener filter estimate applied to the acoustic signals as a function of time. In one embodiment, the filter smoothing module 312 may be mathematically approximated as

$$M(t,\omega) = \lambda_s(t,\omega)W(t,\omega) + (1-\lambda_s(t,\omega))M(t-1,\omega)$$

, where λ_s is a function of the Weiner filter estimate and the primary microphone energy, E_1.

[0033] As shown, the filter smoothing module 312, at time (t) will smooth the Weiner filter estimate using the values of the smoothed Weiner filter estimate from the previous frame at time (t-1). In order to allow for quick response to the acoustic signal changing quickly, the filter smoothing module 312 performs less smoothing on quick changing signals, and more smoothing on slower changing signals. This is accomplished by varying the value of λ_s according to a weighed first order derivative of E_1 with respect to time. If the first order derivative is large and the energy change is large, then λ_s is set to a large value. If the derivative is small then λ_s is set to a smaller value.

[0034] After smoothing by the filter smoothing module 312, the primary acoustic signal is multiplied by the smoothed Weiner filter estimate to estimate the speech. In the above Weiner filter embodiment, the speech estimate is approximated by $S(t,\omega) = X_1(t,\omega) * M(t,\omega)$, where X_1 is the acoustic signal from the primary microphone 106. In exemplary embodiments, the speech estimation occurs in a masking module 314.

[0035] Next, the speech estimate is converted back into time domain from the cochlea domain. The conversion comprises taking the speech estimate, $S(t,\omega)$, and multiplying this with an inverse frequency of the cochlea channels in a frequency synthesis module 316. Once conversion is completed, the signal is output to user.

[0036] It should be noted that the system architecture of the audio processing engine 204 of FIG. 3 is exemplary. Alternative embodiments may comprise more components, less components, or equivalent components and still be within the scope of embodiments of the present invention. Various modules of the audio processing engine 208 may be combined into a single module. For example, the functionalities of the frequency analysis module 302 and energy module 304 may be combined into a single module. Furthermore, the functions of the ILD module 306 may be combined with the functions of the energy module 304 alone, or in combination with the frequency analysis module 302. As a further example, the functionality of the filter module 310 may be combined with the functionality of the filter smoothing module 312.

[0037] Referring now to FIG. 4, a flowchart 400 of an exemplary method for noise suppression utilizing inter-microphone level differences is shown. In step 402, audio signals are received by a primary microphone 106 and a secondary microphone 108 (FIG. 2). In exemplary embodiments, the acoustic signals are converted to digital format for processing.

[0038] Frequency analysis is then performed on the acoustic signals by the frequency analysis module 302 (FIG. 3) in step 404. According to one embodiment, the frequency analysis module 302 utilizes a filter bank to determine individual frequencies present in the complex acoustic signal.

[0039] In step 406, energy estimates for acoustic signals received at both the primary and secondary microphones 106 and 108 are computed. In one embodiment, the energy estimates are determined by an energy module 304 (FIG. 3). The exemplary energy module 304 utilizes a present acoustic signal and a previously calculated energy estimate to determine the present energy estimate.

[0040] Once the energy estimates are calculated, inter-microphone level differences (ILD) are computed in step 408. In one embodiment, the ILD is calculated based on the energy estimates of both the primary and secondary acoustic signals. In exemplary embodiments, the ILD is computed by the ILD module 306 (FIG. 3).

[0041] Based on the calculated ILD, noise is estimated in step 410. According to embodiments of the present invention, the noise estimate is based only on the acoustic signal received at the primary microphone 106. The noise estimate may be based on the present energy estimate of the acoustic signal from the primary microphone 106 and a previously computed noise estimate. In determining the noise estimate, the noise estimation is frozen or slowed down when the ILD increases, according to exemplary embodiments of the present invention.

[0042] In step 412, a filter estimate is computed by the filter module 310 (FIG. 3). In one embodiment, the filter used in the audio processing engine 204 (FIG. 3) is a Wiener filter. Once the filter estimate is determined, the filter estimate may be smoothed in step 414. Smoothing prevents fast fluctuations which may create audio artifacts. The smoothed filter estimate is applied to the acoustic signal from the primary microphone 106 in step 416 to generate a speech estimate.

[0043] In step 418, the speech estimate is converted back to the time domain. Exemplary conversion techniques apply an inverse frequency of the cochlea channel to the speech estimate. Once the speech estimate is converted, the audio signal may now be output to the user in step 420. In some embodiments, the digital acoustic signal is converted to an analog signal for output. The output may be via a speaker, earpieces, or other similar devices.

[0044] The above-described modules can be comprised of instructions that are stored on storage media. The instructions can be retrieved and executed by the processor 202 (FIG. 2). Some examples of instructions include software, program code, and firmware. Some examples of storage media comprise memory devices and integrated circuits. The instructions are operational when executed by the processor 202 to direct the processor 202 to operate in accordance with embodiments of the

present invention. Those skilled in the art are familiar with instructions, processor(s), and storage media.

[0045] The present invention is described above with reference to exemplary embodiments. It will be apparent to those skilled in the art that various modifications may be made and other embodiments can be used without departing from the broader scope of the present invention. Therefore, these and other variations upon the exemplary embodiments are intended to be covered by the present invention.

CLAIMS

What is claimed is:

1. A method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a filter estimate during a frame, the filter estimate based on a noise estimate of the primary acoustic signal, an energy estimate of the primary acoustic signal, and an inter-microphone level difference based on the primary and secondary acoustic signals; and

applying the filter estimate to the primary acoustic signal to produce a speech estimate.

2. The method of claim 1 further comprising determining an energy estimate for each of the acoustic signals during the frame.

3. The method of claim 2 wherein the energy estimate of the primary acoustic signal is approximated as $E_1(t,\omega) = \lambda_E |X_1(t,\omega)|^2 + (1-\lambda_E)E_1(t-1,\omega)$.

4. The method of claim 2 wherein the energy estimate of the secondary acoustic signal is approximated as $E_2(t,\omega) = \lambda_E |X_2(t,\omega)|^2 + (1-\lambda_E)E_2(t-1,\omega)$.

5. The method of claim 2 further comprising using the energy estimates to determine the inter-microphone level difference for the frame.

6. The method of claim 5 wherein the inter-microphone level difference is approximated by $ILD(t,\omega) = \left[1 - 2\frac{E_1(t,\omega)E_2(t,\omega)}{E_1^2(t,\omega) + E_2^2(t,\omega)}\right] * sign(E_1(t,\omega) - E_2(t,\omega))$.

7. The method of claim 5 wherein the inter-microphone level difference is approximated by $ILD(t,\omega)=\frac{E_1(t,\omega)-E_2(t,\omega)}{E_1(t,\omega)+E_2(t,\omega)}$.

8. The method of claim 1 wherein the noise estimate is based on an energy estimate of the primary acoustic signal and the inter-microphone level difference.

9. The method of claim 8 wherein the noise estimate is approximated as $N(t,\omega)=\lambda_I(t,\omega)E_1(t,\omega)+(1-\lambda_I(t,\omega))\min[N(t-1,\omega),E_1(t,\omega)]$.

10. The method of claim 1 further comprising smoothing the filter estimate prior to applying the filter estimate to the primary acoustic signal.

11. The method of claim 10 wherein the smoothing is approximated as $M(t,\omega)=\lambda_s(t,\omega)W(t,\omega)+(1-\lambda_s(t,\omega))M(t-1,\omega)$.

12. The method of claim 1 further comprising converting the speech estimate to a time domain.

13. The method of claim 1 further comprising outputting the speech estimate to a user.

14. The method of claim 1 wherein the filter estimate is based on a Weiner filter.

15. A system for enhancing speech on a device, comprising:

a primary microphone configured to receive a primary acoustic signal;

a secondary microphone located a distance away from the primary microphone and configured to receive a secondary acoustic signal; and

an audio processing engine configured to enhance speech received at the primary microphone, the audio processing engine comprising

a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the primary acoustic signal

and an inter-microphone level difference; and

a filter module configured to determine a filter estimate to be applied to the

primary acoustic signal to generate a filtered acoustic signal, the filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

16. The system of claim 15 wherein the audio processing engine further comprises an energy module configured to determine energy estimates for a frame of the primary and secondary acoustic signals.

17. The system of claim 15 wherein the audio processing engine further comprises an inter-microphone level difference module configured to determine the inter-microphone level difference.

18. The system of claim 15 wherein the audio processing engine further comprises a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

19. The system of claim 15 wherein the audio processing engine further comprises a masking module configured to determine the speech estimate.

20. A computer readable medium having embodied thereon a program, the program being executable by a machine to perform a method for enhancing speech on a device, the method comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining an energy estimate over a frame for each of the acoustic signals;

using the energy estimates to determine an inter-microphone level difference for the frame;

generating a noise estimate based on the energy estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference;

calculating a filter estimate based on the noise estimate and the inter-microphone level difference; and

applying the filter estimate to the primary acoustic signal to produce a speech estimate.

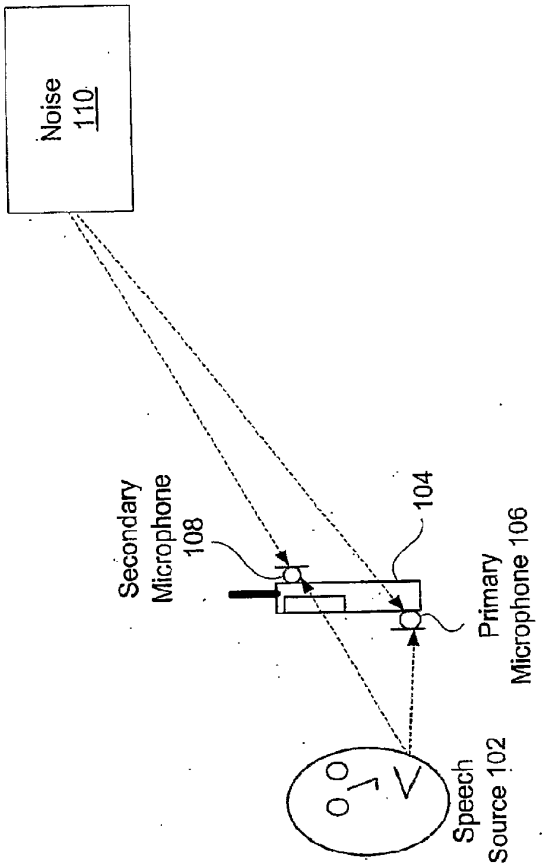

Noise
110
Secondary
Microphone
108
104
Primary
Microphone 106
Speech
Source 102

FIG. 1a

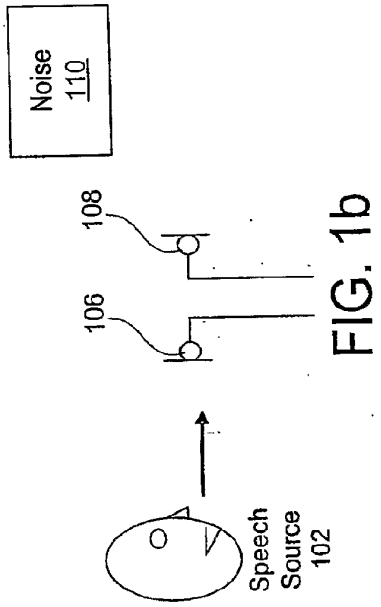

Noise
110
108
106
Speech
Source
102

FIG. 1b

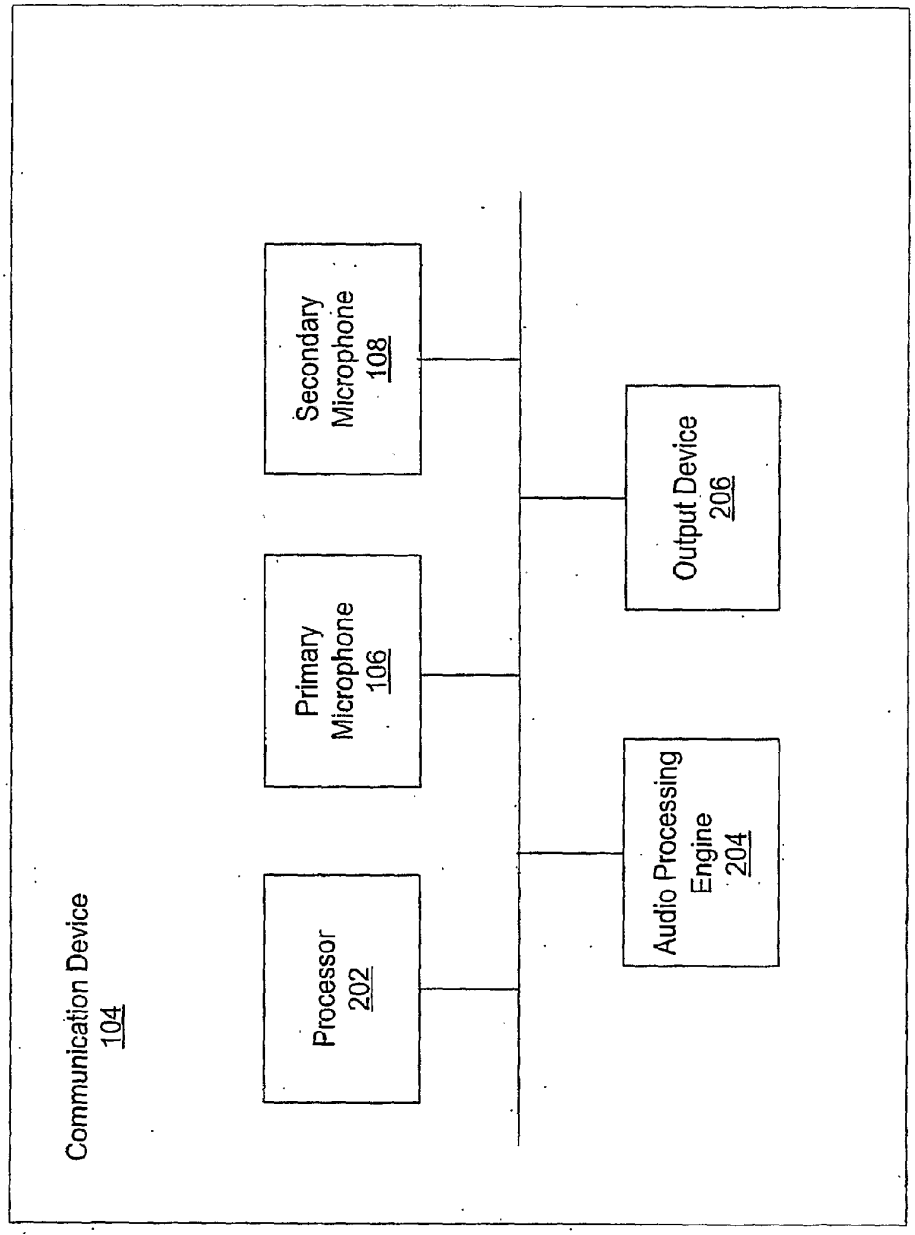

Communication Device
104
Processor
202
Primary Microphone
106
Secondary Microphone
108
Audio Processing Engine
204
Output Device
206

FIG. 2

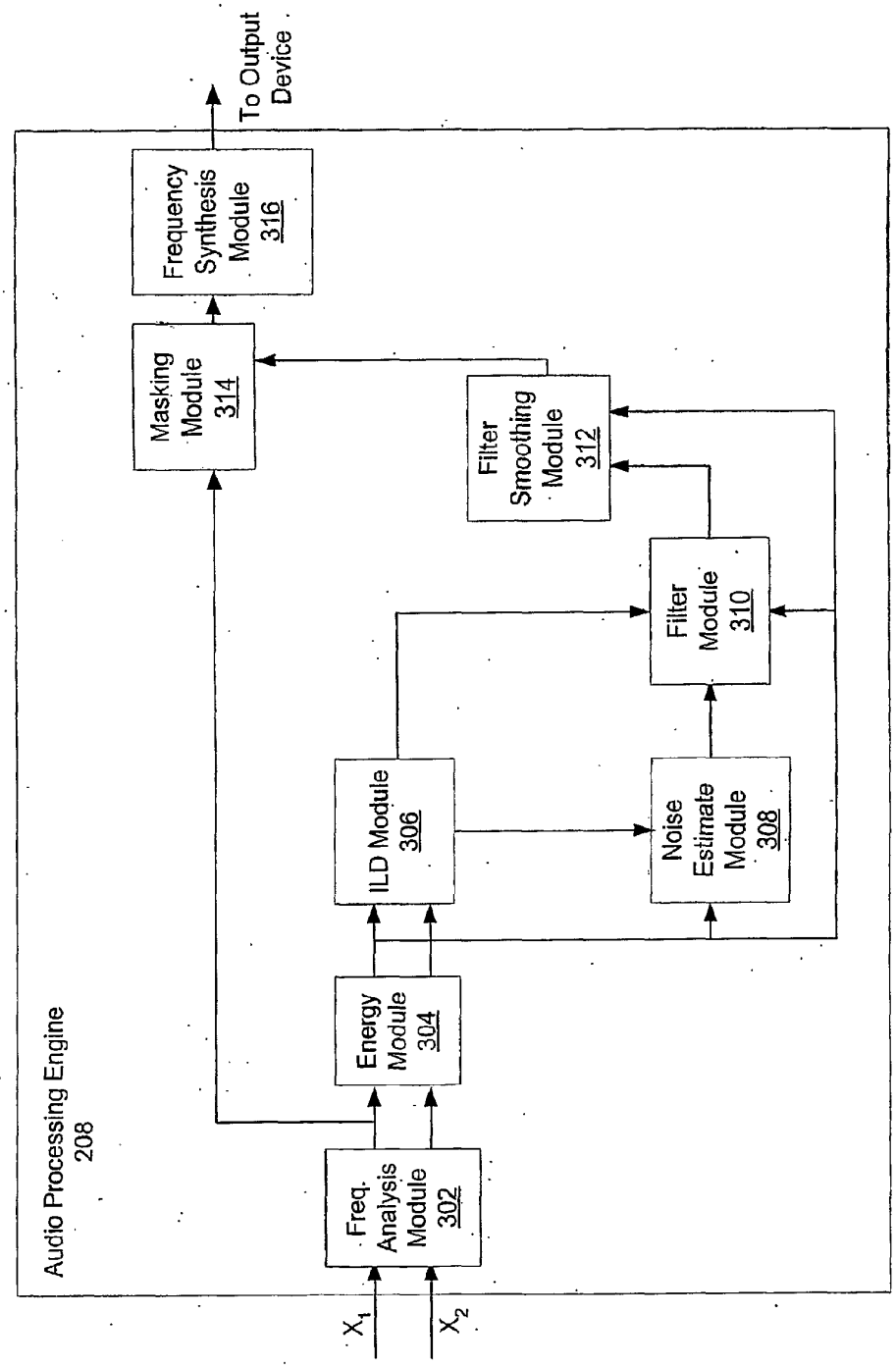

To Output Device
Frequency Synthesis Module 316
Masking Module 314
Filter Smoothing Module 312
Filter Module 310
ILD Module 306
Noise Estimate Module 308
Energy Module 304
Freq. Analysis Module 302
Audio Processing Engine 208
X_1
X_2

FIG. 3

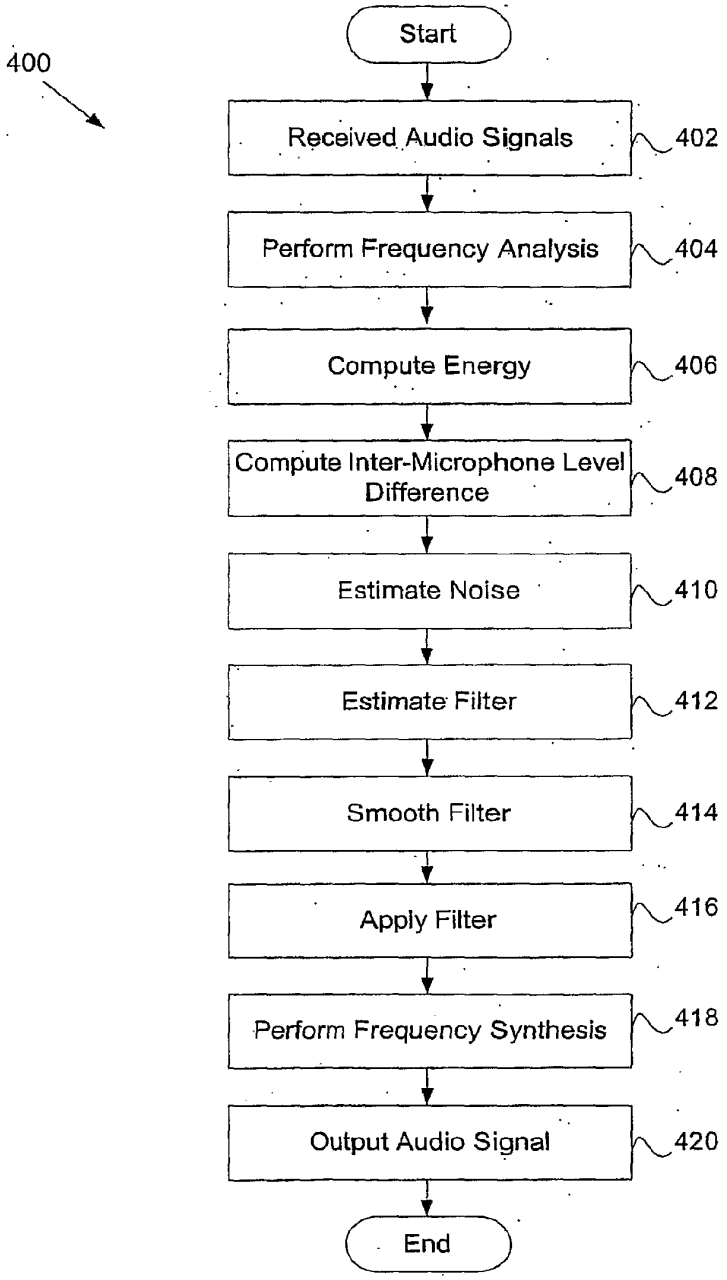

400
Start
Received Audio Signals
402
Perform Frequency Analysis
404
Compute Energy
406
Compute Inter-Microphone Level Difference
408
Estimate Noise
410
Estimate Filter
412
Smooth Filter
414
Apply Filter
416
Perform Frequency Synthesis
418
Output Audio Signal
420
End

FIG. 4

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **11-249693**
(43)Date of publication of application : **17.09.1999**

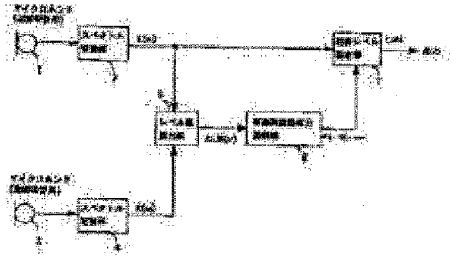
(51)Int.Cl. **G10L 3/02**
G10L 9/00

(21)Application number : **10-049725**
(22)Date of filing : **02.03.1998**

(71)Applicant : **NIPPON TELEGR & TELEPH CORP <NTT>**
(72)Inventor : **TAKANO TOMOHIRO**
MATSUI HIROYUKI

(54) SOUND COLLECTING DEVICE

(57)Abstract:
PROBLEM TO BE SOLVED: To separate/detect only the level of an ambient noise or only a target signal level even when the target signal sound collected by an air guide microphone is mixed with the ambient noise.
SOLUTION: Signals almost containing an ambient noise collected by microphones 1, 2 positioned adjacent to the sound source of a target signal and positioned apart from the second source of the target signal are converted respectively to spectra by spectrum conversion parts 3, 4, and the frequency component excepting the target signal, or the predominant frequency component of the target signal is separated/detected respectively from a level difference of an amplitude spectrum at every frequency component, and the noise level or the target signal level are separated/detected respectively from respective spectrum intensity.



(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号

特開平11－249693

(43)公開日　平成11年(1999) 9月17日

(51)Int.Cl.[6]		識別記号	FI		
Ｇ１０Ｌ	3/02	３０１	Ｇ１０Ｌ	3/02	３０１Ｅ
	9/00			9/00	Ｆ
					Ｅ

審査請求　未請求　請求項の数３　ＯＬ　（全 10 頁）

(21)出願番号　特願平10－49725

(22)出願日　平成10年(1998) 3月2日

(71)出願人　000004226
日本電信電話株式会社
東京都新宿区西新宿三丁目19番２号

(72)発明者　高野　智大
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内

(72)発明者　松井　弘行
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内

(74)代理人　弁理士　草野　卓

(54)【発明の名称】　収音装置

(57)【要約】

【課題】　気導マイクロホンにより収音された目的信号が周囲騒音に埋もれていても周囲騒音のレベルのみまたは目的信号のレベルのみを分離検出可能とする。

【解決手段】　目的信号の音源に近い位置と、前記位置より目的信号の音源から離れた位置に設置されたマイクロホン（１，２）により収音された周囲騒音をほぼ同様に含む信号を各々スペクトルに変換し（３，４）、各周波数成分ごとの振幅スペクトルのレベル差（５）から、目的信号以外の周波数成分、または目的信号の支配的な周波数成分について、そのスペクトル強度から各々騒音レベル、または目的信号のレベルを分離検出する。

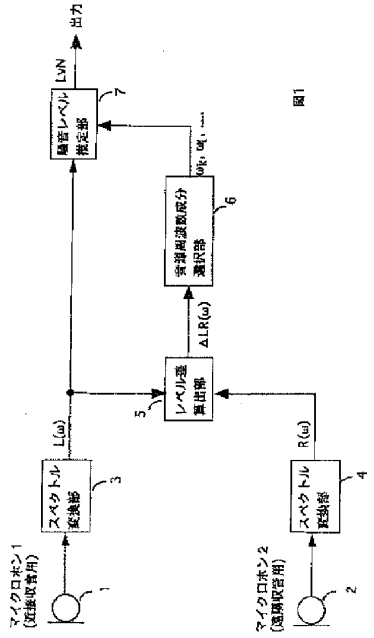


図1

1

【特許請求の範囲】

【請求項１】　目的信号の音源に近い位置に設置された第１マイクロホンと、

前記位置より前記目的信号の音源から離れた位置に設置された第２マイクロホンと、

前記第１、第２マイクロホンの各々の出力信号を振幅スペクトルに変換する第

１、第２スペクトル変換手段と、

前記第１、第２スペクトル変換手段から出力される各周波数成分ごとの振幅スペクトルについて、これらレベル差をそれぞれ計算するレベル差算出手段と、

前記レベル差算出手段により出力される各周波数成分ごとのレベル差と、予め設定されたしきい値とを比較し、目的信号が支配的な周波数成分か否かを判定する音源周波数成分選択手段と、

前記スペクトル変換手段から出力される出力信号の振幅スペクトルから、前記音源周波数成分選択手段により目的信号が支配的と判定されなかった周波数成分を抽出し、その振幅スペクトルから目的信号以外の騒音の振幅スペクトルを推定するスペクトル推定手段とを有することを特徴とする収音装置。

【請求項２】　上記スペクトル推定手段で推定された騒音の振幅スペクトルから騒音レベルを推定する手段を有することを特徴とする請求項１記載の収音装置。

【請求項３】　目的信号の音源に近い位置に設置された第１マイクロホンと、

前記位置より前記目的信号の音源から離れた位置に設置された第２マイクロホンと、

前記第１、第２マイクロホンの各々の出力信号を振幅スペクトルに変換する第１、第２スペクトル変換手段と、

前記第１、第２スペクトル変換手段から出力される各周波数成分ごとの振幅スペクトルについて、レベル差をそれぞれ計算するレベル差算出手段と、

前記レベル差算出手段により出力される各周波数成分ごとのレベル差と、予め設定されたしきい値とを比較し、目的信号が支配的な周波数成分か否かを判定する音源周波数成分選択手段と、

前記スペクトル変換手段から出力される出力信号の振幅スペクトルから、前記音源周波数成分選択手段により目的信号が支配的と判定された周波数成分を抽出し、その振幅スペクトルから目的信号の振幅スペクトルあるいは出力レベルを推定する音源レベル推定手段とを有することを特徴とする収音装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】この発明は、周囲騒音が混在した音源信号に対して、周囲騒音のみのレベル、または目的とする音源のみのレベルを検出する収音装置に関するものである。

【０００２】

2

【従来の技術】一般に音声会議装置や録音装置での収音装置においては、目的とする信号と周囲騒音とが同時に入力され、目的とする信号レベルと周囲騒音のレベルを分離して検出することは困難であった。従来、一部の音声会議装置や拡声電話機の中に、音声の周期性と騒音の非周期性を利用し、零交叉数等から音声の発声又は非発声区間を検出し、騒音の抑圧等を行っているものがある。しかし、これらは、目的とする信号は音声、周囲騒音は非音声である必要があるため、共に周期性のある音声等の信号の場合は誤検出し、性能を著しく劣化させるという問題があった。即ち、目的とする信号や周囲騒音の音源が、周期性等での相違がないと、各々の信号レベルを分離検出することができないため、適用範囲を限定して利用せざるを得ないという問題があった。

【０００３】

【発明が解決しようとする課題】この発明の目的は、目的信号と周囲騒音が周期性等の相違がない音源であっても、各々の信号レベルを分離して検出可能な収音装置を提供することにある。

【０００４】

【課題を解決するための手段】請求項１記載の発明は、目的信号の音源に近い位置に設置された第１マイクロホンと、前記位置より目的信号の音源から離れた位置に設置された第２マイクロホンの各々の出力信号が第１、第２スペクトル変換手段により、振幅スペクトルに変換され、各周波数成分ごとにこれら振幅スペクトル間のレベル差がレベル差算出手段で計算され、各周波数成分ごとのこのレベル差と、予め設定されたしきい値とが音源周波数成分選択手段で比較され、目的信号が支配的な周波数成分か否かの判定がなされ、前記スペクトル変換手段の出力振幅スペクトルから、前記目的信号が支配的と判定されなかった周波数成分が騒音振幅スペクトル推定手段により抽出され、つまり目的信号以外の騒音の振幅スペクトルが抽出される。請求項２の発明では請求項１の発明で抽出された騒音の振幅スペクトルから騒音の出力レベルが騒音レベル推定手段により推定される。

【０００５】請求項３記載の発明は、請求項１の発明と、音源周波数成分選択手段により、目的信号が支配的な周波数成分か否かを判定する構成までは同一である。請求項３の発明では前記スペクトル変換手段から出力される振幅スペクトルから、目的信号が支配的と判定された周波数成分が音源レベル推定手段により抽出され、その振幅スペクトルから目的信号の振幅スペクトルあるいは出力がレベル推定される。

作用

請求項１乃至請求項３記載の各発明においては、目的信号の音源に近い位置と、前記位置より目的信号の音源から離れた位置に設置されたマイクロホンの振幅スペクトルのレベル差によって目的信号が支配的な周波数成分か否かの判定が行われる。この発明の装置構成では、目的

信号について、その音源は一方のマイクロホンより他方のマイクロホンに近く、かつ騒音源は両マイクロホンから可成り離れている状況を想定しているため、２つのマイクロホンの間で安定したレベル差が生じる。一方、周囲騒音については、その音源とマイクロホンの間の距離は、目的信号の音源とマイクロホンの間の距離に比べて長くなると考えてよく、このため、目的信号によって生じるレベル差は、騒音によって生じるレベル差よりも常に大きくなると考えられる。

【０００６】この発明では、上記のように２つのマイクロホンに生じるレベル差が目的信号と騒音とで異なる点に着目して、目的信号が支配的と判定されなかった周波数成分と目的信号が支配的と判定された周波数成分とを分離して、その振幅スペクトルより騒音の振幅スペクトルあるいは目的とする音源の出力レベルを推定する。このように、この発明では音源の周期性等の相違を利用していないため、これらの制約がない目的信号、周囲騒音にも適用できる。

【０００７】

【発明の実施の形態】実施例１

図１は請求項２の発明の実施例の機能構成を示し、図５に請求項２の発明の実施例の処理手順を示す。マイクロホンＩは目的信号の音源に近い位置に在り、マイクロホン２はマイクロホン１より前記目的信号の音源より遠い位置に設けられ、マイクロホン１と２には遠方の騒音などの雑音がほぼ等振幅で到達する。

【０００８】マイクロホン１，２に騒音が重畳した目的信号が各々取り込まれ、それらをディジタル信号として読み込む（Ｓ０２）。読み込まれたマイクロホン１，２の信号を以下では、Ｌ，Ｒとする。スペクトル変換部３，４では、取り込んだ信号Ｌ，ＲをスペクトルＬ（ω_h）、Ｒ（ω_h）（ｈ＝１，２，…，ｎ）に変換する（Ｓ０３）。この変換は、例えば離散的フーリエ変換によって実行される。

【０００９】レベル差算出部５では、Ｌ（ω_h）、Ｒ（ω_h）の各周波数成分について、以下の式で与えられるレベル差ΔＬＲ（ω_h）を計算する（Ｓ０４）。

ΔＬＲ（ω_h）＝２０ｌｏｇ１０（｜Ｌ（ω_h）｜／｜Ｒ（ω_h）｜）

上式中の｜Ｌ（ω_h）｜、｜Ｒ（ω_h）｜は、各々Ｌ，Ｒ信号の振幅スペクトル成分を表わす。

【００１０】音源周波数成分選択部６では、各周波数成分についてΔＬＲ（ω_h）と予め設定されたしきい値Ｔｈ（ω_h）との大小関係より、目的信号が支配的な周波数成分の選択を行う。目的信号が支配的な周波数成分か否かの判定条件は例えば以下の式によって決定される（Ｓ０５）。

ΔＬＲ（ω_h）＞Ｔｈ（ω_h）　→　目的信号が支配的

ΔＬＲ（ω_h）≦Ｔｈ（ω_h）　→　目的信号が支配的でない

つまり、マイクロホン１，２の各出力Ｌ，Ｒは目的信号に対してはそのスペクトル中に有意な差をもつものが表われるが、騒音については、そのスペクトルの大部分はほぼ等しく、ΔＬＲ（ω_h）は小さい値となる。

【００１１】騒音レベル推定部７では、Ｌ（ω_k）（ｈ＝１，２，…，ｎ）より目的信号が支配的でない周波数成分Ｌ（ω_m）（ｍ＝ｋ，１，…）を抽出する（Ｓ０６）。このＬ（ω_m）（ｍ＝ｋ，１，…）より騒音の全帯域にわたる出力レベルＬｖＮを推定する（Ｓ０７）。この推定の方法としては例えば以下の式が考えられる。

ＬｖＮ＝２０ｌｏｇ１０（Σ｜Ｌ（ω_m）｜）

あるいは、以下の式のように全帯域で平均化することによって求めることもできる。

【００１２】ＬｖＮ＝２０ｌｏｇ１０（（ｎ／ｑ）×Σ｜Ｌ（ω_m）｜）

ここで、ｑは目的信号が支配的でないと判定された周波数成分の個数、和Σは目的信号が支配的でない周波数ω_m（ｍ＝ｋ，１，…）に対応するものについてとる。なお、（Ｓ０７）に示した騒音レベルの推定例では全帯域における値として算出されるが、この値は、複数のサブ帯域に分けて求めることにより、各サブ帯域ごとの出力レベルに応じた値として求めることができる。また、（Ｓ０６）の出力が形成する振幅スペクトル包絡より、騒音の振幅スペクトルを推定することも可能である。

【００１３】このように、この実施例では目的とする信号と周囲騒音との識別に周期性の相違等を利用していないので、この制約によらず広く種々の音源（目的信号、周囲騒音）に適用できる。また、騒音と目的信号が重畳している状態でも、それぞれのレベル検出を行なえる特徴がある。図２は、この発明を音声会議等に適用した例で、図１と対応する部分は同一符号を付けてある（以下同様）。周囲騒音レベルに応じて通話レベルや、エコーキャンセラーの学習更新を制御することができる。すなわち、騒音レベル推定部７で推定した騒音レベルを拡声通話制御部８に入力して、騒音レベルが大きい場合は、スピーカ９から放声される音響レベルを大として騒音の影響を小とし、また、マイクロホン１で集音される目的信号レベルを大として相手側で聞きとり易いようにする。あるいは、スピーカ９からの音響信号は、マイクロホン１にとっては騒音であり、この騒音レベルが小さい時はスピーカ９とマイクロホン１との結合も小であるから、エコーキャンセラーの反響路の推定処理を行わない。また騒音レベル推定部７で推定した騒音振幅スペクトルを用いて、スピーカ９とマイクロホン１間の反響路特性の推定に利用することもできる。

【００１４】これらの場合においてこの発明では周囲騒音が例えば周期性のある音源（音声など）であっても騒音として分離し検出できるので、このような周囲騒音に影響を受けないで音声会議装置の通話レベルやエコーキャンセラーの調整ができ、通話性能の優れた拡声通話を

実現できる。騒音振幅スペクトルや騒音出力レベルの推定を行う推定手段７はスペクトル変換部３の出力ではなく、スペクトル変換部４の出力に対し行ってもよい。

実施例２

図３は請求項３の発明の実施例を示す。音源周波数成分選択部６よりの出力として目的信号が支配的な周波数情報を得ること、騒音レベル推定部７の代わりに音源レベル推定部１２を設けたこと、の２点を除けば、図１と同じ動作である。以下で、図７に示す請求項３の発明の実施例の流れ図を用いて音源レベル推定部１２における処理について説明する。

【００１５】音源レベル推定部１２では、Ｌ（ω_h）（ｈ＝１，２，…，ｎ）より目的信号が支配的な周波数成分Ｌ（ω_m）（ｍ＝ｉ，ｊ，…）を抽出する（Ｓ０６）。このＬ（ω_m）（ｍ＝ｉ，ｊ，…）より目的信号の全帯域にわたる出力レベルＬｖＳを推定する（Ｓ０７）。この推定の方法としては例えば以下の式が考えられる。

【００１６】

$$LvS = 20\log_{10}(\Sigma|L(\omega_m)|)$$

ここで、和Σは目的信号が支配的な周波数ω_m（ｍ＝ｉ，ｊ，…）に対応するものについてとる。なお、（Ｓ０７）に示した目的信号の出力レベルの推定例では全帯域における値として算出されるが、この値は、複数のサブ帯域に分けて求めることにより、各サブ帯域ごとの出力レベルに応じた値として求めることができる。また、（Ｓ０６）の出力が形成する振幅スペクトル包絡より、目的信号の振幅スペクトルを推定することも可能である。

【００１７】このように、この発明では目的とする信号と周囲騒音との識別に周期性の相違等を利用していないので、この制約によらず広く種々の音源（目的信号、周囲騒音）に適用できる。また、騒音と目的信号が重畳している状態でも、それぞれのレベル検出を行なえる特徴がある。図４は、この発明を録音装置に適用した例で、この例では音源レベル推定部１０で検出した目的信号のレベルのみに応じて周囲騒音の有無や、周囲騒音のレベルにかかわらず、マイクロホン１の出力である目的信号のゲイン調整をゲイン調整部１３で行って、録音部１４へ供給することが可能である。このため、この装置では、周囲騒音にかかわらず、目的の音源のレベルのみに応じたＡＧＣ（自動ゲインコントロール）動作が可能で、目的音源が明瞭に収音できる。

【００１８】

【発明の効果】以上説明したように、請求項１及び２記載の発明によれば、目的信号の音源に近い位置に設置された第１マイクロホンと、前記位置より目的信号の音源から離れた位置に設置された第２マイクロホンとの各出力中の各同一周波数成分ごとの振幅スペクトルについて、レベル差を計算し、しきい値とを比較することにより、目的信号が支配的な周波数成分か否かを判定して目的信号が支配的と判定されなかった周波数成分を抽出し、その振幅スペクトルから目的信号以外の騒音の振幅スペクトルあるいは騒音の出力レベルを推定しているため、騒音、目的音源の種類によらず騒音の出力レベルを正しく検出することができる。

【００１９】この収音装置を利用すれば、例えば音声会議装置やテレビ会議装置と組み合わせて周囲騒音に応じて通話レベルを制御したり、エコーキャンセラーの学習更新のオン／オフ等を制御することができるので、通話性能に優れた会議装置を提供できるメリットがある。請求項３記載の発明によれば、目的信号の音源に近い位置に設置された第１マイクロホンと、前記位置より目的信号の音源から離れた位置に設置された第２マイクロホンとの各出力の各周波数成分ごとの振幅スペクトルについてレベル差を検出し、この各周波数成分ごとのレベル差と、しきい値とを比較して、目的信号が支配的な周波数成分か否かを判定して、マイクロホンから出力される出力信号より目的信号が支配的と判定された周波数成分を抽出し、その振幅スペクトルから目的信号の振幅スペクトルあるいは出力レベルを推定するため、騒音、目的音源の種類によらず目的とする音源の出力レベルを検出することができる。

【００２０】この収音装置を利用すれば、例えば録音装置において周囲騒音の有無にかかわらず目的とする音源のレベルのみに応じてＡＧＣをかけられるので、目的音源を明瞭に収音できるメリットがある。なお、以上の説明で使用したマイクロホンは、無指向性マイクロホンに限定されるものではなく、例えば、マイクロホン１は、目的信号の音源の方向に指向性を有するマイクロホンを使用し、マイクロホン２は、目的信号の音源と反対の方向に指向性を有するマイクロホンを使用してもよい。

【００２１】この発明は、音声会議装置、録音装置以外に、騒音抑圧装置や、音声認識装置の入力装置として使用し、目的音源のレベルと周囲騒音のレベルを検出し、騒音抑圧や音声認識における耐騒音性の改善等にも利用できる。

【図面の簡単な説明】

【図１】請求項２の発明の実施例の機能的構成を示すブロック図。

【図２】図１に示した実施例を拡声通話制御部に利用した例を示すブロック図。

【図３】請求項３の発明の実施例の機能的構成を示すブロック図。

【図４】図３に示した実施例を録音装置に利用した例を示すブロック図。

【図５】図１に示した実施例の動作を示す流れ図。

【図６】図２に示した適用例の動作を示す流れ図。

【図７】図３に示した実施例の動作を示す流れ図。

【図８】図４に示した適用例の動作を示す流れ図。

【図１】

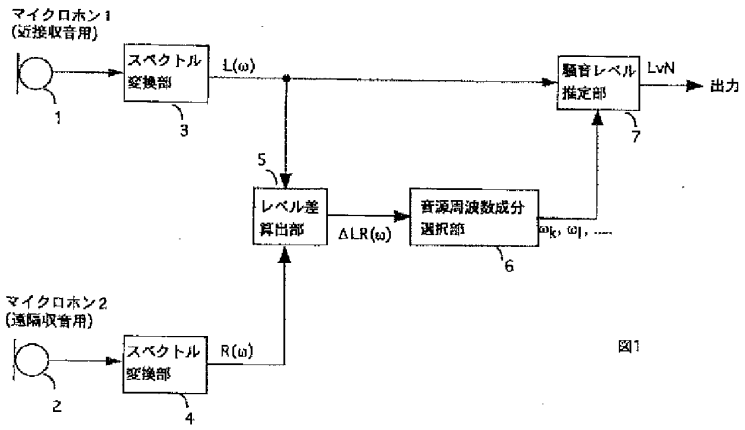

マイクロホン１
(近接収音用)
スペクトル
変換部
L(ω)
騒音レベル
推定部
LvN
出力
レベル差
算出部
ΔLR(ω)
音源周波数成分
選択部
マイクロホン２
(遠隔収音用)
スペクトル
変換部
R(ω)

図1

【図２】

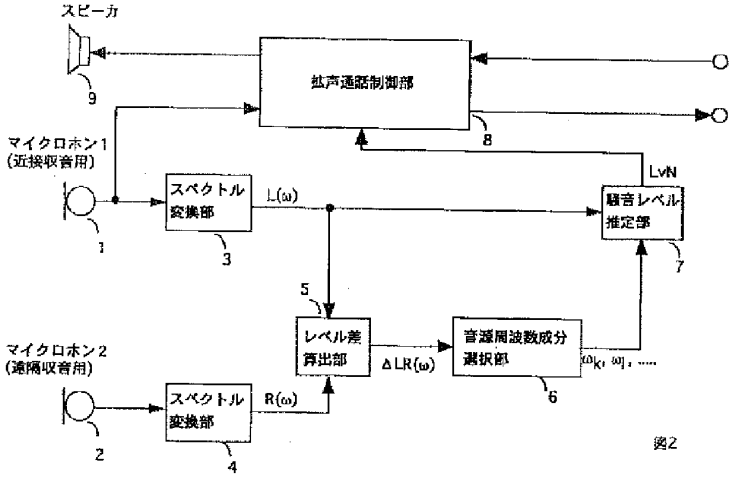

スピーカ
拡声通話制御部
マイクロホン１
(近接収音用)
スペクトル
変換部
L(ω)
LvN
騒音レベル
推定部
レベル差
算出部
ΔLR(ω)
音源周波数成分
選択部
マイクロホン２
(遠隔収音用)
スペクトル
変換部
R(ω)

図2

【図３】

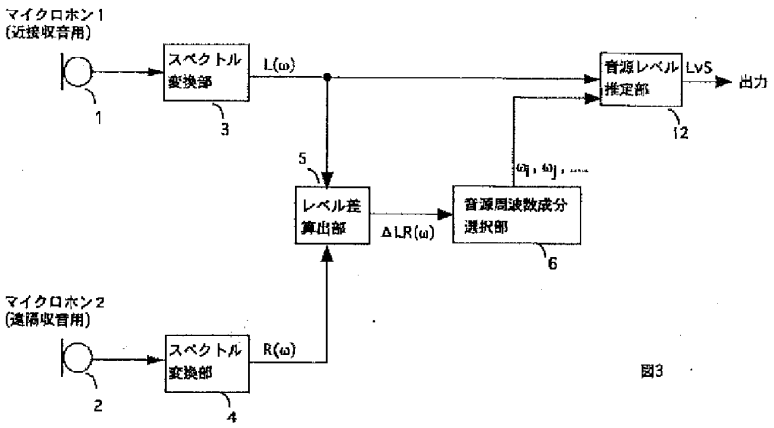

マイクロホン１
(近接収音用)
スペクトル
変換部
L(ω)
1
3
音源レベル
推定部
LvS
出力
12
5
レベル差
算出部
ΔLR(ω)
音源周波数成分
選択部
ω1, ω2, ...
6
マイクロホン２
(遠隔収音用)
スペクトル
変換部
R(ω)
2
4

図3

【図４】

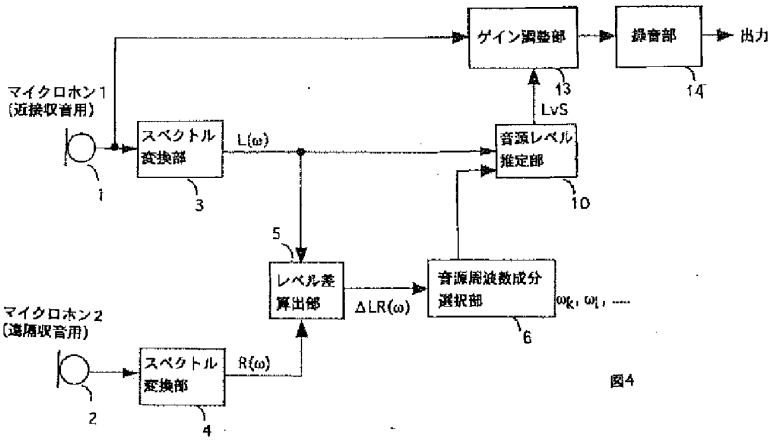

ゲイン調整部
録音部
出力
13
14
LvS
マイクロホン１
(近接収音用)
スペクトル
変換部
L(ω)
1
3
音源レベル
推定部
10
5
レベル差
算出部
ΔLR(ω)
音源周波数成分
選択部
ωk, ω1, ...
6
マイクロホン２
(遠隔収音用)
スペクトル
変換部
R(ω)
2
4

図4

【図５】

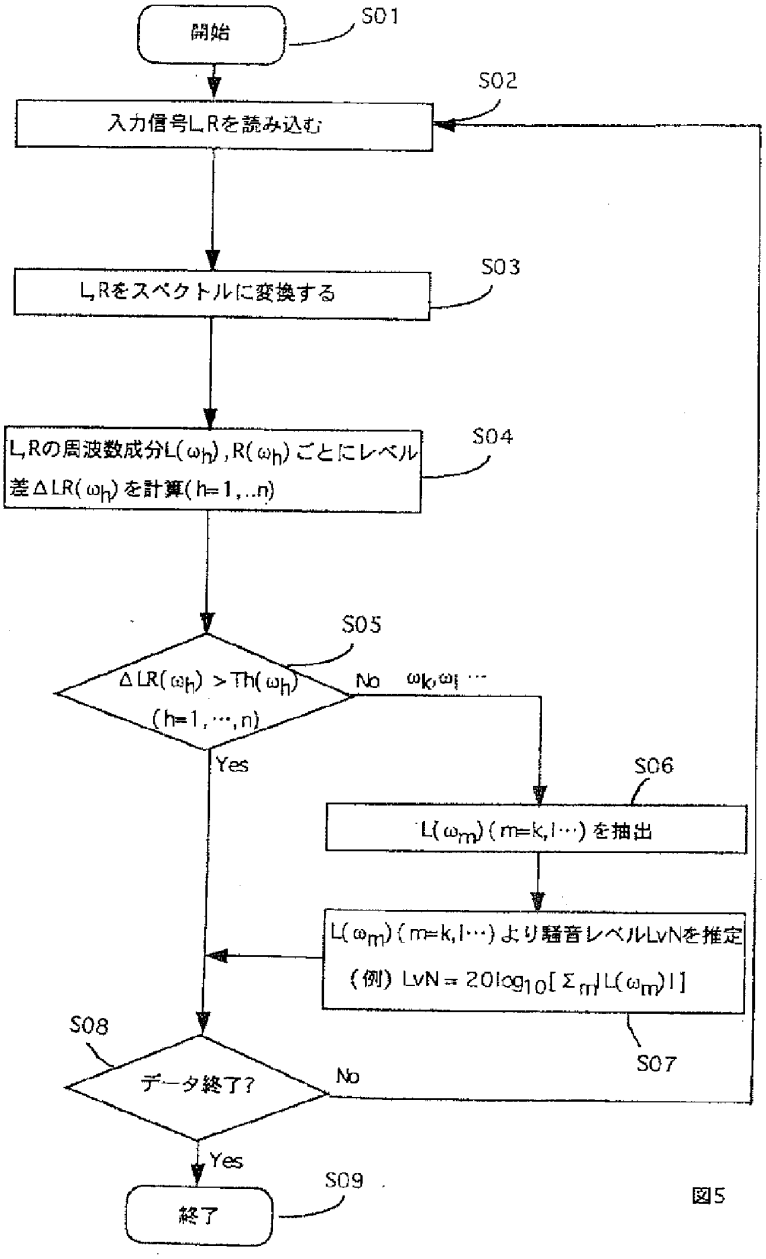

開始
S01
S02
入力信号L,Rを読み込む
S03
L,Rをスペクトルに変換する
S04
L,Rの周波数成分L(ωh),R(ωh)ごとにレベル差ΔLR(ωh)を計算(h=1,..n)
S05
ΔLR(ωh) >Th(ωh) (h=1,…,n)
No ωk,ωl …
Yes
S06
L(ωm)(m=k,l…)を抽出
L(ωm)(m=k,l…)より騒音レベルLvNを推定
(例) LvN = 20log10[Σm|L(ωm)|]
S07
S08
データ終了?
No
Yes
S09
終了

図5

【図６】

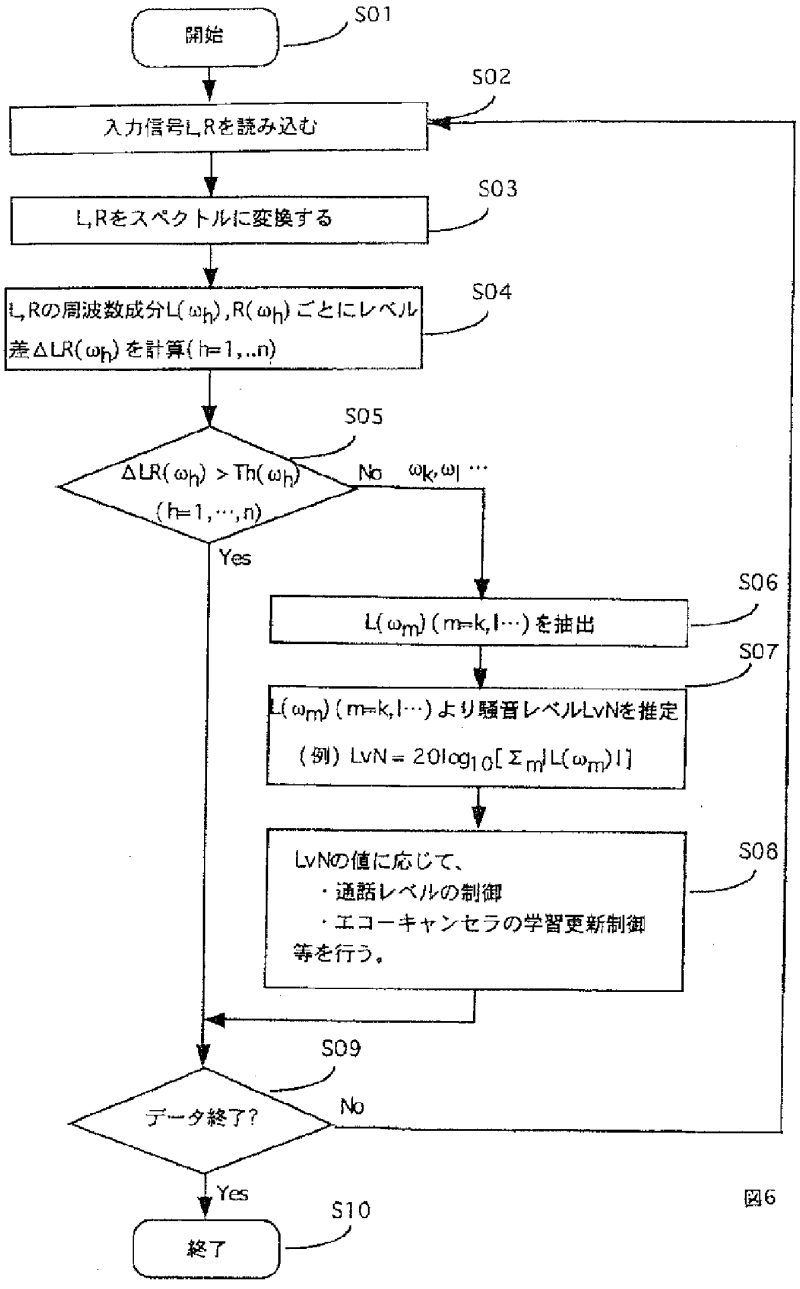

開始
S01
S02
入力信号L,Rを読み込む
S03
L,Rをスペクトルに変換する
S04
L,Rの周波数成分L(ωh),R(ωh)ごとにレベル差ΔLR(ωh)を計算(h=1,..n)
S05
ΔLR(ωh) >Th(ωh) (h=1,⋯,n)
No ωk,ωl ⋯
Yes
S06
L(ωm) (m=k,l⋯) を抽出
S07
L(ωm) (m=k,l⋯) より騒音レベルLvNを推定
(例) LvN = 20log10[Σm|L(ωm)|]
S08
LvNの値に応じて、
・通話レベルの制御
・エコーキャンセラの学習更新制御
等を行う。
S09
データ終了?
No
Yes
S10
終了

図6

【図７】

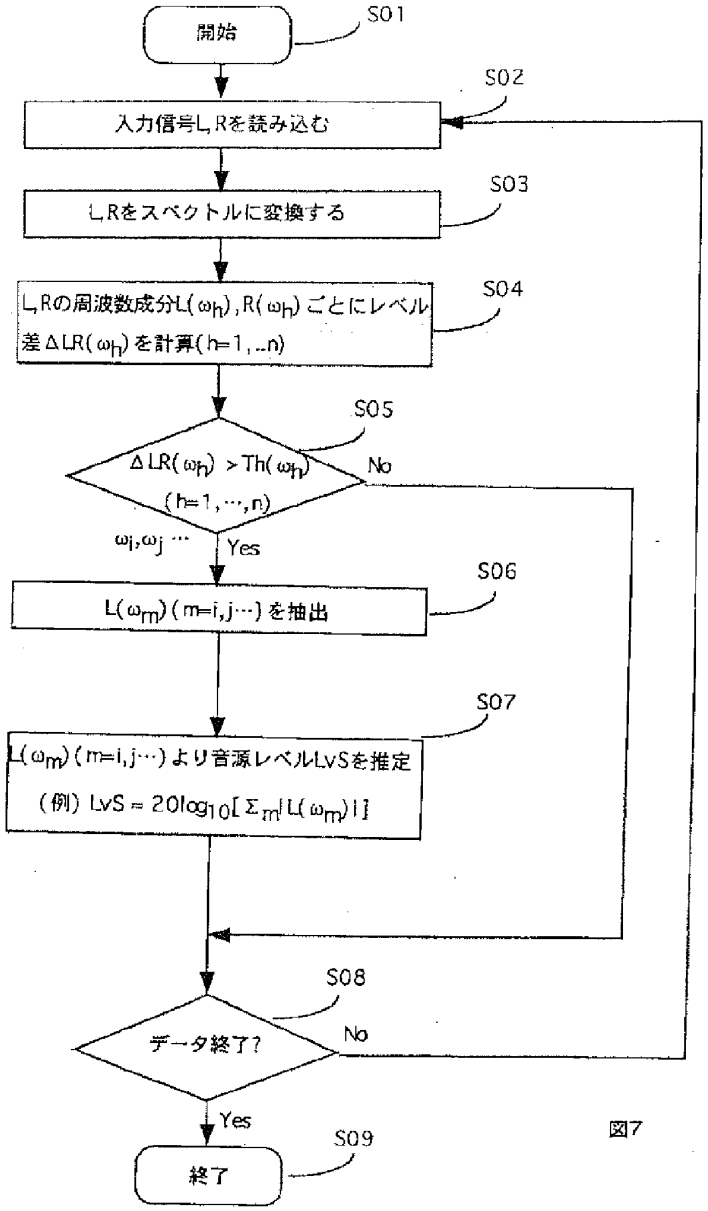

開始
S01
S02
入力信号L,Rを読み込む
S03
L,Rをスペクトルに変換する
S04
L,Rの周波数成分L(ωh),R(ωh)ごとにレベル差ΔLR(ωh)を計算(h=1,..n)
S05
ΔLR(ωh) >Th(ωh) (h=1,…,n)
No
ωi,ωj …
Yes
S06
L(ωm)(m=i,j…)を抽出
S07
L(ωm)(m=i,j…)より音源レベルLvSを推定
(例) LvS = 20log10[Σm|L(ωm)|]
S08
データ終了?
No
Yes
S09
終了

図7

【図８】

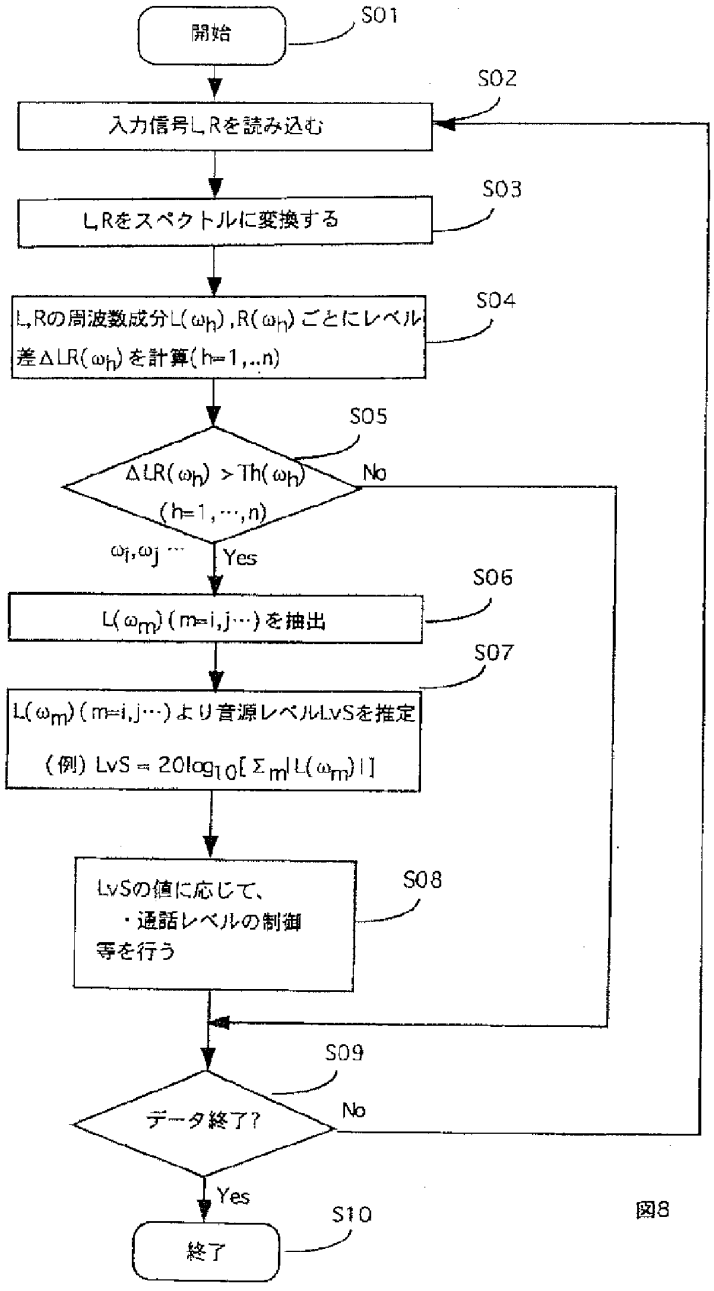

開始
S01
S02
入力信号L,Rを読み込む
S03
L,Rをスペクトルに変換する
S04
L,Rの周波数成分L(ωh),R(ωh)ごとにレベル差ΔLR(ωh)を計算(h=1,..n)
S05
ΔLR(ωh) >Th(ωh)
(h=1,…,n)
No
ωi,ωj …
Yes
S06
L(ωm)(m=i,j…)を抽出
S07
L(ωm)(m=i,j…)より音源レベルLvSを推定
(例) LvS = 20log10[Σm|L(ωm)|]
LvSの値に応じて、
・通話レベルの制御
等を行う
S08
S09
データ終了?
No
Yes
S10
終了

図8

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **10-313497**
(43)Date of publication of application : **24.11.1998**

(51)Int.Cl. **H04R 3/00**
G01S 5/18
G10L 3/02
G10L 7/04
G10L 9/00
H04S 7/00

(21)Application number : **09-252312**
(22)Date of filing : **17.09.1997**

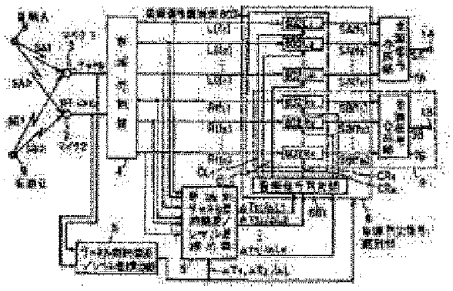
(71)Applicant : **NIPPON TELEGR & TELEPH CORP <NTT>**
(72)Inventor : **AOKI MARIKO**
AOKI SHIGEAKI
MATSUI HIROYUKI
NISHINO YUTAKA

(30)Priority

Priority number	Priority date	Priority country
08246726	**18.09.1996**	**JP**
09 76672	**13.03.1997**	**JP**
09 76695	**13.03.1997**	**JP**

(54) SOUND SOURCE SEPARATION METHOD, SYSTEM AND RECORDING MEDIUM

(57)Abstract:
PROBLEM TO BE SOLVED: To separate sound accurately into respective components and to attain real time processing.
SOLUTION: A difference $\Delta\tau$ in time required for a sound signal to travel from a sound source to microphones 1, 2 is detected from output channel signals L, R of the microphones 1, 2, where the signals L, R are divided into frequency bands L(f1)-L(fn), R(f1)-R(fn) by Fourier transform, and an arrival time difference $\Delta\tau$i (i=1, 2,...n) to the microphones 1, 2 of a corresponding frequency band of the L(f1)-L(fn), R(f1)-R(fn) and a signal level difference ΔLi are detected. Then L(f1)-L(fn), R(f1)-R(fn) are divided into low frequencies $f1<1/(2\Delta\tau)$, medium frequencies $1/(2\Delta\tau)<f1<1/\Delta\tau$, and high frequencies $fi>1/\Delta\tau$, and it is determined from which sound source the L(fi), R(fi) arrive, based on $\Delta\tau$i, in the case of low frequencies, on ΔLi, $\Delta\tau$i in the case of medium frequencies, and on ΔLi in the case of high frequencies, and the result is outputted for each sound source and the output is subject to inverse Fourier transform by each sound source and the results are synthesized.



(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号

特開平10－313497

(43)公開日　平成10年(1998)11月24日

(51)Int.Cl.[6]		識別記号	FI		
Ｈ０４Ｒ	3/00	３２０	Ｈ０４Ｒ	3/00	３２０
Ｇ０１Ｓ	5/18		Ｇ０１Ｓ	5/18	
Ｇ１０Ｌ	3/02	３０１	Ｇ１０Ｌ	3/02	３０１Ｅ
	7/04			7/04	Ａ
	9/00			9/00	Ｈ

審査請求　未請求　請求項の数52　ＯＬ　（全 33 頁）　最終頁に続く

(21)出願番号　特願平9－252312

(22)出願日　平成９年(1997)９月17日

(31)優先権主張番号　特願平8－246726
(32)優先日　平８(1996)９月18日
(33)優先権主張国　日本（ＪＰ）
(31)優先権主張番号　特願平9－76672
(32)優先日　平９(1997)３月13日
(33)優先権主張国　日本（ＪＰ）
(31)優先権主張番号　特願平9－76695
(32)優先日　平９(1997)３月13日
(33)優先権主張国　日本（ＪＰ）

(71)出願人　000004226
日本電信電話株式会社
東京都新宿区西新宿三丁目19番２号
(72)発明者　青木　真理子
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内
(72)発明者　青木　茂明
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内
(72)発明者　松井　弘行
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内
(74)代理人　弁理士　草野　卓

最終頁に続く

(54)【発明の名称】　音源分離方法、装置及び記録媒体

(57)【要約】
【課題】　正確に分離でき、かつ実時間処理を可能とする。
【解決手段】　マイクロホン１，２の出力チャネル信号Ｌ，Ｒから、音源からの音響信号がマイクロホン１，２に到達する時間差Δτを検出し、信号Ｌ，Ｒを、フーリエ変換により各周波数帯域Ｌ（ｆ１）～Ｌ（ｆｎ），Ｒ（ｆ１）～Ｒ（ｆｎ）に分割し、Ｌ（ｆ１）～Ｌ（ｆｎ）とＲ（ｆ１）～Ｒ（ｆｎ）の対応帯域のマイクロホン１，２への到達時間差Δτｉ（ｉ＝１，２，・・・ｎ）と、信号レベル差ΔＬｉとを検出し、Ｌ（ｆ１）～Ｌ（ｆｎ），Ｒ（ｆ１）～Ｒ（ｆｎ）をｆｉ＜１／（２Δτ）の低域と、１／（２Δτ）＜ｆｉ＜１／Δτの中域と、ｆｉ＞１／Δτの高域にわけ、低域ではΔτｉにより、中域ではΔＬｉとΔτｉとにより、高域ではΔＬｉによりＬ（ｆｉ），Ｒ（ｆｉ）が何れの音源から到来したか判定して音源別に出力し、この出力を各音源別に逆フーリエ変換して合成する。

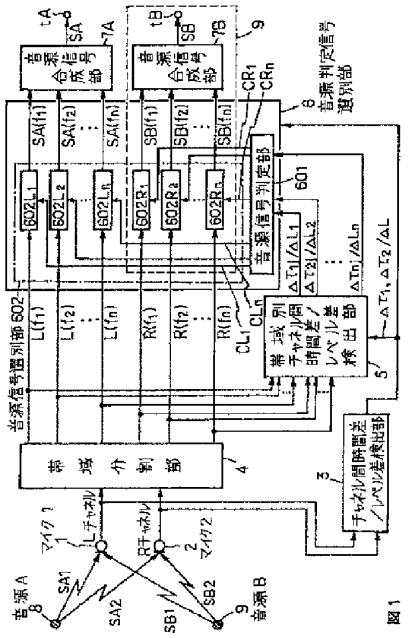


図１

1

【特許請求の範囲】

【請求項１】　互いに離して設けられた複数のマイクロホンを用いて、複数の音源から少なくとも１つの音源を分離する音源分離方法であって、

上記各マイクロホンの各出力チャネル信号を、複数の周波数帯域に分割する帯域分割過程と、

上記帯域分割過程で分割された各出力チャネル信号の各同一帯域ごとに、上記複数のマイクロホンの位置に起因して変化する、マイクロホンに到達する音響信号のパラメータの値の差を、帯域別チャネル間パラメータ値差として検出する帯域別チャネル間パラメータ値差検出過程と、

上記各帯域の帯域別チャネル間パラメータ値差にもとづき、その帯域の上記帯域分割された各出力チャネル信号の何れがいずれの音源から入力された信号であるかを判定する音源信号判定過程と、

上記音源信号判定過程の判定にもとづき、上記帯域分割された各出力チャネル信号から、同一音源から入力された信号を少なくとも１つ選択する音源信号選択過程と、

上記音源信号選択過程で同一音源からの信号として選択された複数の帯域信号を音源信号として合成する音源合成過程とを有することを特徴とする音源分離方法。

【請求項２】　請求項１記載の方法において、

上記帯域分割過程は各出力チャネル信号の各分割された帯域信号は、主として１つの音源の音響信号の成分よりなる程度に、小さく分割することを特徴とする音源分離方法。

【請求項３】　請求項１又は２記載の方法において、

上記帯域別チャネル間パラメータ値差検出過程における上記パラメータ値は音源からの音響信号が上記各マイクロホンに到達するまでの時間であり、上記帯域別チャネル間パラメータ値差は各マイクロホンに到達するまでの時間のマイクロホン間の差である帯域別チャネル間時間差であることを特徴とする音源分離方法。

【請求項４】　請求項３記載の方法において、

上記音響信号が各マイクロホンに到達するまでの時間のマイクロホン間の差をチャネル間時間差として各マイクロホンの出力チャネル信号から検出するチャネル時間差検出過程を有し、

上記音源信号判定過程は、上記各帯域別チャネル間時間差について、上記各チャネル間時間差を照合して、その帯域の上記分割された各出力チャネル信号がいずれの音源から入力された信号であるかを判定することを特徴とする音源分離方法。

【請求項５】　請求項４記載の方法において、

上記チャネル時間差検出過程は各出力チャネル信号間の相互相関を求め、相互相関の各ピークとなるその出力チャネル信号間の各時間差として上記各チャネル間時間差を求めることを特徴とする音源分離方法。

【請求項６】　請求項５記載の方法において、

2

上記帯域別チャネル間時間差は、上記各チャネル間時間差中の、上記分割された各出力チャネルの同一帯域の成分の位相差と対応する時間と最も近いものを求めて、その帯域別チャネル間時間差とすることを特徴とする音源分離方法。

【請求項７】　請求項１又は２記載の方法において、

上記帯域別チャネル間パラメータ値差検出過程における上記パラメータ値は音源からの音響信号が上記各マイクロホンに到達した時の信号レベルであり、上記帯域別チャネル間パラメータ値差は各分割された出力チャネル信号の対応帯域間のレベル差である帯域別チャネル間レベル差であることを特徴とする音源分離方法。

【請求項８】　請求項７記載の方法において、

上記各マイクロホンの出力チャネル信号間のレベル差を、チャネル間レベル差として検出するチャネル間レベル差検出過程と、

上記チャネル間レベル差と、対応する帯域別チャネル間レベル差の全てと比較する比較過程と、

その比較過程で分割帯域の所定数以上が同様の関係にあれば、上記チャネル間レベル差にもとづき、対応する出力チャネル信号の全帯域について同一の音源から入力された信号であると判定し、上記比較過程で所定値以上が同様の関係になければ、上記帯域別にいずれの音源から入力された信号であるかを判定する上記音源信号判定過程を実行することを特徴とする音源分離方法。

【請求項９】　請求項１又は２記載の方法において、

上記パラメータ値は音源からの音響信号が上記マイクロホンに到達するまでの時間と、その音響信号が到達した時の信号レベルであり、上記帯域別チャネル間パラメータ値差として帯域別チャネル間時間差と、帯域別チャネル間レベル差が求められ、

各音源からの音響信号が上記各マイクロホンに到達するまでの時間のマイクロホン間の差を、各マイクロホンの出力チャネル信号から、チャネル時間差として検出するチャネル間時間差検出過程と、

上記チャネル間時間差を基準にして上記分割された各出力チャネル信号を、低域、中域、高域の３つの周波数領域に分け領域分割過程とを有し、

上記音源信号判定過程は、

上記分割された低域の周波数帯域については、上記帯域別チャネル間時間差を利用して対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する過程と、

上記分割された中域の周波数帯域については、上記帯域別チャネル間レベル差と、上記帯域別チャネル間時間差を利用して、対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する過程と、

上記分割された高域の周波数帯域については、上記帯域別チャネル間レベル差を利用して、対応する帯域の分割

された各出力チャネル信号の何れかがいずれの音源からの入力信号であるか判定する過程とからなることを特徴とする音源分離方法。

【請求項１０】　請求項１～９の何れかに記載の方法において、

上記帯域別チャネル間パラメータ値差検出過程において、その互いに差をとるべき、もとのチャネル信号の周波数帯域が異なる場合は、その周波数帯域が互いに重ならない周波数帯域は、上記帯域別チャネル間パラメータ値差検出過程を実行せず、上記音源信号判定過程ではその信号がある帯域を予め知られている広い帯域の音源からの入力信号と判定することを特徴とする音源分離方法。

【請求項１１】　互いに離して設けられた複数のマイクロホンを用いて、複数の音源から少なくとも１つの音源を分離する音源分離方法であって、

上記各マイクロホンの各出力チャネル信号のパワースペクトルを求めるスペクトル分解過程と、

上記各チャネルごとのパワースペクトルを、主としてほぼ１つの音源の成分が含まれるように複数の周波数帯域に分割する帯域分割過程と、

上記各同一帯域ごとに、各チャネル間で分割されたパワースペクトル差を、帯域別チャネル間レベル差として検出する帯域別チャネル間レベル差検出過程と、

上記各帯域の帯域別チャネル間レベル差にもとづき、その帯域の信号が上記出力チャネル信号の何れであるかを判定する音源信号判定過程と、

上記音源信号判定過程の判定にもとづき、上記分割されたパワースペクトルから、同一音源からの信号を少なくとも１つ選択する音源信号選択過程と、

上記音源信号選択過程で同一音源からのものとして選択されたスペクトルを音源信号として合成する音源合成過程とを有することを特徴とする音源分離方法。

【請求項１２】　請求項１１記載の方法において、

上記各マイクロホンの出力チャネル信号間のレベル差をチャネル間レベル差として検出するチャネル間レベル差検出過程と、

上記チャネル間レベル差と、対応する上記帯域別チャネル間レベル差の全てとを比較する比較過程と、

その比較過程で分割帯域の所定数以上が同様の関係であれば、上記チャネル間レベル差にもとづき、対応する出力チャネル信号の全帯域について同一音源から入力された信号であると判定し、上記比較過程で所定値以上が同様の関係になければ、上記音源信号判定過程を実行することを特徴とする音源分離方法。

【請求項１３】　請求項１乃至１２の何れかの方法において、

上記各マイクロホンの出力チャネル信号を、各帯域が主として１つの音源信号成分になる程度に、複数の周波数帯域に分割する第２帯域分割過程と、

上記第２帯域分割過程で分割された各出力チャネル信号の帯域別レベルをそれぞれ検出する帯域別レベル検出過程と、

その帯域別レベル検出過程で検出された各帯域別レベルを同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出する音源状態判定過程と、

その音源状態判定過程で得た発音をしていない音源の検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する合成信号を抑圧する信号抑圧過程とを有することを特徴とする音源分離方法。

【請求項１４】　請求項１３の方法において、

上記音源状態判定過程は、上記各帯域別レベルのチャネル間での比較で、最も大きいチャネルを帯域ごとに決定する過程と、

各チャネルごとに最もレベルが大きい帯域の数を求める過程と、

上記最もレベルが大きい帯域の数が第１基準値を越えるか否か判定する第１判定過程と、

その第１判定過程で第１基準値を越えると判定すると、その越えた最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音している１個の音源を推定する過程と、

その推定された音源以外の音源を発音していないものとして検出する過程とを有することを特徴とする音源分離方法。

【請求項１５】　請求項１～１２の何れかに記載の方法において、

上記各マイクロホンの各出力チャネル信号のパワースペクトルを求めるスペクトル分解過程と、

上記各チャネルごとのパワースペクトルを、主としてほぼ１つの音源の成分が含まれるように周波数帯域を分割して帯域別レベルをそれぞれ検出する帯域別レベル検出過程と、

これら各帯域別レベルを同一帯域について比較し、最大レベルのチャネルを各帯域ごとに決定する過程と、

各チャネルごとの最大レベルの帯域の数を求める過程と、

その帯域の数が第１基準値を越えたか否かを判定する第１判定過程と、

その第１判定過程で第１基準値を越える数と判定すると、その越えたチャネルのマイクロホンが受けもつ、ゾーンから発音している１個の音源を推定する過程と、

その推定された音源以外の音源は発音していないと判定する過程と、

上記音源合成過程で合成された音源信号のうち、上記発音していないと判定された音源と対応する信号を、抑圧する信号抑圧過程とを有することを特徴とする音源分離方法。

【請求項１６】　請求項１４又は１５の方法において、
上記第１判定過程で、第１基準値を越えるものがないと判定されると、上記最もレベルが大きい帯域の数が、上記第１基準値よりも小さい第２基準値以下か否かを判定する第２判定過程と、
その第２判定過程で、第２基準値より小さいと判定されると、その小さいと判定された最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する過程とを有することを特徴とする音源分離方法。

【請求項１７】　請求項１乃至１２の何れかの方法において、
上記各マイクロホンの出力チャネル信号を、各帯域が主として１つの音源信号成分になる程度に、複数の周波数帯域に分割する第２帯域分割過程と、
上記第２帯域分離過程で分割された各出力チャネル信号のそのマイクロホンへの到達時間差を同一帯域ごとに検出する帯域別時間差検出過程と、
この帯域別時間差検出過程で検出された各帯域別到達時間差を、同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出する音源状態判定過程と、
その音源状態判定過程で得た発音をしていない音源の検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する合成信号を抑圧する信号抑圧過程とを有することを特徴とする音源分離方法。

【請求項１８】　請求項３の方法において、
上記帯域別チャネル間時間差を、同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出する音源状態判定過程と、
その音源状態判定過程で得た発音をしていない音源を検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する信号を抑圧する信号抑圧過程とを有することを特徴とする音源分離方法。

【請求項１９】　請求項１７又は１８の方法において、
上記音源状態判定過程は、上記各帯域別到達時間差比較で最も速く音源信号が到達したチャネルを帯域ごとに決定する過程と、
各チャネルごとに最も速く到達した帯域の数が第１基準値を越えるか否かを判定する第１判定過程と、
その第１判定過程が第１基準値を越えると判定すると、その越えた最も速く到達した帯域数と対応するチャネルのマイクロホン位置から発音している１個の音源を推定する過程と、
その推定された音源以外の音源を発音していないものとして検出する過程とを有することを特徴とする音源分離方法。

【請求項２０】　請求項１８の方法において、
上記第１判定過程で、第２基準値を越えるものがないと判定されると、上記最も速く到達する帯域の数が、上記第１基準値よりも小さい第２基準値より小さいか否かを判定する第２判定過程と、
その第２判定過程で、第２基準値より小さいと判定されると、その小さいと判定された最も速い到達時間の帯域数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する過程とを有することを特徴とする音源分離方法。

【請求項２１】　請求項１６又は２０の方法において、
音源が４個以上の場合で、上記第３判定過程で、第３基準値より小さいと判定されると、上記第２基準値を上記第１基準値を越えない範囲内で、順次大きくして、上記第２判定過程と同じ判定を、（Ｍ－２）回以内、Ｍは音源の数、繰返す過程を有することを特徴とする音源分離方法。

【請求項２２】　請求項１３～２１の何れかに記載の方法において、
各出力チャネル信号の全周波数成分のレベルをそれぞれ検出する全帯域レベル検出過程と、
その全帯域レベル検出過程で検出した各チャネルの全周波数成分レベルの何れもが第３基準値以下であるかを判定し、何れかが第３基準値以下でないと判定すると上記音源状態判定過程に移る第３判定過程とを有することを特徴とする音源分離方法。

【請求項２３】　請求項２２の方法において、
上記第３判定過程が第３基準値以下であると判定されると、上記音源合成過程で合成された各音源信号のすべてを抑圧する過程を有することを特徴とする音源分離方法。

【請求項２４】　請求項１３～２３の何れかの方法において、
上記帯域分割過程と上記第２帯域分割過程は同一過程として行われることを特徴とする音源分離方法。

【請求項２５】　互いに離して設けられた複数のマイクロホンを用いて、複数の音源から少なくとも１つの音源を分離する音源分離装置であって、
上記各マイクロホンの各出力チャネル信号を、主として１つの音源の音響信号の成分のみが含まれる程度に複数の周波数帯域に分割すると共に、これら分割された各出力チャネル信号の各同一帯域ごとに、上記複数のマイクロホンの位置に起因して変化する、マイクロホンに到達する音響信号のパラメータの値の差を、帯域別チャネル間パラメータ値差として検出する帯域別チャネル間パラメータ値差検出手段と、
上記各帯域の帯域別チャネル間パラメータ値差にもとづき、その帯域の上記帯域分割された各出力チャネル信号の何れがいずれの音源から入力された信号であるかを判定する音源信号判定手段と、
上記音源信号判定過程の判定にもとづき、上記帯域分割

された各出力チャネル信号から、同一音源から入力された信号を少なくとも１つ選択する音源信号選択手段と、上記音源信号選択過程で同一音源からの信号として選択された、複数の帯域信号を音源信号として合成する音源合成手段とを具備することを特徴とする音源分離装置。

【請求項２６】　請求項２５の装置において、

上記帯域別チャネル間パラメータ値差検出手段における上記パラメータ値は音源からの音響信号が上記各マイクロホンに到達するまでの時間であり、上記帯域別チャネル間パラメータ値差は各マイクロホンに到達するまでの時間のマイクロホン間の差である帯域別チャネル間時間差であることを特徴とする音源分離装置。

【請求項２７】　請求項２５の装置において、

上記音響信号が各マイクロホンに到達するまでの時間のマイクロホン間の差をチャネル間時間差として各マイクロホンの出力チャネル信号から検出するチャネル時間差検出手段を有し、

上記音源信号判定手段は、上記各帯域別チャネル間時間差について、上記各チャネル間時間差を照合して、その帯域の上記分割された各出力チャネル信号がいずれの音源から入力された信号であるかを判定する手段であることを特徴とする音源分離装置。

【請求項２８】　請求項２５の装置において、

上記帯域別チャネル間パラメータ値差検出手段における上記パラメータ値は音源からの音響信号が上記各マイクロホンに到達した時の信号レベルであり、上記帯域別チャネル間パラメータ値差は各分割された出力チャネル信号の対応帯域間のレベル差である帯域別チャネル間レベル差であることを特徴とする音源分離装置。

【請求項２９】　請求項２８の装置において、

上記各マイクロホンの出力チャネル信号間のレベル差を、チャネル間レベル差として検出するチャネル間レベル差検出手段と、

上記チャネル間レベル差と、対応する帯域別チャネル間レベル差の全てと比較する比較手段と、その比較手段で分割帯域の所定数以上が同様の関係にあれば、上記チャネル間レベル差にもとづき、対応する出力チャネル信号の全帯域について同一の音源から入力された信号であると判定し、上記比較手段で所定値以上が同様の関係になければ、上記帯域別にいずれの音源から入力された信号であるかを判定する上記音源信号判定手段を実行する手段を含むことを特徴とする音源分離装置。

【請求項３０】　請求項２５の装置において、

上記パラメータ値は音源からの音響信号が上記マイクロホンに到達するまでの時間と、その音響信号が到達した時の信号レベルであり、上記帯域別チャネル間パラメータ値差として帯域別チャネル間時間差と、帯域別チャネル間レベル差が求められ、

各音源からの音響信号が上記各マイクロホンに到達するまでの時間のマイクロホン間の差と、各マイクロホンの出力チャネル信号から、チャネル時間差として検出するチャネル間時間差検出手段と、

上記チャネル間時間差を基準にして、上記分割された各出力チャネル信号を、低域、中域、高域の３つの周波数領域に分ける領域分割手段とを有し、

上記音源信号判定手段は、

上記分割された低域の周波数帯域については、上記帯域別チャネル間時間差を利用して対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する手段と、

上記分割された中域の周波数帯域については、上記帯域別チャネル間レベル差と、上記帯域別チャネル間時間差を利用して、対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する手段と、

上記分割された高域の周波数帯域については、上記帯域別チャネル間レベル差を利用して、対応する帯域の分割された各出力チャネル信号の何れかがいずれの音源からの入力信号であるか判定する手段とからなることを特徴とする音源分離装置。

【請求項３１】　請求項２５乃至３０の何れかの装置において、

上記帯域分割された各出力チャネル信号の帯域別レベルをそれぞれ検出する帯域別レベル検出手段と、

その帯域別レベル検出手段が検出された各帯域別レベルを同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出する音源状態判定手段と、

その音源状態判定手段で得た発音をしていない音源の検出信号により、上記音源合成手段で合成された音源信号のうち、上記発音していない音源と対応する信号を抑圧する信号抑圧手段とを有することを特徴とする音源分離装置。

【請求項３２】　請求項３１の装置において、

上記音源状態判定手段は、上記各帯域別レベルのチャネル間での比較で、最も大きいチャネルを帯域ごとに決定する手段と、

各チャネルごとに最もレベルが大きい帯域の数を求める手段と、

上記最もレベルが大きい帯域の数が第１基準値を越えるか否か判定する第１判定手段と、

その第１判定手段で第１基準値を越えると判定すると、その越えた最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音している１個の音源を推定する手段と、

その推定された音源以外の音源を発音していないものとして検出する手段とを有することを特徴とする音源分離装置。

【請求項３３】　請求項３１の装置において、

上記第１判定手段で、第１基準値を越えるものがないと

判定されると、上記最もレベルが大きい帯域の数が、上記第１基準値よりも小さい第２基準値以下か否かを判定する第２判定手段と、

その第２判定手段で、第２基準値より小さいと判定されると、その小さいと判定された最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する手段とを有することを特徴とする音源分離装置。

【請求項３４】　請求項２５乃至３０の何れかの装置において、

上記帯域分割された各出力チャネル信号のそのマイクロホンへの到達時間差を同一帯域ごと検出する帯域別時間差検出手段と、

この帯域別時間差検出手段で検出された各帯域別到達時間差を、同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出する音源状態判定手段と、

その音源状態判定手段で得た発音をしていない音源を検出信号により、上記音源合成手段で合成された音源信号のうち、上記発音していない音源と対応する信号を抑圧する信号抑圧手段とを有することを特徴とする音源分離装置。

【請求項３５】　請求項３４の装置において、

上記音源状態判定手段は、上記各帯域別到達時間差比較で最も速く音源信号が到達したチャネルを帯域ごとに決定する手段と、

各チャネルごとに最も速く到達した帯域の数が第１基準値を越えるか否かを判定する第１判定手段と、

その第１判定手段が第２基準値を越えると判定すると、その越えた最も速く到達した帯域数と対応するチャネルのマイクロホン位置から発音している１個の音源を推定する手段と、

その推定された音源以外の音源を発音していないものとして検出する手段とを有することを特徴とする音源分離装置。

【請求項３６】　請求項３５の装置において、

上記第１判定手段で、第１基準値を越えるものがないと判定されると、上記最も速く到達する帯域の数が、上記第１基準値よりも小さい第２基準値以下か否かを判定する第２判定手段と、

その第２判定手段で、第２基準値より小さいと判定されると、その小さいと判定された最も速い到達時間の帯域数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する手段とを有することを特徴とする音源分離装置。

【請求項３７】　請求項３１～３６の何れかに記載の装置において、

各出力チャネル信号の全周波数成分のレベルをそれぞれ検出する全帯域レベル検出手段と、

その全帯域レベル検出手段で検出した各チャネルの全周波数成分レベルの何れもが第３基準値以下であるかを判定し、何れかが第１基準値以下でないと判定すると、上記音源状態判定手段に移る第３判定手段とを有することを特徴とする音源分離装置。

【請求項３８】　互いに離して設けられた複数のマイクロホンを用いて、複数の音源から少なくとも１つの音源を分離する下記過程を有する音源分離方法のプログラムを記録した記録媒体であって、

上記各マイクロホンの各出力チャネル信号を、主に１つの音源の音響信号の成分のみを含む程度に複数の周波数帯域に分割すると共にこれら分割された各出力チャネル信号の各同一帯域ごとに、上記複数のマイクロホンの位置に起因して変化する、マイクロホンに到達する音響信号のパラメータの値の差を、帯域別チャネル間パラメータ値差として検出する帯域別チャネル間パラメータ値差検出過程と、

上記各帯域の帯域別チャネル間パラメータ値差にもとづき、その帯域の上記帯域分割された各出力チャネル信号の何れがいずれの音源から入力された信号であるかを判定する音源信号判定過程と、

上記音源信号判定過程の判定にもとづき、上記帯域分割された各出力チャネル信号から、同一音源から入力された信号を少なくとも１つ選択する音源信号選択過程と、

上記音源信号選択過程で同一音源からの信号として選択された複数の帯域信号を音源信号として合成する音源合成過程とを有するコンピュータにより読出し可能な記録媒体。

【請求項３９】　請求項３８の記録媒体において、

上記帯域別チャネル間パラメータ値差検出過程における上記パラメータ値は音源からの音響信号が上記各マイクロホンに到達するまでの時間であり、上記帯域別チャネル間パラメータ値差は各マイクロホンに到達するまでの時間のマイクロホン間の差である帯域別チャネル間時間差であって、

上記プログラムは上記音響信号が各マイクロホンに到達するまでの時間のマイクロホン間の差をチャネル間時間差として各マイクロホンの出力チャネル信号から検出するチャネル時間差検出過程を有し、

上記音源信号判定過程は、上記各帯域別チャネル間時間差について、上記各チャネル間時間差を照合して、その帯域の上記分割された各出力チャネル信号がいずれの音源から入力された信号であるかを判定することを特徴とする記録媒体。

【請求項４０】　請求項３９の記録媒体において、

上記チャネル時間差検出過程は各出力チャネル信号間の相互相関を求め、相互相関の各ピークとなる、その出力チャネル信号間の各時間差として上記各チャネル間時間差を求めることを特徴とする記録媒体。

【請求項４１】　請求項４０の記録媒体において、

上記帯域別チャネル間時間差は、上記各チャネル間時間

差中の、上記分割された各出力チャネルの同一帯域の成分の位相差と対応する時間と最も近いものを求めて、その帯域別チャネル間時間差とすることを特徴とする記録媒体。

【請求項４２】　請求項３８の記録媒体において、

上記帯域別チャネル間パラメータ値差検出過程における上記パラメータ値は音源からの音響信号が上記各マイクロホンに到達した時の信号レベルであり、上記帯域別チャネル間パラメータ値差は各分割された出力チャネル信号の対応帯域間のレベル差である帯域別チャネル間レベル差であって、

上記プログラムは上記各マイクロホンの出力チャネル信号間のレベル差を、チャネル間レベル差として検出するチャネル間レベル差検出過程と、

上記チャネル間レベル差と、対応する帯域別チャネル間レベル差の全てと比較する比較過程と、その比較過程で分割帯域の所定数以上が同様の関係にあれば、上記チャネル間レベル差にもとづき、対応する出力チャネル信号の全帯域について同一の音源から入力された信号であると判定し、上記比較過程で所定値以上が同様の関係になければ、上記帯域別にいずれの音源から入力された信号であるかを判定する、上記音源信号判定過程と実行する過程とを有することを特徴とする記録媒体。

【請求項４３】　請求項３８の記録媒体において、

上記プログラムは上記パラメータ値は音源からの音響信号が上記マイクロホンに到達するまでの時間と、その音響信号が到達した時の信号レベルであり、上記帯域別チャネル間パラメータ値差として帯域別チャネル間時間差と、帯域別チャネル間レベル差が求められ、

各音源からの音響信号が上記各マイクロホンに到達するまでの時間のマイクロホン間の差を、各マイクロホンの出力チャネル信号から、チャネル時間差として検出するチャネル間時間差検出過程と、

上記チャネル間時間差を基準にして、上記分割された各出力チャネル信号を、低域、中域、高域の３つの周波数領域に分ける領域分割過程とを有し、

上記音源信号判定過程は、

上記分割された低域の周波数帯域については、上記帯域別チャネル間時間差を利用して対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する過程と、

上記分割された中域の周波数帯域については、上記帯域別チャネル間レベル差と、上記帯域別チャネル間時間差を利用して、対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する過程と、

上記分割された高域の周波数帯域については、上記帯域別チャネル間レベル差を利用して、対応する帯域の分割された各出力チャネル信号の何れがいずれの音源からの入力信号であるか判定する過程とからなることを特徴とする記録媒体。

【請求項４４】　請求項３８乃至４３の何れかの記録媒体において、

上記プログラムは上記帯域分割された各出力チャネル信号の帯域別レベルをそれぞれ検出する帯域別レベル検出過程と、

その帯域別レベル検出過程が検出された各帯域別レベルを同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出する音源状態判定過程と、

その音源状態判定過程で得た発音をしていない音源の検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する信号を抑圧する信号抑圧過程とを有することを特徴とする記録媒体。

【請求項４５】　請求項４４の記録媒体において、

上記音源状態判定過程は、上記各帯域別レベルのチャネル間での比較で、最も大きいチャネルを帯域ごとに決定する過程と、

各チャネルごとに最もレベルが大きい帯域の数を求める過程と、

上記最もレベルが大きい帯域の数が第１基準値を越えるか否か判定する第１判定過程と、

その第１判定過程で第１基準値を越えると判定すると、その越えた最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音している１個の音源を推定する過程と、

その推定された音源以外の音源を発音していないものとして検出する過程とを有することを特徴とする記録媒体。

【請求項４６】　請求項４５の記録媒体において、

上記プログラムは上記第１判定過程で、第１基準値を越えるものがないと判定されると、上記最もレベルが大きい帯域の数が、上記第１基準値よりも小さい第２基準値以下か否かを判定する第３判定過程と、

その第２判定過程で、第２基準値より小さいと判定されると、その小さいと判定された最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する過程とを有することを特徴とする記録媒体。

【請求項４７】　請求項３８乃至４３の何れかの記録媒体において、

上記プログラムは上記帯域分割された各出力チャネル信号のそのマイクロホンへの到達時間差を同一帯域ごと検出する帯域別時間差検出過程と、

この帯域別時間差検出過程で検出された各帯域別到達時間差を、同一帯域についてチャネル間で比較した結果にもとづき、発音をしていない音源を検出する音源状態判定過程と、

その音源状態判定過程で得た発音をしていない音源を検

出した検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する信号を抑圧する信号抑圧過程とを有することを特徴とする記録媒体。

【請求項４８】　請求項４７の記録媒体において、

上記音源状態判定過程は、上記各帯域別到達時間差比較で、最も速く音源信号が到達したチャネルを帯域ごとに決定する過程と、

各チャネルごとに最も速く到達した帯域の数が第１基準値を越えるか否かを判定する第１判定過程と、

その第１判定過程が第２基準値を越えると判定すると、その越えた最も速く到達した帯域数と対応するチャネルのマイクロホン位置から発音している１個の音源を推定する過程と、

その推定された音源以外の音源を発音していないものとして検出する過程とを有することを特徴とする記録媒体。

【請求項４９】　請求項４８の記録媒体において、

上記プログラムは上記第１判定過程で、第１基準値を越えるものがないと判定されると、上記最も速く到達する帯域の数が、上記第１基準値よりも小さい第２基準値より小さいか否かを判定する第２判定過程と、

その第３判定過程で、第２基準値より小さいと判定されると、その小さいと判定された最も速い到達時間の帯域数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する過程とを有することを特徴とする記録媒体。

【請求項５０】　請求項４６又は４９の記録媒体において、

音源が４個以上の場合で、上記プログラムは上記第２判定過程で、第２基準値より小さいと判定されると、上記第２基準値を上記第１基準値を越えない範囲内で、順次大きくして、上記第２判定過程と同じ判定を（Ｍ－２）以内、Ｍは音源の数、繰返す過程を有することを特徴とする記録媒体。

【請求項５１】　請求項４４～５０の何れかに記載の記録媒体において、

上記プログラムは各出力チャネル信号の全周波数成分のレベルをそれぞれ検出する全帯域レベル検出過程と、

その全帯域レベル検出過程で検出した各チャネルの全周波数成分レベルの何れもが第３基準値以下であるかを判定し、何れかが第３基準値以下でないと判定すると、上記音源状態判定過程に移る第３判定過程とを有することを特徴とする記録媒体。

【請求項５２】　請求項４７～５０の何れかに記載の記録媒体において、上記プログラムは上記チャネル時間差検出過程は上記帯域時間差検出過程を兼ねていることを特徴とする記録媒体。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】この発明は音声信号源や各種環境音源などの複数の音源から発せられた複数の音響信号が混ざった信号から少なくとも１つの音源の信号を分離抽出する方法、その方法に用いた音源分離装置、およびその方法をコンピュータにより実行するためのプログラを記録した記録媒体に関する。

【０００２】この種の音源分離装置は、例えばテレビ会議における収音装置、騒音環境下で発声した音声信号の伝送のための収音装置、音源の種類を識別する装置の収音装置など各種のものに適用される。従来の音源分離技術は、周波数領域において各信号の基本周波数を推定し、調波構造を抜き出すことにより、同一音源からの成分を集めて合成する方法が用いられてきた。

【０００３】しかしこの方法では、（１）分離可能な信号が、音声の母音や楽音のような調波構造を持つものに限定されるという問題があった、（２）基本周波数の推定は一般に長い処理時間を必要とするため、実時間で音源を分離することは困難であった、（３）調波構造の推定誤りなどにより、抽出された信号に他の音源の周波数成分が混じり、それが雑音として知覚されるため分離精度が不十分であった。

【０００４】

【発明の解決しようとする課題】この発明の目的は調波構造を持たない音源の音響信号でも分離抽出することとし、つまり音源の種類に依存することなく音源分離を可能とし、かつ実時間での音源分離を可能とする方法、装置、及びプログラム記録媒体を提供することにある。

【０００５】この発明の他の目的は分離精度が高く、雑音の混入が少ない音源分離方法、装置及びプログラム記録媒体を提供することにある。

【０００６】

【課題を解決するための手段】この発明の音源分離方法は互いに離して設けられた複数のマイクロホンを用い、上記各マイクロホンの各出力チャネル信号を、帯域分割過程で複数の周波数帯域に分割し、その各帯域には主として１つの音源信号成分のみ存在するようにし、これら分割された各出力チャネル信号の各同一帯域ごとに、上記複数のマイクロホンの位置に起因して変化する、マイクロホンに到達する音響信号のパラメータ、つまりレベル（パワー）、到達時間の値の差を、帯域別チャネル間パラメータ値差として検出し、上記各帯域の帯域別チャネル間パラメータ値差にもとづき、その帯域の上記帯域分割された各出力チャネル信号の何れがいずれの音源から入力された信号であるかを音源信号判定過程で判定し、この音源信号判定過程の判定にもとづき、上記帯域分割された各出力チャネル信号から、同一音源から入力された信号を少なくとも１つ、音源信号選択過程で選択し、その音源信号選択過程で同一音源からの信号として選択された、複数の帯域信号を音源信号として音源合成過程で合成する。

【０００７】この発明の音源分離方法の実施例によれば、上記帯域分割過程で分割された各出力チャネル信号の帯域別レベルをそれぞれ検出し、これらが検出された各帯域別レベルを同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出し、その発音をしていない音源の検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する合成信号を抑圧する。

【０００８】この発明の音源分離方法の他の実施例によると、上記帯域分離過程で分割された各出力チャネル信号のそのマイクロホンへの到達時間差を同一帯域ごとに検出し、これら検出された各帯域別到達時間差を、同一帯域についてチャネル間で比較した結果にもとづき発音をしていない音源を検出し、その発音をしていない音源の検出信号により、上記音源合成過程で合成された音源信号のうち、上記発音していない音源と対応する合成信号を抑圧する。

【０００９】

【発明の実施の形態】図１にこの発明の実施例を示す。マイクロホン１，２が間隔、例えば２０ｃｍ程度をあけて配され、これらマイクロホン１，２はそれぞれ音源Ａ，Ｂからの音響信号を収集して電気信号に変換する。マイクロホン１の出力をＬチャネル信号と、マイクロホン２の出力をＲチャネル信号と称する。Ｌチャネル信号とＲチャネル信号はチャネル間時間差／レベル差検出部３と、帯域分割部４へ供給され、帯域分割部４ではそれぞれ複数の周波数帯域信号に分割されて帯域別チャネル間時間差／レベル差検出部５と音源判定信号選別部６へ供給される。検出部３，５の各検出出力に応じて選別部６において各帯域ごとに何れかのチャネル信号がＡ成分又はＢ成分として選別され、これら選択された帯域ごとのＡ成分信号、Ｂ成分信号はそれぞれ音源信号合成部７Ａ，７Ｂでそれぞれ合成されて、音源Ａ信号と音源Ｂ信号とに分離出力される。音源Ａがマイクロホン２よりマ＊イクロホン１に近いと、音源Ａよりマイクロホン１に到達する信号ＳＡ１は音源Ａよりマイクロホン２に到達する信号ＳＡ２より早く到達し、かつレベルが大きい、また音源Ｂがマイクロホン１よりマイクロホン２に近いと、音源Ｂからマイクロホン１，２にそれぞれ到達する信号ＳＢ１，ＳＢ２は後者が早くマイクロホン２に到達し、レベルも大きい。このようにこの発明では、音源のマイクロホン１，２に対する位置に起因する両マイクロホン１，２に到達する音響信号の変化量、この例では両信号の到達時間差とレベル差を利用する。

【００１０】図１に示した装置は以下に示すように動作する。図２に示すように、マイクロホン１，２に２つの音源Ａ，Ｂからの信号が取り込まれる（Ｓ０１）。チャネル間時間差／レベル差検出部３は、Ｌチャネル信号とＲチャネル信号からチャネル間時間差またはレベル差を検出する。時間差の検出に用いるパラメータとしては、Ｌチャネル信号とＲチャネル信号との相互相関関数を用いた場合で説明する。図３に示すようにまず、Ｌチャネル信号とＲチャネル信号との各サンプルＬ（ｔ），Ｒ（ｔ）を読み込み（Ｓ０２）、これらサンプル間の相互相関関数を算出する（Ｓ０３）。この算出は両チャネル信号が同一サンプル時点についての相互相関を求め、また一方のチャネル信号に対し他方のチャネル信号をサンプル時点を１つだけずらした場合、２つだけずらした場合・・・の各場合の相互相関をそれぞれ求めて相互相関関数を求める。これら相互相関を多数求め、これらをパワーで正規化したヒストグラムを作成する（Ｓ０４）。次に、ヒストグラムの累積度数順位第一位、第二位をそれぞれとる時点差$\Delta\alpha_1$，$\Delta\alpha_2$を求める（Ｓ０５）。これらの時点差$\Delta\alpha_1$，$\Delta\alpha_2$を、次式によりそれぞれチャネル間時間差$\Delta\tau_1$，$\Delta\tau_2$に変換して出力する（Ｓ０６）。

【００１１】

$$\Delta\tau_1 = 1000\times\Delta\alpha_1 / F \quad (1)$$

$$\Delta\tau_2 = 1000\times\Delta\alpha_2 / F \quad (2)$$

ただしＦはサンプリング周波数であり、１０００倍にするのは演算の便宜上値をある程度大きくするためである。時間差$\Delta\tau_1$，$\Delta\tau_2$は、音源Ａ，Ｂそれぞれの信号のＬチャネル信号とＲチャネル信号のチャネル間時間差である。

【００１２】図１、２の説明に戻って帯域分割部４はＬチャネル信号とＲチャネル信号をそれぞれ各周波数帯域の信号Ｌ（ｆ１），Ｌ（ｆ２），…，（ｆｎ）と、信号Ｒ（ｆ１），Ｒ（ｆ２），…，（ｆｎ）に分割する（Ｓ０４）。この分割は例えば各チャネル信号をそれぞれ離散的フーリエ変換して周波数領域信号に変換した後、各周波数帯域に分割することにより行う。この帯域分割は、音源Ａ，Ｂの各信号の周波数特性の差から各帯域において、一方の音源の信号成分のみが主として存在する程度、音声信号の場合は、例えば２０Ｈｚ帯域幅で分割する。音源Ａのパワースペクトルが例えば図４Ａに示すように得られ、音源Ｂのパワースペクトルが図４Ｂに示すように得られ、この各スペクトルが分離できる程度の帯域幅Δｆで分割する。この時、例えば破線で対応するスペクトルを示すように、一方の音源のスペクトルに対し他方の音源のスペクトルは無視できる。またこの図４Ａ、４Ｂから理解されるように帯域幅２Δｆで分離してもよい。つまり、各帯域に１本のスペクトルのみが含まれるようにしなくてもよい。なお、離散的フーリエ変換は例えば２０～４０ｍｓごとに行う。

【００１３】次に、帯域別チャネル間時間差／レベル差検出部５は、例えばＬ（ｆ１）とＲ（ｆ１），…Ｌ（ｆｎ）とＲ（ｆｎ）といった各対応する帯域信号のチャネ

ル間について、帯域別チャネル間時間差またはレベル差を検出する（Ｓ０５）。ここで、帯域別チャネル間時間差は、チャネル間時間差検出部３で検出したチャネル間＊

$$\Delta\tau_1 - \{(\Delta\phi i/(2\pi f i)) + (k i 1/f i)\} = \varepsilon_i 1 \quad (3)$$

$$\Delta\tau_2 - \{(\Delta\phi i/(2\pi f i)) + (k i 2/f i)\} = \varepsilon_i 2 \quad (4)$$

ｉ＝１，２，…，ｎ、Δφｉは信号Ｌ（ｆｉ）と信号Ｒ（ｆｉ）の位相差である。これら式で$\varepsilon_i 1$，$\varepsilon_i 2$が最小になるように整数ｋｉ１，ｋｉ２を決める。次に、その最小値の$\varepsilon_i 1$と$\varepsilon_i 2$とを比べて小さい方のチャネル時間差$\Delta\tau_j$（ｊ＝１，２）を、その帯域ｉのチャネル間時間差$\Delta\tau_{ij}$とする。つまり一方の音源信号のその帯域でのチャネル間時間差とする。

【００１５】音源判定信号選別部６は、帯域別チャネル間時間差／レベル差検出部５で検出された帯域別チャネル間時間差$\Delta\tau_{ij}$～τ_{nj}を用いて各帯域信号Ｌ（ｆ１）～Ｌ（ｆｎ）とＲ（ｆ１）～Ｒ（ｆｎ）との各対応するものについて何れを選択するか判定を音源信号判定部６０１で行う（Ｓ０６）。例えば、チャネル間時間差／レベル差検出部３で算出された時間差$\Delta\tau_1$，$\Delta\tau_2$のうち、$\Delta\tau_1$が、Ｌ側のマイクロホンに近い、音源Ａからの信号のチャネル間時間差であり、$\Delta\tau_2$が、Ｒ側のマイクロホンに近い、音源Ｂからの信号のチャネル間時間差である場合で説明する。

【００１６】この場合、帯域別チャネル間時間差／レベル差検出部５で算出された時間差$\Delta\tau_{ij}$が$\Delta\tau_1$である帯域ｉは、音源信号判定部６０１によりゲート６０２Ｌｉが開とされてＬ側の入力信号Ｌ（ｆｉ）がそのままＳＡ（ｆｉ）として出力され、Ｒ側の帯域ｉの入力信号Ｒ（ｆｉ）は音源信号判定部６０１によりゲート６０２Ｒが閉とされてＳＢ（ｆｉ）は０として出力される。時間差$\Delta\tau_{ij}$が$\Delta\tau_2$となる帯域ｉは、逆に、Ｌ側は信号Ｌ（ｆｉ）はＳＡ（ｆｉ）＝０として出力され、Ｒ側は入力信号Ｒ（ｆｉ）がそのままＳＢ（ｆｉ）として出力される。つまり図１に示すように帯域信号Ｌ（ｆ１）～Ｌ（ｆｎ）はそれぞれゲート６０２Ｌ１～６０２Ｌｎを通じて音源信号合成部７Ａへ供給され、帯域信号Ｒ（ｆ１）～Ｒ（ｆｎ）はそれぞれゲート６０２Ｒ１～６０２Ｒｎを通じて音源信号合成部７へ供給される。音源判定信号選別部６内の音源信号判定部６０１では$\Delta\tau_{ij}$～$\Delta\tau_{nj}$が入力され、$\Delta\tau_{ij}$が$\Delta\tau_1$と判定された帯域ｉについてはゲート制御信号ＣＬｉ＝１とＣＲｉ＝０が生成され、対応するゲート６０２Ｌｉが開、６０２Ｒｉが閉にそれぞれ制御され、$\Delta\tau_{ij}$が$\Delta\tau_2$と判定された帯域ｉについてはゲート制御信号ＣＬｉ＝０と、ＣＲｉ＝１が生成され、対応するゲート６０２Ｌｉが閉、６０２Ｒｉが開にそれぞれ制御される。以上の説明は機能構成であって、実際には例えばデジタルシグナルプロセッサにより処理される。

【００１７】音源信号合成部７Ａで信号ＳＡ（ｆｉ）～ＳＡ（ｆｎ）が合成され、前記帯域分割の例ではそれぞれ逆フーリエ変換され、信号ＳＡとして出力端子t_Aに出力され、また音源信号合成部７Ｂで信号ＳＢ（ｆｉ）～ＳＢ（ｆｎ）が同様に合成されて信号ＳＢとして出力端子t_Bに出力される。以上の説明で明らかなように、この発明装置においては、各チャネル信号の細かく帯域分割した、各帯域成分がそれぞれどの音源からのものであるかを判定し、判定された成分は全て出力する、すなわち、音源Ａ，Ｂの信号の周波数成分が互いに重なっていなければ、特定の周波数帯域を欠落させることなく処理を行うため、調波構造のみ抜き出す従来の方法に比べて音質を高く保ったまま音源Ａ，Ｂの各信号を分離することが可能である。

【００１８】以上の説明は、チャネル間時間差／レベル差検出部３及び帯域別チャネル間時間差／レベル差検出部５で検出した、チャネル間時間差と、帯域別チャネル間時間差のみを利用して、音源判定信号部６０１で判定条件を決定した。次にこの判定条件の決定をチャネル間のレベル差を用いて処理する実施例を説明する。この実施例は図５に示すようにマイクロホン１，２からＬチャネル信号とＲチャネル信号を取込み（Ｓ０２）、これらＬチャネル信号とＲチャネル信号のチャネル間レベル差ΔＬをチャネル間時間差／レベル差検出部３（図１）で検出する（Ｓ０３）。図２中のステップＳ０４と同様に、Ｌチャネル信号、Ｒチャネル信号をそれぞれｎ個の帯域別チャネル信号Ｌ（ｆ１）～Ｌ（ｆｎ），Ｒ（ｆ１）～Ｒ（ｆｎ）に分割し（Ｓ０４）、帯域別チャネル信号Ｌ（ｆ１）～Ｌ（ｆｎ）とＲ（ｆ１）～Ｒ（ｆｎ）との対応帯域、つまりＬ（ｆ１）とＲ（ｆ１），Ｌ（ｆ２）とＲ（ｆ２），…，Ｌ（ｆｎ）とＲ（ｆｎ）について帯域別チャネル間レベル差ΔＬ１，ΔＬ２，…，ΔＬｎを検出する（Ｓ０５）。

【００１９】人間の音声は、２０ｍｓ～４０ｍｓ程度の間は定常状態とみなすことが出来る。そのため、音源信号判定部６０１（図１）においては、２０ｍｓ～４０ｍｓ毎に、チャネル間レベル差ΔＬの対数を取った値の符号と、帯域別チャネル間レベル差ΔＬｉの対数を取った値の符号とが、全帯域のうち何割以上の帯域で、同じ符号（＋又は－）になるのかを算出し、所定値、例えば８割以上の帯域で両者が同じ符号を持てば（Ｓ０６，Ｓ０７）、そこから２０ｍｓ～４０ｍｓの間はチャネル間レベル差ΔＬのみで判定し（Ｓ０８）、同じ符号を持つのが８割以下の帯域であれば、そこから２０ｍｓ～４０ｍｓの間は帯域毎に、帯域別チャネル間レベル差ΔＬｉを用いて判定する（Ｓ０９）。判定の仕方は、全帯域をチャネル間レベル差ΔＬで判定する場合は、ΔＬが正であ

＊時間差$\Delta\tau_1$，$\Delta\tau_2$を利用することにより一意的に検出される。この検出に用いる式は以下のとおりである。

【００１４】

19

れば、Ｌチャネル信号Ｌ（ｔ）がそのまま信号ＳＡとして出力され、Ｒチャネル信号Ｒ（ｔ）は信号ＳＢ＝０として出力される。ΔＬが０以下であれば逆に、Ｌチャネル信号Ｌ（ｔ）は信号ＳＡ＝０として出力され、Ｒチャネル信号Ｒ（ｔ）がそのまま信号ＳＢとして出力される。ただし、これは、チャネル間レベル差としてＬ側からＲ側を引いた値を用いた場合の説明である。また、帯域別チャネル間レベル差ΔＬｉを用いて帯域毎に判定する場合は、各帯域ｆｉごとに帯域別チャネル間レベル差ΔＬｉが正であれば、Ｌ側分割信号Ｌ（ｆｉ）がそのまま信号ＳＡ（ｆｉ）として出力され、Ｒ側分割信号Ｒ（ｆｉ）は信号ＳＢ（ｆｉ）＝０として出力される。レベル差ΔＬｉが０以下であれば逆に、Ｌ側は分割信号Ｌ（ｆｉ）は信号ＳＡ（ｆｉ）＝０として出力され、Ｒ側は分割信号Ｒ（ｆｉ）が信号ＳＢ（ｆｉ）として出力される。以上のようにして音源信号判定部６０１からゲート制御信号ＣＬ１～ＣＬｎ，ＣＲ１～ＣＲｎが出力され、ゲート６０２Ｌ１～６０２Ｌｎ，６０２Ｒ１～６０２Ｒｎがそれぞれ制御される。これも、前者と同様、帯域別チャネル間レベル差として、Ｌ側からＲ側を引いた値を用いた場合の説明である。信号ＳＡ（ｆ１）～ＳＡ（ｆｎ）、信号ＳＢ（ｆ１）～ＳＢ（ｆｎ）は先の実施例と同様にそれぞれ合成された信号ＳＡ，ＳＢとして出力端子t_A，t_Bにそれぞれ出力される（Ｓ１０）。

【００２０】前記実施例では、音源信号判定部６０１で用いる判定条件として、到達時間差とレベル差のうちどちらかの片方のみを利用する。しかし、レベル差のみを利用した場合、低域の周波数帯域ではＬ（ｆｉ）とＲ（ｆｉ）とのレベルが拮抗する場合があり、その場合はレベル差を正確に求めることが困難になる。また、時間差のみを利用した場合は、高い周波数帯域においては、位相の回転が起こるため時間差を正しく算出することが困難な場合がある。これらの点から、低域の周波数帯域では時間差を、高域ではレベル差を判定に用いた方が、全帯域に渡り単一のパラメータを用いるよりも有利である場合がある。

【００２１】そこで、音源信号判定部６０１で帯域別チャネル間時間差と帯域別チャネル間レベル差を共に用いる実施例を図６以下の図面を参照して説明する。この実施例の機能構成のブロックとしては図１と同一であるが、チャネル間時間差／レベル差検出部分３、帯域別チャネル間時間差／レベル差検出部５と音源信号判定部６０１での処理が以下のように異なる。チャネル間時間差／レベル差検出部３は、検出された時間差$\Delta\tau_1$，$\Delta\tau_2$の各絶対値の平均、又は$\Delta\tau_1$，$\Delta\tau_2$が比較的近い＊値であれば、その一方のみなど、一つの時間差$\Delta\tau$を出力する。なおチャネル間時間差$\Delta\tau_1$，$\Delta\tau_2$，$\Delta\tau$をチャネル信号Ｌ（ｔ），Ｒ（ｔ）を周波数軸上で帯域分割する前に算出したが、帯域分割した後に算出することも可能である。

20

【００２２】図５に示すように、Ｌチャネル信号Ｌ（ｔ）、Ｒチャネル信号Ｒ（ｔ）をフレーム（例えば２０～４０ｍｓ）毎に読み込み（Ｓ０２）、帯域分割部４でＬチャネル信号、Ｒチャネル信号をそれぞれ複数の周波数帯域に分割する。この例ではＬチャネル信号Ｌ（ｔ）、Ｒチャネル信号Ｒ（ｔ）にそれぞれハニング窓をかけ（Ｓ０３）、それぞれフーリエ変換を施して分割された信号Ｌ（ｆ１）～Ｌ（ｆｎ）、Ｒ（ｆ１）～Ｒ（ｆｎ）を得る（Ｓ０４）。

【００２３】次に、帯域別チャネル間時間差／レベル差検出部５では分割された信号の周波数ｆｉが１／（２Δτ）（Δτはチャネル時間差）以下の帯域（以下、低域と呼ぶ）であるかを調べ（Ｓ０５）、以下であれば帯域別チャネル間位相差Δφｉを出力し（Ｓ０８）、分割された信号の周波数ｆが１／（２Δτ）より大きく１／Δτ未満の帯域（以下、中域と呼ぶ）であるかがチェックされ（Ｓ０６）、この中域であれば帯域別チャネル間位相差Δφｉ及びレベル差ΔＬｉを出力し（Ｓ０９）、分割された信号の周波数ｆが１／Δτ以上の帯域（以下、高域と呼ぶ）かがチェックされ（Ｓ０７）、高域であれば帯域別チャネル間レベル差ΔＬｉを出力する（Ｓ１０）。

【００２４】音源信号判定部６０１は、帯域別チャネル間時間差／レベル差検出部５で検出された帯域別チャネル間位相差、レベル差を用いてＬ（ｆ１）～Ｌ（ｆｎ）、Ｒ（ｆ１）～Ｒ（ｆｎ）それぞれについて何れを出力するかの判定を行う。なお、位相差Δφｉ、レベル差ΔＬについては、この例では共にＬ側からＲ側の値を引いて算出した値を用いる。

【００２５】低域と判定された信号Ｌ（ｆｉ），Ｒ（ｆｉ）については図７に示すようにまず位相差Δφｉがπ以上かを調べ（Ｓ１５）、π以上であればΔφｉから２πを減算した値をΔφｉとし（Ｓ１７）、ステップＳ１５でΔφｉがπ以上でなければ、－π以下かを調べ（Ｓ１６）、以下であればΔφｉに２πを加算した値をΔφｉとし（Ｓ１８）、ステップＳ１６で－π以下でなければΔφｉをそのまま用いる（Ｓ１９）。ステップＳ１７，Ｓ１８，Ｓ１９で求めた帯域別チャネル間位相差Δφｉを時間差Δσｉに次式で変換する（Ｓ２０）。

【００２６】

$$\Delta\sigma i = 1000 \cdot \Delta\phi i / 2\pi f i \qquad (5)$$

分割された信号Ｌ（ｆｉ），Ｒ（ｆｉ）が中域と判定された場合は図８に示すように帯域別チャネル間レベル差ΔＬ（ｆｉ）を利用して、位相差Δφｉを一意に決定する。即ちΔＬ（ｆｉ）が正かを調べ（Ｓ２３）、正であれば、その帯域別チャネル間位相差Δφｉが正であるかを調べ（Ｓ２４）、正であればそのΔφｉをそのまま出力し（Ｓ２６）、ステップＳ２４で正でなければΔφｉに２πを加算した値をΔφｉとして出力する（Ｓ２

７）。ステップＳ２３でΔＬ（ｆｉ）が正でなければ、その帯域別チャネル間位相差Δφｉが負であるかを調べ（Ｓ２５）、負であれば、そのΔφｉをそのままΔφｉとして出力し（Ｓ２８）、ステップＳ２５で負でなければΔφｉから２πを減算した値をΔφｉとして出力する＊（Ｓ２９）。これらステップＳ２６～Ｓ２９の何れかのΔφｉが次式によりその帯域別チャネル間時間差Δσｉとして演算される（Ｓ３０）。

【００２７】

$$\Delta\sigma i = 1000 \cdot \Delta\phi i / 2\pi f i \qquad (6)$$

以上のようにして低域、中域における帯域別チャネル間時間差Δσｉと、高域における帯域別チャネル間レベル差ΔＬ（ｆｉ）が得られ、これらに応じて音源信号の判別が次のようになされる。図９に示すように低域と中域においては位相差Δφｉを、高域においてはレベル差ΔＬｉを利用して両チャネルの各周波数成分を該当するどちらかの音源の信号として判別する。具体的には、低域と中域においては図７、８でそれぞれ求められた帯域別チャネル間時間差Δσｉが正であるかを調べ（Ｓ３４）、正であれば、その帯域ｉのＬ側チャネル信号Ｌ（ｆｉ）を信号ＳＡ（ｆｉ）として出力し、Ｒ側帯域チャネル信号Ｒ（ｆｉ）を０の信号ＳＢ（ｆｉ）として出力する（Ｓ３６）。ステップＳ３４で帯域別チャネル時間差Δσｉが正でない場合は逆にＳＡ（ｆｉ）として０を出力し、ＳＢ（ｆｉ）としてＲ側チャネル信号Ｒ（ｆｉ）を出力する（Ｓ３７）。

【００２８】また、高域においては、図６中のステップＳ１０で検出した帯域別チャネル間レベル差ΔＬ（ｆｉ）が正であるかを調べ（Ｓ３５）、正であれば信号ＳＡ（ｆｉ）としてＬ側チャネル信号Ｌ（ｆｉ）を出力し、ＳＢ（ｆｉ）として０を出力する（Ｓ３８）。ステップＳ３５でレベル差ΔＬｉが正でなければＳＡ（ｆｉ）として０を出力し、ＳＢ（ｆｉ）としてＲ側帯域チャネル信号Ｒ（ｆｉ）を出力する（Ｓ３９）。

【００２９】以上のようにして各帯域についてＬ側又はＲ側が出力され、音源信号合成部７Ａ，７Ｂでそれぞれ判別した各周波数成分を全帯域に渡り加算し（Ｓ４０）、かつ、加算した各信号を逆フーリエ変換し（Ｓ４１）、その変換した信号ＳＡ，ＳＢを出力する（Ｓ４２）。以上説明したように、この実施例においては、周波数帯域毎に音源分離に有利なパラメータを用いることにより、全帯域に渡り単一のパラメータを用いる場合に比べてより分離性能の高い音源分離を実現することが可能である。

【００３０】この発明は音源の数が３個以上でも適用できる。例として、音源数が3、マイクロホン数が２である場合でマイクロホンへの到達時間差を利用して音源分離する場合を説明する。この場合、チャネル間時間差／レベル差検出部３で各音源についてＬチャネル信号、Ｒチャネル信号のチャネル間時間差を算出する際に、図３に示したように相互相関のパワーで正規化したヒストグラムの、累積度数（ピーク値）第一位から第三位までをとる各時点を求めることによって各音源信号についてのチャネル間時間差Δτ１，Δτ２，Δτ３を算出する。そして、帯域別チャネル間時間差／レベル差検出部５においても、各帯域の帯域別チャネル間時間差をΔτ１からΔτ３のどれかに決定する。この決定の仕方は、前記実施例で述べた計算式（３），（４）と同様である。音源信号判定部６０１では、例として、Δτ１＞０、Δτ２＞０、Δτ３＜０である場合で説明する。ここで、Δτ１，Δτ２，Δτ３はそれぞれ、音源Ａ，Ｂ，Ｃ各信号のチャネル間時間差と仮定し、さらに、これらの値はＬ側からＲ側の値を引いて算出した値と仮定する。この場合、音源ＡはＬ側のマイクロホン１に近く、音源ＢはＲ側のマイクロホン２の近くにある。よって、Ｌチャネルの信号から、帯域別チャネル間時間差がΔτ１となる帯域の信号を加算して音源Ａの信号を、またΔτ２となる帯域を加算して、音源Ｂの信号をそれぞれ分離することが可能である。また、Ｒチャネル信号から、帯域別チャネル間時間差がΔτ３となる帯域の信号を加算して出力することにより、音源Ｃの信号を分離する。

【００３１】上述では音源信号を分離し、分離された各音源信号ＳＡ，ＳＢを各別に出力した。しかし、例えば一方の音源Ａは発話者による音声であり、他方の音源Ｂは騒音のような場合、騒音と混合された音源Ａの信号音を分離抽出し、騒音を抑圧するためにもこの発明を適用することができる。その場合は図１において音源信号合成部７Ａを残し、１点鎖線で示す枠９中の音源信号合成部７Ｂ、ゲート６０２Ｒ１～６０２Ｒｎを省略すればよい。

【００３２】一方の音源Ａが他方の音源Ｂより周波数帯域が広い場合でその各周波数帯域が予め知られている場合は、図１０に示すように図１において帯域分離部１０において、両音源信号の重なっていない周波数帯域を分離する。例えば音源Ａの信号Ａ（ｔ）の周波数帯域はｆ１～ｆｎであるが音源Ｂの信号Ｂ（ｔ）の周波数帯域はｆ１～ｆｎ（ｆｎ＞ｆｍ）の場合、重なっていない帯域ｆｍ＋１～ｆｎの信号をマイクロホン１，２の出力から分離し、この帯域ｆｍ＋１～ｆｎの信号については、音源信号判定部６０１の判定処理、場合によっては帯域別チャネル間時間差／レベル差検出部５の処理を行わず、音源信号判定部６０１は、音源Ｂの信号として選出するチャネル信号ＳＢ（ｔ）として選出するＲの分割された帯域チャネル信号Ｒ（ｆｍ＋１）～Ｒ（ｆｎ）をそれぞれＳＢ（ｆｍ＋１）～ＳＢ（ｆｎ）として出力し、ＳＡ（ｆｍ＋１）～ＳＡ（ｆｎ）は０を出力させるように音源信号選択部６０２を制御する。即ちゲート６０２Ｌｍ＋１～６０２Ｌｎは常閉とし、ゲート６０２Ｒｍ＋１～

６０２Ｒｎは常開とする。

【００３３】上述では各帯域別チャネル間時間差Δσｉ、正か負かにより、また各帯域別チャネル間レベル差ΔＬｉが正か負かにより、つまり、いずれも０をしきい値として、その帯域信号が何れのマイクロホンに近いかを判別した。これはマイクロホン１として結ぶ線の２等分線に対して音源Ａと音源Ｂと左右対称に位置している場合である。この関係にない場合は判別しきい値を以下のように決めればよい。

【００３４】音源Ａの信号がマイクロホン１、マイクロホン２に到達する帯域別チャネル間レベル差をΔL_A、到達する帯域別チャネル間時間差をΔτ_A、音源Ｂの信号がマイクロホン１、マイクロホン２に到達する帯域別チャネル間レベル差をΔL_B、到達する帯域別チャネル間時間差をΔτ_Bとそれぞれする。このとき、帯域別チャネル間レベル差のしきい値ΔＬthは

ΔＬth＝（ΔL_A＋ΔＬｉ）／２

とし、帯域別チャネル間時間差のしきい値Δτthは

Δτth＝（Δτ_A＋Δτ_B）／２

とすればよい。先に述べた実施例ではΔL_B＝－ΔL_A、Δτ_B＝－Δτ_Aの場合でΔＬth＝０、Δτth＝０となる。音源Ａ，Ｂを分離できるように、二つの音源をマイクロホン１，２に対し、互いに異なる側となるように、マイクロホン１，２を位置させ、マイクロホン１，２に対する距離、方向は必ずしも正しくはわかっていない場合があり、しきい値ΔＬth，Δτthを可変として、分離がよく行われるようにΔＬth，Δτthを調整可能としてもよい。

【００３５】前記実施例では部屋の残響や回折の影響により、帯域別チャネル間時間差や帯域別チャネル間レベル差に誤りが生じ、各音源信号を精度よく分離することができない場合がある。このような問題を改善した実施例を次に述べる。図１１に示すように、マイクロホンＭ１，Ｍ２，Ｍ３は、例えば１辺が２０ｃｍの正三角形の頂点の位置に配置されている。マイクロホンＭ１～Ｍ３の指向特性に基づいて空間が分割して設定され、その各分割された空間を音源ゾーンと呼ぶ。全てのマイクロホンＭ１～Ｍ３が無指向で同じ特性を有する場合には、例えば図１２に示すように、ゾーンＺ１～Ｚ６のように６個に分割される。つまり、各マイクロホンＭ１，Ｍ２，Ｍ３と、その中心点Ｃｐをそれぞれ通る直線により、中心点Ｃｐを中心に等角間隔で６分割された６つのゾーンＺ１～Ｚ６が形成される。音源Ａはゾーンｚ３に、音源Ｂはゾーンｚ４に位置している。つまり、１個の音源ゾーンには１個の音源が属するよう、マイクロホンＭ１～Ｍ３の配置や特性に基づいて各音源ゾーンを決定する。

【００３６】図１１において、帯域分割部４１は、マイクロホンＭ１で収音した第１チャネルの音響信号Ｓ１をｎ個の周波数帯域信号Ｓ１（ｆ１）～Ｓ１（ｆｎ）に分割し、分割部４２でマイクロホンＭ２で収音した第２チャネルの音響信号Ｓ２をｎ個の周波数帯域信号Ｓ２（ｆ１）～Ｓ２（ｆｎ）に分割し、帯域分割部４３は、マイクロホンＭ３で収音した第３チャネルの音響信号Ｓ３をｎ個の周波数帯域信号Ｓ３（ｆ１）～Ｓ３（ｆｎ）に分割する。これら各帯域ｆ１～ｆｎは帯域分割部４１～４３で共通であり、このような帯域分割は離散的フーリエ変換器を利用することができる。

【００３７】音源分離部８０は図１乃至図１０を参照して説明した手法を用いて音源信号を分離するものである。ただし図１１ではマイクロホンが３つであるから、この３つのチャネルの信号の各２つの組合せについて同様な処理を行う。従って音源分離部８０内の帯域分割部と帯域分割部４１～４３を兼用することもできる。帯域別レベル（パワー）検出部Ｓ１で帯域分割部４１で得られた各帯域の信号Ｓ１（ｆ１）～Ｓ１（ｆｎ）のレベル（パワー）信号Ｐ（Ｓ１ｆ１）～Ｐ（Ｓ１ｆｎ）が検出され、同様に帯域別レベル検出部５２，５３でそれぞれ帯域分割部４２，４３で得られた各帯域信号Ｓ２（ｆ１）～Ｓ２（ｆｎ），Ｓ３（ｆ１）～Ｓ３（ｆｎ）の各Ｐ（Ｓ２ｆ１）～Ｐ（Ｓ２ｆｎ），Ｐ（Ｓ３ｆ１）～Ｐ（Ｓ３ｆｎ）がそれぞれ検出される。これら帯域別レベル検出もフーリエ変換器で実現できる。つまり各チャネル信号を離散的フーリエ変換によりスペクトルに分解し、その各スペクトルの電力を求めればよい。従って、各チャネル信号について、パワースペクトルを求め、そのパワースペクトルを帯域分割してもよい。各マイクロホンＭ１～Ｍ３の各チャネル信号を、帯域別レベル検出部４００で各帯域に分割すると共にそのレベル（パワー）を出力することになる。

【００３８】一方全帯域レベル検出部６１でマイクロホンＭ１で収音された第１チャネルの音響信号Ｓ１の全周波数成分のレベル（パワー）Ｐ（Ｓ１）が検出され、全帯域レベル検出部６２，６３でそれぞれマイクロホンＭ２，Ｍ３でそれぞれ収音された第２、第３チャネル２，３の各音響信号Ｓ２，Ｓ３の全周波数成分のレベルＰ（Ｓ２），Ｐ（Ｓ３）が検出される。

【００３９】音源状態判定部７０では、コンピュータ処理により、音響を発していない音源ゾーンを判定する。まず、帯域別レベル検出部５０により得られる帯域別レベルＰ（Ｓ１ｆ１）～Ｐ（Ｓ１ｆｎ）、Ｐ（Ｓ２ｆ１）～Ｐ（Ｓ２ｆｎ）、Ｐ（Ｓ３ｆ１）～Ｐ（Ｓ３ｆｎ）を、同一の帯域の信号について相互に比較する。そして各帯域ｆ１～ｆｎ毎に、最も大きなレベルのチャネルを特定する。

【００４０】帯域分割の数ｎを所定数以上にすることにより、前述したように、１つの帯域には１個の音源の音響信号しか含まれないと見なせるようにすることができるので、同一帯域ｆｉ　のレベルＰ（Ｓ１ｆｉ），Ｐ（Ｓ２ｆｉ），Ｐ（Ｓ３ｆｉ）は、同一音源からの音響のレベルと見なすことができる。よって、第１～第３チ

ャネルについて同一の帯域のレベルＰ（Ｓ１ｆｉ），Ｐ（Ｓ２ｆｉ），Ｐ（Ｓ３ｆｉ）に差があるときは、音源に最も近いマイクロホンのチャネルの帯域のレベルが最も大きくなる。

【００４１】前記処理の結果、各帯域ｆ１～ｆｎについて、最もレベルの大きなチャネルがそれぞれ割り当てられる。ｎ個の帯域中で第１～第３各チャネルについて、最もレベルが大きな帯域の合計数χ１，χ２，χ３を算出する。この合計数の値が大きいチャネルのマイクロホンほど、音源に近いとみなすことができる。合計数値が例えば９０ｎ／１００以上程度であればそのチャネルのマイクロホンに音源が近いと判定することができる。しかし、最もレベルが大きい帯域の合計数が５３ｎ／１００、次に合計値が大きい値が４９ｎ／１００の場合はそのそれぞれの対応マイクロホンに音源が近いか明確ではない。従って当該合計数が予め設定した基準値ＴｈＰ、例えばｎ／３程度を越えたとき、当該合計数と対応するチャネルのマイクロホンにその音源が最も近いと判定する。

【００４２】また、この音源状態判定部７０には、全帯域レベル検出部６０で検出された各チャネルのレベルＰ（Ｓ１）～Ｐ（Ｓ３）も入力されていて、そのレベルの全てが予め設定した基準値ＴｈＲ以下の場合には、何れのゾーンにも、音源がないと判定する。この音源状態判定部７０による判定結果に基づき、制御信号を発生して、音源分離部８０で分割された音響信号Ａ，Ｂに対する抑圧を信号抑圧部９０で行う。つまり制御信号ＳＡｉにより音響信号ＳＡを抑圧（減衰ないし削除）し、制御信号ＳＢｉにより音響信号ＳＢを抑圧し、制御信号ＳＡＢｉにより両音響信号ＳＡ，ＳＢを抑圧する。例えば信号抑圧部９０内に常閉スイッチ９Ａ，９Ｂが設けられ、音源分離部８０の出力端子 t_A，t_B が常閉スイッチ９Ａ，９Ｂを通じて、出力端子 t_A'，t_B' に接続され、制御信号ＳＡｉによりスイッチ９Ａが開とされ、制御信号ＳＢｉによりスイッチ９Ｂが開とされ、制御信号ＳＡＢｉによりスイッチ９Ａ，９Ｂが共に開にされる。当然のことであるが、音源分離部８０で行う分離処理するフレームの信号と、信号抑圧部９０での抑圧に用いる制御信号を得るフレームの信号とは同一のものを用いる。抑圧（制御）信号ＳＡｉ，ＳＢｉ，ＳＡＢｉの発生についてわかり易く説明する。

【００４３】いま、図１２に示すように音源Ａ，Ｂが位置している時マイクロホンＭ１～Ｍ３を図に示したように配置し、ゾーンＺ１～Ｚ６を決定し、音源ＡとＢが別個のゾーンＺ３，Ｚ４にそれぞれ位置するようにする。この時、音源Ａのマイクロホン Ｍ１～Ｍ３に対する距離ＳＡ１，ＳＡ２，ＳＡ３は、ＳＡ２＜ＳＡ３＜ＳＡ１となる。また、音源Ｂの各マイクロホンＭ１～Ｍ３に対する距離ＳＢ１，ＳＢ２，ＳＢ３は、ＳＢ３＜ＳＢ２＜ＳＢ１となる。

【００４４】全帯域レベル検出部６０の検出信号Ｐ（Ｓ１）～Ｐ（Ｓ３）のすべてが基準値ＴｈＲよりも小さいとき、音源Ａ，Ｂは発音、例えば発話していないと見なし、制御信号ＳＡＢｉにより、両音響信号ＳＡ，ＳＢを抑圧する。このとき、出力音響信号ＳＡ，ＳＢは無音信号となる（図１３の１０１，１０２）。音源Ａのみが発音しているときは、その音響信号のすべての帯域の周波数成分がマイクロホンＭ２へ一番大きな音圧レベル（パワー）で到達するので、このマイクロホンＭ２のチャネルの合計帯域数χ２が最も多くなる。

【００４５】また、音源Ｂのみが発音しているときは、その音響信号のすべての帯域の周波数成分がマイクロホンＭ３へ一番大きな音圧レベルで到達するので、このマイクロホンＭ３のチャネルの合計帯域数χ３が最も多くなる。さらに、音源Ａ，Ｂが共に発音している場合には、音響信号が最も大きな音圧レベルで到達する帯域数がマイクロホンＭ２とＭ３で拮抗する。

【００４６】したがって、前記した基準値ＴｈＰにより、音響信号があるマイクロホンへ最も大きな音圧レベルで到達する合計帯域数が、当該基準値ＴｈＰを越えた場合、当該マイクロホンが司るゾーンに音源が存在すると判定することにより、発音している音源ゾーンを検出することができる。上記の例では、音源Ａのみが発音しているときは、χ２のみが基準値ＴｈＰを越えて、発音している音源が存在するのはマイクロホンＭ２が司るゾーンＺ３であると検出されるので、制御信号ＳＢｉにより音声信号ＳＢを抑制して、音響信号ＳＡのみを出力させる（図１３の１０３，１０４）。

【００４７】また、音源Ｂのみが発音しているときは、χ３のみが基準値ＴｈＰを越えて、発音している音源が存在するのは、マイクロホンＭ３が司るゾーンＺ４であると検出されるので、制御信号ＳＡｉにより音響信号ＳＡを抑制して、音響信号ＳＢのみを出力させる（図１３の１０５，１０６）。さらに、音源Ａ，Ｂが共に発音していて、χ２，χ３ともに基準値ＴｈＰを越えるときは、例えば音源Ａに優先度を与えて、音源Ａのみが発音していると処理することができる。図１３の処理手順はそのようにしてある。また、χ２，χ３が共に基準値ＴｈＰに達していない場合は、レベルＰ（Ｓ１）～Ｐ（Ｓ３）が基準値ＴｈＲを越えている限り、両音源Ａ，Ｂともに発音していると判断し、制御信号ＳＡｉ，ＳＢｉ，ＳＡＢｉの何れも出力せず、音声抑圧部９０では合成信号ＳＡ，ＳＢに対する抑圧は行われない（図１３の１０７）。

【００４８】以上のようにして、音源分離部８０で分離された音源信号ＳＡ，ＳＢは、音源状態判定部７０によって発音していないと判定された音源に対応するものが、信号抑圧部９０で抑圧され、不要音が抑圧されるようになる。図１２に示した状態に対して、図１４に示すように音源Ｃをゾーン Ｚ６に加えた場合は、図示しない

が音源分離部８０からは、音源Ａに対応する信号ＳＡ、音源Ｂに対応する信号ＳＢの他に、音源Ｃに対応する信号ＳＣを出力する。

【００４９】また、信号抑圧部９０に対して、音源状態判定部７０から、信号ＳＡを抑圧する制御信号ＳＡｉ、信号ＳＢを抑圧する制御信号ＳＢｉの他に、信号ＳＣを抑圧する制御信号ＳＣｉが出力する。また、信号ＳＡとＳＢを抑圧する制御信号ＳＡＢｉの他に、信号ＳＢとＳＣを抑圧する制御信号ＳＢＣｉ、信号ＳＣとＳＡを抑圧する制御信号ＳＣＡｉ、信号ＳＡとＳＢとＳＣの全部を抑圧する制御信号ＳＡＢＣｉが出力する。この音源状態判定部７０は、図１５に示すような処理を行う。

【００５０】まず、レベルＰ（Ｓ１）～Ｐ（Ｓ３）の全部が基準値ＴｈＲを越えていない場合は、いずれの音源Ａ～Ｃも発音していないものと判断して、音源状態判定部７０からＳＡＢＣｉを出力して、信号ＳＡ，ＳＢ，ＳＣのいずれもが抑圧される（図１５の２０１～２０２）。次に、音源Ａ，Ｂ，Ｃがそれぞれ単独で発音している場合は、Ｐ（Ｓ１）～Ｐ（Ｓ３）の何れかはＴｈＲより大となり、前記した音源が２個の場合と同様に、その音源に最も近いマイクロホンのチャネルのレベルが最も大きくなるので、そのチャネルの帯域数χ１，χ２，χ３のいずれかが基準値ＴｈＰを越える。そして、音源Ｃのみが発音している場合は、χ１がＴｈＰを越え、制御信号ＳＡＢｉを出力して信号ＳＡ，ＳＢが抑圧される（図１５の２０３，２０４）。また、音源Ａのみが発音している場合は、制御信号ＳＢＣｉが出力して信号ＳＢ，ＳＣが抑圧される。さらに、音源Ａのみが発音している場合は、制御信号ＳＢＣｉが出力して信号ＳＢ，ＳＣが抑圧される（図１５の２０５～２０８）。

【００５１】次に、３つの音源Ａ～Ｃのうちのいずれか２つが発音する場合は、発音していない音源に対応するゾーンにあるマイクロホンのレベルが最も大きくなる帯域数が、他のマイクロホンのものに比べて小さくなる。例えば、音源Ｃのみが発音していない場合には、マイクロホンＭ１のレベルが最も大きくなる帯域数χ１が、他の２個のマイクロホンＭ２，Ｍ３の帯域数χ２，χ３に比べて小さくなる。

【００５２】よって、予めある基準値ＴｈＱ（＜ＴｈＰ）を設定し、χ１がその基準値ＴｈＱ以下になる場合は、マイクロホンＭ１とマイクロホンＭ３で空間を２分割したゾーンＺ５，Ｚ６の内、マイクロホンＭ１に近いゾーンＺ６では、音源は信号を発していないと判定する。さらに、マイクロホンＭ１とＭ２で空間を２分割したゾーンＺ１，Ｚ２のうちマイクロホンＭ１に近いゾーンＺ１では音源は信号を発していないと判定する。

【００５３】すなわち、ゾーンＺ１，Ｚ６にある音源は信号を発していないと判定するのである。これらのゾーンにある音源は音源Ｃであることから、音源Ｃが信号を発していないと判定される。つまり、音源Ａ，Ｂのみが信号を発していると判定され、制御信号ＳＣｉを生成し、信号ＳＣが抑圧される。図１４に示した状態で３つの音源Ａ～Ｃのうち１つのみが発音していない場合は通常は何れのマイクロホンについても最大となる帯域数χ１，χ２，χ３は基準値ＴｈＰ以下となるため、図１５においてステップ２０３，２０５，２０７を通過し、ステップ２０９で、χ１が基準値ＴｈＱ以下かを調べ、音源Ｃのみが発音していなければ、χ１＜ＴｈＱとなり、制御信号ＳＣｉが生成される（図１５の２１０）。ステップ２０９でχ１がＴｈＱ以下でなければχ２，χ３についても同様にＴｈＱ以下であるかが順次調べられ、ＴｈＱ以下であれば音源Ａのみ、又は音源Ｂのみが発音していないと推定され、それぞれ制御信号ＳＡｉ又はＳＢｉが抑圧される（図１５の２１１～２１４）。

【００５４】ステップ２１３でχ３がＴｈＱ以下でないと判定されると、音源Ａ，Ｂ，Ｃは全て発音していると判定され、何れの制御信号も生成されない（図１５の２１５）。この場合基準値ＴｈＰは２ｎ／３～３ｎ／４程度基準値ＴｈＱはｎ／２～２ｎ／３程度、つまり例えばＴｈＰを２ｎ／３程度にすると、ＴｈＱはｎ／２程度にする。

【００５５】なお、以上の例では、ゾーンをＺ１～Ｚ６の６つに分けたが、図１６に示すように、中心点Ｃｐから各マイクロホン間の中点を通る点線により３つのゾーンＺ１～Ｚ３に分けても同様に音源状態を判定できる。この場合は、例えば、音源Ａのみが発音している場合は、マイクロホンＭ２のチャネルの帯域数χ２が最も大きくなるので、そのマイクロホンＭ２の司るゾーンＺ２に音源があると判定される。また、音源Ｂのみが発音している場合はχ３が最も大きくなり、ゾーンＺ３に音源があると判定される。また、χ１が予め設定した値ＴｈＱ以下である場合には、マイクロホンＭ１とＭ２およびＭ３とそれぞれ２分したうちのゾーンＺ１にある音源は発音していないと判定する。以上の処理により、ゾーンを３分割しても、６分割のときと同様に音源の状態を判定できる。

【００５６】また、基準値ＴｈＲ，ＴｈＰ，ＴｈＱは、全てのマイクロホンＭ１～Ｍ３で同一値を用いた場合で説明したが、マイクロホン毎に適宜変更してもよい。また、以上の説明では、音源が３個でマイクロホンが３個の場合についてであったが、マイクロホンの個数は音源の個数と同数以上であれば、同様に音源ゾーンを検出することができる。

【００５７】例えば、音源が４個の場合には、４個のマイクロホンにより、個々のチャネルのマイクロホンが１個の音源を司るように、図１６の分割方法と同様に４個のゾーンに空間を分割する。このときの音源状態判定は、図１５のステップ２０１～２０８と同様な処理により、４個全部の音源が無音か、いずれか１個が発音しているかを判定する。それらいずれでもないとき、図１５

のステップ２０９～２１４と同様な処理により、４個の内の１個が無音かを判定し、１個の無音もないとき図１５のステップ２１５と同じ処理により全部の音源が発音していると判定する。また、４個の内の３個の音源が発音しているとき（１個が無音のとき）は、そのままとしても良いが、その３個の内のより無音に近い１個を選別するには、次のようにより細かく制御する。すなわち、基準値をＴｈＱからＴｈＳ（ＴｈＰ＞ＴｈＳ＞ＴｈＱ）に換え、図１５の各ステップ２１０，２１２，２１４の各々の次段に図１５のステップ２０９～２１４と同様な処理部分を設けて、３個の内から１個の無音に近い音源を判定する。

【００５８】このように、音源の数が多くなるほど、図１５のステップ２０９～２１４の処理内容を繰り返すことにより、無音又は無音に近い音源を２以上判定することができる。ただし、判定基準値ＴｈＳは処理の繰り返しが増えるほど、ＴｈＰに近付ける。以上の処理動作手順マイクロホンが４個、音源が４個の場合について図１７に示すようになる。まずマイクロホンＭ１～Ｍ４より第１～第４チャネル信号Ｓ１～Ｓ４を取込み（Ｓ０１）、これらチャネル信号Ｓ１～Ｓ４のレベルＰ（Ｓ１）～Ｐ（Ｓ４）をそれぞれ検出し（Ｓ０２）、これらレベルＰ（Ｓ１）～Ｐ（Ｓ４）の何れもが基準値ＴｈＲ以下であるかを調べ（Ｓ０３）、基準値以下であれば制御信号ＳＡＢＣＤｉを生成して合成信号ＳＡ，ＳＢ，ＳＣ（Ｓ１）の出力を抑圧する（Ｓ０４）。ステップＳ０３で何れかが基準値ＴｈＲ以下でなければ、各チャネル信号Ｓ１～Ｓ４をｎ帯域に分割すると共にその各帯域のレベルＰ（Ｓ１ｆｉ），Ｐ（Ｓ２ｆｉ），Ｐ（Ｓ３ｆｉ），Ｐ（Ｓ４ｆｉ）（ｉ＝１，…，ｎ）を求める（Ｓ０５）。各チャネル間で同一帯域ｆｉ　のレベル中の最大のチャネルｆｉＭ（Ｍは１，２，３，４の何れか）を各帯域について決定し（Ｓ０６）、全帯域（ｎ個）中でｆｉ１，ｆｉ２，ｆｉ３，ｆｉ４の各合計値χ１，χ２，χ３，χ４を求める（Ｓ０７）。χ１，χ２，χ３，χ４中の最大のものχ〃を求め（Ｓ０８）、χ〃が基準値ＴｈＰ１（例えばｎ／３）以上であるかを調べ（Ｓ０９）、ＴｈＰ１以上であればチャネルＭと対応して選出した音源信号、音源Ａの信号であれば分離されたチャネルＭ以外の分離されたチャネルの分離音響信号を抑圧する制御信号ＳＢＣＤｉを生成する（Ｓ０１０）。ステップＳ０８から直ちにステップＳ０１０へ移ってもよい。

【００５９】ステップＳ０９で基準値以上でなければχ〃が基準値ＴｈＱ以下のチャネルＭがあるかを調べる（Ｓ０１１）。ＴｈＱ以下のものがなければ、全ての音源が発音しているとみなして、何れの制御信号も発生しない（Ｓ０１２）。ステップＳ０１１でχ〃がＴｈＱ以下のチャネルＭがあれば、これと対応するチャネルＭとして分離された音源信号を抑圧する制御信号ＳＭｉ　を生成する（Ｓ０１３）。

【００６０】制御信号ＳＭｉで抑圧された以外の分離された音源信号中の無音又は無音に近いものを抑圧するには、Ｓを＋１し（Ｓ０１４）（Ｓは予め０に初期化しておく）、ＳがＭ－１（Ｍは音源の数）と一致したかを調べ（Ｓ０１５）、一致していなければ、ＴｈＱを＋ΔＱだけ大としてステップＳ０１１に戻る（Ｓ０１６）。ＳがＭ－１になるまでＴｈＱをＴｈＰを越えない範囲でΔＱづつ増加させステップＳ０１１を実行する。ステップＳ０１５でＭ－１＝Ｓであれば、その時のＴｈＱ以下の各χ〃の各チャネルＭと対応する分離された音源信号を抑圧する各制御信号ＳＭｉを生成する（Ｓ０１３）。必要に応じてステップＳ０１５でＭ－１＝Ｓになる前にステップＳ０１３に移ってもよい。

【００６１】ステップＳ０７でχ１～χ４を計算した後、これらでＴｈＰ２（例えば２ｎ／３）以上のものがあるかを調べ、あればステップＳ０１０に移り、なければステップＳ０１１に移るようにしてもよい（Ｓ０１７）。上述では音源分離の精度を上げるため、マイクロホンＭ１～Ｍ３のチャネル信号Ｓ１～Ｓ３の帯域間レベル差を利用して信号抑圧部９０に対する制御信号を生成したが、帯域間時間差を利用して制御信号を生成することもできる。

【００６２】この例を図１８に、図１１と対応する部分に同一符号を付けて示す。この実施例では帯域分割部４１で得られた各帯域ｆ１～ｆｎの信号Ｓ１（ｆ１）～Ｓ１（ｆｎ）から到達時間差信号Ａｎ（Ｓ１ｆ１）～Ａｎ（Ｓ１ｆｎ）が帯域別時間差検出部１０１で検出され、同様に帯域分割部４２，４３でそれぞれ得られた各帯域の信号Ｓ２（ｆ１）～Ｓ２（ｆｎ）、Ｓ３（ｆ１）～Ｓ３（ｆｎ）からそれぞれ到達時間差信号Ａｎ（Ｓ２ｆ１）～Ａｎ（Ｓ２ｆｎ），Ａｎ（Ｓ３ｆ１）～Ａｎ（Ｓ３ｆｎ）が帯域別時間差検出部１０２，１０３で検出される。

【００６３】これらの到達時間差信号を得る処理は、例えば、フーリエ変換により各帯域の信号の位相（あるいは群遅延）を算出し、同一の帯域ｆｉの信号Ｓ１（ｆｉ），Ｓ２（ｆｉ），Ｓ３（ｆｉ）（ｉ＝１，２，…，ｎ）の位相を相互に比較することで、同一音源信号の到達時間差と対応した信号を得ることができる。この場合も帯域分割部４０での分割は、１つの帯域には１つの音源信号成分しか存在しないとみなせる程度に小さく行う。

【００６４】この到達時間差の表現方法は、例えば、マイクロホンＭ１～Ｍ３のいずれかを基準にしてその基準マイクロホンに対する到達時間差を０に設定しておけば、他のマイクロホンに対する到達時間差はその基準マイクロホンに対して速く到達したか遅く到達したかで判定できるので、正又は負の極性を付した数値で表すことができる。この場合、基準マイクロホンを例えばＭ１と

すると、到達時間差信号Ａｎ（Ｓ１ｆ１）～Ａｎ（Ｓ１ｆｎ）は全て０となる。

【００６５】音源状態判定部１１０では、コンピュータ処理により音声を発していない音源を判定する。まず、帯域別時間差検出部１００により得られる到達時間差信号Ａｎ（Ｓ１ｆ１）～Ａｎ（Ｓ１ｆｎ），Ａｎ（Ｓ２ｆ１）～Ａｎ（Ｓ２ｆｎ），Ａｎ（Ｓ３ｆ１）～Ａｎ（Ｓ３ｆｎ）を、同一の帯域の信号について相互に比較する。これにより各帯域ｆ１～ｆｎ毎に、最も信号が遠く到達するチャネルが決定できる。

【００６６】そこで、各チャネルについて信号が最も遠く到達すると判定された帯域の合計数を算出して、それをチャネル間で比較する。この結果、この合計帯域数の値が大きいチャネルのマイクロホンほど、音源に近いとみなすことができる。そして、あるチャネルについて、当該合計帯域数が予め設定した基準値ＴｈＰを越えたとき、当該のチャネルのマイクロホンが司るゾーンに音源があると判定する。

【００６７】また、この音源状態判定部１１０には、全帯域レベル検出部６０で検出された各チャネルのレベルＰ（Ｓ１）～Ｐ（Ｓ３）も入力され、あるチャネルのレベルが予め設定した基準値ＴｈＲ以下の場合には、そのチャネルのマイクロホンが司るゾーンには、音源がないと判定する。いま図１２に示したように音源Ａ，Ｂに対し、マイクロホンＭ１～Ｍ３を配置したとする。またマイクロホンＭ１のチャネルに対する前記した合計帯域数をχ１、マイクロホンＭ２，Ｍ３の各チャネルに対する合計帯域数をそれぞれχ２，χ３とする。

【００６８】この場合も図１３に示した処理手順と同様にすればよい。即ち、まず、全帯域レベル検出部６０の検出信号Ｐ（Ｓ１）～Ｐ（Ｓ３）のすべてが基準値ＴｈＲよりも小さいとき（１０１）、音源Ａ，Ｂは発音していないと見なし、制御信号ＳＡＢｉを生成して（１０２）、両音源信号ＳＡ，ＳＢを抑圧する。このとき、出力信号ＳＡ′，ＳＢ′は無音信号となる。

【００６９】音源Ａのみが発音しているときは、その音源信号のすべての帯域の周波数成分がマイクロホンＭ２へ一番速く到達するので、このマイクロホンＭ２のチャネルの合計帯域数χ２が最も多くなる。また、音源Ｂのみが発音しているときは、その音源信号のすべての帯域の周波数成分がマイクロホンＭ３へ一番速く到達するので、このマイクロホンＭ３のチャネルの合計帯域数χ３が最も多くなる。

【００７０】さらに、音源Ａ，Ｂが共に発音している場合には、音源信号が最も速く到達する帯域数がマイクロホンＭ２とＭ３で拮抗する。したがって、前記した基準値ＴｈＰにより、音源信号があるマイクロホンへ最も速く到達する合計帯域数が、当該設定値ＴｈＰを越えた場合、当該マイクロホンが司るゾーンに音源が存在し、その音源が発音していると判定する。

【００７１】上記の例では、音源Ａのみが発音しているときは、χ２のみが基準値ＴｈＰを越えて（図３の１０３）、音響を発生している音源が存在するのはマイクロホンＭ２が司るゾーンＺ３であると検出されるので、制御信号ＳＢｉが生成され（１０４）、音響信号ＳＢが抑制され、信号ＳＡのみが出力される。また、音源Ｂのみが発音しているときは、χ３のみが基準値ＴｈＰを越え（１０５）、音を発している音源が存在するのは、マイクロホンＭ３が司るゾーンＺ４であると検出されるので、制御信号ＳＡｉが生成され（１０６）信号ＳＡが抑制されて、信号ＳＢのみが出力される。

【００７２】この例ではＴｈＰは例えばｎ／３程度に設定され、音源Ａ，Ｂが共に発音していて、χ２，χ３ともに基準値ＴｈＰを越えることがある。この場合は図１３の処理手順に示すように一方の音源、この例ではＡを優先させ、音源Ａへ分離信号のみを出力させることもできる。また、χ２，χ３が共に基準値ＴｈＰに達していない場合は、レベルＰ（Ｓ１）～Ｐ（Ｓ３）が基準値ＴｈＲを越えている限り、両音源Ａ，Ｂともに発音していると判断し、制御信号ＳＡｉ，ＳＢｉ，ＳＡＢｉは出力せず（図３の１０７）音声抑圧部９０では音声信号ＳＡ，ＳＢに対する抑圧は行われない。

【００７３】図１２に示した状態に対して図１４に示すように音源Ｃをゾーン Ｚ６に加えた場合、図示しないが音源分離部８０からは、音源Ａに対応する信号ＳＡ、音源Ｂに対応する信号ＳＢの他に、音源Ｃに対応する信号ＳＣが出力する。これと対応して音源状態判定部１１０から、信号ＳＡを抑圧する制御信号ＳＡｉ、信号ＳＢを抑圧する制御信号ＳＢｉの他に、信号ＳＣを抑圧する制御信号ＳＣｉが出力し、また、信号ＳＡとＳＢを抑圧する制御信号ＳＡＢｉの他に、信号ＳＢとＳＣを抑圧する制御信号ＳＢＣｉ、信号ＳＣとＳＡを抑圧する制御信号ＳＣＡｉ、信号ＳＡ，ＳＢ，ＳＣの全部を抑圧する制御信号ＳＡＢＣｉが出力する。そして、この音源状態判定部１１０は先に述べた図１５に示したと同様の処理を行う。

【００７４】まず、レベルＰ（Ｓ１）～Ｐ（Ｓ３）の全部が基準値ＴｈＲを越えていない場合は、いずれの音源Ａ～Ｃも発音していないものと判断して、音源状態判定部１１０からはＳＡＢＣｉが出力して、信号ＳＡ，ＳＢ，ＳＣのいずれもが抑圧される。次に、音源Ａ，Ｂ，Ｃがそれぞれ単独で発音している場合には、前記した音源が２個の場合と同様に、その音源に最も近いマイクロホンのチャネルの到達時間が最も速くなるので、そのチャネルの帯域数χ１，χ２，χ３のいずれかが基準値ＴｈＰを越える。そして、音源Ｃのみが発音している場合は、制御信号ＳＡＢｉが出力して信号ＳＡ，ＳＢが抑圧される。また、音源Ａのみが発音している場合は、制御信号ＳＢＣｉが出力して信号ＳＢ，ＳＣが抑圧される。さらに、音源Ｂのみが鳴っている場合は、制御信号ＳＡ

Ｃｉが出力して信号ＳＡ，ＳＣが抑圧される（図１５の２０３～２０８）。

【００７５】次に、３つの音源Ａ～Ｃのうちのいずれか２つが発音している場合は、発音していない音源に対応するゾーンにあるマイクロホンの到達時間の最も速い帯域数が、他のマイクロホンのものに比べて小さくなる。例えば、音源Ｃのみが鳴っていない場合には、マイクロホンＭ１への到達時間が最も速い帯域数χ１が、他の２個のマイクロホンＭ２，Ｍ３の帯域数χ２，χ３に比べて小さくなる。

【００７６】よって、予めある基準値ＴｈＱ（＜ＴｈＰ）を設定し、χ１がその基準値ＴｈＱ以下になる場合は、マイクロホンＭ１とマイクロホンＭ３で空間を２分割したゾーンＺ５，Ｚ６の内、マイクロホンＭ１に近いゾーンＺ６では、音源は信号を発していないと判定し、さらに、マイクロホンＭ１とＭ２で空間を２分割したゾーンＺ１，Ｚ２のうちマイクロホンＭ１に近いゾーンＺ１では音源は信号を発していないと判定する。

【００７７】すなわち、ゾーンＺ１，Ｚ６にある音源は信号を発していないと判定するのである。これらのゾーンにある音源は音源Ｃであることから、音源Ｃが信号を発していないと判定される。つまり、音源Ａ，Ｂのみが信号を発していると判定され、制御信号ＳＣｉが生成されて信号ＳＣが抑圧される（図１５の２０９～２１０）。音源Ａのみ、音源Ｂのみがそれぞれ信号を発していないゾーンも、同様に判定される（図１５の２１１～２１４）。

【００７８】また、χ１，χ２，χ３がともに基準値ＴｈＱ以下でないと判定されると、音源Ａ，Ｂ，Ｃはその全てが信号を発していると判定される（図１５の２１５）。なお、以上の例では、ゾーンをＺ１～Ｚ６の６つに分けたが、図１６に示したように、３つに分けても同様に音源状態を判定できる。この場合は、例えば、音源Ａのみが発音している場合は、マイクロホンＭ２のチャネルの帯域数χ２が最も大きくなるので、そのマイクロホンＭ２の司るゾーンＺ２に音源があると判定される。また、音源Ｂのみが発音している場合はχ３が最も大きくなり同様にゾーンＺ３に音源があると判定される。また、χ１が予め設定した値ＴｈＱ以下である場合には、マイクロホンＭ１とＭ３で空間を２分したうちのゾーンＺ１にある音源は発音していないと判定し、同じくマイクロホンＭ１とＭ２で空間を分割したうちのゾーンＺ１にある音源は信号を発していないと判定する。以上の処理により、ゾーンを３分割しても、６分割したときと同様に音源の状態を判定できる。

【００７９】以上の場合の基準値ＴｈＰ，ＴｈＱの設定は、先の帯域レベルを利用する場合と同様に行えばよい。また、基準値ＴｈＲ，ＴｈＰ，ＴｈＱは、全てのマイクロホンＭ１～Ｍ３で同一値を用いた場合で説明したが、マイクロホン毎に適宜変更してもよい。また、以上の説明では、音源が３個でマイクロホンが３個の場合についてであったが、マイクロホンの個数は音源の個数と同数以上であれば、同様に音源ゾーンを検出することができる。その処理手順は先に述べた帯域レベルを利用する場合と同様である。従って、例えば音源が４個の場合に４個の内の３個の音源が発音しているとき（１個が無音のとき）は、そのままとしても良いが、その３個の内のより無音に近い１個も選別するには、基準値をＴｈＱからＴｈＳ（ＴｈＰ＞ＴｈＳ＞ＴｈＱ）に換え、図１５の２１０，２１２，２１４の各々の次段に図１５の２０９～２１４と同様な処理部分を設けて、３個の内から１個の無音の音源を判定することも同様である。

【００８０】図１７に示した処理において、そのレベルの代りに時間差を用いれば、図１８に示した到達時間差を利用した不要信号の抑圧に、図１７に示した処理手順も適用できる。上述においては各マイクロホンの出力チャネル信号をまず帯域分割したが、帯域別レベルを利用する場合はまず各チャネルのパワースペクトルを求めた後、帯域分割してもよい。その例を図１９に図１、図１１と対応する部分と同一符号を付けて示し、これらと異なる部分のみを説明する。この例ではマイクロホン１，２よりの各チャネル信号は、パワースペクトル分解部３００により、例えば高速フーリエ変換によりパワースペクトルに変換され、その後、各チャネルごとに帯分割部４で各帯域に分割され、各帯域ではほぼ１つの音源信号のみが主として含まれるようにして帯域別レベルを得る。この場合、音源信号選択部６０２へ供給する各帯域別レベルは、その原スペクトルの位相成分も供給し、音源信号合成部７で音源信号が再生できるようにする。

【００８１】また各帯域別レベルは帯域別チャネル間レベル差検出部５と音源状態判定部７０とへ供給され、これらの部分で図１、図１１で説明したように処理されるその他の動作は図１又は図１１の場合と同一である。図２を参照して説明した実施例において、チャネル間時間差を用いずに、各帯域分割信号ごとに、対応帯域別チャネル間時間差のみを用いて、何れの音源から到来したかを判定してもよい。また図５を参照して説明した実施例において、チャネル間レベル差を用いずに、各帯域分割信号ごとに、対応帯域別チャネル間レベル差のみを用いて、何れの音源から到来したかを判定してもよい。図５を参照した実施例におけるチャネル間レベル差の検出は、対数レベルに変換する前のレベルを用いてもよい。図１中の帯域分割部４、図１１、図１８中の各帯域分割部４０、図２０中の帯域分割部２３３、図２１中の帯域分割部２４１における各周波数帯域の分割は必ずしも同一とする必要はない。要求される精度に応じて、これらの分割数を互いに異ならせてもよい。図２０中の帯域分割部２３３はその後の処理のために、その入力信号のパワースペクトルを先ず求め、その後、複数の周波数帯域に分割してもよい。

【００８２】以下に図６～９に示したこの発明を適用した実験例を示す。図２０に示す３種類の２音源信号の組み合わせにこの発明を適用し、その際に帯域分割部４で与える周波数分解能を変化させ、分離信号を物理的、及び主観的に評価した。分離処理前の混合信号は、チャネル間時間差及びレベル差のみを計算機上で与えて加算することにより作成した。与えたチャネル間時間差、レベル差はそれぞれ、０．４７ｍｓ、２ｄＢである。

【００８３】帯域分割部４の周波数分解能は、約５Ｈｚ，１０Ｈｚ，２０Ｈｚ，４０Ｈｚ，８０Ｈｚの５種類とした。これらの分解能で分離した信号と、原信号（ＯＳ）の計６種類の信号について評価した。なお、信号帯域は約５ｋＨｚである。定量的評価を次のように行った。混合された信号の分離が完全に行われた場合、原信号と分離信号が等しくなる。すなわち、相関係数が１となる。そこで、分離度を計る物理量として、各音について原信号と処理後の信号との相関係数を算出した。

【００８４】結果を、図２２に破線で示す。音声は、いずれの組み合わせについても、周波数分解能が８０Ｈｚになると相関値がかなり低くなるが、それ以外の分解能の場合は顕著な差が見られなかった。鳥の鳴き声については今回用いた周波数分解能の間に顕著な差は見られなかった。主観評価を次のように行った。

【００８５】被験者は、正常な聴力を持つ２０代から３０代の日本人５人とした。各音源について、５種類の周波数分解能の分離音と原音をランダムにヘッドホンでダイオティックに提示し、音質について５段階で評価させた。一つの音の提示時間は約４秒間であった。結果を、図２２に実線で示す。分離音Ｓ１については周波数分解能１０Ｈｚの場合が一番評価が高い。また、全ての条件に対する評価の間に有意差（$\alpha < 0.05$）が存在した。分離音Ｓ２～４、６については周波数分解能２０Ｈｚの評価が最も高いが、２０Ｈｚと１０Ｈｚとの間には有意差はなかった。また、２０Ｈｚの音と５Ｈｚ，４０Ｈｚ，８０Ｈｚの間にはそれぞれ有意差が存在した。これらの結果から、音声については分離する組み合わせの種類によらず、最適な周波数分解能が存在することが分かった。この実験の場合は２０Ｈｚもしくは１０Ｈｚ程度が最適な値である。分離音Ｓ５（鳥の鳴き声）については４０Ｈｚの場合が最も評価が高いが有意差は４０Ｈｚと５Ｈｚ，２０Ｈｚと５Ｈｚの間にしか存在しなかった。なお、いずれの場合についても、分離処理後の音と原音の間には有意差が存在した。

【００８６】図２１、図２３にこの発明の効果を示す。図２１は、分離処理前の男声と女声の混合音声のスペクトル２０１とこの発明による分離処理後の男声Ｓ１、女声Ｓ２の各スペクトル２０２，２０３を表す。図２３は、分離処理前の男声Ｓ１、女声Ｓ２の各原音声の各波形をＡ，Ｂに、混合音声波形をＣに、分離処理後の男声Ｓ１、女声Ｓ２の各波形をＤ，Ｅにそれぞれ示す。図２１からは、不要な成分が抑圧されていることが分かる。さらに、図２３からは、分離処理後の音声が原音声と同程度の品質で復元されていることが分かる。

【００８７】帯域分割の分解能は音声の場合、１０～２０Ｈｚ程度が好ましく、５Ｈｚ以下、５０Ｈｚ以上は好ましくない。帯域分割の手法はフーリエ変換に限らず、帯域フィルタにより分割してもよい。次に図１１に示したレベル差を利用して音源状態を判定して信号抑圧部９０で信号抑圧を行う場合の実験例を示す。２個のマイクロホンを用い、２つの音源Ａ，Ｂをダミーヘッドから距離１．５ｍ、角度差９０度（２つのマイクロホンの中点に対し右４５度、左４５度）の位置に置き、同一の音圧レベルで、残響時間０．２ｓ（５００Ｈｚ）の可変残響室内で収音した。用いた混合音と分離音の組み合せは図２２中のＳ１～Ｓ４である。

【００８８】分離音声Ｓ１～Ｓ４について、無音と判定されたフレームの個数と、原音の無音フレームの個数の比率を算した。その結果は次の通り９０％以上正しく検出された。

	男（S1）	女(S2)	女声１(S3)	女声２(S4)
検出率	９９％	９３％	９２％	９５％

図６～９に示した基本方法と図１１に示した改良方法とのそれぞれで分離した音をランダムにヘッドホンでダイオティックに提示し、雑音の交じり具合の少なさと不連続感の少なさについて評価させた。用いた分離音は前記Ｓ１～Ｓ４であり、被験者は正常な聴力を持つ２０代から３０代の日本人５名である。一つの音の提示時間は約４秒間、各音の試行回数は３回である。その結果、雑音の交じり具合が少ないと評価した率は改良方法が９１．７％、基本方法は８．３％で、改良方法が少ないと判断した回答が格段と多かった。一方不連続感が少ないについては改良方法は２０．０％、基本方法が８０．０％で基本方法の方が少ないと判断する回答が多かったが、改良方法との間に有意な差は見られなかった。

【００８９】次に分離性能を相対評価を行うため、以下の５種類の音の分離度の比較を主観評価により行った。

（１）原音

（２）基本法（計算機）：チャネル間時間差（０．４７ｍｓ）、レベル差（２ｄＢ）を与えて計算機上で加算した混合信号を、基本方法で分離した音。

（３）改良法（実環境）：先の無音区間検出率の実験に用いた条件で収音した混合音を改良方法で分離した音。

（４）基本法（実環境）：先の無音区間検出率の実験に用いた条件で収音した混合音を基本方法で分離した音。

（５）混合音：先の無音区間検出率の実験に用いた条件で収音した混合音。

【００９０】図２０中の最初の２つの混合音に対し、

“原音”上記（１）～（４）の方法で処理した音、“混合音”の計２０種類をランダムにヘッドホンでダイオティックに提示し、分離度について７段階で評価させた。つまり「最も分離されている」を７点、「最も分離されていない」を１点とした。被験者、音の提示時間及び試行回数は、前記雑音の交じり具合の少なさの評価の場合と同一である。

【００９１】この結果を図２４中で、全音源（Ｓ０）をＡに、男声（Ｓ１）をＢに、女声（Ｓ２）をＣに、女声１（Ｓ３）をＤに、女声２（Ｓ４）をＥにそれぞれ示す。全音源について分析した結果（Ｓ０）と、音源の種類毎に分析した結果（Ｓ１）～（Ｓ４）とは、ほぼ同じ傾向を示した。Ｓ０～Ｓ４全ての場合について、“（１）原音”、“（２）基本法（計算機）”、“（３）改良法（実環境）”、“（４）基本法（実環境）”、“（５）混合音”の順に分離精度が高い。つまり実環境では改良方法の方が基本方法より優れている。

【００９２】

【発明の効果】以上述べたようにこの発明によれば複数のマイクロホンからの各チャネル信号を、主な成分が１つの音源信号の成分のみからなる程度に複数の帯域に分割し、これら各同一帯域について、レベル、到達時間を検出し、これらから、各帯域ごとに何れの音源信号かを判定分離することにより、各音源信号を正しく分離することができ、しかも実時間での処理が可能である。

【００９３】特に発音していない音源を検出し、その成分を抑圧することにより、部屋内のような回り込みや、残響がある場所でも、正確に分離することができる。

【図面の簡単な説明】

【図１】この発明の音源分離装置の実施例の機能構成を示すブロック図。

【図２】この発明の音源分離方法の実施例の処理手順を示す流れ図。

【図３】図２中のチャネル間時間差$\Delta\tau_1$，$\Delta\tau_2$を求める処理手順の例を示す流れ図。

【図４】Ａ，Ｂはそれぞれ二つの音源信号のスペクトルの例を示す図である。

【図５】この発明の音源分離方法で、チャネル間レベル差を利用して音源分離を行う実施例の処理手順を示す流れ図。

【図６】この発明音源分離方法で、チャネル間レベル差と、チャネル間到達時間差を利用する実施例の処理手順の一部を示す流れ図。

【図７】図６中のステップＳ０８の続きを示す流れ図。

【図８】図６中のステップＳ０９の続きを示す流れ図。

【図９】図６中のステップＳ１０、図７、図８中のステップＳ２０，Ｓ３０の続きを示す流れ図。

【図１０】周波数帯域が異なる音源信号を分離する実施例の機能構成を示すブロック図。

【図１１】レベル差を利用して不要音源信号を抑圧する構成を付加したこの発明の音源分離装置の実施例の機能構成を示すブロック図。

【図１２】３つのマイクロホンとその受けもつゾーンと、２つの音源の配置例を示す図。

【図１３】発音している音源が１つの場合の音源ゾーンの検出と、抑圧制御信号の生成処理手順の例を示す流れ図。

【図１４】３つのマイクロホンと、その受けもつゾーンと、３つの音源の配置例を示す図。

【図１５】音源が３つの場合の発音音源のゾーン検出と、抑圧制御信号の生成処理手順の例を示す流れ図。

【図１６】３つのマイクロホンによりゾーンを３つに分割した例と、音源の配置例を示す図。

【図１７】この発明の音源分離装置において、発音していない合成音源信号を抑圧する制御信号を生成するための処理手順の例を示す流れ図。

【図１８】到達時間差を利用して不要音源信号を抑圧する構成を付加したこの発明の音源分離装置の実施例の機能構成を示すブロック図。

【図１９】この発明音源分離装置で、パワースペクトルを求めた後、帯域分割を行う場合の実施例の機能構成を示すブロック図。

【図２０】この発明の実験に用いた音源の種類を示す図。

【図２１】図６～図９に示した実施例の方法による処理前と、処理後の音声スペクトルを示す図。

【図２２】図６～図９に示した実施例の方法を用いた主観評価実験の結果を示す図。

【図２３】図６～図９に示した実施例の方法により処理した処理後の音声波形と、その原音声波形を示す図。

【図２４】図６～図９に示した音源分離方法と図１１に示した音源分離装置とについての実験結果を示す図。

【図１】

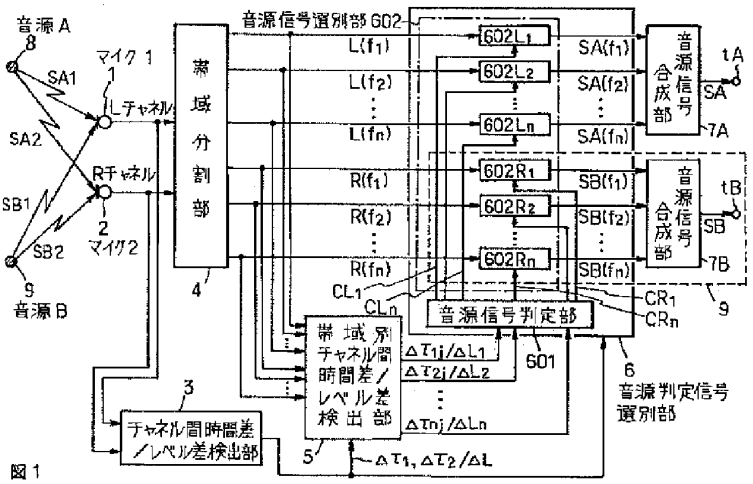


図１

【図４】

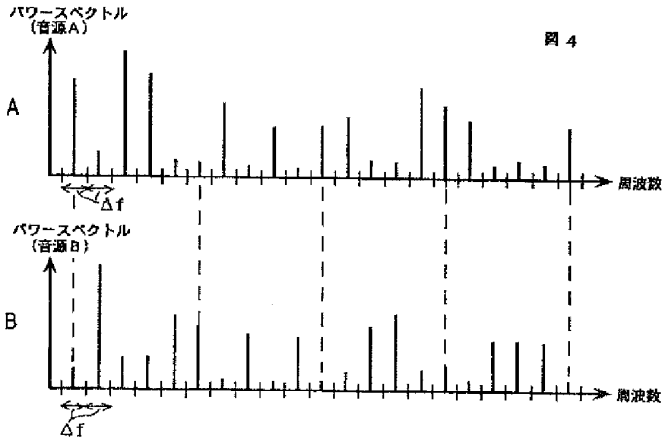


図４

【図１３】

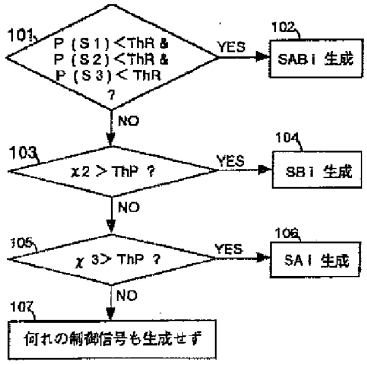


図13

【図２０】

混合信号	分離信号
男声+女声	男声(S1)
	女声(S2)
女声1+女声2 (同一発話者, 発話内容は異なる)	女声1(S3)
	女声2(S4)
おんどりの鳴き声+女声	おんどりの鳴き声(S5)
	女声(S6)

図　20

【図２】

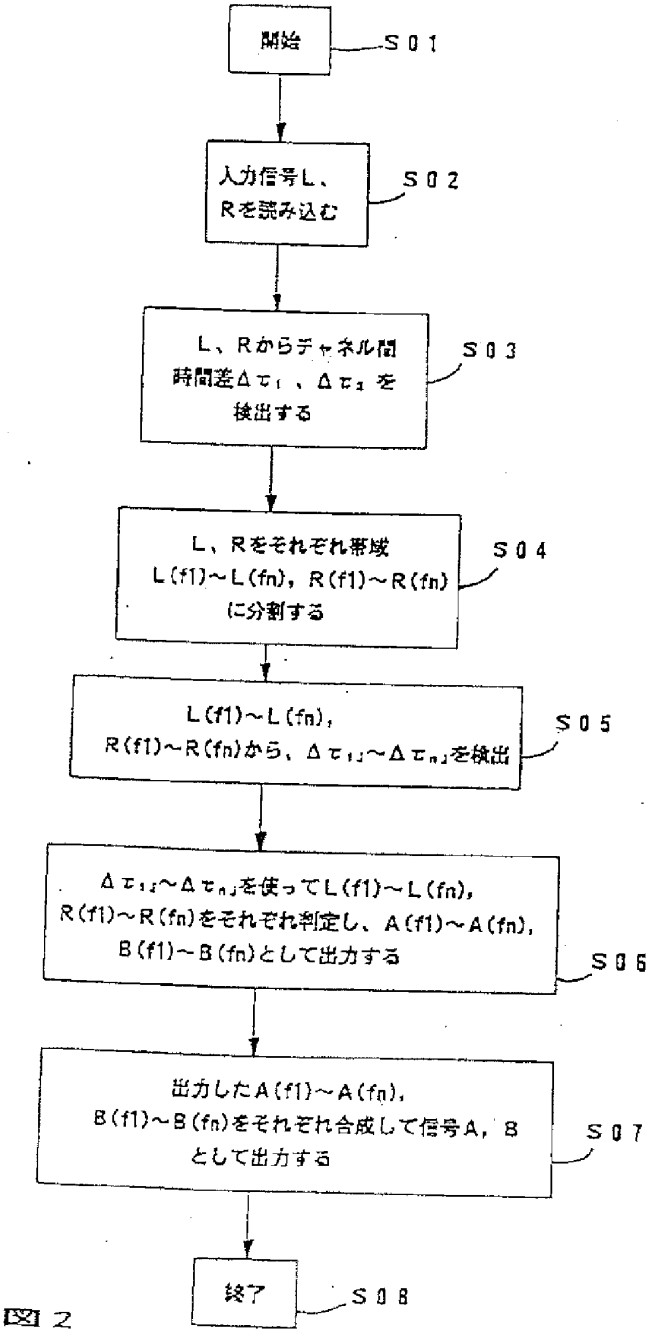
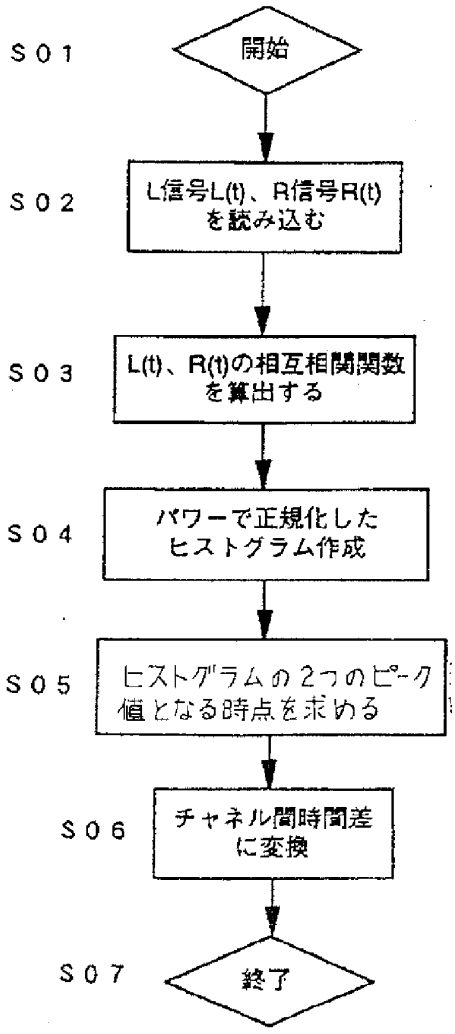

開始
S01
入力信号L、Rを読み込む
S02
L、Rからチャネル間時間差Δτ1、Δτ2を検出する
S03
L、Rをそれぞれ帯域L(f1)～L(fn)，R(f1)～R(fn)に分割する
S04
L(f1)～L(fn)，R(f1)～R(fn)から、Δτ1～Δτnを検出
S05
Δτ1～Δτnを使ってL(f1)～L(fn)，R(f1)～R(fn)をそれぞれ判定し、A(f1)～A(fn)，B(f1)～B(fn)として出力する
S06
出力したA(f1)～A(fn)，B(f1)～B(fn)をそれぞれ合成して信号A，Bとして出力する
S07
終了
S08

図２

【図３】



【図１４】

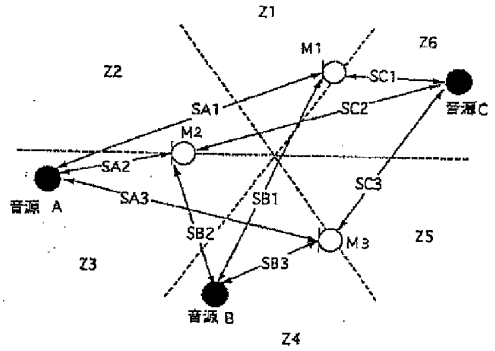


図１４

【図１５】

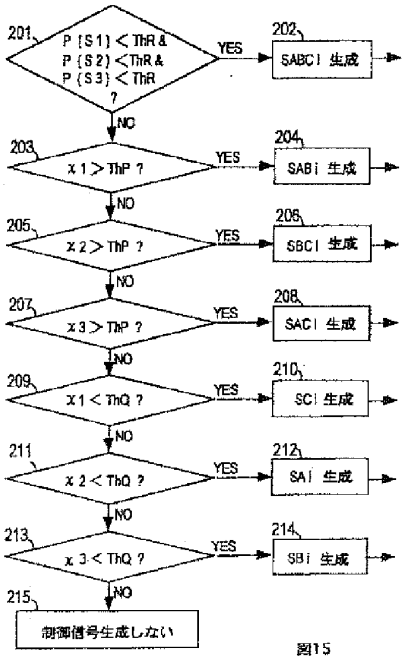


図15

図３

【図５】

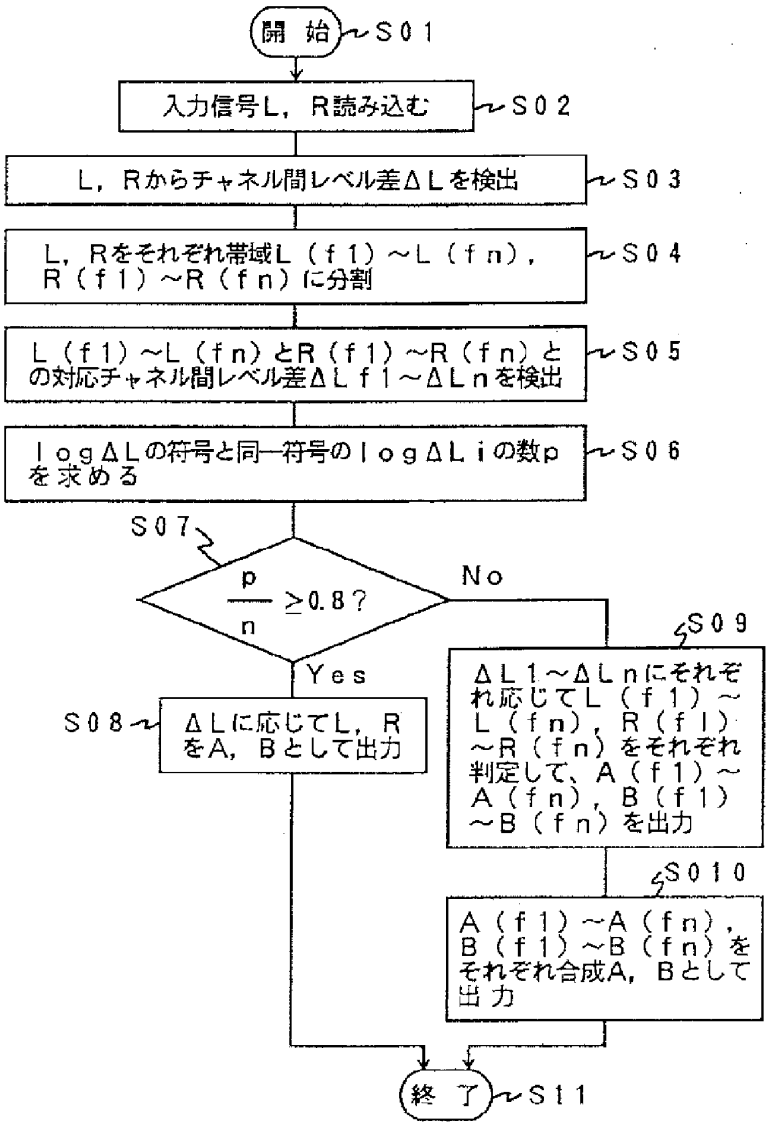

開 始
S01
入力信号L，R読み込む
S02
L，Rからチャネル間レベル差ΔLを検出
S03
L，Rをそれぞれ帯域L（f1）～L（fn），R（f1）～R（fn）に分割
S04
L（f1）～L（fn）とR（f1）～R（fn）との対応チャネル間レベル差ΔLf1～ΔLnを検出
S05
logΔLの符号と同一符号のlogΔLiの数pを求める
S06
S07
p/n ≧0.8？
No
Yes
S08
ΔLに応じてL，RをA，Bとして出力
S09
ΔL1～ΔLnにそれぞれ応じてL（f1）～L（fn），R（f1）～R（fn）をそれぞれ判定して、A（f1）～A（fn），B（f1）～B（fn）を出力
S010
A（f1）～A（fn），B（f1）～B（fn）をそれぞれ合成A，Bとして出力
終 了
S11

図 5

【図６】

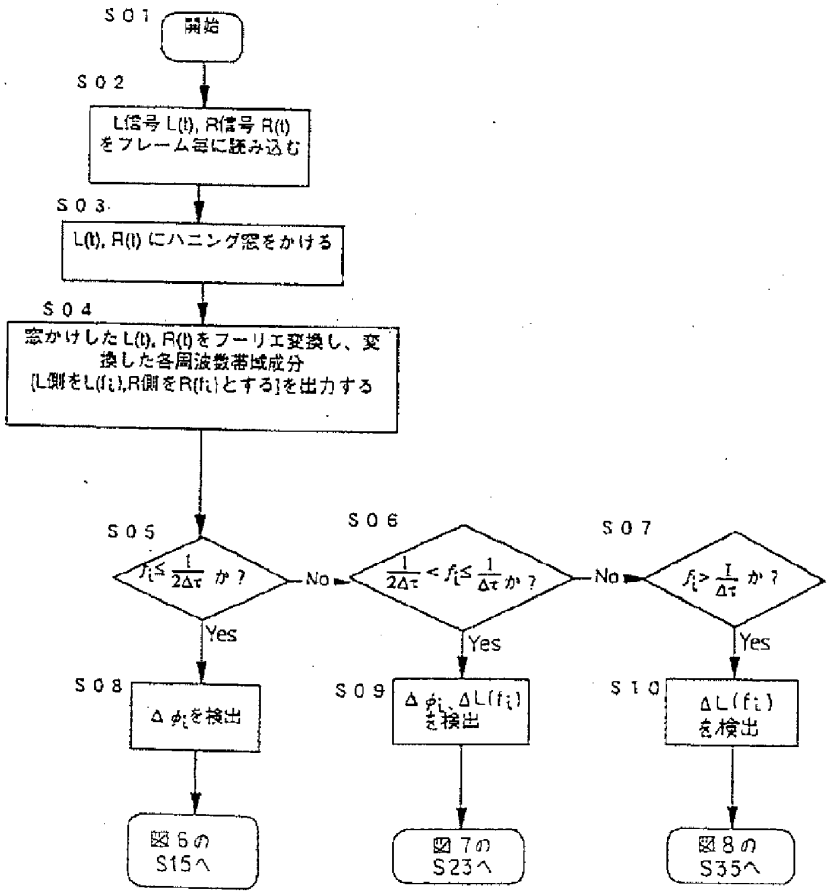

S01
開始
S02
L信号 L(t), R信号 R(t)
をフレーム毎に読み込む
S03
L(t), R(t) にハニング窓をかける
S04
窓かけしたL(t), R(t)をフーリエ変換し、変換した各周波数帯域成分
[L側をL(fi),R側をR(fi)とする]を出力する
S05
fi≦1/2Δτ か？
No
S06
1/2Δτ < fi≦1/Δτ か？
No
S07
fi>1/Δτ か？
Yes
Yes
Yes
S08
Δφiを検出
S09
Δφi、ΔL(fi)
を検出
S10
ΔL(fi)
を検出
図6の
S15へ
図7の
S23へ
図8の
S35へ

図６

【図７】

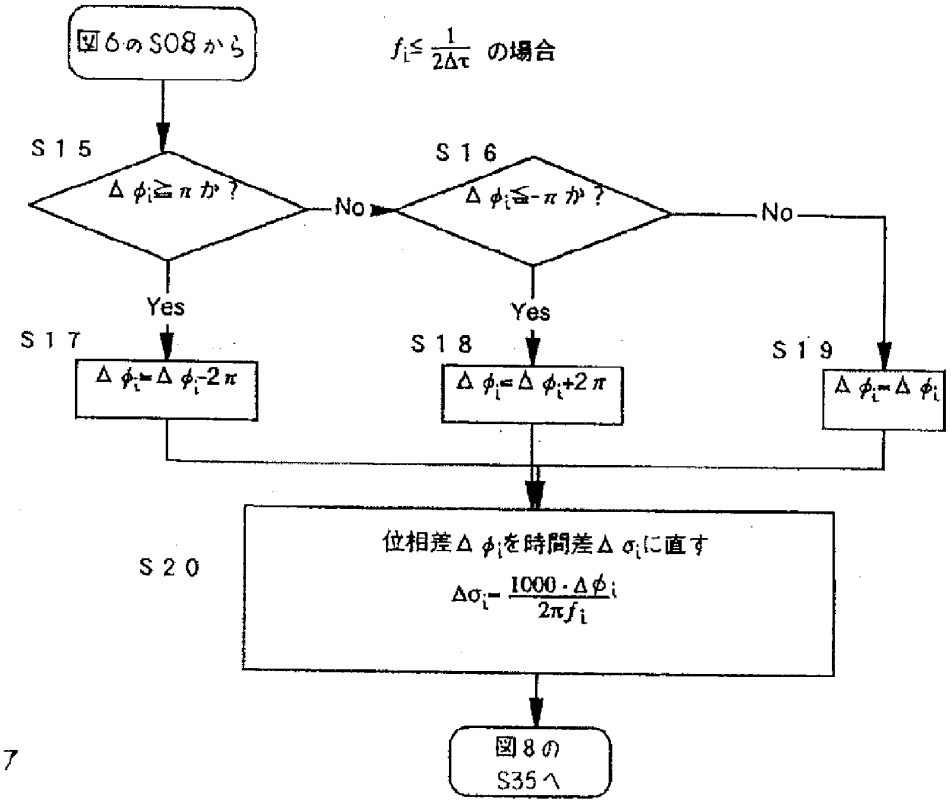

図６のS08から
$f_i \leq \frac{1}{2\Delta\tau}$ の場合
S15
$\Delta\phi_i \geqq \pi$ か？
No
S16
$\Delta\phi_i \leqq -\pi$ か？
No
Yes
Yes
S17
$\Delta\phi_i = \Delta\phi_i - 2\pi$
S18
$\Delta\phi_i = \Delta\phi_i + 2\pi$
S19
$\Delta\phi_i = \Delta\phi_i$
S20
位相差$\Delta\phi_i$を時間差$\Delta\sigma_i$に直す
$\Delta\sigma_i = \frac{1000 \cdot \Delta\phi_i}{2\pi f_i}$
図８のS35へ

図 7

【図１０】

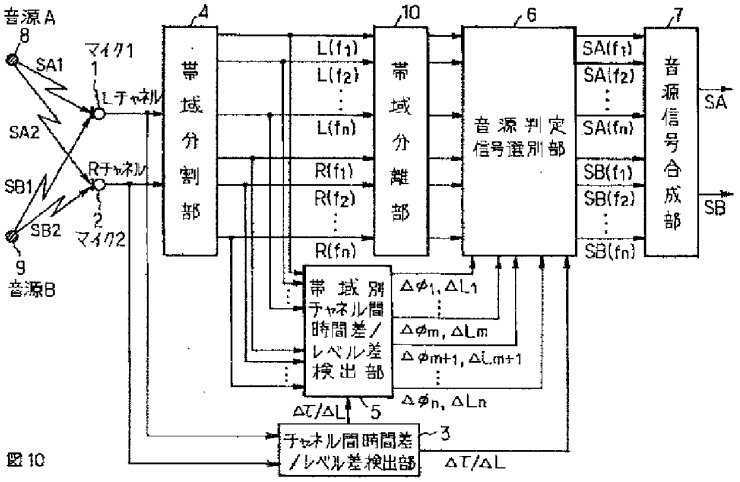

音源A
8
SA1
SA2
SB1
SB2
9
音源B
マイク1
1
Lチャネル
Rチャネル
2
マイク2
4
帯域分割部
L(f1)
L(f2)
L(fn)
R(f1)
R(f2)
R(fn)
10
帯域分離部
6
音源判定信号識別部
SA(f1)
SA(f2)
SA(fn)
SB(f1)
SB(f2)
SB(fn)
7
音源信号合成部
SA
SB
帯域別チャネル間時間差/レベル差検出部
Δφ1, ΔL1
Δφm, ΔLm
Δφm+1, ΔLm+1
Δφn, ΔLn
ΔT/ΔL
5
チャネル間時間差/レベル差検出部
3
ΔT/ΔL

図 10

【図８】

$\frac{1}{2\Delta\tau} < f_i \leq \frac{1}{\Delta\tau}$の場合

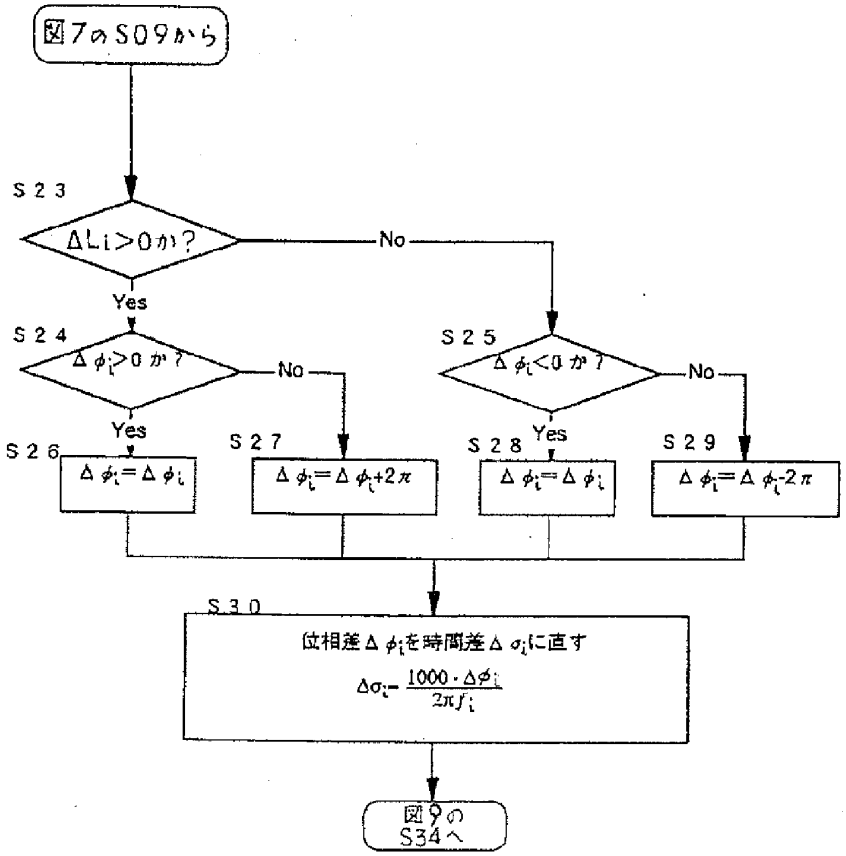


図 8

【図９】

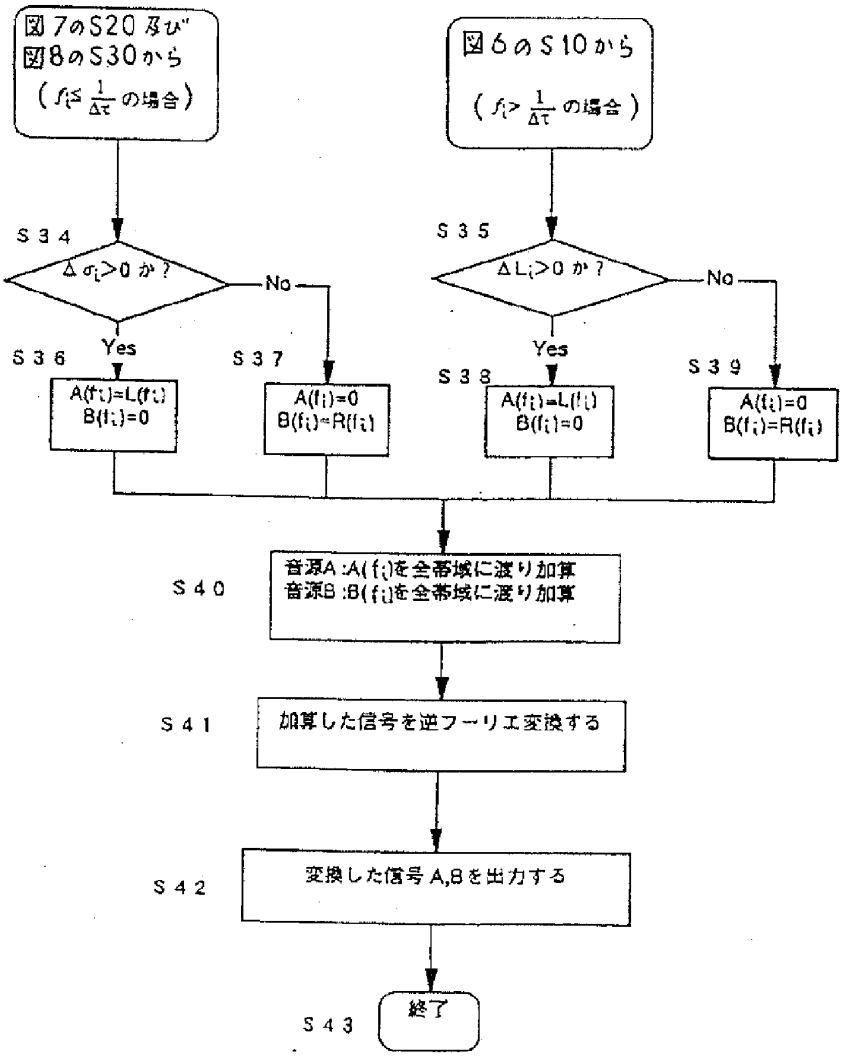

図７のS20及び
図８のS30から
($f_i \leq \frac{1}{\Delta\tau}$ の場合)
図６のS10から
($f_i > \frac{1}{\Delta\tau}$ の場合)
S34
$\Delta\sigma_i > 0$ か？
No
Yes
S35
$\Delta L_i > 0$ か？
No
Yes
S36
$A(f_i)=L(f_i)$
$B(f_i)=0$
S37
$A(f_i)=0$
$B(f_i)=R(f_i)$
S38
$A(f_i)=L(f_i)$
$B(f_i)=0$
S39
$A(f_i)=0$
$B(f_i)=R(f_i)$
S40
音源A：$A(f_i)$を全帯域に渡り加算
音源B：$B(f_i)$を全帯域に渡り加算
S41
加算した信号を逆フーリエ変換する
S42
変換した信号A,Bを出力する
S43
終了

図 9

【図１１】

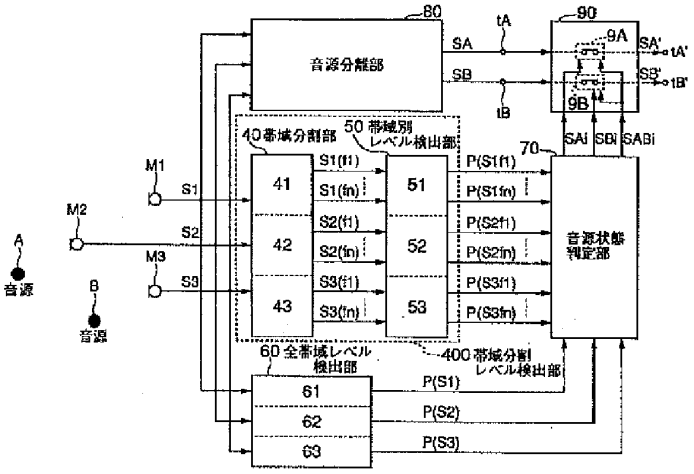


図 11

【図１２】

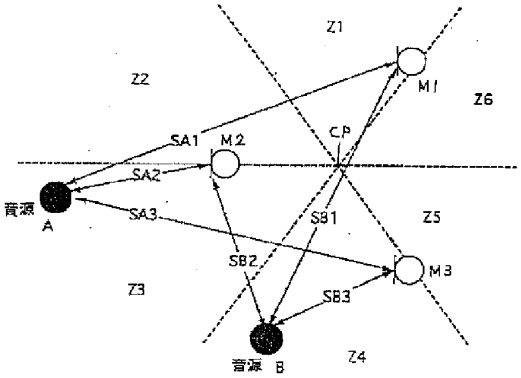


図 12

【図１６】

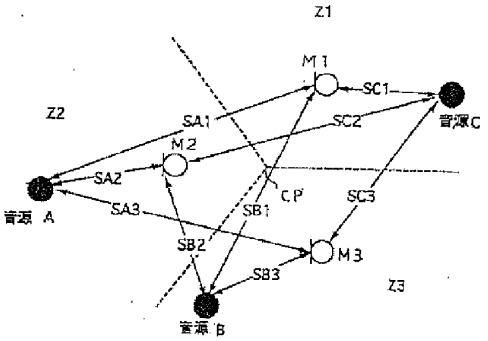


図 16

【図１７】

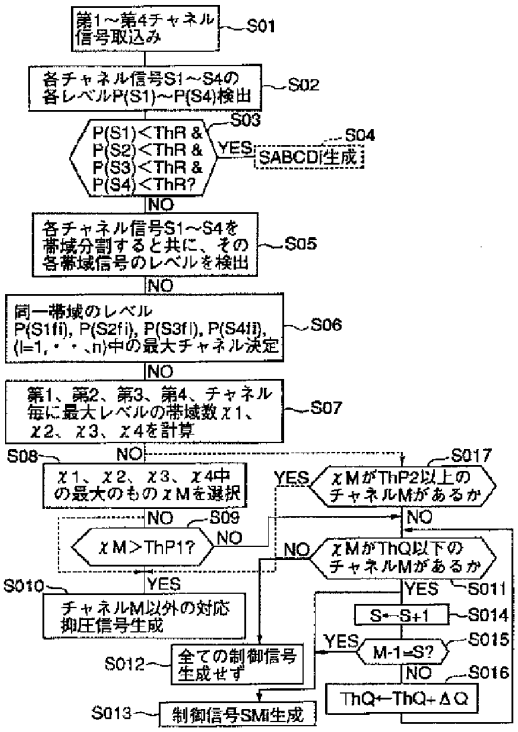


図　17

【図２１】

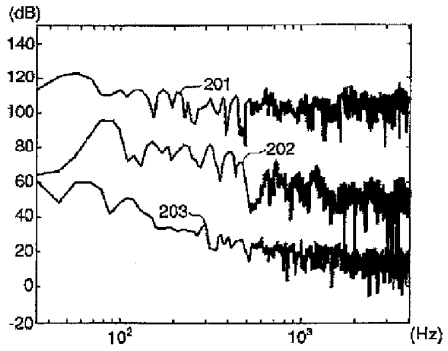


図　21

【図１８】

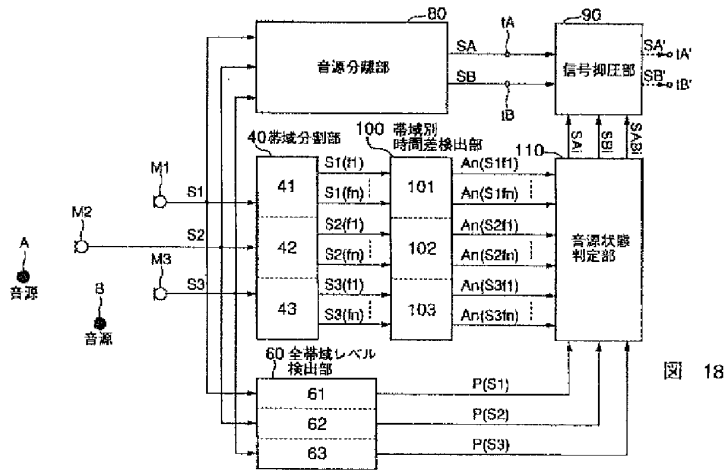


図　18

【図１９】

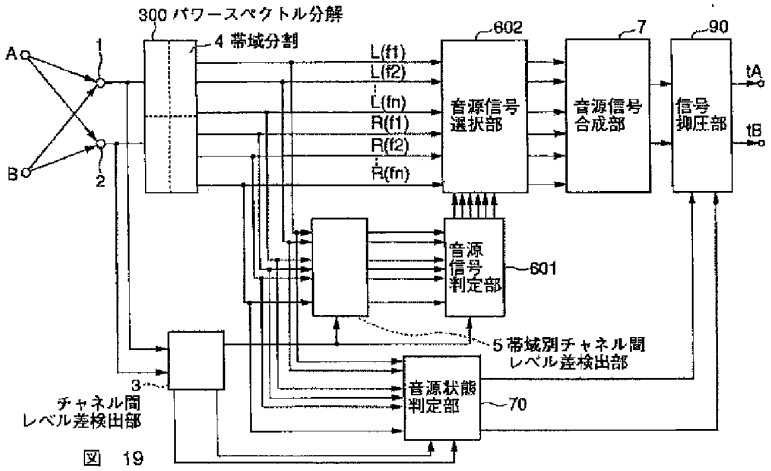


図　19

【図２２】

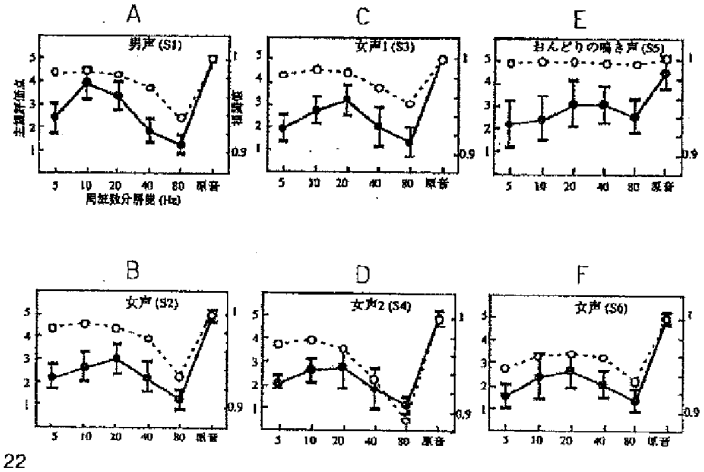


図　22

【図２３】

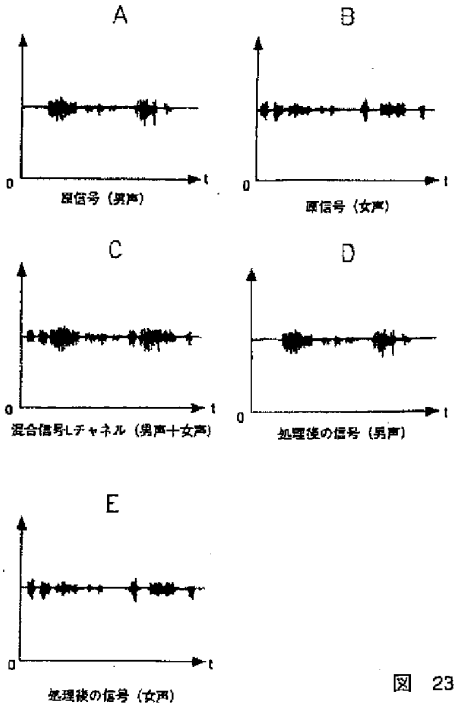


図 23

【図２４】

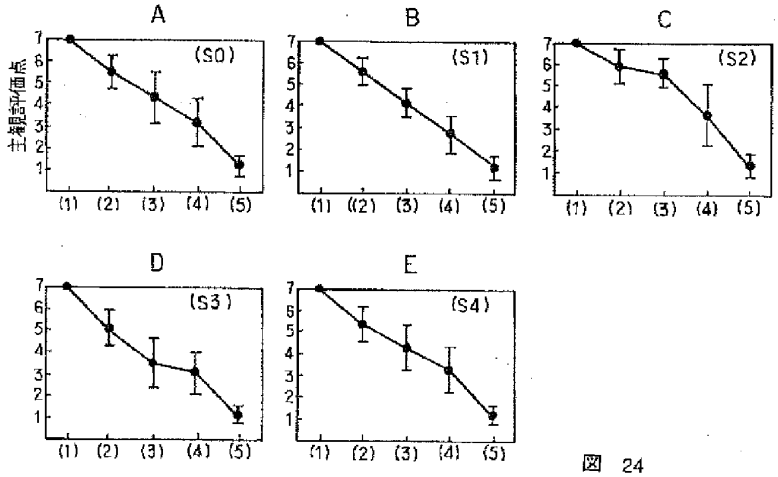


図 24

【手続補正書】
【提出日】平成９年９月２５日
【手続補正１】
【補正対象書類名】明細書
【補正対象項目名】請求項２０
【補正方法】変更
【補正内容】
【請求項２０】　請求項１９の方法において、
上記第１判定過程で、第１基準値を越えるものがないと判定されると、上記最も速く到達する帯域の数が、上記第１基準値よりも小さい第２基準値より小さいか否かを判定する第２判定過程と、
その第２判定過程で、第２基準値より小さいと判定されると、その小さいと判定された最も速い到達時間の帯域数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する過程とを有することを特徴とする音源分離方法。
【手続補正２】
【補正対象書類名】明細書
【補正対象項目名】請求項２１
【補正方法】変更
【補正内容】
【請求項２１】　請求項１６又は２０の方法において、音源が４個以上の場合で、上記第２判定過程で、第２基準値より小さいと判定されると、上記第２基準値を上記第１基準値を越えない範囲内で、順次大きくして、上記第２判定過程と同じ判定を、（Ｍ－２）回以内、Ｍは音源の数、繰返す過程を有することを特徴とする音源分離方法。
【手続補正３】
【補正対象書類名】明細書
【補正対象項目名】請求項３３
【補正方法】変更
【補正内容】
【請求項３３】　請求項３２の装置において、
上記第１判定手段で、第１基準値を越えるものがないと判定されると、上記最もレベルが大きい帯域の数が、上＊記第１基準値よりも小さい第２基準値以下か否かを判定する第２判定手段と、
その第２判定手段で、第２基準値より小さいと判定されると、その小さいと判定された最もレベルが大きい帯域の数と対応するチャネルのマイクロホン位置から、発音していない１個の音源として検出する手段とを有することを特徴とする音源分離装置。
【手続補正４】
【補正対象書類名】明細書
【補正対象項目名】請求項３５
【補正方法】変更
【補正内容】
【請求項３５】　請求項３４の装置において、
上記音源状態判定手段は、上記各帯域別到達時間差比較で最も速く音源信号が到達したチャネルを帯域ごとに決定する手段と、
各チャネルごとに最も速く到達した帯域の数が第１基準値を越えるか否かを判定する第１判定手段と、
その第１判定手段が第１基準値を越えると判定すると、その越えた最も速く到達した帯域数と対応するチャネルのマイクロホン位置から発音している１個の音源を推定する手段と、
その推定された音源以外の音源を発音していないものとして検出する手段とを有することを特徴とする音源分離装置。
【手続補正５】
【補正対象書類名】明細書
【補正対象項目名】請求項４０
【補正方法】変更
【補正内容】
【請求項４０】　請求項３９の記録媒体において、
上記チャネル時間差検出過程は各出力チャネル信号間の相互相関を求め、相互相関の各ピークとなる、その出力チャネル信号間の各時間差を上記各チャネル間時間差として求めることを特徴とする記録媒体。

フロントページの続き

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｈ０４Ｓ　7/00		Ｈ０４Ｓ　7/00	Ｚ

(72)発明者　西野　豊
東京都新宿区西新宿三丁目19番２号　日本電信電話株式会社内

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **2005-195955**
(43)Date of publication of application : **21.07.2005**

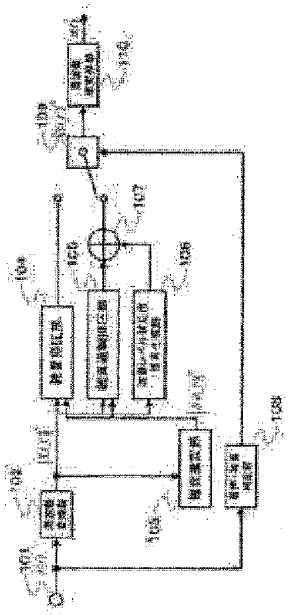
(51)Int.Cl. **G10L 15/20**
G10L 11/02
G10L 15/04
G10L 15/28
G10L 19/00
G10L 21/02
H04M 1/60
H04R 3/00

(21)Application number : **2004-003108**
(22)Date of filing : **08.01.2004**

(71)Applicant : **TOSHIBA CORP**
(72)Inventor : **AMADA KO**
KAWAMURA AKINORI
KOSHIBA AKINORI

(54) DEVICE AND METHOD FOR NOISE SUPPRESSION

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a device and method for noise suppression with which neither musical noise in a noise section nor distortion in a speech section is generated.
SOLUTION: Disclosed is the noise suppressing device which suppresses a noise signal in an input signal wherein the noise signal and a target signal are mixed together. The noise suppressing device comprises a noise estimation part 103 which estimates a noise signal component from the input signal, a speech/noise judgement part 108 which judges a target signal section and a noise signal section from the input signal, a noise suppression part 104 which performs noise suppression based upon a 1st suppression coefficient from the input signal and estimated noise signal, a noise excess suppression part 105 which performs noise suppression on the basis of a 2nd suppression coefficient larger than the 1st suppression coefficient from the input signal and estimated noise signal, and a switching part 109 which performs switching between the output signal of the noise suppression part 104 and the output signal of the noise excess suppression part 105 according to the judgement result of a section judgement means.



(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号

特開2005-195955

(P2005-195955A)

(43)公開日　平成17年7月21日(2005.7.21)

(51)Int.Cl.[7]	F I			テーマコード（参考）
G10L 15/20	G10L	3/02	301D	5D015
G10L 11/02	H04M	1/60	A	5D020
G10L 15/04	H04R	3/00	320	5K027
G10L 15/28	G10L	3/00	513A	
G10L 19/00	G10L	3/02	301E	

審査請求 有　請求項の数 11　OL　（全 18 頁）　最終頁に続く

(21)出願番号　特願2004-3108(P2004-3108)

(22)出願日　平成16年1月8日(2004.1.8)

(71)出願人　000003078
株式会社東芝
東京都港区芝浦一丁目１番１号

(74)代理人　100083161
弁理士　外川　英明

(72)発明者　天田　皇
神奈川県川崎市幸区小向東芝町１番地　株式会社東芝研究開発センター内

(72)発明者　河村　聡典
神奈川県川崎市幸区小向東芝町１番地　株式会社東芝研究開発センター内

(72)発明者　小柴　亮典
神奈川県川崎市幸区小向東芝町１番地　株式会社東芝研究開発センター内

最終頁に続く

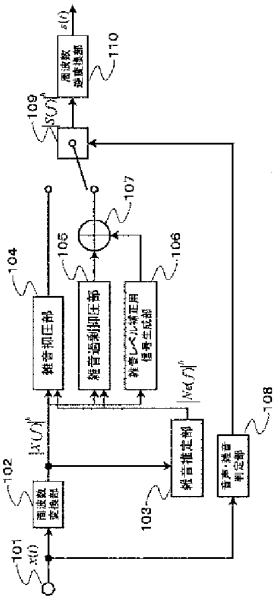
(54)【発明の名称】雑音抑圧装置及び雑音抑圧方法

(57)【要約】

【課題】　雑音区間ではミュージカルノイズが発生せず、音声区間での歪みも発生しない雑音抑圧装置、及び雑音抑圧方法を提供することを目的とする。

【解決手段】　雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、入力信号から雑音信号成分を推定する雑音推定部103と、入力信号から目的信号区間と雑音信号区間を判定する音声・雑音判定部108と、入力信号と推定雑音信号とから第１の抑圧係数に応じて雑音抑圧をする雑音抑圧部104と、入力信号と推定雑音信号とから第１の抑圧係数よりも大きな第２の抑圧係数に応じて雑音抑圧をする雑音過剰抑圧部105と、区間判定手段の判定結果に応じて雑音抑圧部104の出力信号と雑音過剰抑圧部105の出力信号とを切替える切替部109とから構成される。

【選択図】　図１



【特許請求の範囲】
【請求項１】
雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、前記入力信号から雑音信号成分を推定する雑音推定手段と、前記入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて前記入力信号から前記推定雑音信号成分を引き去る雑音抑圧手段とを具備したことを特徴とする雑音抑圧装置。
【請求項２】
雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、前記入力信号から雑音信号成分を推定する雑音推定手段と、前記入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記入力信号と前記推定雑音信号とから第１の抑圧係数に応じて雑音抑圧をする雑音抑圧手段と、前記入力信号と前記推定雑音信号とから前記第１の抑圧係数よりも大きな第２の抑圧係数に応じて雑音抑圧をする雑音過剰抑圧手段と、前記区間判定手段の判定結果に応じて前記雑音抑圧手段の出力信号と前記雑音過剰抑圧手段の出力信号とを切替える切替手段とを具備したことを特徴とする雑音抑圧装置。
【請求項３】
更に、目的信号区間の出力に残留する雑音信号とのレベルの違いを補正する係数を前記入力信号に乗じた補正用信号を生成する補正用信号生成手段と、前記補正用信号と前記雑音過剰抑圧手段の出力とを加算する加算手段とを具備し、前記切替手段は前記雑音抑圧手段の出力信号と前記加算手段の出力信号とを切替えることを特徴とする請求項２記載の雑音抑圧装置。
【請求項４】
雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、前記入力信号から雑音信号成分を推定する雑音推定手段と、前記入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記入力信号と前記推定雑音信号とから第１の抑圧係数を算出する抑圧係数算出手段と、前記入力信号と前記推定雑音信号とから前記第１の抑圧係数よりも大きな第２の抑圧係数を算出する過剰抑圧係数算出手段と、前記区間判定手段の判定結果に応じて前記第１の抑圧係数と前記第２の抑圧係数とを切替える切替手段と、前記切替手段により切替えられた抑圧係数を前記入力信号に乗じる乗算手段とを具備したことを特徴とする雑音抑圧装置。
【請求項５】
更に、前記入力信号から目的信号区間の出力に残留する雑音信号とのレベルの違いを補正する係数を生成する補正用係数生成手段と、前記補正用係数と前記第２の抑圧係数とを加算する加算手段とを具備し、前記切替手段は前記第１の抑圧係数と前記加算手段で加算された係数とを切替えることを特徴とする請求項４記載の雑音抑圧装置。
【請求項６】
前記区間判定手段は、前記入力信号と前記推定雑音信号とから目的信号区間と雑音信号区間とを判定することを特徴とする請求項１乃至請求項５記載の雑音抑圧装置。
【請求項７】
前記補正用信号生成手段は、予め保持された重畳用信号から補正用信号を生成することを特徴とする請求項３記載の雑音抑圧装置。
【請求項８】
雑音信号と目的信号が混合した複数の入力信号から雑音信号を抑圧する雑音抑圧装置において、前記複数の入力信号から目的信号が強調される統合信号を生成する統合信号生成手段と、前記統合信号から雑音信号成分を推定する雑音推定手段と、前記複数の入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて前記統合信号から前記推定雑音信号成分を引き去る雑音抑圧手段とを具備したことを特徴とする雑音抑圧装置。
【請求項９】

雑音信号と目的信号が混合した複数の入力信号から雑音信号を抑圧する雑音抑圧装置において、前記複数の入力信号から目的信号が強調される統合信号を生成する統合信号生成手段と、前記複数の入力信号から目的信号が抑圧された目的音除去信号を生成する目的音除去信号生成手段と、前記統合信号と前記目的音除去信号から雑音信号成分を推定する雑音推定手段と、前記複数の入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて前記統合信号から前記推定雑音信号成分を引き去る雑音抑圧手段とを具備したことを特徴とする雑音抑圧装置。

【請求項１０】

雑音信号と目的信号が混合した複数の入力信号から雑音信号を抑圧する雑音抑圧装置において、前記複数の入力信号から周波数帯域ごとに目的信号が強調されるサブバンド統合信号を生成するサブバンド統合信号生成手段と、前記サブバンド統合信号から各サブバンド毎の雑音信号成分を推定する雑音推定手段と、前記複数の入力信号から各サブバンド毎に目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて各サブバンド毎に前記サブバンド統合信号から前記推定雑音信号成分を引き去る雑音抑圧手段と、各サブバンド毎の雑音抑圧手段の出力信号を合成する合成手段とを具備したことを特徴とする雑音抑圧装置。

【請求項１１】

雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧方法において、前記入力信号から雑音信号成分を雑音推定手段により推定し、前記入力信号から目的信号区間と雑音信号区間を区間判定手段により判定し、前記区間判定手段の判定結果に基づいて雑音抑圧手段により前記入力信号から前記推定雑音信号成分を引き去ることを特徴とする雑音抑圧方法。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、ハンズフリー通話や音声認識等で用いられる雑音抑圧技術の一つであり、入力音響信号から目的とする音声信号を強調して出力する技術に関する。

【背景技術】

【０００２】

実環境での音声認識や携帯電話の実用化に伴い、雑音の重畳した信号から雑音を取り除き音声信号のみを強調する信号処理方法が重要になってきている。スペクトルサブトラクション(Spectral Subtraction：ＳＳ)は、効果的で実現がしやすいためしばしば用いられる（例えば、非特許文献１を参照）。

【０００３】

スペクトルサブトラクションにはミュージカルノイズと呼ばれる聴覚上不自然に聞こえる音が生成される問題がある。これは雑音区間で特に顕著であり、実際にはバラツキが存在する入力信号（雑音信号）からその平均値を引き去ることで、消し残しの成分が不連続に存在することに起因する。この問題を解決する方法として、過剰抑圧を行うという方法がある。過剰抑圧とは推定ノイズよりも大きな値を引去り、雑音の変動成分も含めて抑圧してしまう方法である。なお、減算により負の値になる場合は最小値で置き換えるなどの処理が行われる。しかし、過剰抑圧は、音声区間で抑圧量が過剰になり、音声が歪んでしまうといった問題があった（例えば、非特許文献２を参照）。

【０００４】

また、ミュージカルノイズの発生した区間に何らかの処理を施して目立たなくする方法、例えば入力信号などに小さなゲインをかけて加えるなどの方法もあるが、この方法ではミュージカルノイズが知覚できなくなるまで十分な信号を重畳すると、重畳した信号により雑音レベルが上がり、雑音抑圧の効果が失われかねない問題があった。

【０００５】

【非特許文献１】S.Boll,"Suppression of Acoustic Noise in Speech Using SpectralSubtraction",IEEE Trans., ASSP-27, No.2, pp.113-120,1979

【非特許文献２】Z.Goh,K.Tan and B.T.G.Tan,"Postprocessing Method for Suppressing MusicalNoise Generated by spectral Subtraction",IEEE Trans.,SAP-6, No. 3, May 1998

【発明の開示】

【発明が解決しようとする課題】

【０００６】

上述したように、抑圧係数を大きくして過剰抑圧することは、ミュージカルノイズを押さえる効果はあるものの、音声区間での歪みを生みやすいという問題があった。また、ミュージカルノイズに入力信号を重畳するなどの後処理を用いた手法では、ミュージカルノイズを知覚できなくするに十分な音量を重畳すると、雑音抑圧の効果が失われる問題があった。

【０００７】

本発明は、このような課題を解決するためになされたものであり、雑音区間ではミュージカルノイズが発生せず、音声区間での歪みも発生しない雑音抑圧装置、及び雑音抑圧方法を提供することを目的とする。

【課題を解決するための手段】

【０００８】

上記の課題を解決するために本発明に係る雑音抑圧装置は、雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、前記入力信号から雑音信号成分を推定する雑音推定手段と、前記入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて前記入力信号から前記推定雑音信号成分を引き去る雑音抑圧手段とを具備したことを特徴とする。

【０００９】

また、雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、前記入力信号から雑音信号成分を推定する雑音推定手段と、前記入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記入力信号と前記推定雑音信号とから第１の抑圧係数に応じて雑音抑圧をする雑音抑圧手段と、前記入力信号と前記推定雑音信号とから前記第１の抑圧係数よりも大きな第２の抑圧係数に応じて雑音抑圧をする雑音過剰抑圧手段と、前記区間判定手段の判定結果に応じて前記雑音抑圧手段の出力信号と前記雑音過剰抑圧手段の出力信号とを切替える切替手段とを具備したことを特徴とする。

【００１０】

更に、目的信号区間の出力に残留する雑音信号とのレベルの違いを補正する係数を前記入力信号に乗じた補正用信号を生成する補正用信号生成手段と、前記補正用信号と前記雑音過剰抑圧手段の出力とを加算する加算手段とを具備し、前記切替手段は前記雑音抑圧手段の出力信号と前記加算手段の出力信号とを切替えることを特徴とする。

【００１１】

また、雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧装置において、前記入力信号から雑音信号成分を推定する雑音推定手段と、前記入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記入力信号と前記推定雑音信号とから第１の抑圧係数を算出する抑圧係数算出手段と、前記入力信号と前記推定雑音信号とから前記第１の抑圧係数よりも大きな第２の抑圧係数を算出する過剰抑圧係数算出手段と、前記区間判定手段の判定結果に応じて前記第１の抑圧係数と前記第２の抑圧係数とを切替える切替手段と、前記切替手段により切替えられた抑圧係数を前記入力信号に乗じる乗算手段とを具備したことを特徴とする。

【００１２】

更に、前記入力信号から目的信号区間の出力に残留する雑音信号とのレベルの違いを補正する係数を生成する補正用係数生成手段と、前記補正用係数と前記第２の抑圧係数とを加算する加算手段とを具備し、前記切替手段は前記第１の抑圧係数と前記加算手段で加算された係数とを切替えることを特徴とする。

【００１３】
また、前記区間判定手段は、前記入力信号と前記推定雑音信号とから目的信号区間と雑音信号区間とを判定することを特徴とする。
また、前記補正用信号生成手段は、予め保持された重畳用信号から補正用信号を生成することを特徴とする。
また、雑音信号と目的信号が混合した複数の入力信号から雑音信号を抑圧する雑音抑圧装置において、前記複数の入力信号から目的信号が強調される統合信号を生成する統合信号生成手段と、前記統合信号から雑音信号成分を推定する雑音推定手段と、前記複数の入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて前記統合信号から前記推定雑音信号成分を引き去る雑音抑圧手段とを具備したことを特徴とする。
【００１４】
また、雑音信号と目的信号が混合した複数の入力信号から雑音信号を抑圧する雑音抑圧装置において、前記複数の入力信号から目的信号が強調される統合信号を生成する統合信号生成手段と、前記複数の入力信号から目的信号が抑圧された目的音除去信号を生成する目的音除去信号生成手段と、前記統合信号と前記目的音除去信号から雑音信号成分を推定する雑音推定手段と、前記複数の入力信号から目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて前記統合信号から前記推定雑音信号成分を引き去る雑音抑圧手段とを具備したことを特徴とする。
【００１５】
また、雑音信号と目的信号が混合した複数の入力信号から雑音信号を抑圧する雑音抑圧装置において、前記複数の入力信号から周波数帯域ごとに目的信号が強調されるサブバンド統合信号を生成するサブバンド統合信号生成手段と、前記サブバンド統合信号から各サブバンド毎の雑音信号成分を推定する雑音推定手段と、前記複数の入力信号から各サブバンド毎に目的信号区間と雑音信号区間を判定する区間判定手段と、前記区間判定手段の判定結果に基づいて各サブバンド毎に前記サブバンド統合信号から前記推定雑音信号成分を引き去る雑音抑圧手段と、各サブバンド毎の雑音抑圧手段の出力信号を合成する合成手段とを具備したことを特徴とする。
【００１６】
また、本発明に係る雑音抑圧方法は、雑音信号と目的信号が混合した入力信号から雑音信号を抑圧する雑音抑圧方法において、前記入力信号から雑音信号成分を雑音推定手段により推定し、前記入力信号から目的信号区間と雑音信号区間を区間判定手段により判定し、前記区間判定手段の判定結果に基づいて雑音抑圧手段により前記入力信号から前記推定雑音信号成分を引き去ることを特徴とする。
【発明の効果】
【００１７】
本発明によれば、音声区間に歪みを生むことなく雑音区間に不自然な消し残し音を発生させることなく雑音を抑圧することができる。
【発明を実施するための最良の形態】
【００１８】
以下、図面を参照して本発明の実施の形態について詳細に説明する。
【実施例１】
【００１９】
図１は本発明の第１の実施形態に係る雑音抑圧装置の構成を表すブロック図である。図１に示されるように、第１の実施形態の雑音抑圧装置は、音響信号を入力するための入力端子101と、音響信号を周波数領域に変換する周波数変換部102と、この出力から推定雑音を求める雑音推定部103 と、周波数変換部102 から雑音推定部103の出力を用いて雑音が抑圧された信号を生成する雑音抑圧部104と、同じくより強く雑音が抑圧された信号を生成する雑音過剰抑圧部105と、周波数変換部102の出力から雑音レベルを補正する信号を生成する雑音レベル補正用信号生成部106 と、雑音過剰抑圧部105と雑音レベル補正用信号

生成部106との出力を加算する加算部107と、入力信号から音声区間か雑音区間かを判定する音声・雑音判定部108と、音声・雑音判定部108の出力により雑音抑圧104の出力と加算部107の出力を選択する切替部109 と、この出力を時間領域に変換する周波数逆変換部110から構成される。

【００２０】

入力端子101には、

【００２１】

【数１】

$$x(t)=s(t)+n(t)$$

【００２２】

で表される信号が入力される。ここでｘ（ｔ）はマイクなどで受音した時間波形を表す信号であり、ｓ（ｔ）はその中の目的信号成分（例えば音声）であり、ｎ（ｔ）は非目的信号成分（例えば周囲の雑音）である。入力された信号ｘ（ｔ）は周波数変換部102 においてＤＦＴなどを用いて所定の窓幅で周波数領域に変換されＸ（ｆ）を得る。（ｆ は周波数を表す。）

【００２３】

雑音推定部103では、Ｘ（ｆ）から 雑音信号の推定値Ｎｅ（ｆ）を推定する。この推定には、例えばｓ（ｔ）が音声信号の場合、非発話区間が存在するので、その区間はｘ（ｔ）＝ｎ（ｔ）となり、その区間の平均値をＮｅ（ｆ）とする。これを用いて、

【００２４】

【数２】

$$|Se(f)|=|X(f)|-\alpha|Ne(f)|$$

【００２５】

として、音声の推定値|Se(f)|を得る。これを時間領域にもどすことで、音声のみを推定することができる。|Se(f)|は振幅値のみで位相項がないので、これには入力信号Ｘ（ｆ）の位相項を用いるのが一般的である。（数２）は振幅スペクトルで行う方法であるが、パワースペクトルを用いる方法もあり、一般的な表記を用いると

【００２６】

【数３】

$$|Se(f)|^b=|X(f)|^b-\alpha|Ne(f)|^b$$

【００２７】

と表すことができる。スペクトルサブトラクションをフィルタ演算とみなして、

【００２８】

【数４】

$$Se(f)=\left\{\frac{(|X(f)|^b-\alpha|Ne(f)|^b)}{|X(f)|^b}\right\}^{\left(\frac{1}{a}\right)}X(f)$$

【００２９】

と表記することもできる。（ａ，ｂ）＝（１，１）の場合は振幅スペクトルを用いたスペクトルサブトラクション（数２）と等価になる。また（ａ，ｂ）＝（２，２）の場合はパワースペクトルを用いたスペクトルサブトラクションとなる。さらに、（ａ，ｂ）＝（１

，２）かつα＝１の場合はWienerフィルタの形式となる。これらは実現上においては（数４）で統一的に記述できる同種の手法と見なすことができる。

【００３０】

ここで、一般にＸ（ｆ）は複素数であり、

【００３１】

【数５】

$$X(f)=|X(f)|\exp(j\arg(X(f)))$$

【００３２】

と表される。|X(f)|はＸ（ｆ）の大きさを、ａｒｇ（Ｘ（ｆ））は位相を、ｊは虚数単位である。周波数変換部102からはＸ（ｆ）の大きさが出力されるが、ここでは指数ｂを付加した一般的な表現を用いることにする。その理由は、スペクトルサブトラクションは（数３）で述べたように、いくつかのバリエーションが存在するためである。ｂの値は１または２である場合が多い。雑音推定部103は$|X(f)|^b$から推定雑音$|Ne(f)|^b$を求める。これには$|X(f)|^b$から雑音区間と見なされる区間の平均値を用いる。

【００３３】

例えば、雑音区間において

【００３４】

【数６】

$$|Ne(f,n)|^b=\delta|Ne(f,n-1)|^b+(1-\delta)|X(f)|^b$$

【００３５】

とする方法などがある。ただし、$|Ne(f,n)|^b$は現在フレームの$|Ne(f)|^b$で$|Ne(f,n-1)|^b$は１つ前のフレームの値であり、δは０＜δ＜１なる値で、平滑化の度合いを制御する。音声区間か否かは$|X(f)|^b$の大きさが大きい区間を音声区間とする方法や、$|X(f)|^b$と$|Ne(f,n)|^b$の比率を求め、$|X(f)|^b$がある比率よりも大きくなる区間を音声とするなどの方法がある。

【００３６】

雑音抑圧部104、雑音過剰抑圧部105では周波数変換部102の出力$|X(f)|^b$から雑音推定部103の出力$|Ne(f)|^b$を引き去り出力信号$|S(f)|^b$を出力する。これには（数３）の方法を用いるが、推定雑音|Ne(f)|が入力信号|X(f)|よりも大きい場合などいくつかの処理方法がある。ここでは、

【００３７】

【数７】

$$|Se(f)|^b=Max(|X(f)|^b-\alpha|Ne(f)|^b,\beta|X(f)|^b)$$

【００３８】

を用いることにする。ここで、Ｍａｘ（ｘ，ｙ）はｘ，ｙの大きい方を表し、αは抑圧係数、βはフロアリング係数である。αの値は大きいほど多くの雑音を取り除くことができるので、雑音抑圧効果は大きくなるが、音声が存在する区間では音声成分も引き去られ出力信号に歪みを生じる。βは正の小さな値で演算結果が負になることを抑止する。例えば（α、β）＝（1.0，0.01）等である。

【００３９】

本発明においては雑音抑圧部104の抑圧係数（αｓ）よりも過剰雑音抑圧部105の抑圧係数（αｎ）を大きくしている。また、過剰雑音抑圧部105では大きな抑圧係数が用いられているため雑音の平均的なパワー（雑音レベル）が雑音抑圧部104に比べて下がる。これを補償する手段として雑音レベル補正用信号生成部106を用いている。

【００４０】
ここでは、入力信号$|X(f)|^b$にゲインをかけた信号
【００４１】
【数８】

$$|M(f)|^b = (1-\alpha s)|X(f)|^b$$

【００４２】
を生成し、これを雑音過剰抑圧部105の出力に加算部107において加える。
切替部109では雑音抑圧部104と加算部107の出力を選択して出力信号を生成する。切替えは音声・雑音判定部108の判定結果に基づき、音声区間では雑音抑圧部104の出力を、雑音区間では加算部107の出力を選択する。音声・雑音判定部108での判定方法は様々な方法が存在するが、例えば信号のパワーと閾値を用いて判定する方法などがある。
【００４３】
最後に周波数逆変換部110で切替部109の出力を周波数領域から時間領域に変換され音声の強調された時間信号が得られる。さらに、フレーム単位で処理する場合はオーバーラップアドにより時間的に連続した信号を生成する場合でも適用できる。また周波数逆変換部110を用いずに時間領域に変換せず周波数領域のまま出力してもよい。
【００４４】
雑音過剰抑圧部105と雑音レベル補正用信号生成部106に関して詳細に説明する。上述したようにスペクトルサブトラクションにはミュージカルノイズと呼ばれる雑音区間での引き残しが不自然な音になる現象が存在する。図２を用いてこの現象を模式的に説明する。図２（ａ）は周波数変換した入力信号のある周波数ｆの振幅値（|X(f)|）をフレーム（時刻）ごとに表したものである。ここではｂ＝１として指数部を省略して表す。空白の箱が|X(f)|の雑音成分で斜線の箱が音声成分である。３本の点線のうち中央が雑音推定部により出力された推定雑音の大きさ|Ne(f)|を示し、上側が過剰抑圧を行う場合の値$\alpha n|Ne(f)|$で、下側が通常の抑圧を行う場合の値 $\alpha s|Ne(f)|$である。まず、$\alpha = 1$で抑圧を行うと|Ne(f)|だけ振幅が減少し図２（ｂ）のようになる。これは通常のスペクトルサブトラクションであり、雑音区間のノイズが減少し、音声が強調されている。しかし、雑音区間に引き残し成分が間欠的に存在しミュージカルノイズとなって聞こえる。また、音声区間では引きすぎにより音声成分の一部が欠けてしまう。これは音声の歪みとなって知覚される。
【００４５】
図２（ｃ）は$\alpha n|Ne(f)|$で過剰抑圧を行った場合である。雑音区間は完全に抑圧されミュージカルノイズは発生していないが、音声成分がかなり削られ大きな歪みが発生する。図２（ｄ）は$\alpha s|Ne(f)|$で抑圧を行った場合である。音声成分に歪みは出ていないが、雑音区間に信号が間欠的に残る現象がまだ存在する。本発明は、図２（ｅ）に示すように音声区間と雑音区間を予め区別しておき、音声区間では歪みの生じない図２（ｄ）の方法で抑圧し、雑音区間は過剰抑圧により図２（ｃ）のように強い抑圧を行いミュージカルノイズを完全に除去される。
【００４６】
ところで、図２（ｅ）では雑音区間では雑音は完全に除去されているものの、音声区間では歪みを発生させないかわりに雑音も残っているため、この雑音が知覚され、雑音レベルが不連続に聞こえる場合がある。この問題を解決するため、図２（ｆ）に示したように雑音区間にのみ、入力信号のレベルを低減させた信号を加算することにより雑音レベルを揃える。以上が本発明の模式的な説明である。厳密には雑音と音声を加算した信号の振幅はそれぞれの振幅の和になるとは限らないなど正確な表現になっていない点は考慮しておく必要がある。
【００４７】
本発明ではミュージカルノイズを消しているのは過剰抑圧であり、入力信号の加算は音

声区間との雑音レベルの違いを埋めるために行っている。これは、ミュージカルノイズを入力音声の加算で知覚しにくくする従来の方法とは異なる。従って本発明では、音声区間での抑圧係数を大きくとることにより、雑音区間で付加する信号のレベルを小さくすることが可能であり、この操作によりミュージカルノイズの削減効果が左右されることはない。

【００４８】

一方、従来は加算する信号のレベルとミュージカルノイズの知覚されやすさとは密接な関係があり、加算量を少なくするとミュージカルノイズは知覚されやすくなる。（数８）で用いられている入力信号に対するゲイン（１－αｓ）は次のように求められる。

【００４９】

まず、音声区間で歪みを生じないように抑圧係数αｓが弱めに設定されるので、αｓは１より小さい値となる。したがって、仮に音声区間が雑音のみであった場合、（１－αｓ）の雑音は引かれずに残ることになる。一方雑音区間では過剰抑圧により雑音はゼロになっている。したがって、その差（１－αｓ）分の信号を雑音区間に加えれば音声区間の雑音とレベルがそろうことになる。

【００５０】

ところで、音声区間の抑圧量αｓが１に近い場合、付加する雑音のゲイン（１－αｓ）の値は小さな値となる。このような場合は音声区間と雑音区間の雑音レベルの差が知覚されにくいため、加算そのものを行わないという方法であってもよい。また、分散の大きな雑音の場合は、この方法でもレベル差が完全に補償できないことがあり、その場合は分散を考慮した補償方法を用いることも可能である。

【００５１】

図２（ｇ）は全区間雑音と誤った判定がされた場合の過剰抑圧後の状態を模式的に表している。上述している通り、過剰抑圧を行うと雑音区間ではミュージカルノイズは生じないが、音声区間に大きな歪みを生む。ここではこの後に入力信号の加算を行うため、誤って雑音と判断された音声区間には雑音成分とともに音声成分も加算されることになり、一度生じた歪みを回復させる効果がある（図２（ｈ））。つまり、音声区間を雑音区間と誤った場合でも、音声が誤って抑圧されることがない音声・雑音判定結果の誤りに対して頑健であるという効果がある。

【実施例２】

【００５２】

図３は本発明の第２の実施形態に係る雑音抑圧装置の構成を示すブロック図である。第２の実施形態の雑音抑圧装置は、上述した第１の実施形態におけるスペクトルサブトラクションは伝達関数を乗算する形式にした場合の構成であり、第１の実施形態では（数３）に相当する減算形の抑圧方法であるのに対し、第２の実施形態は（数４）の乗算形に相当する。これらは本質的には同じであるため、以降の実施形態においても（数３）に相当する減算形の方法で実現することも可能である。第２の実施形態と第１の実施形態との違いは、雑音抑圧部104、雑音過剰抑圧部105、雑音レベル補正用信号生成部106が抑圧係数算出部204、過剰抑圧係数算出部205、雑音レベル補正用係数生成部206にそれぞれ置き換わり、切替部 107 の出力である重み係数を入力信号に乗算する乗算部211が加わっている点である。

【００５３】

抑圧係数算出部204では抑圧係数は、

【００５４】

【数９】

$$ws(f)=Max\left(\frac{(|X(f)|^b-\alpha s|Ne(f)|^b)}{|X(f)|^b},\beta\right)^{\left(\frac{1}{a}\right)}$$

【００５５】
と求められ、過剰抑圧係数算出部205では
【００５６】
【数１０】

$$wn(f)=Max\left(\frac{(|X(f)|^b-\alpha n|Ne(f)|^b)}{|X(f)|^b},\beta\right)^{\left(\frac{1}{a}\right)}$$

【００５７】
と求められる。
既に述べたように（ａ，ｂ）＝（１，１）の場合は振幅スペクトルを用いたスペクトルサブトラクションと等価であり、（ａ，ｂ）＝（２，２）の場合はパワースペクトルを用いたスペクトルサブトラクション、（ａ，ｂ）＝（１，２）の場合はWiener フィルタの形式となる。また、抑圧係数は抑圧係数算出部204 ではαｓであり、音声区間に歪みを与えない抑圧量が設定されるのに対して、過剰抑圧係数算出部205ではαｎとなり、雑音区間でミュージカルノイズを十分に除去するため大きな係数が設定される点も第１の実施形態と同様である。
【００５８】
雑音レベル補正用係数生成部206では（数８）に相当する重み係数
【００５９】
【数１１】

$$wo(f)=(1-\alpha s)$$

【００６０】
と求められる。加算部207では 、
【００６１】
【数１２】

$$wno(f)=wn(f)+wo(f)$$

【００６２】
が行われ、音声・雑音判定部208の結果に基づき ws(f) か wno(f) を切替部209で選択して最終的な重み係数 ww(f)を出力する。乗算部211では入力信号のスペクトルＸ（ｆ）にこの重み係数 ww(f)をかけ、出力信号 S(f)を
【００６３】
【数１３】

$$S(f)=ww(f)X(f)$$

【００６４】

と求める。

本実施形態は第１の実施形態の表現を伝達関数が乗算される形式に替えただけのものであるが、|X(f)|の平滑化を行うことで、（数９）（数１０）で求める重み係数の局所的な変動を押さえ重み係数の変化を滑らかにすることができ、音質向上につながる。

【００６５】

一方（数１３）のＸ（ｆ）は平滑化を行うと音がぼやけてしまうので、平滑化を行わない方が望ましい。（数９）（数１０）のＸ（ｆ）の平滑化方法として、例えば（数６）の方法を用いることができる。平滑化に関して本実施例と同等のことを実施例１でも行うことが可能であるが、本実施例の方がより簡単に行える利点がある。

【００６６】

また、第１の実施形態と同様に音声区間の抑圧量αｓが１に近い場合、付加する雑音のゲイン（１－αｓ）の値は小さな値となる。このような場合は音声区間と雑音区間の雑音レベルの差が知覚されにくいため、付加そのものを行わなくてもよい。また、分散の大きな雑音の場合は、この方法でもレベル差が完全に補償できないことがあり、その場合は分散を考慮した補償方法を用いることも可能である。

【実施例３】

【００６７】

図４は本発明の第３の実施形態に係る雑音抑圧装置の構成を表すブロック図である。第２の実施形態の音声・雑音判定部208 が入力信号ｘ（ｔ）に基づき判定を行っているのに対し、本実施例の音声・雑音判定部308は推定雑音 |N(f)|と入力信号|X(f)|に基づき判定を行っている。推定雑音と入力信号との比ＳＮＲは、

【００６８】

【数１４】

$$SNR=\frac{\sum_{f=0}^{M-1}|X(f)|^2}{\sum_{f=0}^{M-1}|N(f)|^2}$$

【００６９】

となる。本実施形態ではこの値を重み係数の切替えに用いている。ＳＮＲは全帯域でなく、音声のパワーが集中している帯域のみで算出するようにしてもよい。

【実施例４】

【００７０】

図５に本発明の第４の実施形態に係る雑音抑圧装置の構成を表すブロック図をしめす。第１の実施形態の雑音レベル補正用信号生成部106が入力信号から補正用の信号を生成しているのに対し、本実施形態の雑音レベル補正用信号生成部406は予め保持している重畳用信号450から生成している。雑音区間を入力信号とは無関係に白色雑音や聴覚的に聞えの良い雑音にしたい場合など効果的である。

【実施例５】

【００７１】

図６は本発明の第５の実施形態に係る雑音抑圧装置の構成を表すブロック図である。本実施形態は第２の実施形態に対して、N個の入力端子501-1～501-Nと、これを周波数領域に変換する周波数変換部502とその出力を統合して１つの信号を出力する統合信号生成部512と、 N個の入力信号から音声・雑音判定508を行う音声・雑音判定部508を備えている点が異なる。

【００７２】

マイクロホンアレーなど複数のマイクを用いて特定の方向の音だけを強調する方法がある。この場合は入力信号が音声か雑音かという問題は特定の方向から到来している信号か

否かという問題に置き換えることができる。音声・雑音判定部では複数の入力信号から信号の到来方向をもとに音声か雑音かの判定を下す。例えば、図７のようにマイク２本で正面から到来する信号を音声信号と見なす場合は、受音信号をX_0（ｆ）、X_1（ｆ）とし

【００７３】

て、

【００７４】

【数１５】

$$Ph = (1/M)\sum_{f=0}^{M-1}\left|\arg(X_0(f)X_1^*(f)\right|$$

【００７５】

を指標として音声区間を検出することができる。

ここで、X_1^*（ｆ）はX_1（ｆ）の共役複素数でａｒｇは位相を取出す演算子、Ｍは周波数の成分数である。正面からの信号は２つのマイクに同じ位相で到来するため、片方を共役複素数にして互いに掛け合わせると、位相項はゼロになる。従って、（数１５）は理想的に正面から到達した信号に関しては、最小値Ｐｈ＝０となる。それ以外の方向に関して正面からずれるに従い値が増加するので、適当な閾値をもとに音声／雑音の区別を行うことができる。なお、マイクの本数が２本以上の場合は、例えば全てのマイクの組み合わせに対して（数１５）を計算するなどの方法がある。

【００７６】

信号統合部512では、複数の入力信号から一つの信号を生成する。例えば遅延和アレーと呼ばれる方法では、入力信号の加算を行う。具体的には、統合された信号Ｘ（ｆ）は入力信号X_1（ｆ）～X_N（ｆ）を用いて、

【００７７】

【数１６】

$$X(f) = 1/N\sum_{i=0}^{N-1}X_i(f)$$

【００７８】

と表される。ここでＮはマイクの本数である。

このようにすることで正面から入力された目的信号は同位相であるため強調され、その他の方向から入力された信号は位相がずれているため弱め合い、その結果目的信号が強調され雑音が抑圧されるので、後段のスペクトルサブトラクションの雑音抑圧効果との相乗効果で、１つのマイクの場合に比べてより高い雑音抑圧性能を実現することができる。

【００７９】

また複数のマイクを使って音声区間の検出を行うので、１つのマイクの場合よりも高い検出能力が実現可能である。例えば、横方向からの妨害音が存在する場合、１つのマイクではこれを音声と区別することは困難であるが、複数マイクであれば（数１５）のように、位相成分を利用して音声信号（正面からの信号）と区別することができる。

【００８０】

なお、周波数変換部502の後に統合信号生成部512が構成されているが、周波数変換部502と統合信号生成部512は逆順であってもよい。

【実施例６】

【００８１】

図８は本発明の第６の実施形態に係る雑音抑圧装置の構成を表すブロック図である。第６の実施形態は第５の実施形態における統合信号生成部612 が目的信号強調部630と目的信号除去部631から構成されている。目的信号強調部630は第５の実施形態と同様に予め設定された目的音方向からの信号（例えば正面）の信号を強調するのに対し、目的信号除去

部631は目的信号強調部630 の目的音方向とは異なる方向（例えば横方向）を目的音方向とする。その結果、目的信号除去部631では正面から到来する音声信号は弱められ、周囲の音が強調されることになる。このように特定の方向に指向性を形成するユニットはビームフォーマと呼ばれることがある。第５の実施形態で説明した遅延和アレーもビームフォーマの１つである。

【００８２】

本実施形態においては目的信号強調部630と 目的信号除去部631を適応形アレーの代表であるGriffith-Jim形のビームフォーマを用いて実現する構成について説明する。

【００８３】

図９にGriffith-Jim形のビームフォーマの一構成例を示す。ビームフォーマの出力Ｘ（ｆ）は、入力信号X_0（ｆ）、X_1（ｆ）と適応フィルタを用いて求められる。入力端子901,902にX_0（ｆ）、X_1（ｆ）がそれぞれ入力される。整相化部903では目的音方向の信号の位相が同位相になるように位相が調整される。その出力は加算部904で加算され、減算部905で減算される。この減算により目的音が消去されるため、残りの信号を適応フィルタ906の入力として加算器904の出力から引き去ることで雑音が除去された信号Ｘ（ｆ）が得られる。

【００８４】

Griffith-Jim形のビームフォーマは妨害音の方向に感度が急峻に落ちる谷状のノッチを作ることが可能であり、この特性は特に目的音信号除去部631 が正面からの音声を妨害音とみなして除去するのに適した性質である。

【００８５】

さらに、目的音信号除去部631の出力信号は雑音推定部603の入力信号としても用いる。雑音推定部603は自力でＸ（ｆ）を観測し音声のない区間を見つけこれを平滑化して推定雑音を生成したが、目的信号除去部631の出力は常に雑音のみであるため、雑音の推定に利用される。このことから、これら２つの信号を利用することでより高精度な雑音推定が可能となる。

【実施例７】

【００８６】

図１０は本発明の第７の実施形態に係る雑音抑圧装置の構成を表すブロック図である。本実施形態は第５の実施形態における統合信号生成部512 の出力Ｘ（ｆ）を帯域分割部740でサブバンドに周波数分割し、各サブバンド毎に雑音抑圧を行う。雑音抑圧はこれまでの実施形態と同様であるが、音声・雑音判定部708 は各サブバンド毎に判定を行う。

【００８７】

音声のスペクトルを周波数方向に眺めると、振幅の出ている区間とそうでない区間が混在しており、山と谷がある。谷の部分の周波数に関しては、雑音区間と考えることができ、雑音レベルの推定や、過剰抑圧といった雑音区間で行う処理が用いられる。サブバンドに分割し各サブバンド毎の雑音／音声の判定に基づいて雑音抑圧を切替えることで、音声区間の品質をより高めることができる。

【００８８】

本実施例では複数の入力信号から統合信号を生成後にサブバンドに分割しているが、入力信号を先にサブバンドに分割した後にサブバンド単位で統合信号を求める構成であってもよい。

【図面の簡単な説明】

【００８９】

【図１】第１の実施形態に係る雑音抑圧装置の構成を表すブロック図。

【図２】入力信号のフレーム毎の振幅を模式した図。

【図３】第２の実施形態に係る雑音抑圧装置の構成を表すブロック図。

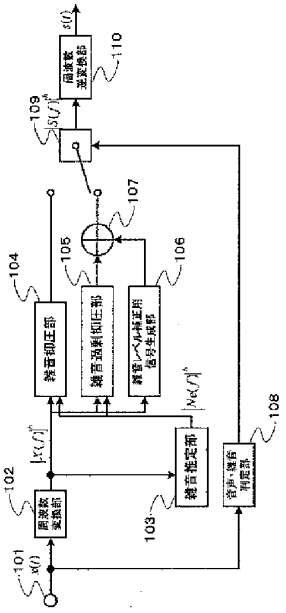
【図４】第３の実施形態に係る雑音抑圧装置の構成を表すブロック図。

【図５】第４の実施形態に係る雑音抑圧装置の構成を表すブロック図。

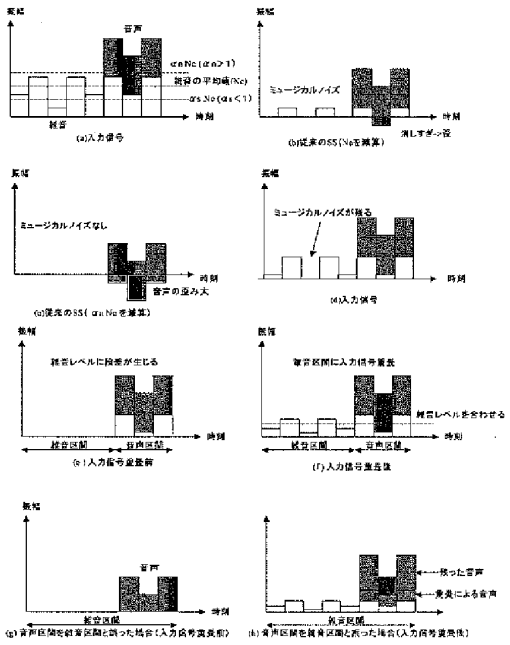
【図６】第５の実施形態に係る雑音抑圧装置の構成を表すブロック図。

【図７】マイクロホンアレイの機能を示す図。
【図８】第６の実施形態に係る雑音抑圧装置の構成を表すブロック図。
【図９】Griffith-Jim形のビームフォーマの一構成例を表すブロック図。
【図１０】第７の実施形態に係る雑音抑圧装置の構成を表すブロック図。
【符号の説明】
【００９０】
101,201,301,401,501-1…501-N, 601-1…601-N, 701-1…701-N,901,902 入力端子
102,202,302,402,502,602,702 周波数変換部
103,203,303,403,503,603 雑音推定部
104,404 雑音抑圧部
204,304,504,604 抑圧係数算出部
105,405 雑音過剰抑圧部
205,305,505,605 過剰抑圧係数算出部
106,406 雑音レベル補正用信号生成部
206,306,506,606 補正レベル補正用係数生成部
107,207,307,407,507,607,707,904,905,907 加算部
108,208,308,408,508,608,708 音声・雑音判定部
109,209,309,409,509,609 切替部
110,210,310,410,510,610,710 周波数逆変換部
211,311,511,611 乗算器
512,612,712 総合信号生成部
450 重畳用信号保存部
630 目的信号強調部
631 目的信号除去部
740 帯域分割部
750 サブバンド雑音抑圧部
760 帯域統合部
903 整相化部
906 適応フィルタ

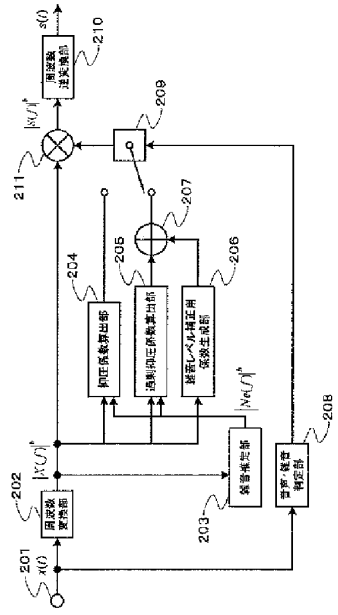
【図１】


周波数逆変換部
雑音初期化部
雑音適応制御部
雑音レベル補正用信号生成部
雑音推定部
音声・雑音判定部
周波数変換部

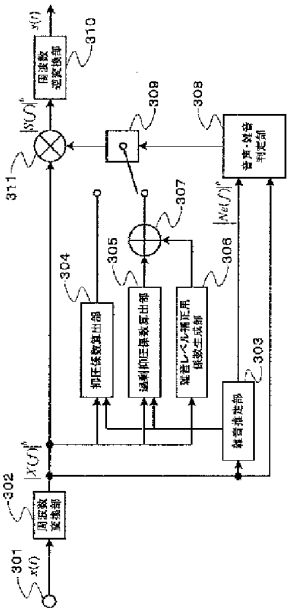
【図２】


音声
雑音
ミュージカルノイズ
ミュージカルノイズなし
ミュージカルノイズが残る
雑音レベルに段差が生じる
雑音区間に入力信号重畳
雑音レベルを合わせる
(a)入力信号
(b)従来のSS(Ne)を減算
(c)従来のSS(αsNe)を減算
(d)入力信号
(e)入力信号重畳前
(f)入力信号重畳後
抜けた音声
発声による音声

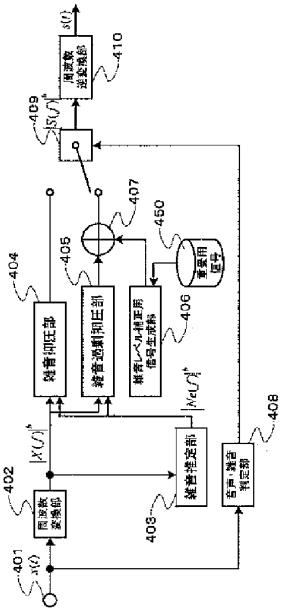
【図３】


周波数逆変換部
雑音適応制御部
過剰抑圧係数算出部
雑音レベル補正用係数生成部
雑音推定部
音声・雑音判定部
周波数変換部

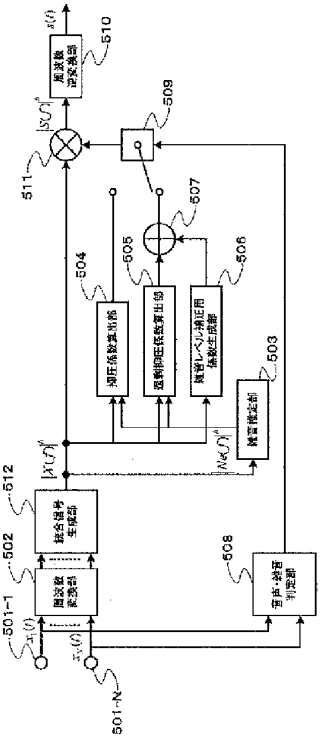
【図４】


周波数逆変換部
雑音適応制御部
過剰抑圧係数算出部
雑音レベル補正用係数生成部
雑音推定部
音声・雑音判定部
周波数変換部

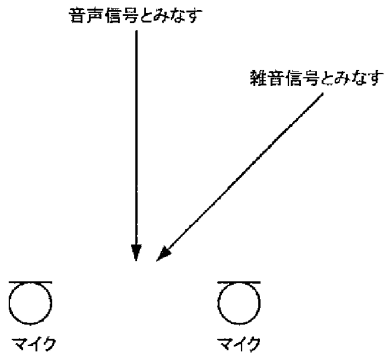
【図５】


雑音抑圧部
雑音適応制御検出部
雑音レベル補正用信号生成部
雑音推定部
音声・雑音判定部
周波数変換部
周波数逆変換部
雑音用信号

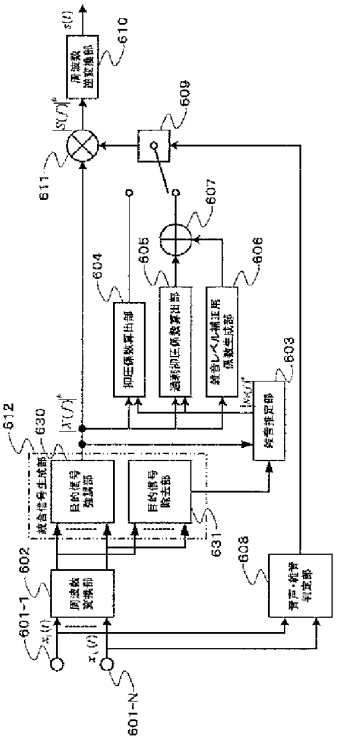
【図６】


周波数逆変換部
抑圧係数算出部
適応制御係数算出部
雑音レベル補正用係数生成部
雑音推定部
総合信号生成部
周波数変換部
音声・雑音判定部

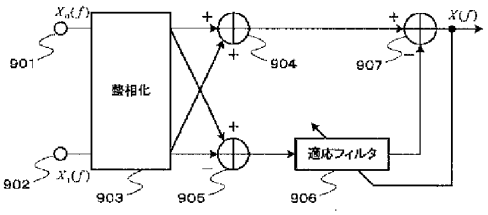
【図７】


音声信号とみなす
雑音信号とみなす
マイク
マイク

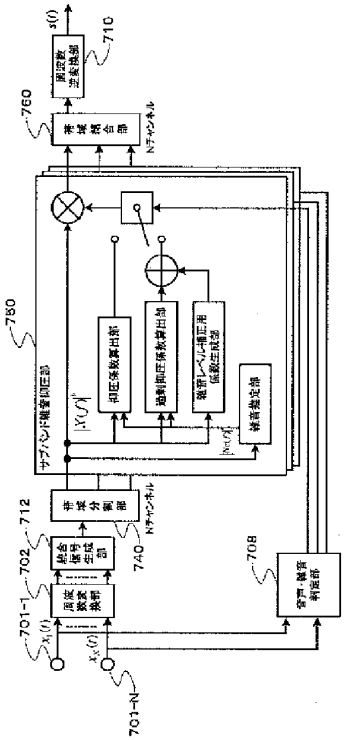
【図８】


周波数逆変換部
抑圧係数算出部
適応制御係数算出部
雑音レベル補正用係数生成部
雑音推定部
目的信号強調部
目的信号除去部
周波数変換部
音声・雑音判定部

【図９】


$X_0(f)$
901
902
$X_1(f)$
整相化
903
904
905
適応フィルタ
906
907
$X(f)$

【図１０】


710
760
750
712
702
701-1
701-N
740
708

フロントページの続き

(51)Int.Cl.[7]		FI			テーマコード（参考）
G10L	21/02	G10L	3/02	301F	
H04M	1/60	G10L	3/02	301G	
H04R	3/00	G10L	7/04	D	
		G10L	9/00	F	
		G10L	3/00	511	

Fターム(参考) 5D015 AA01 BB02 CC03 DD02 DD03 EE05 EE06
5D020 BB07
5K027 BB03 DD12 HH03

AUDIEN_FOR0035

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **2005-110127**

(43)Date of publication of application : **21.04.2005**

(51)Int.Cl. **H04R 3/00**
H04N 5/225
H04R 1/40
H04R 3/04
H04R 5/027

(21)Application number : **2003-343573**

(22)Date of filing : **01.10.2003**

(71)Applicant : **CANON INC**

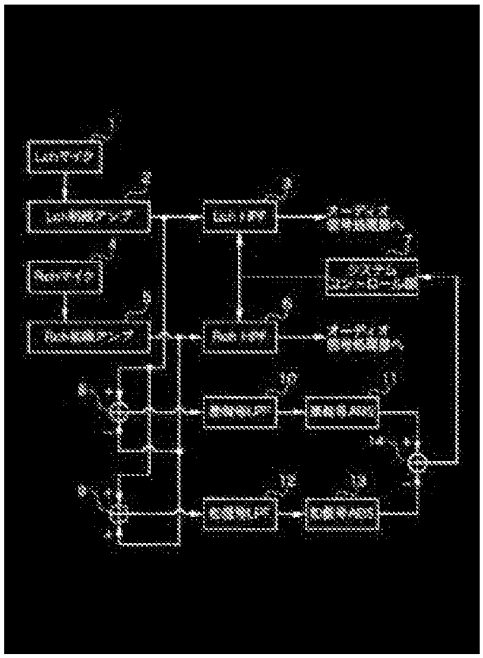
(72)Inventor : **TAKAHASHI KATSUTOSHI**

(54) WIND NOISE DETECTING DEVICE AND VIDEO CAMERA WITH WIND NOISE DETECTING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a wind noise detecting signal having high accuracy in detection by using not only a difference signal of two or more sound collecting means but a summation signal thereof as an information source for detecting the wind noise, and reducing incorrect determination even when the sound pressure level of the acoustic wave is high.

SOLUTION: A difference signal (L-R) is generated by a difference signal generator 8 from signals (L), (R) detected by two microphones, and a summation signal (L+R) is generated by a summation signal generator 9. Each signal is passed through low-pass filters 10, 12 and changed into each modulus by modulus detectors 11, 13. In a wind noise detecting signal generator 14, the modulus of the summation signal (L+R) is subtracted from the modulus of the difference signal (L-R) and the result is generated to a system control unit 7. In this case, 0 is generated if the resulted value is smaller than 0.



(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2005-110127
(P2005-110127A)
(43)公開日　平成17年4月21日(2005.4.21)

(51)Int.Cl.[7]	F I			テーマコード（参考）
H04R 3/00	H04R	3/00	320	5C022
H04N 5/225	H04N	5/225	F	5D011
H04R 1/40	H04R	1/40	320Z	5D020
H04R 3/04	H04R	3/04		
H04R 5/027	H04R	5/027	Z	

審査請求 未請求　請求項の数 5　ＯＬ　（全 9 頁）

(21)出願番号　特願2003-343573 (P2003-343573)
(22)出願日　平成15年10月1日 (2003.10.1)

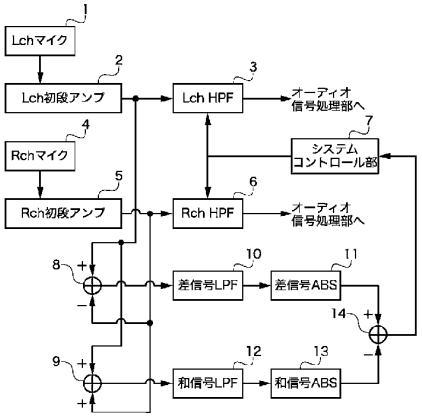
(71)出願人　000001007
キヤノン株式会社
東京都大田区下丸子３丁目３０番２号
(74)代理人　100076428
弁理士　大塚　康徳
(74)代理人　100112508
弁理士　高柳　司郎
(74)代理人　100115071
弁理士　大塚　康弘
(74)代理人　100116894
弁理士　木村　秀二
(72)発明者　高橋　克寿
東京都大田区下丸子３丁目３０番２号　キヤノン株式会社内

Ｆターム(参考)　5C022 AC69　AC72

最終頁に続く

(54)【発明の名称】風雑音検出装置及びそれを有するビデオカメラ装置

(57)【要約】
【課題】　風雑音を検出する情報源として、複数の集音手段の差信号だけではなく、それらの和信号も用いることにより、音波の音圧レベルが高い場合でも、誤判断を少なくし、検出精度の高い風雑音検出信号を得ることを可能にする。
【解決手段】　２つのマイクで検出された信号（Ｌ）（Ｒ）は、差信号生成器８で差信号（Ｌ－Ｒ）、和信号生成器９で和信号（Ｌ＋Ｒ）を生成する。それぞれは低域フィルタ１０、１２を経て、絶対値検出器１１、１３で絶対値に変換される。そして、風雑音検出信号生成器１４は、差信号（Ｌ－Ｒ）の絶対値から和信号（Ｌ＋Ｒ）の絶対値を減じ、その結果をシステムコントロール部７に出力する。ただし、この時の値が０より小さい場合には０を出力する。
【選択図】　図３



【特許請求の範囲】
【請求項１】
複数の集音手段を備え、当該複数の集音手段による信号で風雑音を検出する風雑音検出装置であって、
前記複数の集音手段で得られた信号から差信号を生成する差信号生成手段と、
前記複数の集音手段で得られた信号から和信号を生成する和信号生成手段と、
前記差信号生成手段で生成された差信号を絶対値信号に変換する第１の絶対値変換手段と、
前記和信号生成手段で生成された和信号を絶対値信号に変換する第２の絶対値変換手段と、
前記第１、第２の絶対値変換手段で変換して得た差信号と和信号との差を求め、風雑音検出信号として出力する風雑音検出手段と
を備えることを特徴とする風雑音検出装置。
【請求項２】
前記差信号生成手段及び和信号生成手段それぞれで生成された差信号及び和信号の低周波成分を抽出するフィルタ手段を更に備え、
前記第１の絶対値変換手段、及び、第２の絶対値変換手段それぞれは、低周波成分抽出された差信号、和信号それぞれを絶対値信号に変換することを特徴とする請求項１に記載の風雑音検出装置。
【請求項３】
前記風雑音検出手段からの信号を用いて、音声信号の低周波領域のレベルを連続的に制御する低周波領域連続制御手段を更に備えることを特徴とする請求項１又は２に記載の風雑音検出装置。
【請求項４】
前記風雑音検出手段からの信号のレベルに従い、風雑音のレベルを報知する報知手段を更に備えることを特徴とする請求項１乃至３のいずれか１項に記載の風雑音検出装置。
【請求項５】
請求項１乃至４のいずれか１項に記載の風雑音検出装置を備えることを特徴とするビデオカメラ装置。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は風雑音を除去するための風雑音検出の技術に関するものである。
【背景技術】
【０００２】
ビデオカメラ装置において、マイクが有する振動膜（振動板）が風圧によって振動してしまい、風雑音信号として検出することが知られている。この課題に対し、風雑音を改善する技術が、いくつか既に提案されている（例えば特許文献１）。
【０００３】
以下、図１にしたがって、この風雑音低減技術を説明する。
【０００４】
図１において、１０はＬｃｈ音声信号を集音するＬｃｈマイク、２０はＲｃｈ音声信号を集音するＲｃｈマイク、３０はＬｃｈマイク出力信号を増幅するＬｃｈ初段マイクアンプ、４０はＲｃｈマイク出力信号を増幅するＲｃｈ初段マイクアンプ、５０はＬｃｈ信号を遅延させるＬｃｈマトリックス遅延器、６０はＲｃｈ信号を遅延させるＲｃｈマトリックス遅延器、７０はＬｃｈ信号を減衰させるＬｃｈマトリックス減衰器、８０はＲｃｈ信号を減衰させるＲｃｈマトリックス減衰器である。
【０００５】
また、９０はＲｃｈマトリックス遅延器６０とＲｃｈマトリックス減衰器８０を経由したＲｃｈ信号を、Ｌｃｈ初段マイクアンプの出力信号から減算するＬｃｈマトリックス減

算器、１００はＬｃｈ遅延器５０とＬｃｈ減衰器７０を経由したＬｃｈ信号を、Ｒｃｈ初段マイクアンプの出力信号から減算するＲｃｈ減算器、１１０はＬｃｈ減算器出力の周波数特性を補正するＬｃｈ等化器、１２０はＲｃｈ減算器出力の周波数特性を補正するＲｃｈ等化器である。

【０００６】

１３０は信号処理後のＬｃｈ信号を出力するＬｃｈ出力部、１４０は信号処理後のＲｃｈ信号を出力するＲｃｈ出力部、１５０はＬｃｈ等化器出力の高域信号を抜き出すＬｃｈＨＰＦ、１６０はＲｃｈ等化器出力の高域信号を抜き出すＲｃｈＨＰＦ、１７０はＬｃｈ等化器１１０の出力を減衰させるＬｃｈ等化器出力減衰器、１８０はＬｃｈＨＰＦ出力を減衰させるＬｃｈＨＰＦ出力減衰器、１９０はＲｃｈ等化器１２０の出力を減衰させるＲｃｈ等化器出力減衰器、２００はＲｃｈＨＰＦ出力を減衰させるＲｃｈＨＰＦ出力減衰器、２１０はＬｃｈ等化器出力減衰器の出力とＬｃｈＨＰＦ出力減衰器の出力を加算するＬｃｈ加算器、２２０はＲｃｈ等化器出力減衰器の出力とＲｃｈＨＰＦ出力減衰器の出力を加算するＲｃｈ加算器、２３０はＬｃｈマイク出力とＲｃｈマイク出力から差信号（Ｌ－Ｒ）を生成する差信号生成器、２４０は差信号生成器２３０の出力の低域信号を抜き出すＬＰＦ、２５０はＬＰＦ出力のピーク値を検出するところのピーク検出部である。

【０００７】

尚、これまで説明してきた各部の中で、便箋上風雑音低減に関係しない部分は割愛して以下動作説明する。

【０００８】

近傍に配置される二つのマイク１０及び２０から出力される音声信号は、音波が空気中の波動であるため特に波長が長くなる低周波領域で相関性が強くなる。云い換えれば二つのマイクが出力する信号はほぼ相似なもの言える。

【０００９】

これに対し風雑音は、風がマイク振動膜にランダムに当たることにより発生するもので、２つのマイク出力の相関性が極めて薄く且つ低周波領域でレベルが高いという特性を持つ。

【００１０】

この特性を利用し、２つのマイク出力を差信号生成器２３０に入力し差信号（Ｌ－Ｒ）を生成する。このレベルは同相信号が主成分である音波の場合差信号（Ｌ－Ｒ）は低く、相関性の薄い風雑音の場合は逆にレベルは高くなる。その後、風雑音の主帯域である低周波領域をＬＰＦ２４０で抜き出し、その値をピーク検出部２５０で検出することで、風雑音検出信号を作り出している。風雑音検出信号のレベルにより、Ｌｃｈ等化器出力減衰器１７０（Ｒｃｈ等化器出力減衰器１９０）とＬｃｈＨＰＦ出力減衰器１８０（ＲｃｈＨＰＦ出力減衰器２００）の減衰量を制御しそれぞれの出力信号をＬｃｈ加算器２１０（Ｒｃｈ加算器２２０）で加算することで風雑音の主帯域である低周波領域のレベルを下げ、風雑音を低減している。

【特許文献１】特開平５－７３９２号公報（第６頁、図１）

【発明の開示】

【発明が解決しようとする課題】

【００１１】

しかしながら、上記例には以下の不具合がある。

【００１２】

二つのマイクの出力信号は音波の場合、相関性が強く非常に近似したレベル及び位相関係を持つが、完全に同じではないため、差信号生成器からはあるレベルの差信号（Ｌ－Ｒ）が出力されてしまうことがある。特に音源がマイク正面ではなく横方向に大きくずれた場合で、尚且つ、マイクに印可される音波の音圧レベルが高くなってくると差信号（Ｌ－Ｒ）のレベルも高くなり、風雑音と誤検出することで音波を集音しているのにも関わらず、音声信号の低周波領域のレベルを減衰させてしまう場合があった。

【００１３】

そこで、本発明は、マイクに印可される音波の音圧レベルが非常に高い場合でも風雑音と誤検出するのを防ぎ、高精度な風雑音の検出信号を生成する技術を提供しようとするものである。

【課題を解決するための手段】

【００１４】

かかる課題を解決するため、例えば、本発明の風雑音検出装置は以下の構成を備える。すなわち、

複数の集音手段を備え、当該複数の集音手段による信号で風雑音を検出する風雑音検出装置であって、

前記複数の集音手段で得られた信号から差信号を生成する差信号生成手段と、

前記複数の集音手段で得られた信号から和信号を生成する和信号生成手段と、

前記差信号生成手段で生成された差信号を絶対値信号に変換する第１の絶対値変換手段と、

前記和信号生成手段で生成された和信号を絶対値信号に変換する第２の絶対値変換手段と、

前記第１、第２の絶対値変換手段で変換して得た差信号と和信号との差を求め、風雑音検出信号として出力する風雑音検出手段とを備える。

【発明の効果】

【００１５】

本発明によれば、風雑音を検出する情報源として、複数の集音手段の差信号だけではなく、それらの和信号も用いることにより、音波の音圧レベルが高い場合でも、誤判断を少なくし、検出精度の高い風雑音検出信号を得ることを可能にする。

【発明を実施するための最良の形態】

【００１６】

以下、添付図面にしたがって本発明に係る実施形態を詳細に説明する。

【００１７】

図２は実施形態におけるビデオカメラ全体のブロック構成図を示している。

【００１８】

同図において、１０１は光学画像を取り込むところのレンズ部、１０２はレンズ部１を介して入射した光を結像し画像の電気信号に変換するＣＣＤ部、１０３はＣＣＤ部２からの電気信号を処理するカメラ信号処理部、１０４は音を集音するマイク部、１０５は音声信号を処理するオーディオ信号処理部、１０６はビデオ信号を処理するビデオ信号処理部、１０７はビデオカメラ装置全体の制御を行うシステムコントロール部、１０８はビデオカメラ装置で信号の入出力を行う入出力部、１０９は記録媒体への記録及び再生処理を行う記録再生処理部、１１０は記録媒体にデータを記録或いは記録媒体からデータを再生するメカ部、１１１は表示データを処理する表示信号処理部、１１２は表示信号処理部からのデータを表示する表示部（例えば液晶表示部）である。

【００１９】

図３は、図２におけるマイク部１０４、オーディオ信号処理部１０５及びシステムコントロール部１０７周辺の詳細ブロック図である。

【００２０】

同図において、１はＬｃｈ音声信号を集音するＬｃｈマイク、２はＬｃｈマイク１からの出力信号を増幅するＬｃｈ初段アンプ、３は風雑音検出結果からＬｃｈ音声信号の低周波領域のレベルを制御するＬｃｈＨＰＦ、４はＲｃｈ音声信号を集音するＲｃｈマイク、５はＲｃｈマイク４からの出力信号を増幅するＲｃｈ初段アンプ、６は風雑音検出結果からＲｃｈ音声信号の低周波領域のレベルを制御するＲｃｈＨＰＦ、７は風雑音検出信号生成器１４の出力レベルから、ＬｃｈＨＰＦ３とＲｃｈＨＰＦ６のカットオフ周波数を制御するシステムコントロール部、８はＬｃｈ初段アンプ２とＲｃｈ初段アンプ５の各出力から差信号（Ｌ－Ｒ）を生成する差信号生成器、９はＬｃｈ初段アンプ２とＲｃｈ初段アンプ５の各出力から和信号（Ｌ＋Ｒ）を生成する和信号生成器、１０は差信号中の風の主成

分である低域を抜き出す差信号ＬＰＦ、１１は差信号ＬＰＦ出力の絶対値をとる差信号絶対値検出器、１２は和信号中の風の主成分である低域を抜き出すところの和信号ＬＰＦ、１３は和信号ＬＰＦ出力の絶対値をとる和信号絶対値検出器、１４は風雑音検出信号を生成する風雑音検出信号生成器である。

【００２１】

以下図３について各ブロックの動作について説明する。

【００２２】

Ｌｃｈマイク１とＲｃｈマイク４は互いに近接して配置されている。音波の場合、特に低周波領域においてＬｃｈマイク１の出力とＲｃｈマイク４の出力は、ほぼ同相の相似な信号となる。これに対し、風の場合、マイク振動板への当たる風圧変化はランダムな特性を有するため、Ｌｃｈマイク１の出力とＲｃｈマイク４の出力との間の相関性は小さい、すなわち、位相となって現れる。

【００２３】

Ｌｃｈ初段アンプ２とＲｃｈ初段アンプ５の出力には前記Ｌｃｈマイク１の出力とＲｃｈマイク４の出力が増幅されて出力される。

【００２４】

ここで、出力された信号をそれぞれ（Ｌ）と（Ｒ）と表現することとする。

【００２５】

信号（Ｌ）と（Ｒ）はそれぞれ次段のＬｃｈＨＰＦ３、ＲｃｈＨＰＦ６に入力される。ＬｃｈＨＰＦ３、ＲｃｈＨＰＦ６はシステムコントロール部７からの制御信号により、連続的にカットオフ周波数を振れるような構成となっている。

【００２６】

信号（Ｌ）と（Ｒ）は、差信号生成器８と和信号生成器９に入力される。ここで、信号（Ｌ）と（Ｒ）は差信号生成器８で信号（Ｌ－Ｒ）に、和信号生成器９で信号（Ｌ+Ｒ）に変換される。

【００２７】

生成された信号（Ｌ－Ｒ）は、差信号ＬＰＦ１０で風雑音の主帯域である低域側を抜き出し、その後差信号絶対遅検出器１１で絶対値を算出する。同様に作られた信号（Ｌ+Ｒ）は和信号ＬＰＦ１２で風雑音の主帯域である低域側を抜き出し、その後、和信号絶対遅検出器１３で絶対値を算出する。算出された検出値はそれぞれ風雑音検出信号生成器１４に入力され、信号（Ｌ－Ｒ）の絶対値から、信号（Ｌ+Ｒ）の絶対値が減算された結果が出力される。この風雑音検出信号生成器１４は、０でリミッタが掛かる減算器であり最低値は０となる。すなわち、『絶対値（Ｌ－Ｒ）－絶対値（Ｌ＋Ｒ）＜０』の場合には０を出力する。

【００２８】

かかる構成にした理由を簡単に説明する。

【００２９】

今、マイクに印可されるのが音波の場合で、尚且つそれほど印可レベルも高くない場合を想定する。この場合の各マイクロホンでの検出信号（Ｌ）と（Ｒ）はほぼ相似の信号のため信号（Ｌ－Ｒ）のレベルはあまり大きな値とはならない。したがって、正の値を持つ絶対値（Ｌ＋Ｒ）を、もともと小さな値の絶対値（Ｌ－Ｒ）から減じるわけであるから、風雑音検出信号生成器１４の出力信号は、０もしくは小さな値を出力することが約束される。

【００３０】

ところが音波であっても、例えば音圧レベルが比較的高い場合、より具体的には、音源がビデオカメラの近くに存在し、尚且つ、その音源がビデオカメラの撮影方向に対してずれた場所にあるとき（以下、状況Ｘという）、信号（Ｌ）と（Ｒ）はほぼ同位相であっても、それらの差分信号（Ｌ－Ｒ）（の絶対値）はある程度以上の差が生じ易くなる。したがって、絶対値信号（Ｌ－Ｒ）（の絶対値）のみで、「風雑音」か否かを判断してしまうと、上記のような状況Ｘであっても、「風雑音」と誤判定する、もしくは誤判定しやすく

なる。

【００３１】

しかし、本実施形態の場合、和信号（Ｌ＋Ｒ）も、風雑音か否かの判断基準にしていることに注目されたい。上記状況Ｘの場合、信号（Ｌ）と（Ｒ）とは、ほぼ同位相であるため、和信号（Ｌ＋Ｒ）は差信号（Ｌ－Ｒ）より大きな値となる。つまり、音声の場合、それが如何なる状況にあっても、（Ｌ－Ｒ）＜（Ｌ＋Ｒ）の関係が実施的に約束されることになり、結果的に、風雑音検出信号生成器１４の出力信号は０、もしくは小さな値となる。

【００３２】

一方、風雑音の場合であるが、先に説明したように、２つのマイクロホンから得られる信号（Ｌ）と（Ｒ）（の低周波成分）との間には、相関関係が小さい。したがって、ある任意の期間の平均レベルを比較した場合、信号（Ｌ－Ｒ）の絶対値が、信号（Ｌ＋Ｒ）の絶対値よりもむしろ大きなものとなりやすい。すなわち、風雑音検出信号生成器１４の出力信号は０よりも大きなものなり、問題のない範囲で、風雑音検出を行うことができるようになる。

【００３３】

なお、２つのマイクロホンが、たまたま同時に同位相の風圧変化を受け、尚且つ、その振幅のみが異なる信号を出力した場合、上記の処理によれば、風雑音ではないと判断される可能性がある。しかしながら、そのような状況は、ビデオカメラで撮影している場合における時間軸の局所的なものであり、全体として見た場合には、上記和信号（Ｌ＋Ｒ）までをも参照する構成にする作用効果の方が顕著である。

【００３４】

以上のように、差信号（Ｌ－Ｒ）だけではなく和信号（Ｌ+Ｒ）も検出部のパラメータとして使用することで、風雑音検出信号の精度を大きく向上させることが出来る。

【００３５】

このような過程で生成された風雑音検出信号は、システムコントロール部７でその値を読み込まれ、音声信号のメインルートに配置されたＬｃｈＨＰＦ３とＲｃｈＨＰＦ６のカットオフ周波数を連続的に制御することで、特に低域を主成分とする風雑音を、最適のフィルタ設定で低減することが可能となる。

【００３６】

なお、上記実施形態では、風雑音検出信号生成器１４は、『絶対値（Ｌ－Ｒ）－絶対値（Ｌ＋Ｒ）』を算出するものとして説明したが、絶対値（Ｌ－Ｒ）或いは絶対値（Ｌ＋Ｒ）の一方に係数αを乗算してから減算し、この係数αを適宜不図示のスイッチで変更できるようにし、風雑音か否かの判定基準を調整できるようにしても良い。

【００３７】

また、実施形態では、集音手段としてビデオカメラが通常有するステレオマイクロホンを利用するものとして説明したが、これに限定されるものではない。例えば複数のマイクロホンを有する録音装置に適用しても良いからである。

【００３８】

＜第２の実施形態＞

次に、図４を用いて第２の実施形態を説明する。図４は上記実施形態における図３に相当するものであるが、異なる点は、風雑音低域フィルタ切り換えスイッチＳＷ１８を設けた点と、システムコントロール部７が、表示信号処理部１１を制御して表示部１２の風雑音レベルの表示する点にある。

【００３９】

なお、本第２の実施形態では、ＬｃｈＨＰＦ３とＲｃｈＨＰＦ６は連続的にカットオフ周波数は変化せず、風雑音低減フィルタ切換スイッチ１８によるＯＮ／ＯＦＦの切換タイプとなっている。

【００４０】

図４において、風雑音検出信号生成器１４の出力レベルに応じて、システムコントロー

ル部７は表示信号制御部１５へ風雑音レベルに応じた表示データを送る。そのデータを受けた表示信号制御部１５は表示部１６上に風雑音レベルを示すバーを図示の符号１７に示すように表示する（バーの本数が多いほど、風雑音が大きいことを示す）。撮影者はこのレベル表示を見て必要と判断すると風雑音低減フィルタ切換スイッチ１８を切換て風雑音を低減させる。

【００４１】

撮影者によっては、風雑音のレベルにより自動的に低域のレベルが低減するのを好まない場合があるため、本対応によって風雑音レベル表示値を指標として、撮影者が自分の意志で低域の成分を撮影状態に応じて低減させることが出来るという作用効果を奏する。

【図面の簡単な説明】

【００４２】

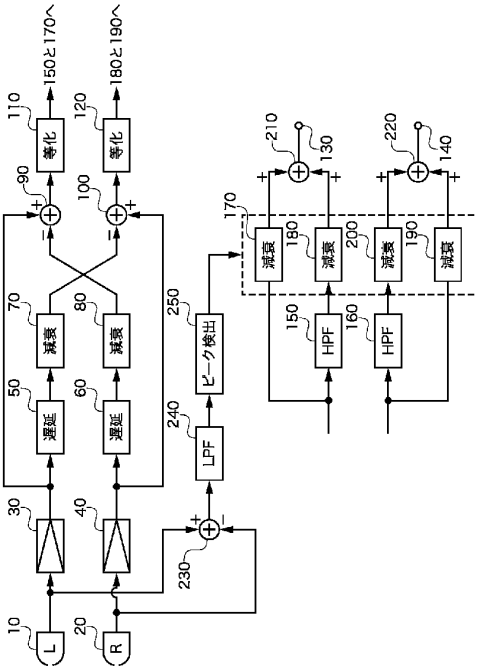
【図１】従来の風雑音検出の構成を示す図である。

【図２】実施形態が適用するビデオカメラ装置全体の構成を示す図である。

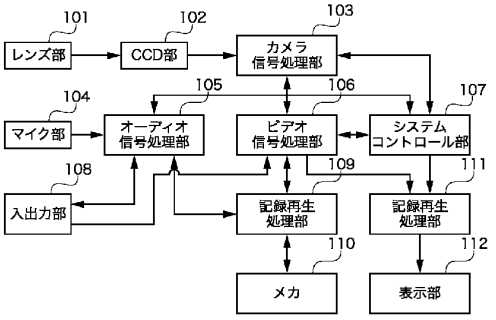
【図３】第１の実施形態が適用する風雑音検出の構成を示す図である。

【図４】第２の実施形態が適用する風雑音検出の構成を示す図である。

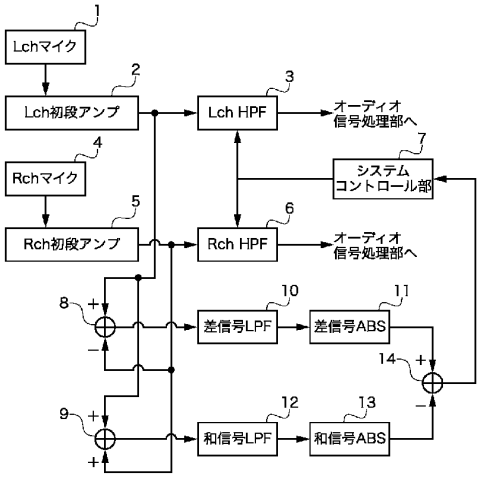
【図１】



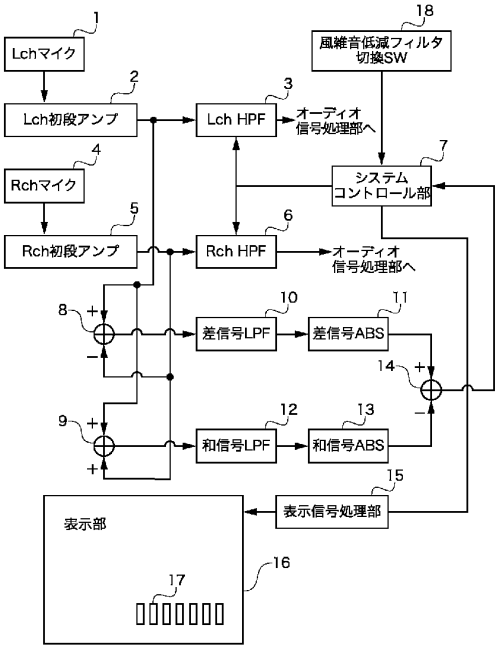
【図２】



【図３】


1
Lchマイク
2
Lch初段アンプ
3
Lch HPF
オーディオ
信号処理部へ
4
Rchマイク
7
システム
コントロール部
5
Rch初段アンプ
6
Rch HPF
オーディオ
信号処理部へ
8
10
差信号LPF
11
差信号ABS
14
9
12
和信号LPF
13
和信号ABS

【図４】


1
Lchマイク
18
風雑音低減フィルタ
切換SW
2
Lch初段アンプ
3
Lch HPF
オーディオ
信号処理部へ
4
Rchマイク
7
システム
コントロール部
5
Rch初段アンプ
6
Rch HPF
オーディオ
信号処理部へ
8
10
差信号LPF
11
差信号ABS
14
9
12
和信号LPF
13
和信号ABS
15
表示信号処理部
表示部
16
17

フロントページの続き

Fターム(参考) 5D011 AB00
5D020 BB07 CE03

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2004 年1 月29 日 (29.01.2004)

PCT

(10) 国際公開番号
WO 2004/010415 A1

(51) 国際特許分類[7]: G10L 19/02

(21) 国際出願番号: PCT/JP2003/007962

(22) 国際出願日: 2003 年6 月24 日 (24.06.2003)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願2002-210945 2002 年7 月19 日 (19.07.2002) JP
特願2002-273010 2002 年9 月19 日 (19.09.2002) JP

(71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒108-8001 東京都 港区 芝五丁目７番１号 Tokyo (JP). 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府 門真市 大字門真１００６番地 Osaka (JP).

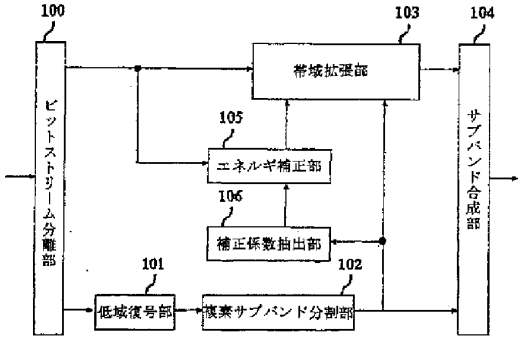
(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 野村 俊之 (NOMURA,Toshiyuki) [JP/JP]; 〒108-8001 東京都 港区 芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 嶋田 修 (SHIMADA,Osamu) [JP/JP]; 〒108-8001 東京都 港区 芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 高見沢 雄一郎 (TAKAMIZAWA,Yuichiro) [JP/JP]; 〒108-8001 東京都 港区 芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 芹沢 昌宏 (SERIZAWA,Masahiro) [JP/JP]; 〒108-8001 東京都 港区 芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 田中 直也 (TANAKA,Naoya) [JP/JP]; 〒572-0055 大阪府

[続葉有]

(54) Title: AUDIO DECODING DEVICE, DECODING METHOD, AND PROGRAM

(54) 発明の名称: オーディオ復号装置と復号方法およびプログラム



100...BIT STREAM SEPARATION UNIT
103...BAND EXTENSION UNIT
105...ENERGY CORRECTION UNIT
106...CORRECTION COEFFICIENT EXTRACTION UNIT
101...LOW BAND DECODING UNIT
102...COMPLEX SUB-BAND DIVISION UNIT
104...SUB-BAND SYNTHESIS UNIT

(57) **Abstract:** An audio decoding device includes an energy correction unit (105) for correcting a target energy of a higher band component and a correction coefficient calculation unit (106) for calculating an energy correction coefficient from a low sub-band signal. These processing units execute the correction processing of the target energy required when performing the band extension processing using only real numbers. Thus, instead of a complex sub-band synthesis filter and a complex band extension unit, it is possible to use a real number sub-band synthesis filter and a band extension unit requiring less calculation amount while maintaining a high sound quality. Thus, it is possible to reduce the necessary calculation amount and the device size.

(57) 要約: 高域成分の目標エネルギを補正するエネルギ補正部（１０５）と、低域サブバンド信号からエネルギ補正係数を算出する補正係数算出部（１０６）を新たに設ける。帯域拡張処理を実数のみで実施する際に必要となる目標

[続葉有]

WO 2004/010415 A1

WO 2004/010415 A1

寝屋川市 御幸東町３０－２３ Osaka (JP). 津島 峰生 (TSUSHIMA,Mineo) [JP/JP]; 〒576-0021 大阪府 交野市 妙見坂５－５－３０２ Osaka (JP). 則松 武志 (NORIMATSU,Takeshi) [JP/JP]; 〒651-1305 兵庫県 神戸市北区 西山２－１２－２７ Hyogo (JP). チョン コク セン (CHONG,Kok Seng) [MY/SG]; 689527 シンガポール チョア チュ カン ノース ７、＃０３－０６、レジェンド グローブ、５０ Singapore (SG). クア キム ハン (KUAH,Kim Hann) [SG/SG]; 760145 シンガポール ＃０４－４３、イシュン ストリート １１、ブロック １４５ Singapore (SG). ネオ スア ホン (NEO,Sua Hong) [SG/SG]; 530959 シンガポール ＃０６－２８４、ホーガン ストリート ９１、ブロック ９５９ Singapore (SG).

(74) 代理人: 宮崎 昭夫, 外(MIYAZAKI,Teruo et al.); 〒107-0052 東京都 港区 赤坂１丁目９番２０号 第１６興和ビル８階 Tokyo (JP).

(81) 指定国(国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) 指定国(広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

２文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

エネルギの補正処理を、これらの処理部が実行する。これにより、高音質を維持しつつ、複素サブバンド合成フィルタおよび複素帯域拡張部の代わりに、より演算量が少ない実数のサブバンド合成フィルタおよび帯域拡張部を用いることができ、必要演算量・装置規模を低減できる。

明細書

オーディオ復号装置と復号方法およびプログラム

技術分野

本発明は、符号化されたオーディオ信号を復号するためのオーディオ復号装置および復号方法に関する。

背景技術

低ビットレートで高音質に符号化可能なオーディオ符号化・復号方式として、ＩＳＯ／ＩＥＣの国際標準方式であるＭＰＥＧ－２　ＡＡＣ（Advanced Audio Coding）が広く知られている。このＭＰＥＧ－２　ＡＡＣに代表される従来のオーディオ符号化復号方式では、まず、時間領域のＰＣＭ信号を複数サンプルまとめて１フレームとし、これをＭＤＣＴ（Modified Discrete Cosine Transform）などの写像変換により周波数領域信号へと変換する。そして、この周波数領域信号を量子化およびハフマン符号化してビットストリームを形成する。量子化においては人間の聴覚特性を考慮して、周波数領域信号の中で知覚されやすい周波数成分については量子化精度を高く、そうでない周波数成分については量子化精度を低くすることにより、限られた符号量で高い音質を実現する。例えば、ＭＰＥＧ－２　ＡＡＣでは９６ｋｂｐｓ程度のビットレートでＣＤと同等音質（４４．１ｋＨｚサンプリング、ステレオ）を実現できる。

これより低いビットレート、例えば４８ｋｂｐｓ程度のビットレートで４４．１ｋＨｚサンプリングのステレオオーディオ信号を符号化しようとする場合、聴覚的に重要度が低い高域周波数成分を符号化対象から外す（量子化値を零とする）ことで限られたビットレートでの主観的音質が最高となるようにする。しかしながらこの場合、高域周波数成分が符号化されていないために音質が劣化し、一般的にはこもった感じの音となってしまう。

このような低ビットレートでの音質劣化を解決する技術として帯域拡張技術が注目を集めている。帯域拡張技術では、ＭＰＥＧ－２　ＡＡＣなどの方式で

低ビットレートに符号化した低域ビットストリームに、僅かな符号量（一般に数ｋｂｐｓ程度）の補助情報である高域ビットストリームを付加してビットストリームを構成する。オーディオ復号装置（デコーダ）では、まず、ＭＰＥＧ－２　ＡＡＣなどの復号方法で低域ビットストリームを復号して、高域周波数成分が含まれていない低域オーディオ信号を得る。そして、帯域拡張技術により、高域ビットストリームが示す補助情報に従った加工を低域オーディオ信号に施して高域周波数成分を生成する。こうして得られた高域周波数成分と、低域ビットストリームを復号して得られた低域オーディオ信号を合成して高域周波数成分を含む復号オーディオ信号を得る。

帯域拡張技術を用いた従来のオーディオ復号装置の一例として、ＭＰＥＧ－２　ＡＡＣ復号装置にＳＢＲと呼ばれる帯域拡張技術を組み合わせた例が、下記の文献１の５．６節に記述されている。ここに記述されているような帯域拡張技術を用いた従来のオーディオ復号装置を図１に示す。

文献１：「Ｄｉｇｉｔａｌ　Ｒａｄｉｏ　Ｍｏｎｄｉａｌｅ（ＤＲＭ）；Ｓｙｓｔｅｍ　Ｓｐｅｃｉｆｉｃａｔｉｏｎ」（ＥＴＳＩ　ＴＳ　１０１　９８０　Ｖ１．１．１）、２００１年９月発行、ｐ．４２－５７

図１に示す従来のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、複素サブバンド分割部４０２、複素帯域拡張部４０３、複素サブバンド合成部４０４から構成される。

ビットストリーム分離部１００は、入力されたビットストリームを分離して低域復号部１０１と複素帯域拡張部４０３へ出力する。ビットストリームには低域信号がＭＰＥＧ－２　ＡＡＣなどの符号化方式で符号化されている低域ビットストリームと、複素帯域拡張部が高域信号を生成するために必要な情報を含む高域ビットストリームが多重化されている。低域ビットストリームは低域復号部１０１へ、高域ビットストリームは複素帯域拡張部４０３へ出力される。

低域復号部１０１は、入力された低域ビットストリームを復号して低域オー

ディオ信号を生成し、複素サブバンド分割部４０２へ出力する。低域復号部１０１はＭＰＥＧ－２　ＡＡＣなど既存のオーディオ復号方式を用いて低域ビットストリームを復号する。

複素サブバンド分割部４０２は、入力された低域オーディオ信号を複素サブバンド分割フィルタによって帯域分割する。帯域分割された低域サブバンド信号は帯域拡張部４０３および複素サブバンド合成部４０４へ出力される。複素サブバンド分割フィルタについては、例えば、従来から広く知られている３２バンド複素ＱＭＦ（Quadrature Mirror Filter：直交鏡像フィルタ）フィルタバンクなどを用いることができる。この場合、３２サブバンドに帯域分割された複素数の低域信号（低域サブバンド信号）が複素帯域拡張部４０３および複素サブバンド合成部４０４へ出力される。３２バンド複素分析ＱＭＦフィルタバンクの計算式は以下のようになる。

$$X_k(m) = \sum_{n=-\infty}^{\infty} h(mM - n)\, x(n) W_{K1}^{-(k+k_0)(n+n_0)}, \quad k = 0,1,\ldots,K1-1 \qquad 402.1$$

$$W_{K1} = e^{j\frac{2\pi}{K1}} \quad \cdots \quad 402.2$$

ここで、x(n)は低域オーディオ信号を表し、Xk(m)はkバンド目の低域サブバンド信号を表し、h(n)は分析用低域通過フィルタを表す。本例ではK1＝64である。

複素帯域拡張部４０３は、入力された高域ビットストリームと低域サブバンド信号から高域のオーディオ信号を表す高域サブバンド信号を生成してサブバンド合成部４０４へ出力する。複素帯域拡張部４０３は、図２に示すように複素高域生成部５００と複素振幅調整部５０１から構成される。複素帯域拡張部４０３は入力端子５０２を介して高域ビットストリームを入力し、入力端子５０４を介して低域サブバンド信号を入力し、出力端子５０３を介して高域サブバンド信号を出力する。

複素高域生成部５００は、低域サブバンド信号と高域ビットストリームを入力とし、低域サブバンド信号の中から高域ビットストリームが指定する帯域の信号を高域サブバンドへと複写する。また、この複写の際には高域ビットストリームが指定する信号加工を施す場合もある。例えば、周波数が低い順にサブバンド０からサブバンド６３まで６４個のサブバンドを仮定し、低域サブバンド信号としてサブバンド０からサブバンド１９までの複素数のサブバンド信号が入力端子５０４を介して供給されているとする。また高域ビットストリームには、サブバンドＡ（Ａ＞１９）をどの低域サブバンド（サブバンド０から１９）から信号を複写して生成するかを表す複写情報、および、複写の際に施す信号加工方法（フィルタリングなどあらかじめ定められた数種類の手段から選択して用いられる）を示す信号加工情報が含まれているとする。複素高域生成部５００では、高域のサブバンドの複素信号（複写加工サブバンド信号と名付ける）を複写情報が示す低域のサブバンドの複素信号と同一とする。さらに、音質向上のために信号加工情報が何らかの信号加工の必要性を表している場合は、複写加工サブバンド信号に信号加工情報が示す信号加工を施す。このようにして生成された複写加工サブバンド信号は複素振幅調整部５０１に出力される。

複素高域生成部５００における信号加工の一例として、音声符号化で一般に良く知られる線形予測逆フィルタを用いることができる。一般に、線形予測逆フィルタのフィルタ係数は入力信号を線形予測することにより算出でき、当該フィルタ係数を用いた前記線形予測逆フィルタは入力信号のスペクトル特性を白色化するように動作することが知られている。このような線形予測逆フィルタを信号加工として用いる理由は、高域サブバンド信号のスペクトル特性を、複写元の低域サブバンド信号のスペクトル特性よりも平坦化するためである。例えば、音声信号などの場合、低域サブバンド信号と高域サブバンド信号のスペクトル特性を比較すると、高域サブバンド信号のスペクトル特性は低域サブバンド信号のスペクトル特性よりも平坦であることが多いため、このような平

坦化処理を備えることにより、高品質な帯域拡張技術が実現できる。

複素振幅調整部５０１は、入力された複写加工サブバンド信号の振幅に対して、高域ビットストリームが指定する補正を施して高域サブバンド信号を生成する。つまり、符号化側における入力信号の高域成分の信号エネルギ（目標エネルギと名付ける）と、帯域拡張部４０３が生成する信号の高域信号エネルギが等価となるような振幅補正を複写加工サブバンド信号に対して施す。そのために、高域ビットストリームには目標エネルギを表す情報が含まれている。生成された高域サブバンド信号は出力端子５０３へ出力される。なお、高域ビットストリームに記載される目標エネルギは、例えば、サブバンド毎にフレーム単位で算出されていると考えることができる。また、入力信号の時間及び周波数方向の特性を考慮して、時間方向にはフレームをさらに分割した時間単位で、周波数方向には複数のサブバンドをまとめたバンド単位で算出しても良い。時間方向のフレームをさらに分割した時間単位で算出する場合には、エネルギの時間変化をより詳細に表すことができる。複数のサブバンドをまとめたバンド単位で算出する場合には、目標エネルギを符号化するのに必要なビット数を削減することができる。ここで、目標エネルギの算出で利用する時間及び周波数方向の分割単位は時間周波数グリッドとし、その情報は高域ビットストリームに記載される。

また、複素振幅調整部５０１の他の実施例として、複写加工サブバンド信号に加えて付加信号を用いて高域サブバンド信号を生成することができる。この場合、高域サブバンド信号のエネルギが目標エネルギとなるように、複写加工サブバンド信号の振幅と付加信号の振幅とが調整される。付加信号の例としては、雑音信号あるいはトーン信号が挙げられる。複写加工サブバンド信号と付加信号の振幅調整用のゲインは、例えば、複写加工サブバンド信号と付加信号とのうちどちらか一方を、生成後の高域サブバンド信号の主成分とし、他の一方を副成分とした上で以下のように算出される。複写加工サブバンド信号を主成分とする場合は、

Gmain=sqrt（R／E／（1＋Q））
Gsub=sqrt（R×Q／N／（1＋Q））

ここで、Gmainは主成分の振幅調整用ゲインであり、Gsubは副成分の振幅調整用ゲインであり、EとNは、それぞれ、複写加工サブバンド信号と付加信号のエネルギを表す。なお、付加信号のエネルギが１に正規化されている場合はN＝１である。また、Rは目標エネルギを表し、Qは主成分と副成分のエネルギ比を表し、RとQは高域ビットストリームに記載されているものとする。なお、sqrt () は平方根である。一方、付加信号を主成分とする場合は、

Gmain=sqrt（R／N（1＋Q））
Gsub=sqrt（R×Q／E／（1＋Q））

となる。以上のように算出した振幅調整用ゲインを用いて、複写加工サブバンド信号と付加信号を重み付け加算することにより高域サブバンド信号は算出できる。

ここで、複素振幅調整部５０１における振幅調整の動作および効果を、図３を用いて詳細に説明する。符号化側における入力信号の高域成分の信号位相（図中、位相A）と、前記低域サブバンド信号から導出した高域サブバンド信号の信号位相（図中、位相B）とは図３に示すように全く異なるが、前記高域サブバンド信号は、その信号エネルギが目標エネルギと同一となるように振幅調整されているため、聴感的な音質劣化を抑えることができる。これは人の聴覚が信号の位相変動よりもエネルギ変動により敏感であるためである。

複素サブバンド合成部４０４は、入力された低域サブバンド信号と高域サブバンド信号を複素サブバンド合成フィルタによって帯域合成する。この帯域合成により生成されたオーディオ信号がオーディオ復号装置の出力となる。複素

サブバンド合成フィルタは、複素サブバンド分割部４０２で用いられている複素サブバンドフィルタに対応したものが用いられる。つまり、ある信号を複素サブバンド分割フィルタで帯域分割し、これを複素サブバンド合成フィルタで帯域合成すると元の信号（複素サブバンド分割フィルタの入力信号）が完全に再構成できるように選ばれる。例えば、複素サブバンド合成フィルタとして式４０２．１に示したような３２バンド複素ＱＭＦ分割フィルタバンク（K1＝６４）が用いられている場合、下記の式４０４．１を用いることができる。

$$x(n) = \sum_{m=-\infty}^{\infty} f(n - mM) \frac{1}{K2} \sum_{k=0}^{K2-1} X_k(m) W_{K2}^{(k+k_0)(n+n_0)} \quad \cdots \quad 404.1$$

ここで、ｆ（ｎ）は合成用低域通過フィルタを表す。本例ではK2＝64である。

また、帯域拡張技術によって低域復号部１０１が出力するオーディオ信号のサンプリング周波数よりも複素サブバンド合成部４０４が出力するオーディオ信号のサンプリング周波数が高い場合は、複素サブバンド合成部４０４が出力するオーディオ信号の低域部（ダウンサンプリングした結果）が低域復号部１０１が出力するオーディオ信号と等しくなるように選ばれる。複素サブバンド合成部４０４では６４バンド複素ＱＭＦ合成フィルタバンク（式４０４．１においてK2＝１２８）を用いることができる。なお、この場合、低域側３２バンドには３２バンド複素分析ＱＭＦフィルタバンクの出力が信号値として用いられる。

上述した従来のオーディオ復号装置では、演算量が大きい複素サブバンド分割部および複素サブバンド合成部を有する上、帯域拡張処理を複素数上で行うため必要演算量・装置規模が大きいという問題があった。

発明の開示

本発明の目的は、高音質を維持しつつ演算量を低減した帯域拡張技術、およびこれを用いたオーディオ復号装置、オーディオ復号方法およびオーディオ復号用プログラムを提供することにある。

上記目的を達成するために、本発明のオーディオ復号装置は、ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、

前記低域サブバンド信号に基づきエネルギ補正係数を算出する補正係数抽出部と、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部とから構成されている。

また、本発明の他のオーディオ復号装置では、前記補正係数抽出部において、前記低域サブバンド信号の信号位相を算出し、前記信号位相によりエネルギ補正係数を算出するようにしてもよい。また、前記補正係数抽出部において、前記低域サブバンド信号の実部のエネルギと前記低域サブバンド信号の信号エネルギとの比をエネルギ補正係数として算出するようにしてもよい。さらに、前記補正係数抽出部において、前記低域サブバンド信号の各サンプルの位相を平均化してエネルギ補正係数を算出するようにしてもよい。さらに、前記補正

係数抽出部において、前記周波数帯域毎に算出したエネルギ補正係数を平滑化するようにしてもよい。

また、本発明の他のオーディオ復号装置では、ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

予め定めたエネルギ補正係数を発生する補正係数発生部と、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部とから構成されている。

また、本発明の他のオーディオ復号装置では、前記補正係数発生部において、乱数を発生させ当該乱数をエネルギ補正係数とするようにしてもよい。さらに、前記補正係数発生部において、周波数帯域ごとに予め定めたエネルギ補正係数を発生させるようにしてもよい。

本発明によるオーディオ復号装置は、高域成分の目標エネルギを補正するエネルギ補正部と、低域サブバンド信号からエネルギ補正係数を算出する補正係数算出部あるいは予め定めた方法でエネルギ補正係数を発生させる補正係数発生部を設けたことを特徴としている。帯域拡張処理を実数のみで実施する際に必要となる目標エネルギの補正処理を、これらの処理部が実行する。これにより、高音質を維持しつつ、複素サブバンド合成フィルタおよび複素帯域拡張部

の代わりに、より演算量が少ない実数のサブバンド合成フィルタおよび帯域拡張部を用いることができ、必要演算量・装置規模を低減できるという効果が得られる。さらに、低域サブバンド信号を用いずにエネルギ補正係数を発生させる補正係数発生部を用いる場合には、サブバンド合成フィルタおよび帯域拡張部に加えて、演算量が少ない実数のサブバンド分割フィルタを用いることができ、より必要演算量・装置規模を低減できるという効果が得られる。

図面の簡単な説明

図１は、従来のオーディオ復号装置の構成を示すブロック図である。

図２は、従来技術における複素帯域拡張部４０３の構成を示すブロック図である。

図３は、従来技術における振幅調整の動作を説明するための図である。

図４は、本発明における振幅調整の動作を説明するための図である。

図５は、エネルギ補正をしない場合の振幅調整の動作を説明するための図である。

図６は、本発明の第１の実施形態のオーディオ復号装置の構成を示すブロック図である。

図７は、本発明の第２の実施形態のオーディオ復号装置の構成を示すブロック図である。

図８は、本発明における帯域拡張部１０３の構成を示すブロック図である。

発明を実施するための最良な形態

次に、本発明の実施の形態について図面を参照して詳細に説明する。

（第１の実施形態）

図６は本発明の第１の実施形態のオーディオ復号装置の構成を示すブロック図である。本実施形態のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、複素サブバンド分割部１０２、帯域拡張部１０３、サ

ブバンド合成部１０４、エネルギ補正部１０５、補正係数抽出部１０６とから構成される。

ビットストリーム分離部１００は、入力されたビットストリームを分離して低域復号部１０１と帯域拡張部１０３とエネルギ補正部１０５へ出力する。ビットストリームには低域信号が符号化されている低域ビットストリームと、帯域拡張部１０３が高域信号を生成するために必要な情報を含む高域ビットストリームが多重化されている。低域ビットストリームは低域復号部１０１へ、高域ビットストリームは帯域拡張部１０３とエネルギ補正部１０５へ出力される。

低域復号部１０１は、入力された低域ビットストリームを復号して低域オーディオ信号を生成し、複素サブバンド分割部１０２へ出力する。低域復号部１０１はＭＰＥＧ－２　ＡＡＣなど既存のオーディオ復号方式を用いて低域ビットストリームを復号する。

複素サブバンド分割部１０２は、入力された低域オーディオ信号を複素サブバンド分割フィルタによって帯域分割する。帯域分割された低域サブバンド信号は帯域拡張部１０３とサブバンド合成部１０４と補正係数抽出部１０６へ出力される。

補正係数抽出部１０６は、低域サブバンド信号から後述する手法でエネルギ補正係数を算出し、エネルギ補正部１０５へ出力する。

エネルギ補正部１０５は、高域ビットストリームに記載される高域成分の目標エネルギをエネルギ補正係数により補正した補正目標エネルギを算出し、帯域拡張部１０３へ出力する。

帯域拡張部１０３は、入力された高域ビットストリームと低域サブバンド信号と補正目標エネルギから高域のオーディオ信号を表す高域サブバンド信号を生成してサブバンド合成部１０４へ出力する。

サブバンド合成部１０４は、入力された低域サブバンド信号と高域サブバンド信号をサブバンド合成フィルタによって帯域合成する。帯域合成により生成

されたオーディオ信号がオーディオ復号装置の出力となる。

このように構成される本発明のオーディオ復号装置と図１の従来のオーディオ復号装置の相違は、図１の複素サブバンド分割部４０２の替わりに図６では複素サブバンド分割部１０２があり、図１の複素サブバンド合成部４０４の替わりに図６ではサブバンド合成部１０４があり、図１の複素帯域拡張部４０３の替わりに図６では帯域拡張部１０３があり、また、本実施形態（図６）には補正係数抽出部１０６とエネルギ補正部１０５が追加されていることにある。他の処理部については、従来方式と同じであり当業者にとってよく知られており、また本発明とは直接関係しないので、その詳細な説明は省略する。以下では従来と手法が異なる複素サブバンド分割部１０２、帯域拡張部１０３、サブバンド合成部１０４、エネルギ補正部１０５、補正係数抽出部１０６について詳細に説明する。

まず、複素サブバンド分割部１０２とサブバンド合成部１０４について説明する。従来はサブバンド分割フィルタとして複素数のサブバンド信号を生成する式４０２．１のようなフィルタバンクを用いていた。また、これに対応する逆変換として式４０４．１のようなフィルタバンクをサブバンド合成フィルタとして用いていた。式４０４．１の出力、あるいは、式４０４．１の出力を式４０２．１の入力信号のサンプリング周波数にダウンサンプリングした信号は式４０２．１の入力と完全に一致する完全再構成性がある。高音質な復号オーディオ信号を得るためには、この完全再構成性の関係がサブバンド分割・合成フィルタに必要となる。

本実施形態においては、演算量を低減するために、従来の複素サブバンド合成部４０４で用いられている複素サブバンド合成フィルタを、実数のサブバンド合成フィルタに置き換えている。しかしながら、単にサブバンド合成フィルタを複素数から実数に変更してしまうと完全再構成性は失われ、音質が劣化する。

複素数のサブバンド分割フィルタと実数のサブバンド合成フィルタの間で完

全再構成性を実現するためには、従来の複素サブバンド分割フィルタの出力に回転演算を施すと良いことが従来から広く知られている。この回転演算は複素数の実数軸と虚数軸を（π÷４）だけ回転させるものであり、一般的に良く知られたＤＦＴからＤＣＴを導出する手法と同じである。例えば、ｋ0＝１／２の場合、式４０２．１の３２バンド複素ＱＭＦ分割フィルタバンクの計算式には、各サブバンドkに対して以下の回転演算（Ｋ＝Ｋ1）を施すと良い。

$$W_K^{-(k+k_0)\frac{3}{4}K} \quad \cdots\cdots \quad 102.1$$

なお、式１０２．１における３／４Ｋは１／４Ｋでも良い。

このように従来の複素サブバンド分割部４０２の後段に式１０２．１の回転演算を行う処理部を付加したものを複素サブバンド分割部１０２としても良いが、これらサブバンド分割フィルタと回転演算処理を合わせた処理と等価な演算をより少ない演算量で実現可能な下式を複素サブバンド分割部１０２で行うと良い。

$$X_k(m) = \sum_{n=-\infty}^{\infty} h(mM-n)x(n)W_{K1}^{-(k+k_0)(n+n_0+\frac{3}{4}K1)}, \quad k = 0,1,\ldots,K1-1$$

$$\cdots\cdot 102.2$$

これに対応する実数のサブバンド合成フィルタとしては、式４０４．１に、下記の式１０４．１に示す変換を施し、実数部分のみを取り出した下記の式１０４．２をサブバンド合成部１０４で用いれば、完全再構成性を実現できるようになる。

$$W_K^{(k+k_0)\frac{3}{4}K} \quad \cdots\cdots \quad 104.1$$

$$x(n) = \sum_{m=-\infty}^{\infty} f(n - mM) \frac{2}{K2} \sum_{k=0}^{K2-1} \mathrm{Re}[X_k(m)] \cos(\frac{2\pi}{K2}(n + n_0)(k + k_0 + \frac{3}{4}K2))$$

・・・104.2

ここで、Re[.]は複素サブバンド信号の実数部のみを取り出すことを表す。

次に帯域拡張部１０３について説明する。帯域拡張部１０３は、入力された高域ビットストリームと低域サブバンド信号と後述する補正目標エネルギから高域のオーディオ信号を表す高域サブバンド信号を生成してサブバンド合成部１０４へ出力する。帯域拡張部１０３は、図８に示すように高域生成部３００と振幅調整部３０１と変換部３０５から構成される。帯域拡張部１０３は入力端子３０２を介して高域ビットストリームを入力し、入力端子３０４を介して低域サブバンド信号を入力し、入力端子３０６を介して補正目標エネルギを入力し、出力端子３０３を介して高域サブバンド信号を出力する。

変換部３０５は、入力端子３０４を介して入力した複素数の低域サブバンド信号から実部のみを取り出し、実数の低域サブバンド信号（以下、断りがない限り低域サブバンド信号は実数とする）に変換し高域生成部３００へ出力する。

高域生成部３００は、低域サブバンド信号と高域ビットストリームを入力とし、低域サブバンド信号の中から高域ビットストリームが指定する帯域の信号を高域サブバンドへと複写する。また、この複写の際には高域ビットストリームが指定する信号加工を施す場合もある。例えば、周波数が低い順にサブバンド0からサブバンド６３まで６４個のサブバンドを仮定し、低域サブバンド信号としてサブバンド0からサブバンド１９までの実数のサブバンド信号が変換部３０５から供給されているとする。また高域ビットストリームには、サブバンドA（A＞１９）をどの低域サブバンド（サブバンド0から１９）から信号を複写して生成するかを表す複写情報、および、複写の際に施す信号加工方法（フィルタリングなどあらかじめ定められた数種類の手段から選択して用いられる）を示す信号加工情報が含まれているとする。高域生成部３００の処理は

、高域のサブバンドの実数信号（複写加工サブバンド信号と名付ける）を複写情報が示す低域のサブバンドの実数信号と同一とする。さらに、音質向上のために信号加工情報が何らかの信号加工の必要性を表している場合は、複写加工サブバンド信号に信号加工情報が示す信号加工を施す。このようにして生成された複写加工サブバンド信号は振幅調整部３０１に出力される。

高域生成部３００における信号加工の一例として、従来技術の複素高域生成部５００と同様に線形予測逆フィルタを用いることができる。当該フィルタの効果は複素高域生成部５００と同様であるため説明を省略する。なお、高域生成処理において線形予測逆フィルタを用いる場合、実数信号上で動作する高域生成部３００は、複素信号上で動作する複素高域生成部５００に比べて、当該フィルタ係数算出に必要な演算量を削減できるという効果がある。

振幅調整部３０１は、高域ビットストリームに従い、入力された複写加工サブバンド信号の振幅を補正目標エネルギと等価となるように補正し、高域サブバンド信号を生成する。生成された高域サブバンド信号は出力端子３０３へ出力される。なお、高域ビットストリームに記載される目標エネルギは、従来例と同様に、例えば、サブバンド毎にフレーム単位で算出されていると考えても良いし、入力信号の時間及び周波数方向の特性を考慮して、時間方向にはフレームをさらに分割した時間単位で、周波数方向には複数のサブバンドをまとめたバンド単位で算出されていると考えても良い。時間方向のフレームをさらに分割した時間単位で算出する場合には、エネルギの時間変化をより詳細に表すことができる。複数のサブバンドをまとめたバンド単位で算出する場合には、目標エネルギを符号化するのに必要なビット数を削減することができる。ここで、目標エネルギの算出で利用する時間及び周波数方向の分割単位は時間周波数グリッドとし、その情報は高域ビットストリームに記載される。

また、振幅調整部３０１の他の実施例として、従来例と同様に、複写加工サブバンド信号に加えて付加信号を用いて高域サブバンド信号を生成する場合においても、高域サブバンド信号のエネルギが補正目標エネルギとなるように、

複写加工サブバンド信号の振幅と付加信号の振幅とが調整される。付加信号の例としては、雑音信号あるいはトーン信号が挙げられる。複写加工サブバンド信号と付加信号の振幅調整用のゲインは、例えば、複写加工サブバンド信号と付加信号とのうちどちらか一方を、生成後の高域サブバンド信号の主成分とし、他の一方を副成分とした上で以下のように算出される。複写加工サブバンド信号を主成分とする場合は、

Gmain＝sqrt（a×R／Er／（1＋Q））
Gsub＝sqrt（a×R×Q／Nr／（1＋Q））

ここで、Gmainは主成分の振幅調整用ゲインであり、Gsubは副成分の振幅調整用ゲインであり、ErとNrは、それぞれ、複写加工サブバンド信号と付加信号のエネルギを表す。なお、これらのエネルギの表記ErとNrが従来例の説明時の表記E、Nと異なるのは、従来例においては複写加工サブバンド信号と付加信号が複素数信号であるのに対して、本発明では実数信号であることを区別するためである。付加信号のエネルギが1に正規化されている場合はNr＝1である。また、Rは目標エネルギを表し、aは後述する補正係数抽出部106で算出されるエネルギ補正係数を表す。したがって、a×Rは補正目標エネルギを表す。Qは主成分と副成分のエネルギ比を表し、RとQは高域ビットストリームに記載されているものとする。なお、sqrt（）は平方根である。一方、付加信号を主成分とする場合は、Gmain、Gsubは下記の式により示される。

Gmain＝sqrt（a×R／Nr／（1＋Q））
Gsub＝sqrt（a×R×Q／Er／（1＋Q））

また、付加信号を主成分とする場合には、複素数の低域サブバンド信号に基づいて算出されたエネルギ補正係数aの代わりに、エネルギ補正係数aと同様

の手法によって付加信号に基づいて算出されたエネルギ補正係数ｂを用いて、Gmain、Gsubを下記の式により示すようにしてもよい。

Gmain＝sqrt（ｂ×R／Nr／（1＋Q））
Gsub＝sqrt（ｂ×R×Q／Er／（1＋Q））

なお、付加信号として、記憶領域に予め格納されている信号を用いる場合には、エネルギ補正係数ｂについても予め算出しておき、定数とすることによって、エネルギ補正係数ｂの算出に関わる演算を省略することができる。以上のように算出した振幅調整用ゲインを用いて、複写加工サブバンド信号と付加信号を重み付け加算することにより高域サブバンド信号は算出できる。

ここで、振幅調整部３０１における振幅調整の動作および効果を、図４を用いて詳細に説明する。実数の高域サブバンド信号（図中、振幅調整後の高域成分の実部）は、その信号エネルギが、入力信号の高域成分の信号エネルギを表す目標エネルギを補正して得られた補正目標エネルギと同一となるように振幅調整される。ここで、図４に示すように補正目標エネルギが変換部３０５で変換される前の複素数の低域サブバンド信号の信号位相（図中、位相Ｂ）を考慮して算出されている場合、複素数の低域サブバンド信号から導出した仮想的な複素数の高域サブバンド信号を鑑みると、仮想的な複素数の高域サブバンド信号の信号エネルギは目標エネルギと等価になっている。なお、本例で利用する複素サブバンド分割部１０２とサブバンド合成部１０４とからなる分析合成系では、上述したようにサブバンド信号の実部のみを用いて、実部と虚部の両方を用いたときと同様に完全再構成が得られる。したがって、実数の高域サブバンド信号が補正目標エネルギと同一な信号エネルギになるように振幅調整されることにより、従来例と同様に、人の聴覚上重要なエネルギ変動は最小限に抑えられるため、聴感的な音質劣化を抑えることができる。一方、振幅調整処理において、補正目標エネルギではなく目標エネルギを用いて振幅調整した場合

の例を図５に示す。図５に示すように実数の高域サブバンド信号が目標エネルギと同一な信号エネルギになるように振幅調整された場合、前記仮想的な複素数の高域サブバンド信号の信号エネルギは目標エネルギよりも大きくなってしまう。この結果、サブバンド合成部１０４で帯域合成されたオーディオ信号の高域成分は、符号化側の入力信号の高域成分よりも大きくなり、音質劣化の原因となる。

以上で帯域拡張部１０３の説明を終える。帯域拡張部１０３の処理を低演算量な実部のみの処理で実現するとともに高音質な復号信号を得るためには、上述したように振幅調整において補正目標エネルギを用いる必要があり、本実施形態では補正係数抽出部１０６とエネルギ補正部１０５において補正目標エネルギを算出する。

補正係数抽出部１０６は、入力された複素数の低域サブバンド信号に基づいてエネルギ補正係数を算出し、エネルギ補正部１０５へ出力する。エネルギ補正係数の算出は、複素数の低域サブバンド信号の信号位相を算出し、これをエネルギ補正係数とすることにより実施できる。例えば、複素数の信号サンプルから成る低域サブバンド信号のエネルギと、その実部のみから算出したエネルギとを算出し、これらエネルギの比をエネルギ補正係数とすることができる。あるいは、低域サブバンド信号を構成する複素数の各信号サンプル値の位相を算出し、これを平均化してエネルギ補正係数を算出することもできる。さらに、上述した手法では、帯域分割された周波数帯域毎にエネルギ補正係数が算出されるが、隣接する帯域のエネルギ補正係数と当該帯域のエネルギ補正係数を平滑化した係数を、当該帯域のエネルギ補正係数とすることもできる。また、予め定めた時定数と前フレームのエネルギ補正係数とを用いて現フレームのエネルギ補正係数を時間方向に平滑化することもできる。このようなエネルギ補正係数の平滑化を行うことにより、エネルギ補正係数の急激な変化を抑制することができ、結果として帯域拡張後のオーディオ信号の品質を向上させることができる。

なお、上述した手法におけるエネルギの算出あるいは、信号サンプル値の位相の平均化は、従来の技術において説明した目標エネルギの時間周波数グリッドに合わせて、当該時間周波数グリッドに含まれる信号サンプルを用いて実施するようにしても良い。帯域拡張後のオーディオ信号の品質を向上させるためには、位相特性を正確に表すエネルギ補正係数を算出する必要があり、このためには、位相特性変化の少ない信号サンプルを用いてエネルギ補正係数を算出することが望ましい。一般的に、時間周波数グリッドは、当該グリッド内の信号の変化が小さくなるように設定される。したがって、時間周波数グリッドに合わせてエネルギ補正係数を算出することにより、位相特性を正確に表すエネルギ補正係数を算出することができ、結果として帯域拡張後のオーディオ信号の品質を向上させることができる。また、本手法は、時間方向あるいは周波数方向のいずれか一方のみについて信号の変化を考慮し、時間方向あるいは周波数方向のどちらか一方のグリッド境界のみによって区切られた範囲に含まれる信号サンプルを用いて実施することもできる。

エネルギ補正部１０５は、高域ビットストリームに記載される入力信号の高域成分の信号エネルギを表す目標エネルギを、補正係数抽出部１０６で算出したエネルギ補正係数により補正して補正目標エネルギを算出し、帯域拡張部１０３へ出力する。

（第２の実施形態）

次に、本発明の第２の実施の形態について図７を用いて詳細に説明する。

図７を参照すると、本発明の第２の実施の形態としてのオーディオ復号装置が示されている。本実施形態のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、サブバンド分割部２０２、帯域拡張部１０３、サブバンド合成部１０４、補正係数発生部２０６、エネルギ補正部１０５とから構成される。

本発明の第２の実施の形態は、本発明の第１の実施の形態と比較して、複素サブバンド分割部１０２がサブバンド分割部２０２に置き換わり、補正係数抽

出部１０６が補正係数発生部２０６に置き換わっているのみであり、他の部分については全く同一である。そこでサブバンド分割部２０２と補正係数発生部２０６について詳細に説明する。

サブバンド分割部２０２は、入力された低域オーディオ信号をサブバンド分割フィルタによって帯域分割する。帯域分割された実数の低域サブバンド信号は帯域拡張部１０３とサブバンド合成部１０４へ出力される。ここで、サブバンド分割部２０２で利用されるサブバンド分割フィルタは式１０２．２の実数の処理部のみから構成され、その出力信号は実数の低域サブバンド信号となる。したがって、帯域拡張部１０３に入力される低域サブバンド信号は実数であるため、変換部３０５においては入力した実数の低域サブバンド信号をそのまま高域生成部３００に出力する。

補正係数発生部２０６は、予め定めた手法でエネルギ補正係数を算出し、エネルギ補正部１０５へ出力する。エネルギ補正係数の算出方法としては、例えば、この補正係数発生部２０６において、乱数を発生させ当該乱数をエネルギ補正係数とすることができる。ここで、発生させる乱数は０から１に正規化される。第1の実施の形態において説明したように、実数の高域サブバンド信号が目標エネルギと同一な信号エネルギになるように振幅調整された場合、復号後のオーディオ信号の高域成分は目標エネルギよりも必ず大きくなってしまうが、０から１に正規化された乱数から導出したエネルギ補正係数を用いることにより、補正目標エネルギは目標エネルギよりも小さくなり得る。この結果、復号後のオーディオ信号の高域成分は目標エネルギよりも必ずしも大きくはならないため、音質改善効果が期待できる。また、周波数帯域ごとに予めエネルギ補正係数を定めておき、帯域拡張部１０３で実施される複写における複写元の周波数帯域および複写先の周波数帯域、あるいは、そのどちらか一方のみに応じて、エネルギ補正係数を発生させることもできる。この場合も予め定めたエネルギ補正係数は０から１の値とする。本手法を用いた場合、乱数を用いてエネルギ補正係数を算出する手法に比べて、人の聴覚特性などを利用でき、より

音質改善効果を得る事ができる。なお、上記２手法を組み合わせて、周波数帯域ごとに乱数の最大値を定めておき、その範囲内に正規化された乱数をエネルギ補正係数とすることもできる。あるいは、周波数帯域ごとに平均値を予め定めておき、その平均値を中心とした乱数を発生させ、エネルギ補正係数を算出することもできる。さらに、周波数帯域毎に算出されたエネルギ補正係数に対して、隣接する帯域のエネルギ補正係数により平滑化を施した係数を、当該帯域のエネルギ補正係数とすることもできる。また、予め定めた時定数と前フレームのエネルギ補正係数とを用いて現フレームのエネルギ補正係数を時間方向に平滑化することもできる。

本発明の第２の実施の形態は、本発明の第１の実施の形態と比較して、低域サブバンド信号の信号位相を考慮しないために復号オーディオ信号の品質は第１の実施の形態よりも低下するが、複素数の低域サブバンド信号を利用する必要がなくなり実数のサブバンド分割フィルタを用いることができるため、更なる低演算量化を実現できる。

なお、本発明が上記各実施形態に限定されず、本発明の技術思想の範囲内において、各実施形態は適宜変更され得ることは明らかである。

また、図には示されていないが、本実施形態のオーディオ復号装置は、上記で説明したオーディオ復号方法を実行するためのプログラムを記録した記録媒体を備えている。この記録媒体は磁気ディスク、半導体メモリまたはその他の記録媒体であってもよい。このプログラムは、記録媒体からオーディオ復号装置に読み込まれ、オーディオ復号装置の動作を制御する。具体的には、オーディオ復号装置内のＣＰＵがこのプログラムの制御によりオーディオ復号装置のハードウェア資源に特定の処理を行うように指示することにより上記の処理が実現される。

請求の範囲

1． ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、

前記低域サブバンド信号に基づきエネルギ補正係数を算出する補正係数抽出部と、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されるオーディオ復号装置。

2． 前記補正係数抽出部は、前記低域サブバンド信号の信号位相を算出し、前記信号位相に基づいてエネルギ補正係数を算出する請求項1記載のオーディオ復号装置。

3． 前記補正係数抽出部は、前記低域サブバンド信号の実部のエネルギと前記低域サブバンド信号の信号エネルギとの比をエネルギ補正係数として算出する請求項1記載のオーディオ復号装置。

4． 前記補正係数抽出部は、前記低域サブバンド信号の各サンプルの位相

を平均化してエネルギ補正係数を算出する請求項１記載のオーディオ復号装置。

５．　前記補正係数抽出部は、前記周波数帯域毎に算出したエネルギ補正係数を平滑化する請求項１から４のいずれか１項記載のオーディオ復号装置。

６．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

予め定めたエネルギ補正係数を発生する補正係数発生部と、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、

から構成されるオーディオ復号装置。

７．　前記補正係数発生部は、乱数を発生させ当該乱数をエネルギ補正係数とする請求項６記載のオーディオ復号装置。

８．　前記補正係数発生部は、周波数帯域ごとに予め定めたエネルギ補正係数を発生させる請求項６記載のオーディオ復号装置。

９．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップと、

前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成するステップと、

前記低域サブバンド信号に基づきエネルギ補正係数を算出するステップと、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するステップと、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

１０．　前記補正目標エネルギの算出において、前記低域サブバンド信号の信号位相を算出し、前記信号位相に基づいてエネルギ補正係数を算出する請求項９記載のオーディオ復号方法。

１１．　前記補正目標エネルギの算出において、前記低域サブバンド信号の実部のエネルギと前記低域サブバンド信号の信号エネルギとの比をエネルギ補正係数として算出する請求項９記載のオーディオ復号方法。

１２．　前記補正目標エネルギの算出において、前記低域サブバンド信号の各サンプルの位相を平均化してエネルギ補正係数を算出する請求項９記載のオ

ーディオ復号方法。

１３．　前記補正目標エネルギの算出において、前記周波数帯域毎に算出したエネルギ補正係数を平滑化する請求項９から１２のいずれか１項記載のオーディオ復号方法。

１４．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップ、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するステップ、

予め定めたエネルギ補正係数を発生するステップ、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するステップ、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

１５．　前記エネルギ補正係数の発生において、乱数を発生させ当該乱数をエネルギ補正係数とする請求項１４記載のオーディオ復号方法。

１６．　前記エネルギ補正係数の発生において、周波数帯域ごとに予め定めたエネルギ補正係数を発生させる請求項１４記載のオーディオ復号方法。

１７．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割処理と、

前記低域サブバンド信号に基づきエネルギ補正係数を算出する補正係数抽出処理と、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正処理と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるためのプログラム。

１８．　前記補正係数抽出処理において、前記低域サブバンド信号の信号位相を算出し、前記信号位相によりエネルギ補正係数を算出する請求項１７記載のプログラム。

１９．　前記補正係数抽出処理において、前記低域サブバンド信号の実部のエネルギと前記低域サブバンド信号の信号エネルギとの比をエネルギ補正係数として算出する請求項１７記載のプログラム。

２０．　前記補正係数抽出処理において、前記低域サブバンド信号の各サン

プルの位相を平均化してエネルギ補正係数を算出する請求項１７記載のプログラム。

２１． 前記補正係数抽出処理において、前記周波数帯域毎に算出したエネルギ補正係数を平滑化する請求項１７から２０のいずれか１項記載のプログラム。

２２． ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割処理と、

予め定めたエネルギ補正係数を発生する補正係数発生処理と、

前記高域ビットストリームに記載される目標エネルギを前記エネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正処理と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるプログラム。

２３． 前記補正係数発生処理において、乱数を発生させ当該乱数をエネルギ補正係数とする請求項２２記載のプログラム。

２４． 前記補正係数発生処理において、周波数帯域ごとに予め定めたエネ

ルギ補正係数を発生させる請求項２２記載のプログラム。

２５．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

前記高域ビットストリームに記載される目標エネルギを予め定められたエネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されるオーディオ復号装置。

２６．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

前記低域サブバンド信号を複写および加工して生成される信号のエネルギ補正係数を出力するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記エネルギ補正係数を用いて振幅

補正を行い、高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されるオーディオ復号装置。

２７．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

前記高域ビットストリームの指示と予め定められたエネルギ補正係数とを用いて補正目標エネルギを算出するエネルギ補正部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されるオーディオ復号装置。

２８．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギ

を前記高域ビットストリームに含まれる指示と予め定められたエネルギ補正係数を用いて振幅補正を行い、高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されるオーディオ復号装置。

２９．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割部と、

前記高域ビットストリームに記載される目標エネルギ（R）を予め定めたエネルギ補正係数（a）により補正した補正目標エネルギ（aR）と前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギ(Er)とが等しくなるように振幅補正する際に、目標エネルギ（R）に代えて前記信号エネルギ（Er）を前記エネルギ補正係数の逆数（1/a）により補正し、高域サブバンド信号を生成する帯域拡張部と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されるオーディオ復号装置。

３０．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップと、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域

サブバンド信号を生成するステップと、

前記高域ビットストリームに記載される目標エネルギを予め定められたエネルギ補正係数により補正し補正目標エネルギを算出するステップと、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

３１．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップと、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するステップと、

前記低域サブバンド信号を複写および加工して生成される信号のエネルギ補正係数を出力するステップと、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記エネルギ補正係数を用いて振幅補正を行い、高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

３２．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップと、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するステップと、

前記高域ビットストリームの指示と予め定められたエネルギ補正係数とを用いて補正目標エネルギを算出するステップと、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

３３．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップと、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するステップと、

前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記高域ビットストリームに含まれる指示と予め定められたエネルギ補正係数を用いて振幅補正を行い、高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

３４．　ビットストリームを低域ビットストリームと高域ビットストリーム

に分離するステップと、

前記低域ビットストリームを復号して低域オーディオ信号を生成するステップと、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するステップと、

前記高域ビットストリームに記載される目標エネルギ（R）を予め定めたエネルギ補正係数（a）により補正した補正目標エネルギ（aR）と前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギ(Er)とが等しくなるように振幅補正する際に、目標エネルギ（R）に代えて前記信号エネルギ（Er）を前記エネルギ補正係数の逆数（1/a）により補正して高域サブバンド信号を生成するステップと、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るステップとを備えたオーディオ復号方法。

３５．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割処理と、

前記高域ビットストリームに記載される目標エネルギを予め定められたエネルギ補正係数により補正し補正目標エネルギを算出するエネルギ補正処理と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成

フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるためのプログラム。

３６．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割処理と、

前記低域サブバンド信号を複写および加工して生成される信号のエネルギ補正係数を出力するエネルギ補正処理と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記エネルギ補正係数を用いて振幅補正を行い、高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるためのプログラム。

３７．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割処理と、

前記高域ビットストリームの指示と予め定められたエネルギ補正係数とを用いて補正目標エネルギを算出するエネルギ補正処理と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写およ

び加工して作成した信号の信号エネルギを前記補正目標エネルギに振幅補正して高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるためのプログラム。

３８．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割処理と、

前記低域サブバンド信号を複写および加工して作成した信号の信号エネルギを前記高域ビットストリームに含まれる指示と予め定められたエネルギ補正係数を用いて振幅補正を行い、高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるためのプログラム。

３９．　ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の実数信号に帯域分割して低域サブバンド信号を生成するサブバンド分割処理と、

前記高域ビットストリームに記載される目標エネルギ（R）を予め定めたエネルギ補正係数（a）により補正した補正目標エネルギ（aR）と前記低域サブバン

ド信号を複写および加工して作成した信号の信号エネルギ(Er)とが等しくなるように振幅補正する際に、目標エネルギ（R）に代えて前記信号エネルギ（Er）を前記エネルギ補正係数の逆数（1/a）により補正し、高域サブバンド信号を生成する帯域拡張処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理とをコンピュータに実行させるためのプログラム。

Fig. 1

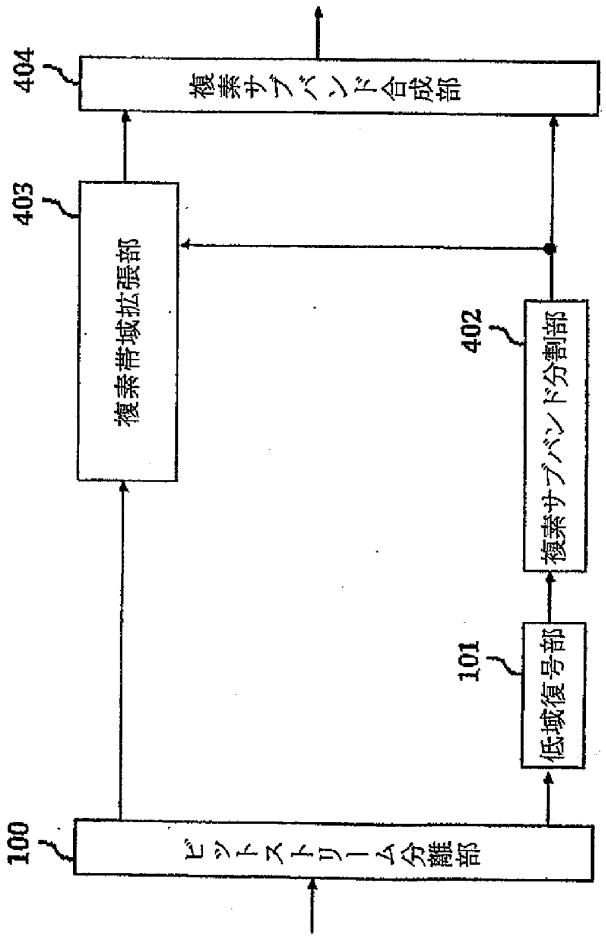

404
複素サブバンド合成部
403
複素帯域拡張部
402
複素サブバンド分割部
101
低域復号部
100
ビットストリーム分離部

Fig. 2

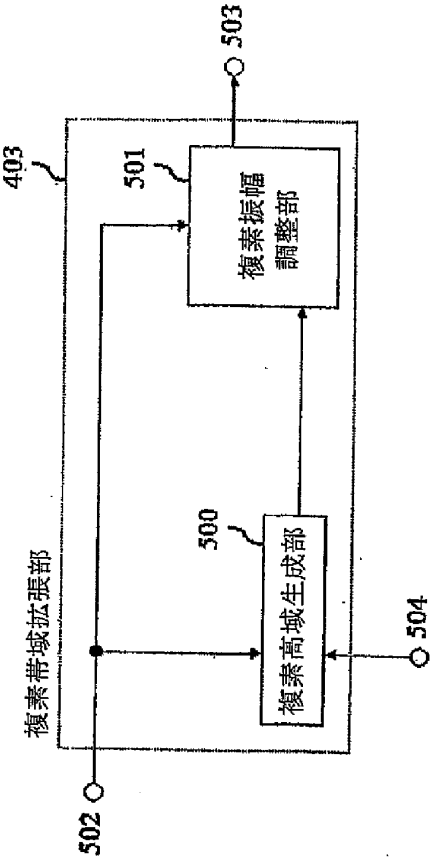


Fig. 3

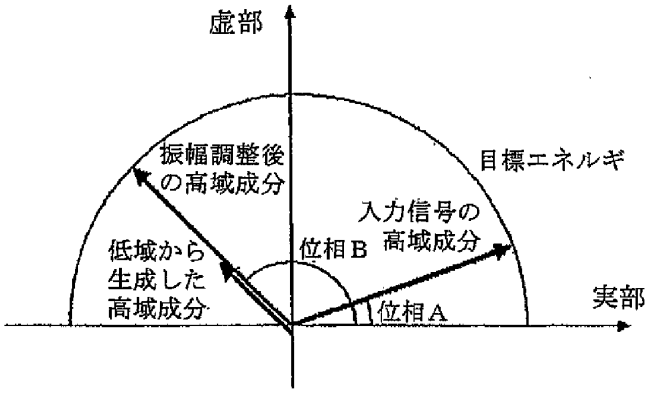

虚部
振幅調整後
の高域成分
目標エネルギ
入力信号の
高域成分
低域から
生成した
高域成分
位相B
位相A
実部

Fig. 4

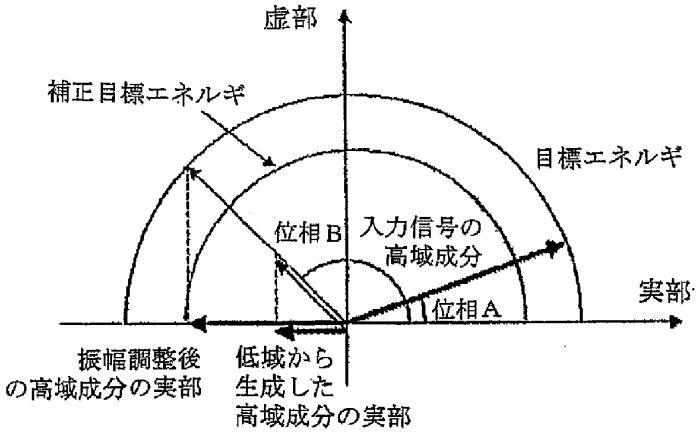

虚部
補正目標エネルギ
目標エネルギ
位相B
入力信号の
高域成分
実部
位相A
振幅調整後
の高域成分の実部
低域から
生成した
高域成分の実部

Fig. 5

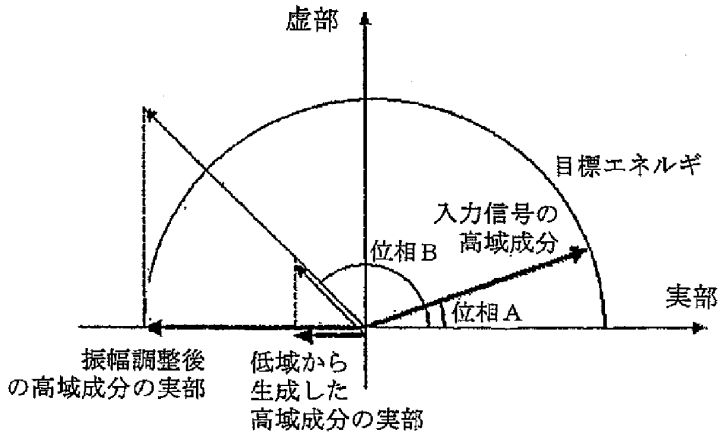

虚部
目標エネルギ
入力信号の
高域成分
位相B
位相A
実部
振幅調整後
の高域成分の実部
低域から
生成した
高域成分の実部

Fig. 6

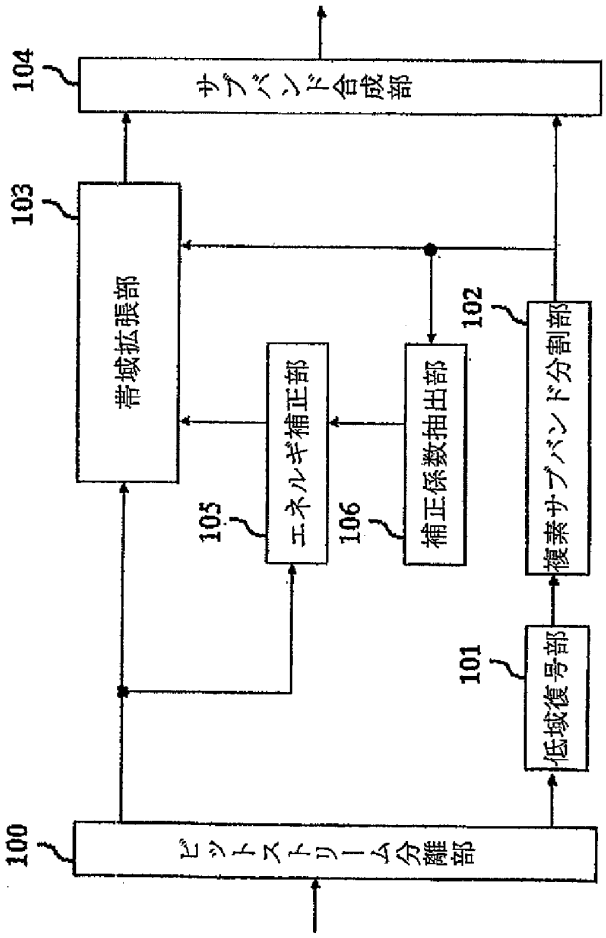

104
サブバンド合成部
103
帯域拡張部
102
複素サブバンド分割部
105
エネルギ補正部
106
補正係数抽出部
101
低域復号部
100
ビットストリーム分離部

Fig. 7

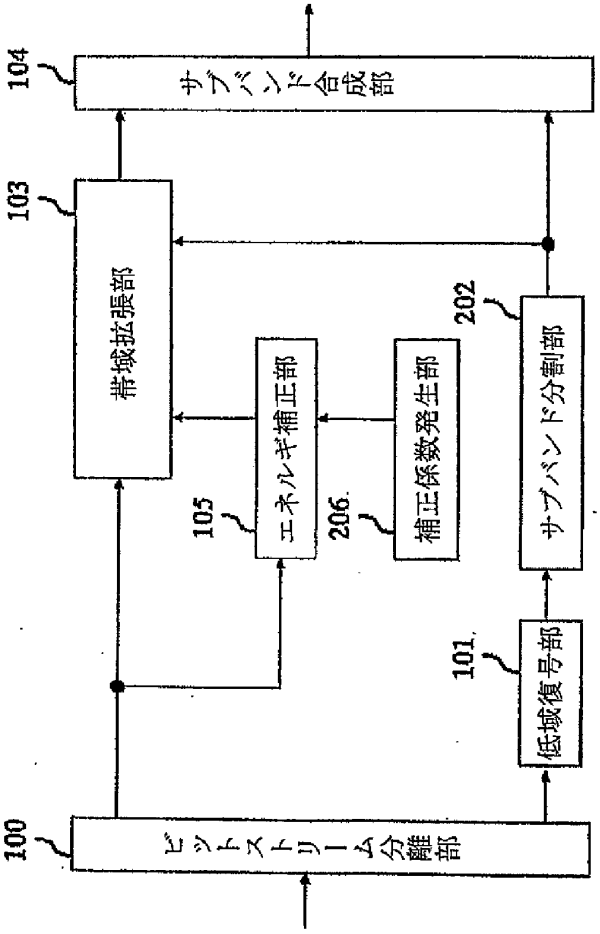
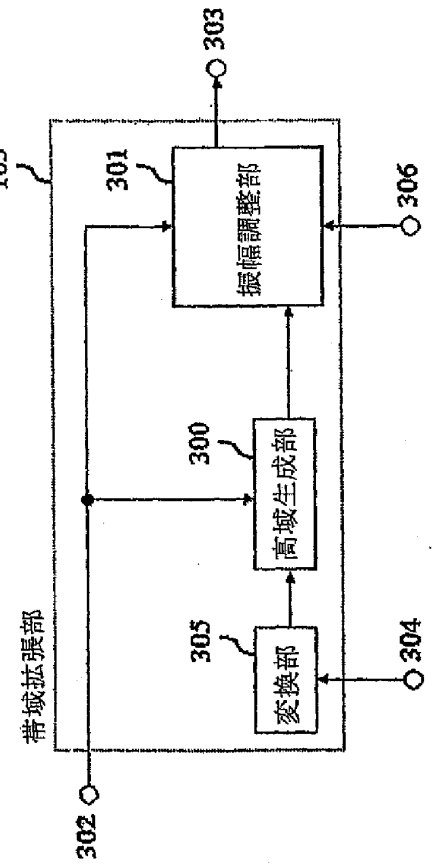

100
ビットストリーム分離部
101
低域復号部
202
サブバンド分割部
206
補正係数発生部
105
エネルギ補正部
103
帯域拡張部
104
サブバンド合成部

Fig. 8



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/07962

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[7] G10L19/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[7] G10L19/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2003
Kokai Jitsuyo Shinan Koho	1971-2003	Jitsuyo Shinan Toroku Koho	1996-2003

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JICST(JOIS)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 9-90992 A (Nippon Telegraph And Telephone Corp.), 04 April, 1997 (04.04.97), (Family: none)	1-39
A	JP 8-305396 A (Matsushita Electric Industrial Co., Ltd.), 22 November, 1996 (22.11.96), & EP 732687 A2 & US 5978759 A1	1-39
A	JP 9-101798 A (Matsushita Electric Industrial Co., Ltd.), 15 April, 1997 (15.04.97), (Family: none)	1-39

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 July, 2003 (17.07.03)	05 August, 2003 (05.08.03)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/07962

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 8-123495 A (Mitsubishi Electric Corp.), 17 May, 1996 (17.05.96), (Family: none)	1-39
A	WO 98/52187 A1 (Hewlett-Packard Co.), 19 November, 1998 (19.11.98), & EP 878790 A1 & JP 13-525079 A	1-39
A	JP 9-127998 A (Sony Corp.), 16 May, 1997 (16.05.97), (Family: none)	1-39
A	Madoka HARA et al., "Taisu Power Spectre-jo deno Senkei Yosoku o Mochiita Kyotaiiki Onsei kara Kotaiiki Onsei no Seiiku", Nihon Onkyo Gakkai Koen Ronbunshu", The Acoustical Society of Japan (ASJ), 1997 March, Shunki I, pages 277 to 278	1-39
A	Mineo TSUSHIMA et al., "Teiiki Kakudai ni yoru Kotaiiki Onsei no Gosei Hoho no Kento", Nihon Onkyo Gakkai Koen Ronbunshu, The Acoustical Society of Japan (ASJ), 1995 March, Shunki I, pages 249 to 250	1-39

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

国際調査報告	国際出願番号　PCT／JP03／07962

A.　発明の属する分野の分類（国際特許分類（IPC））

Int. Cl[7]　　G10L19／02

B.　調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int. Cl[7]　　G10L19／02

最小限資料以外の資料で調査を行った分野に含まれるもの
日本国実用新案公報　　1922－1996年
日本国公開実用新案公報　　1971－2003年
日本国登録実用新案公報　　1994－2003年
日本国実用新案登録公報　　1996－2003年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

JICST（JOIS）

C.　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
A	JP　9－90992　A（日本電信電話株式会社）1997.04.04（ファミリーなし）	1－39
A	JP　8－305396　A（松下電器産業株式会社）1996.11.22 & EP 732687 A2 & US 5978759 A1	1－39
A	JP　9－101798　A（松下電器産業株式会社）1997.04.15（ファミリーなし）	1－39

☒ C欄の続きにも文献が列挙されている。　　☐ パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
17. 07. 03	05.08.03

国際調査機関の名称及びあて先	特許庁審査官（権限のある職員）		
日本国特許庁（ISA／JP） 郵便番号100－8915 東京都千代田区霞が関三丁目4番3号	渡邊　聡	5C	8622
	電話番号　03－3581－1101　内線　3540		

様式PCT／ISA／210（第2ページ）（1998年7月）

国際調査報告	国際出願番号 PCT／JP03／07962

C（続き）. 関連すると認められる文献

引用文献のカテゴリー＊	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
A	JP 8－123495 A（三菱電機株式会社）1996.05.17 （ファミリーなし）	1－39
A	WO98／52187 A1（ヒューレット・パッカード・カンパニー）1998.11.19, & EP 878790 A1 & JP 13-525079 A	1－39
A	JP 9－127998 A（ソニー株式会社）1997.05.16 （ファミリーなし）	1－39
A	原円他，対数パワースペクトル上での線形予測を用いた狭帯域音声から広帯域音声の生成，日本音響学会講演論文集，日本音響学会，１９９７．０３，春季Ｉ，ｐ．２７７－２７８	1－39
A	津島峰生他，帯域拡大による広帯域音声の合成方法の検討，日本音響学会講演論文集，日本音響学会，１９９５．０３，春季Ｉ，ｐ．２４９－２５０	1－39

様式PCT／ISA／210（第2ページの続き）（1998年7月）

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **2004-053895**
(43)Date of publication of application : **19.02.2004**

(51)Int.Cl. **G10L 19/00**
G10L 19/02
H03M 7/30

(21)Application number : **2002-210946**
(22)Date of filing : **19.07.2002**

(71)Applicant : **NEC CORP**
MATSUSHITA ELECTRIC IND CO LTD

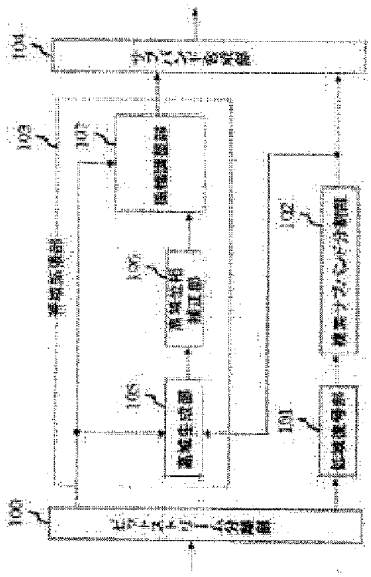
(72)Inventor : **SHIMADA OSAMU**
NOMURA TOSHIYUKI
TAKAMIZAWA YUICHIRO
SERIZAWA MASAHIRO
TANAKA NAOYA
TSUSHIMA MINEO
NORIMATSU TAKESHI

(54) DEVICE AND METHOD FOR AUDIO DECODING, AND PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a technology and a device that reduce an operation quantity while maintaining high sound quality.

SOLUTION: A high-frequency phase correction part 106 is newly provided which corrects the phase of the output of a high-frequency generation part 105. A wide-band phase correction part 106 performs phase correcting processes which are different by subbands needed when a real number subband composition filter performs subband composition of signals divided into subbands by a complex composition subband dividing filter. Consequently, the real number subband composition filter whose operation quantity is small is usable instead of the complex subband composition filter to reduce a necessary operation quantity and a device scale while maintaining high sound quality.



(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2004-53895
(P2004-53895A)
(43)公開日　平成16年2月19日(2004.2.19)

(51)Int.Cl.[7]	F I		テーマコード（参考）
G10L 19/00	G10L 9/18	M	5J064
G10L 19/02	H03M 7/30	A	
H03M 7/30	G10L 7/04	G	

審査請求 未請求　請求項の数 15　OL　（全 17 頁）

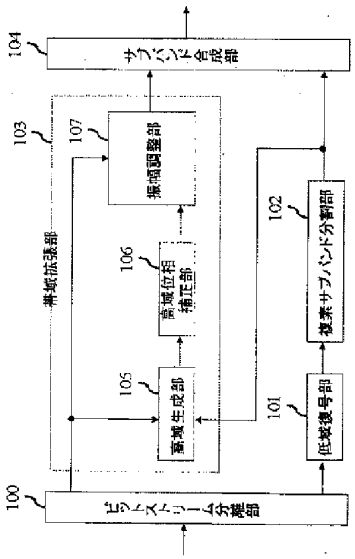
(21)出願番号　特願2002-210946(P2002-210946)
(22)出願日　平成14年7月19日(2002.7.19)

(71)出願人　000004237
日本電気株式会社
東京都港区芝五丁目７番１号
(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真１００６番地
(74)代理人　100088328
弁理士　金田　暢之
(74)代理人　100106297
弁理士　伊藤　克博
(74)代理人　100106138
弁理士　石橋　政幸
(72)発明者　嶋田　修
東京都港区芝五丁目７番１号　日本電気株式会社内

最終頁に続く

(54)【発明の名称】オーディオ復号装置と復号方法およびプログラム

(57)【要約】
【課題】高音質を維持しつつ演算量を低減させる。
【解決手段】高域生成部１０５の出力の位相を補正する高域位相補正部１０６を新たに設ける。複素サブバンド分割フィルタでサブバンド分割された信号を実数のサブバンド合成フィルタでサブバンド合成する際に必要となるサブバンド毎に異なる位相補正処理を、広域位相補正部１０６が行う。これにより、複素サブバンド合成フィルタの代わりに、より演算量が少ない実数のサブバンド合成フィルタを用いることができ、高音質を維持しつつ必要演算量・装置規模を低減できる。
【選択図】　図１



【特許請求の範囲】
【請求項１】
ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成部と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正部と、
前記位相補正複写加工サブバンド信号の信号エネルギを前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整部と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されることを特徴とするオーディオ復号装置。
【請求項２】
ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成部と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正部と、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成するトーン・ノイズ付加部と、
前記トーン・ノイズ信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正トーン・ノイズサブバンド信号を生成するトーン・ノイズ位相補正部と、
前記位相補正複写加工サブバンド信号の信号エネルギと前記位相補正トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整部と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されることを特徴とするオーディオ復号装置。
【請求項３】
ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成部と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正部と、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成するトーン・ノイズ付加部と、

前記位相補正複写加工サブバンド信号の信号エネルギと前記トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整部と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されることを特徴とするオーディオ復号装置。
【請求項４】
前記高域位相補正部が行う位相補正において、前記複写加工サブバンド信号を生成する際に用いられた複写元の周波数と複写先の周波数の差分に応じて位相補正量が決まることを特徴とする請求項１から３のいずれか１項に記載のオーディオ復号装置。
【請求項５】
前記トーン・ノイズ位相補正部が行う位相補正において、前記トーン・ノイズ信号の周波数に応じて位相補正量が決まることを特徴とする請求項２記載のオーディオ復号装置。
【請求項６】
ビットストリームを低域ビットストリームと高域ビットストリームに分離し、
前記低域ビットストリームを復号して低域オーディオ信号を生成し、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成し、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成し、前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成し、
前記位相補正複写加工サブバンド信号の信号エネルギを前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成し、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得ることを特徴とするオーディオ復号方法。
【請求項７】
ビットストリームを低域ビットストリームと高域ビットストリームに分離し、
前記低域ビットストリームを復号して低域オーディオ信号を生成し、前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成し、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成し、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成し、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成し、
前記トーン・ノイズ信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正トーン・ノイズサブバンド信号を生成し、
前記位相補正複写加工サブバンド信号の信号エネルギと前記位相補正トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成し、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得ることを特徴とするオーディオ復号方法。
【請求項８】
ビットストリームを低域ビットストリームと高域ビットストリームに分離し、
前記低域ビットストリームを復号して低域オーディオ信号を生成し、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成し、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成し、

前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成し、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成し、
前記位相補正複写加工サブバンド信号の信号エネルギと前記トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成し、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得ることを特徴とするオーディオ復号方法。

【請求項９】
前記複写加工サブバンド信号に対する位相補正において、前記複写加工サブバンド信号を生成する際に用いられた複写元の周波数と複写先の周波数の差分に応じて位相補正量が決まることを特徴とする請求項６から８のいずれか１項に記載のオーディオ復号方法。

【請求項１０】
前記トーン・ノイズ信号に対する位相補正において、前記トーン・ノイズ信号の周波数に応じて位相補正量が決まることを特徴とする請求項７記載のオーディオ復号方法。

【請求項１１】
ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割処理と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成処理と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正処理と、
前記位相補正複写加工サブバンド信号の信号エネルギを前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整処理と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理をコンピュータに実行させるためのプログラム。

【請求項１２】
ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割処理と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成処理と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正処理と、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成するトーン・ノイズ付加処理と、
前記トーン・ノイズ信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正トーン・ノイズサブバンド信号を生成するトーン・ノイズ位相補正処理と、
前記位相補正複写加工サブバンド信号の信号エネルギと前記位相補正トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理をコンピュータに実行させるためのプログラム。

【請求項１３】

ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離処理と、

前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号処理と、

前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割処理と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成処理と、

前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正処理と、

前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成するトーン・ノイズ付加処理と、

前記位相補正複写加工サブバンド信号の信号エネルギと前記トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整処理と、

前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成処理をコンピュータに実行させるためのプログラム。

【請求項１４】

前記高域位相補正処理が行う位相補正において、前記複写加工サブバンド信号を生成する際に用いられた複写元の周波数と複写先の周波数の差分に応じて位相補正量が決まることをコンピュータにさらに実行させることを特徴とする請求項１１から１３のいずれか１項に記載のプログラム。

【請求項１５】

前記トーン・ノイズ位相補正処理が行う位相補正において、前記トーン・ノイズ信号の周波数に応じて位相補正量が決まることをコンピュータにさらに実行させることを特徴とする請求項１２に記載のプログラム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、符号化されたオーディオ信号を復号するためのオーディオ復号装置および復号方法に関する。

【０００２】

【従来の技術】

低ビットレートで高音質に符号化可能なオーディオ符号化・復号方式として、ＩＳＯ／ＩＥＣの国際標準方式であるＭＰＥＧ－２　ＡＡＣ（Ａｄｖａｎｃｅｄ　Ａｕｄｉｏ　Ｃｏｄｉｎｇ）が広く知られている。このＭＰＥＧ－２　ＡＡＣに代表される従来のオーディオ符号化復号方式では、まず、時間領域のＰＣＭ信号を複数サンプルまとめて１フレームとし、これをＭＤＣＴ（Ｍｏｄｉｆｉｅｄ　Ｄｉｓｃｒｅｔｅ　Ｃｏｓｉｎｅ　Ｔｒａｎｓｆｏｒｍ）などの写像変換により周波数領域信号へと変換する。そして、この周波数領域信号を量子化およびハフマン符号化してビットストリームを形成する。量子化においては人間の聴覚特性を考慮して、周波数領域信号の中で知覚されやすい周波数成分については量子化精度を高く、そうでない周波数成分については量子化精度を低くすることにより、限られた符号量で高い音質を実現する。例えば、ＭＰＥＧ－２　ＡＡＣでは９６ｋｂｐｓ程度のビットレートでＣＤと同等音質（４４．１ｋＨｚサンプリング、ステレオ）を実現できる。

【０００３】
これより低いビットレート、例えば４８ｋｂｐｓ程度のビットレートで４４．１ｋＨｚサンプリングのステレオオーディオ信号を符号化しようとする場合、聴覚的に重要度が低い高域周波数成分を符号化対象から外す（量子化値を零とする）ことで限られたビットレートでの主観的音質が最高となるようにする。しかしながらこの場合、高域周波数成分が符号化されていないために音質が劣化し、一般的にはこもった感じの音となってしまう。
【０００４】
このような低ビットレートでの音質劣化を解決する技術として帯域拡張技術が注目を集めている。帯域拡張技術では、ＭＰＥＧ－２　ＡＡＣなどの方式で低ビットレートに符号化した低域ビットストリームに、僅かな符号量（一般に数ｋｂｐｓ程度）の補助情報である高域ビットストリームを付加してビットストリームを構成する。オーディオ復号装置（デコーダ）では、まず、ＭＰＥＧ－２　ＡＡＣなどの復号方法で低域ビットストリームを復号して、高域周波数成分が含まれていない低域オーディオ信号を得る。そして、帯域拡張技術により、高域ビットストリームが示す補助情報に従った加工を低域オーディオ信号に施して高域周波数成分を生成する。
【０００５】
こうして得られた高域周波数成分と、低域ビットストリームを復号して得られた低域オーディオ信号を合成して高域周波数成分を含む復号オーディオ信号を得る。
【０００６】
帯域拡張技術を用いた従来のオーディオ復号装置の一例として、ＭＰＥＧ－２ＡＡＣ復号装置にＳＢＲと呼ばれる帯域拡張技術を組み合わせた例が、ＥＴＳＩが発行する「Ｄｉｇｉｔａｌ　Ｒａｄｉｏ　Ｍｏｎｄｉａｌｅ（ＤＲＭ）；Ｓｙｓｔｅｍ　Ｓｐｅｃｉｆｉｃａｔｉｏｎ」（ＥＴＳＩ　ＴＳ　１０１　９８０Ｖ１．１．１）の５．６．２節に記述されている。ここに記述されているような帯域拡張技術を用いた従来のオーディオ復号装置を図４および図５に示す。
【０００７】
図４に示す従来のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、複素サブバンド分割部４０２、帯域拡張部４０３、複素サブバンド合成部４０４から構成される。また、帯域拡張部４０３は、高域生成部１０５と振幅調整部１０７から構成される。
【０００８】
ビットストリーム分離部１００は、入力されたビットストリームを分離して低域復号部１０１と帯域拡張部４０３へ出力する。ビットストリームには低域信号がＭＰＥＧ－２　ＡＡＣなどの符号化方式で符号化されている低域ビットストリームと、帯域拡張部が高域信号を生成するために必要な情報を含む高域ビットストリームが多重化されている。低域ビットストリームは低域復号部１０１へ、高域ビットストリームは帯域拡張部４０３へ出力される。
【０００９】
低域復号部１０１は、入力された低域ビットストリームを復号して低域オーディオ信号を生成し、複素サブバンド分割部４０２へ出力する。低域復号部１０１はＭＰＥＧ－２　ＡＡＣなど既存のオーディオ復号方式を用いて低域ビットストリームを復号する。
【００１０】
複素サブバンド分割部４０２は、入力された低域オーディオ信号を複素サブバンド分割フィルタによって帯域分割する。帯域分割された低域サブバンド信号は帯域拡張部４０３および複素サブバンド合成部４０４へ出力される。複素サブバンド分割フィルタについては、例えば、従来から広く知られている３２バンド複素ＱＭＦ（Ｑｕａｄｒａｔｕｒｅ　Ｍｉｒｒｏｒ　Ｆｉｌｔｅｒ：直交鏡像フィルタ）フィルタバンクなどを用いることができる。この場合、３２サブバンドに帯域分割された複素数の低域信号（低域サブバンド信号）が帯域拡張部４０３および複素サブバンド合成部４０４へ出力される。３２バンド複素分析ＱＭＦフィルタバンクの計算式は下記の式４０２．１、４０２．２のようになる。

【００１１】
【数１】

$$X_k(m) = \sum_{n=-\infty}^{\infty} h(mM - n)\, x(n) W_{K1}^{-(k+k_0)(i+n_0)}, \quad k = 0,1,\ldots,K1-1 \quad \cdots \quad 402.1$$

$$W_{K1} = e^{j\frac{2\pi}{K1}} \quad \cdots \quad 402.2$$

【００１２】
ここで、Ｘ（ｎ）は低域オーディオ信号を表し、Ｘｋ（ｍ）はｋバンド目の低域サブバンド信号を表し、ｈ（ｎ）は分析用低域通過フィルタを表す。本例ではＫ１＝６４である。
【００１３】
帯域拡張部４０３は、入力された高域ビットストリームと低域サブバンド信号から高域のオーディオ信号を表す高域サブバンド信号を生成してサブバンド合成部４０４へ出力する。帯域拡張部４０３は、高域生成部１０５と振幅調整部１０７から構成される。
【００１４】
高域生成部１０５は、低域サブバンド信号と高域ビットストリームを入力とし、低域サブバンド信号の中から高域ビットストリームが指定する帯域の信号を高域サブバンドへと複写する。また、この複写の際には高域ビットストリームが指定する信号加工を施す場合もある。例えば、周波数が低い順にサブバンド０からサブバンド６３まで６４個のサブバンドを仮定し、低域サブバンド信号としてサブバンド０からサブバンド１９までの複素数のサブバンド信号が複素サブバンド分割部４０２から供給されているとする。また高域ビットストリームには、サブバンドＡ（Ａ＞１９）をどの低域サブバンド（サブバンド０から１９）から信号を複写して生成するかを表す複写情報、および、複写の際に施す信号加工方法（フィルタリングなどあらかじめ定められた数種類の手段から選択して用いられる）を示す信号加工情報が含まれているとする。高域生成部１０５の処理は、高域のサブバンドの複素信号（複写加工サブバンド信号と名付ける）を複写情報が示す低域のサブバンドの複素信号と同一とする。さらに、音質向上のために信号加工情報が何らかの信号加工の必要性を表している場合は、複写加工サブバンド信号に信号加工情報が示す信号加工を施す。このようにして生成された複写加工サブバンド信号は振幅調整部１０７に出力される。
【００１５】
振幅調整部１０７は、入力された複写加工サブバンド信号の振幅に対して、高域ビットストリームが指定する補正を施して高域サブバンド信号を生成する。つまり、符号化前のオーディオ信号の高域信号エネルギと、帯域拡張部４０３が生成する信号の高域信号エネルギが等価となるような振幅補正を複写加工サブバンド信号に対して施す。そのために、高域ビットストリームには符号化前のオーディオ信号の高域信号エネルギを表す情報が含まれている。生成された高域サブバンド信号は複素サブバンド合成部４０４へ出力される。
【００１６】
複素サブバンド合成部４０４は、入力された低域サブバンド信号と高域サブバンド信号を複素サブバンド合成フィルタによって帯域合成する。この帯域合成により生成されたオーディオ信号がオーディオ復号装置の出力となる。複素サブバンド合成フィルタは、複素サブバンド分割部４０２で用いられている複素サブバンドフィルタに対応したものが用いられる。つまり、ある信号を複素サブバンド分割フィルタで帯域分割し、これを複素サブバンド合成フィルタで帯域合成すると元の信号（複素サブバンド分割フィルタの入力信号）が完全に再構成できるように選ばれる。例えば、複素サブバンド合成フィルタとして式４０２．１に示したような３２バンド複素ＱＭＦ分割フィルタバンク（Ｋ１＝６４）が用いられている場合、
【００１７】

【数２】

$$x(n) = \sum_{m=-\infty}^{\infty} f(n - mM) \frac{1}{K2} \sum_{k=0}^{K2-1} X_k(m) W_{K2}^{(k+k_0)(n+n_0)} \quad \cdots \quad 404.1$$

【００１８】
を用いることができる。ここで、ｆ（ｎ）は合成用低域通過フィルタを表す。本例ではＫ２＝６４である。
【００１９】
また、帯域拡張技術によって低域復号部１０１が出力するオーディオ信号のサンプリング周波数よりも複素サブバンド合成部４０４が出力するオーディオ信号のサンプリング周波数が高い場合は、複素サブバンド合成部４０４が出力するオーディオ信号の低域部（ダウンサンプリングした結果）が低域復号部１０１が出力するオーディオ信号と等しくなるように選ばれる。複素サブバンド合成部４０４では６４バンド複素ＱＭＦ合成フィルタバンク（式４０４．１においてＫ２＝１２８）を用いることができる。なお、この場合、低域側３２バンドには３２バンド複素分析ＱＭＦフィルタバンクの出力が信号値として用いられる。
【００２０】
図５に示す従来のオーディオ復号装置は、図４と比較して、帯域拡張部５０３にトーン・ノイズ付加部２０８が追加されている点で異なっている。
【００２１】
トーン・ノイズ付加部２０８は、高域ビットストリームに含まれているトーン情報とノイズ情報を入力とし、トーン・ノイズ信号を振幅調整部２０７へ出力する。トーン情報は、どの周波数帯域にトーン音（正弦波）を付加するかを表す。一般に高域のトーン音はその低域信号から高音質に生成するのが難しいため、このような別手段で付加するとよい。また、ノイズ情報には、どの周波数にどのエネルギの白色雑音を付加するかを表している。トーン・ノイズ付加部２０８はトーン情報とノイズ情報を元にトーン音と白色雑音を発生させ、これをトーン・ノイズ信号として振幅調整部２０７へ出力する。
【００２２】
振幅調整部２０７では、高域生成部１０５から入力された複写加工サブバンド信号に、トーン・ノイズ付加部２０８から入力されたトーン・ノイズ信号を加算し、そのエネルギが所望のものとなるように、振幅調整部１０７と同様に高域ビットストリームが指定する補正を施して高域サブバンド信号を生成する。生成された高域サブバンド信号は複素サブバンド合成部４０４へ出力される。
【００２３】
【発明が解決しようとする課題】
上述した従来のオーディオ復号装置では、演算量が大きい複素サブバンド分割部および複素サブバンド合成部を有するため、必要演算量・装置規模が大きいという問題点があった。
【００２４】
本発明の目的は、高音質を維持しつつ演算量を低減した帯域拡張技術、およびこれを用いたオーディオ復号装置、オーディオ復号方法およびオーディオ復号用プログラムを提供することにある。
【００２５】
【課題を解決するための手段】
上記目的を達成するために、本発明のオーディオ復号装置は、ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、

前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成部と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正部と、
前記位相補正複写加工サブバンド信号の信号エネルギを前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整部と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されることを特徴とする。

【００２６】

また、本発明の他のオーディオ復号装置は、ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、
前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成部と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正部と、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成するトーン・ノイズ付加部と、
前記トーン・ノイズ信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正トーン・ノイズサブバンド信号を生成するトーン・ノイズ位相補正部と、
前記位相補正複写加工サブバンド信号の信号エネルギと前記位相補正トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整部と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されることを特徴とする。

【００２７】

さらに、本発明の他のオーディオ復号装置は、ビットストリームを低域ビットストリームと高域ビットストリームに分離するビットストリーム分離部と、
前記低域ビットストリームを復号して低域オーディオ信号を生成する低域復号部と、前記低域オーディオ信号を複数の周波数帯域の複素数信号に帯域分割して低域サブバンド信号を生成する複素サブバンド分割部と、
前記高域ビットストリームの指示に従い前記低域サブバンド信号を複写および加工して複写加工サブバンド信号を生成する高域生成部と、
前記複写加工サブバンド信号の位相をその周波数に応じて予め定められた位相補正を施して位相補正複写加工サブバンド信号を生成する高域位相補正部と、
前記高域ビットストリームが指示する周波数を持つトーン信号と前記高域ビットストリームが指示する周波数帯域とエネルギを持つノイズ信号から構成されるトーン・ノイズ信号を生成するトーン・ノイズ付加部と、
前記位相補正複写加工サブバンド信号の信号エネルギと前記トーン・ノイズ信号の信号エネルギの合計を前記高域ビットストリームが指示する信号エネルギに振幅補正して高域サブバンド信号を生成する振幅調整部と、
前記低域サブバンド信号と前記高域サブバンド信号の実数部をサブバンド合成フィルタにより帯域合成して復号オーディオ信号を得るサブバンド合成部と、から構成されることを特徴とする。

【００２８】

また、前記高域位相補正部が行う位相補正において、前記複写加工サブバンド信号を生成する際に用いられた複写元の周波数と複写先の周波数の差分に応じて位相補正量を決めるようにしてもよい。さらに、前記トーン・ノイズ位相補正部が行う位相補正において、前記トーン・ノイズ信号の周波数に応じて位相補正量を決めるようにしてもよい。

【００２９】

本発明によれば、高域生成部の出力の位相を補正する高域位相補正部とトーン・ノイズ付加部の出力の位相を補正するトーン・ノイズ位相補正部を設けたことを特徴としている。複素サブバンド分割フィルタでサブバンド分割された信号を実数のサブバンド合成フィルタでサブバンド合成する際に必要となるサブバンド毎に異なる位相補正処理を、これらの位相補正部が実行する。これにより、高音質を維持しつつ、複素サブバンド合成フィルタの代わりに、より演算量が少ない実数のサブバンド合成フィルタを用いることができ、必要演算量・装置規模を低減することができる。

【００３０】

【発明の実施の形態】

次に、本発明の実施の形態について図面を参照して詳細に説明する。

【００３１】

（第１の実施形態）

図１は本発明の第１の実施形態のオーディオ復号装置の構成を示すブロック図である。

【００３２】

本実施形態のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、複素サブバンド分割部１０２、帯域拡張部１０３、サブバンド合成部１０４とから構成される。さらに、帯域拡張部１０３は、高域生成部１０５、高域位相補正部１０６、振幅調整部１０７から構成される。

【００３３】

ビットストリーム分離部１００は、入力されたビットストリームを分離して低域復号部１０１と帯域拡張部１０３へ出力する。ビットストリームには低域信号が符号化されている低域ビットストリームと、帯域拡張部が高域信号を生成するために必要な情報を含む高域ビットストリームが多重化されている。低域ビットストリームは低域復号部１０１へ、高域ビットストリームは帯域拡張部１０３へ出力される。

【００３４】

低域復号部１０１は、入力された低域ビットストリームを復号して低域オーディオ信号を生成し、複素サブバンド分割部１０２へ出力する。低域復号部１０１はＭＰＥＧ－２　ＡＡＣなど既存のオーディオ復号方式を用いて低域ビットストリームを復号する。

【００３５】

複素サブバンド分割部１０２は、入力された低域オーディオ信号を複素サブバンド分割フィルタによって帯域分割する。帯域分割された低域サブバンド信号は帯域拡張部１０３およびサブバンド合成部１０４へ出力される。

【００３６】

帯域拡張部１０３は、入力された高域ビットストリームと低域サブバンド信号から高域のオーディオ信号を表す高域サブバンド信号を生成してサブバンド合成部１０４へ出力する。

【００３７】

サブバンド合成部１０４は、入力された低域サブバンド信号と高域サブバンド信号をサブバンド合成フィルタによって帯域合成する。帯域合成により生成されたオーディオ信号がオーディオ復号装置の出力となる。

【００３８】

高域生成部１０５は、低域サブバンド信号と高域ビットストリームを入力とし、低域サブバンド信号の中から高域ビットストリームが指定する帯域の信号を加工して複写加工サブバンド信号を生成する。生成された複写加工サブバンド信号は高域位相補正部１０６に出力される。

【００３９】
高域位相補正部１０６は、入力された複写加工サブバンド信号の位相をサブバンド毎に補正し、位相補正複写加工サブバンド信号を生成する。生成された位相補正複写加工サブバンド信号は振幅調整部１０７へ出力される。
【００４０】
振幅調整部１０７は、入力された位相補正複写加工サブバンド信号の振幅に対して、高域ビットストリームが指定する補正を施して高域サブバンド信号を生成する。生成された高域サブバンド信号はサブバンド合成部１０４へ出力される。
【００４１】
このように構成される本発明のオーディオ復号装置と図４の従来のオーディオ復号装置の相違は、図４の複素サブバンド分割部４０２の替わりに図１では複素サブバンド分割部１０２があり、図４の複素サブバンド合成部４０４の替わりに図１ではサブバンド合成部１０４があり、また、本発明（図１）には高域位相補正部１０６が追加されていることにある。他の処理部については、従来方式と同じであり当業者にとってよく知られており、また本発明とは直接関係しないので、その詳細な説明は省略する。以下では従来と手法が異なる複素サブバンド分割部１０２、サブバンド合成部１０４、高域位相補正部１０６について詳細に説明する。
【００４２】
まず、複素サブバンド分割部１０２とサブバンド合成部１０４について説明する。従来はサブバンド分割フィルタとして複素数のサブバンド信号を生成する式４０２．１のようなフィルタバンクを用いていた。また、これに対応する逆変換として式４０４．１のようなフィルタバンクをサブバンド合成フィルタとして用いていた。
【００４３】
式４０４．１の出力、あるいは、式４０４．１の出力を式４０２．１の入力信号のサンプリング周波数にダウンサンプリングした信号は式４０２．１の入力と完全に一致する完全再構成性がある。高音質な復号オーディオ信号を得るためには、この完全再構成性の関係がサブバンド分割・合成フィルタに必要となる。
【００４４】
本発明においては、演算量を低減するために、従来の複素サブバンド合成部（図４～５の４０４）で用いられている複素サブバンド合成フィルタを、実数のサブバンド合成フィルタ（図１～３の１０４）に置き換えている。しかしながら、単にサブバンド合成フィルタを複素数から実数に変更してしまうと完全再構成性は失われ、音質が劣化する。
【００４５】
複素数のサブバンド分割フィルタと実数のサブバンド合成フィルタの間で完全再構成性を実現するためには、従来の複素サブバンド分割フィルタ（図４～５の４０２）の出力に回転演算を施すと良いことが従来から広く知られている。この回転演算は複素数の実数軸と虚数軸を（$\pi \div 4$）だけ回転させるものであり、一般的に良く知られたＤＦＴからＤＣＴを導出する手法と同じである。例えば、ｋ０＝１／２の場合、式４０２．１の３２バンド複素ＱＭＦ分割フィルタバンクの計算式には、各サブバンドｋに対して以下の回転演算（Ｋ＝Ｋ１）を施すと良い。
【００４６】
【数３】

$$W_K^{-(k+k_0)\frac{3}{4}K} \quad \cdots\cdots \quad 102.1$$

【００４７】
なお、式１０２．１における３／４Ｋは１／４Ｋでも良い。
【００４８】
このように従来の複素サブバンド分割部（図４～５の４０２）の後段に式１０２．１の回

転演算を行う処理部を付加したものを複素サブバンド分割部（図１～３の１０２）としても良いが、これらサブバンド分割フィルタと回転演算処理を合わせた処理と等価な演算をより少ない演算量で実現可能な下式を複素サブバンド分割部（図１～３の１０２）で行うと良い。

【００４９】

【数４】

$$X_k(m) = \sum_{n=-\infty}^{\infty} h(mM - n)x(n)W_{K1}^{-(k+k_0)(n+n_0+\frac{3}{4}K1)},\ k = 0,1,\ldots,K1-1 \cdots 102.2$$

【００５０】

これに対応する実数のサブバンド合成フィルタとしては、式４０４．１に

【００５１】

【数５】

$$W_K^{(k+k_0)\frac{3}{4}K} \cdots\cdots 104.1$$

【００５２】

という変換を施し、実数部分のみを取り出した

【００５３】

【数６】

$$x(n) = \sum_{m=-\infty}^{\infty} f(n - mM)\frac{2}{K2}\sum_{k=0}^{K2-1}\mathrm{Re}[X_k(m)]\cos(\frac{2\pi}{K2}(n+n_0)(k+k_0+\frac{3}{4}K2)) \cdots 104.2$$

【００５４】

をサブバンド合成部（図１～３の１０４）で用いれば、完全再構成性を実現できるようになる。ここで、Ｒｅ［．］は複素サブバンド信号の実数部のみを取り出すことを表す。

【００５５】

回転演算は式１０２．２で表されるようにバンド番号ｋの関数となっているため、式１０２．２を用いた複素サブバンド分割部１０２と式１０４．２を用いたサブバンド合成部１０４の間に帯域拡張処理が入ると位相の不整合が起こる。すなわち、帯域拡張部（図１の１０３、図２の２０３、図３の３０３）から出力される高域サブバンド信号は、低域のバンド番号に応じた回転演算が施された低域サブバンド信号を複写加工して生成されたものであり、サブバンド合成部において高域のバンド番号に応じた逆回転演算が施されると、サブバンド合成部の出力信号における高域成分の位相に狂いが生じる。

【００５６】

そこで高域位相補正部１０４が以下のように位相の補正をする。

【００５７】

【数７】

$$W_K^{-(k2-k1+k_0)\frac{3}{4}K} \cdots\cdots 106.1$$

【００５８】

ここで、ｋ２は複写先の高域のバンド番号を表し、ｋ１は複写元の低域バンド番号を表す。すなわち、高域生成部１０５がサブバンドｋ２の信号を、サブバンドｋ１の信号を複写加工して生成した場合に高域位相補正部１０６において式１０６．１の位相補正処理を行うということである。さらに、式１０６．１はＣＯＳ関数及びＳＩＮ関数の特性を利用し

て低演算量で実現できる。

【００５９】

（第２の実施形態）

次に、本発明の第２の実施の形態について図２を用いて詳細に説明する。

【００６０】

図２を参照すると、本発明の第２の実施の形態としてのオーディオ復号装置が示されている。本実施形態のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、複素サブバンド分割部１０２、帯域拡張部２０３、サブバンド合成部１０４とから構成される。さらに、帯域拡張部２０３は、高域生成部１０５、高域位相補正部１０６、振幅調整部２０７、トーン・ノイズ付加部２０８、トーン・ノイズ位相補正部２０９から構成される。

【００６１】

本発明の第２の実施の形態は、本発明の第１の実施の形態と比較して、トーン・ノイズ付加部２０８とトーン・ノイズ位相補正部２０９が追加され、振幅調整部２０７がトーン・ノイズ位相補正部の出力信号をも入力信号としていることのみであり、他の部分については全く同一である。さらにトーン・ノイズ付加部２０８および振幅調整部２０７は図５に示した従来技術と全く同一であり、本発明の第２の実施の形態は、トーン・ノイズ位相補正部２０９を追加したことのみである。そこでトーン・ノイズ位相補正部２０９について詳細に説明する。

【００６２】

トーン・ノイズ位相補正部２０９は高域位相補正部１０６と同様に、例えば　ｋ０　＝　１／２　の場合、サブバンド合成部１０４における逆回転演算を考慮して以下の回転演算

【００６３】

【数８】

$$W_K^{-(k2+k_0)\frac{3}{4}K} \quad \cdots\cdots \quad 209.1$$

【００６４】

を施す必要がある。なお、式２０９．１はｃｏｓ関数及びｓｉｎ関数の特性を利用することにより簡略化でき、トーン・ノイズ位相補正部２０９の処理は低演算量で実現できる。さらに、トーン信号およびノイズ信号、あるいはそのどちらかをトーン・ノイズ付加部２０８においてテーブル化して蓄える場合は、位相補正を施したトーン信号およびノイズ信号、あるいはそのどちらかをテーブル化して蓄えることにより、トーン・ノイズ位相補正部２０９を備えるのと同等な効果をより低演算量で得ることができる。

【００６５】

（第３の実施形態）

次に、本発明の第３の実施の形態について図３を用いて詳細に説明する。

【００６６】

図３を参照すると、本発明の第３の実施の形態としてのオーディオ復号装置が示されている。本実施形態のオーディオ復号装置は、ビットストリーム分離部１００、低域復号部１０１、複素サブバンド分割部１０２、帯域拡張部３０３、サブバンド合成部１０４とから構成される。さらに、帯域拡張部２０３は、高域生成部１０５、高域位相補正部１０６、振幅調整部２０７、トーン・ノイズ付加部２０８から構成される。

【００６７】

本発明の第３の実施の形態は、本発明の第２の実施の形態と比較して、トーン・ノイズ位相補正部２０９が無くなり、トーン・ノイズ付加部２０８の出力が、トーン・ノイズ位相補正部２０９を通さずに振幅調整部２０７の入力信号となっていることのみであり、他の部分については全く同一である。トーン・ノイズ信号に対して位相補正を行わないために復号オーディオ信号の品質は他の実施の形態よりも低下するが、更なる低演算量化を実現

できる。

【００６８】

なお、本発明が上記各実施形態に限定されず、本発明の技術思想の範囲内において、各実施形態は適宜変更され得ることは明らかである。

【００６９】

また、図には示されていないが、本実施形態のオーディオ復号装置は、上記で説明したオーディオ復号方法を実行するためのプログラムを記録した記録媒体を備えている。この記録媒体は磁気ディスク、半導体メモリまたはその他の記録媒体であってもよい。このプログラムは、記録媒体からオーディオ復号装置に読み込まれ、オーディオ復号装置の動作を制御する。具体的には、オーディオ復号装置内のＣＰＵがこのプログラムの制御によりオーディオ復号装置のハードウェア資源に特定の処理を行うように指示することにより上記の処理が実現される。

【００７０】

【発明の効果】

以上説明したように、本発明によれば、高域生成部の出力に位相補正処理を行う高域位相補正部と、トーン・ノイズ付加部の出力に位相補正処理を行うトーン・ノイズ位相補正部のいずれか、または両方を新たに追加するという基本構成によって、サブバンド合成フィルタを複素フィルタバンクから演算量の少ない実数のフィルタバンクに変更しても高音質な復号オーディオ信号を得ることができるという効果が得られる。

【図面の簡単な説明】

【図１】本発明の第１の実施形態のオーディオ復号装置の構成を示すブロック図である。

【図２】本発明の第２の実施形態のオーディオ復号装置の構成を示すブロック図である。

【図３】本発明の第３の実施形態のオーディオ復号装置の構成を示すブロック図である。

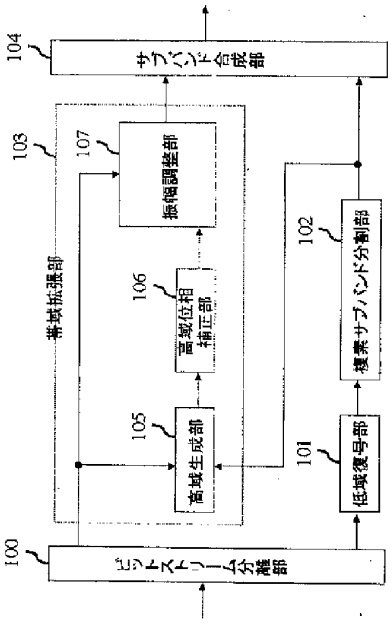
【図４】第１の従来のオーディオ復号装置の構成を示すブロック図である。

【図５】第２の従来のオーディオ復号装置の構成を示すブロック図である。

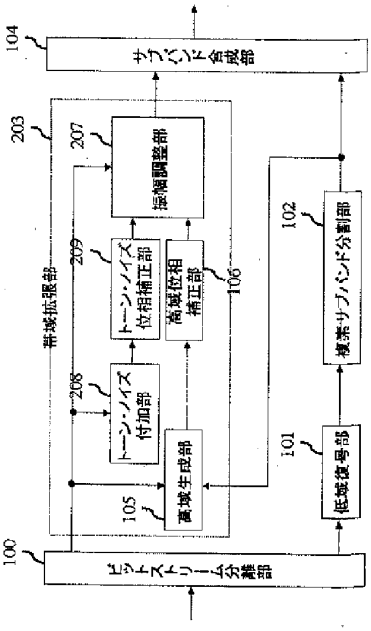
【符号の説明】

１００　ビットストリーム分離部
１０１　低域復号部
１０２、４０２　複素サブバンド分割部
１０３、２０３、３０３、４０３、５０３　帯域拡張部
１０４、４０４　サブバンド合成部
１０５　高域生成部
１０６　高域位相補正部
１０７、２０７　振幅調整部
２０８　トーン・ノイズ付加部
２０９　トーン・ノイズ位相補正部

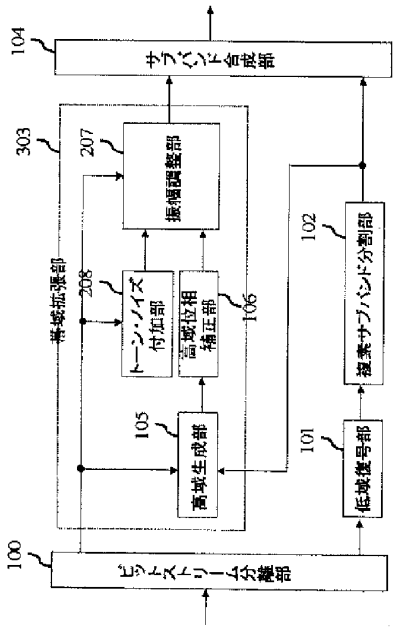
【図１】


104
サブバンド合成部
103
107
振幅調整部
帯域拡張部
106
高域位相補正部
105
高域生成部
102
複素サブバンド分割部
101
低域復号部
100
ビットストリーム分離部

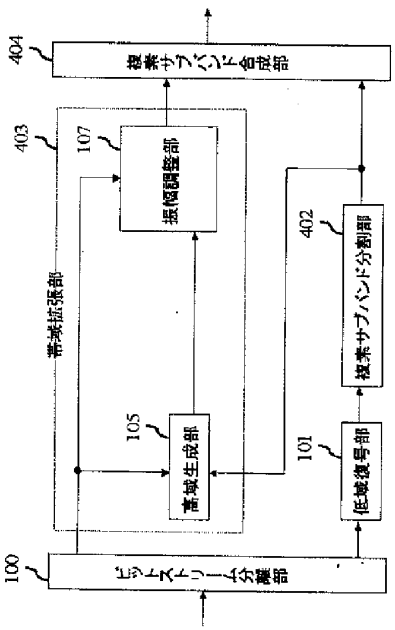
【図２】


104
サブバンド合成部
203
207
振幅調整部
帯域拡張部
209
トーン・ノイズ位相補正部
高域位相補正部
106
208
トーン・ノイズ付加部
105
高域生成部
102
複素サブバンド分割部
101
低域復号部
100
ビットストリーム分離部

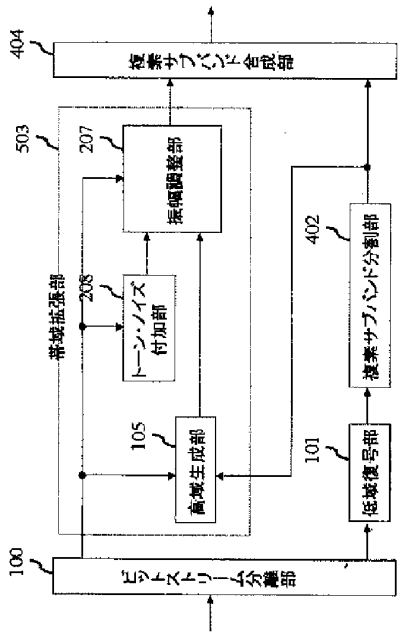
【図３】


104
サブバンド合成部
303
207
振幅調整部
帯域拡張部
208
トーン・ノイズ付加部
高域位相補正部
106
105
高域生成部
102
複素サブバンド分割部
101
低域復号部
100
ビットストリーム分離部

【図４】


404
複素サブバンド合成部
403
107
振幅調整部
帯域拡張部
105
高域生成部
402
複素サブバンド分割部
101
低域復号部
100
ビットストリーム分離部

【図５】


404
複素サブバンド合成部
503
207
振幅調整部
帯域拡張部
208
トーン・ノイズ付加部
105
高域生成部
402
複素サブバンド分割部
101
低域復号部
100
デコーダ入力部

フロントページの続き

(72)発明者 野村 俊之
東京都港区芝五丁目７番１号 日本電気株式会社内
(72)発明者 高見沢 雄一郎
東京都港区芝五丁目７番１号 日本電気株式会社内
(72)発明者 芹沢 昌宏
東京都港区芝五丁目７番１号 日本電気株式会社内
(72)発明者 田中 直也
大阪府門真市大字門真１００６番地 松下電器産業株式会社内
(72)発明者 津島 峰生
大阪府門真市大字門真１００６番地 松下電器産業株式会社内
(72)発明者 則松 武志
大阪府門真市大字門真１００６番地 松下電器産業株式会社内
Ｆターム(参考) 5J064 AA00 BA16 BB12 BC11 BC15 BC18 BD01

Espacenet

Bibliographic data: JP2005518118 (A) — 2005-06-16

FILTER SET FOR FREQUENCY ANALYSIS

Inventor(s):

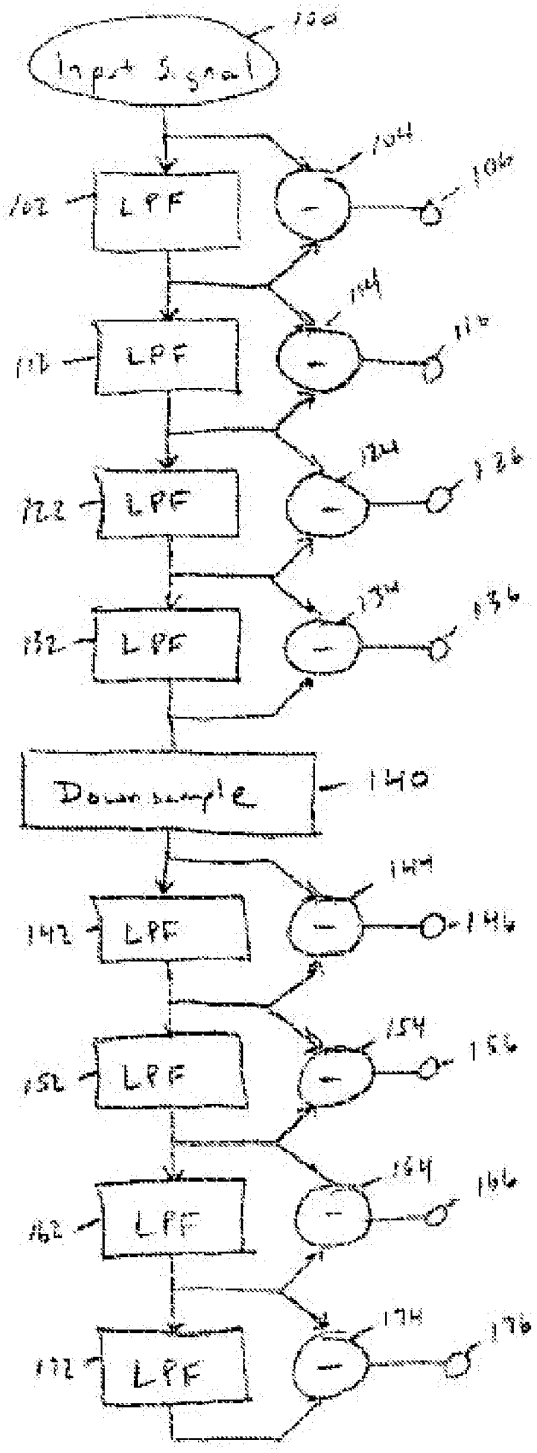
Applicant(s):

Classification:	**- international:**	***H03H17/02; H03H17/04;*** (IPC1-7): H03H17/02
	- European:	H03H17/02
Application number:	JP20030568555T 20030211	
Priority number(s):	US20020074991 20020213; WO2003US04124 20030211	
Also published as:	WO03069499 (A1) WO03069499 (A9) US2005216259 (A1) US2005228518 (A1) EP1474755 (A1) more	

Abstract not available for JP2005518118 (A)
Abstract of corresponding document: WO03069499 (A1)

A system and method are disclosed for analyzing an input signal (100) into a plurality of frequency components. In one embodiment, the input signal (100) is processed with a first set of low pass filters (102-132) to derive a first set of frequency components wherein the first set of low pass filters (102-132) are arranged serially in a chain having a first low pass filter (102) and a last low pass filter (132), the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter (132). The output of the last low pass filter (132) is downsampled (140) to produce a downsampled signal. The downsampled signal is processed with a second set of low pass filters (142-172) to derive a second set of frequency components.

Last updated: 28.02.2012 Worldwide Database 5.7.36; 93p


Input Signal
LPF
LPF
LPF
LPF
Downsample
LPF
LPF
LPF
LPF

(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2005-518118
(P2005-518118A)
(43)公表日 平成17年6月16日(2005.6.16)

(51)Int.Cl.[7]	F I		テーマコード（参考）
H03H 17/02	H03H 17/02	615K	
	H03H 17/02	613C	
	H03H 17/02	615J	

審査請求 未請求　予備審査請求 未請求　（全 18 頁）

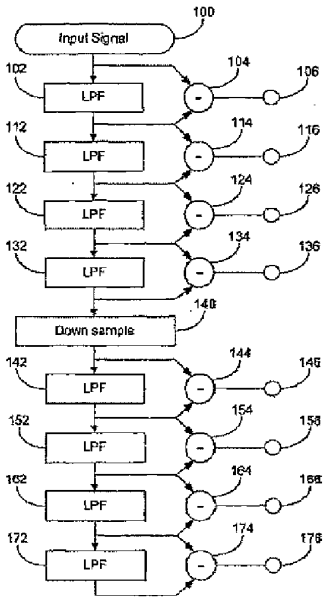
(21)出願番号	特願2003-568555 (P2003-568555)
(86)(22)出願日	平成15年2月11日 (2003.2.11)
(85)翻訳文提出日	平成16年10月8日 (2004.10.8)
(86)国際出願番号	PCT/US2003/004124
(87)国際公開番号	WO2003/069499
(87)国際公開日	平成15年8月21日 (2003.8.21)
(31)優先権主張番号	10/074,991
(32)優先日	平成14年2月13日 (2002.2.13)
(33)優先権主張国	米国(US)

(71)出願人 504308349
オーディエンス・インコーポレーテッド
AUDIENCE INCORPORATED
アメリカ合衆国 カリフォルニア州９４０２２ ロス・アルトス，エル・カミノ・リアル，４９０６，スイート２０５
(74)代理人 110000028
特許業務法人明成国際特許事務所
(72)発明者 ワッツ・ロイド
アメリカ合衆国 カリフォルニア州９４０４１ マウンテン・ビュー，チキータ・アベニュー，５００，＃２

最終頁に続く

(54)【発明の名称】周波数解析のためのフィルタセット

(57)【要約】
【課題】
【解決手段】　入力信号（１００）を解析して複数の周波数成分に分離するためのシステムおよび方法が開示されている。一実施形態では、入力信号（１００）が第１の組のローパスフィルタ（１０２～１３２）で処理されることにより、第１の組の周波数成分が取得される。なお、第１の組のローパスフィルタ（１０２～１３２）は、第１のローパスフィルタ（１０２）と最後のローパスフィルタ（１３２）とを有するチェーン内に直列に配置され、各ローパスフィルタの出力は、最後のローパスフィルタ（１３２）に至るまで、チェーン内の次のローパスフィルタに供給される。最後のローパスフィルタ（１３２）の出力がダウンサンプリング（１４０）されることにより、ダウンサンプリングされた信号が生成される。ダウンサンプリングされた信号が第２の組のローパスフィルタ（１４２～１７２）で処理されることにより、第２の組の周波数成分が取得される。



【特許請求の範囲】
【請求項１】
入力信号を解析して複数の周波数成分に分離する方法であって、
前記信号を第１の組のローパスフィルタで処理して、第１の組の周波数成分を取得する工程であって、前記第１の組のローパスフィルタは、最初のローパスフィルタと最後のローパスフィルタとを有するチェーン内に直列に配置され、各ローパスフィルタの出力は、前記最後のローパスフィルタに至るまで、前記チェーン内の次のローパスフィルタに供給される、工程と、
前記最後のローパスフィルタの出力をダウンサンプリングして、ダウンサンプリングされた信号を生成する工程と、
前記ダウンサンプリングされた信号を第２の組のローパスフィルタで処理して、第２の組の周波数成分を取得する工程と、
を備える、方法。
【請求項２】
請求項１に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記周波数成分は、各ローパスフィルタへの入力から当該ローパスフィルタの出力を減算することにより取得される、方法。
【請求項３】
請求項１に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記第２の組のローパスフィルタは、前記第１の組のローパスフィルタとは異なるＱ値を有する、方法。
【請求項４】
請求項１に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記第２の組のローパスフィルタは、前記第１の組のローパスフィルタよりも鋭度の低いＱ値を有する、方法。
【請求項５】
請求項１に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記第２の組のローパスフィルタは、実質的に人間の臨界帯域幅に従って、前記第１の組のローパスフィルタのＱ値と異なるＱ値を有する、方法。
【請求項６】
入力信号を解析して複数の周波数成分に分離する方法であって、
前記信号を第１のローパスフィルタで処理して、第１のローパスフィルタ信号を生成する工程と、
前記入力信号から前記第１のローパスフィルタ信号を減算して、第１の周波数成分を取得する工程と、
前記信号を第２のローパスフィルタで処理して、第２のローパスフィルタ信号を生成する工程と、
前記第１のローパスフィルタ信号から前記第２のローパスフィルタ信号を減算して、第２の周波数成分を取得する工程と、
を備える、方法。
【請求項７】
入力信号を解析して複数の周波数成分に分離する方法であって、
前記信号を第１のローパスフィルタで処理して、第１のローパスフィルタ信号を生成する工程と、
前記入力信号から前記第１のローパスフィルタ信号を減算して、第１の周波数成分を取得する工程と、
前記ローパスフィルタ信号を第２のローパスフィルタで処理して、第２のローパスフィルタ信号を生成する工程と、
前記第１のローパスフィルタ信号から前記第２のローパスフィルタ信号を減算して、第２の周波数成分を取得する工程と、

を備える、方法。
【請求項８】
入力信号を解析して複数の周波数成分に分離する方法であって、
前記信号の一部を第１の出力周波数チャネルに分離するよう構成されている第１のフィルタを用いて、前記信号を処理する工程と、
前記信号の一部を第２の出力周波数チャネルに分離するよう構成されている第２のフィルタを用いて、前記信号を処理する工程と、
を備え、
前記第１の周波数チャネルは、前記第２の周波数チャネルよりも高い周波数を強調し、
前記第２のフィルタは、前記第１のフィルタと異なるＱ値を有する、方法。
【請求項９】
請求項８に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記第２のフィルタは、前記第１のフィルタよりも鋭度の低いＱ値を有する、方法。
【請求項１０】
請求項８に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記第２のフィルタは、実質的に人間の臨界帯域幅に従って、前記第１のフィルタのＱ値と異なるＱ値を有する、方法。
【請求項１１】
請求項８に記載の入力信号を解析して複数の周波数成分に分離する方法であって、前記フィルタは、ローパスフィルタである、方法。
【請求項１２】
入力信号を解析して複数の周波数成分に分離するためのシステムであって、
第１の組の周波数成分を取得するよう構成された第１の組のローパスフィルタであって、前記第１の組のローパスフィルタは、最初のローパスフィルタと最後のローパスフィルタとを有するチェーン内に直列に配置され、各ローパスフィルタの出力は、前記最後のローパスフィルタに至るまで、前記チェーン内の次のローパスフィルタに供給される、第１の組のローパスフィルタと、
前記最後のローパスフィルタの出力をダウンサンプリングして、ダウンサンプリングされた信号を生成するよう構成されたダウンサンプラと、
前記ダウンサンプリングされた信号を処理して、第２の組の周波数成分を取得するよう構成された第２の組のローパスフィルタと、
を備える、システム。
【請求項１３】
請求項１２に記載の入力信号を解析して複数の周波数成分に分離するためのシステムであって、前記周波数成分は、各ローパスフィルタへの入力から当該ローパスフィルタの出力を減算することにより取得される、システム。
【請求項１４】
請求項１２に記載の入力信号を複数の周波数成分に分離して解析するためのシステムであって、前記第２の組のローパスフィルタは、前記第１の組のローパスフィルタとは異なるＱ値を有する、システム。
【請求項１５】
請求項１２に記載の入力信号を複数の周波数成分に分離して解析するためのシステムであって、前記第２の組のローパスフィルタは、前記第１の組のローパスフィルタよりも鋭度の低いＱ値を有する、システム。
【請求項１６】
請求項１２に記載の入力信号を解析して複数の周波数成分に分離するためのシステムであって、前記第２の組のローパスフィルタは、実質的に臨界帯域に従って、前記第１の組のローパスフィルタのＱ値と異なるＱ値を有する、システム。
【請求項１７】
請求項１２に記載の入力信号を解析して複数の周波数に分離するためのシステムであっ

て、前記システムは音声認識システムで用いられるものである、システム。
【請求項１８】
請求項１２に記載の入力信号を解析して複数の周波数に分離するためのシステムであって、前記システムは音声ストリーム分離に用いられるものである、システム。
【請求項１９】
請求項１２に記載の入力信号を解析して複数の周波数に分離するためのシステムであって、前記システムは音声定位に用いられるものである、システム。
【請求項２０】
入力信号を解析して複数の周波数成分に分離するためのシステムであって、
第１のローパスフィルタ信号を出力する第１のローパスフィルタと、
前記入力信号から前記第１のローパスフィルタ信号を減算して、第１の周波数成分を取得するよう構成された第１のプロセッサと、
第２のローパスフィルタ信号を出力する第２のローパスフィルタと、
前記第１のローパスフィルタ信号から前記第２のローパスフィルタ信号を減算して、第２の周波数成分を取得するよう構成された第２のプロセッサと、
を備える、システム。
【請求項２１】
入力信号を解析して複数の周波数成分に分離するためのシステムであって、
第１のローパスフィルタ信号を出力する第１のローパスフィルタと、
前記入力信号から前記第１のローパスフィルタ信号を減算して、第１の周波数成分を取得するよう構成された第１のプロセッサと、
前記ローパスフィルタ信号を処理して、第２のローパスフィルタ信号を生成するよう構成された第２のローパスフィルタと、
前記第１のローパスフィルタ信号から前記第２のローパスフィルタ信号を減算して、第２の周波数成分を取得するよう構成された第２のプロセッサと、
を備える、システム。
【請求項２２】
入力信号を解析して複数の周波数成分に分離するためのシステムであって、
前記信号を処理するよう構成された第１のフィルタであって、前記信号の一部を第１の出力周波数チャネルに分離するよう構成されている第１のフィルタと、
前記信号を処理するよう構成された第２のフィルタであって、前記信号の一部を第２の出力周波数チャネルに分離するよう構成されている第２のフィルタと、
を備え、
前記第１の周波数チャネルは、前記第２の周波数チャネルよりも高い周波数を強調し、
前記第２のフィルタは、前記第１のフィルタと異なるＱ値を有する、システム。
【請求項２３】
請求項２２に記載の入力信号を解析して複数の周波数成分に分離するためのシステムであって、前記第２のフィルタは、前記第１のフィルタよりも鋭度の低いＱ値を有する、システム。
【請求項２４】
請求項２２に記載の入力信号を解析して複数の周波数成分に分離するためのシステムであって、前記第２のフィルタは、実質的に臨界帯域に従って、前記第１のフィルタのＱ値と異なるＱ値を有する、システム。
【請求項２５】
請求項２２に記載の入力信号を解析して複数の周波数成分に分離するためのシステムであって、前記フィルタは、ローパスフィルタである、システム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、一般に、信号処理に関する。信号を解析して周波数成分に分離するためのシ

ステムおよび方法が開示されている。

【背景技術】

【０００２】

信号を解析する際に有用な工程として、信号を周波数成分に分離する方法がある。ここ最近、時間領域信号を解析して周波数成分に分離する際には、高速フーリエ変換（ＦＦＴ）アルゴリズムが用いられている。各種の処理、特に、音声信号の処理に対しては、高周波数における高い時間分解能と低周波数における高いスペクトル分解能とを実現しつつ、信号を解析して周波数成分に分離することが望ましい。これを実現するために、数多くの技術が提案されている。そのような技術としては、１組のフィルタを用いて、解析対象の信号を、異なるチャネルすなわち周波数成分に分離するシステムが挙げられる。そのようなフィルタセットは、概ね、生物学上の蝸牛と同じように動作するものであり、蝸牛は、異なる周波数チャネルに対応する一連のフィルタ出力信号を含む。

【０００３】

フィルタセットは、アナログフィルタとデジタルフィルタのいずれで実現されてもよい。従来のフィルタセットは、フィルタ設計の際の実際的な検討事項によって制限されていた。例えば、各チャネル出力を分離するための高次バンドパスフィルタを用いると、費用がかかる。ハイパスフィルタおよびローパスフィルタを用いて、様々な方法が実施されてきたが、音声認識、音声信号の音源分離、および音声信号のストリーム分離など、種々の重要な用途に対して、信号のリアルタイム処理を可能にする、より効率的な技術が必要とされている。

【発明を実施するための最良の形態】

【０００４】

以下では、本発明の好ましい実施形態を詳細に説明する。本発明は、好ましい実施形態を用いて説明されているが、どの実施形態にも限定されないことを理解されたい。逆に、本発明の範囲は、添付の請求項によってのみ限定され、本発明は、数多くの代替、変形、および等価物を包含する。本発明の完全な理解を促すために、以下の説明では、例として数多くの具体的な詳細事項が示されている。本発明は、これらの具体的な詳細事項の一部またはすべてがなくとも、請求項に従って実施することができる。明確にするために、本発明に関連する技術分野で周知の技術事項については、本発明を不必要に不明瞭にしないよう、詳細に説明していない。

【０００５】

多くの特徴を備えた周波数解析のためのフィルタカスケードが以下に開示されている。様々な実施形態において、それらの特徴は、別個あるいは組み合わせて実現されている。例えば、一部の実施形態では、各周波数成分は、ローパスフィルタへの入力からその出力を減算することにより算出される。これにより、バンドパス信号が導出される。一部の実施形態では、ローパスフィルタは、チェーン接続、すなわちカスケード接続されており、各フィルタ出力は、フィルタセット内の次のフィルタ入力に供給される。フィルタセット内の最後のフィルタの出力はダウンサンプリングされ、フィルタセット自体は、全体としては１つの高次アンチエイリアスフィルタとして機能する。次いで、より低い周波数成分からなるダウンサンプリングされたフィルタセット出力は、より効率的に処理されてよい。カスケード内のフィルタは、フィルタのＱ値が周波数と共に変化するよう設計されてよい。

【０００６】

予め参照として組み込まれた米国特許出願番号０９／５３４，６８２（以後、「６８２出願」とする）は、周波数解析のためのデジタルフィルタカスケードを開示している。カスケード内のフィルタは、チェーン接続されており、複数組のフィルタは、オクターブに分けられ、オクターブ間のダウンサンプリングを行う。フィルタのパラメータは、異なるオクターブ内の対応する複数のフィルタ間で共有される。本明細書で説明するように、フィルタのパラメータが、オクターブ間でＱ値すなわちフィルタの鋭度（シャープネス）を変更するようにオクターブ間で変更された場合に、利点を実現できる。一実施形態では、

Q値は、実質的に臨界帯域幅に従って変更される。

【０００７】

図１は、入力信号を解析して複数の周波数成分に分離するために一実施形態で用いられるフィルタネットワークを示すブロック図である。入力信号１００は、ローパスフィルタ（ＬＰＦ）１０２に供給される。ＬＰＦ１０２の出力は、減算器１０４によって入力信号１００から減算される。したがって、ノード１０６における出力は、ＬＰＦ１０２前後の信号の差分である。その出力は、入力信号の上限周波数の遮断がどのようなものであっても、ＬＰＦ１０２の遮断周波数よりも高い周波数の帯域すなわちチャネルを強調する。ＬＰＦ１０２の出力は、同様に、ＬＰＦ１１２の入力に向けられ、ＬＰＦ１１２の入力および出力の間の差分が、減算器１１４によって算出され、ノード１１６で出力される。ノード１１６の出力は、ＬＰＦ１０２およびＬＰＦ１１２の遮断周波数の間の周波数を強調する別の周波数チャネルである。同様に、ＬＰＦ１２２、ＬＰＦ１３２、および減算器１２４、１３４は、ノード１２６および１３６において他の周波数チャネルを出力する。ノードの出力は、必要に応じてさらに処理されてよい。例えば、一部の実施形態では、それらの出力は半波整流され、一部の実施形態では、ダイナミックレンジを縮小または拡大するために、出力のゲインが調整される。

【０００８】

他の実施形態では、２次以上のデジタルまたはアナログフィルタを用いてもよい。フィルタの性質は、もちろん、一般に所与の周波数帯域を強調すると共に一般的なバンドパス特性を有する各チャネル出力の正確な性質を決定する。全体として、チャネル出力は、信号の周波数成分を示す。各ＬＰＦの入出力の減算のため、各チャネル出力は、周波数の帯域すなわちスライスを示し、すべての出力の和は、入力信号全体を示す。

【０００９】

各ＬＰＦの出力は、次のＬＰＦの入力に供給され、ローパスフィルタのチェーンを形成するため、チェーン内の最後のＬＰＦの出力は、最後のフィルタの次数よりもはるかに高次なフィルタの特性を有する。この高次フィルタ効果は、チェーン内の最後のフィルタの出力がダウンサンプリングされる場合に有効に利用することができる。基本的に、周波数チャネルを分離するために用いられるローパスフィルタのチェーンは、信号がダウンサンプリングされる際にアンチエイリアスフィルタの機能を実行する高次フィルタの機能を全体として実現する。

【００１０】

そのことの例は、ダウンサンプラ１４０がＬＰＦ１３２の出力をダウンサンプリングする図１に示されている。図示のために、４つのフィルタのみをチェーン内に示していることに注意されたい。ほとんどの実施形態では、ダウンサンプリングの前に周波数範囲を処理するために、５以上のフィルタが用いられる。ダウンサンプラ１４０から出力されたダウンサンプリング済みの信号は、次に、ＬＰＦ１４２、ＬＰＦ１４２、ＬＰＦ１４２、およびＬＰＦ１４２を備える別のローパスフィルタのチェーンによって処理され、ノード１４６、１５６、１６６、および１７６において、減算器１４４、１５４、１６４、および１７４により、周波数チャネル出力が導出される。

【００１１】

一実施形態では、個々の２次のフィルタが用いられ、６０のフィルタのチェーンが、ダウンサンプリングの前に信号の１オクターブを処理する。ダウンサンプリングは、すべての他のサンプルを単に切り捨てる方法や、それ以外の任意の適切な技術によって実行されてよい。ダウンサンプリングの量は、ナイキスト基準によって決定される。必要に応じて、適切な量のオーバーサンプリングが実行されてよい。フィルタのチェーンの複合的な効果は、非常に高次なアンチエイリアスフィルタの効果と等しい。したがって、信号のダウンサンプリングを実行することにより、高価な高次アンチエイリアスフィルタを用いることなく、より低い周波数オクターブの処理を高速化することができる。

【００１２】

ローパスフィルタをチェーン接続する利点は、一部の実施形態では、周波数帯域を算出

するための減算器を用いることなく実現されることに注意されたい。各ローパスフィルタの出力は、各周波数チャネルにおけるエネルギを示すよう直接的に用いられてよい。各チェーン内の最後のフィルタの出力はダウンサンプリングされるため、フィルタチェーン自体は、１つのアンチエイリアスフィルタの機能を実行する。

【００１３】

図２は、ローパスフィルタが、それらの入出力においてチェーン接続されていない別の実施形態を示す図である。入力信号２００は、ローパスフィルタ２０２、２０４、２０６、および２０８に供給される。各ローパスフィルタの入力と出力との間の差分は、減算器２１２、２１４、２１６、および２１８によって算出される。ここでも、算出された差分は、入力信号の周波数帯域すなわちチャネルの解析を示す。しかしながら、各フィルタの出力は、次のフィルタの入力に供給されないので、上述したチェーン内の最後のフィルタの出力において見られる高次フィルタ効果は、実現されない。

【００１４】

フィルタカスケードは、アナログフィルタとデジタルフィルタのどちらを用いて実現されてもよい。一実施形態では、フィルタは、所望のチャネル分解能を実現するよう設計された遮断周波数を有するデジタルフィルタとして実装される。各フィルタは、それに関連する１組の係数（a_0、a_1、a_2、b_1、b_2）を有する。各フィルタの出力は、以下の関数に従って算出される：

【００１５】

【数１】

Equation 1. $$y_n = a_0x_n + a_1x_{n-1} + a_2x_{n-2} - b_1y_{n-1} - b_2y_{n-2}$$

【００１６】

ここで、フィルタ出力y_nは、時間ｎにおける入力データx_n、以前の入力x_{n-1}およびx_{n-2}、および以前の出力y_{n-1}およびy_{n-2}の関数である。図３は、この信号の流れを示すブロック図である。フィルタの出力y_nは、カスケード内の次のフィルタの入力x_nに提供される。

【００１７】

フィルタ応答Ｈ（ｚ）は、以下の式で与えられる：

【００１８】

【数２】

Equation 2. $$H(z) = \frac{a_0 + a_1z^{-1} + a_2z^{-2}}{1 + b_1z^{-1} + b_2z^{-2}}$$

$$\text{and } z = e^{i*(\omega/\omega_s)},\ \omega = 2\pi f,\quad \omega_s = 2\pi f_s$$

【００１９】

ここで、f_sは、サンプリング周波数である。

【００２０】

式２の伝達関数に上記の代入を行うと、フィルタ係数a_0、a_1、a_2、b_1、b_2、およびサンプリングレートf_sの関数であるフィルタ応答Ｈ（ｆ）が得られる。

【００２１】

６８２出願に記載されているように、フィルタ係数は、複数組のフィルタ間で再利用されてよく、その際、フィルタの応答は、複数組のフィルタの間のダウンサンプリングの結果として変更される。図の実施形態では、フィルタは、複数のオクターブにわたって一様に分配され、１オクターブ当たりのフィルタの数は６０となっている。６０のオブジェクトがコンピュータで作成される。各オブジェクトは、上述の１組の係数を有しており、さ

らに、１オクターブずつ離れた周波数で動作する１０のフィルタに対応して、１０組の状態変数を有する。第１の組の状態変数を用いる６０のオブジェクトは、第１のオクターブ群のフィルタに対応し、第２の組の状態変数を用いる（そして、比較的低い周波数でサンプリングを行う）６０のオブジェクトは、第２のオクターブ群のフィルタに対応し、以下も、同様に対応している。別の実施形態では、各オブジェクトは、１組の係数と、１組だけの状態変数とを備えており、単一の周波数で実行される。この場合、６００のフィルタを表現するために、６００のオブジェクトが必要となる。

【００２２】

第１のオクターブのフィルタは、最も高いオクターブ（すなわち、２０ｋＨｚから１０ｋＨｚ）の周波数に調整され、４４．１ｋＨｚでサンプリングされる。これは、ナイキストのサンプリング基準を満たしている。第２のオクターブのフィルタは、第１のオクターブ内の対応するフィルタの半分の周波数に調整され、１０ｋＨｚから５ｋＨｚの範囲となる。第２のオクターブのフィルタは、２２．０５ｋＨｚ、すなわち第１のサンプリング周波数の半分の周波数でサンプリングされる。各フィルタの係数は、メモリに格納され、フィルタのための計算で用いられる。カスケードの応答は、個々のフィルタの応答（それぞれは弱い応答であるが、合計すると、はるかに強い応答になる）の和である。フィルタの係数は、所望の応答によって決定される。

【００２３】

音声信号は、各フィルタに通されると、次のフィルタに供給される前にサンプリングおよびフィルタリングされる。図４は、フィルタの配列を示すブロック図である。最初のオクターブが終わると、信号は、上述のように第１のオクターブの半分のサンプリングレートでサンプリングを行うフィルタを備える次のオクターブの第１のフィルタに通される。連続的なオクターブが、同じ係数「２」を用いて同様にダウンサンプリングされる。この構成では、各段は、後ろの段のためのアンチエイリアスフィルタとして機能し、高周波数を十分に除去して、エイリアシングなしにダウンサンプリングを実行することを可能にする。余分なアンチエイリアスフィルタは必要にならない。

【００２４】

連続的なオクターブの各々をダウンサンプリングすると、システムの計算の複雑さが大幅に低減される。さらに、フィルタ係数に求められる精度が低くなるため、各係数を表現するために必要なビットが少なくなる。デジタルローパスフィルタは、フィルタ係数を表現するために必要な数値精度が、遮断周波数とサンプリング周波数との比で決まるという特徴を有する。所与のサンプリング周波数に対しては、低い遮断周波数を有するフィルタは、より高い遮断周波数を有するフィルタよりも精度の高い係数を必要とする。連続ダウンサンプリング技術を用いない場合には、４４ｋＨｚのサンプリングレートで遮断周波数の最も低いフィルタ（３０Ｈｚ）を表現するために、非常に精度の高いフィルタ係数（２３ビットのオーダー）が必要になる。連続ダウンサンプリング技術を用いると、多くのダウンサンプリング工程の後に、最も低いオクターブにおいてサンプリングレートが非常に低くなるため、より低い精度の係数（１２ビットのオーダー）を用いて、３０Ｈｚの遮断フィルタを表現することができる。このように精度が低減されることで、ハードウェアの複雑さが低減され（少ないメモリ、小さいレジスタ、精度の低い算術演算子）、カスタムハードウェアの実装におけるコスト全体が低減される。

【００２５】

６８２出願に記載された実施形態では、各フィルタが、１、２、またはそれ以上高いオクターブもしくは低いオクターブのフィルタとフィルタパラメータを共有することで、必要な記憶域を低減している。例えば、第１のオクターブに属する最も周波数の高いフィルタ４０は、第２のオクターブに属する最も周波数の高いフィルタ５０や、第３のオクターブに属する最も周波数の高いフィルタ６０などとフィルタ係数を共有する。第１のオクターブに属する２番目に周波数の高いフィルタ４２は、第２および第３のオクターブに属する２番目に周波数の高いフィルタ５２および６２、およびその他の対応するフィルタすべて（１、２、またはそれ以上低いオクターブの周波数に調整されたもの）と、フィルタ係

数を共有する。

【００２６】

あるいは、以下で説明するように、各オクターブ内でフィルタパラメータを変えることにより、低周波数における遅延を改善できることが確認された。多くのシステムにとって、このことは、低周波数での遅延の改善による利点が、記憶装置の必要性の増大を相殺するため、異なるオクターブに属する対応するフィルタ間でフィルタパラメータを共有するのに好ましい。

【００２７】

一実施形態では、フィルタ係数は、フィルタによって処理される周波数帯域（周波数遮断によって決定されたもの）に応じて、所望のＱ値（品質係数、すなわち、鋭度もしくは周波数選択度）を実現するよう調整される。カスケード内でフィルタ係数を再利用すると、カスケードのＱ値が一定になり、フィルタ応答はすべて、同じ形状（Ｑ値）を有することになる。この「定Ｑ値」構成は、概念的に単純であることやフィルタ係数を共有するという利点を有するが、低周波数では大きな遅延を生じる。例えば、すべての周波数で４周期の位相の蓄積を有する定Ｑ設計では、２０ｋＨｚタップにおける遅延は、２００μｓであり、２０Ｈｚタップにおける遅延は、２００ｍｓである。低周波数での動作が速いと、カスケードの応答時間を改善するのに好ましく、これは、低いオクターブにおけるフィルタのフィルタ係数を変更することにより実現できる。

【００２８】

図５は、極のみを有する２次のフィルタ応答の例を示す図である。フィルタは、時定数Ｔａｕおよび品質係数Ｑによって表現されてもよいし、前述したようにフィルタ係数ｂ１およびｂ２によって表現されてよい。Ｔａｕは、中心周波数f_cの逆数であって、ピークの位置を表すものであり、Ｑ値は、ピークの鋭さを表すものである。図５からわかるように、Ｑ値の値が高いほど、ピークは鋭くなり、f_cはピークの生じる位置を示している。フィルタの式は、以下の通りである：

【００２９】

【数３】

$$Vout(z) = \frac{1}{1 - b_1 z^{-1} - b_2 z^{-2}} Vin(z)$$

and

$$Vout(s) = \frac{1}{1 + Tau\, s/Q + Tau^2 s^2} Vin(s)$$

【００３０】

Ｔａｕ及びＱ値と、ｂ１及びｂ２との関係は、参照として本明細書に組み込まれる「Ｌｙｏｎ’ｓ　Ｃｏｃｈｌｅａｒ　Ｍｏｄｅｌ」、Ａｐｐｌｅ　Ｔｅｃｈｎｉｃａｌ　Ｒｅｐｏｒｔ　＃１３、Ｍａｌｃｏｌｍ　Ｓｌａｎｅｙ著、１９８８、に記載されている。フィルタのフィルタ係数は、中心周波数f_c＝１／Ｔａｕおよびフィルタのｑ値から決定することができる。

【００３１】

フィルタは、極だけでなくゼロ点を有するよう設計されてもよく、そのようなシステムのための式は、以下の通りである：

【００３２】

【数４】

$$Vout(s)=\frac{1+Tau_z s/Q_z+Tau_z^2 s^2}{1+Tau_p s/Q_p+Tau_p^2 s^2}*Vin(s)$$

【００３３】

図６は代表的なフィルタ応答の図であり、Ｑｐは極のＱ値を表し、Q_zはゼロ点のＱ値を、f_{cp}は極の中心周波数（f_pとも呼ぶ）を、f_{cz}はゼロ点の中心周波数（f_zとも呼ぶ）を表す。ゼロ点は、ゲインの降下を停止し、位相を０に戻す。ゼロ点が極に近いほど、この効果は早く生じる。ゼロ点が極に非常に近い場合には、位相軌跡は、ゼロ点によって戻されるまでにあまり遠ざからない（１周期の内のほんの少し）ことがある。この特性は、カスケードを通しての位相の蓄積の総量、延いては、カスケードの遅延応答を抑制する鍵となる。

【００３４】

６００個のフィルタを、極だけを有する６００の部分のカスケードで用いる場合、各フィルタは、最良の周波数で１／４周期の位相の蓄積を引き起こし、その結果、多大な遅延が発生する。一実施形態では、フィルタカスケードは、中心周波数が、カスケードを通じて幾何級数的（exponentially）に減少するよう構成される。Ｑ値は、カスケードを通じて徐々に減少し、遅延が問題とならない高周波数において鋭い応答を実現し、より早い応答の代償として、ある程度の鋭度の損失が許容される低周波数においては早い応答を実現する。Ｑ値が一定ではないこの実施例は、例えば、潜水艦のパッシブソナー、音声認識、音楽の録音、音声ストリームの分離、および音声定位において用いられる信号処理システムに対して特に有効である。この方法は、ダウンサンプリングされたフィルタカスケードに限定されず、ダウンサンプリングされないフィルタカスケードで用いられてもよいことに注意されたい。

【００３５】

Ｑ値の一定なフィルタカスケードの設計は、遮断周波数の範囲とオクターブ当たりのタップ数とを選択する工程（例えば、２０Ｈｚから２０ｋＨｚまでの周波数範囲、６００のタップおよび１０のオクターブ（６０タップ／オクターブ）を選択）を含む。これにより、各タップのf_pが決定される。カスケードを通して所望される鋭度および遅延に基づいて、Q_p、Q_z、および$f_{ratio}=f_z/f_p$に対して固定値が選択される。一実施形態では、定Ｑ設計に用いられる値は、$Q_p=7.0$、$Q_z=7.5$、$f_{ratio}=1.03$であってよい。別の実施形態では、それらの値は、$Q_p=23$、$Q_z=26$、$f_{ratio}=1.01$であってよい。

【００３６】

１０オクターブで６００のタップを用いる可変Ｑフィルタカスケードについては、一実施形態は、最も高いオクターブにおける４４．１ｋＨｚのサンプリングレートおよび２倍オーバーサンプリングと共に、以下の値を用いてよい：$Q_p=7.0$、$Q_z=7.0$、$f_{ratio}=1.03$。これらの値は、最初の３６０のタップに用いられ、次の２４０のタップについては、タップ６００（周波数の最も低いタップ）における値、すなわち$Q_p=1.6$、$Q_z=1.6$、$f_{ratio}=1.1$、に至るまで線形的に変更される。この結果、低周波数においてフィルタ応答が広く、時間応答が大幅に早い設計が実現される。

【００３７】

別の実施形態では、Q_p、Q_z、およびf_{ratio}パラメータは、適切な心理物理的臨界帯域幅およびラウドネス知覚曲線にフィルタ応答を適合させるよう選択される。臨界帯域幅は、フィルタ応答曲線の調整幅であり、その範囲内では、信号成分が、互いに相互作用することができる。臨界帯域幅曲線は、Ｒｏｓｓｉｎｇ、１９８２、「Ｔｈｅ Ｓｃｉｅｎｃｅ ｏｆ Ｓｏｕｎｄ」（Ａｄｄｉｓｏｎ－Ｗｅｓｌｅｙ、Ｒｅａｄｉｎｇ、ＭＡ）に記載されており、この開示は、参照として本明細書に組み込まれる。臨界帯域幅は、低周波数における１００Ｈｚ弱から、高周波数における２半音と３半音（１２％および１９％

）との間の周波数まで変化する。ラウドネス知覚は、異なる周波数に対するフィルタの感度を表す。例えば、２０Ｈｚにおける可聴閾値は、約６５ｄＢであり、１ｋＨｚにおける閾値よりも高い。

【００３８】

可変Ｑフィルタカスケードの一実施形態は、以下のパラメータを用いる：

【００３９】

【数５】

$$\text{Tap 0: } Q_p=7.0,\ Q_z=7.0,\ f_{ratio}=1.03$$
$$\text{Tap 300: } Q_p=11.0,\ Q_z=11.0,\ f_{ratio}=1.03$$
$$\text{Tap 360: } Q_p=9.0,\ Q_z=11.0,\ f_{ratio}=1.03$$
$$\text{Tap 600: } Q_p=1.6,\ Q_z=1.6,\ f_{ratio}=1.01$$

【００４０】

ただし、値の指定されているタップの間のパラメータについては、線形補間される。このように、パラメータが区分的に線形変化することにより、心理物理的な臨界帯域幅およびラウドネス知覚曲線との良好な適合が実現される。図７は、臨界帯域に従って設計されたフィルタに対するフィルタ応答を示す図である。フィルタ応答は、中周波数において鋭く、低周波数において広く、臨界帯域幅曲線に対応する。フィルタは、中周波数において比較的感度が高く、低周波数においては約６５ｄＢだけ感度が低く、ラウドネス知覚パラメータに適合している。

【００４１】

図８は、臨界帯域に従って設計されたフィルタに対する位相特性を示す図である。フィルタの位相特性は、中から高周波数において約２周期の位相の蓄積が存在するが、低周波数においてはあまり存在しないというものである。その結果、より早い応答が必要とされる低周波数において、応答が早くなる。

【００４２】

信号を周波数成分に分離して解析するためのフィルタカスケードを説明した。様々な実施形態において、フィルタカスケードは、異なる技術を用いて、高周波数における時間分解能と低周波数におけるスペクトル分解能とを改善する。したがって、開示されたフィルタカスケードの各実施形態は、音声認識システムの構成要素として特に有用である。さらに、フィルタカスケードは、音声ストリームの分離や音声定位に有用である。

【００４３】

図９Ａは、本明細書に記載されたフィルタセットが音声認識システムで用いられる様子を示す図である。フィルタセット９０２に音声信号が入力され、特徴抽出部９０４によってフィルタセットの出力が解析される。特徴は、音素データベース９０８に含まれる音素と特徴とを一致させる音素分類部９０６によって分類される。単語データベース９１０にアクセスする単語検索ブロック９０９により、音素に基づいて単語が導き出される。

【００４４】

図９Ｂは、本明細書に記載されたフィルタセットが、音声ストリーム分離システムで用いられる様子を示す図である。音声ストリーム分離システムは、例えば、Ｌｌｏｙｄ Ｗａｔｔｓにより２００１年６月２１日に提出された米国特許出願Ｎｏ．６０／３００，０１２（代理人整理番号ＡＮＳＣＰ００３＋）「ＲＯＢＵＳＴ ＨＥＡＲＩＮＧ ＳＹＳＴＥＭＳ ＦＯＲ ＩＮＴＥＬＬＩＧＥＮＴ ＭＡＣＨＩＮＥＳ」に記載されているようなシステムであり、その出願は、参照として本明細書に組み込まれる。フィルタセット９１２に音声信号が入力され、特徴を抽出する１組の特徴抽出部９１４によってフィルタセットの出力が解析される。特徴は、特徴分類プロセッサ９１６によって、関連した音声信号の別個のストリームに分類される。

【００４５】

図９Ｃは、本明細書に記載されたフィルタセットが、空間相関器または音声定位システムで用いられる様子を示す図である。そのシステムは、例えば、Ｌｌｏｙｄ Ｗａｔｔｓ

により２００１年１１月１４日に提出された米国特許出願Ｎｏ．１０／００４，１４１（代理人整理番号ＡＮＳＣＰ００５）「ＣＯＭＰＵＴＡＴＩＯＮ　ＯＦ　ＭＵＬＴＩ－ＳＥＮＳＯＲ　ＴＩＭＥ　ＤＥＬＡＹＳ」に記載されているようなシステムであり、その出願は、参照として本明細書に組み込まれる。右チャネル用フィルタセット９２２に右チャネル音声信号が入力され、左チャネル用フィルタセット９２４に左チャネル音声信号が入力される。フィルタセットの出力は、バイノーラルプロセッサ９２６によって相関され、左右のチャネル入力信号の間の時間遅延が決定される。音声の発生源の方向は、時間遅延から決定できる。

【００４６】

理解を深めるために、ある程度詳細に本発明について説明したが、添付の特許請求の範囲内で、ある程度の変更や変形を行ってもよいことは明らかである。本発明のプロセスおよび装置を実施する多くのほかの方法が存在することに注意されたい。したがって、上述の実施形態は、例示のためのものであって限定を意図しないものと見なされ、本発明は、本明細書に記載された詳細事項に限定されず、添付の特許請求の範囲と均等物との範囲内で変形されてもよい。

【図面の簡単な説明】

【００４７】

【図１】入力信号を複数の周波数成分に分離して解析するために一実施形態で用いられるフィルタネットワークを示すブロック図。

【図２】ローパスフィルタが、それらの入出力においてチェーン接続されていない別の実施形態を示す図。

【図３】フィルタ方程式の信号フローのグラフ。

【図４】フィルタの配列を示すブロック図。

【図５】極のみを有する２次部分のフィルタ応答の例を示す図。

【図６】Q_pが極のＱ値、Q_zがゼロ点のＱ値、f_{cp}が極の中心周波数（f_pとも呼ぶ）、f_{cz}がゼロ点の中心周波数（f_zとも呼ぶ）を表す代表的なフィルタ応答の図。

【図７】臨界帯域に従って設計されたフィルタに対するフィルタ応答を示す図。

【図８】臨界帯域に従って設計されたフィルタに対する位相特性を示す図。

【図９Ａ】本明細書に記載されたフィルタセットが音声認識システムで用いられる様子を示す図。

【図９Ｂ】本明細書に記載されたフィルタセットが音声ストリーム分離システムで用いられる様子を示す図。

【図９Ｃ】本明細書に記載されたフィルタセットが空間相関器または音声定位システムで用いられる様子を示す図。

【符号の説明】

【００４８】

４０…フィルタ
４２…フィルタ
５０…フィルタ
５２…フィルタ
６０…フィルタ
６２…フィルタ
１００…入力信号
１０２…ローパスフィルタ
１０４…減算器
１０６…ノード
１１２…ローパスフィルタ
１１４…減算器
１１６…ノード
１２２…ローパスフィルタ

１２４…減算器
１２６…ノード
１３２…ローパスフィルタ
１３４…減算器
１３６…ノード
１４０…ダウンサンプラ
１４２…ローパスフィルタ
１４４…減算器
１４６…ノード
１５２…ローパスフィルタ
１５４…減算器
１５６…ノード
１６２…ローパスフィルタ
１６４…減算器
１６６…ノード
１７２…ローパスフィルタ
１７４…減算器
１７６…ノード
２００…入力信号
２０２…ローパスフィルタ
２０４…ローパスフィルタ
２０６…ローパスフィルタ
２０８…ローパスフィルタ
２１２…減算器
２１４…減算器
２１６…減算器
２１８…減算器
９０２…フィルタセット
９０４…特徴抽出部
９０６…音素分類部
９０８…音素データベース
９０９…単語検索ブロック
９１０…単語データベース
９１２…フィルタセット
９１４…特徴抽出部
９１６…特徴分類プロセッサ
９２２…右チャネル用フィルタセット
９２４…左チャネル用フィルタセット
９２６…バイノーラルプロセッサ

【図１】

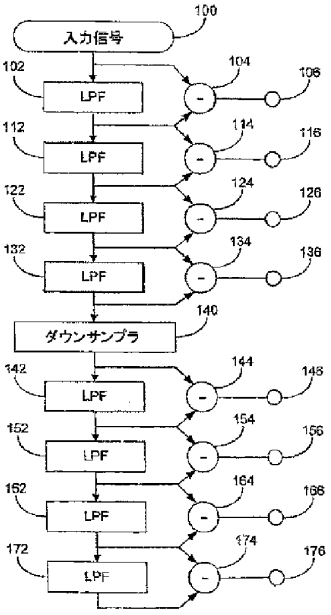
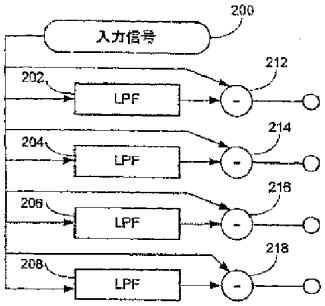


Figure 1

【図２】



【図３】

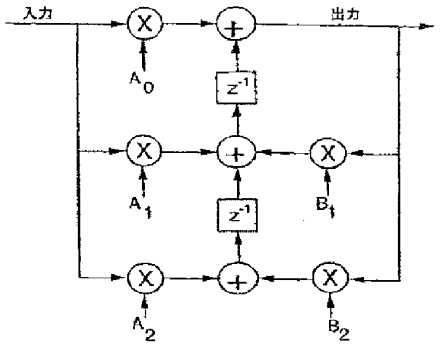


Figure 3

【図４】

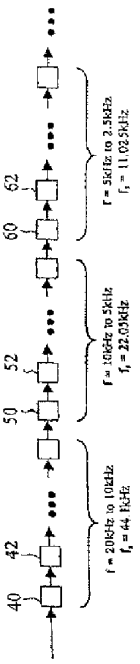


Figure 4

【図５】

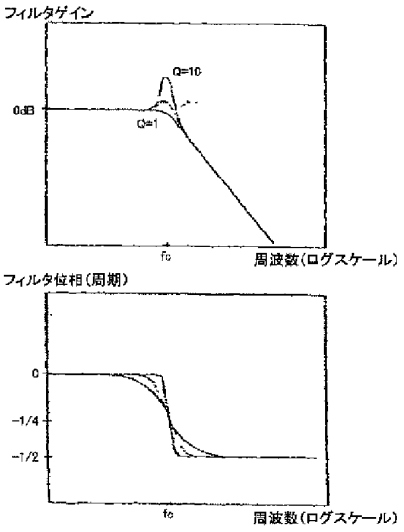


Figure 5

【図６】

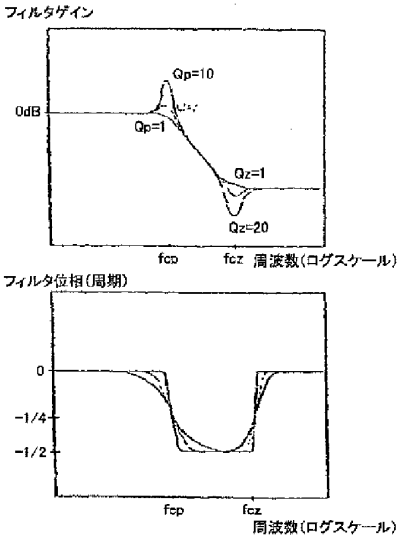


Figure 6

【図７】

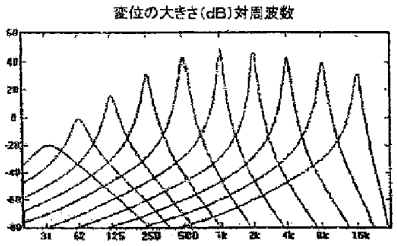


Figure 7

【図８】

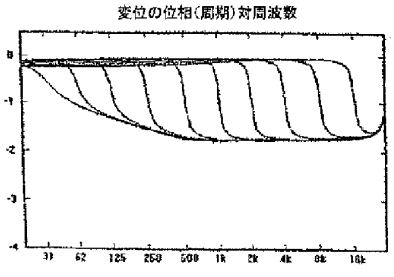


Figure 8

【図９Ａ】

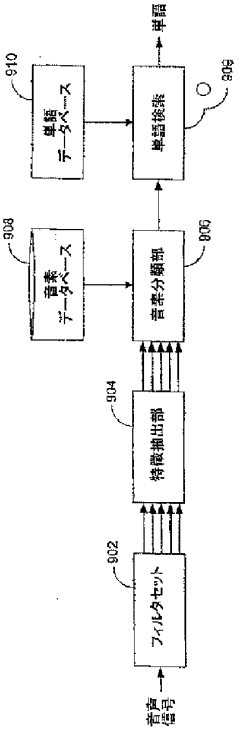


Figure 9A

【図９Ｂ】

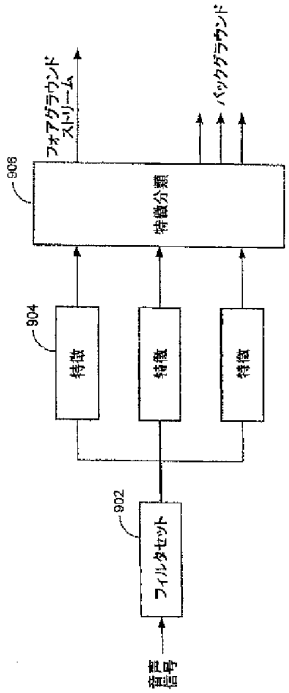


Figure 9B

【図９Ｃ】

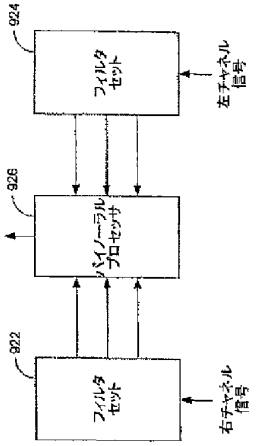


Figure 9C

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/04124

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G06F 17/00; H03G 5/00; G10L 19/00
US CL : 700/94; 381/98; 704/500
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 700/94; 381/98; 704/500

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	US 2002/0147595 A1 (BAUMGARTE) 10 October 2002 (10.10.2002), figures1, 2 and 7, page 3 and 5.	1,3-5, 8-12, 14-19, 22-25
A	US 6,513,004 B1 (RIGAZIO et al) 28 January 2003 (28.01.2003), entire document.	1-25
A	US 4,536,844 A (LYON) 20 August 1985 (20.08.1985), entire document.	1-25
A	US 5,473,759 A (SLANEY et al) 05 December 1995 (05.12.1995), entire document.	1-25
A	KATES, JAMES M. "A Time-Domain Digital Cochlear Model" IEEE Transactions of Signal Processing, December 1991, Volume 39 Number 12, pages 2573-2592.	1-25

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 May 2003 (05.05.2003)	29 MAY 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Forester W Isen Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HU,IE,IT,LU,MC,NL,PT,SE,SI,SK,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,PL,PT,RO,RU,SD,SE,SG,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZM,ZW

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 August 2003 (21.08.2003)

PCT

(10) International Publication Number
WO 03/069499 A1

(51) International Patent Classification[7]: G06F 17/00, H03G 5/00, G10L 19/00

(21) International Application Number: PCT/US03/04124

(22) International Filing Date: 11 February 2003 (11.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/074,991 13 February 2002 (13.02.2002) US

(71) Applicant: AUDIENCE, INC. [US/US]; 4906 El Camino Real, Suite 205, Los Altos, CA 94022 (US).

(72) Inventor: WATTS, Lloyd; 500 Chiquita Avenue #2, Mountain View, CA 94041 (US).

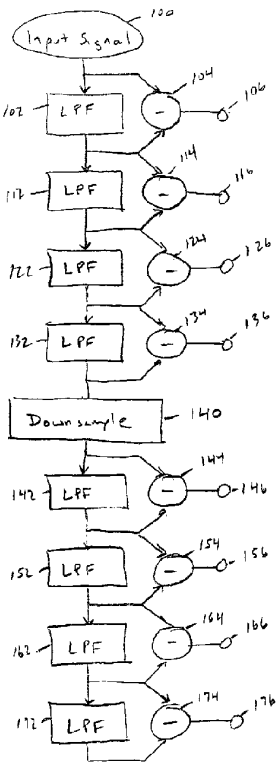
(74) Agent: VAN PELT, Lee; Van Pelt & Yi LLP, 10050 N. Foothill Blvd., Suite 200, Cupertino, CA 95014 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: FILTER SET FOR FREQUENCY ANALYSIS



(57) Abstract: A system and method are disclosed for analyzing an input signal (100) into a plurality of frequency components. In one embodiment, the input signal (100) is processed with a first set of low pass filters (102-132) to derive a first set of frequency components wherein the first set of low pass filters (102-132) are arranged serially in a chain having a first low pass filter (102) and a last low pass filter (132), the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter (132). The output of the last low pass filter (132) is downsampled (140) to produce a downsampled signal. The downsampled signal is processed with a second set of low pass filters (142-172) to derive a second set of frequency components.

WO 03/069499 A1

WO 03/069499 A1

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

FILTER SET FOR FREQUENCY ANALYSIS

Cross reference to related applications

This application claims priority to co-pending U.S. Patent Application No. 09/534,682 (Attorney Docket No. ANSCP001) entitled EFFICIENT COMPUTATION OF LOG-FREQUENCY-SCALE DIGITAL FILTER CASCADE filed March 24, 2000, which is incorporated herein by reference for all purposes.

Field of the Invention

The present invention relates generally to signal processing. A system and method for analyzing a signal into frequency components is disclosed.

Background Of The Invention

A useful step in analyzing a signal is the separation of the signal into frequency components. For some time, the fast Fourier transform or FFT algorithm has been used to analyze a time domain signal into its frequency components. For various types of processing, and in particular for processing audio signals, it would be desirable to analyze a signal into its frequency components with improved temporal resolution at high frequencies and better spectral resolution at low frequencies. Numerous techniques have been proposed for accomplishing this. Included among such techniques are systems that use a set of filters to separate the signal being analyzed into different channels or frequency components. Such filter sets operate roughly in a manner that is analogous to a biological cochlea, which includes a series of filtered output signals that correspond to different frequency channels.

Filter sets may be implemented with analog or digital filters. Previous instantiations of filter sets have been limited by practical considerations in designing filters. For example, high order bandpass filters to separate each channel output are expensive to implement. Various approaches have been implemented using combinations of high pass and low pass filters; however, more efficient techniques are needed to allow real time processing of signals for various important applications including speech recognition, source separation of audio signals and stream separation of audio signals.

Brief Description Of The Drawings

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals designate like structural elements, and in which:

Figure 1 is a block diagram illustrating a filter network used in one embodiment for analyzing an input signal into a plurality of frequency components.

Figure 2 is a diagram illustrating an alternative embodiment wherein the low pass filters are not chained together at their inputs and outputs.

Figure 3 is a signal flow graph of a filter equation.

Figure 4 is a block diagram illustrating the arrangement of the filters.

Figure 5 is a diagram illustrating an example of the filter response of a second-order section with poles only.

Figure 6 is a diagram illustrating a typical filter response where Q_p is the Q of the pole, Q_z is the Q of the zero, f_{cp} is the center frequency of the pole (also referred to as f_p), and f_{cz} is the center frequency of the zero (also referred to as f_z).

Figure 7 is a diagram illustrating filter responses for filters designed according to the critical band.

Figure 8 is a diagram illustrating the phase characteristics for filters designed according to the critical band.

Figure 9A is a diagram illustrating how a filter set as described herein is used in a voice recognition system.

Figure 9B is a diagram illustrating how a filter set as described herein is used in an audio stream separation system.

Figure 9C is a diagram illustrating how a filter set as described herein is used in a spatial correlator or sound localization system.

Detailed Description

A detailed description of a preferred embodiment of the invention is provided below. While the invention is described in conjunction with that preferred embodiment, it should be understood that the invention is not limited to any one embodiment. On the contrary, the scope of the invention is limited only by the appended claims and the invention encompasses numerous alternatives, modifications and equivalents. For the purpose of example, numerous specific details are set forth in the following description in order to provide a thorough understanding of the present invention. The present invention may be practiced according to the claims without some or all of these specific details. For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the present invention is not unnecessarily obscured.

A filter cascade for frequency analysis is disclosed that includes a number of features. In various embodiments, the features are implemented either separately or together. For example, in some embodiments, each frequency component is computed by subtracting the output of a low pass filter from the input to the filter. In this manner a bandpass signal is derived. In some embodiments, low pass filters are chained or cascaded with each filter output being fed to the next filter input in a filter set. The output of the last filter in the set is downsampled, with the filter set itself collectively acting as a high order antialiasing filter. The downsampled filter set output comprised of lower frequency components may then be more efficiently processed. Filters in the cascade may be designed so that the Q of the filters varies with frequency.

U.S. Patent Application No. 09/534,682 which was previously incorporated by reference (hereinafter, "the 682 application") discloses a digital filter cascade for frequency analysis. The filters in the cascade are chained together and sets of filters are separated into octaves with downsampling between octaves. Filter parameters are shared among corresponding filters in different octaves. As described herein, advantages may be realized if filter parameters are varied among octaves in a manner that varies the Q, or sharpness of the filters among octaves. In one embodiment, the Q is varied substantially according to critical bandwidth.

Figure 1 is a block diagram illustrating a filter network used in one embodiment for analyzing an input signal into a plurality of frequency components. An input signal 100 is fed to a low pass filter (LPF) 102. The output of LPF 102 is subtracted from input signal 100 by a subtractor 104. The output at node 106 thus represents the difference between the signal before and after LPF 102. It emphasizes a band or channel of frequencies above the cutoff frequency of LPF 102 and whatever

the upper frequency cutoff of the input signal happens to be. The output of LPF 102 is similarly directed to the input of LPF 112 and the difference between the input and the output of LFP 112 is computed by a subtractor 114 and output at node 116. The output at node 116 represents another frequency channel that emphasizes frequencies between the cutoff frequencies of LPF 102 and LPF 112. In a similar manner, LPF 122 and LPF 132 and subtractor 124 and 134 output other frequency channels at nodes 126 and 136. The output of the nodes may be further processed as is appropriate. For example, in some embodiments, the outputs are half wave rectified and in some embodiments, the gain of the outputs is adjusted to compress or expand the dynamic range.

In different embodiments, second order or higher digital or analog filters may be used. The nature of the filters, of course determines the exact nature of each channel output that generally emphasizes a given frequency band and thus has a general bandpass character. Collectively, the channel outputs represent the frequency components of the signal. Because of the subtraction of each LPF input and output, each channel output represents a band or slice of frequencies and the sum of all the outputs represents the entire input signal.

Because the output of each LPF is fed to the input of the next LPF, forming a chain of low pass filters, the output of the last LPF in the chain has characteristics of a much higher order filter than the order of the last filter. This higher order filtering effect may be exploited when the output of the last filter in the chain is downsampled. Essentially, the chain of low pass filters used to separate out frequency channels collectively act as a high order filter that performs the function of an anti aliasing filter when the signal is downsampled.

An example of this is depicted in Figure 1 where downsampler 140 downsamples the output from LPF 132. It should be noted that only four filters are shown in the chain for the purpose of illustration. In most embodiments, more than four filters would be used to process a frequency range before downsampling. The downsampled signal output from downsampler 140 is then processed by another chain of low pass filters that includes LPF 142, LPF 142, LPF 142, LPF 142 and frequency channel outputs are derived by subtractors 144, 154, 164 and 174 at nodes 146, 156, 166 and 176.

In one embodiment, second order individual filters are used and a chain of 60 filters process one octave of the signal before downsampling. Downsampling may be implemented by simply discarding every other sample or any other appropriate technique. The amount of downsampling is determined by the Nyquist criterion. A suitable amount of oversampling may be done as desired. The combined effect of the chain of filters is that of a very high order anti aliasing filter. Thus, downsampling the signal may be done to speed the processing of lower frequency octaves without requiring an expensive high order anti aliasing filter.

It should be noted that the benefit of chaining the low pass filters is realized in certain embodiments without implementing the subtractors to calculate the frequency bands. The output of each low pass filter may be used directly to represent the energy in each frequency channel. The output of the last filter in each chain is downsampled with the filter chain itself performing the function of an antialiasing filter.

Figure 2 is a diagram illustrating an alternative embodiment wherein the low pass filters are not chained together at their inputs and outputs. Input signal 200 is fed into low pass filters 202, 204, 206, and 208. The difference between the input and the output of each low pass filter is calculated by subtractors 212, 214, 216 and 218.

Again, the differences calculated represent an analysis of the frequency bands or channels of the input signal. However, because the output of each filter is not fed to the input of the next filter, the higher order filter effect in the output of the last filter in the chain described above is not realized.

The filter cascade may be implemented using either analog or digital filters. In one embodiment, the filters are implemented as digital filters with cutoff frequencies designed to produce the desired channel resolution. Each filter has a set of coefficients (a_0, a_1, a_2, b_1, b_2) associated with it. The output of each filter is calculated according to the following function:

Equation 1. $$y_n = a_0x_n + a_1x_{n-1} + a_2x_{n-2} - b_1y_{n-1} - b_2y_{n-2}$$

where the filter output y_n is a function of the input data x_n at time n, previous inputs x_{n-1} and x_{n-2}, and previous outputs y_{n-1} and y_{n-2}. Figure 3 is a block diagram illustrating this signal flow. The output of the filter y_n is passed to the input x_n of the next filter in the cascade.

The filter response $H(z)$ is given by the following:

Equation 2. $$H(z) = \frac{a_0 + a_1z^{-1} + a_2z^{-2}}{1 + b_1z^{-1} + b_2z^{-2}}$$

and $z = e^{i*(\omega/\omega s)}$, $\omega = 2\pi f$, $\omega_s = 2\pi f_s$

where f_s is the sampling frequency.

Substitution of the above into the transfer function of Equation 2 produces a filter response $H(f)$, which is a function of the filter coefficients a_0, a_1, a_2, b_1, b_2 and the sampling rate f_s.

As described in the 682 application, the filter coefficients may be reused between sets of filters with the response of the filters being altered as a result of downsampling between the sets of filters. In the embodiment shown, the filters are evenly distributed over the octaves, resulting in 60 filters per octave. 60 objects are created in a computer. Each object has a set of coefficients as described above, and additionally has ten sets of state variables, corresponding to ten filters running at frequencies that are whole octaves apart. The 60 objects using their first sets of state variables correspond to the first octave group of filters, while the 60 objects using their second sets of state variables (and sampling at a lower frequency) correspond to the second octave group of filters, and so on. In another embodiment, each object contains a set of coefficients, but only one set of state variables, and is run at a single frequency. In this case, 600 objects are required to represent 600 filters.

The filters in the first octave are tuned to the frequencies in the highest octave, 20 kHz to 10 kHz, and are sampled at 44.1 kHz, which satisfies the Nyquist sampling criterion. The filters in the second octave are tuned to half of the frequencies of the corresponding filters in the first octave, and range from 10 kHz to 5 kHz. These filters in the second octave are sampled at 22.05 kHz, half of the first sampling frequency. Coefficients for each filter are stored in memory and applied in the computations for the filters. The cascade response is the sum of responses of individual filters (which are all weak responses by themselves, but when summed, produce a much stronger response). The coefficients of the filters are determined by the desired response.

As the audio signal is passed through each filter, the signal is sampled and filtered before being passed to the next filter. Figure 4 is a block diagram illustrating the arrangement of the filters. At the end of the first octave, the signal is passed into

the first filter in the next octave, which comprises filters sampling at half the sampling rate of the first octave, as stated above. Successive octaves are downsampled in a similar manner, using the same factor of two. In this configuration, each stage acts as an anti-aliasing filter for later stages, removing the high frequencies sufficiently to allow downsampling without aliasing. No extra anti-aliasing filters are required.

Downsampling each successive octave significantly decreases the computational complexity of the system. In addition, the required precision for filter coefficients is lower, and thus, fewer bits are required to represent each coefficient. Digital low-pass filters have the property that the numerical precision required to represent the filter coefficients depends on the ratio between the cutoff frequency and the sampling

frequency. For a given sampling frequency, a filter with a low cutoff frequency will require higher-precision coefficients than a filter with a higher cutoff frequency. Without the successive downsampling technique, very high-precision filter coefficients (on the order of 23 bits) are required to represent the lowest-cutoff-frequency filters (30 Hz) at the 44 kHz sampling rate. With the successive downsampling technique, lower-precision coefficients (on the order of 12 bits) can be used to represent the 30-Hz cutoff filters, since the sampling rate is much lower in the lowest octave after many downsampling steps. This reduced precision results in lower hardware complexity (less memory, smaller registers, lower-precision arithmetic operators) and thus lower overall cost in a custom hardware implementation.

In the embodiment described in the 682 application, each filter shares filter parameters with filters that are one, two, or more octaves higher or lower, resulting in reduced storage requirements. For example, the highest frequency filter 40 in the first

octave shares filter coefficients with the highest frequency filter 50 in the second octave, the highest frequency filter 60 in the third octave, and so on. The second-highest frequency filter 42 in the first octave shares filter coefficients with the second-highest frequency filters 52 and 62 in the second and third octaves, and with all other corresponding filters (tuned to frequencies that are one, two, or more octaves lower).

Alternatively, it has been determined that the delay at low frequencies can be improved by changing the filter parameters within each octave as described below. For many systems, this is preferable to sharing filter parameters between corresponding filters in different octaves because the benefit from improved delay at low frequencies offsets increased memory storage requirements.

In one embodiment, filter coefficients are tuned to produce a desired Q (quality factor, or degree of sharpness or frequency selectivity) depending on the frequency band (determined by the frequency cutoff) being processed by the filter. Reusing filter coefficients in the cascade results in a cascade with constant Q, and all the filter responses will have the same shape (Q). This "constant-Q" configuration has the advantages of conceptual simplicity and shared filter coefficients, but has significant delays at low frequencies. For example, for a constant-Q design with a phase accumulation of four cycles at all frequencies, the delay at the 20 kHz tap will be 200 μs, while the delay at the 20 Hz tap will be 200 ms. Faster performance at low frequencies is desirable to improve the response time of the cascade, which may be accomplished by changing the filter coefficients of the filters in lower octaves.

Figure 5 is a diagram illustrating an example of the filter response of a second-order section with poles only. The filter may be described in terms of the time constant Tau and quality factor Q, or in terms of filter coefficients $\mathbf{b}_1$ and $\mathbf{b}_2$

mentioned previously. Tau is the inverse of the center frequency $\mathbf{f_c}$ and describes where the peak is, while Q describes how sharp the peak is. As can be seen from Figure 5, a higher Q results in a sharper peak, while $\mathbf{f_c}$ indicates where the peak occurs. The equations for the filter are as follows:

$$Vout(z) = \frac{1}{1 - b_1 z^{-1} - b_2 z^{-2}} Vin(z)$$

and

$$Vout(s) = \frac{1}{1 + Tau\ s/Q + Tau^2 s^2} Vin(s)$$

where the relationship between Tau, Q and $\mathbf{b_1}$, $\mathbf{b_2}$ are given in the "Lyon's Cochlear Model" Apple Technical Report #13 by Malcolm Slaney ©1988 which is herein incorporated by reference. The filter coefficients for the filter can be determined from the center frequency f_c=1/Tau, and the Q of the filter.

The filters may be designed to have zeros as well as poles, and the equation for such a system is given by

$$Vout(s) = \frac{1 + Tau_z\ s/Q_z + Tau_z^2 s^2}{1 + Tau_p\ s/Q_p + Tau_p^2 s^2} * Vin(s)$$

Figure 6 is a diagram illustrating a typical filter response where Q_p is the Q of the pole, Q_z is the Q of the zero, f_{cp} is the center frequency of the pole (also referred to as f_p), and f_{cz} is the center frequency of the zero (also referred to as f_z). The zeros arrest the dropping gain, and reverse the phase back up to zero. The closer the zero is to the pole, the sooner these effects occur. If the zero is very close to the pole, the phase trajectory may not get very far (a small fraction of a cycle) before the zero

reverses it. This property is the key to controlling the total amount of phase accumulation through the cascade, and hence the delay response of the cascade.

If 600 filters are used, and implemented with a cascade of 600 poles-only sections, each one would contribute a quarter-cycle of phase accumulation at its best frequency, resulting in a large amount of delay. In one embodiment, the filter cascade is configured so that the center frequencies decrease exponentially through the cascade. The Q's decrease gradually through the cascade, to give sharp responses at high frequencies, where delay is not an issue, and to give fast responses at low frequencies, where some loss of sharpness is acceptable in return for faster response. This implementation of nonconstant Q filters is particularly useful for signal processing systems used, for example in submarine passive sonar, speech recognition, music transcription, audio stream separation and sound localization. It should be noted that this approach is not limited to downsampled filter cascades, and may be used with filter cascades with no downsampling.

Design of a filter cascade with constant-Q involves choosing the range of cutoff frequencies and the number of taps per octave, such as a frequency range of 20 Hz to 20 kHz, 600 taps, 10 octaves (60 taps/octave). This determines f_p for each tap. Fixed values are chosen for Q_p, Q_z, and $f_{ratio}=f_z/f_p$, based on the sharpness and delay desired through the cascade. In one embodiment, values used for a constant-Q design may be $Q_p = 7.0$, $Q_z = 7.5$, and $f_{ratio} = 1.03$. In another embodiment, the values may be $Q_p = 23$, $Q_z = 26$, and $f_{ratio} = 1.01$.

For a variable-Q filter cascade using 600 taps in 10 octaves, one embodiment may employ the following values: $Q_p = 7.0$, $Q_z = 7.0$, and $f_{ratio} = 1.03$, with a sampling

rate of 44.1 kHz and 2x oversampling in the highest octave. These values are used for the first 360 taps, and then varied linearly over the next 240 taps to $Q_p = 1.6$, $Q_z = 1.6$, and $f_{ratio} = 1.1$ at tap 600 (the lowest frequency tap). This results in a design with broader filter responses at low frequencies, but much faster time response.

In another embodiment, the Q_p, Q_z, and f_{ratio} parameters are selected to match the filter responses to appropriate psychophysical critical bandwidth and loudness perception curves. Critical bandwidth is the tuning width of the filter response curves, within which signal components can interact with each other. Critical bandwidth curves are given in Rossing, 1982, "The Science of Sound" (Addison-Wesley, Reading, MA), the disclosure of which is hereby incorporated by reference. The critical bandwidth varies from a little less than 100 Hz at low frequencies to between two and three musical semitones (12% to 19%) at high frequencies. Loudness perception describes how sensitive the filters are to different frequencies. For example, the threshold of audibility at 20 Hz is about 65 dB higher than at 1 kHz.

One embodiment of a variable-Q filter cascade uses the following parameters:

Tap 0: $Q_p = 7.0$, $Q_z = 7.0$, $f_{ratio} = 1.03$
Tap 300: $Q_p = 11.0$, $Q_z = 11.0$, $f_{ratio} = 1.03$
Tap 360: $Q_p = 9.0$, $Q_z = 11.0$, $f_{ratio} = 1.03$
Tap 600: $Q_p = 1.6$, $Q_z = 1.6$, $f_{ratio} = 1.01$

with linear interpolation of parameters between the specified taps. This piecewise linear variation of the parameters gives a good fit to the psychophysical critical bandwidth and loudness perception curves. Figure 7 is a diagram illustrating filter responses for filters designed according to the critical band. The filter responses are sharp at mid-range frequencies, and very broad at low frequencies, corresponding to the critical bandwidth curve. The filters are more sensitive at mid-range

frequencies, and about 65 dB less sensitive at low frequencies, so as to match the loudness perception parameters.

Figure 8 is a diagram illustrating the phase characteristics for filters designed according to the critical band. The phase characteristics of the filters are such that there are about two cycles of phase accumulation at mid-to-high frequencies, but much less at low frequencies. This results in a faster response at low frequencies, where it is needed.

A filter cascade for analyzing a signal into frequency components has been described. In various embodiments, the filter cascade utilizes different techniques to improve temporal resolution at high frequencies and spectral resolution at low frequencies. As a result, each of the disclosed filter cascade embodiments are particularly useful as a component of a voice recognition system. In addition, the filter cascade is useful for audio stream separation and sound localization.

Figure 9A is a diagram illustrating how a filter set as described herein is used in a voice recognition system. An audio signal is input to a filter set 902 and the output of the filter set is analyzed by a feature extractor 904. The features are classified by a phoneme classifier 906 that matches features with phonemes included in a phoneme database 908. Words are derived based on the phonemes by a word search block 909 that access a word database 910.

Figure 9B is a diagram illustrating how a filter set as described herein is used in an audio stream separation system such as is described in United States Patent Application No. 60/300,012 (Attorney Docket No.ANSCP003+) by Lloyd Watts (filed June 21, 2001,) entitled: ROBUST HEARING SYSTEMS FOR INTELLIGENT MACHINES which is herein incorporated by reference. An audio signal is input to a filter set 912 and the output of the filter set is analyzed by a set of

feature extractors 914 that extract features. The features are grouped by feature grouping processor 916 into separate streams of associated audio signals.

Figure 9C is a diagram illustrating how a filter set as described herein is used in a spatial correlator or sound localization system such as is described in United States Patent Application No. 10/004,141 (Attorney Docket No. ANSCP005) by Lloyd Watts (filed November 14, 2001) entitled: COMPUTATION OF MULTI-SENSOR TIME DELAYS which is herein incorporated by reference. A right channel audio signal is input to a right channel filter set 922 and a left channel audio signal is input to a left channel filter set 924. The outputs of the filter sets are correlated by a binaural processor 926 to determine the time delay between the left and right channel input signals. The direction from which a sound emanates may be determined from the time delay.

Although the foregoing invention has been described in some detail for purposes of clarity of understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims. It should be noted that there are many alternative ways of implementing both the process and apparatus of the present invention. Accordingly, the present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein, but may be modified within the scope and equivalents of the appended claims.

WHAT IS CLAIMED IS:

CLAIMS

1. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first set of low pass filters to derive a first set of frequency components wherein the first set of low pass filters are arranged serially in a chain having a first low pass filter and a last low pass filter, the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter;

downsampling the output of the last low pass filter to produce a downsampled signal;

processing the downsampled signal with a second set of low pass filters to derive a second set of frequency components.

2. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the frequency components are derived by subtracting the output of each low pass filter from the input to the low pass filter.

3. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the second set of low pass filters have a different Q than the first set of low pass filters.

4. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the second set of low pass filters have a Q that is less sharp than the first set of low pass filters.

5. A method of analyzing an input signal into a plurality of frequency components as recited in claim 1 wherein the second set of low pass filters have a Q that differs from the Q of the first set of low pass filters substantially according to human critical bandwidth.

6. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first low pass filter to produce a first low pass filtered signal;

subtracting the first low pass filtered signal from the input signal to derive a first frequency component;

processing the signal with a second low pass filter to produce a second low pass filtered signal; and

subtracting the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

7. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first low pass filter to produce a first low pass filtered signal;

subtracting the first low pass filtered signal from the input signal to derive a first frequency component;

processing the low pass filtered signal with a second low pass filter to produce a second low pass filtered signal; and

subtracting the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

8. A method of analyzing an input signal into a plurality of frequency components comprising:

processing the signal with a first filter wherein the first filter is configured to separate part of the signal into a first output frequency channel; and

processing the signal with a second filter wherein the second filter is configured to separate part of the signal into a second output frequency channel wherein the first frequency channel emphasizes higher frequencies than the second frequency channel; and wherein the second filter has a different Q than the first filter.

9. A method of analyzing an input signal into a plurality of frequency components as recited in claim 8 wherein the second filter has a Q that is less sharp than the first filter.

10. A method of analyzing an input signal into a plurality of frequency components as recited in claim 8 wherein the second filter has a Q that differs from the Q of the first filter substantially according to human critical bandwidth.

11. A method of analyzing an input signal into a plurality of frequency components as recited in claim 8 wherein the filters are low pass filters.

12. A system for analyzing an input signal into a plurality of frequency components comprising:

a first set of low pass filters configured to derive a first set of frequency components wherein the first set of low pass filters are arranged serially in a chain

having a first low pass filter and a last low pass filter, the output of each low pass filter being fed to the next low pass filter in the chain until the last low pass filter;

a downsampler configured to downsample the output of the last low pass filter to produce a downsampled signal;

a second set of low pass filters configured to process the downsampled signal to derive a second set of frequency components.

13. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the frequency components are derived by subtracting the output of each low pass filter from the input to the low pass filter.

14. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the second set of low pass filters have a different Q than the first set of low pass filters.

15. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the second set of low pass filters have a Q that is less sharp than the first set of low pass filters.

16. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the second set of low pass filters have a Q that differs from the Q of the first set of low pass filters substantially according to critical band.

17. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the system is used in a voice recognition system.

18. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the system is used for audio stream separation

19. A system for analyzing an input signal into a plurality of frequency components as recited in claim 12 wherein the system is used for sound localization.

20. A system for analyzing an input signal into a plurality of frequency components comprising:

a first low pass filter that outputs a first low pass filtered signal;

a first processor configured to subtract the first low pass filtered signal from the input signal to derive a first frequency component;

a second low pass filter that outputs a second low pass filtered signal; and

a second processor configured to subtract the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

21. A system for analyzing an input signal into a plurality of frequency components comprising:

a first low pass filter that outputs a first low pass filtered signal;

a first processor configured to subtract the first low pass filtered signal from the input signal to derive a first frequency component;

a second low pass filter configured to process the low pass filtered signal to produce a second low pass filtered signal; and

a second processor configured to subtract the second low pass filtered signal from the first low pass filtered signal to derive a second frequency component.

22. A system for analyzing an input signal into a plurality of frequency components comprising:

a first filter configured to process the signal wherein the first filter is configured to separate part of the signal into a first output frequency channel; and

a second filter configured to process the signal wherein the second filter is configured to separate part of the signal into a second output frequency channel wherein the first frequency channel emphasizes higher frequencies than the second frequency channel; and wherein the second filter has a different Q than the first filter.

23. A system for analyzing an input signal into a plurality of frequency components as recited in claim 22 wherein the second filter has a Q that is less sharp than the first filter.

24. A system for analyzing an input signal into a plurality of frequency components as recited in claim 22 wherein the second filter has a Q that differs from the Q of the first filter substantially according to a critical band.

25. A system for analyzing an input signal into a plurality of frequency components as recited in claim 22 wherein the filters are low pass filters.

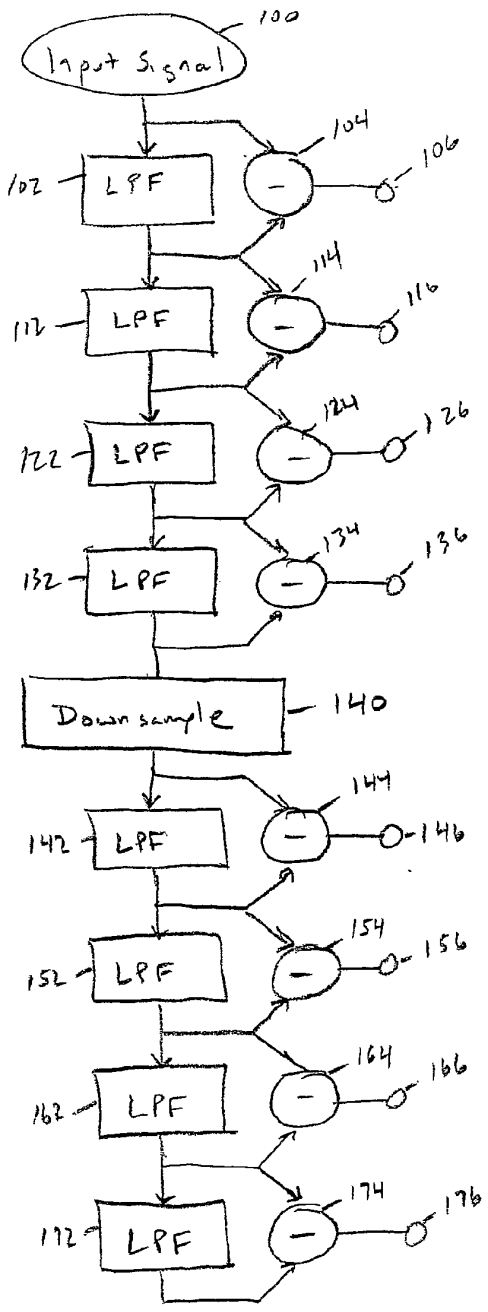

100
Input Signal
104
102
LPF
106
114
112
LPF
116
124
122
LPF
126
134
132
LPF
136
Downsample
140
144
142
LPF
146
154
152
LPF
156
164
162
LPF
166
174
172
LPF
176

Figure 1

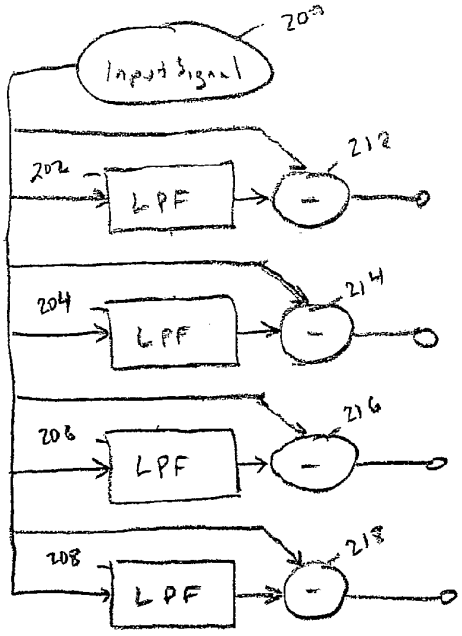


Figure 2

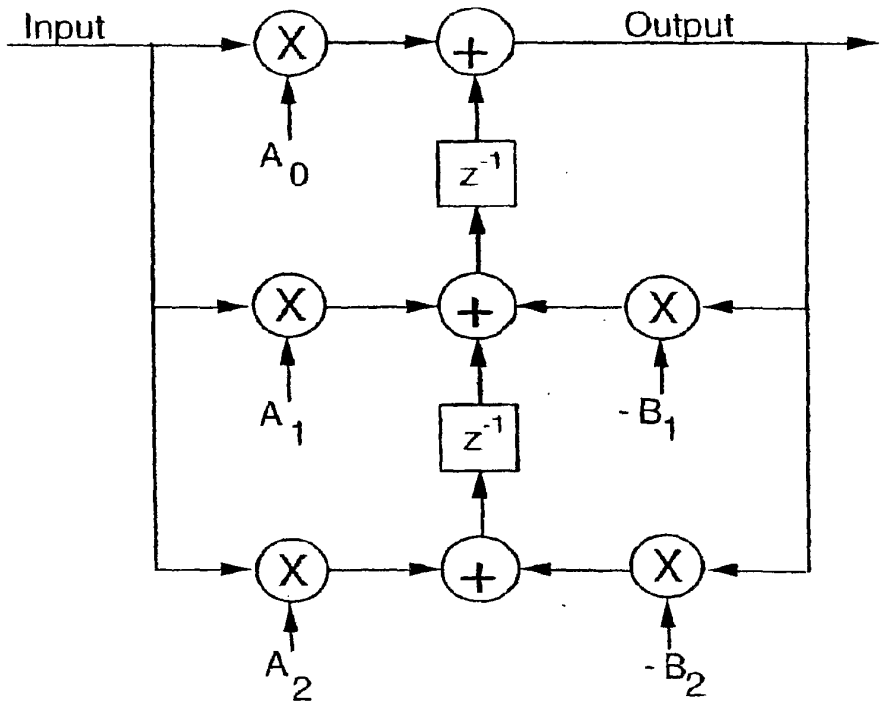


Figure 3

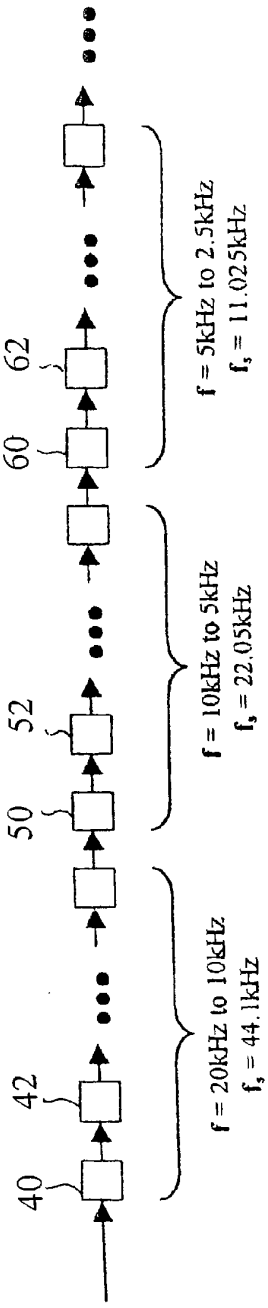

40
42
50
52
60
62
f = 20kHz to 10kHz
f_s = 44.1kHz
f = 10kHz to 5kHz
f_s = 22.05kHz
f = 5kHz to 2.5kHz
f_s = 11.025kHz

Figure 4

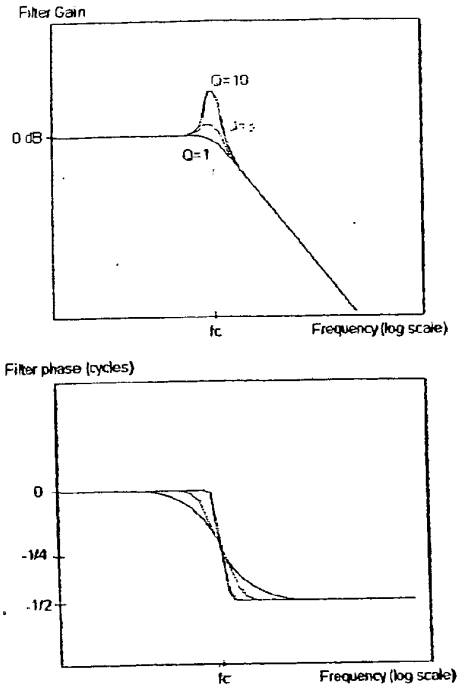


Figure 5

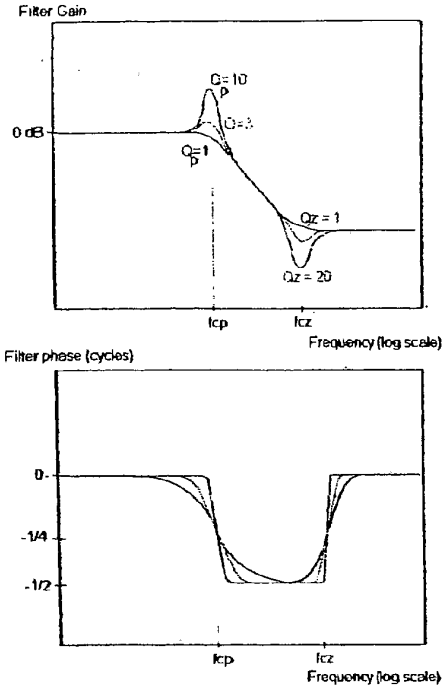


Figure 6

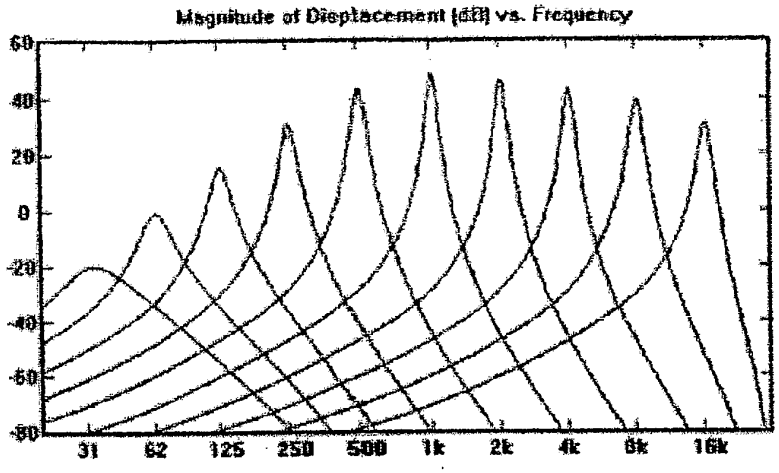


Figure 7

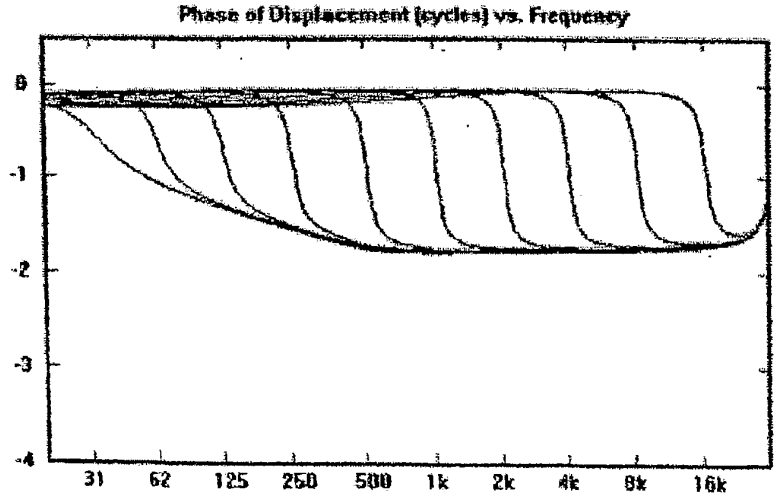


Figure 8

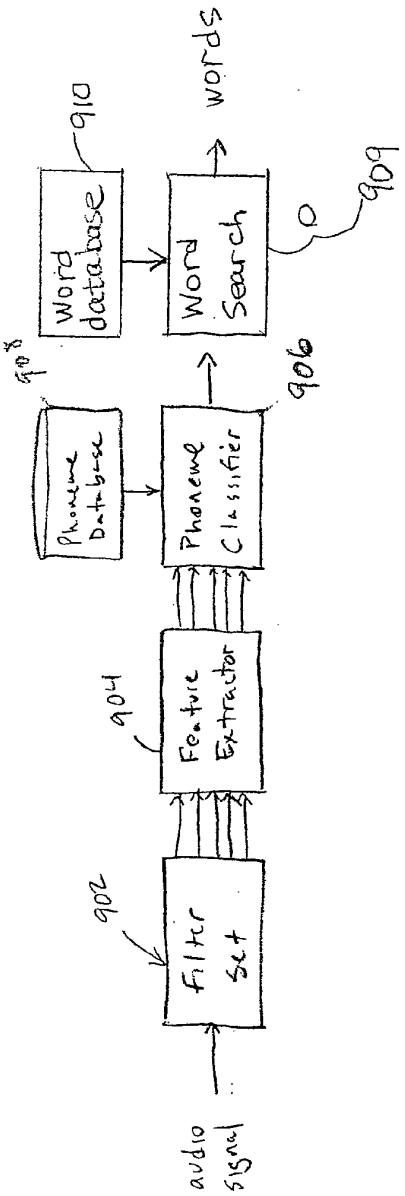

audio signal
902
filter set
904
Feature Extractor
906
Phoneme Classifier
908
Phoneme Database
909
Word Search
910
Word database
words

Figure 9A

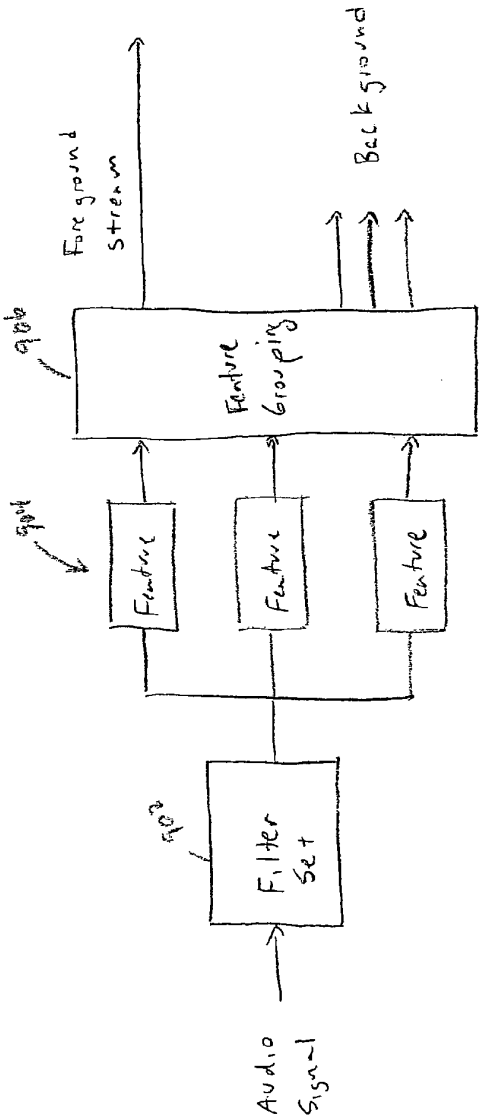

Audio Signal
Filter Set
Feature
Feature
Feature
Feature Groupings
Foreground stream
Background

Figure 9B

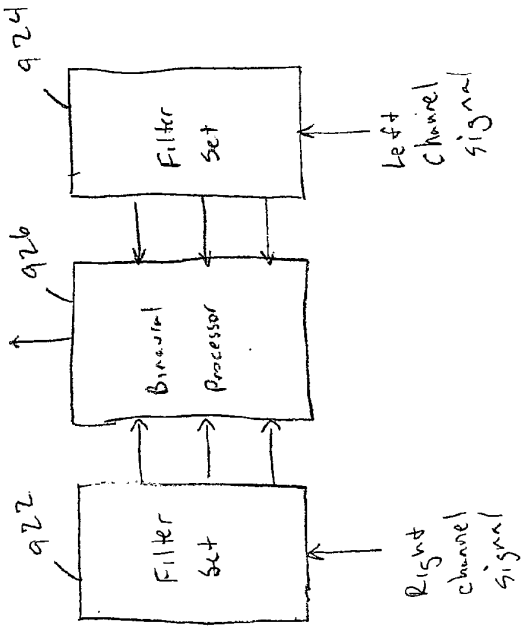

924
Filter Set
Left channel signal
926
Binaural Processor
922
Filter Set
Right channel signal

Figure 9C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/04124

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/00; H03G 5/00; G10L 19/00
US CL : 700/94; 381/98; 704/500

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 700/94; 381/98; 704/500

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	US 2002/0147595 A1 (BAUMGARTE) 10 October 2002 (10.10.2002), figures1, 2 and 7, page 3 and 5.	1,3-5, 8-12, 14-19, 22-25
A	US 6,513,004 B1 (RIGAZIO et al) 28 January 2003 (28.01.2003), entire document.	1-25
A	US 4,536,844 A (LYON) 20 August 1985 (20.08.1985), entire document.	1-25
A	US 5,473,759 A (SLANEY et al) 05 December 1995 (05.12.1995), entire document.	1-25
A	KATES, JAMES M. "A Time-Domain Digital Cochlear Model" IEEE Transactions of Siganl Processing, December 1991, Volume 39 Number 12, pages 2573-2592.	1-25

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search 05 May 2003 (05.05.2003)	Date of mailing of the international search report 29 MAY 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Forester W Isen Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

Espacenet

Bibliographic data: JP2004531767 (A) — 2004-10-14

SPEECH FEATURE EXTRACTION SYSTEM

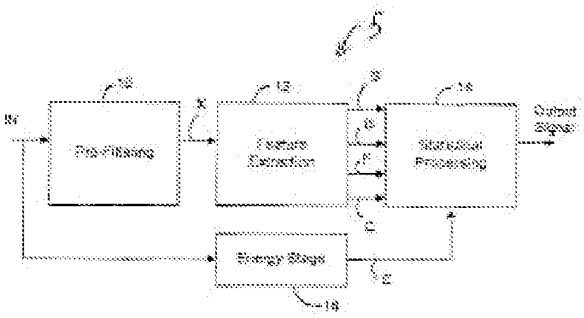
Inventor(s):

Applicant(s):

Classification:	**- international:**	***G10L15/02; G10L15/20; G10L21/02; G10L19/02;*** (IPC1-7): G10L15/02; G10L21/02
	- European:	G10L15/02
Application number:	JP20030505912T 20020614	
Priority number(s):	US20010882744 20010615; WO2002US19182 20020614	
Also published as:	JP4177755 (B2) US6493668 (B1) US2002198711 (A1) US2003014245 (A1) US7013274 (B2) more	

Abstract not available for JP2004531767 (A)
Abstract of corresponding document: US6493668 (B1)

A system suitable for use in a speech recognition system or other voice processing system extracts features related to the frequency and amplitude characteristics of an input speech signal using a plurality of complex band pass filters by processing the outputs of adjacent bandpass filters.



Last updated: 28.02.2012 Worldwide Database 5.7.36; 93p

(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2004-531767
(P2004-531767A)
(43)公表日　平成16年10月14日(2004.10.14)

(51)Int.Cl.[7]	F I			テーマコード（参考）
G10L 21/02	G10L	3/02	301G	5D015
G10L 15/02	G10L	3/00	515E	
	G10L	9/16	301A	

審査請求 未請求　予備審査請求 有　（全 51 頁）

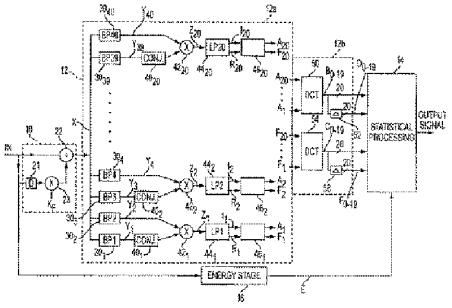
(21)出願番号　特願2003-505912(P2003-505912)
(86)(22)出願日　平成14年6月14日(2002.6.14)
(85)翻訳文提出日　平成15年12月15日(2003.12.15)
(86)国際出願番号　PCT/US2002/019182
(87)国際公開番号　WO2002/103676
(87)国際公開日　平成14年12月27日(2002.12.27)
(31)優先権主張番号　09/882,744
(32)優先日　平成13年6月15日(2001.6.15)
(33)優先権主張国　米国(US)

(71)出願人　503459361
ブランドマン，　イーガル
アメリカ合衆国　カリフォルニア　９４３０６，　パロ　アルト，　ドナルド　ドライブ　４１２０
(74)代理人　100078282
弁理士　山本　秀策
(74)代理人　100062409
弁理士　安村　高明
(74)代理人　100113413
弁理士　森下　夏樹
(72)発明者　ブランドマン，　イーガル
アメリカ合衆国　カリフォルニア　９４３０６，　パロ　アルト，　ドナルド　ドライブ　４１２０

最終頁に続く

(54)【発明の名称】発話特徴抽出システム

(57)【要約】
音声認識システムまたは他の音声処理システム（１４）における使用に対して適したシステム（１２ａ）は、隣接する周波数フィルタ（４２）の出力を処理することによって複数の複素バンドパスフィルタ（３０）を用いて入力音声信号（ＩＮ）の周波数（４６、Ｆ）および振幅（４６、Ａ）特性に関連付けられた特徴を抽出する。本発明の装置は、入力発話信号を受信するように適用され、かつ第１および第２信号のそれぞれを提供する第１および第２のバンドパスフィルタと、第２のバンドパスフィルタに接続され、かつ第２の信号の共役である第３の信号を提供する共役回路と、乗算器と、フィルタ手段とを含む。
【選択図】図２



【特許請求の範囲】
【請求項1】
周波数および振幅特徴を有する入力発話信号から特徴を抽出するための発話処理システムにおける使用のための装置であって、
該入力発話信号を受信するように適用され、かつ第１および第２信号のそれぞれを提供する第１および第２のバンドパスフィルタと、
該第２のバンドパスフィルタに接続され、かつ該第２の信号との共役である第３の信号を提供する共役回路と、
該第１のバンドパスフィルタおよび該共役回路に接続され、かつ該第１および該第３の信号の積である第４の信号を提供する乗算器と、
該入力信号の該周波数特徴に関連した第５の信号、および該入力信号の振幅特徴に関連した第６の信号を提供するために該乗算器に接続されたフィルタ手段と
を含む、装置。
【請求項2】
前記フィルタは、ロウパスフィルタを含む、請求項１に記載の装置。
【請求項3】
前記入力信号と前記第１のおよび第２のバンドパスフィルタとの間に接続されたバンドパスフィルタをさらに含む、請求項１に記載の装置。
【請求項4】
前記ロウパスフィルタは、実数部Ｒおよび虚数部Ｉを有する信号を提供し、前記フィルタ手段は、Ａ＝１ｏｇ　ＲおよびＦ＝Ｉ／Ｓｑｒｔ（Ｒ２＋Ｉ２）を計算するための回路を含む、請求項２に記載の装置。
【請求項5】
前記第１および第２のバンドパスフィルタの各々は、遅延エレメント、複素乗算器、および加算器を含む、請求項１に記載の装置。
【請求項6】
前記加算器は、前記入力発話信号を受信し、出力信号を出力し、該加算器は、入力発話信号および前記複素係数によって乗算された出力信号の遅延されたサンプルを含む信号を加算する、請求項５に記載の装置。
【請求項7】
特徴拡張回路をさらに含む、請求項４に記載の装置。
【請求項8】
前記特徴拡張回路の出力を統計的に処理し、該出力を複数の所定のモデルと比較するための回路をさらに含む、請求項７に記載の装置。
【請求項9】
前記特徴拡張回路は、ＡおよびＦの離散コサイン変換を計算するための回路を含む、請求項７に記載の装置。
【請求項10】
前記特徴拡張回路は、Ａの連続的な離散コサイン変換間の差を計算するため、およびＦの連続的な離散コサイン変換間の差を計算するための回路をさらに含む、請求項７に記載の装置。
【請求項11】
音を電気信号に変換するためのトランスデューサと、
前記電気信号をデジタル信号に変換するためのサンプラーと
をさらに含む、請求項１に記載の装置。
【請求項12】
デジタルプロセッサをさらに含み、前記第１および第２のバンドパスフィルタ、前記共役回路、前記乗算器、ならびに、前記フィルタ手段は、該デジタルプロセッサ上で実行されるように適応されたアルゴリズムをさらに含む、請求項１に記載の装置。
【請求項13】

前記第１のバンドパスフィルタの周波数は、対数スケールに従って選択され、前記第２のバンドパスフィルタの周波数は、該第１のバンドパスフィルタの周波数未満の所定の周波数で選択される、請求項１に記載の装置。
【請求項14】
前記第２のフィルタの中心の周波数は、前記第１のフィルタの中心の周波数未満の１００ヘルツである、請求項１３に記載の装置。
【請求項15】
前記第１および第２のフィルタの帯域は重なる、請求項１３に記載の装置。
【請求項16】
前記第１および第２のフィルタの帯域幅は同じである、請求項１３に記載の装置。
【請求項17】
前記第１のバンドパスフィルタの周波数は、メルスケールに従って選択され、前記第２のバンドパスフィルタの周波数は、該第１のバンドパスフィルタの周波数未満の所定の周波数において選択される、請求項１に記載の装置。
【請求項18】
前記第２のフィルタの中心の周波数は、前記第１のフィルタの中心の周波数未満の１００ヘルツである、請求項１７に記載の装置。
【請求項19】
前記第１および第２のフィルタの帯域は重なる、請求項１７に記載の装置。
【請求項20】
前記メルスケールは、式
【数1】

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700} \right)$$

によって定義される、請求項１７に記載の装置。
【請求項21】
前記第１および第２のフィルタの帯域幅は同じである、請求項１７に記載の装置。
【請求項22】
前記第１および第２のバンドパスフィルタは、高速フーリエ変換を用いてインプリメントされる、請求項１に記載の装置。
【請求項23】
発話処理デバイスにおける使用のために入力発話信号から特徴を抽出するための方法であって、
該入力信号を、第１の周波数帯域における第１の信号および第２の周波数帯域における第２の信号を分離するステップと、
該第１の信号との共役を提供するステップと、
第３の信号を提供するために、該第１の信号の共役と第２の信号の共役を乗算するステップと、
該入力発話信号における周波数特徴に関連した周波数成分および該入力発話信号における振幅特徴に関連した振幅成分を生成するために、該第３の信号を処理するステップと
を包含する、方法。
【請求項24】
前記第１および第２の信号を生成するために、前記入力信号を処理する前に、前記発話信号をハイパスフィルタリングするステップをさらに包含する、請求項２３に記載の方法。
【請求項25】

前記第３の信号は、該第３の信号をロウパスフィルタリングするステップをさらに包含する、請求項２３に記載の方法。
【請求項26】
前記第３の信号をロウパスフィルタリングするステップは、実数部Ｒおよび虚数部Ｉを有する信号を提供し、該第３の信号を処理するステップは、Ａ＝１ｏｇ　ＲおよびＦ＝Ｉ／Ｓｑｒｔ（Ｒ２＋Ｉ２）を計算するステップを包含する、請求項２５に記載の方法。
【請求項27】
一連の特徴ベクトルを生成するための特徴強化技術を用いて前記第３の信号を処理するステップをさらに包含する、請求項２６に記載の方法。
【請求項28】
前記一連の特徴ベクトルを複数の所定のモデルと比較することによって該一連の特徴ベクトルを統計的に処理するステップをさらに包含する、請求項２７に記載の方法。
【請求項29】
ＡおよびＦの離散コサイン変換を計算するステップをさらに包含する、請求項２７に記載の方法。
【請求項30】
Ａの連続的な離散コサイン変換間およびＦの連続的な離散コサイン変換間の差を計算するステップをさらに包含する、請求項２９に記載の方法。
【請求項31】
前記差を複数の所定のパターンパターンと比較するために、連続的に離散コサイン変換の間の差を統計的に処理するステップをさらに包含する、請求項３０に記載の方法。
【請求項32】
音を電気的信号に変換するステップと、
デジタル信号を供給するために、該電気信号をサンプリングするステップであって、前記入力発話信号は、該デジタル信号を包含する、請求項２３に記載の方法。
【請求項33】
対数スケールに従って前記第１の周波数を選択するステップ、および前記第１の周波数帯域の周波数未満の所定の周波数で前記第２の周波数帯域の周波数を選択するステップを包含する、請求項２３に記載の方法。
【請求項34】
前記第２のフィルタの中心の周波数は、前記第１の周波数帯域の中心の周波数未満の１００ヘルツである、請求項３３に記載の方法。
【請求項35】
前記第１および第２の周波数帯域を重ね合わせるステップを包含する、請求項３３に記載の方法。
【請求項36】
前記第１の周波数帯域の周波数をメルスケールに従って選択するステップ、および該第１の周波数帯域の周波数未満の所定の周波数における前記第２の周波数帯域の周波数を選択するステップを包含する、請求項２３に記載の方法。
【請求項37】
前記第２の周波数帯域の中心周波数は、前記第１の周波数帯域の中心周波数未満の１００ヘルツである、請求項３６に記載の装置。
【請求項38】
前記第１および第２の周波数帯域を重ね合わせることを含む、請求項３６に記載の装置。

【発明の詳細な説明】
【技術分野】
【０００１】
（発明の背景）
本発明は、発話認識、音声識別、音声認証システムにおいて使用するための発話特徴抽出システムに関する。より詳細には、本発明は、低減されたエラーレートを有する発話認識

システムまたは他の発話処理システムを作成するために使用され得る発話特徴処理システムに関する。

【背景技術】

【０００２】

一般的に、発話認識システムは、発話者の音声信号を解析することによって発話された語を識別することを試みる装置である。発話は、特徴が抽出される電気的形態に変換される。次いで、システムは、公知の発話ユニットに関連付けられた以前に格納されたモデルのシーケンスに特徴のシーケンスを整合させることを試みる。特定の規則に従って、特徴のシーケンスがモデルのシーケンスに対応する場合、対応する語は、発話認識システムによって認識されているとみなされる。

【０００３】

しかし、バックグランド音（ラジオ、車の騒音、または他の付近の発話者）は、音声から有用な特徴を抽出することが困難であり得る。さらに、異なるマイクロホン、電話受話器または電話回線の使用等の周囲の条件の変化は、システム性能と干渉し得る。あるいは、発話者の距離、発話者間の差異、発話者の抑揚または強調、および発話者の健康であっても、システム性能に悪影響を与え得る。いくつかのこれらの問題のさらなる説明は、Ｒｉｃｈａｒｄ Ａ．Ｑｕｉｎｎｅｌｌによる、「Ｓｐｅｅｃｈ Ｒｅｃｏｇｎｉｔｉｏｎ：Ｎｏ Ｌｏｎｇｅｒ ａ Ｄｒｅａｍ，Ｂｕｔ Ｓｔｉｌｌ ａ Ｃｈａｌｌｅｎｇｅ，」ＥＤＭ Ｍａｇａｚｉｎｅ，１９９５年1月１９日、４１～４６頁を参照のこと。

【０００４】

ほとんどの発話認識システムでは、発話特徴は、ケプストラム解析によって抽出され、ケプストラム解析は、特定の周波数帯域におけるエネルギーを測定することを含む。この解析の結果は、これらの帯域の信号の振幅を反映する。連続的な期間にわたるこれらの振幅変化の解析は、振幅変調信号としてモデル化され得る。

【０００５】

人間の耳が、受信された発話信号の周波数変調および振幅変調に敏感であるが、この周波数変調された量は、ケプストラム解析を実行するシステムにおいて部分的のみ反映される。

【０００６】

従って、発話の周波数変調特徴および以前から公知の振幅変調特徴をキャプチャすることが可能である発話特徴抽出システムを提供することが望ましい。

【０００７】

入力発話信号の周波数変調特徴に関する情報を提供する特徴抽出システムを組み込む発話認識および他の発話処理システムを提供することもまた望ましい。

【発明の開示】

【課題を解決するための手段】

【０００８】

（発明の要旨）

上記観点から、本発明の目的は、発話の周波数変調特徴および以前から公知の振幅変調特徴をキャプチャすることが可能な発話特徴抽出システムを提供することである。

【０００９】

本発明のさらなる目的は、入力発話信号の周波数変調特徴に関する情報を提供する特徴抽出システムを組み込む発話認識および他の発話処理システムを提供することである。

【００１０】

本発明は、発話の周波数変調特徴および振幅特徴を反映する発話特徴抽出システムを提供する。これは、一実施形態では、線形周波数スケール（「線形スケール」）に従って、隣接する周波数帯域において配置された複数の複素バンドパスフィルタを含むフィルタ抽出ステージによってなされる。複数の複素バンドパスフィルタは、複数のペアに分割される。１つのペアは、隣接する周波数帯域において２つの複素バンドパスフィルタを含む。各ペアに対して、より高い周波数（「一次周波数」）においてフィルタの出力は、より低い

周波数（「二次フィルタ」）におけるフィルタの出力の共役によって増倍される。生成した信号は、ロウパスフィルタリングされる。

【００１１】

別の実施形態では、特徴抽出位相は、対数（または指数）周波数スケール（「対数スケール」）に従って配置された複数の複素バンドパスフィルタを含む。フィルタ対の一次フィルタは、対数スケールに沿った種々の周波数において集中される。各ペアの一次フィルタに対応する二次フィルタは、一次フィルタ未満の所定の周波数において集中される。各ペアに対して、一次フィルタの出力は、二次フィルタの出力の共役によって増倍される。生成した信号は、ロウパスフィルタリングされる。

【００１２】

さらに別の実施形態では、複数のバンドパスフィルタがメルスケール（ｍｅｌ－ｓｃａｌｅ）に従って配置される。フィルタ対の一次フィルタは、メルスケールに沿った種々の周波数において集中される。各ペアの一次フィルタに対応する二次フィルタは、一次フィルタ未満の所定の周波数において集中される。各ペアに対して、一次フィルタの出力は、二次フィルタの出力の共役によって増倍される。生成した信号は、ロウパスフィルタリングされる。

【００１３】

さらに別の実施形態では、複数のバンドパスフィルタが、上述の線形およびログスケールメルスケール（ｍｅｌ－ｓｃａｌｅ）に従って配置される。バンドパスフィルタのペアの部分は、線形スケールに従って隣接する周波数帯域に配置される。これらの対の各々に対して、一次フィルタの出力は、二次フィルタの出力の共役によって増倍される。生成した信号は、ロウパスフィルタリングされる。

【００１４】

バンドパスフィルタの残りの対の一次フィルタは、対数スケールに沿った種々の周波数において集中され、一次フィルタに対応する二次フィルタは、一次フィルタ未満の所定の周波数で集中される。これらの対の各々に対して、一次フィルタの出力は、二次フィルタの出力の結合によって増倍される。生成した信号は、ロウパスフィルタリングされる。

【００１５】

上述の実施形態に対して、ロウパスフィルタの各々は、２つの成分（ロウパスフィルタ出力が生成された隣接するバンドパスフィルタを通過した信号の周波数に実質的に敏感であるＦＭ成分、および、隣接するバンドパスフィルタを通過した信号の振幅に実質的に敏感であるＡＭ成分）を計算するために処理される。ＦＭ成分は、ロウパスフィルタ出力を生成するために使用された隣接するバンドパスフィルタの出力の位相差を反映する。

【００１６】

次いで、ＡＭおよびＦＭ成分は、離散コサイン変換、メルスケール変換、平均規格化、デルタおよび加速度解析、線形判別分析、主成分解析等の公知の特徴分析を用いて処理され、統計的処理、あるいは他の認識または識別方法に対して適用可能な発話特徴を生成する。代替的な実施形態では、複数の複素バンドパスフィルタは、高速フーリエ変換（ＦＦＴ）または他のデジタル信号処理（ＤＳＰ）技術を用いてインプリメントされ得る。

【００１７】

さらに、本発明の方法および装置は、発話認識システムにおけるケプストラル解析を実行することに加えて使用され得る。本発明の上記および他の目的が、添付の図面を考慮して、以下の詳細な説明の考慮によって明らかになる。その図面では、同様な参照符号は同様の部品を指す。

【発明を実施するための最良の形態】

【００１８】

（発明の詳細な説明）

図１を参照すると、例示的なスピーチ認識システム５の一般的な記載は、本発明のスピーチ抽出システムを取り込むことが記載される。当業者に明らかなように、本発明のスピーチ特徴抽出システムはまた、話し手の識別、認証、および他の音声処理システムにおいて

用いられ得る。

【００１９】

システム５は、プリフィルタリングステージ１０、特徴抽出ステージ１２、統計処理ステージ１４、およびエネルギーステージ１６の４つのステージを含む。

【００２０】

プリフィルタリングステージ１０、統計処理ステージ１４、およびエネルギーステージ１６は、当業者に公知のスピーチ処理技術を利用し、本発明の部分を形成しない。特徴抽出ステージ１２は、本発明のスピーチ特徴抽出システムを取り込み、以下に記載されるように、当業者に公知である特徴拡張技術を含む。

【００２１】

音声スピーチ信号は、マイクロホン、受話器、または他のデバイスによって電気的な信号に変換され、システム５に入力スピーチ信号として提供される。本発明の好適な実施形態において、電気的な信号は、サンプリングされるか、またはデジタル化され、音声スピーチを表すデジタル信号（ＩＮ）を提供する。プリフィルタリングステージ１０は、音声信号ＩＮの高周波コンポーネントを増幅して、プリフィルタリングされた信号は、次に、特徴抽出ステージ１２に提供される。

【００２２】

特徴抽出ステージ１２は、プリフィルタリングされた信号Ｘを処理して、スピーチ認識のために役立ち得る入力信号ＩＮの特性に関係した特徴ベクトルのシーケンスを生成する。特徴抽出ステージ１２の出力は、特徴ベクトルのシーケンスと好適な統計なモデルとを比較して、入力信号ＩＮにおいてワードまたは他のスピーチユニットを識別する統計処理ステージ１４によって用いられる。特徴ベクトルは、例えば、Ｊｅｌｉｎｅｋの「Ｓｔａｔｉｓｔｉｃａｌ　Ｍｅｔｈｏｄｓ　ｆｏｒ　Ｓｐｅｅｃｈ　Ｒｅｃｏｇｎｉｔｉｏｎ」、Ｔｈｅ　ＭＩＴ　Ｐｒｅｓｓ、１９９７、ｐｐ．１５－３７に記載される隠れマルコフモデル（Ｈｉｄｄｅｎ　Ｍａｒｋｏｖ　Ｍｏｄｅｌ；ＨＭＭ）といった公知の技術を用いてモデルが比較される。統計処理ステージ１４の出力は、認識されたワードまたは特定のアプリケーションによる他の適切な出力である。

【００２３】

ステージ１４における統計処理は、ローカルに実行されるか、またはステージ１０、１２の処理が実行される場合に関連して、遠隔位置で実行され得る。例えば、特徴ベクトルのシーケンスは、統計処理のために遠隔サーバに転送され得る。

【００２４】

図１の例示のスピーチ認識システムは、好適には、入力信号ＩＮのフレームにおいて総エネルギーを示す出力信号を提供するエネルギーステージ１６をさらに含む。統計処理ステージ１４は、この総エネルギー情報を用いて、入力信号に含まれるスピーチの改善された認識を提供し得る。

【００２５】

ここで図２を参照すると、プリフィルタリングステージ１０および特徴抽出ステージ１２は、より詳細に記載される。プリフィルタリングステージ１０は、入力信号の高周波コンポーネントを増幅するハイパスフィルタである。プリフィルタリングステージ１０は、１サンプル遅延素子２１、乗算器２３、および加算器２４を含む。乗算器２３は、典型的に－０．９７のある値を有する定常のk_fによって１サンプル遅延された信号を乗算する。プリフィルタリングステージ１０の出力Ｘは、バンドパスフィルタ30_1、30_2、．．．30_nのバンクにおけるサンプリングレートにおいて出力される。

【００２６】

一実施形態において、バンドパスフィルタ30_1、30_2、．．．30_nは、近接した周波数バンドにおいて位置付けられる。図５のグラフ７２において示されるように線形周波数スケール（「線形スケール」）６８によってバンドパスフィルタ30_1、30_2、．．．30_nの間隔があけられる。用語「線形周波数スケール」は、通常かつ従来の意味（すなわち、実際の周波数分割は不規則に間隔があけられる）に従って本明細書中に用いられ

る。複数の複素バンドパスフィルタ３０$_1$、３０$_2$、．．．３０$_n$は、ペアＰ$_{1-2}$に分割される。ペア（Ｐ$_1$またはＰ$_2$）は、隣接した周波数バンドにそれぞれ２つの複素バンドパスフィルタ（３０$_{1-2}$または３０$_{3-4}$）を含む。全てのペア（Ｐ$_1$またはＰ$_2$）のため、高周波数バンド（３０$_2$または３０$_4$）におけるフィルタの出力（「一次フィルタ」として以下に言及される）は、より低い周波数バンド（３０$_1$または３０$_3$）におけるフィルタの出力の共役によって乗算される（「二次フィルタ」として以下に言及される）。結果として生じる信号は、ロウパスフィルタリングされる。

【００２７】

バンドパスフィルタ３０$_1$、３０$_2$、．．．３０$_n$の数および周波数バンドの幅は、好適には、スピーチ処理システムのアプリケーションによって選択される。例えば、電話技術のアプリケーションに役立つシステムは、中心周波数からほぼ１００Ｈｚ離れた周波数を有する約４０のバンドパスフィルタ３０$_1$、３０$_2$、．．．３０$_n$を利用する。例えば、フィルタ３０$_1$は５０Ｈｚの中心周波数を有し、フィルタ３０$_2$は１５０Ｈｚの中心周波数を有し、フィルタ３０$_3$は２５０Ｈｚの中心周波数を有し得、ゆえに、フィルタ３０$_{40}$の中心周波数は３９５０Ｈｚである。各フィルタのバンド幅は、数百ヘルツであり得る。

【００２８】

別の実施形態において、図６のグラフ７０に示されるように、バンドパスフィルタ３０$_1$、３０$_2$、．．．３０$_{108}$は、例えば、対数的（または指数関数的）周波数スケール７４（「対数スケール」）といった非線形周波数スケールによって構成される。用語対数的周波数は、通常かつ従来の意味によって本明細書中に用いられる。

【００２９】

経験的な証明は、図５の線形スケール６８の代わりに図６の対数スケール７４を用いることが音声認識性能を改良することを示唆する。すなわち、なぜなら、人間の耳は、音声スペクトルを超える周波数を非線形に消滅するからである。線形スケール６８の代わりに対数スケールを用いる別の利点は、対数スケール７４がさらなるバンドパスフィルタ３０$_1$、３０$_2$、．．．３０$_n$を用いることなく周波数スペクトルのより広い領域をカバーし得る。

【００３０】

バンドパスフィルタ３０$_{1-108}$のペアＰ$_{1-54}$は、対数スケール７４によって間隔があけられる。ペアＰ$_1$はフィルタ３０$_1$および３０$_2$を含み、ペアＰ$_{10}$はフィルタ３０$_{19}$を含み、ペアＰ$_{54}$はフィルタ３０$_{107}$および３０$_{108}$を含む。この構成において、フィルタ３０$_1$、３０$_{20}$および３０$_{108}$は一次フィルタであり、フィルタ３０$_{19}$、および３０$_{107}$は二次フィルタである。

【００３１】

１つの好適な実施形態において、一次フィルタ３０$_2$、３０$_{20}$、．．．３０$_{108}$は、対数スケール７４に沿った多様な周波数において集中され、二次フィルタ３０$_1$、３０$_3$、．．．３０$_{107}$は、それぞれ、一次フィルタ３０$_2$、３０$_4$、．．．３０$_{108}$に対応する１００ヘルツ（Ｈｚ）未満に集中される。図６のグラフ７０を作成するために例示のＭＡＴＬＡＢコードは以下に示される。

【００３２】

【数２】

```
v=(2.^([26.715:0.25:40]/3.345)); % バンドパスフィルタペアに対する中心周波数を生成する
f(2:2:2*length(v))=v+50;          % 一次フィルタ中心周波数
f(1:2:2*length(v))=v-50;          % 二次フィルタ中心周波数
semilogy([v'+50 v'-50],'.');grid  % 対数スケール上の中心周波数をプロットする
```

別の実施形態において、一次フィルタ３０$_{2}$、３０$_{20}$、・・・３０$_{108}$および二次フィルタ３０$_{1}$、３０$_{3}$、・・・３０$_{107}$の中心周波数は、別々および独立したアルゴリズムを用いて配置され得、二次フィルタ３０$_{1}$、３０$_{3}$、・・・３０$_{107}$がそれらの対応する一次フィルタ３０$_{2}$、３０$_{20}$、・・・３０$_{108}$未満の１００ヘルツ（Ｈｚ）に集中される。

【００３３】

一実施形態において、バンドパスフィルタ３０$_{1-108}$は、三角形状である。別の実施形態において、バンドパスフィルタ３０$_{1-108}$は、特定の音声認識システムの必要性による多様な形状であり得る。

【００３４】

対数スケール７４は、０～４０００Ｈｚの領域に示され得る。全てのペアＰ$_{1-54}$に対して、一次フィルタ３０$_{2}$、３０$_{20}$、・・・３０$_{108}$の出力は、二次フィルタ３０$_{1}$、３０$_{3}$、・・・３０$_{107}$の出力の共役によって乗算される。結果として生じる信号は、ロウパスフィルタリングされる。

【００３５】

ペアＰ$_{1-54}$は、より低い周波数がより高い周波数よりもより高い集中ペアＰ$_{1-54}$の共役を含むように構成される。例えば、５００～１０００Ｈｚの周波数領域において７つのペア（Ｐ$_{16-22}$）があり、３０００～３５００Ｈｚの周波数領域において３つだけのペア（Ｐ$_{49-51}$）がある。従って、より低い周波数においてオーバーサンプリングがあり、この実施形態は、より高い周波数において少なくともいくつかのサンプリングをさらに実行する。対数スケール７４に沿ったペアＰ$_{1-54}$の集中は、特定の音声認識システムの必要性によって変化され得る。

【００３６】

デジタル信号処理設計の当業者に明らかであるように、前述の実施形態のバンドパスフィルタは、任意の多くのソフトウェアまたはハードウェア技術を用いてインプリメントされ得る。例えば、複数の複素フィルタは、高速フーリエ変換（ＦＦＴ）、チャープＺ変換、他の周波数ドメイン分析技術を用いてインプリメントされ得る。

【００３７】

代替の実施形態において、図７に記載されるように、バンドパスフィルタ３０$_{1}$、３０$_{2}$、・・・３０$_{n}$は、例えば、メルスケールといった非線形周波数スケールによって構成される。メルスケール８０は、音声認識システムの技術において周知であり、典型的に等式

【００３８】

【数３】

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right)$$

によって定義される。ここで、ｆは、線形スケール６８による周波数を表し、Ｍｅｌ（ｆ）は、それに対応するメルスケール８０の周波数を表す。

【００３９】

図７は、メルスケール８０によって間隔があけられたバンドパスフィルタ３０$_{1-9}$を示すグラフ８４の一実施形態を示す。中心周波数（ＣＦ$_{1-9}$）は、上述の等式を用いることによって計算されるＭｅｌ（ｆ）値である。典型的に、フィルタ３０$_{1-9}$は、ゼロからナイキスト（Ｎｙｑｕｉｓｔ）周波数までの全体の周波数領域に広がる。一実施形態において、フィルタ３０$_{1-9}$は、同じバンド幅を有する。別の実施形態において、フィルタ３０$_{1-9}$は異なるバンド幅を有し得る。

【００４０】

図８に記載される、さらに別の実施形態において、バンドパスフィルタ３０$_{1-8}$は線形６８および非線形７４の周波数スケールの組み合わせによって間隔があけられる。バンドパスフィルタ３０$_{1-4}$（Ｐ$_{1-2}$）は、線形周波数６８によって隣接した周波数バンド

に構成される。

【００４１】

一次フィルタ３０$_6$および３０$_8$は、対数スケール７４に沿って集中される。二次フィルタ３０$_5$および３０$_7$は、３０$_6$および３０$_8$の中心周波数未満の１００Ｈｚの周波数において集中される。これらのペア（Ｐ$_1$またはＰ$_2$）のそれぞれのため、一次フィルタ（３０$_6$および３０$_8$）の出力は、二次フィルタ（３０$_5$および３０$_7$）の出力の共役によってそれぞれ乗算され、結果的に生じる信号は、ロウパスフィルタリングされる。

【００４２】

図２を再度参照すると、ブロック４０$_{1-20}$は、バンドパスフィルタ３０$_1$、３０$_2$、．．．３０$_{n-1}$の出力信号の複素共役を提供する。乗算器ブロック４２$_{1-20}$は、近似のより高い周波数のバンドパスフィルタ３０$_2$、３０$_4$、３０$_6$、．．．３０$_{40}$の出力によって複素共役を乗算する。つぎに、出力信号Ｚ$_{1-20}$は、一連のロウパスフィルタ４４$_{1-22}$を通される。ロウパスフィルタの出力は、典型的に、特徴フレームレートにおいてのみ生成される。例えば、８ｋＨｚの入力スピーチサンプリングレートにおいて、ロウパスフィルタの出力は、１０ミリ秒毎に１度の特徴フレームレートにおいて計算されるだけである。

【００４３】

ロウパスフィルタ４４$_{1-22}$の各出力は、実数のコンポーネントＲおよび虚数のコンポーネントＩを有する複素信号である。ブロック４６$_{1-20}$は、ロウパスフィルタ出力の実数および虚数のコンポーネントを処理し、等式（１）および（２）、

【００４４】

【数４】

$$A_i = \log R_i \qquad (1)$$

$$F_i = \frac{I_i}{\sqrt{R_i^2 + I_i^2}} \qquad (2)$$

に示されるように出力信号Ａ$_{1-20}$およびＦ$_{1-20}$を提供し、ここで、Ｒ$_i$およびＩ$_i$は対応するロウパスフィルタ出力の実数および虚数のコンポーネントである。出力信号Ａ$_i$は、ロウパスフィルタ出力の振幅の関数であり、Ｆ$_i$はロウパスフィールドルタ出力が生成された隣接したバンドパスフィルタによって通過される信号の周波数の関数である。入力信号の振幅および周波数を示す２つのセットの信号を計算することによって、本発明のスピーチ特徴抽出システムを組み込むスピーチ認識システムは、減少されるエラーレートを提供するように期待される。

【００４５】

次に、振幅信号Ａ$_{1-20}$および周波数信号Ｆ$_{1-20}$は、例えば、離散コサイン変換、メルスケール変換、平均規格化、デルタおよび加速度分析、線形判別分析、ならびに、当該分野において本来公知である主要コンポーネント分析を用いて、特徴拡張コンポーネント１２ｂにおける従来の特徴拡張技術を用いて処理される。本発明のスピーチ抽出システムに組み込む本発明のスピーチ認識システムの好適な実施形態は、本明細書中で以下に記載されるように、離散コサイン変換およびデルタ特徴技術を利用する。

【００４６】

さらに図２を参照すると、特徴拡張コンポーネント１２ｂは、出力信号Ａ$_{1-20}$およびＦ$_{1-20}$を受信し、それぞれ離散コサイン変換（ＤＣＴ）ブロック５０および５４を用いてそれらの信号を処理する。ＤＣＴ５０および５４は、信号Ａ$_{1-20}$およびＦ$_{1-20}$の分散行列を対角化するように試みる。これは、ＤＣＴ５０の出力信号Ｂ$_{0-19}$およ

びＤＣＴ５４の出力信号C_{0-19}における特徴を無相関するために役立つ。次に、出力信号B_{0-19}およびC_{0-19}の各セットは統計的な処理ステージ１４に入力される。出力信号B_{0-19}を提供するように入力信号A_{0-20}上のＤＣＴ５０によって実行される関数は等式（３）によって示され、出力信号C_{0-19}を提供するように入力信号F_{1-20}上のＤＣＴ５４によって実行される関数は等式（４）によって示される。

【００４７】

【数５】

$$B_r = D(r) \sum_{n=0}^{N-1} A_{n+1} \bullet \cos\frac{(2n+1)\pi r}{2N} \qquad (3)$$

$$C_r = D(r) \sum_{n=0}^{N-1} F_{n+1} \bullet \cos\frac{(2n+1)\pi r}{2N} \qquad (4)$$

等式（３）および（４）において、Ｎは入力信号ベクトルＡおよびＦの長さ（例えば、図２においてＮ＝２０）に等しく、ｎは０からＮ－１のインデックス（例えば、図２の実施形態においてｎ＝０～１９）であり、ｒは出力信号ＢおよびＣのインデックス（例えば、図２の実施形態においてｒ＝０～１９）である。従って、各ベクトル出力信号B_rのため、入力信号A_{1-20}の各ベクトルは、コサイン関数およびＤ（ｒ）によって乗算され、等式（３）において示されるように共に加算される。各ベクトル出力信号C_rのため、入力信号S_{1-20}の各ベクトルは、コサイン関数およびＤ（ｒ）によって乗算され、等式（４）において示されるように共に加算される。Ｄ（ｒ）は、等式

【００４８】

【数６】

$$D(r) = \frac{1}{\sqrt{N}} \qquad r = 0 \text{に対して} \qquad (5)$$

$$D(r) = \frac{\sqrt{2}}{\sqrt{N}} \qquad r > 0 \text{ に対して} \qquad (6)$$

によって得られる係数である。

【００４９】

出力信号B_{0-19}およびC_{0-19}はまた、それぞれデルタブロック５２および５６に入力される。デルタブロック５２および５６のそれぞれは、連続的な特徴フレーム間の特徴ベクトル値の測定間の差を取得し、この差がスピーチ認識性能を拡張するように用いられ得る。いくつかの差の式は、当該分野に公知であるように、デルタブロック５２および５６によって用いられ得る。例えば、デルタブロック５２および５６は、２つの連続的な特徴フレーム間の差を取得し得る。デルタブロック５２および５６の出力信号は、統計処理ステージ１４に入力される。

【００５０】

式

【００５１】

【数７】

$$E(nT) = \log\left(\frac{\sum_{i=0}^{K-1} I\,N^2_{(n-i)T}}{K}\right) \qquad (7)$$

によって、入力スピーチ信号ＩＮの各フレームの総エネルギーの対数（Ｅで表される）を計算するための図２のエネルギーステージ１６は、既に公知の技術である。

【００５２】

式（７）は、エネルギーブロック１６が前のＬサンプリング間隔（例えば、Ｋ＝２２０、Ｔ＝１／８０００秒）の間の入力信号ＩＮの値の正方形の和を取得し、Ｋで和を割って、最終的な結果の対数を取得することを示す。エネルギーブロック１６はフレーム毎（例えば、１０ミリ秒）にこの計算を実行し、統計処理ブロック１４に入力としての結果を提供する。

【００５３】

ここで、図３を参照すると、本発明の特徴抽出システムにおける使用に適した例示の複素バンドパスフィルタ３０’が記載される。フィルタ３０’は、加算器３１、乗算器３２、および１サンプル遅延素子３３を含む。乗算器３２は、１サンプル遅延出力Ｙに複素係数Ｇを掛けて、結果として、入力信号Ｘに足されて、出力信号Ｙを生成する。

【００５４】

本発明の特徴抽出システムの代替の実施形態は、図４に関して記載される。図４の実施形態は、図２の実施形態に類似しており、フィルタリングステージ１０、統計処理ステージ１４、およびエネルギーステージ１６を含み、上述されるように実質的に動作する。しかし、図４の実施形態は、特徴抽出ステージ１２’が特徴抽出システム１２a’内でさらなるの回路部を含むという前述される実施形態とは異なり、特徴ベクトルはさらなる情報を含む。

【００５５】

例えば、特徴抽出ステージ１２ａ’は、バンドパスフィルタ３０１－４１のバンクおよび共役ブロック４０１－４０を含む。各バンドパスフィルタの出力は、乗算器４２１－４０によるより低い隣接したバンドパスフィルタの出力の共役と組み合わされ、計算ブロック４６１－４０は、ベクトルが２０の代わりに４０の素子の長さを有することを除いて、上述されるようにベクトルＡおよびＦを計算する。ＤＣＴ５０および５４、ならびに、特徴拡張コンポーネント１２ｂ’のデルタブロック５２および５６は、４０の素子入力ベクトルを承認して、４０の素子ベクトルを統計処理ブロック１４に出力する。バンドパスフィルタ３０１－４１が例えば対数スケールまたはメルスケールといった非線型周波数スケールによって構成される場合、図４に示される構成は適用できないことは理解される。

【００５６】

本発明は、意図された音声処理アプリケーションによってバンドパスフィルタ３０の任意の数、ならびに、共役ブロック４０、乗算器４２、ロウパスフィルタ４４、およびブロック４６の対応する数を含み、各ロウパスフィルタの出力信号ＡおよびＦを提供し得る。特徴抽出ステージを含む。さらに、信号ＡおよびＦは、重み付された様態で組み合わされ得、信号の一部のみが使用され得る。例えば、一周波数ドメインにおける振幅信号、ならびに、別の周波数ドメインにおける振幅と周波数との組み合わせのみを用いることは利点があり得る。

【００５７】

本発明の好適な例示の実施形態が上述され、多様な変化および変更が本発明から逸脱することなくなされることは、当業者によって理解され、本発明の実際の意図および範囲内に含まれるような全ての変化および変更をカバーすることが、添付の特許請求の範囲において意図される。

【図面の簡単な説明】

【００５８】

【図１】図１は、本発明の発話特徴抽出システムを組み込む例示的な発話認識システムのブロック図である。

【図２】図２は、図１の発話認識システムの詳細なブロック図である。

【図３】図３は、本発明の特徴抽出システムをインプリメントするために適するバンドパスフィルタの詳細なブロック図である。

【図４】図４は、本発明の代替的な発話特徴抽出システムを含む発話認識の代替的な実施形態の詳細なブロック図である。

【図５】図５は、線形的な周波数スケールに従って配置されたバンドパスフィルタ周波数を示すグラフである。

【図６】図６は、対数周波数スケールに従って配置されたバンドパスフィルタのペアを示すグラフである。

【図７】図７は、メルスケールに従って配置されたバンドパスフィルタのペアを示すグラフである。

【図８】図８は、線形および対数周波数スケールの組み合わせに従って配置されたバンドパスフィルタ周波数を示すグラフである。

【図1】

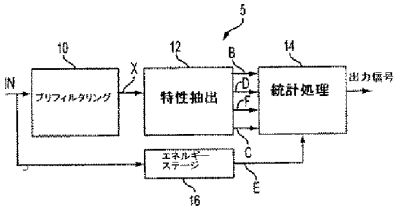


FIG. 1

【図2】

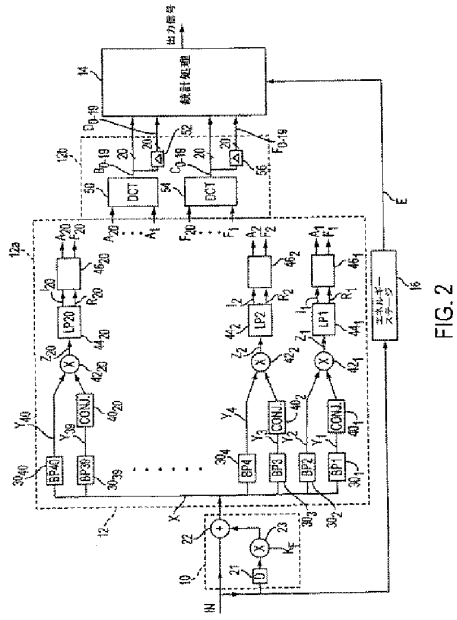


FIG. 2

【図4】

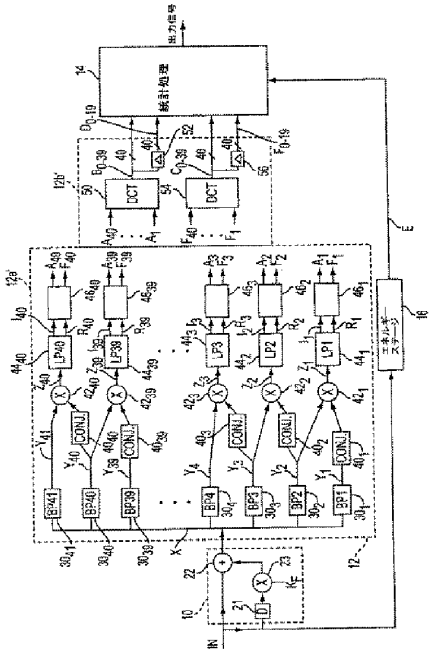


FIG. 4

【図5】

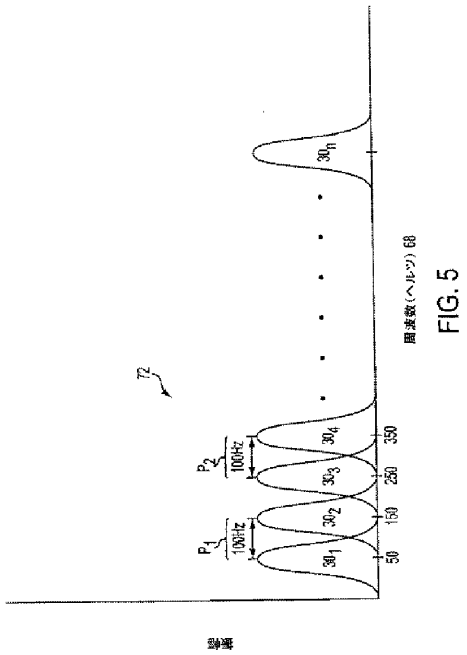


FIG. 5

【図6】

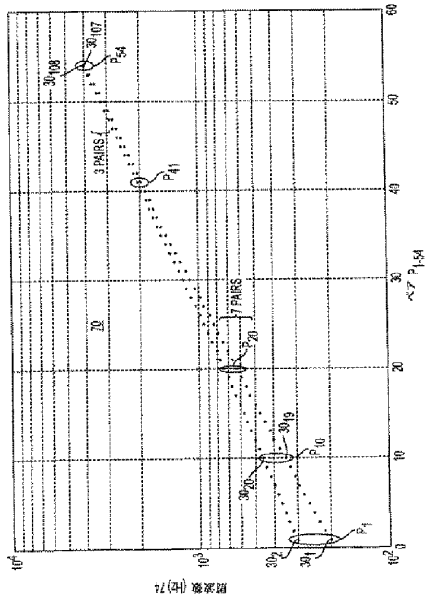


FIG. 6

【図7】

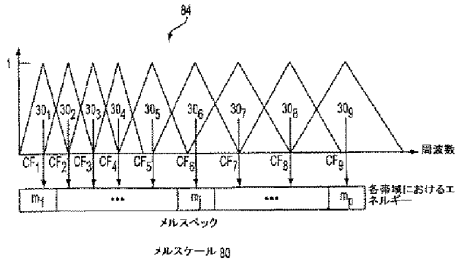


FIG. 7

【図8】

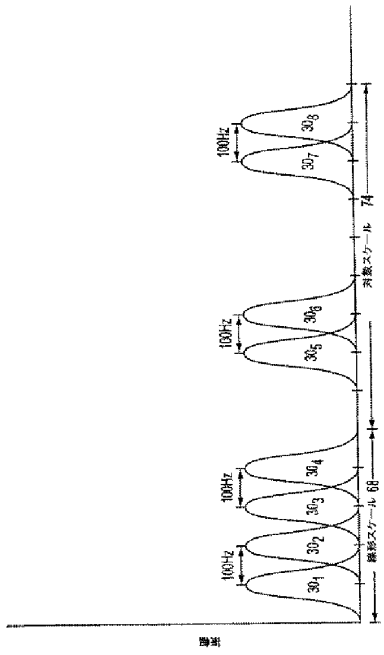

100Hz
100Hz
100Hz
100Hz
30_1
30_2
30_3
30_4
30_5
30_6
30_7
30_8
68
74
周波数スケール
周波数スケール
振幅

FIG. 8

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2002 (27.12.2002)

PCT

(10) International Publication Number
WO 02/103676 A1

(51) International Patent Classification⁷: G10L 15/02

(21) International Application Number: PCT/US02/19182

(22) International Filing Date: 14 June 2002 (14.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/882,744 15 June 2001 (15.06.2001) US

(71) Applicant and
(72) Inventor: BRANDMAN, Vigal [US/US]; 4120 Donald Drive, Palo Alto, CA 94306 (US).

(74) Agents: PISANO, Nicola, A. et al.; c/o Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

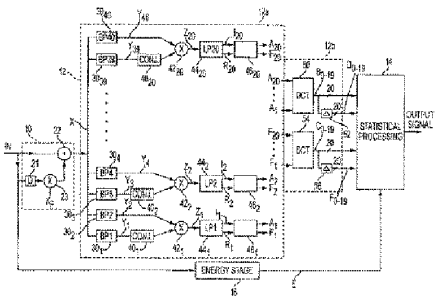
(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH FEATURE EXTRACTION SYSTEM



(57) Abstract: A system (12a) suitable for use in a speech recognition system or other voice processing system (14) extracts features related to the frequency (46, F) and amplitude (46, A) characteristics of an input speech signal (IN) using a plurality of complex band pass filters (30) by processing the outputs of adjacent bandpass filters (42).

WO 02/103676 A1

SPEECH FEATURE EXTRACTION SYSTEM

Background of the Invention

This invention relates to a speech feature extraction system for use in speech recognition, voice identification or voice authentication systems. More specifically, this invention relates to a speech feature extraction system that can be used to create a speech recognition system or other speech processing system with a reduced error rate.

Generally, a speech recognition system is an apparatus that attempts to identify spoken words by analyzing the speaker's voice signal. Speech is converted into an electronic form from which features are extracted. The system then attempts to match a sequence of features to previously stored sequence of models associated with known speech units. When a sequence of features corresponds to a sequence of models in accordance with specified rules, the corresponding words are deemed to be recognized by the speech recognition system.

However, background sounds such as radios, car noise, or other nearby speakers can make it difficult to extract useful features from the speech. In addition, a change in the ambient conditions such as the use of a

different microphone, telephone handset or telephone line can interfere with system performance. Also, a speaker's distance from the microphone, differences between speakers, changes in speaker intonation or emphasis, and even a speaker's health can adversely impact system performance. For a further description of some of these problems, see Richard A. Quinnell, "Speech Recognition: No Longer a Dream, But Still a Challenge," EDN Magazine, January 19, 1995, p. 41-46.

In most speech recognition systems, the speech features are extracted by cepstral analysis, which generally involves measuring the energy in specific frequency bands. The product of that analysis reflects the amplitude of the signal in those bands. Analysis of these amplitude changes over successive time periods can be modeled as an amplitude modulated signal.

Whereas the human ear is a sensitive to frequency modulation as well as amplitude modulation in received speech signals, this frequency modulated content is only partially reflected in systems that perform cepstral analysis.

Accordingly, it would be desirable to provide a speech feature extraction system capable of capturing the frequency modulation characteristics of speech, as well as previously known amplitude modulation characteristics.

It also would be desirable to provide speech recognition and other speech processing systems that incorporate feature extraction systems that provide information on frequency modulation characteristics of the input speech signal.

Summary of the Invention

In view of the foregoing, it is an object of the present invention to provide a speech feature extraction system capable of capturing the frequency modulation characteristics of speech, as well as previously known amplitude modulation characteristics.

It also is an object of this invention to provide speech recognition and other speech processing systems that incorporate feature extraction systems that provide information on frequency modulation characteristics of the input speech signal.

The present invention provides a speech feature extraction system that reflects frequency modulation characteristics of speech as well as amplitude characteristics. This is done by a feature extraction stage which, in one embodiment, includes a plurality of complex band pass filters arranged in adjacent frequency bands according to a linear frequency scale ("linear scale"). The plurality of complex band pass filters are divided into pairs. A pair includes two complex band pass filters in adjacent frequency bands. For every pair, the output of the filter in the higher frequency band ("primary frequency") is multiplied by the conjugate of the output of the filter in the lower frequency band ("secondary filter"). The resulting signal is low pass filtered.

In another embodiment, the feature extraction phase includes a plurality of complex band pass filters arranged according to a logarithmic (or exponential) frequency scale ("log scale"). The primary filters of the filter pairs are centered at various frequencies

along the log scale. The secondary filter corresponding to the primary filter of each pair is centered at a predetermined frequency below the primary filter. For every pair, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

In yet another embodiment, the plurality of band pass filters are arranged according to a mel-scale. The primary filters of the filter pairs are centered at various frequencies along the mel-scale. The secondary filter corresponding to the primary filter of each pair is centered at a predetermined frequency below the primary filter. For every pair, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

In still another embodiment, the plurality of band pass filters are arranged according to a combination of the linear and log scale embodiments mentioned above. A portion of the pairs of the band pass filters are arranged in adjacent frequency bands according to a linear scale. For each of these pairs, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

The primary filters of the remaining pairs of band pass filters are centered at various frequencies along the log scale and the secondary filters corresponding to the primary filters are centered a predetermined frequency below the primary filters. For

every pair, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

For the embodiments described above, each of the low pass filter outputs is processed to compute two components: a FM component that is substantially sensitive to the frequency of the signal passed by the adjacent band pass filters from which the low pass filter output was generated, and an AM component that is substantially sensitive to the amplitude of the signal passed by the adjacent band pass filters. The FM component reflects the difference in the phase of the outputs of the adjacent band pass filters used to generate the lowpass filter output.

The AM and FM components are then processed using known feature enhancement techniques, such as discrete cosine transform, mel-scale translation, mean normalization, delta and acceleration analysis, linear discriminant analysis and principal component analysis, to generate speech features suitable for statistical processing or other recognition or identification methods. In an alternative embodiment, the plurality of complex band pass filters can be implemented using a Fast Fourier Transform (FFT) of the speech signal or other digital signal processing (DSP) techniques.

In addition, the methods and apparatus of the present invention may be used in addition to performing cepstral analysis in a speech recognition system.

Brief Description of the Drawings

The above and other objects and advantages of the present invention will be apparent upon consideration of the following detailed description, taken in conjunction with the accompanying drawings, in which like reference characters refer to like parts throughout, and in which:

FIG. 1 is a block diagram of an illustrative speech recognition system incorporating the speech feature extraction system of the present invention;

FIG. 2 is a detailed block diagram of the speech recognition system of FIG. 1; and

FIG. 3 is a detailed block diagram of a band pass filter suitable for implementing the feature extraction system of the present invention; and

FIG. 4 is a detailed block diagram of an alternative embodiment of a speech recognition including an alternative speech feature extraction system of the present invention; and

FIG. 5 illustrates is a graph showing band pass filter frequencies spaced according to a linear frequency scale; and

FIG. 6 is a graph showing pairs of band pass filters spaced according to a logarithmic frequency scale; and

FIG. 7 is a graph showing band pass filter frequency pairs spaced according to a mel-scale; and

FIG. 8 is a graph showing band pass filter frequencies spaced according to a combination of linear and logarithmic frequency scales.

Detailed Description of the Invention

Referring to FIG. 1, a generalized depiction of illustrative speech recognition system 5 is described that incorporates the speech extraction system of the present invention. As will be apparent to one of ordinary skill in the art, the speech feature extraction system of the present invention also may be used in speaker identification, authentication and other voice processing systems.

System 5 illustratively includes four stages: pre-filtering stage 10, feature extraction stage 12, statistical processing stage 14, and energy stage 16.

Pre-filtering stage 10, statistical processing stage 14 and energy stage 16 employ speech processing techniques known in the art and do not form part of the present invention. Feature extraction stage 12 incorporates the speech feature extraction system of the present invention, and further includes feature enhancement techniques which are known in the art, as described hereinafter.

Audio speech signal is converted into an electrical signal by a microphone, telephone receiver or other device, and provided as an input speech signal to system 5. In a preferred embodiment of the present invention, the electrical signal is sampled or digitized to provide a digital signal (IN) representative of the audio speech. Pre-filtering stage 10 amplifies the high frequency components of audio signal IN, and the prefiltered signal is then provided to feature extraction stage 12.

Feature extraction stage 12 processes pre-filtered signal X to generate a sequence of feature vectors related to characteristics of input signal IN that may be useful for speech recognition. The output of feature extraction stage 12 is used by statistical processing stage 14 which compares the sequence of feature vectors to predefined statistical models to identify words or other speech units in the input signal IN. The feature vectors are compared to the models using known techniques, such as the Hidden Markov Model (HMM) described in Jelinek, "Statistical Methods for Speech Recognition," The MIT Press, 1997, pp. 13-37. The output of statistical processing stage 14 is the recognized word, or other suitable output depending upon the specific application.

Statistical processing at stage 14 may be performed locally, or at a remote location relative to where the processing of stages 10, 12, and 16 are performed. For example, the sequence of feature vectors may be transmitted to a remote server for statistical processing.

The illustrative speech recognition system of FIG. 1 preferably also includes energy stage 16 which provides an output signal indicative of the total energy in a frame of input signal IN. Statistical processing stage 14 may use this total energy information to provide improved recognition of speech contained in the input signal.

Referring now to FIG. 2, pre-filtering stage 10 and feature extraction stage 12 are described in greater detail. Pre-filtering stage 10 is a high pass filter

that amplifies high frequency components of the input signal. Pre-filtering stage 10 comprises one-sample delay element 21, multiplier 23 and adder 24. Multiplier 23 multiplies the one-sample delayed signal by constant K_E, which typically has a value of -0.97. The output of pre-filtering stage 10, X, is input at the sampling rate into a bank of band pass filters 30_1, 30_2, ... 30_n.

In one embodiment, the band pass filters 30_1, 30_2, ... 30_n are positioned in adjacent frequency bands. The spacing of the band pass filters 30_1, 30_2, ... 30_n is done according to a linear frequency scale ("linear scale") 68 as shown in graph 72 of FIG. 5. The term "linear frequency scale" is used in this specification in accordance with its ordinary and accustomed meaning, i.e., the actual frequency divisions are uniformly spaced. The plurality of complex band pass filters 30_1, 30_2, ... 30_n are divided into pairs P_{1-2}. A pair (P_1 or P_2) includes two complex band pass filters (30_{1-2} or 30_{3-4}) respectively in adjacent frequency bands. For every pair (P_1 or P_2), the output of the filter in the higher frequency band (30_2 or 30_4) (referred to hereinafter as the "primary filter") is multiplied by the conjugate of the output of the filter in the lower frequency band (30_1 or 30_3) (referred to hereinafter as the "secondary filter"). The resulting signal is low pass filtered.

The number of band pass filters 30_1, 30_2, ... 30_n and width of the frequency bands preferably are selected according to the application for the speech processing system. For example, a system useful in telephony applications preferably will employ about forty band pass filters 30_1, 30_2, ... 30_n having center

frequencies approximately 100 Hz apart. For example, filter 30_1 may have a center frequency of 50 Hz, filter 30_2 may have a center frequency of 150 Hz, filter 30_3 may have a center frequency of 250 Hz, and so on, so that the center frequency of filter 30_{40} is 3950 Hz. The bandwidth of each filter may be several hundred Hertz.

In another embodiment, as illustrated in the graph 70 of FIG. 6, the band pass filters 30_1, 30_2, ... 30_{108} are arranged according to a non-linear frequency scale such as a logarithmic (or exponential) frequency scale 74 ("log scale"). The term logarithmic frequency scale is used in this specification according to its ordinary and accustomed meaning.

Empirical evidence suggests that using log scale 74 of FIG. 6 instead of linear scale 68 of FIG. 5 improves voice recognition performance. That is so because the human ear resolves frequencies non-linearly across the audio spectrum. Another advantage using log scale 74 instead of linear scale 68 is that log scale 74 can cover a wider range of frequency spectrum without using additional band pass filters 30_1, 30_2, ... 30_n.

Pairs P_{1-54} of band pass filters 30_{1-108} are spaced according to log scale 74. Pair P_1 includes filters 30_1 and 30_2, pair P_{10} includes filters 30_{19} and 30_{20}, and pair P_{54} includes filters 30_{107} and 30_{108}. In this arrangement, filters 30_2, 30_{20} and 30_{108} are the primary filters and filters 30_1, 30_{19} and 30_{107} are the secondary filters.

In one preferred embodiment, primary filters 30_2, 30_{20}...30_{108} are centered at various frequencies along log scale 74, while secondary filters 30_1, 30_3...30_{107} are

centered 100 hertz (Hz) below corresponding primary filters 30_2, 30_4...30_{108} respectively. An exemplary MATLAB code to generate graph 70 of FIG. 6 is shown below.

```
v=(2.^([26.715:0.25:40]/3.345)); % Generate center frequencies for bandpass filter pairs
f(2:2:2*length(v))=v+50;         % Primary filters center frequencies
f(1:2:2*length(v))=v-50;         % Secondary filters center frequencies
semilogy([v'+50 v'-50],'.');grid % Plot center frequencies on a logarithmic scale'
```

In another embodiment, the center frequencies of primary filters 30_2,30_{20}...30_{108} and secondary filters 30_1, 30_3...30_{107} may be placed using separate and independent algorithms, while ensuring that secondary filters 30_1, 30_3...30_{107} are centered 100 hertz (Hz) below their corresponding primary filters 30_2, 30_4...30_{108}, respectively.

In one embodiment, band pass filters 30_{1-108} are of triangular shape. In other embodiments, band pass filters 30_{1-108} can be of various shapes depending on the requirements of the particular voice recognition systems.

Log scale 74 is shown to range from 0-4000 Hz. For every pair P_{1-54}, the output of the primary filter 30_2, 30_4...30_{108} is multiplied by the conjugate of the output of secondary filter 30_1, 30_3...30_{107}. The resulting signal is low pass filtered.

The pairs P_{1-54} are arranged such that the lower frequencies include a higher concentration of pairs P_{1-54} than the higher frequencies. For example, there are 7 pairs (P_{16-22}) in the frequency range of 500-1000 Hz while there are only 3 pairs (P_{52-54}) in the frequency range of

3000-3500 Hz. Thus, although there is over sampling at the lower frequencies, this embodiment also performs at least some sampling at the higher frequencies. The concentration of pairs P_{1-84} along log scale 74 can be varied depending on the needs of a particular voice recognition system.

As will be apparent to one of ordinary skill in the art of digital signal processing design, the band pass filters of the preceding embodiments may be implemented using any of a number of software or hardware techniques. For example, the plurality of complex filters may be implemented using a Fast Fourier Transform (FFT), Chirp-Z transform, other frequency domain analysis techniques.

In an alternative embodiment, depicted in FIG. 7, band pass filters 30_1, 30_2, ... 30_n are arranged according to a non-linear frequency scale, such as mel-scale 80. Mel-scale 80 is well known in the art of voice recognition systems and is typically defined by the equation,

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right)$$

where f represents frequency according to the linear scale 68 and Mel(f) represents its corresponding mel-scale 80 frequency.

FIG. 7 illustrates one embodiment of a graph 84 showing band pass filters 30_{1-9} spaced according to mel-scale 80. The center frequencies (CF_{1-9}) are the Mel(f) values calculated by using the above equation. Typically, filters 30_{1-9} are spread over the whole

frequency range from zero up to the Nyquist frequency. In one embodiment, filters 30_{1-9} have the same band width. In another embodiment, filters 30_{1-9} may have different bandwidths.

In yet another embodiment, depicted in FIG. 8, the band pass filters 30_{1-8} are spaced according to a combination of linear 68 and non-linear 74 frequency scales. Band pass filters 30_{1-4} (P_{1-2}) are arranged in adjacent frequency bands according to linear frequency 68.

Primary filters 30_6 and 30_8 are centered along log scale 74. Secondary filters 30_5 and 30_7 are centered at frequencies of 100 Hz below the center frequencies of 30_6 and 30_8 respectively. For each of these pairs (P_1 or P_2), the output of the primary filter (30_6 or 30_8) is multiplied by the conjugate of the output of the secondary filter (30_5 and 30_7), respectively and the resulting signal is low pass filtered.

Referring again to FIG. 2, blocks 40_{1-20} provide the complex conjugate of the output signal of band pass filter 30_1, 30_3, ... 30_{n-1}. Multiplier blocks 42_{1-20} multiply the complex conjugates by the outputs of an adjacent higher frequency band pass filter 30_2, 30_4, 30_6, ... 30_{40} to provide output signals Z_{1-20}. Output signals Z_{1-20} then are passed through a series of low pass filters 44_{1-20}. The outputs of the low pass filters typically are generated only at the feature frame rate. For example, at a input speech sampling rate of 8 kHz, the output of the low pass filters is only computed at a feature frame rate of once every 10 msec.

Each output of low pass filters 44_{1-20} is a complex signal having real component R and imaginary component I. Blocks 46_{1-20} process the real and imaginary components of the low pass filter outputs to provide output signals A_{1-20} and F_{1-20} as shown in equations (1) and (2):

$$A_i = \log R_i \quad (1)$$

$$F_i = \frac{I_i}{\sqrt{R_i^2 + I_i^2}} \quad (2)$$

wherein R_i and I_i are the real and imaginary components of the corresponding low pass filter output. Output signals A_i are a function of the amplitude of the low pass filter output and signals F_i are a function of the frequency of the signal passed by the adjacent band pass filters from which the low pass filter output was generated. By computing two sets of signals that are indicative of the amplitude and frequency of the input signal, the speech recognition system incorporating the speech feature extraction system of the present invention is expected to provide reduced error rate.

The amplitude and frequency signals A_{1-20} and F_{1-20} then are processed using conventional feature enhancement techniques in feature enhancement component 12b, using, for example, discrete cosine transform, mel-scale translation, mean normalization, delta and acceleration analysis, linear discriminant analysis and principal component analysis techniques that are per se known in the art. A preferred embodiment of a speech

recognition system of the present invention incorporating the speech extraction system of the present invention employs a discrete cosine transform and delta features technique, as described hereinafter.

Still referring to FIG. 2, feature enhancement component 12b receives output signals A_{1-20} and F_{1-20}, and processes those signals using discrete cosine transform (DCT) blocks 50 and 54, respectively. DCTs 50 and 54 attempt to diagonalize the co-variance matrix of signals A_{1-20} and F_{1-20}. This helps to uncorrelate the features in output signals B_{0-19} of DCT 50 and output signals C_{0-19} of DCT 54. Each set of output signals B_{0-19} and C_{0-19} then are input into statistical processing stage 14. The function performed by DCT 50 on input signals A_{1-20} to provide output signals B_{0-19} is shown by equation (3), and the function performed by DCT 54 on input signals F_{1-20} to provide output signals C_{0-19} is shown by equation (4).

$$B_r = D(r) \sum_{n=0}^{N-1} A_{n+1} \cdot \cos\frac{(2n+1)\pi r}{2N} \qquad (3)$$

$$C_r = D(r) \sum_{n=0}^{N-1} F_{n+1} \cdot \cos\frac{(2n+1)\pi r}{2N} \qquad (4)$$

In equations (3) and (4), N equals the length of the input signal vectors A and F (e.g., N = 20 in FIG. 2), n is an index from 0 to N - 1 (e.g., n = 0 to 19 in the embodiment of FIG. 2), and r is the index of output signals B and C (e.g., r = 0 to 19 in the embodiment of FIG. 2). Thus, for each vector output signal B_r, each vector of input signals A_{1-20} are multiplied by a cosine function and D(r) and summed

together as shown in equation (3). For each vector output signal C_z, each vector of input signals S_{1-20} are multiplied by a cosine function and D(r) and summed together as shown in equation (4). D(r) are coefficients that are given by the following equations:

$$D(r)=\frac{1}{\sqrt{N}} \quad for\ r=0 \tag{5}$$

$$D(r)=\frac{\sqrt{2}}{\sqrt{N}} \quad for\ r>0 \tag{6}$$

Output signals E_{0-19} and C_{0-19} also are input into delta blocks 52 and 56, respectively. Each of delta blocks 52 and 56 takes the difference between measurements of feature vector values between consecutive feature frames and this difference may be used to enhance speech recognition performance. Several difference formulas may be used by delta blocks 52 and 56, as are known in the art. For example, delta blocks 52 and 56 may take the difference between two consecutive feature frames. The output signals of delta blocks 52 and 56 are input into statistical processing stage 14.

Energy stage 16 of FIG. 2 is a previously known technique for computing the logarithm of the total energy (represented by E) of each frame of input speech signal IN, according to the following equation:

$$E(iT)=\log\left(\frac{\sum_{i=0}^{K-1} I\,N_{i0-qT}^{2}}{K}\right) \tag{7}$$

Equation 7 shows that energy block 16 takes the sum of the squares of the values of the input signal IN during the previous K sampling intervals (e.g., K = 220, T = 1/8000 seconds), divides the sum by K, and takes the logarithm of the final result. Energy block 16 performs this calculation every frame (e.g., 10 msec), and provides the result as an input to statistical processing block 14.

Referring now to FIG. 3, illustrative complex bandpass filter 30' suitable for use in the feature extraction system of the present invention is described. Filter 30' comprises adder 31, multiplier 32 and one-sample delay element 33. Multiplier 32 multiples the one-sample delayed output Y by complex coefficient G and the resultant is added to the input signal X to generate an output signal Y.

An alternative embodiment of the feature extraction system of the present invention is described with respect to FIG. 4. The embodiment of FIG. 4 is similar to the embodiment of FIG. 2 and includes pre-filtering stage 10, statistical processing stage 14, and energy stage 16 the operate substantially as described above. However, the embodiment of FIG. 4 differs from the previously described embodiment in that feature extraction stage 12' includes additional circuitry within feature extraction system 12a, so that the feature vectors include additional information.

For example, feature extraction stage 12a' includes a bank of 41 band pass filters 301-41 and conjugate blocks 401-40. The output of each band pass filter is combined with the conjugate of the output of a

lower adjacent band pass filter by multipliers 421-40. Low pass filters 441-40, and computation blocks 461-40 compute vectors A and F as described above, except that the vectors have a length of forty elements instead of twenty. DCTs 50 and 54, and delta blocks 52 and 56 of feature enhancement component 12b' each accept the forty element input vectors and output forty element vectors to statistical processing block 14. It is understood that the arrangement illustrated in FIG. 4 is not applicable if the band pass filters 301-41 arranged according to a non-linear frequency scale such as a log scale or a mel-scale.

The present invention includes feature extraction stages which may include any number of band pass filters 30, depending upon the intended voice processing application, and corresponding numbers of conjugate blocks 40, multipliers 42, low pass filters 44 and blocks 46 to provide output signals A and F for each low pass filter. In addition, signals A and F may be combined in a weighted fashion or only part of the signals may be used. For example, it may be advantageous to use only the amplitude signals in one frequency domain, and a combination of the amplitude and frequency signals in another.

While preferred illustrative embodiments of the invention are described above, it will be apparent to one skilled in the art that various changes and modifications may be made therein without therein without departing from the invention, and it is intended in the appended claims to cover all such changes and modifications which fall within the true spirit and scope of the invention.

What is claimed is:

1. Apparatus for use in a speech processing system for extracting features from an input speech signal having frequency and amplitude characteristics, the apparatus comprising:

first and second band pass filters adapted to receive the input speech signal and providing respectively, first and second signals;

a conjugate circuit coupled to the second band pass filter and providing a third signal that is the conjugate of the second signal;

a multiplier coupled to the first band pass filter and to the conjugate circuit and providing a fourth signal that is the product of the first and third signals; and

filter means coupled to the multiplier for providing a fifth signal related to the frequency characteristics of the input signal, and a sixth signal related to amplitude characteristics of the input signal.

2. The apparatus of claim 1 wherein the filter means includes a low pass filter.

3. The apparatus of claim 1 further comprising a high pass filter coupled between the input signal and the first and second band pass filters.

4. The apparatus of claim 2 wherein the low pass filter provides a signal having a real part, R, and an imaginary part, I, and the filter means includes circuitry for calculating:

A = log R and

F = I / Sqrt{R2 + I2}.

5. The apparatus of claim 1 wherein the first and second band pass filters each comprises a delay element, a complex multiplier and an adder.

6. The apparatus of claim 5 wherein the adder receives the input speech signals and outputs an output signal, the adder adding the input speech signal and a signal comprising a delayed sample of the output signal multiplied by the complex coefficient.

7. The apparatus of claim 4 further comprising feature enhancement circuitry.

8. The apparatus of claim 7 further comprising circuitry for statistically processing an output of the feature enhancement circuitry to compare that output to a plurality of predetermined models.

9. The apparatus of claim 7 wherein the feature enhancement circuitry comprises circuitry for computing a discrete cosine transform of A and F.

10. The apparatus of claim 7 wherein the feature enhancement circuitry further comprises circuitry for calculating a difference between successive discrete cosine transforms of A and for calculating a difference between successive discrete cosine transforms of F.

11. The apparatus of claim 1 further comprising:

a transducer for converting sound into an electrical signal; and

a sampler for converting the electrical signal into a digital signal, wherein the input speech signal comprises the digital signal.

12. The apparatus of claim 1 further comprising a digital processor, wherein the first and second band pass filters, the conjugate circuit, the multiplier, and the filter means comprise algorithms adapted to be executed on the digital processor.

13. The apparatus of claim 1, wherein a frequency of the first bandpass filter is selected according to a logarithmic scale and a frequency of the second bandpass filter is selected at a predetermined frequency below the frequency of the first bandpass filter.

14. The apparatus of claim 13, wherein a center frequency of the second filter is one hundred hertz less than a center frequency of the first filter.

15. The apparatus of claim 13, wherein the bands of the first and second filters overlap.

16. The apparatus of claim 13, wherein the bandwidths of the first and second filters are the same.

17. The apparatus of claim 1, wherein a frequency of the first bandpass filter is selected according to a mel-scale and a frequency of the second bandpass filter is selected at a predetermined frequency below the frequency of the first bandpass filter.

18. The apparatus of claim 17, wherein a center frequency of the second filter is one hundred hertz less than a center frequency of the first filter.

19. The apparatus of claim 17, wherein the bands of the first and second filters overlap.

20. The apparatus of claim 17, wherein the mel-scale is defined by the equation

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right).$$

21. The apparatus of claim 17, wherein the bandwidths of the first and second filters are the same.

22. The apparatus of claim 1, wherein the first and second band pass filters are implemented using a Fast Fourier Transform.

23. A method for extracting features from an input speech signal for use in a speech processing device, the method comprising:

separating the input signal into a first signal in a first frequency band and a second signal in a second frequency band;

providing a conjugate of the first signal;

multiplying the conjugate of the first signal with the second signal to provide a third signal; and

processing the third signal to generate a frequency component related to frequency features in the input speech signal and an amplitude component related to amplitude features in the input speech signal.

24. The method of claim 23 further comprising high pass filtering the input speech signal prior to processing the input signal to generate the first and second signals.

25. The method of claim 23 wherein processing the third signal further comprises low pass filtering the third signal.

26. The method of claim 25 wherein low pass filtering the third signal provides a signal having a real part, R, and an imaginary part, I, and processing the third signal includes calculating:

A = log R and

F = I / Sqrt(R2 + I2).

27. The method of claim 26 further comprising processing the third signal using feature enhancement techniques to generate a series of feature vectors.

28. The method of claim 27 further comprising statistically processing the series of feature vectors by comparing the series of feature vectors to a plurality of predetermined models.

29. The method of claim 27 further comprising calculating a discrete cosine transform of A and F.

30. The method of claim 29 further comprising calculating a difference between successive discrete cosine transforms of A and between successive discrete cosine transforms of S.

31. The method of claim 30 further comprising statistically processing the differences between successive discrete cosine transforms to compare the differences to a plurality of predetermined patterns.

32. The method of claim 23 further comprising:

converting a sound into an electrical signal; and

sampling the electrical signal to provide a digital signal, wherein the input speech signal comprises the digital signal.

33. A method of claim 23, comprising selecting a frequency of the first frequency band according to a logarithmic scale and selecting a frequency of the second frequency band at a predetermined frequency below the frequency of the first frequency band.

34. The method of claim 33, wherein a center frequency of the second frequency band is one hundred hertz less than a center frequency of the first frequency band.

35. The method of claim 33, comprising overlapping the first and second frequency bands.

36. The method of claim 23, comprising selecting a frequency of the first frequency band according to a mal-scale and selecting the frequency of the second frequency band at a predetermined frequency below the frequency of the the first frequency band.

37. The apparatus of claim 36, wherein a center frequency of the second frequency band is one hundred hertz less than a center frequency of the first frequency band.

38. The apparatus of claim 36, comprising overlapping the first and second frequency bands.

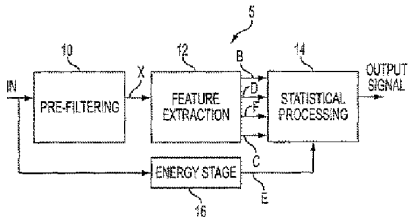

5
10
12
14
B
IN
X
PRE-FILTERING
FEATURE EXTRACTION
STATISTICAL PROCESSING
OUTPUT SIGNAL
D
F
C
ENERGY STAGE
16
E

FIG. 1

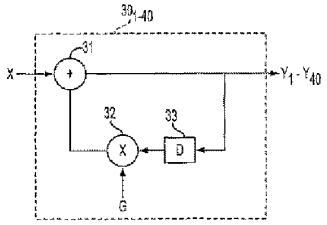

39
40
31
X
+
$Y_1 - Y_{40}$
32
33
X
D
G

FIG. 3

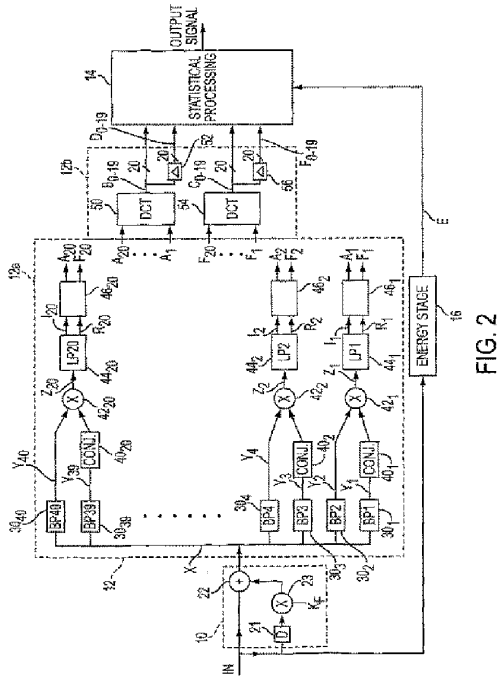

OUTPUT SIGNAL
STATISTICAL PROCESSING
14
DCT
DCT
ENERGY STAGE
16
LP20
LP2
LP1
CONJ
BP40
BP39
BP4
BP3
BP2
BP1
IN

FIG. 2

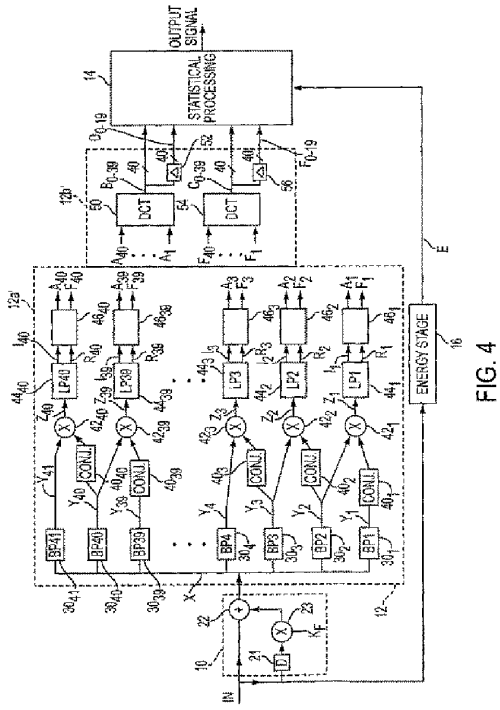

OUTPUT SIGNAL
STATISTICAL PROCESSING
DCT
DCT
ENERGY STAGE
CONJ.
IN

FIG. 4

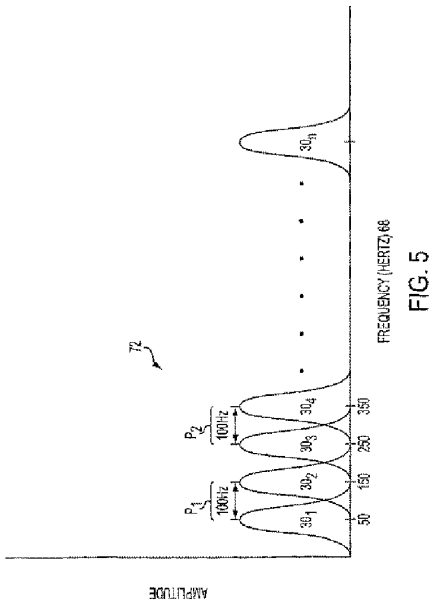

AMPLITUDE
FREQUENCY (HERTZ) 68
72
P_1
100Hz
P_2
100Hz
30_1
30_2
30_3
30_4
30_n
50
150
250
350

FIG. 5

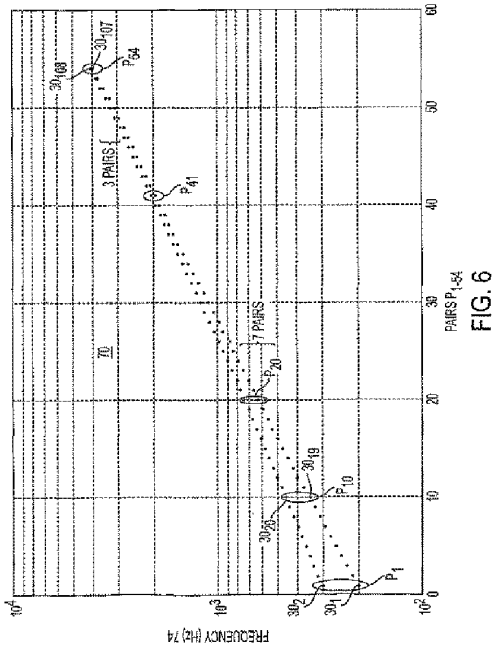

3 PAIRS
7 PAIRS
P41
P64
P20
P10
P1
30168
30107
30119
30125
30-2
30-1
FREQUENCY (Hz) 74
PAIRS P1-64

FIG. 6

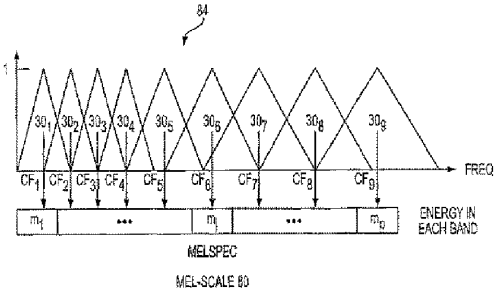

84
1
30_1 30_2 30_3 30_4 30_5 30_6 30_7 30_8 30_9
FREQ
CF_1 CF_2 CF_3 CF_4 CF_5 CF_6 CF_7 CF_8 CF_9
m_1 ••• m_i ••• m_Q
ENERGY IN EACH BAND
MELSPEC
MEL-SCALE 80

FIG. 7

11/

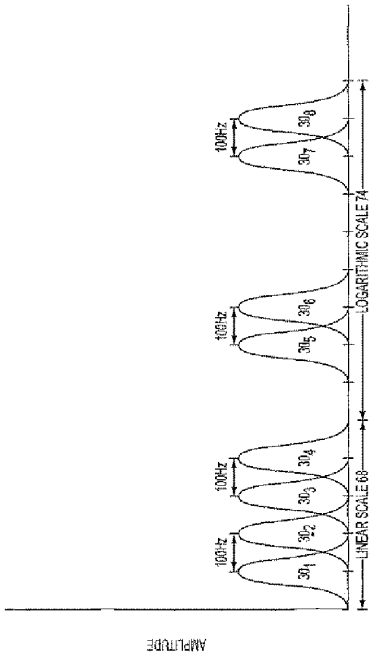

AMPLITUDE
100Hz
100Hz
100Hz
LINEAR SCALE 68
LOGARITHMIC SCALE 74
30_1
30_2
30_3
30_4
30_5
30_6
30_7
30_8

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/19182

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G10L 15/02
US CL : 704/243, 235, 206; 708/300
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
U.S. : 704/243, 235, 206, 231, 268; 708/300-323; 370/484; 375/285.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,300,229 A (HIROSAKI) 10 November 1981 (10.11.1981), ALL	1-38
A	US 4,729,112 A (MILLAR) 01 March 1988 (01.03.1988), ALL	1-38
A	US 4,669,216 A (CLAASEN et al.) 21 April 1987 (21.04.1987), ALL	1-38
A	US 4,221,934 A (SCHIFF) 09 September 1980 (09.09.1980), ALL	1-38

☐ Further documents are listed in the continuation of Box C.　☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 September 2002 (08.09.2002)	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer David D. Knepper Telephone No. (703) 306-0377

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

PCT/US02/19182

Continuation of B. FIELDS SEARCHED Item 3:
EAST: (speech OR voice) AND (Frequency AND Amplitude) AND (conjugate OR imaginary) AND ((band NEAR1 pass NEAR1 filter) OR EPD)

Form PCT/ISA/210 (second sheet) (July 1998)

(81)指定国　　AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZM,ZW

Ｆターム(参考) 5D015 AA01　BB02　CC03　CC11　FF04

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2002 (27.12.2002)

PCT

(10) International Publication Number
WO 02/103676 A1

(51) International Patent Classification[7]: G10L 15/02

(21) International Application Number: PCT/US02/19182

(22) International Filing Date: 14 June 2002 (14.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/882,744 15 June 2001 (15.06.2001) US

(71) Applicant and
(72) Inventor: **BRANDMAN, Yigal** [US/US]; 4120 Donald Drive, Palo Alto, CA 94306 (US).

(74) Agents: **PISANO, Nicola, A.** et al.; c/o Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

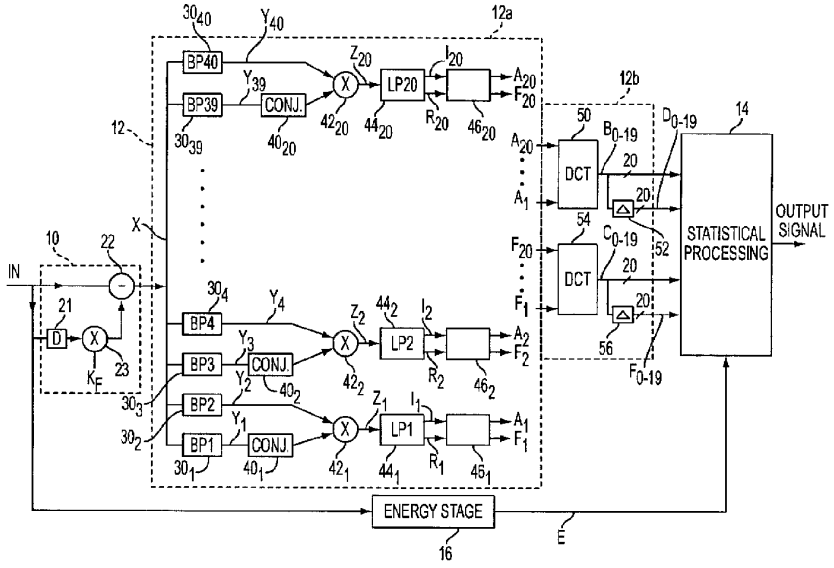
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH FEATURE EXTRACTION SYSTEM



(57) Abstract: A system (12a) suitable for use in a speech recognition system or other voice processing system (14) extracts features related to the frequency (46, F) and amplitude (46, A) characteristics of an imput speech signal (IN) using a plurality of complex band pass filters (30) by processing the outputs of adjacent bandpass filters (42).

WO 02/103676 A1

SPEECH FEATURE EXTRACTION SYSTEM

Background of the Invention

This invention relates to a speech feature extraction system for use in speech recognition, voice identification or voice authentication systems. More specifically, this invention relates to a speech feature extraction system that can be used to create a speech recognition system or other speech processing system with a reduced error rate.

Generally, a speech recognition system is an apparatus that attempts to identify spoken words by analyzing the speaker's voice signal. Speech is converted into an electronic form from which features are extracted. The system then attempts to match a sequence of features to previously stored sequence of models associated with known speech units. When a sequence of features corresponds to a sequence of models in accordance with specified rules, the corresponding words are deemed to be recognized by the speech recognition system.

However, background sounds such as radios, car noise, or other nearby speakers can make it difficult to extract useful features from the speech. In addition, a change in the ambient conditions such as the use of a

different microphone, telephone handset or telephone line can interfere with system performance. Also, a speaker's distance from the microphone, differences between speakers, changes in speaker intonation or emphasis, and even a speaker's health can adversely impact system performance. For a further description of some of these problems, see Richard A. Quinnell, "Speech Recognition: No Longer a Dream, But Still a Challenge," EDN Magazine, January 19, 1995, p. 41-46.

In most speech recognition systems, the speech features are extracted by cepstral analysis, which generally involves measuring the energy in specific frequency bands. The product of that analysis reflects the amplitude of the signal in those bands. Analysis of these amplitude changes over successive time periods can be modeled as an amplitude modulated signal.

Whereas the human ear is a sensitive to frequency modulation as well as amplitude modulation in received speech signals, this frequency modulated content is only partially reflected in systems that perform cepstral analysis.

Accordingly, it would be desirable to provide a speech feature extraction system capable of capturing the frequency modulation characteristics of speech, as well as previously known amplitude modulation characteristics.

It also would be desirable to provide speech recognition and other speech processing systems that incorporate feature extraction systems that provide information on frequency modulation characteristics of the input speech signal.

Summary of the Invention

In view of the foregoing, it is an object of the present invention to provide a speech feature extraction system capable of capturing the frequency modulation characteristics of speech, as well as previously known amplitude modulation characteristics.

It also is an object of this invention to provide speech recognition and other speech processing systems that incorporate feature extraction systems that provide information on frequency modulation characteristics of the input speech signal.

The present invention provides a speech feature extraction system that reflects frequency modulation characteristics of speech as well as amplitude characteristics. This is done by a feature extraction stage which, in one embodiment, includes a plurality of complex band pass filters arranged in adjacent frequency bands according to a linear frequency scale ("linear scale"). The plurality of complex band pass filters are divided into pairs. A pair includes two complex band pass filters in adjacent frequency bands. For every pair, the output of the filter in the higher frequency band ("primary frequency") is multiplied by the conjugate of the output of the filter in the lower frequency band ("secondary filter"). The resulting signal is low pass filtered.

In another embodiment, the feature extraction phase includes a plurality of complex band pass filters arranged according to a logarithmic (or exponential) frequency scale ("log scale"). The primary filters of the filter pairs are centered at various frequencies

along the log scale. The secondary filter corresponding to the primary filter of each pair is centered at a predetermined frequency below the primary filter. For every pair, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

In yet another embodiment, the plurality of band pass filters are arranged according to a mel-scale. The primary filters of the filter pairs are centered at various frequencies along the mel-scale. The secondary filter corresponding to the primary filter of each pair is centered at a predetermined frequency below the primary filter. For every pair, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

In still another embodiment, the plurality of band pass filters are arranged according to a combination of the linear and log scale embodiments mentioned above. A portion of the pairs of the band pass filters are arranged in adjacent frequency bands according to a linear scale. For each of these pairs, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

The primary filters of the remaining pairs of band pass filters are centered at various frequencies along the log scale and the secondary filters corresponding to the primary filters are centered a predetermined frequency below the primary filters. For

every pair, the output of the primary filter is multiplied by the conjugate of the output of the secondary filter. The resulting signal is low pass filtered.

For the embodiments described above, each of the low pass filter outputs is processed to compute two components: a FM component that is substantially sensitive to the frequency of the signal passed by the adjacent band pass filters from which the low pass filter output was generated, and an AM component that is substantially sensitive to the amplitude of the signal passed by the adjacent band pass filters. The FM component reflects the difference in the phase of the outputs of the adjacent band pass filters used to generate the lowpass filter output.

The AM and FM components are then processed using known feature enhancement techniques, such as discrete cosine transform, mel-scale translation, mean normalization, delta and acceleration analysis, linear discriminant analysis and principal component analysis, to generate speech features suitable for statistical processing or other recognition or identification methods. In an alternative embodiment, the plurality of complex band pass filters can be implemented using a Fast Fourier Transform (FFT) of the speech signal or other digital signal processing (DSP) techniques.

In addition, the methods and apparatus of the present invention may be used in addition to performing cepstral analysis in a speech recognition system.

Brief Description of the Drawings

The above and other objects and advantages of the present invention will be apparent upon consideration of the following detailed description, taken in conjunction with the accompanying drawings, in which like reference characters refer to like parts throughout, and in which:

FIG. 1 is a block diagram of an illustrative speech recognition system incorporating the speech feature extraction system of the present invention;

FIG. 2 is a detailed block diagram of the speech recognition system of FIG. 1; and

FIG. 3 is a detailed block diagram of a band pass filter suitable for implementing the feature extraction system of the present invention; and

FIG. 4 is a detailed block diagram of an alternative embodiment of a speech recognition including an alternative speech feature extraction system of the present invention; and

FIG. 5 illustrates is a graph showing band pass filter frequencies spaced according to a linear frequency scale; and

FIG. 6 is a graph showing pairs of band pass filters spaced according to a logarithmic frequency scale; and

FIG. 7 is a graph showing band pass filter frequency pairs spaced according to a mel-scale; and

FIG. 8 is a graph showing band pass filter frequencies spaced according to a combination of linear and logarithmic frequency scales.

Detailed Description of the Invention

Referring to FIG. 1, a generalized depiction of illustrative speech recognition system 5 is described that incorporates the speech extraction system of the present invention. As will be apparent to one of ordinary skill in the art, the speech feature extraction system of the present invention also may be used in speaker identification, authentication and other voice processing systems.

System 5 illustratively includes four stages: pre-filtering stage 10, feature extraction stage 12, statistical processing stage 14, and energy stage 16.

Pre-filtering stage 10, statistical processing stage 14 and energy stage 16 employ speech processing techniques known in the art and do not form part of the present invention. Feature extraction stage 12 incorporates the speech feature extraction system of the present invention, and further includes feature enhancement techniques which are known in the art, as described hereinafter.

Audio speech signal is converted into an electrical signal by a microphone, telephone receiver or other device, and provided as an input speech signal to system 5. In a preferred embodiment of the present invention, the electrical signal is sampled or digitized to provide a digital signal (IN) representative of the audio speech. Pre-filtering stage 10 amplifies the high frequency components of audio signal IN, and the prefiltered signal is then provided to feature extraction stage 12.

Feature extraction stage 12 processes pre-filtered signal X to generate a sequence of feature vectors related to characteristics of input signal IN that may be useful for speech recognition. The output of feature extraction stage 12 is used by statistical processing stage 14 which compares the sequence of feature vectors to predefined statistical models to identify words or other speech units in the input signal IN. The feature vectors are compared to the models using known techniques, such as the Hidden Markov Model (HMM) described in Jelinek, "Statistical Methods for Speech Recognition," The MIT Press, 1997, pp. 15-37. The output of statistical processing stage 14 is the recognized word, or other suitable output depending upon the specific application.

Statistical processing at stage 14 may be performed locally, or at a remote location relative to where the processing of stages 10, 12, and 16 are performed. For example, the sequence of feature vectors may be transmitted to a remote server for statistical processing.

The illustrative speech recognition system of FIG. 1 preferably also includes energy stage 16 which provides an output signal indicative of the total energy in a frame of input signal IN. Statistical processing stage 14 may use this total energy information to provide improved recognition of speech contained in the input signal.

Referring now to FIG. 2, pre-filtering stage 10 and feature extraction stage 12 are described in greater detail. Pre-filtering stage 10 is a high pass filter

that amplifies high frequency components of the input signal. Pre-filtering stage 10 comprises one-sample delay element 21, multiplier 23 and adder 24. Multiplier 23 multiplies the one-sample delayed signal by constant K_f, which typically has a value of -0.97. The output of pre-filtering stage 10, X, is input at the sampling rate into a bank of band pass filters 30_1, 30_2, ... 30_n.

In one embodiment, the band pass filters 30_1, 30_2, ... 30_n are positioned in adjacent frequency bands. The spacing of the band pass filters 30_1, 30_2, ... 30_n is done according to a linear frequency scale ("linear scale") 68 as shown in graph 72 of FIG. 5. The term "linear frequency scale" is used in this specification in accordance with its ordinary and accustomed meaning, i.e., the actual frequency divisions are uniformly spaced. The plurality of complex band pass filters 30_1, 30_2, ... 30_n are divided into pairs P_{1-2}. A pair (P_1 or P_2) includes two complex band pass filters (30_{1-2} or 30_{3-4}) respectively in adjacent frequency bands. For every pair (P_1 or P_2), the output of the filter in the higher frequency band (30_2 or 30_4) (referred to hereinafter as the "primary filter") is multiplied by the conjugate of the output of the filter in the lower frequency band (30_1 or 30_3) (referred to hereinafter as the "secondary filter"). The resulting signal is low pass filtered.

The number of band pass filters 30_1, 30_2, ... 30_n and width of the frequency bands preferably are selected according to the application for the speech processing system. For example, a system useful in telephony applications preferably will employ about forty band pass filters 30_1, 30_2, ... 30_n having center

frequencies approximately 100 Hz apart. For example, filter 30_1 may have a center frequency of 50 Hz, filter 30_2 may have a center frequency of 150 Hz, filter 30_3 may have a center frequency of 250 Hz, and so on, so that the center frequency of filter 30_{40} is 3950 Hz. The bandwidth of each filter may be several hundred Hertz.

In another embodiment, as illustrated in the graph 70 of FIG. 6, the band pass filters 30_1, 30_2, ... 30_{108} are arranged according to a non-linear frequency scale such as a logarithmic (or exponential) frequency scale 74 ("log scale"). The term logarithmic frequency scale is used in this specification according to its ordinary and accustomed meaning.

Empirical evidence suggests that using log scale 74 of FIG. 6 instead of linear scale 68 of FIG. 5 improves voice recognition performance. That is so because the human ear resolves frequencies non-linearly across the audio spectrum. Another advantage using log scale 74 instead of linear scale 68 is that log scale 74 can cover a wider range of frequency spectrum without using additional band pass filters 30_1, 30_2, ... 30_n.

Pairs P_{1-54} of band pass filters 30_{1-108} are spaced according to log scale 74. Pair P_1 includes filters 30_1 and 30_2, pair P_{10} includes filters 30_{19} and 30_{20}, and pair P_{54} includes filters 30_{107} and 30_{108}. In this arrangement, filters $30_2, 30_{20}$ and 30_{108} are the primary filters and filters 30_1, 30_{19} and 30_{107} are the secondary filters.

In one preferred embodiment, primary filters $30_2, 30_{20}...30_{108}$ are centered at various frequencies along log scale 74, while secondary filters 30_1, $30_3...30_{107}$ are

centered 100 hertz (Hz) below corresponding primary filters 30_2, 30_4...30_{108} respectively. An exemplary MATLAB code to generate graph 70 of FIG. 6 is shown below.

```
v=(2.^([26.715:0.25:40]/3.345)); % Generate center frequencies for bandpass filter pairs
f(2:2:2*length(v))=v+50;         % Primary filters center frequencies
f(1:2:2*length(v))=v-50;         % Secondary filters center frequencies

semilogy([v'+50 v'-50],'.');grid % Plot center frequencies on a logarithmic scale'
```

In another embodiment, the center frequencies of primary filters 30_2,30_{20}...30_{108} and secondary filters 30_1, 30_3...30_{107} may be placed using separate and independent algorithms, while ensuring that secondary filters 30_1, 30_3...30_{107} are centered 100 hertz (Hz) below their corresponding primary filters 30_2, 30_4...30_{108}, respectively.

In one embodiment, band pass filters 30_{1-108} are of triangular shape. In other embodiments, band pass filters 30_{1-108} can be of various shapes depending on the requirements of the particular voice recognition systems.

Log scale 74 is shown to range from 0-4000 Hz. For every pair P_{1-54}, the output of the primary filter 30_2, 30_4...30_{108} is multiplied by the conjugate of the output of secondary filter 30_1, 30_3...30_{107}. The resulting signal is low pass filtered.

The pairs P_{1-54} are arranged such that the lower frequencies include a higher concentration of pairs P_{1-54} than the higher frequencies. For example, there are 7 pairs (P_{16-22}) in the frequency range of 500-1000 Hz while there are only 3 pairs (P_{49-51}) in the frequency range of

3000-3500 Hz. Thus, although there is over sampling at the lower frequencies, this embodiment also performs at least some sampling at the higher frequencies. The concentration of pairs P_{1-54} along log scale 74 can be varied depending on the needs of a particular voice recognition system.

As will be apparent to one of ordinary skill in the art of digital signal processing design, the band pass filters of the preceding embodiments may be implemented using any of a number of software or hardware techniques. For example, the plurality of complex filters may be implemented using a Fast Fourier Transform (FFT), Chirp-Z transform, other frequency domain analysis techniques.

In an alternative embodiment, depicted in FIG. 7, band pass filters 30_1, 30_2, ... 30_n are arranged according to a non-linear frequency scale, such as mel-scale 80. Mel-scale 80 is well known in the art of voice recognition systems and is typically defined by the equation,

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right)$$

where f represents frequency according to the linear scale 68 and Mel(f) represents its corresponding mel-scale 80 frequency.

FIG. 7 illustrates one embodiment of a graph 84 showing band pass filters 30_{1-9} spaced according to mel-scale 80. The center frequencies (CF_{1-9}) are the Mel(f) values calculated by using the above equation. Typically, filters 30_{1-9} are spread over the whole

frequency range from zero up to the Nyquist frequency. In one embodiment, filters 30_{1-9} have the same band width. In another embodiment, filters 30_{1-9} may have different bandwidths.

In yet another embodiment, depicted in FIG. 8, the band pass filters 30_{1-8} are spaced according to a combination of linear 68 and non-linear 74 frequency scales. Band pass filters 30_{1-4} (P_{1-2}) are arranged in adjacent frequency bands according to linear frequency 68.

Primary filters 30_6 and 30_8 are centered along log scale 74. Secondary filters 30_5 and 30_7 are centered at frequencies of 100 Hz below the center frequencies of 30_6 and 30_8 respectively. For each of these pairs (P_1 or P_2), the output of the primary filter (30_6 or 30_8) is multiplied by the conjugate of the output of the secondary filter (30_5 and 30_7), respectively and the resulting signal is low pass filtered.

Referring again to FIG. 2, blocks 40_{1-20} provide the complex conjugate of the output signal of band pass filter 30_1, 30_3, ... 30_{n-1}. Multiplier blocks 42_{1-20} multiply the complex conjugates by the outputs of an adjacent higher frequency band pass filter 30_2, 30_4, 30_6, ... 30_{40} to provide output signals Z_{1-20}. Output signals Z_{1-20} then are passed through a series of low pass filters 44_{1-20}. The outputs of the low pass filters typically are generated only at the feature frame rate. For example, at a input speech sampling rate of 8 kHz, the output of the low pass filters is only computed at a feature frame rate of once every 10 msec.

Each output of low pass filters 44_{1-20} is a complex signal having real component R and imaginary component I. Blocks 46_{1-20} process the real and imaginary components of the low pass filter outputs to provide output signals A_{1-20} and F_{1-20} as shown in equations (1) and (2):

$$A_i = \log R_i \quad (1)$$

$$F_i = \frac{I_i}{\sqrt{R_i^2 + I_i^2}} \quad (2)$$

wherein R_i and I_i are the real and imaginary components of the corresponding low pass filter output. Output signals A_i are a function of the amplitude of the low pass filter output and signals F_i are a function of the frequency of the signal passed by the adjacent band pass filters from which the low pass filter output was generated. By computing two sets of signals that are indicative of the amplitude and frequency of the input signal, the speech recognition system incorporating the speech feature extraction system of the present invention is expected to provide reduced error rate.

The amplitude and frequency signals A_{1-20} and F_{1-20} then are processed using conventional feature enhancement techniques in feature enhancement component 12b, using, for example, discrete cosine transform, mel-scale translation, mean normalization, delta and acceleration analysis, linear discriminant analysis and principal component analysis techniques that are per se known in the art. A preferred embodiment of a speech

recognition system of the present invention incorporating the speech extraction system of the present invention employs a discrete cosine transform and delta features technique, as described hereinafter.

Still referring to FIG. 2, feature enhancement component 12b receives output signals A_{1-20} and F_{1-20}, and processes those signals using discrete cosine transform (DCT) blocks 50 and 54, respectively. DCTs 50 and 54 attempt to diagonalize the co-variance matrix of signals A_{1-20} and F_{1-20}. This helps to uncorrelate the features in output signals B_{0-19} of DCT 50 and output signals C_{0-19} of DCT 54. Each set of output signals B_{0-19} and C_{0-19} then are input into statistical processing stage 14. The function performed by DCT 50 on input signals A_{1-20} to provide output signals B_{0-19} is shown by equation (3), and the function performed by DCT 54 on input signals F_{1-20} to provide output signals C_{0-19} is shown by equation (4).

$$B_r = D(r) \sum_{n=0}^{N-1} A_{n+1} \bullet \cos\frac{(2n+1)\pi r}{2N} \qquad (3)$$

$$C_r = D(r) \sum_{n=0}^{N-1} F_{n+1} \bullet \cos\frac{(2n+1)\pi r}{2N} \qquad (4)$$

In equations (3) and (4), N equals the length of the input signal vectors A and F (e.g., N = 20 in FIG. 2), n is an index from 0 to N - 1 (e.g., n = 0 to 19 in the embodiment of FIG. 2), and r is the index of output signals B and C (e.g., r = 0 to 19 in the embodiment of FIG. 2). Thus, for each vector output signal B_r, each vector of input signals A_{1-20} are multiplied by a cosine function and D(r) and summed

together as shown in equation (3). For each vector output signal C_r, each vector of input signals S_{1-20} are multiplied by a cosine function and D(r) and summed together as shown in equation (4). D(r) are coefficients that are given by the following equations:

$$D(r)=\frac{1}{\sqrt{N}} \quad for\ r = 0 \tag{5}$$

$$D(r)=\frac{\sqrt{2}}{\sqrt{N}} \quad for\ r>0 \tag{6}$$

Output signals B_{0-19} and C_{0-19} also are input into delta blocks 52 and 56, respectively. Each of delta blocks 52 and 56 takes the difference between measurements of feature vector values between consecutive feature frames and this difference may be used to enhance speech recognition performance. Several difference formulas may be used by delta blocks 52 and 56, as are known in the art. For example, delta blocks 52 and 56 may take the difference between two consecutive feature frames. The output signals of delta blocks 52 and 56 are input into statistical processing stage 14.

Energy stage 16 of FIG. 2 is a previously known technique for computing the logarithm of the total energy (represented by E) of each frame of input speech signal IN, according to the following equation:

$$E(nT)=\log\left(\frac{\sum_{i=0}^{K-1} I\,N^2_{(n-i)T}}{K}\right) \tag{7}$$

Equation 7 shows that energy block 16 takes the sum of the squares of the values of the input signal IN during the previous K sampling intervals (e.g., K = 220, T = 1/8000 seconds), divides the sum by K, and takes the logarithm of the final result. Energy block 16 performs this calculation every frame (e.g., 10 msec), and provides the result as an input to statistical processing block 14.

Referring now to FIG. 3, illustrative complex bandpass filter 30' suitable for use in the feature extraction system of the present invention is described. Filter 30' comprises adder 31, multiplier 32 and one-sample delay element 33. Multiplier 32 multiples the one-sample delayed output Y by complex coefficient G and the resultant is added to the input signal X to generate an output signal Y.

An alternative embodiment of the feature extraction system of the present invention is described with respect to FIG. 4. The embodiment of FIG. 4 is similar to the embodiment of FIG. 2 and includes pre-filtering stage 10, statistical processing stage 14, and energy stage 16 the operate substantially as described above. However, the embodiment of FIG. 4 differs from the previously described embodiment in that feature extraction stage 12' includes additional circuitry within feature extraction system 12a, so that the feature vectors include additional information.

For example, feature extraction stage 12a' includes a bank of 41 band pass filters 301-41 and conjugate blocks 401-40. The output of each band pass filter is combined with the conjugate of the output of a

lower adjacent band bass filter by multipliers 421-40. Low pass filters 441-40, and computation blocks 461-40 compute vectors A and F as described above, except that the vectors have a length of forty elements instead of twenty. DCTs 50 and 54, and delta blocks 52 and 56 of feature enhancement component 12b' each accept the forty element input vectors and output forty element vectors to statistical processing block 14. It is understood that the arrangement illustrated in FIG. 4 is not applicable if the band pass filters 301-41 arranged according to a non-linear frequency scale such as a log scale or a mel-scale.

The present invention includes feature extraction stages which may include any number of band pass filters 30, depending upon the intended voice processing application, and corresponding numbers of conjugate blocks 40, multipliers 42, low pass filters 44 and blocks 46 to provide output signals A and F for each low pass filter. In addition, signals A and F may be combined in a weighted fashion or only part of the signals may be used. For example, it may be advantageous to use only the amplitude signals in one frequency domain, and a combination of the amplitude and frequency signals in another.

While preferred illustrative embodiments of the invention are described above, it will be apparent to one skilled in the art that various changes and modifications may be made therein without therein without departing from the invention, and it is intended in the appended claims to cover all such changes and modifications which fall within the true spirit and scope of the invention.

What is claimed is:

1. Apparatus for use in a speech processing system for extracting features from an input speech signal having frequency and amplitude characteristics, the apparatus comprising:

first and second band pass filters adapted to receive the input speech signal and providing respectively, first and second signals;

a conjugate circuit coupled to the second band pass filter and providing a third signal that is the conjugate of the second signal;

a multiplier coupled to the first band pass filter and to the conjugate circuit and providing a fourth signal that is the product of the first and third signals; and

filter means coupled to the multiplier for providing a fifth signal related to the frequency characteristics of the input signal, and a sixth signal related to amplitude characteristics of the input signal.

2. The apparatus of claim 1 wherein the filter means includes a low pass filter.

3. The apparatus of claim 1 further comprising a high pass filter coupled between the input signal and the first and second band pass filters.

4. The apparatus of claim 2 wherein the low pass filter provides a signal having a real part, R, and an imaginary part, I, and the filter means includes circuitry for calculating:

$A = \log R$ and

F = I / Sqrt(R2 + I2).

5. The apparatus of claim 1 wherein the first and second band pass filters each comprises a delay element, a complex multiplier and an adder.

6. The apparatus of claim 5 wherein the adder receives the input speech signals and outputs an output signal, the adder adding the input speech signal and a signal comprising a delayed sample of the output signal multiplied by the complex coefficient.

7. The apparatus of claim 4 further comprising feature enhancement circuitry.

8. The apparatus of claim 7 further comprising circuitry for statistically processing an output of the feature enhancement circuitry to compare that output to a plurality of predetermined models.

9. The apparatus of claim 7 wherein the feature enhancement circuitry comprises circuitry for computing a discrete cosine transform of A and F.

10. The apparatus of claim 7 wherein the feature enhancement circuitry further comprises circuitry for calculating a difference between successive discrete cosine transforms of A and for calculating a difference between successive discrete cosine transforms of F.

11. The apparatus of claim 1 further comprising:

a transducer for converting sound into an electrical signal; and

a sampler for converting the electrical signal into a digital signal, wherein the input speech signal comprises the digital signal.

12. The apparatus of claim 1 further comprising a digital processor, wherein the first and second band pass filters, the conjugate circuit, the multiplier, and the filter means comprise algorithms adapted to be executed on the digital processor.

13. The apparatus of claim 1, wherein a frequency of the first bandpass filter is selected according to a logarithmic scale and a frequency of the second bandpass filter is selected at a predetermined frequency below the frequency of the first bandpass filter.

14. The apparatus of claim 13, wherein a center frequency of the second filter is one hundred hertz less than a center frequency of the first filter.

15. The apparatus of claim 13, wherein the bands of the first and second filters overlap.

16. The apparatus of claim 13, wherein the bandwidths of the first and second filters are the same.

17. The apparatus of claim 1, wherein a frequency of the first bandpass filter is selected according to a mel-scale and a frequency of the second bandpass filter is selected at a predetermined frequency below the frequency of the first bandpass filter.

18. The apparatus of claim 17, wherein a center frequency of the second filter is one hundred hertz less than a center frequency of the first filter.

19. The apparatus of claim 17, wherein the bands of the first and second filters overlap.

20. The apparatus of claim 17, wherein the mel-scale is defined by the equation

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700} \right).$$

21. The apparatus of claim 17, wherein the bandwidths of the first and second filters are the same.

22. The apparatus of claim 1, wherein the first and second band pass filters are implemented using a Fast Fourier Transform.

23. A method for extracting features from an input speech signal for use in a speech processing device, the method comprising:

separating the input signal into a first signal in a first frequency band and a second signal in a second frequency band;

providing a conjugate of the first signal;

multiplying the conjugate of the first signal with the second signal to provide a third signal; and

processing the third signal to generate a frequency component related to frequency features in the input speech signal and an amplitude component related to amplitude features in the input speech signal.

24. The method of claim 23 further comprising high pass filtering the input speech signal prior to processing the input signal to generate the first and second signals.

25. The method of claim 23 wherein processing the third signal further comprises low pass filtering the third signal.

26. The method of claim 25 wherein low pass filtering the third signal provides a signal having a real part, R, and an imaginary part, I, and processing the third signal includes calculating:

$$A = \log R \text{ and}$$

$$F = I / \mathrm{Sqrt}(R2 + I2).$$

27. The method of claim 26 further comprising processing the third signal using feature enhancement techniques to generate a series of feature vectors.

28. The method of claim 27 further comprising statistically processing the series of feature vectors by comparing the series of feature vectors to a plurality of predetermined models.

29. The method of claim 27 further comprising calculating a discrete cosine transform of A and F.

30. The method of claim 29 further comprising calculating a difference between successive discrete cosine transforms of A and between successive discrete cosine transforms of S.

31. The method of claim 30 further comprising statistically processing the differences between successive discrete cosine transforms to compare the differences to a plurality of predetermined patterns.

32. The method of claim 23 further comprising:

converting a sound into an electrical signal; and

sampling the electrical signal to provide a digital signal, wherein the input speech signal comprises the digital signal.

33. A method of claim 23, comprising selecting a frequency of the first frequency band according to a logarithmic scale and selecting a frequency of the second frequency band at a predetermined frequency below the frequency of the first frequency band.

34. The method of claim 33, wherein a center frequency of the second frequency band is one hundred hertz less than a center frequency of the first frequency band.

35. The method of claim 33, comprising overlapping the first and second frequency bands.

36. The method of claim 23, comprising selecting a frequency of the first frequency band according to a mel-scale and selecting the frequency of the second frequency band at a predetermined frequency below the frequency of the the first frequency band.

37. The apparatus of claim 36, wherein a center frequency of the second frequency band is one hundred hertz less than a center frequency of the first frequency band.

38. The apparatus of claim 36, comprising overlapping the first and second frequency bands.

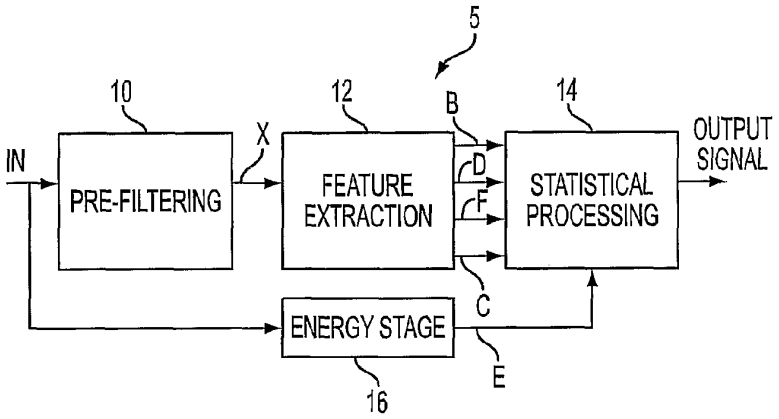

5
10
12
14
B
X
IN
PRE-FILTERING
FEATURE EXTRACTION
STATISTICAL PROCESSING
OUTPUT SIGNAL
D
F
C
ENERGY STAGE
E
16

FIG. 1

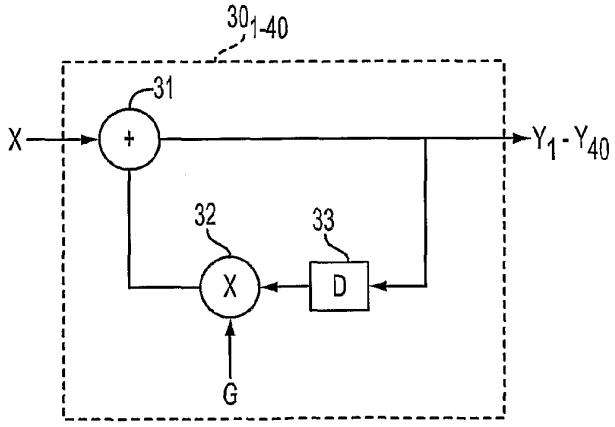

30 1-40
31
X
+
Y_1 - Y_{40}
32
33
X
D
G

FIG. 3

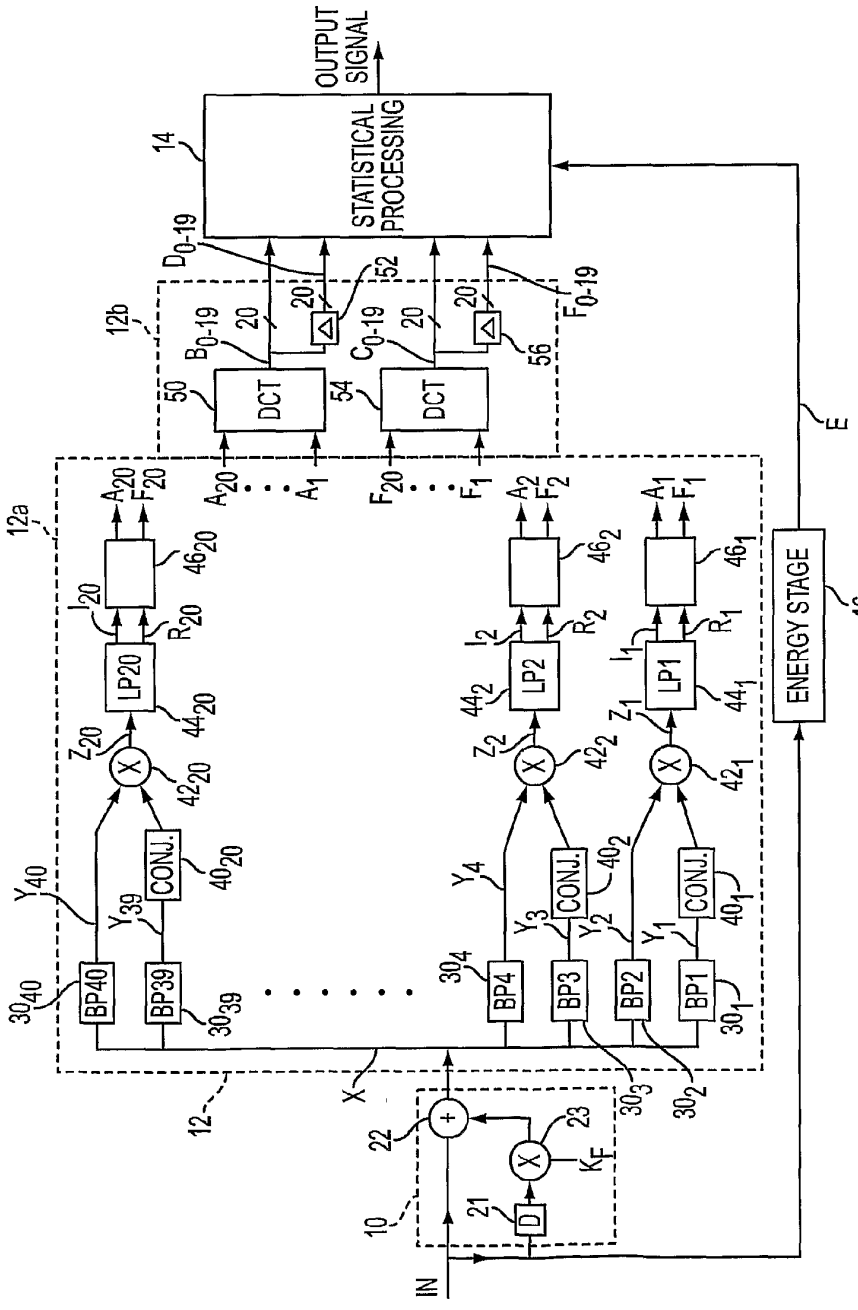

OUTPUT SIGNAL
STATISTICAL PROCESSING
14
D_{0-19}
B_{0-19}
C_{0-19}
F_{0-19}
12b
DCT
DCT
50
54
52
56
20
E
ENERGY STAGE
16
12a
LP20
LP2
LP1
44_{20}
44_2
44_1
46_{20}
46_2
46_1
I_{20}
I_2
I_1
R_{20}
R_2
R_1
Z_{20}
Z_2
Z_1
42_{20}
42_2
42_1
CONJ.
40_{20}
40_2
40_1
Y_{40}
Y_{39}
Y_4
Y_3
Y_2
Y_1
BP40
BP39
BP4
BP3
BP2
BP1
30_{40}
30_{39}
30_4
30_3
30_2
30_1
A_{20} ... A_1
F_{20} ... F_1
A_2 F_2
A_1 F_1
12
X
22
23
K_F
21
D
10
IN

FIG. 2

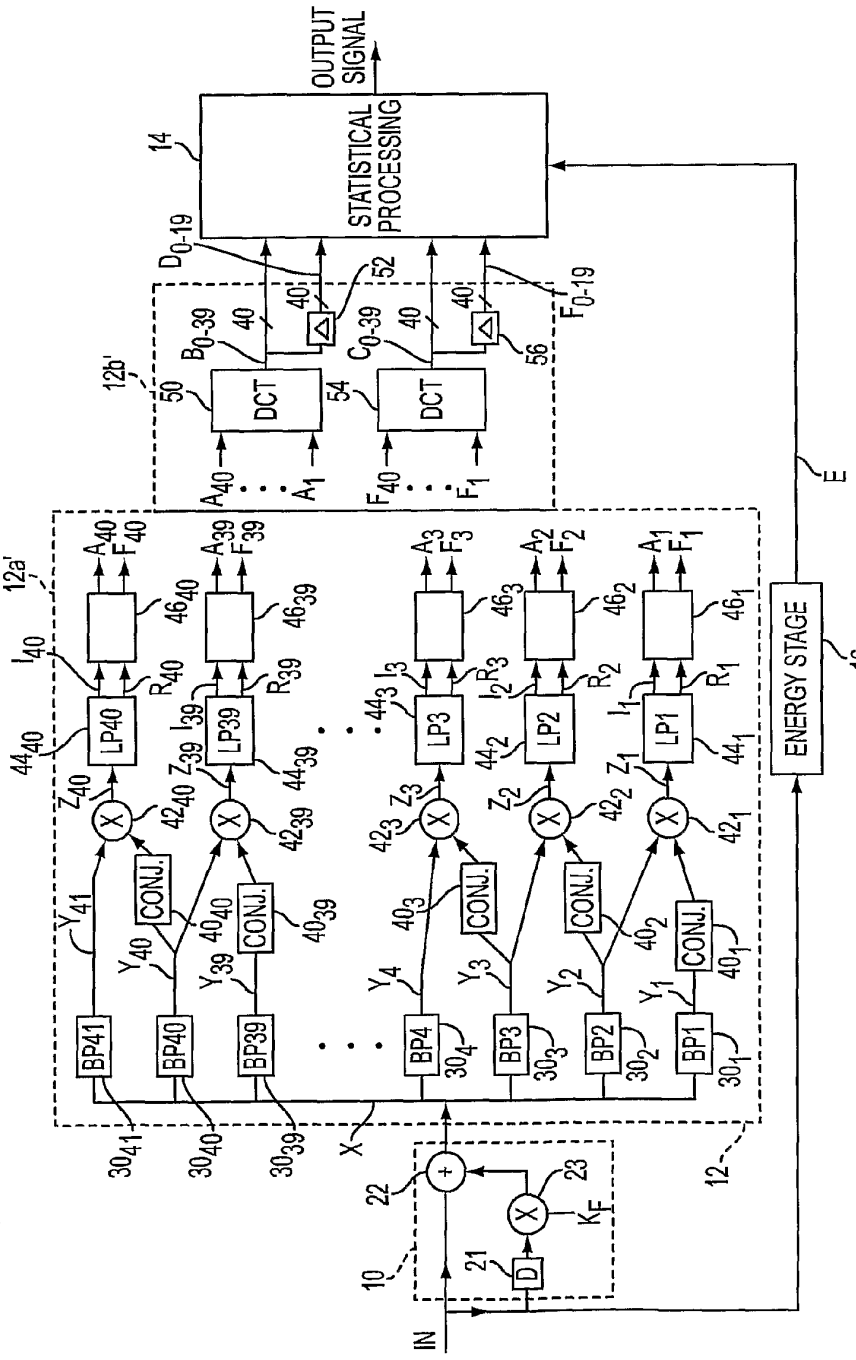

OUTPUT SIGNAL
STATISTICAL PROCESSING
14
D_{0-19}
F_{0-19}
B_{0-39}
C_{0-39}
DCT
DCT
12b'
12a'
A_{40} ... A_1
F_{40} ... F_1
E
ENERGY STAGE
16
LP40
LP39
LP3
LP2
LP1
BP41
BP40
BP39
BP4
BP3
BP2
BP1
CONJ.
X
12
10
IN
22
23
K_F
21
D

FIG. 4

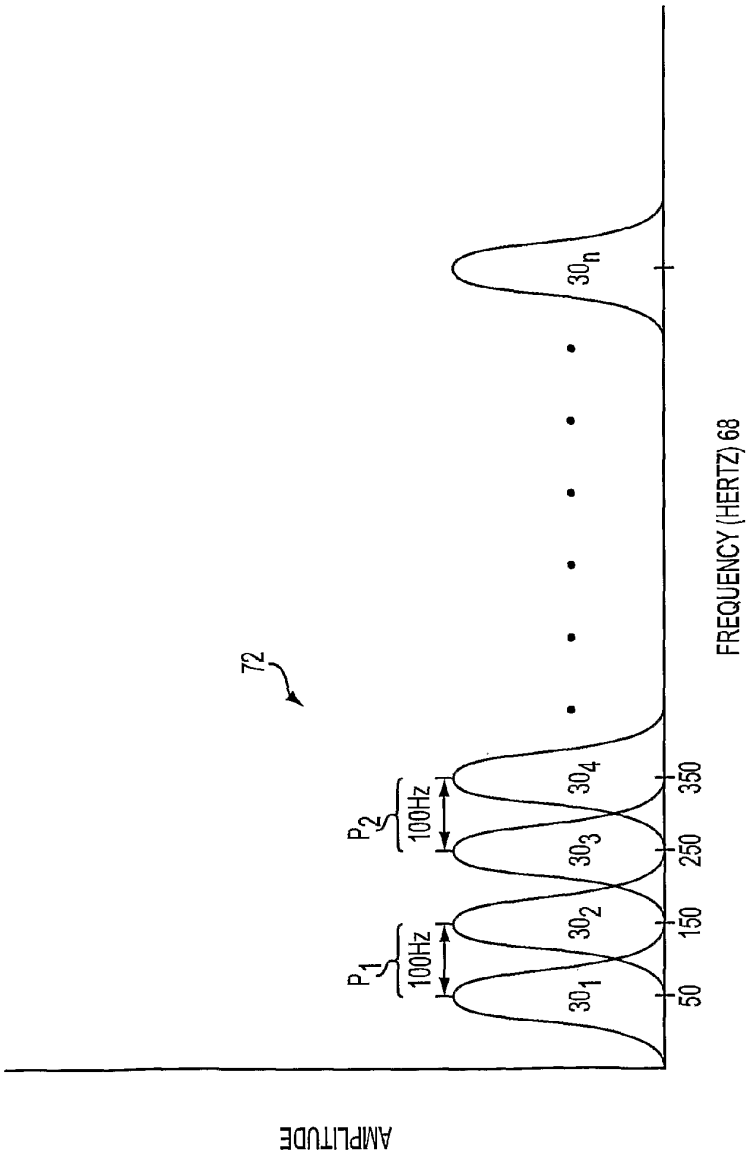

AMPLITUDE
72
30_1
30_2
30_3
30_4
30_n
P_1
100Hz
P_2
100Hz
50
150
250
350
FREQUENCY (HERTZ) 68

FIG. 5

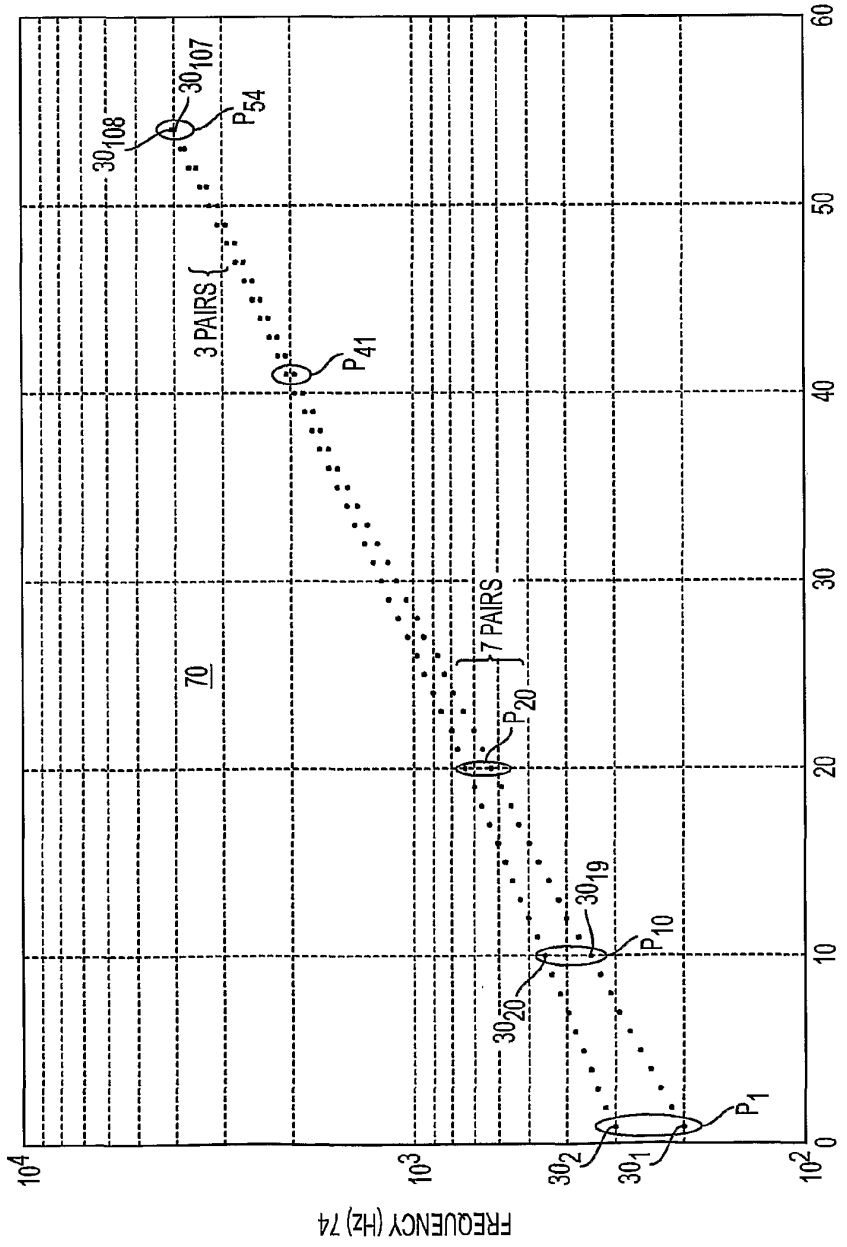

FREQUENCY (HZ) 74
10^4
10^3
10^2
0
10
20
30
40
50
60
PAIRS P_{1-54}
30_1
30_2
30_{20}
30_{19}
30_{108}
30_{107}
P_1
P_{10}
P_{20}
P_{41}
P_{54}
3 PAIRS
7 PAIRS
70

FIG. 6

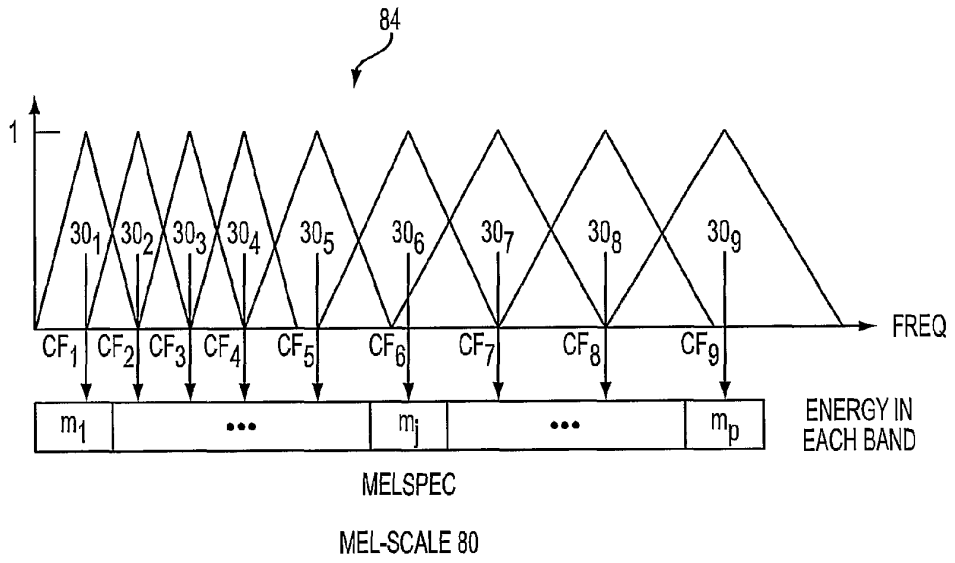

84
1
30_1 30_2 30_3 30_4 30_5 30_6 30_7 30_8 30_9
FREQ
CF_1 CF_2 CF_3 CF_4 CF_5 CF_6 CF_7 CF_8 CF_9
m_1 ••• m_j ••• m_p
ENERGY IN EACH BAND
MELSPEC
MEL-SCALE 80

FIG. 7

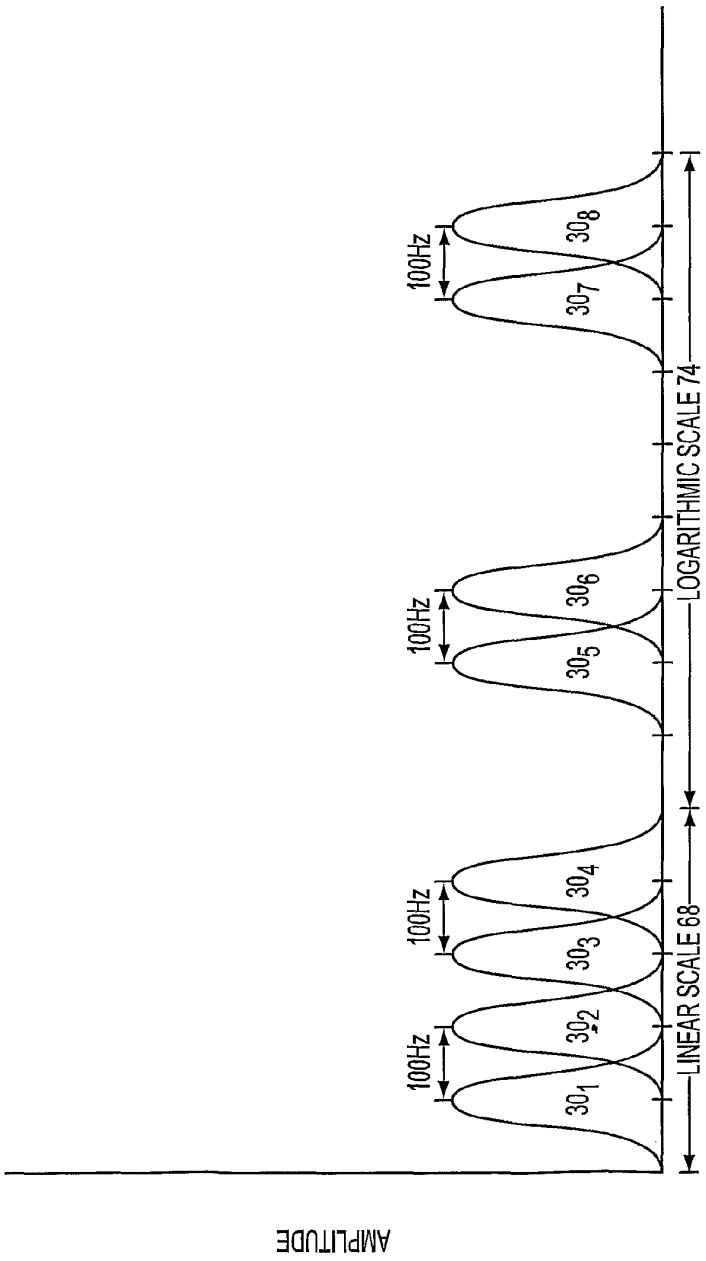

AMPLITUDE
LINEAR SCALE 68
LOGARITHMIC SCALE 74
100Hz
100Hz
100Hz
30_1
30_2
30_3
30_4
30_5
30_6
30_7
30_8

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/19182

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G10L 15/02
US CL : 704/243, 205, 206; 708/300
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 704/243, 205, 206, 231, 268; 708/300-323; 370/484; 375/285.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 4,300,229 A (HIROSAKI) 10 November 1981 (10.11.1981), ALL	1-38
A	US 4,729,112 A (MILLAR) 01 March 1988 (01.03.1988), ALL	1-38
A	US 4,660,216 A (CLAASEN et al.) 21 April 1987 (21.04.1987), ALL	1-38
A	US 4,221,934 A (SCHIFF) 09 September 1980 (09.09.1980), ALL	1-38

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"B" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 September 2002 (08.09.2002)	
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer David D. Knepper Telephone No. (703) 306-037[illegible]

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

PCT/US02/19182

Continuation of B. FIELDS SEARCHED Item 3:
EAST: (speech OR voice) AND (Frequency AND Amplitude) AND (conjugate OR imaginary) AND ((band NEAR1 pass NEAR1 filter) OR BPF)

Espacenet

Bibliographic data: JP2005148274 (A) — 2005-06-09

SIGNAL ANALYZING METHOD AND SIGNAL COMPOSING METHOD FOR COMPLEX INDEX MODULATION FILTER BANK, AND PROGRAM THEREFOR AND RECORDING MEDIUM THEREFOR

Inventor(s): NAGAI KIYOTAKA; USAMI AKIRA ±

Applicant(s): MATSUSHITA ELECTRIC IND CO LTD ±

Classification:
- **international:** ***G10L11/00; G10L19/02; G10L19/14; H03H17/02;*** (IPC1-7): G10L11/00; G10L19/02; G10L19/14
- **European:** G10L19/02S; H03H17/02E; H03H17/02F8A

Application number: JP20030383535 20031113

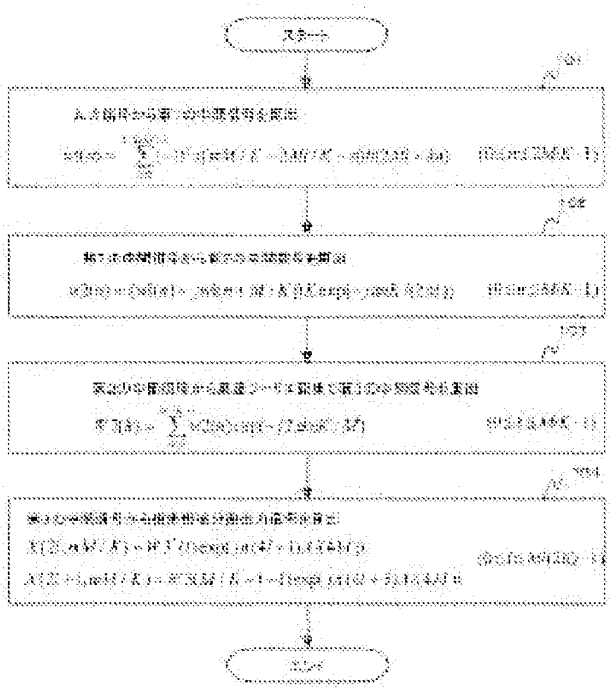
Priority number(s): JP20030383535 20031113

Also published as: JP4396233 (B2) US2005108002 (A1) US7433907 (B2)

Abstract of JP2005148274 (A)

PROBLEM TO BE SOLVED: To achieve low power consumption or faster operation of a complex index modulation filter bank by reducing its computational complexity. ; SOLUTION: Provided are a step 101 of calculating a 1st intermediate signal from an input signal, a step 102 of calculating a 2nd intermediate signal from the 1st intermediate signal, a step 103 of calculating a 3rd intermediate signal from the 2nd intermediate signal through fast Fourier transformation, and a step 104 of calculating a complex band output signal from the 3rd intermediate signal. Consequently, reduced is the operation quantity needed to realize the complex index modulation filter bank. ; COPYRIGHT: (C)2005,JPO&NCIPI

Last updated: 28.02.2012 Worldwide Database 5.7.36; 93p



(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2005-148274
(P2005-148274A)
(43)公開日　平成17年6月9日(2005.6.9)

(51)Int.Cl.[7]	F I		テーマコード（参考）
G10L 19/14	G10L 7/04	H	
G10L 11/00	G10L 7/04	G	
G10L 19/02	G10L 9/16		

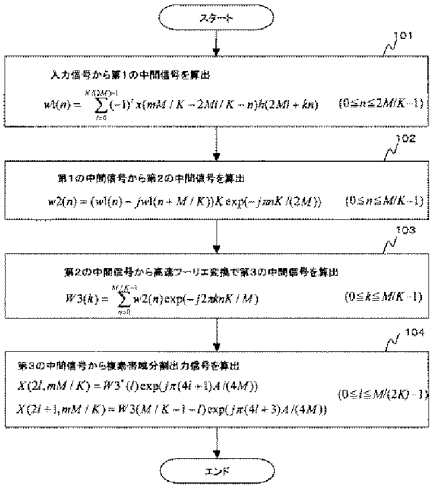
審査請求 未請求　請求項の数 9　OL　（全 38 頁）

(21)出願番号　特願2003-383535 (P2003-383535)
(22)出願日　平成15年11月13日 (2003.11.13)

(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真１００６番地
(74)代理人　100097445
弁理士　岩橋　文雄
(74)代理人　100103355
弁理士　坂口　智康
(74)代理人　100109667
弁理士　内藤　浩樹
(72)発明者　永井　清隆
大阪府門真市大字門真１００６番地　松下電器産業株式会社内
(72)発明者　宇佐見　陽
大阪府門真市大字門真１００６番地　松下電器産業株式会社内

(54)【発明の名称】複素指数変調フィルタバンクの信号分析方法、信号合成方法、そのプログラム及びその記録媒体

(57)【要約】
【課題】複素指数変調フィルタバンクにおいて、演算量を削減することにより、低消費電力化あるいは高速化を実現する。
【解決手段】入力信号から第１の中間信号を算出するステップ１０１と、第１の中間信号から第２の中間信号を算出するステップ１０２と、第２の中間信号から高速フーリエ変換で第３の中間信号を算出するステップ１０３と、第３の中間信号から複素帯域出力信号を算出するステップ１０４とを備える。これにより、複素指数変調フィルタバンクの実現に必要な演算量を削減することができる。
【選択図】図１



【特許請求の範囲】
【請求項1】
サンプリング周波数ｆｓの入力信号をＫ倍（ただし、Ｋは帯域分割数Ｍの約数で１を含む正整数）にアップサンプリングして、Ｍ個の等帯域幅のサンプリング周波数ｆｓＫ／Ｍの複素帯域分割信号に分割して出力する信号分析方法あって、直線位相非巡回形のプロトタイプフィルタのフィルタ係数をｈ（ｎ）（ただし、Ｎはフィルタ次数、０≦ｎ≦Ｎ）、最初と最後のフィルタ係数の値を零、サンプリング時刻ｎにおける入力信号をｘ（ｎ）とするとき、
入力信号から第１の中間信号ｗ１（ｎ）を（数１）によって算出するステップと、
【数1】

$$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+Kn)\quad(0\leqq n\leqq 2M/K-1)$$

第１の中間信号から第２の中間信号ｗ２（ｎ）を（数２）（ただし、ｊは虚数単位）によって算出するステップと、
【数2】

$$w2(n)=(w1(n)-jw1(n+M/K))K\exp(-j\pi nK/(2M))\quad(0\leqq n\leqq M/K-1)$$

第２の中間信号から高速フーリエ変換で第３の中間信号Ｗ３（ｋ）を（数３）によって算出するステップと、
【数3】

$$W3(k)=\sum_{n=0}^{M/K-1}w2(n)\exp(-j2\pi knK/M)\quad(0\leqq k\leqq M/K-1)$$

第３の中間信号から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を、ｋが偶数（ｋ＝２１）の場合には（数４）（ただし、＊は共役複素数、Ａは信号分析の位相）、ｋが奇数（ｋ＝２１＋１）の場合には（数５）、によって算出するステップを備えたことを特徴とする複素指数変調フィルタバンクの信号分析方法。
【数4】

$$X(2l,mM/K)=W3^*(l)\exp(j\pi(4l+1)A/(4M))\quad(0\leqq l\leqq M/(2K)-1)$$

【数5】

$$X(2l+1,mM/K)=W3(M/K-1-l)\exp(j\pi(4l+3)A/(4M))\quad(0\leqq l\leqq M/(2K)-1)$$

【請求項2】
サンプリング周波数ｆｓの入力信号をＫ倍（ただし、Ｋは帯域分割数Ｍの約数で１を含む正整数）にアップサンプリングして、Ｍ個の等帯域幅のサンプリング周波数ｆｓＫ／Ｍの複素帯域分割信号に分割して出力する信号分析方法あって、直線位相非巡回形のプロトタイプフィルタのフィルタ係数をｈ（ｎ）（ただし、Ｎはフィルタ次数、０≦ｎ≦Ｎ）、最初と最後のフィルタ係数の値を零、サンプリング時刻ｎにおける入力信号をｘ（ｎ）とするとき、
入力信号から第１の中間信号ｗ１（ｎ）を（数６）によって算出するステップと、
【数6】

$$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+Kn)\quad(0\leqq n\leqq 2M/K-1)$$

第１の中間信号から第２の中間信号ｗ２（ｎ）を（数７）（ただし、ｊは虚数単位）によって算出するステップと、
【数7】

$$w2(n)=(w1(n)+jw1(n+M/K))K\exp(j\pi mK/(2M)) \quad (0\leqq n\leqq M/K-1)$$

第２の中間信号から逆高速フーリエ変換で第３の中間信号Ｗ３（ｋ）を（数８）によって算出するステップと、

【数8】

$$W3(k)=\sum_{n=0}^{M/K-1} w2(n)\exp(j2\pi knK/M) \quad (0\leqq k\leqq M/K-1)$$

第３の中間信号から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を、ｋが偶数（ｋ＝２１）の場合には（数９）（ただし、Ａは信号分析の位相）、ｋが奇数（ｋ＝２１＋１）の場合には（数１０）（ただし、＊は共役複素数）、によって算出するステップを備えたことを特徴とする複素指数変調フィルタバンクの信号分析方法。

【数9】

$$X(2l,mM/K)=W3(l)\exp(j\pi(4l+1)A/(4M)) \quad (0\leqq l\leqq M/(2K)-1)$$

【数10】

$$X(2l+1,mM/K)=W3^{*}(M/K-1-l)\exp(j\pi(4l+3)A/(4M)) \quad (0\leqq l\leqq M/(2K)-1)$$

【請求項3】

信号分析の位相が、第３の中間信号から複素帯域分割出力信号を算出するステップで乗算を発生しないように設定されていることを特徴とする請求項１ないし２のいずれかに記載の複素指数変調フィルタバンクの信号分析方法。

【請求項4】

Ｍ個の等帯域幅のサンプリング周波数ｆｓＫ／Ｍ（ただし、Ｋは帯域分割数Ｍの約数で１を含む正整数）の複素帯域分割入力信号を合成し、１／Ｌ倍（ただし、ＬはＭの約数で１を含む正整数）にダウンサンプリングして、サンプリング周波数ｆｓＫ／Ｌの信号を出力する信号合成方法であって、直線位相非巡回形のプロトタイプフィルタのフィルタ係数をｈ（ｎ）（ただし、Ｎはフィルタ次数、０≦ｎ≦Ｎ）、最初と最後のフィルタ係数の値を零、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とするとき、

複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数１１）（ただし、ｊは虚数単位、＊は共役複素数、Ｂは信号合成の位相）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数１２）、によって算出するステップと、

【数11】

$$W1(k)=X^{*}(2k,mM/K)\exp(-j\pi kB/M) \quad (0\leqq k\leqq M/(2L)-1)$$

【数12】

$$W1(k)=X(2M/L-1-2k,mM/K)\exp(j\pi B(1/L-k/M)) \quad (M/(2L)\leqq k\leqq M/L-1)$$

第１の中間信号から高速フーリエ変換で第２の中間信号ｗ２（ｎ）を（数１３）によって算出するステップと、

【数13】

$$w2(n)=\sum_{k=0}^{M/L-1} W1(k)\exp(-j2\pi knL/M) \quad (0\leqq n\leqq M/L-1)$$

０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数１４）（ただし、Ｒｅ（ｘ）は複素数ｘの実部、Ｉｍ（ｘ）はｘの虚部）によって算出するステップと、

【数14】

$$w3(n)=\mathrm{Re}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$$
$$w3(n+M/L)=\mathrm{Im}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\} \qquad (0\leqq n\leqq M/L-1)$$

第３の中間信号からサンプリング時刻ｍＭ／Ｋ＋ｎＬ／Ｋ（ただし、０≦ｎ≦Ｍ／Ｌ－１））における出力信号ｘ（ｍＭ／Ｋ＋ｎＬ／Ｋ）を（数１５）によって算出するステップを備えたことを特徴とする複素指数変調フィルタバンクの信号合成方法。

【数15】

$$x(mM/K+nL/K)=\sum_{i=0}^{N/(2M)-1}(-1)^i\{w3(4Mi/L+n)h(2Mi+Ln)+w3(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$$

【請求項5】

Ｍ個の等帯域幅のサンプリング周波数ｆｓＫ／Ｍ（ただし、Ｋは帯域分割数Ｍの約数で１を含む正整数）の複素帯域分割入力信号を合成し、１／Ｌ倍（ただし、ＬはＭの約数で１を含む正整数）にダウンサンプリングして、サンプリング周波数ｆｓＫ／Ｌの信号を出力する信号合成方法であって、直線位相非巡回形のプロトタイプフィルタのフィルタ係数をｈ（ｎ）（ただし、Ｎはフィルタ次数、０≦ｎ≦Ｎ）、最初と最後のフィルタ係数の値を零、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とするとき、

複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数１６）（ただし、ｊは虚数単位、Ｂは信号合成の位相）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数１７）（ただし、＊は共役複素数）、によって算出するステップと、

【数16】

$$W1(k)=X(2k,mM/K)\exp(j\pi kB/M) \quad (0\leqq k\leqq M/(2L)-1)$$

【数17】

$$W1(k)=X^*(2M/L-1-2k,mM/K)\exp(j\pi B(k/M-1/L)) \quad (M/(2L)\leqq k\leqq M/L-1)$$

第１の中間信号から逆高速フーリエ変換で第２の中間信号ｗ２（ｎ）を（数１８）によって算出するステップと、

【数18】

$$w2(n)=\sum_{k=0}^{M/L-1}W1(k)\exp(j2\pi knL/M) \quad (0\leqq n\leqq M/L-1)$$

０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数１９）（ただし、Ｒｅ（ｘ）は複素数ｘの実部、Ｉｍ（ｘ）はｘの虚部）によって算出するステップと、

【数19】

$$w3(n)=\mathrm{Re}\{w2(n)(1/M)\exp(j\pi(2Ln+B)/(4M))\}$$
$$w3(n+M/L)=-\mathrm{Im}\{w2(n)(1/M)\exp(j\pi(2Ln+B)/(4M))\} \qquad (0\leqq n\leqq M/L-1)$$

第３の中間信号からサンプリング時刻ｍＭ／Ｋ＋ｎＬ／Ｋ（ただし、０≦ｎ≦Ｍ／Ｌ－１））における出力信号ｘ（ｍＭ／Ｋ＋ｎＬ／Ｋ）を（数２０）によって算出するステップを備えたことを特徴とする複素指数変調フィルタバンクの信号合成方法。

【数20】

$$x(mM/K+nL/K)=\sum_{i=0}^{N/(2M)-1}(-1)^i\{w3(4Mi/L+n)h(2Mi+Ln)+w3(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$$

【請求項6】
信号合成の位相が、複素帯域分割入力信号から第１の中間信号を算出するステップで乗算が発生しないように、設定されていることを特徴とする請求項４ないし５のいずれかに記載の複素指数変調フィルタバンクの信号合成方法。
【請求項7】
信号分析の位相Ａと信号合成の位相Ｂが（数２１）（ただし、Ｎはプロトタイプフィルタの次数、Ｍは帯域分割数、Ｐは任意の整数）を満足することを特徴とする請求項１ないし６のいずれかに記載の複素指数変調フィルタバンクの信号分析と合成方法。
【数21】

$$A+B+2N=8MP$$

【請求項8】
請求項１ないし７のいずれかに記載の複素指数変調フィルタバンクの信号分析と合成方法を、コンピュータまたはデジタルシグナルプロセッサに実行させるためのプログラム。
【請求項9】
請求項１ないし７のいずれかに記載の複素指数変調フィルタバンクの信号分析と合成方法を、コンピュータまたはデジタルシグナルプロセッサに実行させるためのプログラムを記録したコンピュータで読み取り可能な記録媒体。
【発明の詳細な説明】
【技術分野】
【0001】
本発明は、演算量を削減して低消費電力のあるいは高速な複素指数変調フィルタバンクの信号分析と信号合成方法に関するものである。
【背景技術】
【0002】
近年、信号を複数の帯域分割信号に分割する信号分析フィルタバンクと前記帯域分割信号を合成して元の信号を再生する信号合成フィルタバンクは、オーディオ信号や画像信号のサブバンド符号化による高能率符号化を実現する信号分析と信号合成手段として注目されている。特に、少ない情報量で広帯域再生を実現するオーディオ信号帯域拡張技術において、複素指数変調フィルタバンクが注目されている。複素指数変調フィルタバンクは、ＭＰＥＧ（Moving Picture Experts Group）－１オーディオ等で使用されてきたコサイン変調フィルタバンクと異なり、帯域毎にゲインを変えても折り返し歪が発生しないという利点を有している。したがって各種デジタルイコライザーとしても利用可能である。しかしながら、複素指数変調フィルタバンクは、帯域分割信号を複素数で扱うので、帯域分割信号を実数で扱うコサイン変調フィルタバンクと比較して演算量が多いという欠点を有している。
【0003】
例えば、ＭＰＥＧ（Moving Picture Experts Group）ＡＡＣ（Advanced Audio Coding）の帯域拡張技術であるＡＡＣ＋ＳＢＲ（Spectral Band Replication）では、複素指数変調フィルタバンクが用いられている（非特許文献１参照）。
【0004】
以下に従来の複素指数変調フィルタバンクの信号分析と信号合成方法について説明する。図１２は信号分析フィルタバンクと信号合成フィルタバンクの構成を示すブロック図である。図１２において、１２０１は信号分析フィルタバンク、１２０２は信号合成フィルタバンク、Ｍは帯域分割数、１２０３はサンプリング周波数Ｋ倍補間器、１２０４は分析用帯域通過フィルタ、１２０５は間引き器、１２０６は補間器、１２０７は合成用帯域通

過フィルタ、１２０８は加算器、１２０９はサンプリング周波数１／Ｌ倍間引き器である。

【0005】

信号分析フィルタバンク１２０１は、サンプリング周波数Ｋ倍補間器１２０３とＭ個の分析用帯域通過フィルタ１２０４とＭ個の間引き器１２０５とから成り、信号合成フィルタバンク１２０２は、Ｍ個の補間器１２０６とＭ個の合成用帯域通過フィルタ１２０７と加算器１２０８とサンプリング周波数１／Ｌ倍間引き器１２０９とから成る。分析用帯域通過フィルタ１２０４と合成用帯域通過フィルタ１２０７は、互いにペアをなす。ここで、ＫとＬは帯域分割数Ｍの約数で１を含む正整数である。なお、図１２でサンプリング周波数Ｋ倍補間器１２０３、あるいはサンプリング周波数１／Ｌ倍間引き器１２０９を備えていない構成（すなわち、Ｋあるいは Ｌの値が１の構成）も可能である。

【0006】

最初に信号分析フィルタバンク１２０１の動作について説明する。サンプリング周波数Ｋ倍補間器１２０３では、サンプリング周波数ｆｓの入力信号に対して、１個のデータ毎に（Ｋ－１）個の零データを挿入することにより、サンプリング周波数をＫ倍上昇して、Ｋｆｓにする。次に、全帯域をＭ個の等帯域幅の帯域に分割する分析用帯域通過フィルタ１２０４によって帯域通過信号とし、間引き器１２０５でＭ個のデータ毎に（Ｍ－１）個のデータを間引き、１個のデータを出力することによってサンプリング周波数ｆｓＫ／Ｍの帯域分割信号に変換し、出力する。入力信号のサンプリング周波数をＫ倍にしているので、第Ｍ／Ｋ帯域以上の帯域分割信号は零である。

【0007】

次に、信号合成フィルタバンク１２０２の動作について説明する。信号分析フィルタバンク１２０１から出力されたサンプリング周波数ｆｓＫ／Ｍの帯域分割信号を入力とし、補間器１２０６で１個のデータ毎に（Ｍ－１）個の零データを挿入することによってサンプリング周波数をＭ倍のＫｆｓに上昇し、Ｍ個の等帯域幅の合成用帯域通過フィルタ１２０７により帯域通過信号とした後、加算器１２０８によって合成し、サンプリング周波数Ｋｆｓの信号を再生する。次に、サンプリング周波数１／Ｌ倍間引き器１２０９で、Ｌ個のデータ毎に（Ｌ－１）個のデータを間引き、サンプリング周波数ｆｓＫ／Ｌの信号を出力する。

【0008】

オーディオ信号や画像信号のサブバンド符号化では、信号分析フィルタバンクと信号合成フィルタバンクとの間で帯域分割信号の周波数方向の分布の偏りや人間の聴覚特性あるいは視覚特性を利用して情報圧縮を行い、高能率符号化を実現する。

【0009】

特許文献１に記載されているように複素指数変調フィルタバンクは、プロトタイプフィルタを複素指数変調することにより構成される。特許文献１より、直線位相非巡回形プロトタイプフィルタのフィルタ係数をｈ（ｎ）（ただし、０≦ｎ≦Ｎ、Ｎはフィルタ次数）とするとき、複素指数変調信号分析フィルタバンクの第ｋ帯域のｎサンプルのフィルタ係数ｈａ（ｋ，ｎ）は（数１）（ただし、ｊは虚数単位、Ａは信号分析の位相）で与えられる。

【0010】

【数1】

$$ha(k,n) = Kh(Kn)\exp(j\pi(k+0.5)(2Kn+A)/(2M))$$

【0011】

したがって、最初と最後のプロトタイプフィルタのフィルタ係数の値を零とするとき、信号分析フィルタバンクのサンプル時刻ｎにおける入力信号をｘ（ｎ）とすれば、第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプル時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）は（数２）で与えられる。

【0012】

【数2】

$$X(k,mM/K)=K\sum_{n=0}^{N/K-1}h(Kn)\exp(j\pi(k+0.5)(2Kn+A)/(2M))x(mM/K-n)$$

【0013】

（数２）を直接計算すると演算量が多くなるので、従来、以下のように演算量を削減した複素指数変調信号分析フィルタバンクが用いられてきた。図１３は従来の複素指数変調信号分析フィルタバンクの分析方法の処理ステップを示すフローチャートである。ステップ１３０１で、サンプリング時刻ｎの入力信号ｘ（ｎ）から中間信号ｗ（ｎ）を（数３）によって算出する。

【0014】

【数3】

$$w(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+Kn)\quad(0\leqq n\leqq 2M/K-1)$$

【0015】

次に、ステップ１３０２で中間信号ｗ（ｎ）から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を（数４）（ただし、Ａは信号分析の位相）によって算出する。

【0016】

【数4】

$$X(k,mM/K)=K\sum_{n=0}^{2M/K-1}w(n)\exp(j\pi(k+0.5)(2Kn+A)/(2M))\quad(0\leqq k\leqq M/K-1)$$

【0017】

例えば、非特許文献１の６０頁と６２頁に、帯域分割数Ｍが６４、プロトタイプフィルタのフィルタ次数Ｎが６４０、アップサンプリングの倍率Ｋが２、信号の位相Ａが－１の場合の複素指数変調信号分析フィルタバンクの例が記載されている。ただし、非特許文献１では、プロトタイプフィルタのフィルタ係数ｈ（ｎ）の代わりに、ｈ（ｎ）から（数５）で算出したｃ（ｎ）を用いている。

【0018】

【数5】

$$c(n)=(-1)^{INT(n/2M)}h(n)\quad(0\leqq n\leqq N-1)$$

【0019】

ここで、ＩＮＴ（ｘ）は、ｘの小数点以下を切り捨て整数化する関数である。

【0020】

図１３の従来例では、中間信号を導入することにより、（数２）を直接計算する場合と比較して演算量を削減することができる。

【0021】

ここで、図１３の複素指数変調信号分析フィルタバンクを実現するのに必要な演算量を実数加算回数と実数乗算回数を評価する。

【0022】

ステップ１３０１で、ｈ（ｎ）の代わりに（数５）であらかじめ算出したｃ（ｎ）をテーブルに記憶して用いることにすると、実数加算回数は（Ｎ／２Ｍ－１）（２Ｍ／Ｋ）＝Ｎ／Ｋ－２Ｍ／Ｋ回、実数乗算回数は（Ｎ／２Ｍ）（２Ｍ／Ｋ）＝Ｎ／Ｋ回である。

【0023】

ステップ１３０２で、Ｋｅｘｐ（ｊπ（ｋ＋０．５）（２Ｋｎ＋Ａ）／（２Ｍ））をあらかじめ算出してテーブルに記憶して用いることにすると、実数加算回数は２（２Ｍ／Ｋ－１）（Ｍ／Ｋ）＝４（Ｍ／Ｋ）^{2}－２（Ｍ／Ｋ）回であり、実数乗算回数は２（２Ｍ／

Ｋ）（Ｍ／Ｋ）＝４（Ｍ／Ｋ）２回である。

【0024】

プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，アップサンプリングの倍率Ｋを１とすれば、ステップ１３０１で、実数加算回数が５１２回、実数乗算回数が６４０回、ステップ１３０２で、実数加算回数が１６２５６回、実数乗算回数が１６３８４回、ステップ１３０１とステップ１３０２を合わせた合計で、実数加算回数が１６７６８回、実数乗算回数が１７０２４回である。

【0025】

次に、従来の複素指数変調信号合成フィルタバンクについて説明する。特許文献１より、複素指数変調信号合成フィルタバンクの第ｋ帯域のｎサンプルのフィルタ係数ｈｓ（ｋ，ｎ）は（数６）（ただし、Ｂは信号合成の位相）で与えられる。

【0026】

【数6】

$$hs(k,n) = (1/M)h(Ln)\exp(j\pi(k+0.5)(2Ln+B)/(2M))$$

【0027】

ここで、信号合成の位相Ｂは信号分析の位相Ａと（数７）（ただし、Ｐは任意の整数）の関係式を満足する。

【0028】

【数7】

$$A + B + 2N = 8MP$$

【0029】

入力の複素帯域分割信号を（数６）のフィルタ係数でたたみこんだ信号の有効帯域（第０帯域から第（Ｍ／Ｌ－１）帯域までの帯域）での総和の実数部が信号合成フィルタバンク１２０２の出力である。最初と最後のプロトタイプフィルタのフィルタ係数の値を零とするとき、複素指数変調信号合成フィルタバンクの第ｋ帯域のサンプル時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とすれば、サンプリング時刻ｍＭ／Ｋ＋ｎＬ／Ｋにおける出力信号ｘ（ｍＭ／Ｋ＋ｎＬ／Ｋ）は（数８）（ただし、Ｒｅ（ｘ）は複素数ｘの実部）で与えられる。

【0030】

【数8】

$$x(mM/K+nL/K) = \mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/L-1}\sum_{l=0}^{N/L-1}X(k,mM/K+L(n-l))h(Ll)\exp(j\pi(k+0.5)(2Ll+B)/(2M))\right\}$$

$$(0 \leqq n \leqq M/L-1)$$

【0031】

（数８）を直接計算すると演算量が多くなるので、複素指数変調信号分析フィルタバンクの場合と同様に、従来、以下のように演算量を削減した複素指数変調信号合成バンクが用いられてきた。図１４は従来の複素指数変調信号合成フィルタバンクの合成方法の処理ステップを示すフローチャートである。ステップ１４０１で、０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する中間信号ｗ（ｎ）をｗ（ｎ＋２Ｍ／Ｌ）にシフトし、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号Ｘ（ｋ，ｍＭ／Ｋ）から０≦ｎ≦２Ｍ／Ｌ－１に対する中間信号ｗ（ｎ）を（数９）によって算出する。

【0032】

【数9】

$$w(n)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/L-1}X(k,mM/K)\exp(j\pi(k+0.5)(2Ln+B)/(2M))\right\}\quad(0\leqq n\leqq 2M/L-1)$$

【0033】

次に、ステップ１４０２で中間信号ｗ（ｎ）からサンプリング時刻ｍＭ／Ｋ＋ｎＬ／Ｋ（ただし、０≦ｎ≦Ｍ／Ｌ－１））における出力信号ｘ（ｍＭ／Ｋ＋ｎＬ／Ｋ）を（数１０）によって算出する。

【0034】

【数10】

$$x(mM/K+nL/K)=\sum_{i=0}^{N/(2M)-1}(-1)^i\{w(4Mi/L+n)h(2Mi+Ln)+w(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$$

$$(0\leqq n\leqq M/L-1)$$

【0035】

非特許文献１の６０頁、６１頁、６３頁に、帯域分割数Ｍが６４、プロトタイプフィルタのフィルタ次数Ｎが６４０、アップサンプリングの倍率Ｋが２、ダウンサンプリングの倍率１／Ｌが１、信号合成の位相Ｂが－２５５の場合の複素指数変調信号合成フィルタバンクの例が記載されている。

【0036】

図１４の複素指数変調信号合成フィルタバンクを実現するのに必要な演算量を、実数加算回数と実数乗算回数を評価する。

【0037】

ステップ１４０１で、（１／Ｍ）ｅｘｐ（ｊπ（ｋ＋０．５）（２Ｌｎ＋Ｂ）／（２Ｍ））をあらかじめ算出し、テーブルに記憶して用いることにすると、実数加算回数は（Ｍ／Ｌ－１）（２Ｍ／Ｌ）＋（Ｍ／Ｌ）（２Ｍ／Ｌ）＝４（Ｍ／Ｌ）^{2}－２（Ｍ／Ｌ）回であり、実数乗算回数は２（Ｍ／Ｌ）（２Ｍ／Ｌ）＝４（Ｍ／Ｌ）^{2}回である。

【0038】

ステップ１４０２で、ｈ（ｎ）の代わりに（数５）で算出したｃ（ｎ）をテーブルに記憶して用いるとすると、実数加算回数は（Ｎ／Ｍ－１）（Ｍ／Ｌ）＝Ｎ／Ｌ－Ｍ／Ｌ回、実数乗算回数は（Ｎ／２Ｍ）２（Ｍ／Ｌ）＝Ｎ／Ｌ回である。

【0039】

プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，ダウンサンプリングの倍率１／Ｌを１とすれば、ステップ１４０１で、実数加算回数が１６２５６回、実数乗算回数が１６３８４回、ステップ１４０２で、実数加算回数が５７６回、実数乗算回数が６４０回、ステップ１４０１とステップ１４０２を合わせた合計で、実数加算回数が１６８３２回、実数乗算回数が１７０２４回である。

【特許文献1】国際公開第０２／０８０３６２号パンフレット

【非特許文献1】イソ／アイイーシー１４４９６－３：２００１（ＩＳＯ／ＩＥＣ１４４９３－３：２００１）、インフォメーション　テクノロジー　コーディング　オブ　オーディオ・ビジュアル　オブジェクト　パート３　オーディオ（Ｉｎｆｏｒｍａｔｉｏｎ　ｔｅｃｈｎｏｌｏｇｙ　－　Ｃｏｄｉｎｇ　ｏｆ　ａｕｄｉｏ－ｖｉｓｕａｌ　ｏｂｊｅｃｔｓ　－　Ｐａｒｔ　３）、２００３年３月　（ＦＤＡＭ１の本文、帯域拡張（Ｂａｎｄｗｉｄｔｈ　Ｅｘｔｅｎｓｉｏｎ））

【発明の開示】

【発明が解決しようとする課題】

【0040】

しかしながら、上記従来の複素指数変調フィルタバンクの信号分析方法と信号合成方法

では、演算量が多い、という問題点を有していた。したがって、デジタルシグナルプロセッサのように限られた数の乗算器と加算器を時分割で使ってフィルタバンクを実装するときの動作クロック周波数が高くなり、消費電力が大きい、という課題を有していた。また、コンピュータを使ってソフトウェアでフィルタバンクを実行するときの処理時間が大きいという課題を有していた。

【0041】

本発明は上記従来の問題点を解決するもので、複素指数変調フィルタバンク実行時の演算量を削減し、消費電力を低減した信号分析方法と信号合成方法を提供することを目的とする。また、複素指数変調フィルタバンク実行時の処理時間を短縮し、高速化した信号分析方法と信号合成方法を提供することを目的とする。

【課題を解決するための手段】

【0042】

この課題を解決するために、本発明の信号分析方法は、サンプリング周波数ｆｓの入力信号をＫ倍にアップサンプリングして、Ｍ個の等帯域幅のサンプリング周波数ｆｓＫ／Ｍの複素帯域分割信号に分割して出力する複素指数変調フィルタバンクの信号分析方法あって、入力信号から第１の中間信号を算出するステップと、第１の中間信号から第２の中間信号を算出するステップと、第２の中間信号からＭ／Ｋ点の高速フーリエ変換あるいは逆高速フーリエ変換で第３の中間信号を算出するステップと、第３の中間信号から複素帯域分割出力信号を算出するステップと、を備えたことを特徴とする。

【0043】

また、本発明の信号合成方法は、Ｍ個の等帯域幅のサンプリング周波数ｆｓＫ／Ｍの複素帯域分割入力信号を合成し、１／Ｌ倍にダウンサンプリングして、サンプリング周波数ｆｓＫ／Ｌの信号を出力する複素指数変調フィルタバンクの信号合成方法であって、複素帯域分割入力信号から第１の中間信号を算出するステップと、第１の中間信号からＭ／Ｌ点の高速フーリエ変換あるいは逆高速フーリエ変換で第２の中間信号を算出するステップと、第３の中間信号をシフトし、次に第２の中間信号から第３の中間信号を算出するステップと、第３の中間信号から出力信号を算出するステップと、を備えたことを特徴とする。

【0044】

これにより、複素指数関数の周期性を利用して、高速フーリエ変換あるいは逆高速フーリエ変換を効率的に適用することができるので、複素指数変調フィルタバンクの演算量を削減することができる。

【0045】

また、上記信号分析方法と信号合成方法は、実数部と虚数部を分離せずに複素数として扱うので、信号分析の位相あるいは信号合成の位相を演算量が小さくなるように設定することにより、演算量をさらに削減することができる。

【発明の効果】

【0046】

本発明によれば、複素指数変調フィルタバンク実行時の演算量を削減することができる。したがって、デジタルシグナルプロセッサやＬＳＩでフィルタバンクを実装するときの動作クロック周波数を低減することができ、低消費電力で複素指数変調フィルタバンクを実現することができる。また、複素指数変調フィルタバンクをコンピュータを使ってソフトウェアで実行する時の処理時間を短縮することができ、処理の高速化を実現することができる。

【発明を実施するための最良の形態】

【0047】

本発明の複素指数変調フィルタバンクの信号分析方法は、サンプリング周波数ｆｓの信号をＫ倍（ただし、Ｋは帯域分割数Ｍの約数で１を含む正整数）にアップサンプリングしてＭ個の等帯域幅のサンプリング周波数がｆｓＫ／Ｍの複素帯域分割信号に分割し、また、本発明の複素指数変調フィルタバンクの信号合成方法は、Ｍ個の等帯域幅のサンプリン

グ周波数がｆｓＫ／Ｍの複素帯域分割信号を合成し、１／Ｌ倍（ただし、ＬはＭの約数で１を含む正整数）にダウンサンプリングして、サンプリング周波数ｆｓＫ／Ｌの信号を出力する。複素指数変調フィルタバンクはプロトタイプフィルタを複素指数変調することによって構成される。以下の実施の形態では、直線位相非巡回形プロトタイプフィルタのフィルタ次数をＮ、フィルタ係数をｈ（ｎ）（０≦ｎ≦Ｎ）、最初と最後のフィルタ係数の値を零（ｈ（０）＝ｈ（Ｎ）＝０）とする。

【0048】

以下本発明の実施の形態について、図面を参照しながら説明する。

【0049】

（実施の形態１）

図１は、本発明の実施の形態１における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャートである。

【0050】

ステップ１０１では、入力信号から第１の中間信号を算出する。すなわち、サンプリング時刻ｎにおける入力信号をｘ（ｎ）とするとき、入力信号から第１の中間信号ｗ１（ｎ）（ただし、０≦ｋ≦２Ｍ／Ｋ－１）を（数１１）によって算出する。

【0051】

【数11】

$$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+Kn)\quad(0\leqq n\leqq 2M/K-1)$$

【0052】

ステップ１０１は、図１３の従来例のステップ１３０１と同一である。ｈ（ｎ）の代わりに（数１２）であらかじめ算出したｃ（ｎ）をテーブルに記憶して用いることにすると、（数１１）は（数１３）となる。

【0053】

【数12】

$$c(n)=(-1)^{INT(n/2M)}h(n)\quad(0\leqq n\leqq N-1)$$

【0054】

【数13】

$$w1(n)=\sum_{i=0}^{N/(2M)-1}x(mM/K-2Mi/K-n)c(2Mi+Kn)\quad(0\leqq n\leqq 2M/K-1)$$

【0055】

したがって、ステップ１０１を実行するのに必要な実数加算回数は（Ｎ／２Ｍ－１）（２Ｍ／Ｋ）＝Ｎ／Ｋ－２Ｍ／Ｋ回、実数乗算回数は（Ｎ／２Ｍ）（２Ｍ／Ｋ）＝Ｎ／Ｋ回である。

【0056】

ステップ１０２では、第１の中間信号から（数１４）（ただし、ｊは虚数単位）によって第２の中間信号ｗ２（ｎ）（ただし、０≦ｎ≦２Ｍ／Ｋ－１）を算出する。

【0057】

【数14】

$$w2(n)=(w1(n)-jw1(n+M/K))K\exp(-j\pi nK/(2M))\quad(0\leqq n\leqq M/K-1)$$

【0058】

ここで、Ｋｅｘｐ（－ｊπｎＫ／（２Ｍ））をテーブルに記憶して用いることにすると、１回の複素数同士の乗算（複素乗算）は、２回の実数加算と４回の実数乗算で実行できるので、ステップ１０２を実行するのに必要な実数加算回数は２Ｍ／Ｋ回、実数乗算回数は４Ｍ／Ｋ回である。

【0059】
ステップ１０３では第２の中間信号から高速フーリエ変換（ＦＦＴ：Fast Fourier Transform）で（数１５）によって第３の中間信号Ｗ３（ｋ）（ただし、０≦ｎ≦Ｍ／Ｋ－１）を算出する。
【0060】
【数15】

$$W3(k)=\sum_{n=0}^{M/K-1} w2(n)\exp(-j2\pi knK/M) \quad (0\leqq k\leqq M/K-1)$$

【0061】
Ｍ／Ｋ点のＦＦＴでは、（Ｍ／（２Ｋ））ｌｏｇ$_2$（Ｍ／Ｋ）回のバタフライ演算が必要である。１回のバタフライ演算には２回の複素加算と１回の複素乗算が含まれるので、これを実数演算に直すと、６回の実数加算と４回の実数乗算が必要である。したがって、ステップ１０３を実行するのに必要な実数加算回数は（３Ｍ／Ｋ）ｌｏｇ$_2$（Ｍ／Ｋ）回で、実数乗算回数は（２Ｍ／Ｋ）ｌｏｇ$_2$（Ｍ／Ｋ）回である。
【0062】
ステップ１０４で、第３の中間信号から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を、ｋが偶数（ｋ＝２ｌ）の場合には（数１６）（ただし、＊は共役複素数、Ａは信号分析の位相）、ｋが奇数（ｋ＝２ｌ＋１）の場合には（数１７）、によって算出する。
【0063】
【数16】

$$X(2l,mM/K)=W3^{*}(l)\exp(j\pi(4l+1)A/(4M)) \quad (0\leqq l\leqq M/(2K)-1)$$

【0064】
【数17】

$$X(2l+1,mM/K)=W3(M/K-1-l)\exp(j\pi(4l+3)A/(4M)) \quad (0\leqq l\leqq M/(2K)-1)$$

【0065】
入力信号をＫ倍にアップサンプリングしているので、複素帯域分割出力信号の有効な帯域は第０帯域から第（Ｍ／Ｋ－１）帯域までで、第Ｍ／Ｋ帯域以上の複素帯域分割出力信号は零である。
【0066】
ステップ１０４を実行するには、Ｍ／Ｋ回の複素乗算が必要なので、２Ｍ／Ｋ回の実数加算と４Ｍ／Ｋ回の実数乗算が必要である。
【0067】
したがって、図１の実施の形態１を実行するのに必要な演算回数を、上記各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｋ＋２Ｍ／Ｋ＋（３Ｍ／Ｋ）ｌｏｇ$_2$（Ｍ／Ｋ）回、実数乗算回数はＮ／Ｋ＋８Ｍ／Ｋ＋（２Ｍ／Ｋ）ｌｏｇ$_2$（Ｍ／Ｋ）回である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，アップサンプリングの倍率Ｋを１とすれば、実施の形態１を実行するのに必要な演算量は、実数加算回数が１９２０回、実数乗算回数が１９２０回である。これを従来例（実数加算回数１６７６８回、実数乗算回数１７０２４回）と比較すると、実施の形態１では、実数加算回数、実数乗算回数をともに従来例の１／８以下に削減することができる。
【0068】
実施の形態１の動作についてさらに説明する。
【0069】
（数１４）を（数１５）に代入し、式を変形すると（数１８）が成立する。（数１８）の最後の式を（数１６）に代入すると（数１９）が成立する。
【0070】

【数18】

$$W3(k)=\sum_{n=0}^{M/K-1}(w1(n)-jw1(n+M/K)K\exp(-j\pi(2k+0.5)nK/M)$$
$$=\sum_{n=0}^{M/K-1}w1(n)K\exp(-j\pi(2k+0.5)nK/M)+\sum_{n=0}^{M/K-1}(-j)w1(n+M/K)K\exp(-j\pi(2k+0.5)nK/M)$$
$$=\sum_{n=0}^{M/K-1}w1(n)K\exp(-j\pi(2k+0.5)nK/M)+\sum_{n=0}^{M/K-1}w1(n+M/K)K\exp(-j\pi(2k+0.5)(n+M/K)K/M)$$

$$W3(k)=\sum_{n=0}^{2M/K-1}w1(n)K\exp(-j\pi(2k+0.5)nK/M)$$

【0071】

【数19】

$$X(2l,mM/K)=W3^{*}(l)\exp(j\pi(4l+1)A/(4M))$$
$$=\exp(j\pi(4l+1)A/(4M))\sum_{n=0}^{2M/K-1}w1(n)K\exp(j\pi(2l+0.5)nK/M)$$
$$=K\sum_{n=0}^{2M/K-1}w1(n)\exp(j\pi(2l+0.5)(2Kn+A)/(2M))$$

【0072】

（数１９）は（数４）で、ｗをｗ１に置き換え、ｋ＝２１としたときの式と一致する。同様にして（数１７）は（数４）でｗをｗ１に置き換え、ｋ＝２１＋１としたときの式と一致することが説明できる。したがって、実施の形態１は、図１３の従来例と同一の結果を算出することができる。

【0073】

図２は、実施の形態１の複素指数変調フィルタバンクの信号分析方法を実行する装置の構成を示すブロック図である。図２で、２０１は第１の中間信号生成部、２０２は第２の中間信号生成部、２０３はＦＦＴ部（第３の中間信号生成部）、２０４は帯域分割出力信号生成部である。第１の中間信号生成部２０１では、入力信号とプロトタイプフィルタのフィルタ係数から（数１３）を実行して第１の中間信号を出力する。第２の中間信号生成部２０２では、第１の中間信号生成部２０１の出力した第１の中間信号から（数１４）を実行して第２の中間信号を出力する。ＦＦＴ部２０３では、第２の中間信号生成部２０２の出力した第２の中間信号に対してＭ／Ｋ点の高速フーリエ変換で（数１５）を実行して結果を第３の中間信号として出力する。帯域分割出力信号生成部２０４では、ＦＦＴ部２０３の出力した第３の中間信号から（数１６）と（数１７）にしたがって複素帯域分割出力信号を算出し、出力する。

【0074】

以上のように実施の形態１の複素指数変調フィルタバンクの信号分析方法では、複素指数関数の周期性を利用して高速フーリエ変換を効率的に適用することにより、複素指数変調信号分析フィルタバンクの実行に必要な演算量を削減することができる。これによってデジタルシグナルプロセッサやＬＳＩで複素指数変調信号分析フィルタバンクを実装するときの動作クロック周波数を低減することができ、低消費電力で複素指数変調信号分析フィルタバンクを実現することができる。また、コンピュータを使ってソフトウェアで複素指数変調信号分析フィルタバンクを実行する時の処理時間を短縮することができ、処理の高速化を実現することができる。

【0075】

（実施の形態２）

図３は、本発明の実施の形態２における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャートである。

【0076】
実施の形態２は、実施の形態１において信号分析の位相Ａを８ＭＰ（ただし、Ｐは任意の整数）に設定したものである。
【0077】
図３のステップ３０１、３０２、３０３は、図１のステップ１０１、１０２、１０３、とそれぞれ同一であり、説明を省略する。
【0078】
ステップ３０４で、第３の中間信号から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を、ｋが偶数（ｋ＝２ｌ）の場合には（数２０）、ｋが奇数（ｋ＝２ｌ＋１）の場合には（数２１）、によって算出する。
【0079】
【数20】

$$X(2l,mM/K)=W3^{*}(l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0080】
【数21】

$$X(2l+1,mM/K)=W3(M/K-1-l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0081】
実施の形態１の図１のステップ１０４を実行するには、Ｍ／Ｋ回の複素乗算が必要で、２Ｍ／Ｋ回の実数加算と４Ｍ／Ｋ回の実数乗算が必要であったが、ステップ３０４では複素乗算は不要であり、その分演算量を削減できる。
【0082】
したがって、図３の実施の形態２を実行するのに必要な演算回数を、図３の各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｋ＋（３Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回、実数乗算回数はＮ／Ｋ＋４Ｍ／Ｋ＋（２Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，アップサンプリングの倍率Ｋを１とすれば、実施の形態２を実行するのに必要な演算量は、実数加算回数が１７９２回、実数乗算回数が１６６４回である。これは実施の形態１と比較して、実数加算回数が１２８回、実数乗算回数が２５６回少ない。
【0083】
以上のように実施の形態２の複素指数変調フィルタバンクの信号分析方法では、信号分析の位相を特定の値に設定することによって、実施の形態１と比較して演算量をさらに削減することができる。
【0084】
なお、実施の形態２では、信号分析の位相Ａを８ＭＰに設定したが、信号の位相Ａを次の（２）、（３）、（４）のいずれかに設定してもよい。
（１）Ａ＝８ＭＰ
（２）Ａ＝２Ｍ＋８ＭＰ
（３）Ａ＝４Ｍ＋８ＭＰ
（４）Ａ＝６Ｍ＋８ＭＰ
信号分析の位相Ａを（２）、（３）、（４）のいずれに設定した場合においても、第３の中間信号から複素帯域分割出力信号を算出するステップで複素乗算は不要となり、実施の形態１と比較して演算量を削減することができる。
【0085】
（２）は信号分析の位相Ａを２Ｍ＋８ＭＰに設定した場合で、ｋが偶数（ｋ＝２１）のときには（数２２）、ｋが奇数（ｋ＝２１＋１）のときには（数２３）、によって複素帯域分割出力信号を算出する。
【0086】

【数22】

$$X(2l,mM/K) = jW3^{*}(l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0087】

【数23】

$$X(2l+1,mM/K) = -jW3(M/K-1-l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0088】

（数２２）で、虚数単位ｊを掛けることは、実数部を虚数部とし、かつ虚数部の符号を反転して実数部とすることであり、乗算は不要である。同様に（数２３）で、－ｊを掛けることは実数部の符号を反転して虚数部とし、かつ虚数部を実数部とすることであり、乗算は不要である。

【0089】

（３）は信号分析の位相Ａを４Ｍ＋８ＭＰに設定した場合で、ｋが偶数（ｋ＝２１）のときには（数２４）、ｋが奇数（ｋ＝２１＋１）のときには（数２５）、によって複素帯域分割出力信号を算出する。

【0090】

【数24】

$$X(2l,mM/K) = -W3^{*}(l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0091】

【数25】

$$X(2l+1,mM/K) = -W3(M/K-1-l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0092】

（４）は信号分析の位相Ａを６Ｍ＋８ＭＰに設定した場合で、ｋが偶数（ｋ＝２１）のときには（数２６）、ｋが奇数（ｋ＝２１＋１）のときには（数２７）、によって複素帯域分割出力信号を算出する。

【0093】

【数26】

$$X(2l,mM/K) = -jW3^{*}(l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0094】

【数27】

$$X(2l+1,mM/K) = jW3(M/K-1-l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0095】

以上のように実施の形態２の複素指数変調フィルタバンクの信号分析方法では、信号分析の位相を、第３の中間信号から複素帯域分割出力信号を算出するステップで乗算が発生しないように設定することによって、実施の形態１と比較して演算量をさらに削減することができる。

【0096】

（実施の形態３）

図４は、本発明の実施の形態３における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャートである。実施の形態１では高速フーリエ変換を効率的に適用することにより、複素指数変調信号分析フィルタバンクの実行に必要な演算量を削減したが、実施の形態３では、逆高速フーリエ変換（ＩＦＦＴ：Inverse Fast Fourier Transform）を効率的に適用することにより複素指数変調信号分析フィルタバンクの実行に必要な演算量を削減する。図４においてステップ４０１は、図１のステップ１０１と同一であり、説明を省略する。

【0097】
ステップ４０２では、第１の中間信号から（数２８）によって第２の中間信号ｗ２（ｎ）（ただし、０≦ｎ≦２Ｍ／Ｋ－１）を算出する。
【0098】
【数28】

$$w2(n) = (w1(n) + jw1(n + M/K))K\exp(j\pi nK/(2M)) \quad (0 \leqq n \leqq M/K - 1)$$

【0099】
ここで、Ｋｅｘｐ（ｊπｎＫ／（２Ｍ））をテーブルに記憶して用いることにすると、ステップ４０２を実行するのに必要な複素乗算回数はＭ／Ｋ回であり、これを実数演算に直すと実数加算回数は２Ｍ／Ｋ回、実数乗算回数は４Ｍ／Ｋ回である。
【0100】
ステップ４０３では第２の中間信号から逆高速フーリエ変換で（数２９）によって第３の中間信号Ｗ３（ｋ）（ただし、０≦ｎ≦Ｍ／Ｋ－１）を算出する。
【0101】
【数29】

$$W3(k) = \sum_{n=0}^{M/K-1} w2(n)\exp(j2\pi knK/M) \quad (0 \leqq k \leqq M/K - 1)$$

【0102】
Ｍ／Ｋ点のＩＦＦＴは、ＦＦＴと同様に（Ｍ／（２Ｋ））$\log_2$（Ｍ／Ｋ）回のバタフライ演算が必要である。したがって、ステップ４０３を実行するのに必要な実数加算回数は（３Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回で、実数乗算回数は（２Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回である。
【0103】
ステップ４０４で、第３の中間信号から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を、ｋが偶数（ｋ＝２ｌ）の場合には（数３０）（ただし、Ａは信号分析の位相）、ｋが奇数（ｋ＝２ｌ＋１）の場合には（数３１）、によって算出する。
【0104】
【数30】

$$X(2l, mM/K) = W3(l)\exp(j\pi(4l+1)A/(4M)) \quad (0 \leqq l \leqq M/(2K) - 1)$$

【0105】
【数31】

$$X(2l+1, mM/K) = W3^*(M/K - 1 - l)\exp(j\pi(4l+3)A/(4M)) \quad (0 \leqq l \leqq M/(2K) - 1)$$

【0106】
ステップ４０４を実行するには、Ｍ／Ｋ回の複素乗算が必要なので、２Ｍ／Ｋ回の実数加算と４Ｍ／Ｋ回の実数乗算が必要である。
【0107】
したがって、図４の実施の形態３を実行するのに必要な演算回数を、上記各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｋ＋２Ｍ／Ｋ＋（３Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回、実数乗算回数はＮ／Ｋ＋８Ｍ／Ｋ＋（２Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回である。これは、実施の形態１と同じ演算量である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４、アップサンプリングの倍率Ｋを１とすれば、実施の形態３を実行するのに必要な演算量は、実数加算回数が１９２０回、実数乗算回数が１９２０回である。これを従来例（実数加算回数１６７６８回、実数乗算回数１７０２４回）と比較すると、実施の形態３では、実数加算回数、実数乗算回数をともに従来例の１／８以下に削減することができる。

【0108】
実施の形態３のステップ４０２、４０３、４０４が、従来例の図１３のステップ１３０２と同一の結果を出力することができることは、実施の形態１で説明したのと同様に、（数２８）を（数２９）に代入し、代入結果をさらに（数３０）と（数３１）に代入することにより説明できる。
【0109】
以上のように実施の形態３の複素指数変調フィルタバンクの信号分析方法では、複素指数関数の周期性を利用して逆高速フーリエ変換を効率的に適用することにより、複素指数変調信号分析フィルタバンクの実行に必要な演算量を削減することができる。これによってデジタルシグナルプロセッサやＬＳＩで複素指数変調信号分析フィルタバンクを実装するときの動作クロック周波数を低減することができ、低消費電力で複素指数変調信号分析フィルタバンクを実現することができる。また、コンピュータを使ってソフトウェアで複素指数変調信号分析フィルタバンクを実行する時の処理時間を短縮することができ、処理の高速化を実現することができる。
【0110】
（実施の形態４）
図５は、本発明の実施の形態４における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャートである。
【0111】
実施の形態４は、実施の形態３において信号分析の位相Ａを８ＭＰ（ただし、Ｐは任意の整数）に設定したものである。
【0112】
図５のステップ５０１、５０２、５０３は、図４のステップ４０１、４０２、４０３とそれぞれ同一であり、説明を省略する。
【0113】
ステップ５０４で、第３の中間信号から第ｋ帯域（ただし、０≦ｋ≦Ｍ／Ｋ－１）のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割出力信号Ｘ（ｋ，ｍＭ／Ｋ）を、ｋが偶数（ｋ＝２ｌ）の場合には（数３２）、ｋが奇数（ｋ＝２ｌ＋１）の場合には（数３３）、によって算出する。
【0114】
【数32】

$$X(2l,mM/K)=W3(l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0115】
【数33】

$$X(2l+1,mM/K)=W3^{*}(M/K-1-l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0116】
図４のステップ４０４を実行するには、Ｍ／Ｋ回の複素乗算が必要で、２Ｍ／Ｋ回の実数加算と４Ｍ／Ｋ回の実数乗算が必要であったが、ステップ５０４では複素乗算は不要であり、その分演算量を削減できる。
【0117】
したがって、図５の実施の形態４を実行するのに必要な演算回数を、図５の各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｋ＋（３Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回、実数乗算回数はＮ／Ｋ＋４Ｍ／Ｋ＋（２Ｍ／Ｋ）$\log_2$（Ｍ／Ｋ）回である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，アップサンプリングの倍率Ｋを１とすれば、実施の形態４を実行するのに必要な演算量は、実数加算回数が１７９２回、実数乗算回数が１６６４回である。これは実施の形態３と比較して、実数加算回数が１２８回、実数乗算回数が２５６回少ない。
【0118】

以上のように実施の形態４の複素指数変調フィルタバンクの信号分析方法では、信号分析の位相を特定の値に設定することによって、実施の形態３と比較して演算量をさらに削減することができる。

【0119】

なお、実施の形態４では、信号分析の位相Ａを８ＭＰに設定したが、信号の位相Ａを次の（２）、（３）、（４）のいずれかに設定してもよい。

（１）Ａ＝８ＭＰ

（２）Ａ＝２Ｍ＋８ＭＰ

（３）Ａ＝４Ｍ＋８ＭＰ

（４）Ａ＝６Ｍ＋８ＭＰ

信号分析の位相Ａを（２）、（３）、（４）のいずれに設定した場合においても、第３の中間信号から複素帯域分割出力信号を算出するステップで複素乗算は不要となり、実施の形態３と比較して演算量を削減することができる。

【0120】

（２）は信号分析の位相Ａを２Ｍ＋８ＭＰに設定した場合で、ｋが偶数（ｋ＝２１）のときには（数３４）、ｋが奇数（ｋ＝２１＋１）のときには（数３５）、によって複素帯域分割出力信号を算出する。

【0121】

【数34】

$$X(2l,mM/K)=jW3(l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0122】

【数35】

$$X(2l+1,mM/K)=-jW3^{*}(M/K-1-l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0123】

（数３４）で、虚数単位ｊを掛けることは、実数部を虚数部とし、かつ虚数部の符号を反転して実数部とすることであり、乗算は不要である。同様に（数３５）で、－ｊを掛けることは実数部の符号を反転して虚数部とし、かつ虚数部を実数部とすることであり、乗算は不要である。

【0124】

（３）は信号分析の位相Ａを４Ｍ＋８ＭＰに設定した場合で、ｋが偶数（ｋ＝２１）のときには（数３６）、ｋが奇数（ｋ＝２１＋１）のときには（数３７）、によって複素帯域分割出力信号を算出する。

【0125】

【数36】

$$X(2l,mM/K)=-W3(l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0126】

【数37】

$$X(2l+1,mM/K)=-W3^{*}(M/K-1-l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0127】

（４）は信号分析の位相Ａを６Ｍ＋８ＭＰに設定した場合で、ｋが偶数（ｋ＝２１）のときには（数３８）、ｋが奇数（ｋ＝２１＋１）のときには（数３９）、によって複素帯域分割出力信号を算出する。

【0128】

【数38】

$$X(2l,mM/K)=-jW3(l)\quad(0\leqq l\leqq M/(2K)-1)$$

【0129】
【数39】

$$X(2l+1,mM/K) = jW3^{*}(M/K-1-l) \quad (0 \leqq l \leqq M/(2K)-1)$$

【0130】
以上のように実施の形態４の複素指数変調フィルタバンクの信号分析方法では、信号分析の位相を、第３の中間信号から複素帯域分割出力信号を算出するステップで乗算が発生しないように設定することによって、実施の形態３と比較して演算量をさらに削減することができる。
【0131】
（実施の形態５）
図６は、本発明の実施の形態５における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャートである。
【0132】
ステップ６０１では、複素帯域分割入力信号から第１の中間信号を算出する。すなわち、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とするとき、複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数４０）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数４１）、によって算出する。
【0133】
【数40】

$$W1(k) = X^{*}(2k,mM/K)\exp(-j\pi kB/M) \quad (0 \leqq k \leqq M/(2L)-1)$$

【0134】
【数41】

$$W1(k) = X(2M/L-1-2k,mM/K)\exp(j\pi B(1/L-k/M)) \quad (M/(2L) \leqq k \leqq M/L-1)$$

【0135】
Ｂは信号合成の位相であり、信号分析の位相Ａと（数４２）（ただし、Ｐは任意の整数）の関係式を満足する。
【0136】
【数42】

$$A+B+2N = 8MP$$

【0137】
ここで、ｅｘｐ（－ｊπｋＢ／Ｍ）とｅｘｐ（－ｊπＢ（１／Ｌ－ｋ／Ｍ））をテーブルに記憶して用いることにすると、ステップ６０１を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。
【0138】
ステップ６０２で、第１の中間信号から高速フーリエ変換で（数４３）によって第２の中間信号ｗ２（ｎ）（ただし、０≦ｎ≦Ｍ／Ｌ－１）を算出する。
【0139】
【数43】

$$w2(n) = \sum_{k=0}^{M/L-1} W1(k)\exp(-j2\pi knL/M) \quad (0 \leqq n \leqq M/L-1)$$

【0140】
ステップ６０２でＭ／Ｌ点のＦＦＴを実行するのに必要な実数加算回数は（３Ｍ／Ｌ）ｌｏｇ$_2$（Ｍ／Ｌ）回で、実数乗算回数は（２Ｍ／Ｌ）ｌｏｇ$_2$（Ｍ／Ｌ）回である。

【0141】
ステップ６０３で、０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、次に、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数４４）（ただし、Ｒｅ（ｘ）は複素数ｘの実部、Ｉｍ（ｘ）はｘの虚部）によって算出する。
【0142】
【数44】

$$w3(n) = \mathrm{Re}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$$
$$w3(n+M/L) = \mathrm{Im}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\} \qquad (0 \leqq n \leqq M/L-1)$$

【0143】
ステップ６０３で、（１／Ｍ）ｅｘｐ（－ｊπ（２Ｌｌ＋Ｂ）／（４Ｍ））をテーブルに記憶して用いることにすると、ステップ６０３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。
【0144】
ステップ６０４で第３の中間信号からサンプリング時刻ｍＭ／Ｋ＋ｎＬ／Ｋ（ただし、０≦ｎ≦Ｍ／Ｌ－１））における出力信号ｘ（ｍＭ／Ｋ＋ｎＬ／Ｋ）を（数４５）によって算出する。
【0145】
【数45】

$$x(mM/K+nL/K) = \sum_{i=0}^{N/(2M)-1}(-1)^i\{w3(4Mi/L+n)h(2Mi+Ln)+w3(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$$

【0146】
ステップ６０４は、図１４の従来例のステップ１４０２と同一である。
【0147】
ｈ（ｎ）の代わりに（数４６）であらかじめ算出したｃ（ｎ）をテーブルに記憶して用いることにすると、（数４５）は（数４７）となる。
【0148】
【数46】

$$c(n) = (-1)^{INT(n/2M)}h(n) \quad (0 \leqq n \leqq N-1)$$

【0149】
【数47】

$$x(mM/K+nL/K) = \sum_{i=0}^{N/(2M)-1}\{w3(4Mi/L+n)c(2Mi+Ln)+w3(4Mi/L+3M/L+n)c(2Mi+M+Ln)\}$$

$$(0 \leqq n \leqq M/L-1)$$

【0150】
したがって、ステップ６０４を実行するのに必要な実数加算回数は（Ｎ／Ｍ－１）（Ｍ／Ｌ）＝Ｎ／Ｌ－Ｍ／Ｌ回、実数乗算回数は（Ｎ／２Ｍ）２（Ｍ／Ｌ）＝Ｎ／Ｌ回である。
【0151】
図６の実施の形態５を実行するのに必要な演算回数を、上記各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｌ＋３Ｍ／Ｌ＋（３Ｍ／Ｌ）$\log_2$（Ｍ／Ｌ）回、実数乗算回数はＮ／Ｌ＋８Ｍ／Ｌ＋（２Ｍ／Ｌ）$\log_2$（Ｍ／Ｌ）回である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，ダウンサンプリングの倍率１／Ｌ

を１とすれば、実施の形態５を実行するのに必要な演算量は、実数加算回数が１９８４回、実数乗算回数が１９２０回である。これを従来例（実数加算回数１６８３２回、実数乗算回数１７０２４回）と比較すると、実施の形態５では、実数加算回数、実数乗算回数をともに従来例の１／８以下に削減することができる。

【0152】

実施の形態５の動作についてさらに説明する。

【0153】

（数４０）と（数４１）を（数４３）に代入し、式を変形すると（数４８）が成立する。（数４８）を（数４４）の上の式に代入し、式を変形すると（数４９）が成立する。

【0154】

【数48】

$$w2(n)=\sum_{k=0}^{M/(2L)-1}X^{*}(2k,mM/K)\exp(-j\pi k(2Ln+B)/M)+\sum_{k=-M/(2L)}^{M/L-1}X(2M/L-1-2k,mM/K)\exp(j\pi B/L)\exp(-j\pi k(2Ln+B)/M)$$

【0155】

【数49】

$$w3(n)=\mathrm{Re}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$$

$$w3(n)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/(2L)-1}X^{*}(2k,mM/K)\exp(-j\pi(2k+0.5)(2Ln+B)/(2M))\right\}+\mathrm{Re}\left\{(1/M)\sum_{k=-M/(2L)}^{M/L-1}X(2M/L-1-2k,mM/K)\exp(j\pi B/L)\exp(-j\pi(2k+0.5)(2Ln+B)/(2M))\right\}$$

$$w3(n)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/(2L)-1}X(2k,mM/K)\exp(j\pi(2k+0.5)(2Ln+B)/(2M))\right\}+\mathrm{Re}\left\{(1/M)\sum_{k=-M/(2L)}^{M/L-1}X(2M/L-1-2k,mM/K)\exp(j\pi(2M/L-2k-0.5)(2Ln+B)/(2M))\right\}$$

$$w3(n)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/(2L)-1}X(2k,mM/K)\exp(j\pi(2k+0.5)(2Ln+B)/(2M))\right\}+\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/(2L)-1}X(2k+1,mM/K)\exp(j\pi(2k+1+0.5)(2Ln+B)/(2M))\right\}$$

$$w3(n)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/L-1}X(k,mM/K)\exp(j\pi(k+0.5)(2Ln+B)/(2M))\right\}$$

【0156】

（数４９）の最後の式は、（数９）でｗをｗ３で置き換え、ｎの値の範囲を０≦ｎ≦Ｍ／Ｌ－１としたときの式と一致する。

【0157】

同様に（数４８）を（数４４）の下の式に代入し、式を変形すると（数５０）が成立する。

【0158】

【数50】

$$w3(n+M/L)=\mathrm{Im}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$$

$$w3(n+M/L)=\mathrm{Re}\{(-j)w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$$

$$w3(n+M/L)=\mathrm{Re}\left\{(-j)(1/M)\sum_{k=0}^{M/(2L)-1}X^{*}(2k,mM/K)\exp(-j\pi(2k+0.5)(2Ln+B)/(2M))\right\}$$
$$+\mathrm{Re}\left\{(-j)(1/M)\sum_{k=M/(2L)}^{M/L-1}X(2M/L-1-2k,mM/K)\exp(j\pi B/L)\exp(-j\pi(2k+0.5)(2Ln+B)/(2M))\right\}$$

$$w3(n+M/L)=\mathrm{Re}\left\{j(1/M)\sum_{k=0}^{M/(2L)-1}X(2k,mM/K)\exp(j\pi(2k+0.5)(2Ln+B)/(2M))\right\}$$
$$+\mathrm{Re}\left\{(-j)(1/M)\sum_{k=M/(2L)}^{M/L-1}X(2M/L-1-2k,mM/K)\exp(j\pi(2M/L-2k-0.5)(2Ln+B)/(2M))\right\}$$

$$w3(n+M/L)=\mathrm{Re}\left\{j(1/M)\sum_{k=0}^{M/(2L)-1}X(2k,mM/K)\exp(j\pi(2k+0.5)(2Ln+B)/(2M))\right\}$$
$$+\mathrm{Re}\left\{(-j)(1/M)\sum_{k=0}^{M/(2L)-1}X(2k+1,mM/K)\exp(j\pi(2k+1+0.5)(2Ln+B)/(2M))\right\}$$

$$w3(n+M/L)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/(2L)-1}X(2k,mM/K)\exp(j\pi(2k+0.5)(2L(n+M/L)+B)/(2M))\right\}$$
$$+\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/(2L)-1}X(2k+1,mM/K)\exp(j\pi(2k+1+0.5)(2L(n+M/L)+B)/(2M))\right\}$$

$$w3(n+M/L)=\mathrm{Re}\left\{(1/M)\sum_{k=0}^{M/L-1}X(k,mM/K)\exp(j\pi(k+0.5)(2L(n+M/L)+B)/(2M))\right\}$$

【0159】

（数５０）の最後の式は（数９）でｎの値の範囲をＭ／Ｌ≦１≦２Ｍ／Ｌ－１とし、ｎをｎ＋Ｍ／Ｌ、ｗをｗ３で置き換えたときの式と一致する。したがって、実施の形態５は、図１４の従来例と同一の結果を算出することができる。

【0160】

図７は、実施の形態５の複素指数変調フィルタバンクの信号合成方法を実行する装置のブロック図である。図７で、７０１は第１の中間信号生成部、７０２はＦＦＴ部（第２の中間信号生成部）、７０３は第３の中間信号生成部、７０４は出力信号生成部である。第１の中間信号生成部では、複素帯域分割入力信号から（数４０）と（数４１）を実行して第１の中間信号を出力する。ＦＦＴ部７０２では、第１の中間信号生成部７０１の出力した第１の中間信号に対してＭ／Ｌ点の高速フーリエ変換で（数４３）を実行して第２の中間信号を出力する。第３の中間信号生成部７０３では、ＦＦＴ部７０２の出力した第２の中間信号から（数４４）を実行し、第３の中間信号を出力する。出力信号生成部７０４では、第３の中間信号生成部７０３の出力した第３の中間信号とプロトタイプフィルタのフィルタ係数から（数４７）によって出力信号を算出し、出力する。

【0161】

以上のように実施の形態５の複素指数変調フィルタバンクの信号合成方法では、複素指数関数の周期性を利用して高速フーリエ変換を効率的に適用することにより、複素指数変調信号合成フィルタバンクの実行に必要な演算量を削減することができる。これによってデジタルシグナルプロセッサやＬＳＩで複素指数変調信号合成フィルタバンクを実装するときの動作クロック周波数を低減することができ、低消費電力で複素指数変調信号合成フ

ィルタバンクを実現することができる。また、コンピュータを使ってソフトウェアで複素指数変調信号合成フィルタバンクを実行する時の処理時間を短縮することができ、処理の高速化を実現することができる。

【0162】

なお、実施の形態５では、ステップ６０３で定数１／Ｍを掛けるようにしたが、これを別のステップ、例えばステップ６０２で掛けるように変形してもよい。

【0163】

（実施の形態６）

図８は、本発明の実施の形態６における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャートである。

【0164】

実施の形態６は、実施の形態５において信号合成の位相ＢをＭＬに設定したものである。

【0165】

図８のステップ８０２と８０４は、図６のステップ６０２と６０４とそれぞれ同一であり、説明を省略する。

【0166】

ステップ８０１では、複素帯域分割入力信号から第１の中間信号を算出する。すなわち、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とするとき、複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数５１）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数５２）、によって算出する。

【0167】

【数51】

$$W1(k) = X^{*}(2k, mM/K) \quad (0 \leqq k \leqq M/(2L) - 1)$$

【0168】

【数52】

$$W1(k) = X(2M/L - 1 - 2k, mM/K) \quad (M/(2L) \leqq k \leqq M/L - 1)$$

【0169】

実施の形態５の図６のステップ６０１を実行するには、Ｍ／Ｌ回の複素乗算が必要で、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要であったが、ステップ８０１では複素乗算は不要であり、その分演算量を削減できる。

【0170】

ステップ８０３で、０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、次に、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数５３）によって算出する。

【0171】

【数53】

$$\begin{aligned} w3(n) &= \mathrm{Re}\{w2(n)(1/M)\exp(-j\pi L(n+M)/(2M))\} \\ w3(n+M/L) &= \mathrm{Im}\{w2(n)(1/M)\exp(-j\pi L(n+M)/(2M))\} \end{aligned} \quad (0 \leqq n \leqq M/L - 1)$$

【0172】

ステップ８０３で、（１／Ｍ）ｅｘｐ（－ｊπ（ｎ＋Ｍ）／（２Ｍ））をテーブルに記憶して用いることにすると、ステップ８０３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。この演算回数は図６のステップ６０３と同一である。

【0173】

したがって、図８の実施の形態６を実行するのに必要な演算回数を、上記各ステップの

演算回数の総和で求めると、実数加算回数はＮ／Ｌ＋Ｍ／Ｌ＋（３Ｍ／Ｌ）ｌｏｇ$_2$（Ｍ／Ｌ）回、実数乗算回数はＮ／Ｌ＋４Ｍ／Ｌ＋（２Ｍ／Ｌ）ｌｏｇ$_2$（Ｍ／Ｌ）回である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，ダウンサンプリングの倍率１／Ｌを１とすれば、実施の形態６を実行するのに必要な演算量は、実数加算回数が１８５６回、実数乗算回数が１６６４回である。これは実施の形態５と比較して、実数加算回数が１２８回、実数乗算回数が２５６回少ない。

【0174】

以上のように実施の形態６の複素指数変調フィルタバンクの信号合成方法では、信号合成の位相を、複素帯域分割入力信号から第１の中間信号を算出するステップで乗算が発生しないように設定することによって、実施の形態５と比較して演算量をさらに削減することができる。

【0175】

（実施の形態７）

図９は、本発明の実施の形態７における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャートである。実施の形態５では高速フーリエ変換を効率的に適用することにより、複素指数変調信号合成フィルタバンクの実現に必要な演算量を削減したが、実施の形態７では、逆高速フーリエ変換を効率的に適用することにより複素指数変調信号合成フィルタバンクの実行に必要な演算量を削減する。図９においてステップ９０４は図６のステップ６０４と同一であり、説明を省略する。

【0176】

ステップ９０１では複素帯域分割信号から第１の中間信号を算出する。すなわち、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とするとき、複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数５４）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数５５）、によって算出する。ここで、Ｂは信号合成の位相で信号分析の位相Ａと（数４２）の関係式を満足する。

【0177】

【数54】

$$W1(k) = X(2k, mM/K)\exp(j\pi kB/M) \quad (0 \leqq k \leqq M/(2L)-1)$$

【0178】

【数55】

$$W1(k) = X^{*}(2M/L-1-2k, mM/K)\exp(j\pi B(k/M-1/L)) \quad (M/(2L) \leqq k \leqq M/L-1)$$

【0179】

ｅｘｐ（ｊπｋＢ／Ｍ）とｅｘｐ（ｊπＢ（ｋ／Ｍ）－１／Ｌ）をテーブルに記憶して用いることにすると、ステップ９０１を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。

【0180】

ステップ９０２で、第１の中間信号から逆高速フーリエ変換で（数５６）によって第２の中間信号ｗ２（ｎ）（ただし、０≦ｎ≦Ｍ／Ｌ－１）を算出するステップと、

【0181】

【数56】

$$w2(n) = \sum_{k=0}^{M/L-1} W1(k)\exp(j2\pi knL/M) \quad (0 \leqq n \leqq M/L-1)$$

【0182】

ステップ９０２でＭ／Ｌ点のＩＦＦＴを実行するのに必要な実数加算回数は（３Ｍ／Ｌ）ｌｏｇ$_2$（Ｍ／Ｌ）回で、実数乗算回数は（２Ｍ／Ｌ）ｌｏｇ$_2$（Ｍ／Ｌ）回である。

【0183】

ステップ９０３で、０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、次に、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数５７）によって算出する。

【0184】

【数57】

$$w3(n) = \mathrm{Re}\{w2(n)(1/M)\exp(j\pi(2Ln+B)/(4M))\}$$
$$w3(n+M/L) = -\mathrm{Im}\{w2(n)(1/M)\exp(j\pi(2Ln+B)/(4M))\} \qquad (0 \leq n \leq M/L-1)$$

【0185】

ステップ９０３で、（１／Ｍ）ｅｘｐ（ｊπ（２Ｌｎ＋Ｂ）／（４Ｍ））をテーブルに記憶して用いることにすると、ステップ９０３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。

【0186】

したがって、図９の実施の形態７を実行するのに必要な演算回数を、上記各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｌ＋３Ｍ／Ｌ＋（３Ｍ／Ｌ）$\log_2$（Ｍ／Ｌ）回、実数乗算回数はＮ／Ｌ＋８Ｍ／Ｌ＋（２Ｍ／Ｌ）$\log_2$（Ｍ／Ｌ）回である。これは、実施の形態５と同じ演算量である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４、ダウンサンプリングの倍率１／Ｌを１とすれば、実施の形態７を実行するのに必要な演算量は、実数加算回数が１９８４回、実数乗算回数が１９２０回である。これを従来例（実数加算回数１６８３２回、実数乗算回数１７０２４回）と比較すると、実施の形態７では、実数加算回数、実数乗算回数をともに従来例の１／８以下に削減することができる。

【0187】

実施の形態７のステップ９０１、９０２、９０３が、従来例の図１４のステップ１４０１と同一の結果を出力することができることは、実施の形態５で説明したのと同様に、（数５４）と（数５５）を（数５６）に代入し、代入結果をさらに（数５７）に代入することにより説明できる。

【0188】

以上のように実施の形態７の複素指数変調フィルタバンクの信号合成方法では、複素指数関数の周期性を利用して逆高速フーリエ変換を効率的に適用することにより、複素指数変調信号合成フィルタバンクの実行に必要な演算量を削減することができる。これによってデジタルシグナルプロセッサやＬＳＩで複素指数変調信号合成フィルタバンクを実装するときの動作クロック周波数を低減することができ、低消費電力で複素指数変調信号合成フィルタバンクを実現することができる。また、コンピュータを使ってソフトウェアで複素指数変調信号合成フィルタバンクを実行する時の処理時間を短縮することができ、処理の高速化を実現することができる。

【0189】

（実施の形態８）

図１０は、本発明の実施の形態８における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャートである。

【0190】

実施の形態８は、実施の形態７において信号合成の位相ＢをＭＬに設定したものである。

【0191】

図１０のステップ１００２と、１００４は、図９のステップ９０２と９０４とそれぞれ同一であり、説明を省略する。

【0192】

ステップ１００１では、複素帯域分割入力信号から第１の中間信号を算出する。すなわち、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ

／Ｋ）とするとき、複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数５８）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数５９）、によって算出する。

【0193】

【数58】

$$W1(k) = X(2k, mM/K) \quad (0 \leqq k \leqq M/(2L) - 1)$$

【0194】

【数59】

$$W1(k) = X^{*}(2M/L - 1 - 2k, mM/K) \quad (M/(2L) \leqq k \leqq M/L - 1)$$

【0195】

図９のステップ９０１を実行するには、Ｍ／Ｌ回の複素乗算が必要で、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要であったが、ステップ１００１では複素乗算は不要であり、その分演算量を削減できる。

【0196】

ステップ１００３で、０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、次に、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数６０）によって算出する。

【0197】

【数60】

$$\begin{aligned} w3(n) &= \mathrm{Re}\{w2(n)(1/M)\exp(j\pi L(n+M)/(2M))\} \\ w3(n+M/L) &= -\mathrm{Im}\{w2(n)(1/M)\exp(j\pi L(n+M)/(2M))\} \end{aligned} \quad (0 \leqq n \leqq M/L - 1)$$

【0198】

ステップ１００３で、（１／Ｍ）ｅｘｐ（ｊπＬ（１＋Ｍ）／（２Ｍ））をテーブルに記憶して用いることにすると、ステップ１００３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。この演算回数は図９のステップ９０３と同一である。

【0199】

したがって、図１０の実施の形態８を実行するのに必要な演算回数を、上記各ステップの演算回数の総和で求めると、実数加算回数はＮ／Ｌ＋Ｍ／Ｌ＋（３Ｍ／Ｌ）$\log_2$（Ｍ／Ｌ）回、実数乗算回数はＮ／Ｌ＋４Ｍ／Ｌ＋（２Ｍ／Ｌ）$\log_2$（Ｍ／Ｌ）回である。プロトタイプフィルタの次数Ｎを６４０、帯域分割数Ｍを６４，ダウンサンプリングの倍率１／Ｌを１とすれば、実施の形態８を実行するのに必要な演算量は、実数加算回数が１８５６回、実数乗算回数が１６６４回である。これは実施の形態７と比較して、実数加算回数が１２８回、実数乗算回数が２５６回少ない。

【0200】

以上のように実施の形態８の複素指数変調フィルタバンクの信号合成方法では、信号合成の位相を、複素帯域分割入力信号から第１の中間信号を算出するステップで乗算が発生しないように設定することによって、実施の形態７と比較して演算量をさらに削減することができる。

【0201】

（実施の形態９）

図１１は、本発明の実施の形態９における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャートである。

【0202】

図１１のステップ１１０２と１１０４は、図６のステップ６０２と６０４とそれぞれ同一であり、説明を省略する。

【0203】

ステップ１１０１では、複素帯域分割入力信号から第１の中間信号を算出する。すなわち、第ｋ帯域のサンプリング時刻ｍＭ／Ｋにおける複素帯域分割入力信号をＸ（ｋ，ｍＭ／Ｋ）とするとき、複素帯域分割入力信号から第１の中間信号Ｗ１（ｋ）を、０≦ｋ≦Ｍ／（２Ｌ）－１に対しては（数６１）、Ｍ／（２Ｌ）≦ｋ≦Ｍ／Ｌ－１に対しては（数６２）、によって算出する。

【0204】

【数61】

$$W1(k) = X^{*}(2k, mM/K)\exp(-j\pi kB(4k+1)/(4M)) \quad (0 \leqq k \leqq M/(2L)-1)$$

【0205】

【数62】

$$W1(k) = X(2M/L-1-2k, mM/K)\exp(j\pi B(1/L-(4k+1)/(4M)))$$

$$(M/(2L) \leqq k \leqq M/L-1)$$

【0206】

（数６１）と（数４０）を、（数６２）と（数４１）を比較することにより、ステップ１１０１はステップ６０１にｅｘｐ（－ｊπＢ／（４Ｍ））を掛けたものであることがわかる。

【0207】

ステップ１１０３で、０≦ｎ≦２（Ｎ－Ｍ）／Ｌ－１に対する第３の中間信号ｗ３（ｎ）をｗ３（ｎ＋２Ｍ／Ｌ）にシフトし、次に、０≦ｎ≦Ｍ／Ｌに対する第３の中間信号ｗ３（ｎ）とｗ３（ｎ＋Ｍ／Ｌ）を第２の中間信号から（数６３）によって算出する。

【0208】

【数63】

$$w3(n) = \mathrm{Re}\{w2(n)(1/M)\exp(-j\pi Ln/(2M))\}$$
$$w3(n+M/L) = \mathrm{Im}\{w2(n)(1/M)\exp(-j\pi Ln/(2M))\} \quad (0 \leqq n \leqq M/L-1)$$

【0209】

（数６３）と（数４４）を比較することにより、ステップ１１０３はステップ６０３にｅｘｐ（ｊπＢ／（４Ｍ））を掛けたものであることがわかる。

【0210】

すなわち、図１１の実施の形態９は、図６のステップ６０１で定数ｅｘｐ（－ｊπＢ／（４Ｍ））を掛け、ステップ６０３で前記定数の逆数ｅｘｐ（ｊπＢ／（４Ｍ））を掛けたものに相当する。前のステップで定数を掛け、後のステップで前記定数の逆数を掛ければ、元と同じ結果を得ることができる。したがって、図１１の実施の形態９は、図６の実施の形態５と同一の結果を得ることができる。

【0211】

なお、上記各実施の形態で、同様に、前のステップで定数を掛け、後のステップで前記定数の逆数を掛けるようにステップを変形してもよい。

【0212】

なお、上記各実施の形態における信号分析と合成方法は、コンピュータまたはデジタルシグナルプロセッサに実行させるためのプログラムとして実現することができ、これをコンピュータで読み取り可能な記録媒体に記録してもよい。

【産業上の利用可能性】

【0213】

以上のように、本発明にかかる複素指数変調フィルタバンクの信号分析と信号合成方法は、オーディオ信号や画像信号の高能率符号化やデジタルイコラーザー等の信号分析と信号合成方法として有用であり、特に演算量を削減し、低消費電力のあるいは高速な複素指数変調フィルタバンクを実現するのに適している。

【図面の簡単な説明】

【0214】

【図1】本発明の実施の形態１における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャート

【図2】本発明の実施の形態１における複素指数変調フィルタバンクの信号分析方法を実行する装置の構成を示すブロック図

【図3】本発明の実施の形態２における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャート

【図4】本発明の実施の形態３における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャート

【図5】本発明の実施の形態４における複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャート

【図6】本発明の実施の形態５における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャート

【図7】本発明の実施の形態５における複素指数変調フィルタバンクの信号合成方法を実行する装置の構成を示すブロック図

【図8】本発明の実施の形態６における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャート

【図9】本発明の実施の形態７における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャート

【図10】本発明の実施の形態８における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャート

【図11】本発明の実施の形態９における複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャート

【図12】信号分析フィルタバンクと信号合成フィルタバンクの構成を示すブロック図

【図13】従来の複素指数変調フィルタバンクの信号分析方法の処理ステップを示すフローチャート

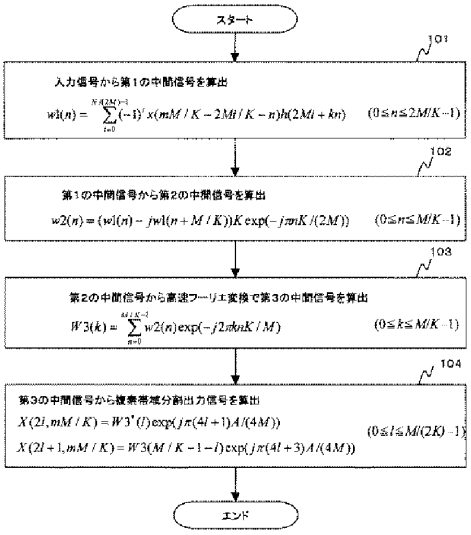
【図14】従来の複素指数変調フィルタバンクの信号合成方法の処理ステップを示すフローチャート

【符号の説明】

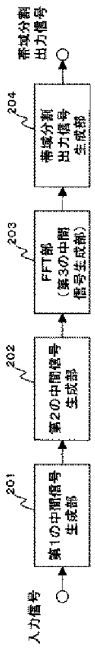
【0215】

２０１　第１の中間信号生成部
２０２　第２の中間信号生成部
２０３　ＦＦＴ部（第３の中間信号生成部）
２０４　帯域分割出力信号生成部
７０１　第１の中間信号生成部
７０２　ＦＦＴ部（第２の中間信号生成部）
７０３　第３の中間信号生成部
７０４　出力信号生成部

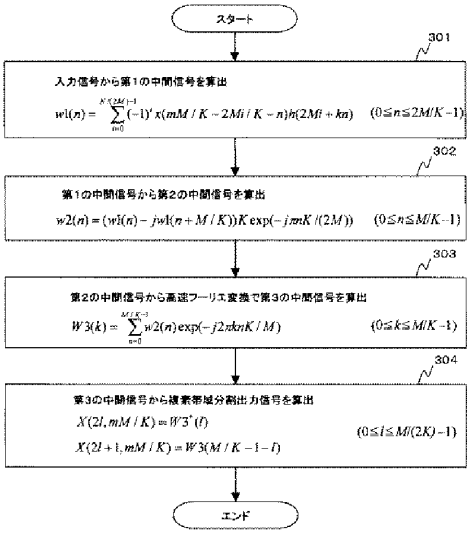
【図1】


スタート
101
入力信号から第1の中間信号を算出
$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+kn)$　$(0\le n\le 2M/K-1)$
102
第1の中間信号から第2の中間信号を算出
$w2(n)=(w1(n)-jw1(n+M/K))K\exp(-j\pi nK/(2M))$　$(0\le n\le M/K-1)$
103
第2の中間信号から高速フーリエ変換で第3の中間信号を算出
$W3(k)=\sum_{n=0}^{M/K-1}w2(n)\exp(-j2\pi knK/M)$　$(0\le k\le M/K-1)$
104
第3の中間信号から複素帯域分割出力信号を算出
$X(2l,mM/K)=W3^*(l)\exp(j\pi(4l+1)A/(4M))$
$X(2l+1,mM/K)=W3(M/K-1-l)\exp(j\pi(4l+3)A/(4M))$　$(0\le l\le M/(2K)-1)$
エンド

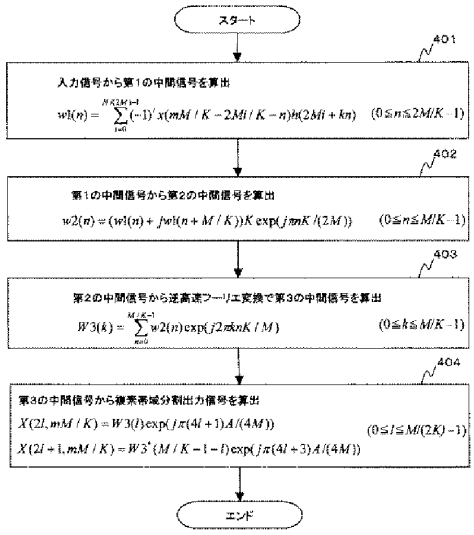
【図2】


帯域分割出力信号
204
帯域分割出力信号生成部
203
FFT部（第3の中間信号生成部）
202
第2の中間信号生成部
201
第1の中間信号生成部
入力信号

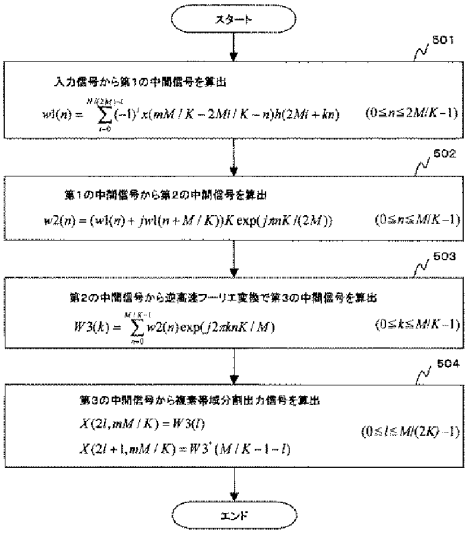
【図3】


スタート
301
入力信号から第1の中間信号を算出
$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+kn)$　$(0\le n\le 2M/K-1)$
302
第1の中間信号から第2の中間信号を算出
$w2(n)=(w1(n)-jw1(n+M/K))K\exp(-j\pi nK/(2M))$　$(0\le n\le M/K-1)$
303
第2の中間信号から高速フーリエ変換で第3の中間信号を算出
$W3(k)=\sum_{n=0}^{M/K-1}w2(n)\exp(-j2\pi knK/M)$　$(0\le k\le M/K-1)$
304
第3の中間信号から複素帯域分割出力信号を算出
$X(2l,mM/K)=W3^*(l)$
$X(2l+1,mM/K)=W3(M/K-1-l)$　$(0\le l\le M/(2K)-1)$
エンド

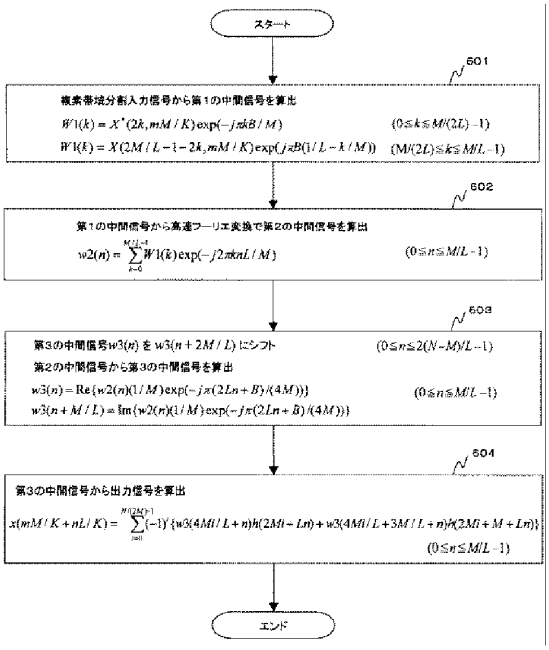
【図4】


スタート
401
入力信号から第1の中間信号を算出
$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+kn)$　$(0\le n\le 2M/K-1)$
402
第1の中間信号から第2の中間信号を算出
$w2(n)=(w1(n)+jw1(n+M/K))K\exp(j\pi nK/(2M))$　$(0\le n\le M/K-1)$
403
第2の中間信号から逆高速フーリエ変換で第3の中間信号を算出
$W3(k)=\sum_{n=0}^{M/K-1}w2(n)\exp(j2\pi knK/M)$　$(0\le k\le M/K-1)$
404
第3の中間信号から複素帯域分割出力信号を算出
$X(2l,mM/K)=W3(l)\exp(j\pi(4l+1)A/(4M))$
$X(2l+1,mM/K)=W3^*(M/K-1-l)\exp(j\pi(4l+3)A/(4M))$　$(0\le l\le M/(2K)-1)$
エンド

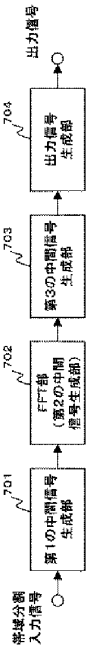
【図5】


スタート
501
入力信号から第1の中間信号を算出
$w1(n)=\sum_{i=0}^{N/(2M)-1}(-1)^i x(mM/K-2Mi/K-n)h(2Mi+kn)$　$(0\le n\le 2M/K-1)$
502
第1の中間信号から第2の中間信号を算出
$w2(n)=(w1(n)+jw1(n+M/K))K\exp(j\pi nK/(2M))$　$(0\le n\le M/K-1)$
503
第2の中間信号から逆高速フーリエ変換で第3の中間信号を算出
$W3(k)=\sum_{n=0}^{M/K-1}w2(n)\exp(j2\pi knK/M)$　$(0\le k\le M/K-1)$
504
第3の中間信号から複素帯域分割出力信号を算出
$X(2l,mM/K)=W3(l)$
$X(2l+1,mM/K)=W3^*(M/K-1-l)$　$(0\le l\le M/(2K)-1)$
エンド

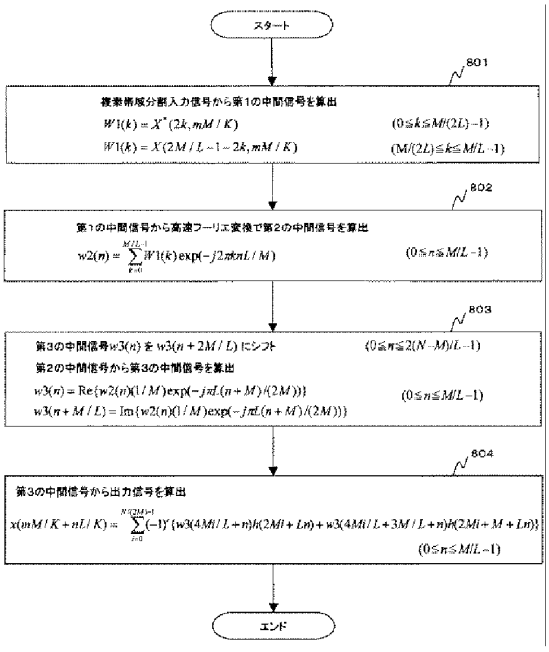
【図6】


スタート
501
複素帯域分割入力信号から第1の中間信号を算出
$W1(k)=X^*(2k,mM/K)\exp(-j\pi kB/M)$　$(0\le k\le M/(2L)-1)$
$W1(k)=X(2M/L-1-2k,mM/K)\exp(j\pi B(1/L-k/M))$　$(M/(2L)\le k\le M/L-1)$
502
第1の中間信号から高速フーリエ変換で第2の中間信号を算出
$w2(n)=\sum_{k=0}^{M/L-1}W1(k)\exp(-j2\pi knL/M)$　$(0\le n\le M/L-1)$
503
第3の中間信号$w3(n)$を$w3(n+2M/L)$にシフト　$(0\le n\le 2(N-M)/L-1)$
第2の中間信号から第3の中間信号を算出
$w3(n)=\mathrm{Re}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$　$(0\le n\le M/L-1)$
$w3(n+M/L)=\mathrm{Im}\{w2(n)(1/M)\exp(-j\pi(2Ln+B)/(4M))\}$
504
第3の中間信号から出力信号を算出
$x(mM/K+nL/K)=\sum_{i=0}^{N/(2M)-1}(-1)^i\{w3(4Mi/L+n)h(2Mi-Ln)+w3(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$
$(0\le n\le M/L-1)$
エンド

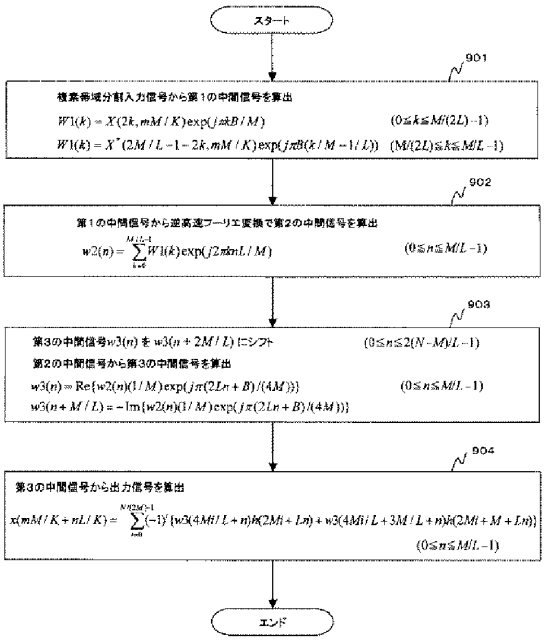
【図7】


出力信号
704
出力信号生成部
703
第3の中間信号生成部
702
FFT部（第2の中間信号生成部）
701
第1の中間信号生成部
帯域分割入力信号

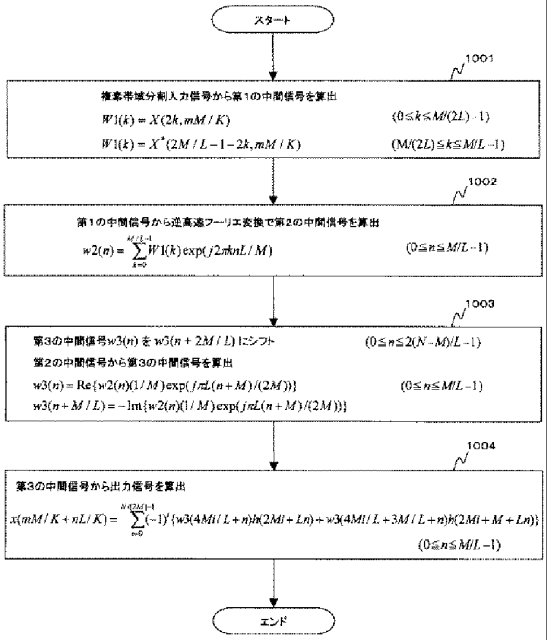
【図8】


スタート
801
複素帯域分割入力信号から第1の中間信号を算出
$W1(k)=X^*(2k,mM/K)$　$(0\le k\le M/(2L)-1)$
$W1(k)=X(2M/L-1-2k,mM/K)$　$(M/(2L)\le k\le M/L-1)$
802
第1の中間信号から高速フーリエ変換で第2の中間信号を算出
$w2(n)=\sum_{k=0}^{M/L-1}W1(k)\exp(-j2\pi knL/M)$　$(0\le n\le M/L-1)$
803
第3の中間信号$w3(n)$を$w3(n+2M/L)$にシフト　$(0\le n\le 2(N-M)/L-1)$
第2の中間信号から第3の中間信号を算出
$w3(n)=\mathrm{Re}\{w2(n)(1/M)\exp(-j\pi L(n+M)/(2M))\}$
$w3(n+M/L)=\mathrm{Im}\{w2(n)(1/M)\exp(-j\pi L(n+M)/(2M))\}$　$(0\le n\le M/L-1)$
804
第3の中間信号から出力信号を算出
$x(mM/K+nL/K)=\sum_{i=0}^{N/(2M)-1}(-1)^i\{w3(4Mi/L+n)h(2Mi+Ln)+w3(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$
$(0\le n\le M/L-1)$
エンド

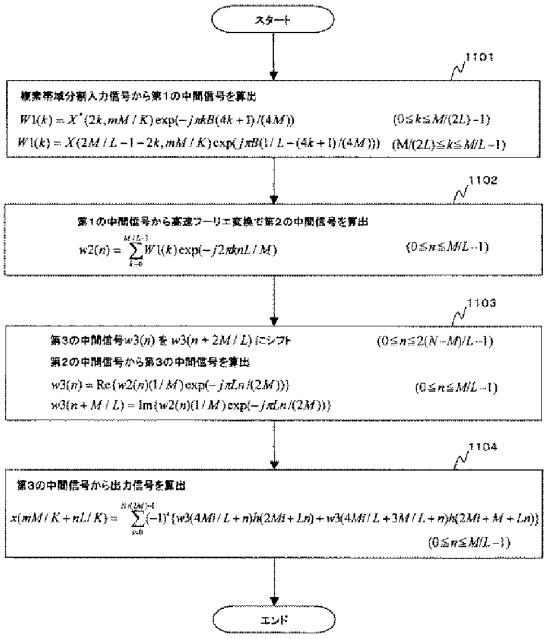
【図9】



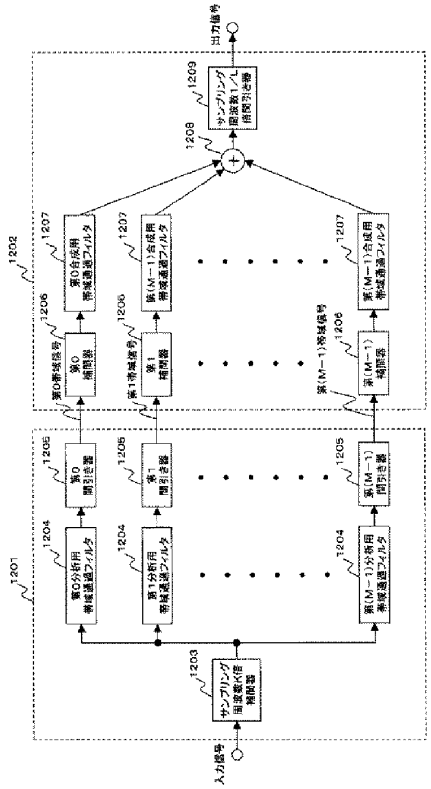
【図10】



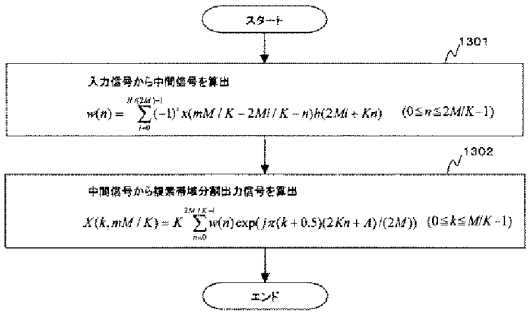
【図11】



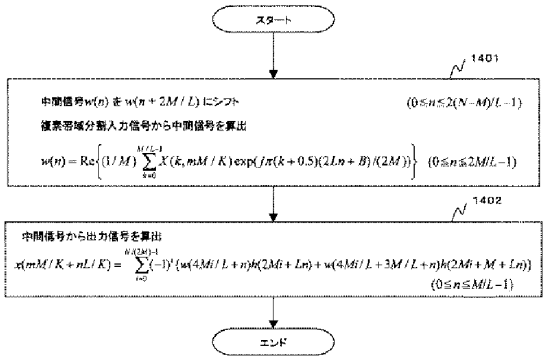
【図12】



【図13】



【図14】



【手続補正書】
【提出日】平成16年10月29日(2004.10.29)
【手続補正1】
【補正対象書類名】明細書
【補正対象項目名】００５０
【補正方法】変更
【補正の内容】
【0050】
ステップ１０１では、入力信号から第１の中間信号を算出する。すなわち、サンプリング時刻ｎにおける入力信号をｘ（ｎ）とするとき、入力信号から第１の中間信号ｗ１（ｎ）（ただし、０≦ｎ≦２Ｍ／Ｋ－１）を（数１１）によって算出する。
【手続補正2】
【補正対象書類名】明細書
【補正対象項目名】００５６
【補正方法】変更
【補正の内容】
【0056】
ステップ１０２では、第１の中間信号から（数１４）（ただし、ｊは虚数単位）によって第２の中間信号ｗ２（ｎ）（ただし、０≦ｎ≦Ｍ／Ｋ－１）を算出する。
【手続補正3】
【補正対象書類名】明細書
【補正対象項目名】００５９
【補正方法】変更
【補正の内容】
【0059】

ステップ１０３では第２の中間信号から高速フーリエ変換（ＦＦＴ：Fast Fourier Transform）で（数１５）によって第３の中間信号Ｗ３（ｋ）（ただし、０≦ｋ≦Ｍ／Ｋ－１）を算出する。

【手続補正4】
【補正対象書類名】明細書
【補正対象項目名】００９７
【補正方法】変更
【補正の内容】
【0097】
ステップ４０２では、第１の中間信号から（数２８）によって第２の中間信号ｗ２（ｎ）（ただし、０≦ｎ≦Ｍ／Ｋ－１）を算出する。

【手続補正5】
【補正対象書類名】明細書
【補正対象項目名】０１００
【補正方法】変更
【補正の内容】
【0100】
ステップ４０３では第２の中間信号から逆高速フーリエ変換で（数２９）によって第３の中間信号Ｗ３（ｋ）（ただし、０≦ｋ≦Ｍ／Ｋ－１）を算出する。

【手続補正6】
【補正対象書類名】明細書
【補正対象項目名】０１４３
【補正方法】変更
【補正の内容】
【0143】
ステップ６０３で、（１／Ｍ）ｅｘｐ（－ｊπ（２Ｌｎ＋Ｂ）／（４Ｍ））をテーブルに記憶して用いることにすると、ステップ６０３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。

【手続補正7】
【補正対象書類名】明細書
【補正対象項目名】０１７２
【補正方法】変更
【補正の内容】
【0172】
ステップ８０３で、（１／Ｍ）ｅｘｐ（－ｊπＬ（ｎ＋Ｍ）／（２Ｍ））をテーブルに記憶して用いることにすると、ステップ８０３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。この演算回数は図６のステップ６０３と同一である。

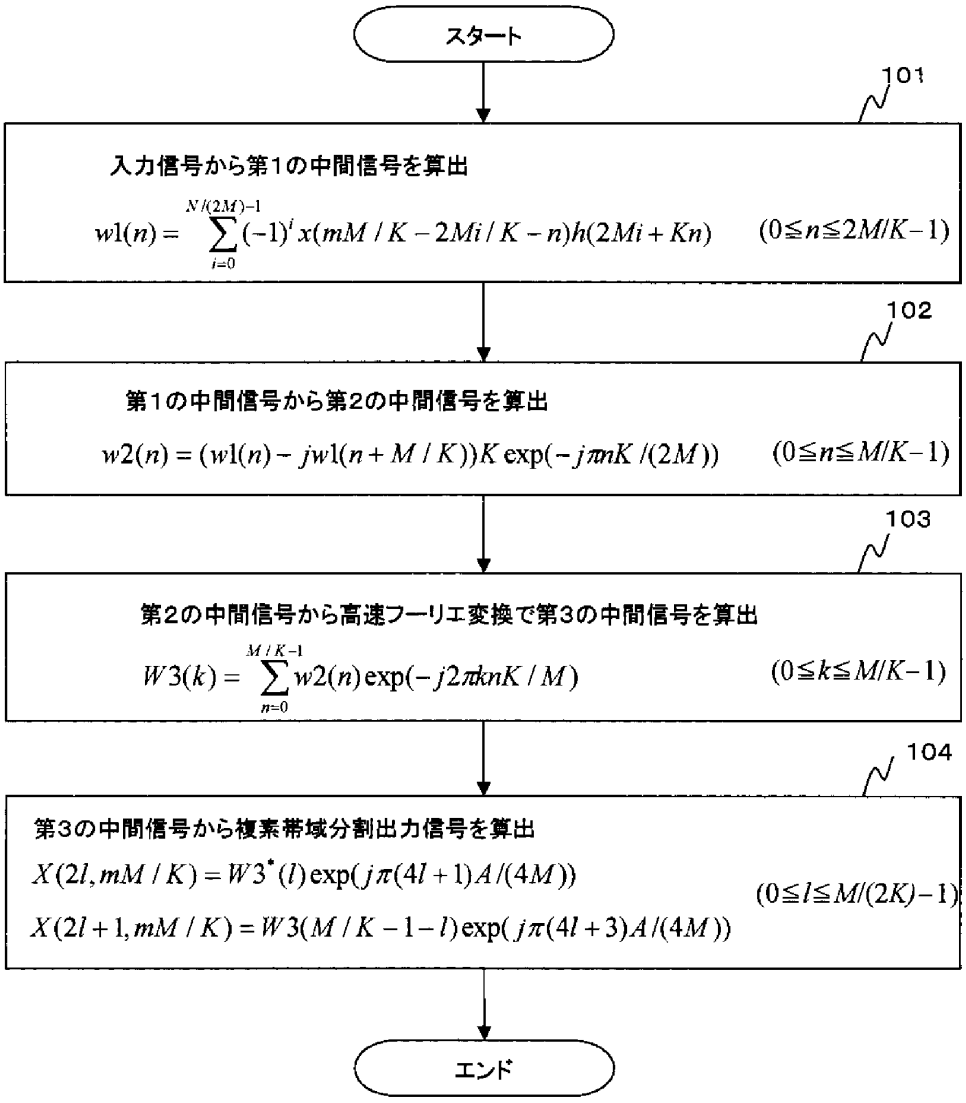
【手続補正8】
【補正対象書類名】明細書
【補正対象項目名】０１９８
【補正方法】変更
【補正の内容】
【0198】
ステップ１００３で、（１／Ｍ）ｅｘｐ（ｊπＬ（ｎ＋Ｍ）／（２Ｍ））をテーブルに記憶して用いることにすると、ステップ１００３を実行するには、Ｍ／Ｌ回の複素乗算が必要で、これを実数演算に直すと、２Ｍ／Ｌ回の実数加算と４Ｍ／Ｌ回の実数乗算が必要である。この演算回数は図９のステップ９０３と同一である。

【手続補正9】

【補正対象書類名】明細書
【補正対象項目名】０２０４
【補正方法】変更
【補正の内容】
【0204】
【数61】

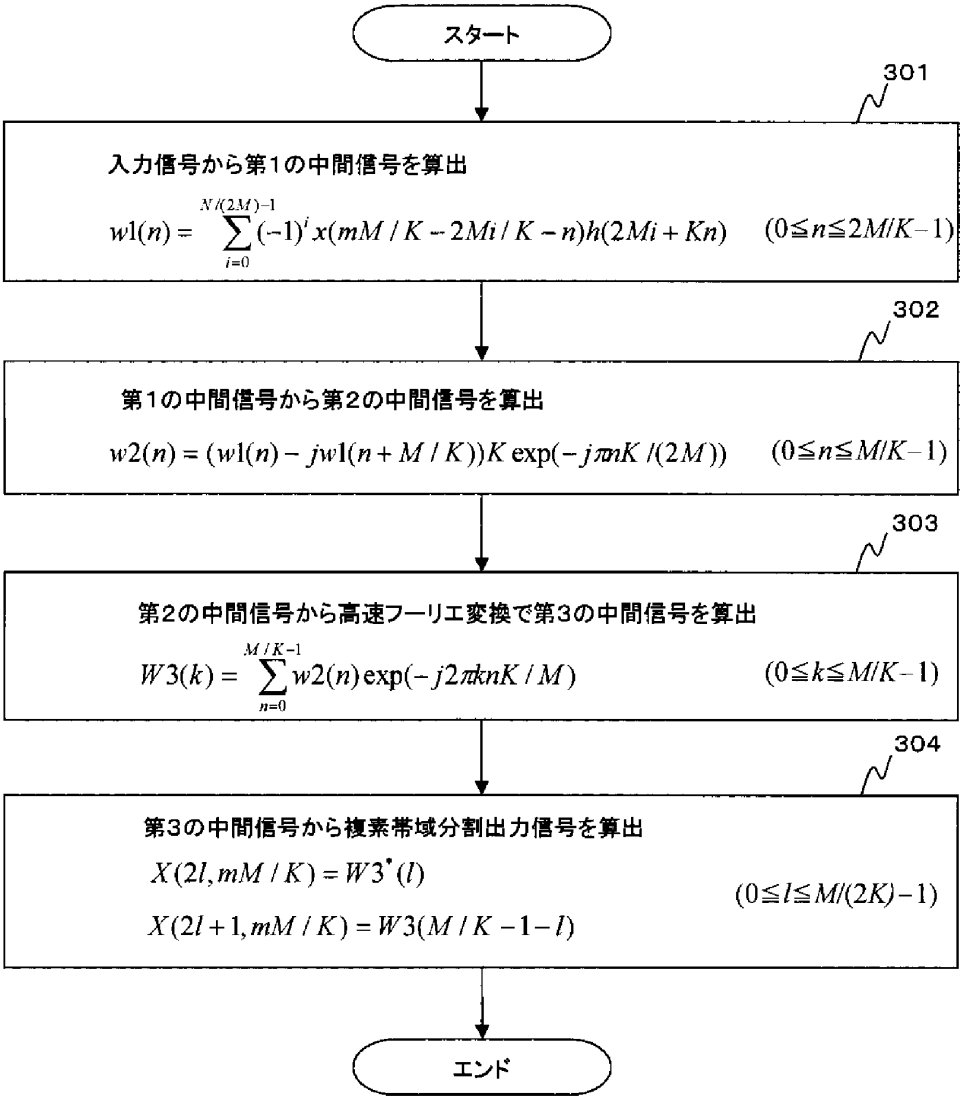
$$W1(k) = X^*(2k, mM/K)\exp(-j\pi B(4k+1)/(4M)) \quad (0 \leqq k \leqq M/(2L)-1)$$

【手続補正10】
【補正対象書類名】図面
【補正対象項目名】図１
【補正方法】変更
【補正の内容】
【図1】



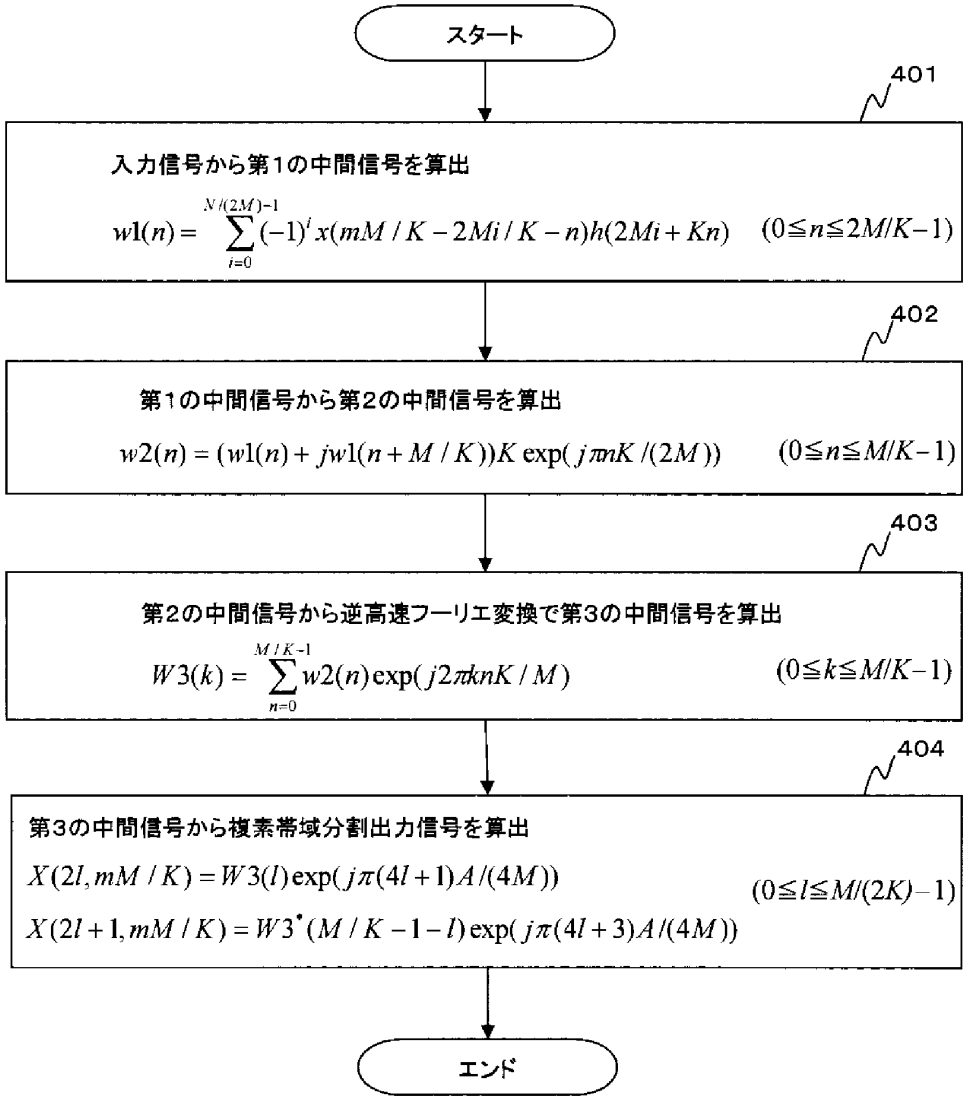
【手続補正11】
【補正対象書類名】図面

【補正対象項目名】図３
【補正方法】変更
【補正の内容】
【図3】



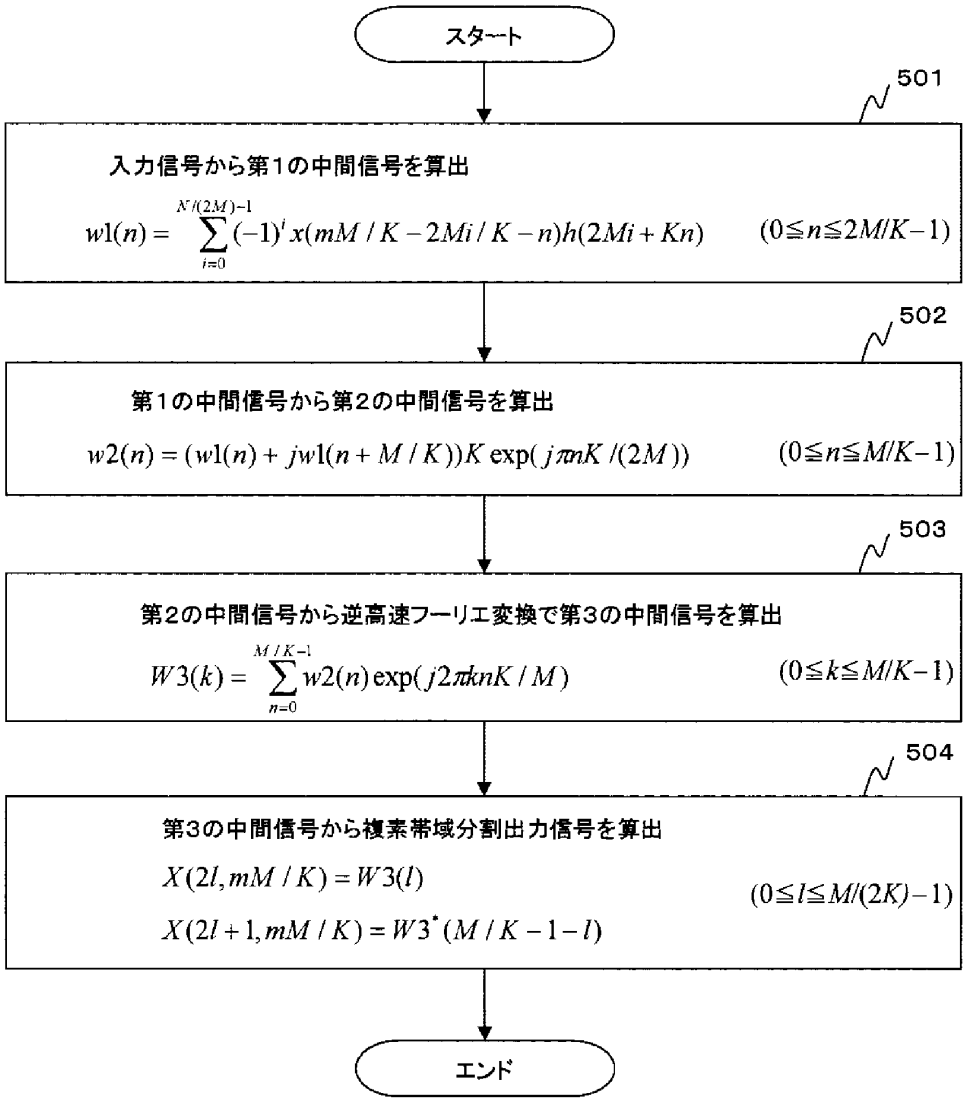
【手続補正12】
【補正対象書類名】図面
【補正対象項目名】図４
【補正方法】変更
【補正の内容】

【図4】



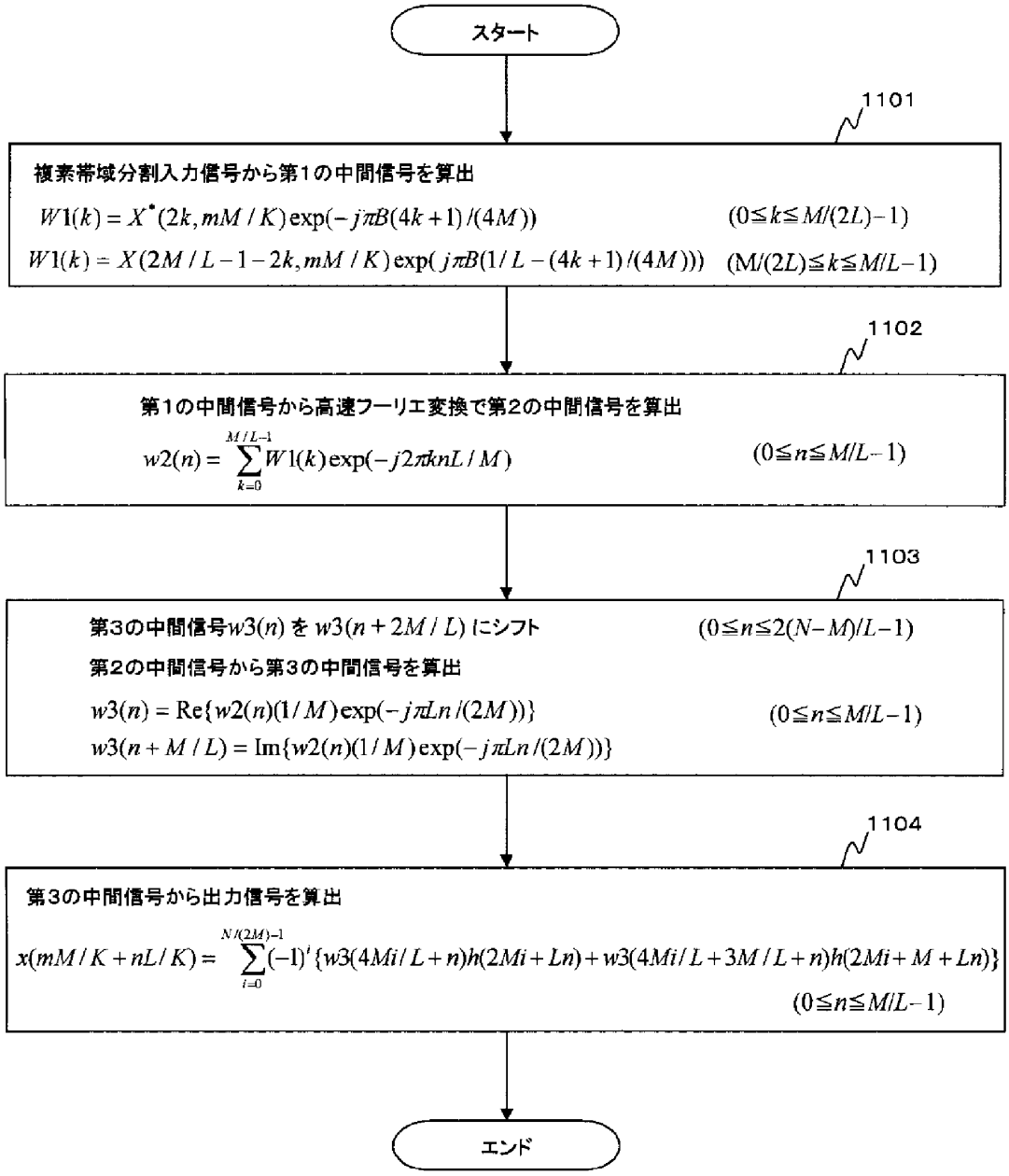
【手続補正13】
【補正対象書類名】図面
【補正対象項目名】図5
【補正方法】変更
【補正の内容】

【図5】



【手続補正14】
【補正対象書類名】図面
【補正対象項目名】図１１
【補正方法】変更
【補正の内容】

【図11】


スタート
1101
複素帯域分割入力信号から第1の中間信号を算出
$W1(k) = X^*(2k, mM/K)\exp(-j\pi B(4k+1)/(4M))$ $(0 \leqq k \leqq M/(2L)-1)$
$W1(k) = X(2M/L-1-2k, mM/K)\exp(j\pi B(1/L-(4k+1)/(4M)))$ $(M/(2L) \leqq k \leqq M/L-1)$
1102
第1の中間信号から高速フーリエ変換で第2の中間信号を算出
$w2(n) = \sum_{k=0}^{M/L-1} W1(k)\exp(-j2\pi knL/M)$ $(0 \leqq n \leqq M/L-1)$
1103
第3の中間信号$w3(n)$を$w3(n+2M/L)$にシフト $(0 \leqq n \leqq 2(N-M)/L-1)$
第2の中間信号から第3の中間信号を算出
$w3(n) = \mathrm{Re}\{w2(n)(1/M)\exp(-j\pi Ln/(2M))\}$ $(0 \leqq n \leqq M/L-1)$
$w3(n+M/L) = \mathrm{Im}\{w2(n)(1/M)\exp(-j\pi Ln/(2M))\}$
1104
第3の中間信号から出力信号を算出
$x(mM/K+nL/K) = \sum_{i=0}^{N/(2M)-1}(-1)^i\{w3(4Mi/L+n)h(2Mi+Ln)+w3(4Mi/L+3M/L+n)h(2Mi+M+Ln)\}$
$(0 \leqq n \leqq M/L-1)$
エンド

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 October 2002 (10.10.2002)

PCT

(10) International Publication Number
WO 02/080362 A1

(51) International Patent Classification[7]: H03H 17/00

(21) International Application Number: PCT/SE02/00626

(22) International Filing Date: 28 March 2002 (28.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0101175-8 2 April 2001 (02.04.2001) SE

(71) Applicant *(for all designated States except US)*: **CODING TECHNOLOGIES SWEDEN AB** [SE/SE]; Döbelnsgatan 64, S-113 52 Stockholm (SE).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **EKSTRAND, Per** [SE/SE]; Södermannagatan 45, S-116 40 Stockholm (SE).

(74) Agents: **ÖRTENBLAD, Bertil** et al.; Noréns Patentbyrå AB, Box 10198, S-100 55 Stockholm (SE).

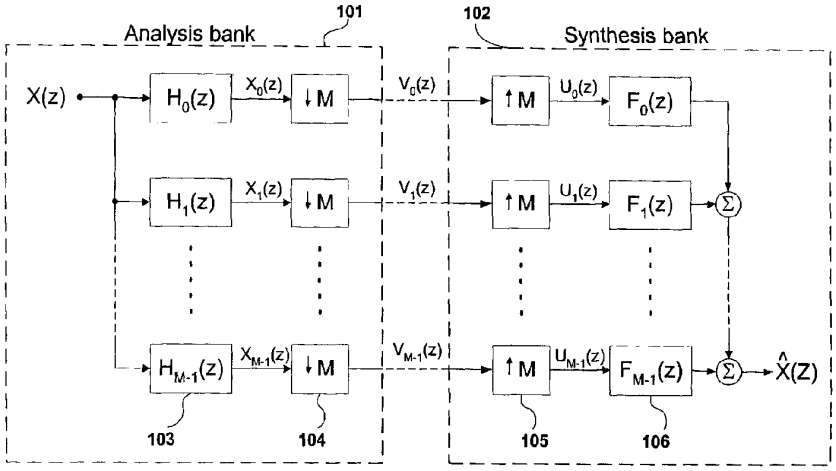
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** ALIASING REDUCTION USING COMPLEX-EXPONENTIAL MODULATED FILTERBANKS



(57) **Abstract:** The present invention proposes a new method and apparatus for the improvement of digital filterbanks, by a complex extension of cosine modulated digital filterbanks. The invention employs complex-exponential modulation of a low-pass prototype filter and a new method for optimizing the characteristics of this filter. The invention substantially reduces artifacts due to aliasing emerging from independent modifications of subband signals, for example when using a filterbank as an spectral equalizer. The invention is preferably implemented in software, running on a standard PC or a digital signal processor (DSP), but can also be hardcoded on a custom chip. The invention offers essential improvements for various types of digital equalizers, adaptive filters, multiband companders and spectral envelope adjusting filterbanks used in high frequency reconstruction (HFR) systems.

WO 02/080362 A1

ALIASING REDUCTION USING COMPLEX-EXPONENTIAL MODULATED FILTERBANKS

TECHNICAL FIELD

The present invention relates to the area of subsampled digital filterbanks and provides a new method and apparatus for substantial reduction of impairments emerging from modifications, for example quantization or attenuation, of the spectral coefficients or subband signals of digital filterbanks. The invention is applicable to digital equalizers ["An Efficient 20 Band Digital Audio Equalizer" A. J. S. Ferreira, J. M. N. Viera, AES preprint, 98th Convention 1995 February 25-28 Paris, N.Y., USA], adaptive filters [Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments, and Application to Acoustic Echo Cancellation" A. Gilloire, M. Vetterli, IEEE Transactions on Signal Processing, vol. 40, no. 8, August, 1992], multiband companders, and to audio coding systems using high frequency reconstruction (HFR), where a digital filterbank is used for the adaptive adjustment of the spectral envelope, such as the Spectral Band Replication (SBR) system [WO 98/57436].

BACKGROUND OF THE INVENTION

A digital filter bank is a collection of two or more parallel digital filters. The analysis filter bank splits the incoming signal into a number of separate signals named subband signals (or spectral coefficients). The filter bank is critically sampled (or maximally decimated) when the total number of subband samples per unit time is the same as that for the input signal. The synthesis filter bank combines these subband signals into an output signal. A popular type of critically sampled filterbanks is the cosine modulated filterbank. The filters in the cosine modulated system are obtained by cosine modulation of a low-pass filter, a so-called prototype filter. The cosine modulated banks offer very effective implementations and are often used in natural audio codecs ["Introduction to Perceptual Coding" K. Brandenburg, AES, Collected Papers on Digital Audio Bitrate Reduction, 1996]. However, any attempt to alter the subband samples or spectral coefficients, e.g. by applying an equalizing gain curve or quantizing the samples, renders severe aliasing artifacts in the output signal.

SUMMARY OF THE INVENTION

The present invention shows that impairments emerging from modifications of the subband signals can be significantly reduced by extending a cosine modulated filterbank with an imaginary sine modulated part, forming a complex-exponential modulated filterbank. The sine extension eliminates the main alias terms present in the cosine modulated filterbank. Moreover, the invention presents a method, referred to as alias term minimization (ATM), for optimization of the prototype filter. The complex-exponential modulation creates complex-valued subband signals that can be interpreted as the analytic signals of the signals obtained from the real part of the filterbank, i.e. the underlying cosine modulated filterbank. This feature provides an inherent measure of the instantaneous energy of the subband signals.

The main steps for operation of the complex-exponential modulated filterbank according to the present invention are:

1. The design of a symmetric low-pass prototype filter with cutoff frequency $\pi/2M$, optimized for a desired aliasing rejection and passband flatness;
2. The construction of an M-channel filterbank by complex-exponential modulation of the optimized prototype filter;
3. The filtering of a real-valued time domain signal through the analysis part of the filterbank;
4. The modification of the complex-valued subband signals according to a desired, possibly time-varying, equalizer setting;
5. The filtering of the modified complex-valued subband samples through the synthesis part of the filterbank; and
6. The computation of the real part of the complex-valued time domain output signal obtained from the synthesis part of the filterbank.

The most attractive applications of the invention are improvement of various types of digital equalizers, adaptive filters, multiband companders and adaptive envelope adjusting filterbanks used in HFR systems.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described by way of illustrative examples, not limiting the scope nor the spirit of the invention, with reference to the accompanying drawings, in which:

Fig.1 illustrates the analysis and synthesis sections of a digital filterbank;

Fig.2 shows the magnitudes in a composite alias component matrix of a cosine modulated filterbank;

Fig.3 shows the magnitudes in a composite alias component matrix of a complex-exponential modulated filterbank;

Fig.4 illustrates wanted terms and main alias terms in a cosine modulated filterbank adjusted for a band-pass filter response;

Fig.5 shows the attenuation of alias gain terms for different realizations of complex-exponential modulated filterbanks;

Fig.6 illustrates the analysis part of a complex-exponential modulated filterbank system according to the present invention; and

Fig.7 illustrates the synthesis part of a complex-exponential modulated filterbank system according to the present invention.

DESCRIPTION OF PREFERRED EMBODIMENTS

It should be understood that the present invention is applicable on a range of implementations that incorporates digital filterbanks other than those explicitly mentioned in this patent.

Digital Filterbanks

A digital filter bank is a collection of two or more parallel digital filters that share a common input or a common output ["Multirate Systems and Filter Banks" P.P. Vaidyanathan Prentice Hall: Englewood Cliffs, NJ, 1993]. When sharing a common input the filter bank is called an analysis bank. The analysis bank splits the incoming signal into M separate signals called subband signals. The analysis filters are denoted $H_k(z)$, where $k = 0 \ldots M\text{-}1$. The filter bank is critically sampled (or maximally decimated) when the subband signals are decimated by a factor M. The total number of subband samples per unit time is then the same as the number of samples per unit time for the input signal. The synthesis bank combines these subband signals into a common output signal. The synthesis filters are denoted $F_k(z)$, for $k = 0 \ldots M\text{-}1$. A maximally decimated filter bank with M channels (subbands) is shown in Fig.1. The analysis part 101 produces the signals $V_k(z)$, which constitute the signals to be transmitted, stored or modified, from the input signal $X(z)$. The synthesis part 102 recombines the signals $V_k(z)$ to the output signal $\hat{X}(z)$.

The recombination of $V_k(z)$ to obtain the approximation $\hat{X}(z)$ of the original signal $X(z)$ is subject to several errors. One of these is aliasing, due to the decimation and interpolation of the subbands. Other errors are phase and amplitude distortion.

Following the notations of Fig.1, the outputs of the analysis filters $H_k(z)$ 103 are

$$X_k(z) = H_k(z)X(z) \tag{1}$$

where $k = 0 \ldots M\text{-}1$. The decimators 104 give the outputs

$$V_k(z) = \frac{1}{M}\sum_{l=0}^{M-1} X_k(z^{1/M}W^l) = \frac{1}{M}\sum_{l=0}^{M-1} H_k(z^{1/M}W^l)X(z^{1/M}W^l) \tag{2}$$

where $W = e^{-i2\pi/M}$. The outputs of the interpolators 105 are given by

$$U_k(z) = V_k(z^M) = \frac{1}{M}\sum_{l=0}^{M-1} H_k(zW^l)X(zW^l) \tag{3}$$

and the sum of the signals obtained from the synthesis filters 106 can be written as

$$\begin{aligned}\hat{X}(z) &= \sum_{k=0}^{M-1} F_k(z)U_k(z) = \sum_{k=0}^{M-1} F_k(z)\frac{1}{M}\sum_{l=0}^{M-1} H_k(zW^l)X(zW^l) = \\ &= \frac{1}{M}\sum_{l=0}^{M-1} X(zW^l)\sum_{k=0}^{M-1} H_k(zW^l)F_k(z) = \frac{1}{M}\sum_{l=0}^{M-1} X(zW^l)A_l(z)\end{aligned} \tag{4}$$

where

$$A_l(z) = \sum_{k=0}^{M-1} H_k(zW^l)F_k(z) \tag{5}$$

is the gain for the lth alias term $X(zW^l)$. Eq.(4) can be rewritten as

$$\hat{X}(z)=\frac{1}{M}\left\{X(z)A_0(z)+\sum_{l=1}^{M-1}X(zW^l)A_l(z)\right\} \tag{6}$$

The last sum on the right hand side (RHS) constitutes the sum of all non-wanted alias terms. Canceling all aliasing, that is forcing this sum to zero by means of proper choices of $H_k(z)$ and $F_k(z)$, gives

$$\hat{X}(z)=\frac{1}{M}X(z)A_0(z)=\frac{1}{M}X(z)\sum_{k=0}^{M-1}H_k(z)F_k(z)=X(z)T(z) \tag{7}$$

where

$$T(z)=\frac{1}{M}\sum_{k=0}^{M-1}H_k(z)F_k(z) \tag{8}$$

is the overall transfer function or distortion function. Choosing the synthesis filter $F_k(z)$ so that

$$F_k(z)=z^{-N}\tilde{H}_k(z) \tag{9}$$

where N is the order of the analysis filter, results in the following transfer function:

$$T(z)=\frac{1}{M}\sum_{k=0}^{M-1}H_k(z)F_k(z)=\frac{z^{-N}}{M}\sum_{k=0}^{M-1}H_k(z)\tilde{H}_k(z) \tag{10}$$

The notation $\tilde{H}(z)$ is the Z-transform of the time-reversed and complex-conjugated sequence $h_k(n)$. Eq.(10) evaluated on the unit circle gives

$$T(e^{jw})=\frac{e^{-jwN}}{M}\sum_{k=0}^{M-1}H_k(e^{jw})H_k^*(e^{-jw})=\frac{e^{-jwN}}{M}\sum_{k=0}^{M-1}\left|H_k(e^{jw})\right|^2 \tag{11}$$

Eq.(11) shows that $T(z)$ has linear phase, and thus has no phase distortion. Further, if the last sum on the RHS is a constant, there is no amplitude distortion. The overall transfer function is in this case simply a delay with a constant scale factor c, i.e.

$$T(z)=cz^{-N} \tag{12}$$

which substituted into Eq.(7) gives

$$\hat{X}(z)=cz^{-N}X(z) \tag{13}$$

The type of filters that satisfy Eq.(13) are said to have the perfect reconstruction (PR) property.

Cosine Modulated Filterbanks

In a cosine modulated filterbank the analysis filters $h_k(n)$ are cosine modulated versions of a symmetric low-pass prototype filter $p_0(n)$ as

$$h_k(n) = 2p_0(n)\ \cos\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \tag{14}$$

where M is the number of channels, $k = 0 \ldots M\text{-}1$, N is the prototype filter order and $n = 0 \ldots N$. The sum of the real-valued prototype filter coefficients is assumed to be unity as

$$\sum_{n=0}^{N} p_0^2(n) = 1 \tag{15}$$

Following the same notation, the synthesis filters are given by

$$f_k(n) = 2p_0(N-n)\ \cos\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \tag{16}$$

The analysis filter bank produces real-valued subband samples for real-valued input signals. The subband samples are downsampled by a factor M, which makes the system critically sampled. Depending on the choice of the prototype filter, the filterbank may constitute a near perfect reconstruction system, a so called pseudo QMF bank [US5436940], or a perfect reconstruction (PR) system. An example of a PR system is the modulated lapped transform (MLT) ["Lapped Transforms for Efficient Transform/Subband Coding" H.S. Malvar, IEEE Trans ASSP, vol. 38, no. 6, 1990]. One inherent property of the cosine modulation is that every filter has two passbands; one in the positive frequency range and one corresponding passband in the negative frequency range.

Eq.(5) can be written in matrix form as

$$\mathbf{a} = \mathbf{H}\,\mathbf{f} \tag{17}$$

or explicitly as

$$\begin{bmatrix} A_0(z) \\ A_1(z) \\ \vdots \\ A_{M-1}(z) \end{bmatrix} = \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(Wz) & H_1(Wz) & & H_{M-1}(Wz) \\ \vdots & & \ddots & \vdots \\ H_0(W^{M-1}z) & H_1(W^{M-1}z) & \cdots & H_{M-1}(W^{M-1}z) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} \tag{18}$$

The matrix **H** is called the alias component (AC) matrix. For a closer examination of this equation, **f** may be written as

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ \cdot \\ \cdot \\ \cdot \\ F_{M-1}(z) \end{bmatrix} = \begin{bmatrix} F_0(z) & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & F_1(z) & & & & \cdot \\ \cdot & & \cdot & & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & F_{M-1}(z) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \tag{19}$$

or in a compact form as

$$\mathbf{f} = \mathbf{F}\,\mathbf{e} \tag{20}$$

Substituting Eq.(20) in Eq.(17), the alias gains can be written as $\mathbf{a} = \mathbf{HF}\,\mathbf{e}$, where the product

$$\mathbf{HF} = \mathbf{U} \tag{21}$$

is a $M \times M$ matrix, hereinafter referred to as the composite alias component matrix.

For a cosine modulated system, the dominant terms in the composite alias component matrix are the first row and four diagonals. The three-dimensional plot of Fig.2 illustrates the magnitudes of the components in this matrix. The first row holds the terms from the transfer function, Eq.(8), while the four diagonals primarily consist of the main alias terms, i.e. the aliasing due to overlap between filters and their closest neighbors. It is easily seen that the main alias terms emerge from overlap in frequency between either the filters negative passbands with frequency modulated versions of the positive passbands, or reciprocally, the filters positive passbands with frequency modulated versions of the negative passbands. Summing the terms of the rows in the composite alias component matrix, i.e. calculating the alias gains, results in cancellation of the main alias terms. The aliasing is canceled in a pairwise manner, where the first main alias term is canceled by the second in the same row. Superimposed on the main alias terms are also other smaller alias terms. If the prototype filter characteristics is so that the transition- and stop-band of the filters have substantial overlap with their modulated versions, these alias terms will be large. As an example, the second and the last row consists of alias terms induced by the overlap of filters with their closest modulated versions. For a PR system, these smaller alias terms also cancels completely when summing the terms for the alias gains. In the pseudo QMF system, however, these terms remain.

Complex-Exponential Modulated Filterbanks

Extending the cosine modulation to complex-exponential modulation according to the present invention yields the analysis filters $h_k(n)$ as

$$h_k(n) = p_0(n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)\left(n - \frac{N}{2} - \frac{M}{2}\right)\right\} \tag{22}$$

using the same notation as before. This can be viewed as adding an imaginary part to the real-valued filterbank, where the imaginary part consists of sine modulated versions of the same prototype filter. Considering a real-valued input signal, the output from the filter bank can be interpreted as a set of subband signals, where the real and the imaginary parts are Hilbert transforms of each other. The resulting subbands are thus the analytic signals of the real-valued output obtained from the cosine modulated filterbank. Hence, due to the complex-valued representation, the subband signals are oversampled by a factor two.

The synthesis filters are extended in the same way as

$$f_k(n) = p_0(N-n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)\left(n - \frac{N}{2} + \frac{M}{2}\right)\right\} \tag{23}$$

Eq.(22) and (23) implies that the output from the synthesis bank is complex-valued. Using matrix notation, where $\mathbf{C}_a$ is a matrix with analysis filters from Eq.(14), and $\mathbf{S}_a$ is a matrix with filters as

$$h_k(n) = p_0(n)\ \sin\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \qquad (24)$$

the filters of Eq.(22) is obtained as $\mathbf{C_a} + \mathrm{j}\,\mathbf{S_a}$. In these matrices, k is the row index and n is the column index. Analogously, the matrix $\mathbf{C_s}$ has synthesis filters from Eq.(16), and $\mathbf{S_s}$ is a matrix with filters as

$$f_k(n) = p_0(N-n)\ \sin\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \qquad (25)$$

Eq.(23) can thus be written $\mathbf{C_s} + \mathrm{j}\,\mathbf{S_s}$, where k is the column index and n is the row index. Denoting the input signal **x**, the output signal **y** is found from

$$\mathbf{y} = (\mathbf{C_s} + \mathrm{j}\,\mathbf{S_s})\,(\mathbf{C_a} + \mathrm{j}\,\mathbf{S_a})\,\mathbf{x} = (\mathbf{C_s}\mathbf{C_a} - \mathbf{S_s}\mathbf{S_a})\,\mathbf{x} + \mathrm{j}\,(\mathbf{C_s}\mathbf{S_a} + \mathbf{S_s}\mathbf{C_a})\,\mathbf{x} \qquad (26)$$

As seen from Eq.(26), the real part consists of two terms; the output from the ordinary cosine modulated filterbank and an output from a sine modulated filterbank. It is easily verified that if a cosine modulated filterbank has the PR property, then its sine modulated version, with a change of sign, constitutes a PR system as well. Thus, by taking the real part of the output, the complex-exponential modulated system offers the same reconstruction accuracy as the corresponding cosine modulated version.

The complex-exponential modulated system can be extended to handle also complex-valued input signals. By extending the number of channels to $2M$, i.e. adding the filters for the negative frequencies, and keeping the imaginary part of the output signal, a pseudo QMF or a PR system for complex-valued signals is obtained.

Examining the composite alias component matrix from Eq.(21), the main alias diagonals vanish for the complex-exponential modulated filterbank. This is easily understood since the complex-exponential modulated filterbank has only one passband for every filter. In other words, the filterbank is free from main alias terms, and do not rely on the pairwise aliasing cancellation mentioned above. The composite alias component matrix has the dominant terms on the first row only. Fig.3 shows the magnitude of the components in the resulting matrix. Depending on the prototype filter characteristics, the terms on rows 1 through M-1, are more or less attenuated. The absence of main alias terms makes the aliasing cancellation constraint from the cosine (or sine) modulated filterbank obsolete in the complex-exponential modulated version. Both the analysis and synthesis filters can thus be found from

$$h_k(n) = f_k(n) = p_0(n)\ \exp\left\{i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\} \qquad (27)$$

since $p_0(n) = p_0(N-n)$ for a symmetric prototype filter. As before, M is the number of channels, $k = 0 \ldots M$-1, N is the prototype filter order and $n = 0 \ldots N$.

Referring to Eq.(4), the Z-transform of the real part of the output signal $\hat{x}(n)$ is

$$Z\left\{\mathrm{Re}\left(\hat{x}(n)\right)\right\} = \hat{X}_R(z) = \frac{\hat{X}(z) + \hat{X}*(z)}{2} \qquad (28)$$

The notation $\hat{X}*(z)$ is the Z-transform of the complex-conjugated sequence $\hat{x}(n)$. From Eq.(4), it follows that the transform of the real part of the output signal is

$$2M\hat{X}_R(z) = \sum_{l=0}^{M-1}\left(X(zW^l)A_l(z) + X(zW^{-l})A_l*(z)\right) \tag{29}$$

where it is used that the input signal $x(n)$ is real-valued. Eq.(29) may after manipulation be written

$$\begin{aligned} 2M\hat{X}_R(z) = {} & X(z)\{A_0(z) + A_0*(z)\} + \\ & + X(-z)\{A_{M/2}(z) + A_{M/2}*(z)\} + \\ & + \sum_{l=1}^{M/2-1} X(zW^l)\{A_l(z) + A_{M-l}*(z)\} + \\ & + \sum_{l=1}^{M/2-1} X(zW^{-l})\{A_{M-l}(z) + A_l*(z)\} \end{aligned} \tag{30}$$

By inspecting Eq.(30), and recalling the transform of Eq.(28), it is obvious that the real part of $a_0(n)$ must be a dirac pulse for a PR system. Moreover, the real part of $a_{M/2}(n)$ must be zero and the alias gains, for $l = 1 \ldots M/2\text{-}1$, must satisfy

$$A_{M-l}(z) = -A_l*(z) \tag{31}$$

In pseudo QMF systems, Eq.(31) holds true only approximately. Moreover, the real part of $a_0(n)$ is not exactly a dirac-pulse, nor is the real part of $a_{M/2}$(n) exactly zero.

Modification of subband signals

Changing the gains of the channels in a cosine modulated filterbank, i.e. using the analysis/synthesis system as an equalizer, renders severe distortion due to the main alias terms. Say for example that the intention is to adjust an eight-channel filterbank for a band-pass response, where except for the second and third channel all subband signals are set to zero. The composite alias component matrix from Eq.(21) is then an 8 × 8 matrix where all elements are zero except for the elements on the second and third column, Fig.4. There are seven significant alias terms left as indicated in the figure. The aliasing from rows three and five will be canceled, since the main alias terms in these rows have the same gains, i.e. the pairwise cancellation is working intentionally. In rows two, four and six however, there are single alias terms, since their corresponding alias terms have zero gain. The alias cancellation will hence not work as intended, and the aliasing in the output signal will be substantial.

From this example it is obvious that a substantial improvement is achieved when using complex-exponential modulated filterbanks as equalizers. The 8-channel system depicted in Fig.4 has a prototype filter of order 128. The total aliasing attenuation is only 16 dB in the above equalizer example. Moving to complex-exponential modulation gives an aliasing attenuation of 95 dB. Due to the non-existing main

alias terms, the resulting aliasing is dependent only on the suppression of the alias terms emanating from overlap between filters and their modulated versions. It is thus of great importance to design the prototype filter for maximum suppression of the alias gains terms. The first term on the RHS of Eq.(30) evaluated on the unit circle gives the error energy e_t of the transfer function as

$$e_t = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\left|A_0(e^{j\omega}) + A_0^*(e^{-j\omega})\right|}{2M} - 1\right)^2 d\omega \qquad (32)$$

The energy of the total aliasing e_a can be calculated by evaluating all remaining terms on the RHS of Eq.(30) on the unit circle as

$$\begin{aligned} e_a = {} & \frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega}) + A_{M/2}^*(e^{-j\omega})\right|^2 d\omega + \\ & + \frac{1}{8\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left\{\left|A_l(e^{j\omega}) + A_{M-l}^*(e^{-j\omega})\right|^2 + \right. \\ & \left. + \left|A_{M-l}(e^{j\omega}) + A_l^*(e^{-j\omega})\right|^2\right\} d\omega \end{aligned} \qquad (33)$$

Due to symmetry, Eq.(9) and the fact

$$P_0(z) = z^{-N}\tilde{P}_0(z) \qquad (34)$$

the terms within the curly braces of the sum in Eq.(33) are equal. The total aliasing energy thus has $M/2$-1 terms as

$$\begin{aligned} e_a = {} & \frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega}) + A_{M/2}^*(e^{-j\omega})\right|^2 d\omega + \\ & + \frac{1}{4\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left|A_l(e^{j\omega}) + A_{M-l}^*(e^{-j\omega})\right|^2 d\omega \end{aligned} \qquad (35)$$

The minimization of the alias gain terms is done by optimizing the prototype filter. This is preferably accomplished by minimizing a composite objective function, using standard nonlinear optimization algorithms, for example the Downhill Simplex Method ["Numerical Recipes in C, The Art of Scientific Computing, Second Edition" W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Cambridge University Press, NY, 1992]. For alias term minimization (ATM) of the prototype filter according to the invention, the objective function looks like

$$\varepsilon_{tot}(\alpha) = \alpha\varepsilon_t + (1-\alpha)\varepsilon_a \qquad (36)$$

During the optimization, a random equalization curve is applied to the filterbank when calculating ε_a, i.e. the analysis and synthesis filters are multiplied with gainfactors g_k as

$$H_k^{(eq)}(z) = g_k H_k(z) \quad \text{and} \quad F_k^{(eq)}(z) = g_k F_k(z) \tag{37}$$

and the resulting filters $H_k^{(eq)}$ and $F_k^{(eq)}$, $k = 0 \ldots M\text{-}1$, are used when calculating the alias gain terms $A_l(z)$ for $l = 1 \ldots M\text{-}1$.

In Fig.5, the alias gains of five different complex-exponential modulated systems are compared. Four of these are 8-channel systems and one is a 64-channel system. All of the systems have prototype filter lengths of 128. The dotted trace and the solid trace with stars shows alias components for two pseudo QMF systems, where one is alias term minimized. The dashed and the dash-dotted traces are the components for two 8-channel perfect reconstruction systems, where again one of the systems is alias term minimized. The solid trace is the alias components for a complex-exponential modulated lapped transform (MLT). All the systems are adjusted for band-pass responses according to the example above, and the results are tabulated in Table 1. The rejection of total aliasing is calculated as the inverse of Eq.(33). The passband flatness is calculated as the inverse of Eq.(32) with the integration interval adjusted for the bandpass response.

Table 1

System	Rejection of total aliasing	Passband flatness
8-ch Pseudo QMF ATM N=128	114,7 dB	98,1 dB
8-ch Pseudo QMF N=128	95,4 dB	87,6 dB
8-ch PR ATM N=127	77,3 dB	132,7 dB
8-ch PR N=127	55,0 dB	93,6 dB
64-ch MLT N=127	38,5 dB	87,1 dB

As seen from the numbers in Table 1, a substantial improvement is achieved when moving from the 64-channel MLT to the 8-channel PR systems. The MLT is a perfect reconstruction system and has only $(N+1) / 2M = 1$ coefficient per polyphase component. The number of coefficients for the 8-channel PR systems are $128 / 16 = 8$. This enables filters with higher stopband attenuation and higher rejection of alias terms. Moreover, it is seen that alias term minimization of the PR system rejects the aliasing and improves the passband flatness significantly. Comparing the pseudo QMF systems and the PR systems, it is clear that the aliasing rejection improves by 40 dB while almost preserving the passband flatness. An additional alias rejection of approximately 20 dB and improved passband flatness of 10 dB is achieved when minimizing the alias terms. Thus, it is clear that the perfect reconstruction constraint imposes limitations to a filterbank used in an equalization system. A pseudo QMF system can always be designed for adequate reconstruction accuracy, since all practical digital implementations have limited resolution in the numeric representation. For both the pseudo QMF and the PR system, it is obvious that an optimum system is built on a prototype filter that has large rejection of the stopband. This imposes the usage of prototype filters with relative lengths longer than the windows used in the MLT.

A great advantage of the complex-exponential modulated system is that the instantaneous energy is easily calculated since the subband signals constitute the analytic signals of the real-valued subband signals

obtained from a cosine modulated filterbank. This is a property of great value in for example adaptive filters, automatic gain controls (AGC), in multiband companders, and in spectral band replication systems (SBR), where a filterbank is used for the spectral envelope adjustment. The averaged energy within a subband k might be calculated as

$$E_k(m) = \sum_{n=-L+1}^{L-1} |v_k(mL+n)|^2 w(n) \qquad (38)$$

where $v_k(n)$ is the subband samples of channel k, and $w(n)$ is a window of length $2L$-1 centered around $n = 0$. This measure is then used as an input parameter for the adaptation or gain-calculation algorithms.

Practical implementations

Using a standard PC or DSP, real-time operation of a complex-exponential modulated filterbank is possible. The filterbank may also be hard-coded on a custom chip. Fig.6 shows the structure for an effective implementation of the analysis part of a complex-exponential modulated filterbank system. The analogue input signal is first fed to an A/D converter 601. The digital time domain signal is fed to a shift register holding $2M$ samples shifting M samples at a time 602. The signals from the shift register are then filtered through the polyphase coefficients of the prototype filter 603. The filtered signals are subsequently combined 604 and in parallel transformed with a DCT-IV 605 and a DST-IV 606 transform. The outputs from the cosine and sine transforms constitute the real and the imaginary parts of the subband samples respectively. The gains of the subband samples are modified according to the current spectral envelope adjuster setting 607.

An effective implementation of the synthesis part of a complex-exponential modulated system is shown in Fig.7. The subband samples are first multiplied with complex-valued twiddlefactors 701, and the real part is modulated with a DCT-IV 702 and the imaginary part with a DST-IV 703 transform. The outputs from the transforms are combined 704 and fed through the polyphase components of the prototype filter 705. The time domain output signal is obtained from the shift register 706. Finally, the digital output signal is converted back to an analogue waveform 707.

The above-described embodiments are merely illustrative for the principles of the complex-exponential modulated filterbank systems according to the present invention. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.

CLAIMS

1. A method for estimation of energy measures of coefficients or subband signals obtained from a digital filterbank, **characterized by:**

- optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- building an M-channel analysis filterbank by complex-exponential modulation of said prototype filter; where said filterbank has filter coefficients as

$$h_k(n) = p_0(n) \ \exp\left\{ i \ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\},$$

and $n = 0, 1, \ldots, N$, and $k = 0,1, \ldots, M\text{-}1$;
- filtering a real-valued time domain signal through said filterbank; and
calculating the squared absolute values of the complex-valued subband signals obtained from said filtering.

2. A method according to claim **1, characterized in** that said analysis filterbank is used to estimate average energy measures as

$$E_k(m) = \sum_{n=-L+1}^{L-1} |v_k(mL+n)|^2 w(n)$$

where $v_k(n)$ is said complex-valued subband signal of channel k, and $w(n)$ is a window of length $2L\text{-}1$ centered around $n = 0$.

3. A method for reduction of aliasing emerging from modifications of the coefficients or subband signals obtained from a digital filterbank, **characterized by:**

- optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- building an M-channel filterbank by complex-exponential modulation of said prototype filter; where said filterbank has analysis and synthesis filter coefficients as

$$h_k(n) = f_k(n) = p_0(n) \ \exp\left\{ i \ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\},$$

and $n = 0 \ldots N$, and $k = 0 \ldots M\text{-}1$;
- filtering a real-valued time domain signal through the analysis part of said filterbank;
- modifying the complex-valued subband signals obtained from said filtering;
- filtering said modified complex-valued subband signals through the synthesis part of said filterbank; and
- taking the real part of the complex-valued time domain output signal, where said output signal is the sum of the signals obtained from said synthesis filtering.

4. A method according to claims **1 - 3, characterized in** that said low-pass prototype filter has a filter order N that is higher than $2M$-1, where M is the number of channels in said digital filterbank.

5. A method according to claims **1 – 4, characterized in** that said optimization of said low-pass prototype filter is accomplished by minimizing a composite objective function $\varepsilon_{tot}(\alpha)$ as

$$\varepsilon_{tot}(\alpha) = \alpha\varepsilon_t + (1-\alpha)\varepsilon_a$$

where α is a weighting constant, ε_t is the passband error and ε_a is the aliasing error.

6. A method according to claim **5, characterized in** that said passband error ε_t is calculated as

$$e_t = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\left|A_0(e^{j\omega}) + A_0^*(e^{-j\omega})\right|}{2M} - 1\right)^2 d\omega,$$

and said aliasing error ε_a is calculated as

$$e_a = \frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega}) + A_{M/2}^*(e^{-j\omega})\right|^2 d\omega + \\ + \frac{1}{4\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left|A_l(e^{j\omega}) + A_{M-l}^*(e^{-j\omega})\right|^2 d\omega,$$

where $A_l(e^{j\omega})$, $l = 0 \ldots M/2$, are the alias gain terms evaluated on the unit circle.

7. A method according to claim **6, characterized in** that said alias gain terms are calculated using gain adjusted versions of said analysis and synthesis filters according to

$$h_k^{(eq)}(n) = g_k h_k(n) \quad \text{and} \quad f_k^{(eq)}(n) = g_k f_k(n),$$

where g_k represents randomized or deterministic real-valued gainfactors, $n = 0 \ldots N$, and $k = 0 \ldots M$-1.

8. A method according to claims **5 – 7, characterized in** that said low-pass prototype filter has a filter order of at least a number N, where N is chosen so that said composite objective function is smaller than a predefined value.

9. A method according to claims **1 – 2, characterized in** that said analysis filterbank is used for the estimation of energy measures in a high frequency reconstruction system.

10. A method according to claims **3 – 8, characterized in** that said complex-exponential modulated filterbank is used as the envelope adjusting filterbank in a high frequency reconstruction system.

11. A method according to claims **3 – 8, characterized in** that said complex-exponential modulated filterbank is used as the filters in a digital equalizing system.

12. A method according to claims **3 – 8, characterized in** that said complex-exponential modulated filterbank is used in an adaptive filtering system.

13. A method according to claims **3 – 8, characterized in** that said complex-exponential modulated filterbank is used in a multiband compander system.

14. An apparatus for estimation of energy measures of coefficients or subband signals obtained from a digital filterbank, **characterized by:**

- means for optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- means for building an M-channel analysis filterbank by complex-exponential modulation of said prototype filter; where said filterbank has filter coefficients as

$$h_k(n) = p_0(n) \ \exp\left\{ i \ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\},$$

and $n = 0, 1, \ldots, N$, and $k = 0,1, \ldots, M\text{-}1$;

- means for filtering a real-valued time domain signal through said filterbank; and

means for calculation of the squared absolute values of the complex-valued subband signals obtained from said filtering.

15. An apparatus for reduction of aliasing emerging from modifications of the coefficients or subband signals obtained from a digital filterbank, **characterized by:**

- means for optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- means for building an M-channel filterbank by complex-exponential modulation of said prototype filter; where said filterbank has analysis and synthesis filter coefficients as

$$h_k(n) = f_k(n) = p_0(n) \ \exp\left\{ i \ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\},$$

and $n = 0, 1, \ldots, N$, and $k = 0,1, \ldots, M\text{-}1$;

- means for filtering a real-valued time domain signal through the analysis part of said filterbank;
- means for modifying the complex-valued subband signals obtained from said filtering;
- means for filtering said modified complex-valued subband signals through the synthesis part of said filterbank; and

means for taking the real part of the complex-valued time domain output signal, where said output signal is the sum of the signals obtained from said synthesis filtering.

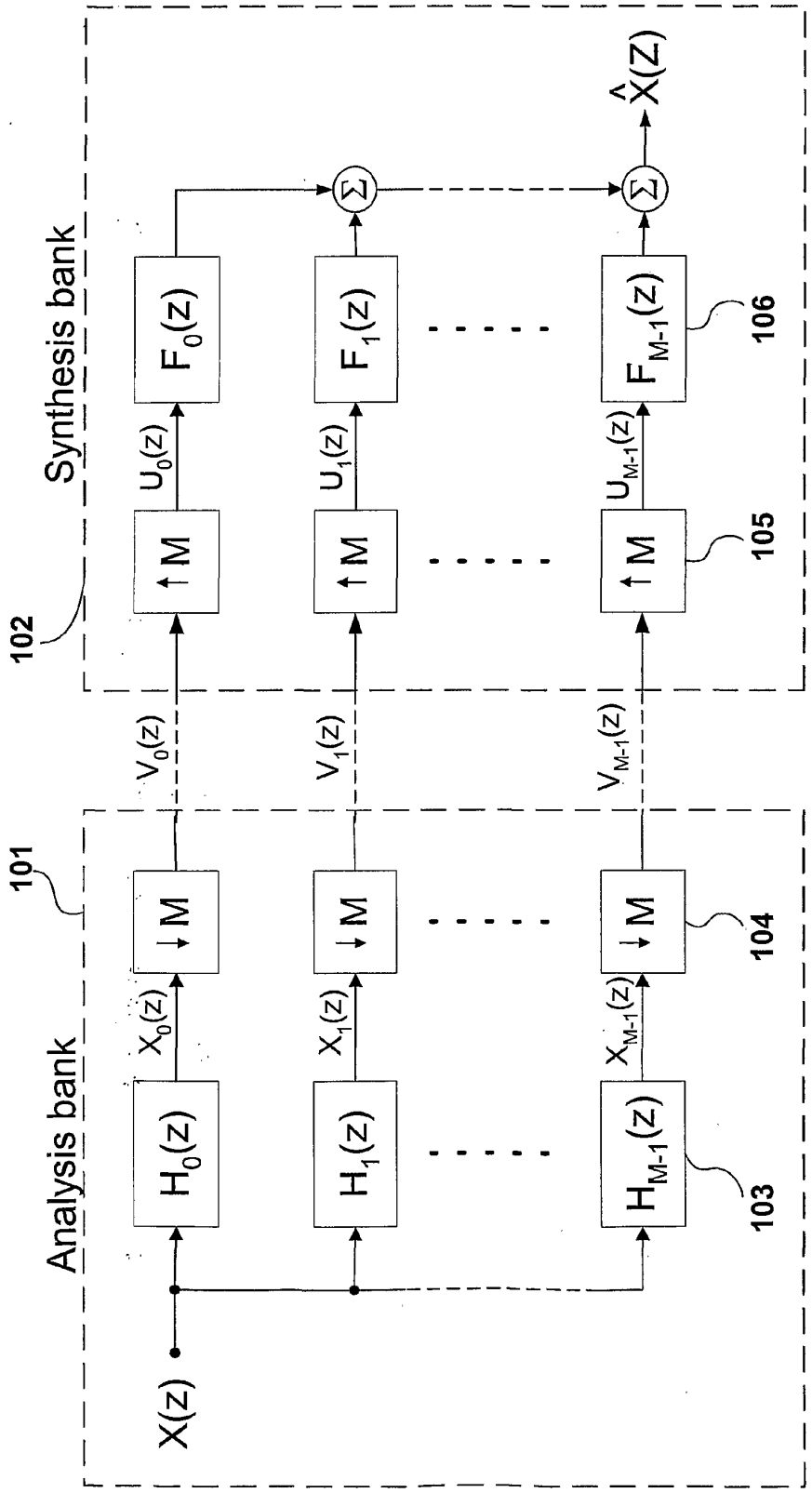

Analysis bank
101
X(z)
H0(z)
H1(z)
HM-1(z)
103
X0(z)
X1(z)
XM-1(z)
↓M
104
V0(z)
V1(z)
VM-1(z)
Synthesis bank
102
↑M
105
U0(z)
U1(z)
UM-1(z)
F0(z)
F1(z)
FM-1(z)
106
Σ
X̂(Z)

Fig. 1

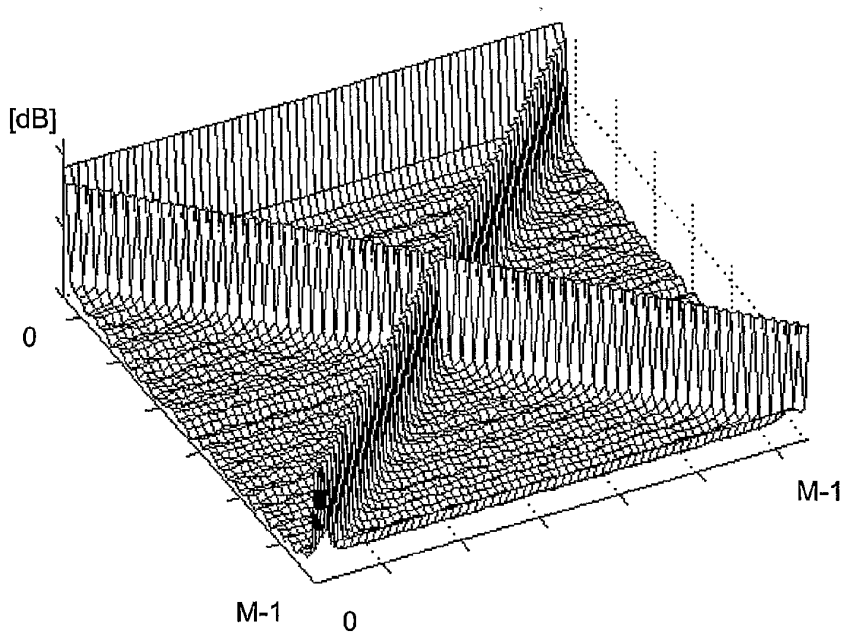


Fig. 2

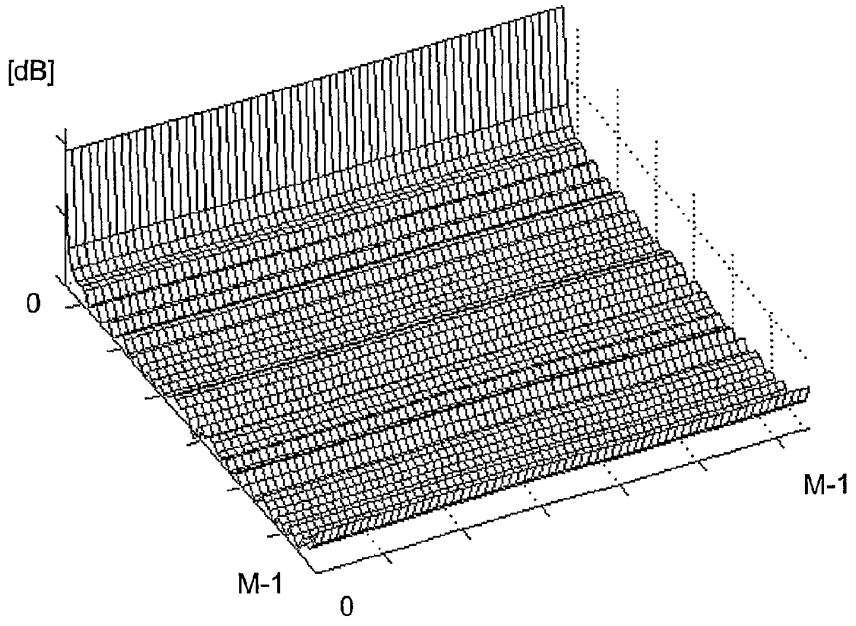


Fig. 3

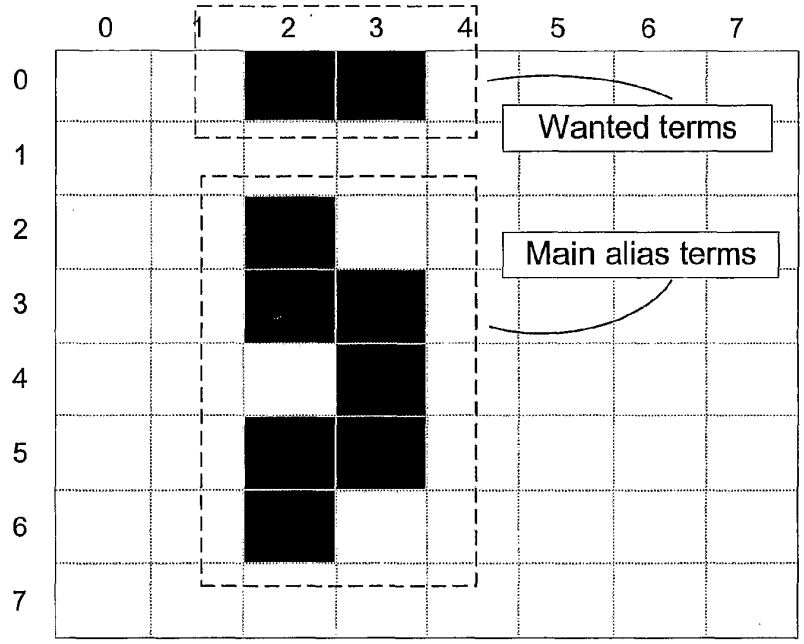

0
1
2
3
4
5
6
7
Wanted terms
Main alias terms

Fig. 4

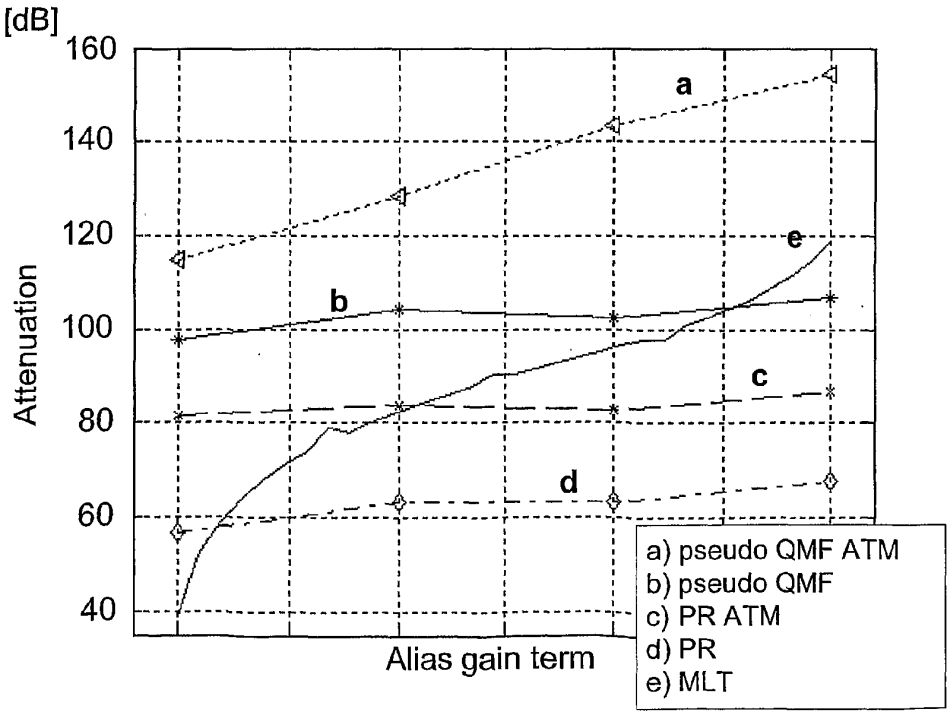

[dB]
160
140
120
100
80
60
40
Attenuation
Alias gain term
a
b
c
d
e
a) pseudo QMF ATM
b) pseudo QMF
c) PR ATM
d) PR
e) MLT

Fig. 5

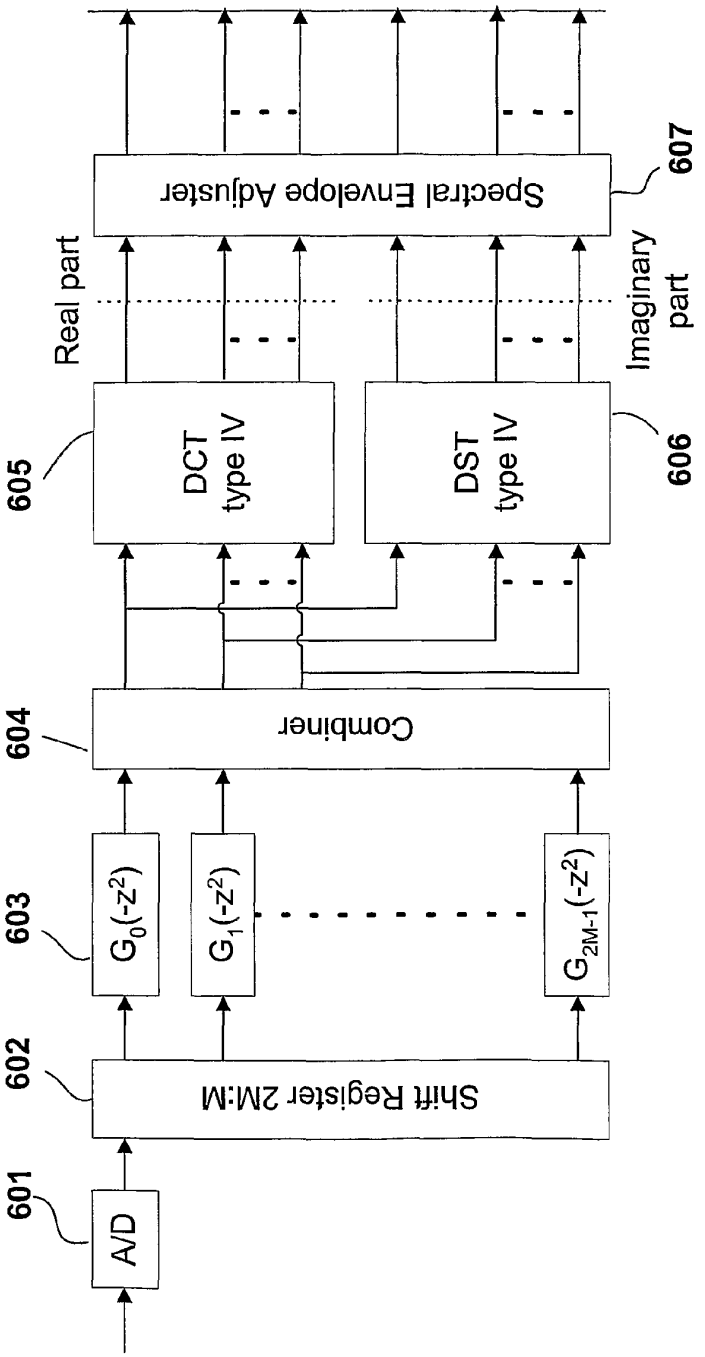

601
A/D
602
Shift Register 2M:M
603
$G_0(-z^2)$
$G_1(-z^2)$
$G_{2M-1}(-z^2)$
604
Combiner
605
DCT type IV
606
DST type IV
607
Spectral Envelope Adjuster
Real part
Imaginary part

Fig. 6

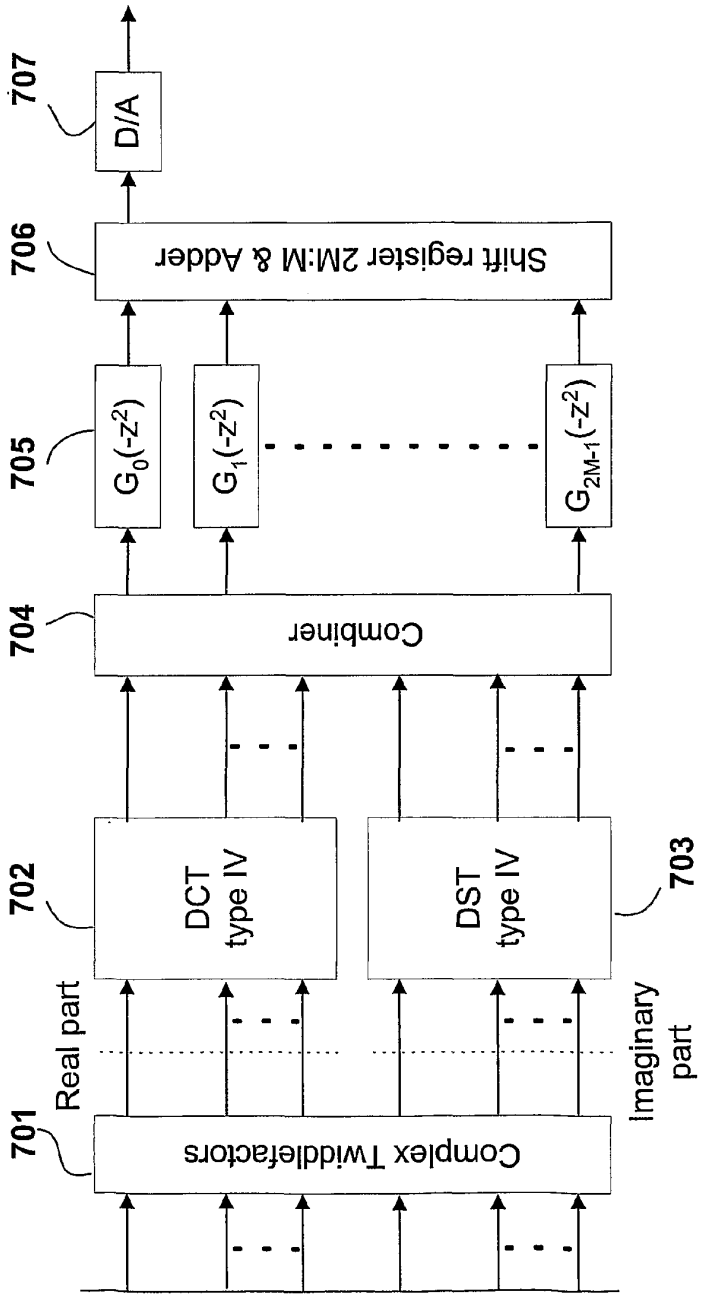

707
D/A
706
Shift register 2M:M & Adder
705
$G_0(-z^2)$
$G_1(-z^2)$
$G_{2M-1}(-z^2)$
704
Combiner
702
DCT type IV
DST type IV
703
Real part
Imaginary part
701
Complex Twiddlefactors

Fig. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 02/00626

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H03H 17/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H03H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5732189 A (JOHNSTON ET AL), 24 March 1998 (24.03.98), column 5, line 11 - column 9, line 63	1-15
A	US 5285498 A (JOHNSTON), 8 February 1994 (08.02.94), column 3, line 1 - column 4, line 9	1-15
A	WO 0141320 A1 (SOLANA TECHNOLOGY DEVELOPMENT CORPORATION), 7 June 2001 (07.06.01), See the whole document	1-15
A	US 5568142 A (VELAZQUEZ ET AL), 22 October 1996 (22.10.96), See the whole document	1-15

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 July 2002	19 -07- 2002
Name and mailing address of the ISA Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Stefan Hultquist/SN Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

10/06/02

International application No.
PCT/SE 02/00626

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5732189 A	24/03/98	NONE	
US 5285498 A	08/02/94	CA 2090159 A	03/09/93
		DE 69326492 D,T	04/05/00
		EP 0559383 A,B	08/09/93
		JP 3258424 B	18/02/02
		JP 6075590 A	18/03/94
		KR 9707661 B	15/05/97
		US 5481614 A	02/01/96
WO 0141320 A1	07/06/01	AU 1837200 A	12/06/01
US 5568142 A	22/10/96	CA 2203156 A	02/05/96
		EP 0787385 A	06/08/97
		JP 10507891 T	28/07/98
		WO 9613097 A	02/05/96

Form PCT/ISA/210 (patent family annex) (July 1998)

AUDIEN.FOR0044

Espacenet

Bibliographic data: JP2004533155 (A) — 2004-10-28

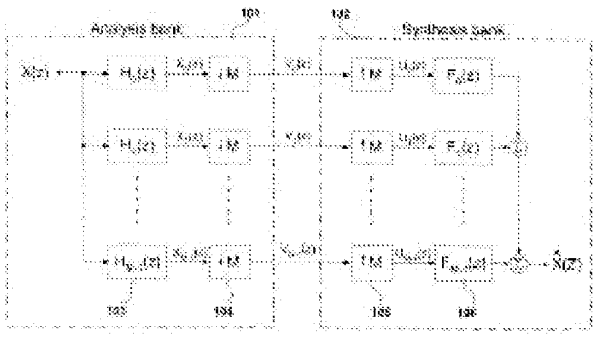
ALIASING REDUCTION USING COMPLEX-EXPONENTIAL MODULATED FILTERBANKS

Inventor(s):

Applicant(s):

Classification:	**- international:**	***G10L11/00; H03H17/02; H03H21/00; H03M7/30;*** (IPC1-7): G10L11/00; H03H17/02; H03H21/00; H03M7/30
	- European:	H03H17/02F8A
Application number:	JP20020578650T 20020328	
Priority number(s):	SE20010001175 20010402; WO2002SE00626 20020328	
Also published as:	JP3977744 (B2) WO02080362 (A1) US2003016772 (A1) US7242710 (B2) JP2006203881 (A) more	

Abstract not available for JP2004533155 (A)
Abstract of corresponding document: WO02080362 (A1)



The present invention proposes a new method and apparatus for the improvement of digital filterbanks, by a complex extension of cosine modulated digital filterbanks. The invention employs complex-exponential modulation of a low-pass prototype filter and a new method for optimizing the characteristics of this filter. The invention substantially reduces artifacts due to aliasing emerging from independent modifications of subband signals, for example when using a filterbank as an spectral equalizer. The invention is preferably implemented in software, running on a standard PC or a digital signal processor (DSP), but can also be hardcoded on a custom chip. The invention offers essential improvements for various types of digital equalizers, adaptive filters, multiband companders and spectral envelope adjusting filterbanks used in high frequency reconstruction (HFR) systems.

AUDIEN.FOR0044

Last updated: 28.02.2012 Worldwide Database 5.7.36; 93p

(19)日本国特許庁(JP)　(12)公表特許公報(A)　(11)特許出願公表番号
特表2004-533155
(P2004-533155A)
(43)公表日　平成16年10月28日(2004.10.28)

(51)Int.Cl.[7]	F I		テーマコード（参考）
H03H 17/02	H03H 17/02	601G	5J023
G10L 11/00	H03H 17/02	613C	5J064
H03H 21/00	H03H 21/00		
H03M 7/30	H03M 7/30	A	
	G10L 7/04	A	

審査請求 有　予備審査請求 有　（全 40 頁）

(21)出願番号　特願2002-578650(P2002-578650)
(86)(22)出願日　平成14年3月28日(2002.3.28)
(85)翻訳文提出日　平成15年9月26日(2003.9.26)
(86)国際出願番号　PCT/SE2002/000626
(87)国際公開番号　WO2002/080362
(87)国際公開日　平成14年10月10日(2002.10.10)
(31)優先権主張番号　0101175-8
(32)優先日　平成13年4月2日(2001.4.2)
(33)優先権主張国　スウェーデン(SE)

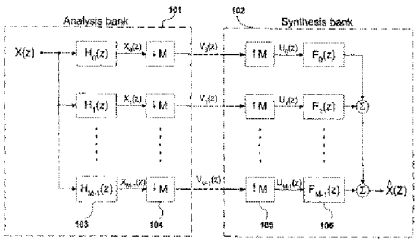
(71)出願人　502112267
コーディング　テクノロジーズ　アクチボラゲット
スウェーデン国　ストックホルム、ドベルンスガタン　６４
(74)代理人　100085464
弁理士　野口　繁雄
(72)発明者　ペル・エクストランド
スウェーデン国　Ｓ－１１６４０　ストックホルム　ゼーデルマンナガタン　４５

Fターム(参考)　5J023 DB01 DB07 DC03 DD01
5J064 AA01 BA16 BC11 BC15 BC18

最終頁に続く

(54)【発明の名称】複素指数変調フィルタバンクを使用するエイリアシングの低減

(57)【要約】
本発明は、余弦変調デジタル・フィルタバンクの複合的拡張によってデジタル・フィルタバンクを改良するための新しい方法及び装置を提案する。本発明は、低域通過プロトタイプ・フィルタの複素指数変調及びこのフィルタの特性を最適化するための新規方法を使用する。本発明は、例えばフィルタバンクをスペクトル・イコライザとして使用する際にサブバンド信号の個々の修正から発生するエイリアシングに起因するアーティファクトを大幅に低減する。本発明は、標準的なＰＣ又はデジタル信号プロセッサ（ＤＳＰ）上で実行されるソフトウェアで実現されることが好適であるが、カスタム・チップ上にハードコード化されることも可能である。本発明は、高周波数再構成（ＨＦＲ）システムにおいて使用される様々なタイプのデジタル・イコライザ、適応型フィルタ、マルチバンド・コンパンダ及びスペクトル包絡線調整フィルタバンクを本質的に改善するものである。



【特許請求の範囲】
【請求項1】
デジタル・フィルタバンクから取得される係数又はサブバンド信号のエネルギー測度を推定する方法であって、
完全再構成特性を有するようには制約されない、フィルタ次数Ｎを有する低域通過対称プロトタイプ・フィルタp_0（ｎ）を最適化することと、
Ｍチャネルの分析フィルタバンクであって、該フィルタバンクが、

$$h_k(n) = p_0(n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\}$$

で表わされるフィルタ係数を有し、ｎ＝０，１，．．．，Ｎでありかつｋ＝０，１，．．．，Ｍ－１であるものを前記プロトタイプ・フィルタの複素指数変調によって構築することと、
前記フィルタバンクを介して実数値の時間領域信号をフィルタリングすることと、
前記フィルタリングにより取得された複素数値のサブバンド信号の絶対値二乗を計算することを特徴とする方法。
【請求項2】
前記分析フィルタバンクは、

$$E_k(m) = \sum_{n=-L+1}^{L-1} |v_k(mL+n)|^2 w(n)$$

で表わされる平均エネルギー測度の推定に使用されることを特徴とする請求項１に記載の方法。ここで、v_k（ｎ）はチャネルｋの前記複素数値サブバンド信号であり、ｗ（ｎ）はｎ＝０近傍を中心とする長さ２Ｌ－１の窓である。
【請求項3】
デジタル・フィルタバンクから取得される係数又はサブバンド信号の修正によって発生するエイリアシングを低減する方法であって、
完全再構成特性を有するようには制約されない、フィルタ次数Ｎを有する低域通過対称プロトタイプ・フィルタp_0（ｎ）を最適化することと、
Ｍチャネルのフィルタバンクであって、該フィルタバンクが、

$$h_k(n) = f_k(n) = p_0(n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\}$$

で表わされる分析及び合成フィルタ係数を有し、ｎ＝０．．．Ｎでありかつｋ＝０．．．Ｍ－１であるものを前記プロトタイプ・フィルタの複素指数変調によって構築することと、
前記フィルタバンクの分析部分を介して実数値の時間領域信号をフィルタリングすることと、
前記フィルタリングにより取得された複素数値のサブバンド信号を修正することと、
前記修正された複素数値のサブバンド信号を前記フィルタバンクの合成部分を介してフィルタリングすることと、
複素数値の時間領域出力信号であって、該出力信号が前記合成フィルタリングにより取得された信号の和であるものの実数部を取り出すことを特徴とする方法。
【請求項4】
前記低域通過プロトタイプ・フィルタは２Ｍ－１より高次のフィルタ次数Ｎを有し、Ｍは前記デジタル・フィルタバンクにおけるチャネル数であることを特徴とする請求項１から３に記載の方法。
【請求項5】
前記低域通過プロトタイプ・フィルタの前記最適化は、

$$\varepsilon_{tot}(\alpha)=\alpha\varepsilon_t+(1-\alpha)\varepsilon_a$$

で表わされる合成目的関数ε_{tot}（α）の最小化によって達成されることを特徴とする請求項１から４に記載の方法。ここで、αは重み定数であり、ε_tは通過帯域誤差であり、ε_aはエイリアシング誤差である。

【請求項6】

前記通過帯域誤差ε_tは、

$$e_t=\frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\left|A_0(e^{j\omega})+A_0^{*}(e^{-j\omega})\right|}{2M}-1\right)^2 d\omega$$

として計算され、

かつ前記エイリアシング誤差ε_aは、

$$e_a=\frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega})+A_{M/2}^{*}(e^{-j\omega})\right|^2 d\omega+$$
$$+\frac{1}{4\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left|A_l(e^{j\omega})+A_{M-l}^{*}(e^{-j\omega})\right|^2 d\omega$$

として計算されることを特徴とする請求項５に記載の方法。ここで、$A_l(e^{j\omega})$，l＝０．．．M／２，は単位円上で評価されたエイリアス利得項である。

【請求項7】

前記エイリアス利得項は、

$$h_k^{(eq)}(n)=g_k h_k(n)\quad 及び\quad f_k^{(eq)}(n)=g_k f_k(n)$$

による前記分析フィルタ及び合成フィルタの利得調整バージョンを使用して計算されることを特徴とする請求項６に記載の方法。ここで、g_kは確率的又は確定的な実数値利得係数を表し、ｎ＝０．．．Ｎでありかつｋ＝０．．．Ｍ－１である。

【請求項8】

前記低域通過プロトタイプ・フィルタは少なくとも数値Ｎのフィルタ次数を有することを特徴とする請求項５から７に記載の方法。ここで、Ｎは前記合成目的関数が予め定義された値より小さくなるように選定される。

【請求項9】

前記分析フィルタバンクは、高周波数再構成システムにおいてエネルギー測度の推定に使用されることを特徴とする請求項１から２に記載の方法。

【請求項10】

前記複素指数変調フィルタバンクは、高周波数再構成システムにおいて包絡線調整フィルタバンクとして使用されることを特徴とする請求項３から８に記載の方法。

【請求項11】

前記複素指数変調フィルタバンクは、デジタルイコライジングシステムにおいてフィルタとして使用されることを特徴とする請求項３から８に記載の方法。

【請求項12】

前記複素指数変調フィルタバンクは、適応型フィルタリング・システムにおいて使用されることを特徴とする請求項３から８に記載の方法。

【請求項13】

前記複素指数変調フィルタバンクは、マルチバンド・コンパンダ・システムにおいて使用されることを特徴とする請求項３から８に記載の方法。

【請求項14】

デジタル・フィルタバンクから取得される係数又はサブバンド信号のエネルギー測度を推定するための装置であって、

完全再構成特性を有するようには制約されない、フィルタ次数Ｎを有する低域通過対称プロトタイプ・フィルタp_0（ｎ）を最適化するための手段と、

Ｍチャネルの分析フィルタバンクであって、該フィルタバンクが、

$$h_k(n) = p_0(n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\}$$

で表されるフィルタ係数を有し、ｎ＝０，１，．．．．，Ｎでありかつｋ＝０，１，．．．，Ｍ－１であるものを前記プロトタイプ・フィルタの複素指数変調によって構築するための手段と、

前記フィルタバンクを介して実数値の時間領域信号をフィルタリングするための手段と、

前記フィルタリングにより取得された複素数値のサブバンド信号の絶対値二乗を計算するための手段を特徴とする装置。

【請求項15】

デジタル・フィルタバンクから取得される係数又はサブバンド信号の修正によって発生するエイリアシングを低減するための装置であって、

完全再構成特性を有するようには制約されない、フィルタ次数Ｎを有する低域通過対称プロトタイプ・フィルタp_0（ｎ）を最適化するための手段と、

Ｍチャネルのフィルタバンクであって、該フィルタバンクが、

$$h_k(n) = f_k(n) = p_0(n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\}$$

で表わされる分析及び合成フィルタ係数を有し、ｎ＝０，１，．．．，Ｎでありかつｋ＝０，１，．．．，Ｍ－１であるものを前記プロトタイプ・フィルタの複素指数変調によって構築するための手段と、

前記フィルタバンクの分析部分を介して実数値の時間領域信号をフィルタリングするための手段と、

前記フィルタリングにより取得された複素数値のサブバンド信号を修正するための手段と、

前記修正された複素数値のサブバンド信号を前記フィルタバンクの合成部分を介してフィルタリングするための手段と、

複素数値の時間領域出力信号であって、該出力信号が前記合成フィルタリングにより取得された信号の和であるものの実数部を取り出すための手段を特徴とする装置。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、サブサンプリングされるデジタル・フィルタバンクに関し、かつデジタル・フィルタバンクのスペクトル係数又はサブバンド信号の修正、例えば量子化又は減衰、から発生する影響を実質的に低減させる新しい方法及び装置を提供する。本発明は、デジタル・イコライザ［Ａ．Ｊ．Ｓ．Ｆｅｒｒｅｉｒａ，Ｊ．Ｍ．Ｎ．Ｖｉｅｒａ著「効率的な２０バンド・デジタル音声イコライザ」ＡＥＳプレプリント、第９８回大会、１９９５年２月２５－２８日、米国ニューヨーク州パリス］、適応型フィルタ［Ａ．Ｇｉｌｌｏｉｒｅ，Ｍ．Ｖｅｔｔｅｒｌｉ著「クリティカル・サンプリングによるサブバンドにおける適応型フィルタリング：分析、実験及び音響キャンセリングへの適用」信号処理に関するＩＥ

ＥＥ会誌第４０巻第８号、１９９２年８月］、マルチバンド・コンパンダ及び高周波数再構成（high frequency reconstruction）（ＨＦＲ）を使用する音声コーディング・システムに適用可能である。そこでは、デジタル・フィルタバンクは、スペクトル・バンド複製（ＳＢＲ）システム［ＷＯ　９８／５７４３６］のように、スペクトル包絡線の適応調整に使用されている。

【背景技術】

【０００２】

デジタル・フィルタバンクは、２つ又はそれ以上の並列するデジタル・フィルタの集合体である。分析フィルタバンクは、受信信号をサブバンド信号（又はスペクトル係数）と称される多数の別々の信号に分割する。該フィルタバンクは、単位時間当たりのサブバンド・サンプルの合計数が入力信号のそれと同じである場合にクリティカル・サンプリングを実行される（又は最大限にデシメート（decimate）される）。合成フィルタバンクは、これらのサブバンド信号を結合して出力信号にする。クリティカル・サンプリングを実行されるフィルタバンクの一般的なタイプは、余弦変調フィルタバンクである。余弦変調システムにおけるフィルタは、所謂プロトタイプ・フィルタである低域通過フィルタの余弦変調によって取得される。余弦変調されたバンクは極めて効果的な実施を提供し、自然音声コーディック［Ｋ．Ｂｒａｎｄｅｎｂｕｒｇ著「知覚コーディング入門」ＡＥＳ、１９９６年度デジタル音声ビットレート低減に関する論文集、１９９６年］に使用されることが多い。しかしながら、例えば等化利得曲線の印加又はサンプルの量子化によってサブバンド・サンプル又はスペクトル係数を変更する試みはどれも、出力信号に重度のエイリアシング・アーティファクトをもたらす。

【発明の開示】

【０００３】

本発明は、正弦変調された虚数部で余弦変調フィルタバンクを拡張して複素指数変調フィルタバンクを形成することにより、サブバンド信号の修正から発生する障害は大幅に低減可能であることを証明する。正弦拡張は、余弦変調されたフィルタバンクに存在する主要なエイリアス項を排除する。さらに本発明は、プロトタイプ・フィルタを最適化するためのエイリアス項最小化（ＡＴＭ）と呼ばれる方法を提示する。複素指数変調は、フィルタバンクの実数部、即ち基礎的な余弦変調フィルタバンクから取得される信号の解析信号として解釈されることが可能な複素数値化されたサブバンド信号を生成する。この機能は、サブバンド信号の瞬間エネルギーの固有測度を提示する。

【０００４】

本発明による複素指数変調フィルタバンクを動作させるための主要ステップは、次の通りである。

１．所望されるエイリアシング阻止及び通過帯域の平坦度に合わせて最適化される、カットオフ周波数π／２Ｍを有する対称性低域通過プロトタイプ・フィルタの設計、

２．最適化されたプロトタイプ・フィルタの複素指数変調によるＭチャネル・フィルタバンクの構成、

３．フィルタバンクの分析部分を介する、実数値時間領域信号のフィルタリング、

４．所望される、おそらくは時間的に変動するイコライザ設定による複素数値化されたサブバンド信号の修正、

５．フィルタバンクの合成部分を介する、修正された複素値サブバンド・サンプルのフィルタリング、

６．フィルタバンクの合成部分から取得される複素数値化された時間領域出力信号の実数部の計算。

【０００５】

本発明の最も魅力あるアプリケーションは、ＨＦＲシステムにおいて使用される様々なタイプのデジタル・イコライザ、適応型フィルタ、マルチバンド・コンパンダ及び適応型包絡線調整フィルタバンクの改良である。

【発明を実施するための最良の形態】

【０００６】
次に、添付の図面を参照して、本発明を発明の範囲及び精神を限定することなく例示的に説明する。
【０００７】
本発明は、本特許において明示的に記述されているもの以外のデジタル・フィルタバンクを組み込む実施範囲にも適用可能であることは理解されなければならない。
【０００８】
デジタル・フィルタバンク
デジタル・フィルタバンクは、共通の入力又は共通の出力を共用する２つ又はそれ以上の並列するデジタル・フィルタの集合体である［Ｐ．Ｐ．Ｖａｉｄｙａｎａｔｈａｎ著「マルチレート・システムとフィルタバンク」Ｐｒｅｎｔｉｃｅ　Ｈａｌｌ：Ｅｎｇｌｅｗｏｏｄ　Ｃｌｉｆｆｓ，ＮＪ，１９９３年］。共通の入力を共用する場合、フィルタバンクは分析バンクと呼ばれる。分析バンクは受信する信号を、サブバンド信号と呼ばれるＭ個の別々の信号に分割する。分析フィルタは、H_k（ｚ）で示される。但し、ｋ＝０．．．Ｍ－１である。サブバンド信号がM毎にデシメートされるとき、フィルタバンクはクリティカル・サンプリングを実行される（又は最大限にデシメートされる）。したがって、単位時間当たりの合計サブバンド・サンプル数は、入力信号に関する単位時間当たりのサンプル数と同じである。合成バンクは、これらのサブバンド信号を結合して１つの共通の出力信号にする。合成フィルタはF_k（ｚ）で示され、ｋ＝０．．．Ｍ－１である。図１は、Ｍチャネル（サブバンド）を有する最大限にデシメートされたフィルタバンクを示している。分析部分１０１は、入力信号Ｘ（ｚ）から信号V_k（ｚ）を生成する。信号V_k（ｚ）は、送信され、記憶され、又は修正される信号を構成する。合成部分１０２は、信号V_k（ｚ）を再び結合して出力信号

$$\hat{X}(z)$$

にする。
【０００９】
原信号Ｘ（ｚ）の近似値

$$\hat{X}(z)$$

を求めるためのV_k（ｚ）の再結合は、幾つかの誤差を伴う。その１つは、サブバンドのデシメーション及び補間に起因するエイリアシングである。その他の誤差としては、位相及び振幅の歪みがある。
【００１０】
図１の表記法に従えば、分析フィルタH_k（ｚ）１０３の出力は、

$$X_k(z)=H_k(z)X(z) \tag{1}$$

である。但し、ｋ＝０．．．Ｍ－１である。デシメータ１０４が与える出力は、

$$V_k(z)=\frac{1}{M}\sum_{l=0}^{M-1}X_k(z^{1/M}W^l)=\frac{1}{M}\sum_{l=0}^{M-1}H_k(z^{1/M}W^l)X(z^{1/M}W^l) \tag{2}$$

であり、$W=e^{-i2\pi/M}$である。インターポレータ１０５の出力は、

$$U_k(z)=V_k(z^M)=\frac{1}{M}\sum_{l=0}^{M-1}H_k(zW^l)X(zW^l) \tag{3}$$

で与えられ、合成フィルタ１０６から求められる信号の和は、

$$\hat{X}(z)=\sum_{k=0}^{M-1}F_k(z)U_k(z)=\sum_{k=0}^{M-1}F_k(z)\frac{1}{M}\sum_{l=0}^{M-1}H_k(zW^l)X(zW^l)=$$
$$=\frac{1}{M}\sum_{l=0}^{M-1}X(zW^l)\sum_{k=0}^{M-1}H_k(zW^l)F_k(z)=\frac{1}{M}\sum_{l=0}^{M-1}X(zW^l)A_l(z) \quad (4)$$

と書き表すことができる。但し、

$$A_l(z)=\sum_{k=0}^{M-1}H_k(zW^l)F_k(z) \quad (5)$$

は、１番目のエイリアス項Ｘ（ｚW^{l}）の利得である。式（４）は、

$$\hat{X}(z)=\frac{1}{M}\left\{X(z)A_0(z)+\sum_{l=1}^{M-1}X(zW^l)A_l(z)\right\} \quad (6)$$

と書き直すことができる。右辺（ＲＨＳ）にある最後の和は、望まれないエイリアス項の和を構成する。全てのエイリアシングを除去する、即ちH_k（ｚ）及びF_k（ｚ）の適正な選定によってこの和をゼロにすると、

$$\hat{X}(z)=\frac{1}{M}X(z)A_0(z)=\frac{1}{M}X(z)\sum_{k=0}^{M-1}H_k(z)F_k(z)=X(z)T(z) \quad (7)$$

になる。但し、

$$T(z)=\frac{1}{M}\sum_{k=0}^{M-1}H_k(z)F_k(z) \quad (8)$$

は全体の伝達関数又は歪み関数である。合成フィルタF_k（ｚ）を、

$$F_k(z)=z^{-N}\tilde{H}_k(z^{-1}) \quad (9)$$

但しＮは分析フィルタの次数、となるように選定すると、次のような伝達関数が得られる。

$$T(z)=\frac{1}{M}\sum_{k=0}^{M-1}H_k(z)F_k(z)=\frac{z^{-N}}{M}\sum_{k=0}^{M-1}H_k(z)\tilde{H}_k(z) \quad (10)$$

表記法

$\tilde{H}(z)$

は、時間反転および複素共役系列h_k（ｎ）のＺ変換である。式（１０）を単位円上で評価すると、

$$T(e^{jw}) = \frac{e^{-jwN}}{M} \sum_{k=0}^{M-1} H_k(e^{jw}) H_k^*(e^{-jw}) = \frac{e^{-jwN}}{M} \sum_{k=0}^{M-1} \left| H_k(e^{jw}) \right|^2 \qquad (11)$$

になる。式（１１）は、Ｔ（ｚ）が線形位相を有し、よって位相歪みを全く生じないことを示す。さらに、右辺の最終的な和が定数であれば、振幅の歪みは存在しない。この場合、全体の伝達関数は一定の換算係数ｃを有する単一遅延であり、即ち、

$$T(z) = cz^{-N} \qquad (12)$$

である。これを式（７）に代入すると、

$$\hat{X}(z) = cz^{-N} X(z) \qquad (13)$$

になる。式（１３）を満足させるフィルタ・タイプは、完全再構成（ＰＲ）特性を有するものとされる。

【００１１】

余弦変調フィルタバンク

余弦変調フィルタバンクにおいては、分析フィルタh_k（ｎ）は、

$$h_k(n) = 2p_0(n) \cos\left\{ \frac{\pi}{2M}(2k+1)\left(n - \frac{N}{2} - \frac{M}{2}\right) \right\} \qquad (14)$$

のように、対称性低域通過プロトタイプ・フィルタp_0（ｎ）の余弦変調バージョンである。但し、Ｍはチャネルの数であり、ｋ＝０．．．Ｍ－１であり、Ｎはプロトタイプ・フィルタ次数であり、ｎ＝０．．．Ｎである。実数値プロトタイプ・フィルタ係数の和は、

$$\sum_{n=0}^{N} P_0^2(n) = 1 \qquad (15)$$

のように１と仮定する。同じ表記法によれば、合成フィルタは、

$$f_k(n) = 2p_0(N-n) \cos\left\{ \frac{\pi}{2M}(2k+1)\left(n - \frac{N}{2} + \frac{M}{2}\right) \right\} \qquad (16)$$

で与えられる。分析フィルタバンクは、実数値入力信号の実数値サブバンド・サンプルを生成する。サブバンド・サンプルはＭ毎にダウンサンプリングされ、これによりシステムはクリティカル・サンプリングを実行される。プロトタイプ・フィルタの選定に依存して、フィルタバンクは、ほぼ完全な再構成システム、所謂擬似ＱＭＦバンク［米国特許第５，４３６，９４０号］又は完全再構成（ＰＲ）システムを構成する可能性がある。ＰＲシステムの一例は、変調重複変換（ＭＬＴ）である［Ｈ．Ｓ．Ｍａｌｖａｒ著「効率的変換／サブバンド・コーディングのための重複変換」ＩＥＥＥ、ＡＳＳＰ会誌第３８巻第６号、１９９０年］。余弦変調固有の特性の１つは、あらゆるフィルタが、一方は正の周波数範囲にありかつ他方は負の周波数範囲内の対応する通過帯域にある２つの通過帯域を有することにある。

【００１２】

式（５）は、行列形式で、

$$\mathbf{a} = \mathbf{H}\,\mathbf{f} \tag{17}$$

のように、あるいは明示的に、

$$\begin{bmatrix} A_0(z) \\ A_1(z) \\ \vdots \\ A_{M-1}(z) \end{bmatrix} = \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(Wz) & H_1(Wz) & & H_{M-1}(Wz) \\ \vdots & & \ddots & \vdots \\ H_0(W^{M-1}z) & H_1(W^{M-1}z) & \cdots & H_{M-1}(W^{M-1}z) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} \tag{18}$$

と書き表すことができる。行列

$\mathbf{H}$

は、エイリアス成分（ＡＣ）行列と呼ばれる。この式をより厳密に検討するために、
は、
$\mathbf{f}$

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ . \\ . \\ . \\ F_{M-1}(z) \end{bmatrix} = \begin{bmatrix} F_0(z) & 0 & . & . & . & 0 \\ 0 & F_1(z) & & & & . \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & 0 \\ 0 & . & . & . & 0 & F_{M-1}(z) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ . \\ . \\ . \\ 1 \end{bmatrix} \tag{19}$$

と書き表すことが可能であり、もしくはコンパクトな形式で、

$$\mathbf{f} = \mathbf{F}\,\mathbf{e} \tag{20}$$

と書き表すことができる。式（１７）に式（２０）を代入すると、エイリアス利得を

$\mathbf{a} = \mathbf{HF}\,\mathbf{e}$

と書き表すことができる。但し、積、

$$\mathbf{HF} = \mathbf{U} \tag{21}$$

はＭ×Ｍの行列であり、以後これを複合エイリアス成分行列と称す。

【００１３】

余弦変調システムの場合、複合エイリアス成分行列における支配項は第１の行と４つの対角項上にある。三次元プロットである図２は、この行列における成分の大きさを示している。第１の行は伝達関数である式（８）の項を保有し、４つの対角項は主として主要なエイリアス項、即ちフィルタとそれらの最近傍フィルタとのオーバーラップに起因するエイリアシングで構成される。主要なエイリアス項は、正の通過帯域の周波数変調バージョンを有するフィルタの負の通過帯域間、又は逆に負の通過帯域の周波数変調バージョンを有するフィルタの正の通過帯域間の何れかの周波数オーバーラップから発現することは容易に分かる。複合エイリアス成分行列における複数行の項を合計する、即ちエイリアス利得を計算すると、主要なエイリアス項が相殺される結果となる。エイリアシングは２つ１組

で相殺され、第１の主要エイリアス項が同じ行内の第２のエイリアス項によって相殺される。主要エイリアス項には、他のより小さいエイリアス項も重畳される。プロトタイプ・フィルタ特性が、フィルタの遷移および遮断帯域がその変調バージョンとの実質的オーバーラップを有する類のものであれば、これらのエイリアス項は大きなものになる。一例として、第２かつ最終の行は、フィルタとそれらの最近傍の変調バージョンとのオーバーラップによって導入されたエイリアス項で構成される。ＰＲシステムの場合、これらのより小さいエイリアス項もまた、エイリアス利得項を合計すると完全に相殺される。しかしながらＱＭＦシステムでは、これらの項は残存する。

【００１４】

複素指数変調フィルタバンク

余弦変調を本発明による複素指数変調にまで拡張すると、前と同じ表記法を使用して、r/>

$$h_k(n) = p_0(n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \qquad (22)$$

として表わされる分析フィルタh_k（ｎ）が生じる。これは、虚数部を実数値のフィルタバンクに加算したものとして捉えることができる。但し、虚数部は同じプロトタイプ・フィルタの正弦変調バージョンより成る。実数値の入力信号を考慮すると、フィルタバンクからの出力は、実数部及び虚数部が互いのヒルベルト変換である一組のサブバンド信号として解釈することができる。したがって、得られるサブバンドは、余弦変調フィルタバンクから取得される実数値出力の解析信号である。故に、複素数値表現に起因して、サブバンド信号は２倍にオーバーサンプリングされる。

【００１５】

合成フィルタも同様に、

$$f_k(n) = p_0(N-n)\ \exp\left\{ i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \qquad (23)$$

のように拡張される。式（２２）及び（２３）は、合成バンクからの出力が複素値であることを意味している。行列表記法を使用して、

$\mathbf{C_a}$

は式（１４）の分析フィルタを有する行列であり、

は、

$\mathbf{S_a}$

$$h_k(n) = p_0(n)\ \sin\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \qquad (24)$$

で表されるフィルタを有する行列とすると、式（２２）のフィルタは

$\mathbf{C_a} + \mathrm{j}\,\mathbf{S_a}$

として求められる。これらの行列では、ｋは行の指標であり、ｎは列の指標である。同様に、行列

$\mathbf{C_s}$
は式（１６）の合成フィルタを有し、
は、
$\mathbf{S_s}$

$$f_k(n) = p_0(N-n)\ \sin\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \qquad (25)$$

で表されるフィルタを有する行列である。したがって、式（２３）は

$\mathbf{C_s + j\,S_s}$
と書き表すことができる。但し、ｋは列の指標であり、ｎは行の指標である。入力信号をｘで示すと、

$$\mathbf{y = (C_s + j\,S_s)\,(C_a + j\,S_a)\,x = (C_sC_a - S_sS_a)\,x + j\,(C_sS_a + S_sC_a)\,x} \qquad (26)$$

から、出力信号ｙが求められる。
式（２６）から分かるように、実数部は、一般的な余弦変調フィルタバンクからの出力及び正弦変調フィルタバンクからの出力である２項より成る。余弦変調フィルタバンクがＰＲ特性を保有すれば、その正弦変調バージョンもまたＰＲシステムを構成することは、容易に検証される。したがって、出力の実数部を利用することにより、複素指数変調システムは対応する余弦変調バージョンと同じ再構成精度をもたらす。
【００１６】
複素指数変調システムは、複素数値の入力信号をも処理するように拡張されることが可能である。チャネル数を２Ｍに拡大する、即ち負の周波数用のフィルタを追加しかつ出力信号の虚数部を維持することにより、複素数値信号の擬似ＱＭＦ又はＰＲシステムが取得される。
【００１７】
式（２１）の複合エイリアス成分行列を精査すると、複素指数変調フィルタバンクでは主要なエイリアス対角項が消える。複素指数変調フィルタバンクはあらゆるフィルタが通過帯域を１つしか保有しないことから、これは容易に理解される。言い替えれば、このフィルタバンクは主要なエイリアス項を保有せず、上述の２つ１組のエイリアシングの相殺のシステムは不要である。複合エイリアス成分行列は、第１の行にのみ支配項を有する。図３は、得られる行列の成分の大きさを示している。プロトタイプ・フィルタの特性に依存して、１乃至（Ｍ－１）行上の項は多かれ少なかれ減衰される。複素指数変調バージョンにおいて、主要なエイリアス項が存在していないことは、余弦（又は正弦）変調フィルタバンクに基づくエイリアシング相殺の制約を不要にする。したがって、分析及び合成フィルタは共に、対称プロトタイプ・フィルタではp_0（ｎ）＝p_0（Ｎ－ｎ）であるため、r/>

$$h_k(n) = f_k(n) = p_0(n)\ \exp\left\{i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\} \qquad (27)$$

から求めることができる。先にも述べたように、Ｍはチャネル数であり、ｋ＝０．．．Ｍ－１であり、Ｎはプロトタイプ・フィルタ次数であり、ｎ＝０．．．Ｎである。
【００１８】
式（４）を参照すると、出力信号

$\hat{x}(n)$

の実数部のＺ変換は、

$$Z\{\mathrm{Re}(\hat{x}(n))\} = \hat{X}_R(z) = \frac{\hat{X}(z) + \hat{X}*(z)}{2} \tag{28}$$

である。表記法

$\hat{X}*(z)$

は、複素共役系列

$\hat{x}(n)$

のＺ変換である。式（４）から、出力信号の実数部の当該変換は、

$$2M\hat{X}_R(z) = \sum_{l=0}^{M-1}\left(X(zW^l)A_l(z) + X(zW^{-l})A_l*(z)\right) \tag{29}$$

であることになる。但し、入力信号ｘ（ｎ）は実数値であることが利用されている。式（２９）は、

$$\begin{aligned} 2M\hat{X}_R(z) = {} & X(z)\{A_0(z) + A_0*(z)\} + \\ & + X(-z)\{A_{M/2}(z) + A_{M/2}*(z)\} + \\ & + \sum_{l=1}^{M/2-1} X(zW^l)\{A_l(z) + A_{M-l}*(z)\} + \\ & + \sum_{l=1}^{M/2-1} X(zW^{-l})\{A_{M-l}(z) + A_l*(z)\} \end{aligned} \tag{30}$$

と書き表すことができる。式（３０）を検分しかつ式（２８）の変換を想起すると、ＰＲシステムの場合、a_0（ｎ）の実数部は必ずディラック・パルスになることは明白である。さらに、$a_{M/2}$（ｎ）の実数部は必ずゼロであり、ｌ＝１．．．Ｍ／２－１に関するエイリアス利得は、必ず、

$$A_{M-l}(z) = -A_l*(z) \tag{31}$$

を満足させる。擬似ＱＭＦシステムでは、式（３１）は近似的でしか成立しない。さらに、a_0（ｎ）の実数部は厳密にはディラック・パルスではなく、$a_{M/2}$（ｎ）の実数部も厳密にはゼロではない。

【００１９】

サブバンド信号の修正

余弦変調フィルタバンクにおけるチャネル利得を変更する、即ち分析／合成システムをイコライザとして使用すると、主要エイリアス項に起因して甚だしい歪みが生じる。例えば、８チャネルのフィルタバンクを帯域通過応答用に調整しようとして、第２及び第３のチャネル以外の全てのサブバンド信号がゼロに設定されるものとする。すると、式（２１）の複合エイリアス成分行列は、図４が示すように第２及び第３の列を除く全てのエレメントがゼロである８×８行列になる。図が示すように、７つの重大なエイリアス項が残存す

る。第３行及び５行のエイリアシングは、これらの行における主要エイリアス項が同じ利得を有する、即ち２つ１組の相殺が意図的に働くことから相殺される。しかしながら第２行、４行及び６行には、その対応するエイリアス項が利得ゼロを有することから、単一のエイリアス項が存在する。故に、エイリアス相殺は意図された通りには働かず、出力信号においてエイリアシングが実在することになる。

【００２０】

この例から、複素指数変調フィルタバンクをイコライザとして使用すると、実質的改良が達成されることは明白である。図４が示す８チャネル・システムは、次数１２８のプロトタイプ・フィルタを有する。上述の例では、エイリアシングの全減衰量は１６ｄＢでしかないが、複素指数変調に移行すると、エイリアシングの減衰量は９５ｄＢになる。存在しない主要エイリアス項に起因して、結果として得られるエイリアシングは、フィルタとその変調バージョンとのオーバーラップから生じるエイリアス項の抑圧にのみ依存する。したがって、エイリアス利得項を最大限に抑圧するようにプロトタイプ・フィルタを設計することが極めて重要である。単位円上で評価された式（３０）の右辺の第１項は、

$$e_t = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\left|A_0(e^{j\omega})+A_0^{*}(e^{-j\omega})\right|}{2M}-1\right)^2 d\omega \tag{32}$$

で表わされる伝達関数の誤差エネルギーe_tをもたらす。全エイリアシングのエネルギーe_aは、

$$\begin{aligned} e_a = {} & \frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega})+A_{M/2}{}^{*}(e^{-j\omega})\right|^2 d\omega + \\ & + \frac{1}{8\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left\{\left|A_l(e^{j\omega})+A_{M-l}{}^{*}(e^{-j\omega})\right|^2 + \right. \\ & \left. + \left|A_{M-l}(e^{j\omega})+A_l{}^{*}(e^{-j\omega})\right|^2\right\} d\omega \end{aligned} \tag{33}$$

のように、単位円上で式（３０）の右辺の残り全ての項を評価して計算することができる。対称性、式（９）及び、

$$P_0(z) = z^{-N}\tilde{P}_0(z) \tag{34}$$

という事実に起因して、式（３３）における和の中括弧内の複数項は等しい。したがって、エイリアシングの全エネルギーは（Ｍ／２－１）項を有し、

$$\begin{aligned} e_a = {} & \frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega})+A_{M/2}{}^{*}(e^{-j\omega})\right|^2 d\omega + \\ & + \frac{1}{4\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left|A_l(e^{j\omega})+A_{M-l}{}^{*}(e^{-j\omega})\right|^2 d\omega \end{aligned} \tag{35}$$

となる。エイリアス利得項の最小化は、プロトタイプ・フィルタの最適化によって実行される。好適には、これは、例えばダウンヒル・シンプレックス法［Ｗ．Ｈ．Ｐｒｅｓｓ、Ｓ．Ａ．Ｔｅｕｋｏｌｓｋｙ、Ｗ．Ｔ．Ｖｅｔｔｅｒｌｉｎｇ、Ｂ．Ｐ．Ｆｌａｎｎｅｒ

ｙ共著「Ｎｕｍｅｒｉｃａｌ　Ｒｅｃｉｐｅｓ　ｉｎ　Ｃ，Ｔｈｅ　Ａｒｔ　ｏｆ　Ｓｃｉｅｎｔｉｆｉｃ　Ｃｏｍｐｕｔｉｎｇ：第２版」ケンブリッジ大学出版、ＮＹ、１９９２年］のような標準的な非線形最適化アルゴリズムを使用して合成目的関数を最小化することにより達成される。本発明によるプロトタイプ・フィルタのエイリアス項最小化（ＡＴＭ）の場合、目的関数は、

$$\varepsilon_{tot}(\alpha)=\alpha\varepsilon_t+(1-\alpha)\varepsilon_a \tag{36}$$

のようなものである。最適化の間は、ε_aの計算に際して、即ち分析及び合成フィルタが、

$$H_k^{(eq)}(z)=g_kH_k(z)\quad\text{及び}\quad F_k^{(eq)}(z)=g_kF_k(z) \tag{37}$$

のように利得ファクタg_kで乗算される際にはフィルタバンクにランダムな等化曲線が適用され、エイリアス利得項A_1（ｚ）をｌ＝１．．．Ｍ－１に関して計算する際には、結果として得られるフィルタ$H_k^{(eq)}$及び$F_k^{(eq)}$、但しｋ＝０．．．Ｍ－１、が使用される。

【００２１】

図５では、５つの異なる複素指数変調システムのエイリアス利得が比較されている。５つのうちの４つは８チャネル・システムであり、１つは６４チャネル・システムである。これらのシステムは全て、１２８のプロトタイプ・フィルタ長を有する。星印のついた点線軌跡及び実線軌跡は２つの擬似ＱＭＦシステムのエイリアス成分を示し、一方はエイリアス項が最小化されている。破線軌跡及び一点破線軌跡は、２つの８チャネル完全再構成システムの成分に関するものであり、この場合も一方のシステムはエイリアス項が最小化されている。実線軌跡は、複素指数変調重複変換（ＭＬＴ）のエイリアス成分である。全てのシステムは、上述の例による帯域通過応答に合わせて調整され、結果は表１に示されている。全エイリアシングの阻止は、式（３３）の逆数として計算される。通過帯域の平坦度は、帯域通過応答に合わせて調整された積分間隔をもって、式（３２）の逆数として計算される。

【００２２】

【表１】

システム	全エイリアシングの阻止	通過帯域の平坦度
8-ch 擬似 QMF ATM N=128	114.7 dB	98.1 dB
8-ch 擬似 QMF N=128	95.4 dB	87.6 dB
8-ch PR ATM N=127	77.3 dB	132.7 dB
8-ch PR N=127	55.0 dB	93.6 dB
64-ch MLT N=127	38.5 dB	87.1 dB

【００２３】

表１における数値から分かるように、６４チャネルのＭＬＴから８チャネルのＰＲシステムに移行すると実質的改良が達成される。ＭＬＴは完全再構成システムであり、ポリフェーズ成分当たり（Ｎ＋１）／２Ｍ＝１個の係数しか保有しない。８チャネルＰＲシステムの係数の数は、１２８／１６＝８である。これは、より高度の遮断帯域減衰及びより高度のエイリアス項の阻止を有するフィルタを可能にする。さらに、ＰＲシステムのエイリアス項最小化はエイリアシングを阻止し、かつ通過帯域の平坦化を大幅に向上させることが分かる。擬似ＱＭＦシステムとＰＲシステムとを比較すると、通過帯域の平坦度はほぼ維持されるが、エイリアシングの阻止は明らかに４０ｄＢほど向上する。エイリアス項を最

小化すれば、エイリアス阻止の約２０ｄＢの改善及び通過帯域平坦度の１０ｄＢの向上が達成される。したがって、完全再構成の制約は、等化システムにおいて使用されるフィルタバンクを限定することは明白である。実際のデジタル実現は全て、数値表現では限定された分解能を有することから、擬似ＱＭＦシステムは、常に十分な再構成精度を見込んで設計することが可能である。擬似ＱＭＦ及びＰＲシステムに関しては何れも、最適なシステムは遮断帯域の大幅な阻止を有するプロトタイプ・フィルタを基礎として構築されることが明白である。このため、プロトタイプ・フィルタは、ＭＬＴで使用される窓より長い相対長さを有するものを使用しなければならない。

【００２４】

複素指数変調システムの卓越した優位点は、サブバンド信号が余弦変調フィルタバンクから取得される実数値のサブバンド信号の解析信号を構成することから、瞬間エネルギーが容易に計算されることにある。これは、例えば適応型フィルタ、自動利得制御（ＡＧＣ）、マルチバンド・コンパンダ及びフィルタバンクを使用してスペクトル包絡線が調整されるスペクトル・バンド複製システム（ＳＢＲ）において極めて価値のある特性である。サブバンドｋにおける平均エネルギーは、

$$E_k(m) = \sum_{n=-L+1}^{L-1} \left| v_k(mL+n) \right|^2 w(n) \qquad (38)$$

で計算される。但し、v_k（ｎ）はチャネルｋのサブバンド・サンプルであり、ｗ（ｎ）はｎ＝０近傍を中心とする長さ（２Ｌ－１）の窓である。次にこの測度は、適応化又は利得計算アルゴリズム用の入力パラメータとして使用される。

【００２５】

実用的な実施例

標準型のＰＣ又はＤＳＰを使用すれば、複素指数変調フィルタバンクのリアルタイム処理が可能である。フィルタバンクはまた、カスタム・チップにハードコード化される場合もある。図６は、複素指数変調フィルタバンク・システムの分析部分の効果的な実施のための構成を示している。まず、アナログ入力信号がＡ／Ｄ変換器６０１に供給される。デジタルの時間領域信号は、一度に２Ｍ個のサンプルを保持しかつＭ個のサンプルをシフトするシフト・レジスタ６０２に供給される。シフト・レジスタからの信号は、次にプロトタイプ・フィルタ６０３のポリフェーズ係数を介してフィルタリングされる。フィルタリングされた信号は続いて６０４で結合され、ＤＣＴ－ＩＶ変換器６０５及びＤＳＴ－ＩＶ変換器６０６により並列して変換される。余弦変換器及び正弦変換器からの出力は、各々サブバンド・サンプルの実数部及び虚数部を構成する。サブバンド・サンプルの利得は、カレント・スペクトル包絡線調整器設定６０７により修正される。

【００２６】

図７は、複素指数変調システムの合成部分の効果的な実施を示している。サブバンド・サンプルは、まず複素数値の回転因子７０１で乗算され、実数部はＤＣＴ－ＩＶ変換器７０２で変調され、虚数部はＤＳＴ－ＩＶ変換器７０３で変調される。両変換器からの出力は７０４で結合され、プロトタイプ・フィルタ７０５のポリフェーズ成分を介して供給される。シフト・レジスタ７０６からは、時間領域出力信号が取得される。最後に、デジタル出力信号が元のアナログ波形７０７に変換される。

【００２７】

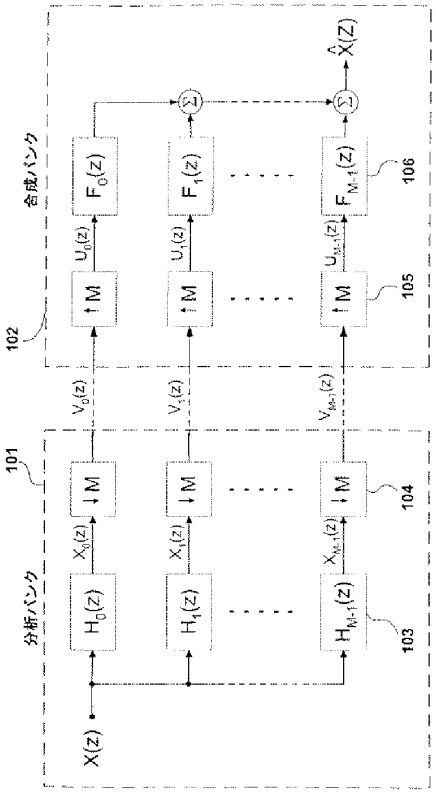
上述の実施形態は、本発明による複素指数変調フィルタバンク・システムの原理を単に例示したものである。当業者には、本明細書に記述された装置及び詳細事項の修正及び変更が明らかとなることは理解される。したがって本発明は、本明細書における実施形態の記述及び説明によって提示された特定の詳細事項ではなく、添付の請求の範囲に記載された範囲によってのみ限定されるべきものである。

【図面の簡単な説明】

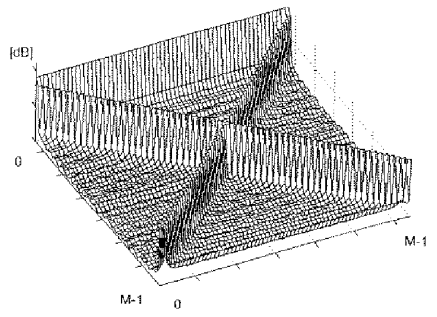
【００２８】

【図１】デジタル・フィルタバンクの分析セクション及び合成セクションを示している。
【図２】余弦変調フィルタバンクの複合エイリアス成分行列における大きさを示している。
【図３】複素指数変調フィルタバンクの複合エイリアス成分行列における大きさを示している。
【図４】帯域通過フィルタ応答用に調整された余弦変調フィルタバンクにおける希望される項及び主要なエイリアス項を示している。
【図５】異なる複素指数変調フィルタバンクの実現のためのエイリアス利得項の減衰を示している。
【図６】本発明による複素指数変調フィルタバンク・システムの分析部分を示している。
【図７】本発明による複素指数変調フィルタバンク・システムの合成部分を示している。

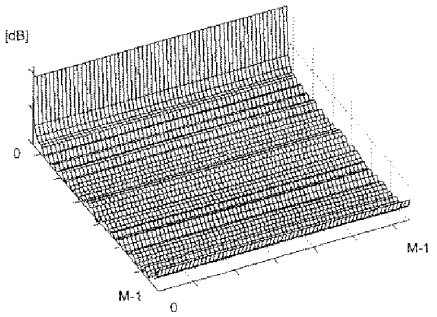
【図1】



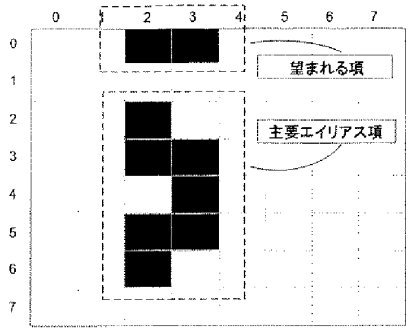
【図2】



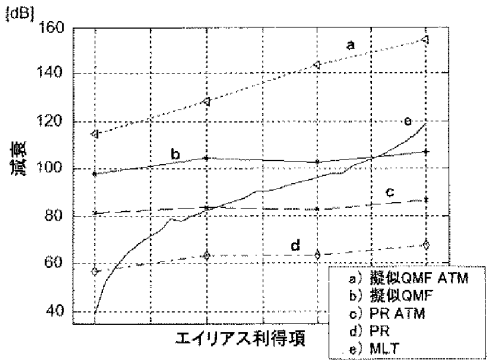
【図3】



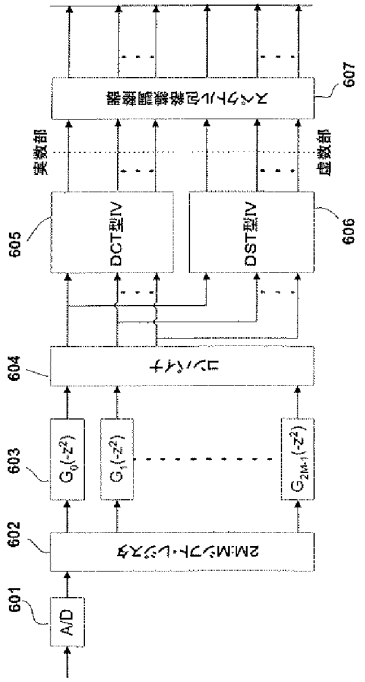
【図4】


望まれる項
主要エイリアス項

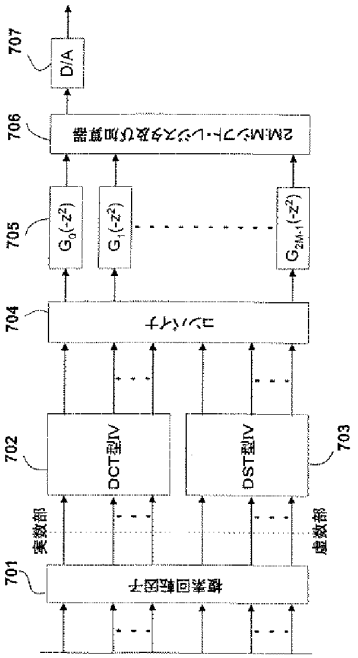
【図5】


[dB]
減衰
エイリアス利得項
a) 擬似QMF ATM
b) 擬似QMF
c) PR ATM
d) PR
e) MLT

【図6】


A/D
2M:ポリフェーズ
コンバイナ
DCT型IV
DST型IV
実数部
虚数部

【図7】


D/A
2M:ポリフェーズ
コンバイナ
DCT型IV
DST型IV
実数部
虚数部

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 October 2002 (10.10.2002)

PCT

(10) International Publication Number
WO 02/080362 A1

(51) International Patent Classification⁷: H03H 17/00

(21) International Application Number: PCT/SE02/00626

(22) International Filing Date: 28 March 2002 (28.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0101175-8 2 April 2001 (02.04.2001) SE

(71) Applicant (*for all designated States except US*): CODING TECHNOLOGIES SWEDEN AB [SE/SE]; Döbelngatan 64, S-113 52 Stockholm (SE).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): EKSTRAND, Per [SE/SE]; Södermannagatan 45, S-116 40 Stockholm (SE).

(74) Agents: ÖRTENBLAD, Bertil et al.; Noréns Patentbyrå AB, Box 10198, S-100 55 Stockholm (SE).

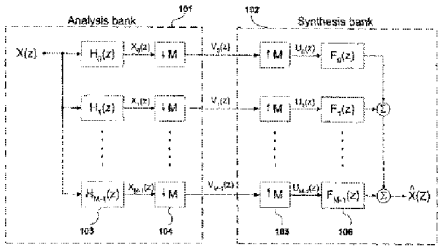
(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ALIASING REDUCTION USING COMPLEX-EXPONENTIAL MODULATED FILTERBANKS



(57) Abstract: The present invention proposes a new method and apparatus for the improvement of digital filterbanks, be a complex extension of cosine modulated digital filterbanks. The invention employs a complex-exponential modulation of a low-pass prototype filter and a new method for optimizing the characteristics of this filter. The invention substantially reduces artifacts due to aliasing emerging from independent modifications of subband signals, for example when using a filterbank as an spectral equalizer. The invention is preferably implemented in software, running on a standard PC or a digital signal processor (DSP), but can also be hardcoded on a custom chip. The invention offers essential improvements for various types of digital equalizers, adaptive filters, multiband companders and spectral envelope adjusting filterbanks used in high frequency reconstruction (HFR) systems.

WO 02/080362 A1

ALIASING REDUCTION USING COMPLEX-EXPONENTIAL MODULATED FILTERBANKS

TECHNICAL FIELD

The present invention relates to the area of subsampled digital filterbanks and provides a new method and apparatus for substantial reduction of impairments emerging from modifications, for example quantization or attenuation, of the spectral coefficients or subband signals of digital filterbanks. The invention is applicable to digital equalizers ["An Efficient 20 Band Digital Audio Equalizer" A. J. S. Ferreira, J. M. N. Viera, AES preprint, 98th Convention 1995 February 25-28 Paris, N.Y., USA], adaptive filters [Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments, and Application to Acoustic Echo Cancellation" A. Gilloire, M. Vetterli, IEEE Transactions on Signal Processing, vol. 40, no. 8, August, 1992], multiband companders, and to audio coding systems using high frequency reconstruction (HFR), where a digital filterbank is used for the adaptive adjustment of the spectral envelope, such as the Spectral Band Replication (SBR) system [WO 98/57436].

BACKGROUND OF THE INVENTION

A digital filter bank is a collection of two or more parallel digital filters. The analysis filter bank splits the incoming signal into a number of separate signals named subband signals (or spectral coefficients). The filter bank is critically sampled (or maximally decimated) when the total number of subband samples per unit time is the same as that for the input signal. The synthesis filter bank combines these subband signals into an output signal. A popular type of critically sampled filterbanks is the cosine modulated filterbank. The filters in the cosine modulated system are obtained by cosine modulation of a low-pass filter, a so-called prototype filter. The cosine modulated banks offer very effective implementations and are often used in natural audio codecs ["Introduction to Perceptual Coding" K. Brandenburg, AES, Collected Papers on Digital Audio Bitrate Reduction, 1996]. However, any attempt to alter the subband samples or spectral coefficients, e.g. by applying an equalizing gain curve or quantizing the samples, renders severe aliasing artifacts in the output signal.

SUMMARY OF THE INVENTION

The present invention shows that impairments emerging from modifications of the subband signals can be significantly reduced by extending a cosine modulated filterbank with an imaginary sine modulated part, forming a complex-exponential modulated filterbank. The sine extension eliminates the main alias terms present in the cosine modulated filterbank. Moreover, the invention presents a method, referred to as alias term minimization (ATM), for optimization of the prototype filter. The complex-exponential modulation creates complex-valued subband signals that can be interpreted as the analytic signals of the signals obtained from the real part of the filterbank, i.e. the underlying cosine modulated filterbank. This feature provides an inherent measure of the instantaneous energy of the subband signals.

The main steps for operation of the complex-exponential modulated filterbank according to the present invention are:

1. The design of a symmetric low-pass prototype filter with cutoff frequency $\pi/2M$, optimized for a desired aliasing rejection and passband flatness;
2. The construction of an M-channel filterbank by complex-exponential modulation of the optimized prototype filter;
3. The filtering of a real-valued time domain signal through the analysis part of the filterbank;
4. The modification of the complex-valued subband signals according to a desired, possibly time-varying, equalizer setting;
5. The filtering of the modified complex-valued subband samples through the synthesis part of the filterbank; and
6. The computation of the real part of the complex-valued time domain output signal obtained from the synthesis part of the filterbank.

The most attractive applications of the invention are improvement of various types of digital equalizers, adaptive filters, multiband companders and adaptive envelope adjusting filterbanks used in HFR systems.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described by way of illustrative examples, not limiting the scope nor the spirit of the invention, with reference to the accompanying drawings, in which:

Fig.1 illustrates the analysis and synthesis sections of a digital filterbank;

Fig.2 shows the magnitudes in a composite alias component matrix of a cosine modulated filterbank;

Fig.3 shows the magnitudes in a composite alias component matrix of a complex-exponential modulated filterbank;

Fig.4 illustrates wanted terms and main alias terms in a cosine modulated filterbank adjusted for a band-pass filter response;

Fig.5 shows the attenuation of alias gain terms for different realizations of complex-exponential modulated filterbanks;

Fig.6 illustrates the analysis part of a complex-exponential modulated filterbank system according to the present invention; and

Fig.7 illustrates the synthesis part of a complex-exponential modulated filterbank system according to the present invention.

DESCRIPTION OF PREFERRED EMBODIMENTS

It should be understood that the present invention is applicable on a range of implementations that incorporates digital filterbanks other than those explicitly mentioned in this patent.

Digital Filterbanks

A digital filter bank is a collection of two or more parallel digital filters that share a common input or a common output ["Multirate Systems and Filter Banks" P.P. Vaidyanathan Prentice Hall: Englewood Cliffs, NJ, 1993]. When sharing a common input the filter bank is called an analysis bank. The analysis bank splits the incoming signal into M separate signals called subband signals. The analysis filters are denoted $H_k(z)$, where $k = 0 \ldots M-1$. The filter bank is critically sampled (or maximally decimated) when the subband signals are decimated by a factor M. The total number of subband samples per unit time is then the same as the number of samples per unit time for the input signal. The synthesis bank combines these subband signals into a common output signal. The synthesis filters are denoted $F_k(z)$, for $k = 0 \ldots M-1$. A maximally decimated filter bank with M channels (subbands) is shown in Fig.1. The analysis part 101 produces the signals $V_k(z)$, which constitute the signals to be transmitted, stored or modified, from the input signal $X(z)$. The synthesis part 102 recombines the signals $V_k(z)$ to the output signal $\hat{X}(z)$.

The recombination of $V_k(z)$ to obtain the approximation $\hat{X}(z)$ of the original signal $X(z)$ is subject to several errors. One of these is aliasing, due to the desimation and interpolation of the subbands. Other errors are phase and amplitude distortion.

Following the notations of Fig.1, the outputs of the analysis filters $H_k(z)$ 103 are

$$X_k(z) = H_k(z)X(z) \tag{1}$$

where $k = 0 \ldots M-1$. The decimators 104 give the outputs

$$V_k(z) = \frac{1}{M}\sum_{l=0}^{M-1} X_k(z^{1/M}W^l) = \frac{1}{M}\sum_{l=0}^{M-1} H_k(z^{1/M}W^l)X(z^{1/M}W^l) \tag{2}$$

where $W = e^{-j2\pi/M}$. The outputs of the interpolators 105 are given by

$$U_k(z) = V_k(z^M) = \frac{1}{M}\sum_{l=0}^{M-1} H_k(zW^l)X(zW^l) \tag{3}$$

and the sum of the signals obtained from the synthesis filters 106 can be written as

$$\hat{X}(z) = \sum_{k=0}^{M-1} F_k(z)U_k(z) = \sum_{k=0}^{M-1} F_k(z)\frac{1}{M}\sum_{l=0}^{M-1} H_k(zW^l)X(zW^l) = \tag{4}$$

$$= \frac{1}{M}\sum_{l=0}^{M-1} X(zW^l)\sum_{k=0}^{M-1} H_k(zW^l)F_k(z) = \frac{1}{M}\sum_{l=0}^{M-1} X(zW^l)A_l(z)$$

where

$$A_l(z) = \sum_{k=0}^{M-1} H_k(zW^l)F_k(z) \tag{5}$$

is the gain for the lth alias term $X(zW^l)$. Eq.(4) can be rewritten as

$$\hat{X}(z) = \frac{1}{M}\left\{X(z)A_0(z) + \sum_{l=1}^{M-1} X(zW^l)A_l(z)\right\} \quad (6)$$

The last sum on the right hand side (RHS) constitutes the sum of all non-wanted alias terms. Canceling all aliasing, that is forcing this sum to zero by means of proper choices of $H_k(z)$ and $F_k(z)$, gives

$$\hat{X}(z) = \frac{1}{M}X(z)A_0(z) = \frac{1}{M}X(z)\sum_{k=0}^{M-1} H_k(z)F_k(z) = X(z)T(z) \quad (7)$$

where

$$T(z) = \frac{1}{M}\sum_{k=0}^{M-1} H_k(z)F_k(z) \quad (8)$$

is the overall transfer function or distortion function. Choosing the synthesis filter $F_k(z)$ so that

$$F_k(z) = z^{-N}\tilde{H}_k(z) \quad (9)$$

where N is the order of the analysis filter, results in the following transfer function:

$$T(z) = \frac{1}{M}\sum_{k=0}^{M-1} H_k(z)F_k(z) = \frac{z^{-N}}{M}\sum_{k=0}^{M-1} H_k(z)\tilde{H}_k(z) \quad (10)$$

The notation $\tilde{H}(z)$ is the Z-transform of the time-reversed and complex-conjugated sequence $h_k(n)$. Eq.(10) evaluated on the unit circle gives

$$T(e^{j\omega}) = \frac{e^{-j\omega N}}{M}\sum_{k=0}^{M-1} H_k(e^{j\omega})H_k^*(e^{-j\omega}) = \frac{e^{-j\omega N}}{M}\sum_{k=0}^{M-1}\left|H_k(e^{j\omega})\right|^2 \quad (11)$$

Eq.(11) shows that $T(z)$ has linear phase, and thus has no phase distortion. Further, if the last sum on the RHS is a constant, there is no amplitude distortion. The overall transfer function is in this case simply a delay with a constant scale factor c, i.e.

$$T(z) = cz^{-N} \quad (12)$$

which substituted into Eq.(7) gives

$$\hat{X}(z) = cz^{-N}X(z) \quad (13)$$

The type of filters that satisfy Eq.(13) are said to have the perfect reconstruction (PR) property.

Cosine Modulated Filterbanks

In a cosine modulated filterbank the analysis filters $h_k(n)$ are cosine modulated versions of a symmetric low-pass prototype filter $p_0(n)$ as

$$h_k(n) = 2p_0(n)\ \cos\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \quad (14)$$

where M is the number of channels, $k = 0 \ldots M-1$, N is the prototype filter order and $n = 0 \ldots N$. The sum of the real-valued prototype filter coefficients is assumed to be unity as

$$\sum_{n=0}^{N} p_0^2(n) = 1 \quad (15)$$

Following the same notation, the synthesis filters are given by

$$f_k(n) = 2p_0(N-n)\ \cos\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \quad (16)$$

The analysis filter bank produces real-valued subband samples for real-valued input signals. The subband samples are downsampled by a factor M, which makes the system critically sampled. Depending on the choice of the prototype filter, the filterbank may constitute a near perfect reconstruction system, a so called pseudo QMF bank [US5436940], or a perfect reconstruction (PR) system. An example of a PR system is the modulated lapped transform (MLT) ["Lapped Transforms for Efficient Transform/Subband Coding" H.S. Malvar, IEEE Trans ASSP, vol. 38, no. 6, 1990]. One inherent property of the cosine modulation is that every filter has two passbands; one in the positive frequency range and one corresponding passband in the negative frequency range.

Eq.(3) can be written in matrix form as

$$\mathbf{a} = \mathbf{H}\,\mathbf{f} \quad (17)$$

or explicitly as

$$\begin{bmatrix} A_0(z) \\ A_1(z) \\ \vdots \\ A_{M-1}(z) \end{bmatrix} = \begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(Wz) & H_1(Wz) & & H_{M-1}(Wz) \\ \vdots & & \ddots & \vdots \\ H_0(W^{M-1}z) & H_1(W^{M-1}z) & \cdots & H_{M-1}(W^{M-1}z) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} \quad (18)$$

The matrix H is called the alias component (AC) matrix. For a closer examination of this equation, f may be written as

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ \cdot \\ \cdot \\ \cdot \\ F_{M-1}(z) \end{bmatrix} = \begin{bmatrix} F_0(z) & 0 & \cdots & 0 \\ 0 & F_1(z) & & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & & 0 \\ 0 & \cdots & 0 & F_{M-1}(z) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \quad (19)$$

or in a compact form as

$$\mathbf{f} = \mathbf{F}\,\mathbf{e} \tag{20}$$

Substituting Eq.(20) in Eq.(17), the alias gains can be written as $\mathbf{a} = \mathbf{H}\mathbf{F}\,\mathbf{e}$, where the product

$$\mathbf{H}\mathbf{F} = \mathbf{U} \tag{21}$$

is a $M \times M$ matrix, hereinafter referred to as the composite alias component matrix.

For a cosine modulated system, the dominant terms in the composite alias component matrix are the first row and four diagonals. The three-dimensional plot of Fig.2 illustrates the magnitudes of the components in this matrix. The first row holds the terms from the transfer function, Eq.(8), while the four diagonals primarily consist of the main alias terms, i.e. the aliasing due to overlap between filters and their closest neighbors. It is easily seen that the main alias terms emerge from overlap in frequency between either the filters negative passbands with frequency modulated versions of the positive passbands, or reciprocally, the filters positive passbands with frequency modulated versions of the negative passbands. Summing the terms of the rows in the composite alias component matrix, i.e. calculating the alias gains, results in cancellation of the main alias terms. The aliasing is canceled in a pairwise manner, where the first main alias term is canceled by the second in the same row. Superimposed on the main alias terms are also other smaller alias terms. If the prototype filter characteristics is so that the transition- and stop-band of the filters have substantial overlap with their modulated versions, these alias terms will be large. As an example, the second and the last row consists of alias terms induced by the overlap of filters with their closest modulated versions. For a PR system, these smaller alias terms also cancels completely when summing the terms for the alias gains. In the pseudo QMF system, however, these terms remain.

Complex-Exponential Modulated Filterbanks

Extending the cosine modulation to complex-exponential modulation according to the present invention yields the analysis filters $h_k(n)$ as

$$h_k(n) = p_0(n)\ \exp\left\{i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \tag{22}$$

using the same notation as before. This can be viewed as adding an imaginary part to the real-valued filterbank, where the imaginary part consists of sine modulated versions of the same prototype filter. Considering a real-valued input signal, the output from the filter bank can be interpreted as a set of subband signals, where the real and the imaginary parts are Hilbert transforms of each other. The resulting subbands are thus the analytic signals of the real-valued output obtained from the cosine modulated filterbank. Hence, due to the complex-valued representation, the subband signals are oversampled by a factor two.

The synthesis filters are extended in the same way as

$$f_k(n) = p_0(N-n)\ \exp\left\{i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \tag{23}$$

Eq.(22) and (23) implies that the output from the synthesis bank is complex-valued. Using matrix notation, where $\mathbf{C}_a$ is a matrix with analysis filters from Eq.(14), and $\mathbf{S}_s$ is a matrix with filters as

$$h_k(n) = p_0(n)\ \sin\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}-\frac{M}{2})\right\} \qquad (24)$$

the filters of Eq.(22) is obtained as $\mathbf{C_a} + \mathrm{j}\,\mathbf{S_a}$. In these matrices, k is the row index and n is the column index. Analogously, the matrix $\mathbf{C_s}$ has synthesis filters from Eq.(16), and $\mathbf{S_s}$ is a matrix with filters as

$$f_k(n) = p_0(N-n)\ \sin\left\{\frac{\pi}{2M}(2k+1)(n-\frac{N}{2}+\frac{M}{2})\right\} \qquad (25)$$

Eq.(23) can thus be written $\mathbf{C_s} + \mathrm{j}\,\mathbf{S_s}$, where k is the column index and n is the row index. Denoting the input signal $\mathbf{x}$, the output signal $\mathbf{y}$ is found from

$$\mathbf{y} = (\mathbf{C_s} + \mathrm{j}\,\mathbf{S_s})(\mathbf{C_a} + \mathrm{j}\,\mathbf{S_a})\,\mathbf{x} = (\mathbf{C_s}\mathbf{C_a} - \mathbf{S_s}\mathbf{S_a})\,\mathbf{x} + \mathrm{j}\,(\mathbf{C_s}\mathbf{S_a} + \mathbf{S_s}\mathbf{C_a})\,\mathbf{x} \qquad (26)$$

As seen from Eq.(26), the real part consists of two terms; the output from the ordinary cosine modulated filterbank and an output from a sine modulated filterbank. It is easily verified that if a cosine modulated filterbank has the PR property, then its sine modulated version, with a change of sign, constitutes a PR system as well. Thus, by taking the real part of the output, the complex-exponential modulated system offers the same reconstruction accuracy as the corresponding cosine modulated version.

The complex-exponential modulated system can be extended to handle also complex-valued input signals. By extending the number of channels to $2M$, i.e. adding the filters for the negative frequencies, and keeping the imaginary part of the output signal, a pseudo QMF or a PR system for complex-valued signals is obtained.

Examining the composite alias component matrix from Eq.(21), the main alias diagonals vanish for the complex-exponential modulated filterbank. This is easily understood since the complex-exponential modulated filterbank has only one passband for every filter. In other words, the filterbank is free from main alias terms, and do not rely on the pairwise aliasing cancellation mentioned above. The composite alias component matrix has the dominant terms on the first row only. Fig 3 shows the magnitude of the components in the resulting matrix. Depending on the prototype filter characteristics, the terms on rows 1 through M-1, are more or less attenuated. The absence of main alias terms makes the aliasing cancellation constraint from the cosine (or sine) modulated filterbank obsolete in the complex-exponential modulated version. Both the analysis and synthesis filters can thus be found from

$$h_k(n) = f_k(n) = p_0(n)\ \exp\left\{i\ \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\} \qquad (27)$$

since $p_0(n) = p_0(N-n)$ for a symmetric prototype filter. As before, M is the number of channels, $k = 0 \ldots M$-1, N is the prototype filter order and $n = 0 \ldots N$.

Referring to Eq.(4), the Z-transform of the real part of the output signal $\hat{x}(n)$ is

$$Z\left\{\mathrm{Re}\left\{\hat{x}(n)\right\}\right\} = \hat{X}_R(z) = \frac{\hat{X}(z) + \hat{X}^*(z)}{2} \qquad (28)$$

The notation $\hat{X}*(z)$ is the Z-transform of the complex-conjugated sequence $\hat{x}(n)$. From Eq.(4), it follows that the transform of the real part of the output signal is

$$2M\hat{X}_R(z) = \sum_{l=0}^{M-1}\left(X(zW^l)A_l(z) + X(zW^{-l})A_l*(z)\right) \tag{29}$$

where it is used that the input signal $x(n)$ is real-valued. Eq.(29) may after manipulation be written

$$\begin{aligned} 2M\hat{X}_R(z) = {} & X(z)\{A_0(z) + A_0*(z)\} + \\ & + X(-z)\{A_{M/2}(z) + A_{M/2}*(z)\} + \\ & + \sum_{l=1}^{M/2-1} X(zW^l)\{A_l(z) + A_{M-l}*(z)\} + \\ & + \sum_{l=1}^{M/2-1} X(zW^{-l})\{A_{M-l}(z) + A_l*(z)\} \end{aligned} \tag{30}$$

By inspecting Eq.(30), and recalling the transform of Eq.(28), it is obvious that the real part of $a_0(n)$ must be a dirac pulse for a PR system. Moreover, the real part of $a_{M/2}(n)$ must be zero and the alias gains, for $l = 1 \ldots M/2-1$, must satisfy

$$A_{M-l}(z) = -A_l*(z) \tag{31}$$

In pseudo QMF systems, Eq.(31) holds true only approximately. Moreover, the real part of $a_0(n)$ is not exactly a dirac-pulse, nor is the real part of $a_{M/2}$(n) exactly zero.

Modification of subband signals

Changing the gains of the channels in a cosine modulated filterbank, i.e. using the analysis/synthesis system as an equalizer, renders severe distortion due to the main alias terms. Say for example that the intention is to adjust an eight-channel filterbank for a band-pass response, where except for the second and third channel all subband signals are set to zero. The composite alias component matrix from Eq.(21) is then an 8 × 8 matrix where all elements are zero except for the elements on the second and third column, Fig.4. There are seven significant alias terms left as indicated in the figure. The aliasing from rows three and five will be canceled, since the main alias terms in these rows have the same gains, i.e. the pairwise cancellation is working intentionally. In rows two, four and six however, there are single alias terms, since their corresponding alias terms have zero gain. The alias cancellation will hence not work as intended, and the aliasing in the output signal will be substantial.

From this example it is obvious that a substantial improvement is achieved when using complex-exponential modulated filterbanks as equalizers. The 8-channel system depicted in Fig.4 has a prototype filter of order 128. The total aliasing attenuation is only 16 dB in the above equalizer example. Moving to complex-exponential modulation gives an aliasing attenuation of 95 dB. Due to the non-existing main

alias terms, the resulting aliasing is dependent only on the suppression of the alias terms emanating from overlap between filters and their modulated versions. It is thus of great importance to design the prototype filter for maximum suppression of the alias gains terms. The first term on the RHS of Eq.(30) evaluated on the unit circle gives the error energy e_t of the transfer function as

$$e_t = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\left|A_0(e^{j\omega}) + A_0^*(e^{-j\omega})\right|}{2M} - 1\right)^2 d\omega \qquad (32)$$

The energy of the total aliasing e_a can be calculated by evaluating all remaining terms on the RHS of Eq.(30) on the unit circle as

$$\begin{aligned} e_a = &\frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega}) + A_{M/2}^*(e^{-j\omega})\right|^2 d\omega + \\ &+\frac{1}{8\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left\{\left|A_l(e^{j\omega}) + A_{M-l}^*(e^{-j\omega})\right|^2 + \right. \\ &\left. +\left|A_{M-l}(e^{j\omega}) + A_l^*(e^{-j\omega})\right|^2\right\} d\omega \end{aligned} \qquad (33)$$

Due to symmetry, Eq.(9) and the fact

$$F_0(z) = z^{-N}\tilde{F}_0(z) \qquad (34)$$

the terms within the curly braces of the sum in Eq.(33) are equal. The total aliasing energy thus has $M/2-1$ terms as

$$\begin{aligned} e_a = &\frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega}) + A_{M/2}^*(e^{-j\omega})\right|^2 d\omega + \\ &+\frac{1}{4\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left|A_l(e^{j\omega}) + A_{M-l}^*(e^{-j\omega})\right|^2 d\omega \end{aligned} \qquad (35)$$

The minimization of the alias gain terms is done by optimizing the prototype filter. This is preferably accomplished by minimizing a composite objective function, using standard nonlinear optimization algorithms, for example the Downhill Simplex Method ["Numerical Recipes in C, The Art of Scientific Computing, Second Edition" W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Cambridge University Press, NY, 1992]. For alias term minimization (ATM) of the prototype filter according to the invention, the objective function looks like

$$e_{tot}(\alpha) = \alpha e_t + (1-\alpha)e_a \qquad (36)$$

During the optimization, a random equalization curve is applied to the filterbank when calculating e_a, i.e. the analysis and synthesis filters are multiplied with gainfactors g_k as

$$H_k^{(eq)}(z)=g_k H_k(z) \quad \text{and} \quad F_k^{(eq)}(z)=g_k F_k(z) \tag{37}$$

and the resulting filters $H_k^{(eq)}$ and $F_k^{(eq)}$, $k = 0 \ldots M-1$, are used when calculating the alias gain terms $A_l(z)$ for $l = 1 \ldots M-1$.

In Fig.5, the alias gains of five different complex-exponential modulated systems are compared. Four of these are 8-channel systems and one is a 64-channel system. All of the systems have prototype filter lengths of 128. The dotted trace and the solid trace with stars shows alias components for two pseudo QMF systems, where one is alias term minimized. The dashed and the dash-dotted traces are the components for two 8-channel perfect reconstruction systems, where again one of the systems is alias term minimized. The solid trace is the alias components for a complex-exponential modulated lapped transform (*MLT*). All the systems are adjusted for band-pass responses according to the example above, and the results are tabulated in Table 1. The rejection of total aliasing is calculated as the inverse of Eq.(33). The passband flatness is calculated as the inverse of Eq.(32) with the integration interval adjusted for the bandpass response.

Table 1

System	Rejection of total aliasing	Passband flatness
8-ch Pseudo QMF ATM N=128	114,7 dB	98,1 dB
8-ch Pseudo QMF N=128	95,4 dB	87,6 dB
8-ch PR ATM N=127	77,3 dB	132,7 dB
8-ch PR N=127	55,0 dB	93,6 dB
64-ch MLT N=127	38,5 dB	87,1 dB

As seen from the numbers in Table 1, a substantial improvement is achieved when moving from the 64-channel *MLT* to the 8-channel PR systems. The MLT is a perfect reconstruction system and has only $(N+1) / 2M = 1$ coefficient per polyphase component. The number of coefficients for the 8-channel PR systems are 128 / 16 = 8. This enables filters with higher stopband attenuation and higher rejection of alias terms. Moreover, it is seen that alias term minimization of the PR system rejects the aliasing and improves the passband flatness significantly. Comparing the pseudo QMF systems and the PR systems, it is clear that the aliasing rejection improves by 46 dB while almost preserving the passband flatness. An additional alias rejection of approximately 20 dB and improved passband flatness of 10 dB is achieved when minimizing the alias terms. Thus, it is clear that the perfect reconstruction constraint imposes limitations to a filterbank used in an equalization system. A pseudo QMF system can always be designed for adequate reconstruction accuracy, since all practical digital implementations have limited resolution in the numeric representation. For both the pseudo QMF and the PR system, it is obvious that an optimum system is built on a prototype filter that has large rejection of the stopband. This imposes the usage of prototype filters with relative lengths longer than the windows used in the MLT.

A great advantage of the complex-exponential modulated system is that the instantaneous energy is easily calculated since the subband signals constitute the analytic signals of the real-valued subband signals

obtained from a cosine modulated filterbank. This is a property of great value in for example adaptive filters, automatic gain controls (AGC), in multiband companders, and in spectral band replication systems (SBR), where a filterbank is used for the spectral envelope adjustment. The averaged energy within a subband k might be calculated as

$$E_k(m) = \sum_{n=-L+1}^{L-1} |v_k(mL+n)|^2 w(n) \qquad (38)$$

where $v_k(n)$ is the subband samples of channel k, and $w(n)$ is a window of length $2L-1$ centered around $n=0$. This measure is then used as an input parameter for the adaptation or gain-calculation algorithms.

Practical implementations

Using a standard PC or DSP, real-time operation of a complex-exponential modulated filterbank is possible. The filterbank may also be hard-coded on a custom chip. Fig.6 shows the structure for an effective implementation of the analysis part of a complex-exponential modulated filterbank system. The analogue input signal is first fed to an A/D converter 601. The digital time domain signal is fed to a shift register holding $2M$ samples shifting M samples at a time 602. The signals from the shift register are then filtered through the polyphase coefficients of the prototype filter 603. The filtered signals are subsequently combined 604 and in parallel transformed with a DCT-IV 605 and a DST-IV 606 transform. The outputs from the cosine and sine transforms constitute the real and the imaginary parts of the subband samples respectively. The gains of the subband samples are modified according to the current spectral envelope adjuster setting 607.

An effective implementation of the synthesis part of a complex-exponential modulated system is shown in Fig.7. The subband samples are first multiplied with complex-valued twiddlefactors 701, and the real part is modulated with a DCT-IV 702 and the imaginary part with a DST-IV 703 transform. The outputs from the transforms are combined 704 and fed through the polyphase components of the prototype filter 705. The time domain output signal is obtained from the shift register 706. Finally, the digital output signal is converted back to an analogue waveform 707.

The above-described embodiments are merely illustrative for the principles of the complex-exponential modulated filterbank systems according to the present invention. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.

CLAIMS

1. A method for estimation of energy measures of coefficients or subband signals obtained from a digital filterbank, **characterized by:**

- optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- building an M-channel analysis filterbank by complex-exponential modulation of said prototype filter; where said filterbank has filter coefficients as

$$h_k(n) = p_0(n) \exp\left\{ i \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\},$$

and $n = 0, 1, \ldots, N$, and $k = 0,1, \ldots, M-1$;

- filtering a real-valued time domain signal through said filterbank; and

calculating the squared absolute values of the complex-valued subband signals obtained from said filtering.

2. A method according to claim 1, **characterized in** that said analysis filterbank is used to estimate average energy measures as

$$E_k(n) = \sum_{n=-L+1}^{L-1} |v_k(nL+n)|^2 w(n)$$

where $v_k(n)$ is said complex-valued subband signal of channel k, and $w(n)$ is a window of length $2L-1$ centered around $n = 0$.

3. A method for reduction of aliasing emerging from modifications of the coefficients or subband signals obtained from a digital filterbank, **characterized by:**

- optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- building an M-channel filterbank by complex-exponential modulation of said prototype filter; where said filterbank has analysis and synthesis filter coefficients as

$$h_k(n) = f_k(n) = p_0(n) \exp\left\{ i \frac{\pi}{2M}(2k+1)(n-\frac{N}{2}) \right\},$$

and $n = 0 \ldots N$, and $k = 0 \ldots M-1$;

- filtering a real-valued time domain signal through the analysis part of said filterbank;
- modifying the complex-valued subband signals obtained from said filtering;
- filtering said modified complex-valued subband signals through the synthesis part of said filterbank; and
- taking the real part of the complex-valued time domain output signal, where said output signal is the sum of the signals obtained from said synthesis filtering.

4. A method according to claims **1 – 3, characterized in** that said low-pass prototype filter has a filter order N that is higher than $2M$-1, where M is the number of channels in said digital filterbank.

5. A method according to claims **1 – 4, characterized in** that said optimization of said low-pass prototype filter is accomplished by minimizing a composite objective function $\varepsilon_{tot}(\alpha)$ as

$$\varepsilon_{tot}(\alpha) = \alpha\varepsilon_t + (1-\alpha)\varepsilon_a$$

where α is a weighting constant, ε_t is the passband error and ε_a is the aliasing error.

6. A method according to claim **5, characterized in** that said passband error ε_t is calculated as

$$e_t = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{\left|A_0(e^{j\omega}) + A_0^*(e^{-j\omega})\right|}{2M} - 1\right)^2 d\omega,$$

and said aliasing error ε_a is calculated as

$$e_a = \frac{1}{8\pi M^2}\int_{-\pi}^{\pi}\left|A_{M/2}(e^{j\omega}) + A_{M/2}^*(e^{-j\omega})\right|^2 d\omega + \frac{1}{4\pi M^2}\sum_{l=1}^{M/2-1}\int_{-\pi}^{\pi}\left|A_l(e^{j\omega}) + A_{M-l}^*(e^{-j\omega})\right|^2 d\omega$$

where $A_l(e^{j\omega})$, $l = 0 \ldots M/2$, are the alias gain terms evaluated on the unit circle.

7. A method according to claim **6, characterized in** that said alias gain terms are calculated using gain adjusted versions of said analysis and synthesis filters according to

$$h_k^{(eq)}(n) = g_k h_k(n) \quad \text{and} \quad f_k^{(eq)}(n) = g_k f_k(n),$$

where g_k represents randomized or deterministic real-valued gainfactors, $n = 0 \ldots N$, and $k = 0 \ldots M$-1.

8. A method according to claims **5 – 7, characterized in** that said low-pass prototype filter has a filter order of at least a number N, where N is chosen so that said composite objective function is smaller than a predefined value.

9. A method according to claims **1 – 2, characterized in** that said analysis filterbank is used for the estimation of energy measures in a high frequency reconstruction system.

10. A method according to claims **3 – 8, characterized in** that said complex-exponential modulated filterbank is used as the envelope adjusting filterbank in a high frequency reconstruction system.

11. A method according to claims 3 – 8, **characterized in** that said complex-exponential modulated filterbank is used as the filters in a digital equalizing system.

12. A method according to claims 3 – 8, **characterized in** that said complex-exponential modulated filterbank is used in an adaptive filtering system.

13. A method according to claims 3 – 8, **characterized in** that said complex-exponential modulated filterbank is used in a multiband compander system.

14. An apparatus for estimation of energy measures of coefficients or subband signals obtained from a digital filterbank, **characterized by:**
- means for optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- means for building an M-channel analysis filterbank by complex-exponential modulation of said prototype filter; where said filterbank has filter coefficients as

$$h_k(n) = p_0(n) \exp\left\{i \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\},$$

and $n = 0, 1, \ldots, N$, and $k = 0,1, \ldots, M-1$;
- means for filtering a real-valued time domain signal through said filterbank; and
means for calculation of the squared absolute values of the complex-valued subband signals obtained from said filtering.

15. An apparatus for reduction of aliasing emerging from modifications of the coefficients or subband signals obtained from a digital filterbank, **characterized by:**
- means for optimizing a symmetric low-pass prototype filter $p_0(n)$, having filter order N, not constrained to have the perfect reconstruction property;
- means for building an M-channel filterbank by complex-exponential modulation of said prototype filter; where said filterbank has analysis and synthesis filter coefficients as

$$h_k(n) = f_k(n) = p_0(n) \exp\left\{i \frac{\pi}{2M}(2k+1)(n-\frac{N}{2})\right\},$$

and $n = 0, 1, \ldots, N$, and $k = 0,1, \ldots, M-1$;
- means for filtering a real-valued time domain signal through the analysis part of said filterbank;
- means for modifying the complex-valued subband signals obtained from said filtering;
- means for filtering said modified complex-valued subband signals through the synthesis part of said filterbank; and
means for taking the real part of the complex-valued time domain output signal, where said output signal is the sum of the signals obtained from said synthesis filtering.

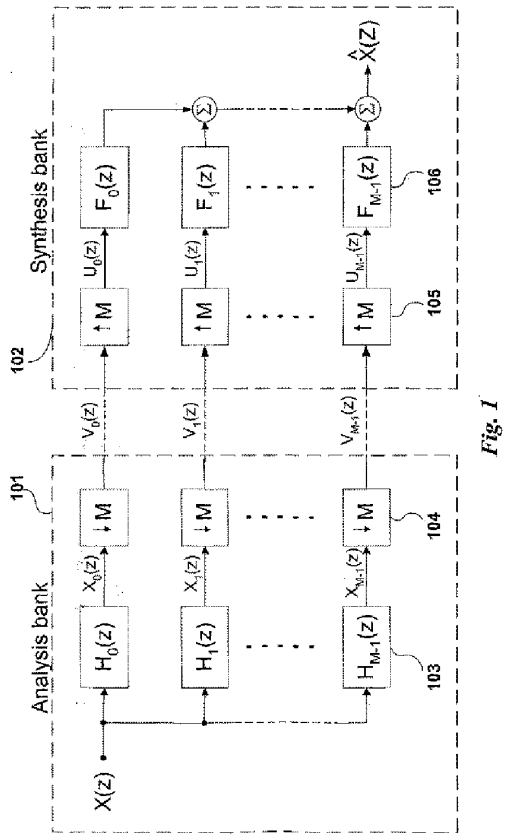

Analysis bank
Synthesis bank
X(z)
$\hat{X}(z)$
$H_0(z)$
$H_1(z)$
$H_{M-1}(z)$
$F_0(z)$
$F_1(z)$
$F_{M-1}(z)$
↓M
↑M
$X_0(z)$
$X_1(z)$
$X_{M-1}(z)$
$V_0(z)$
$V_1(z)$
$V_{M-1}(z)$
$U_0(z)$
$U_1(z)$
$U_{M-1}(z)$
101
102
103
104
105
106

Fig. 1

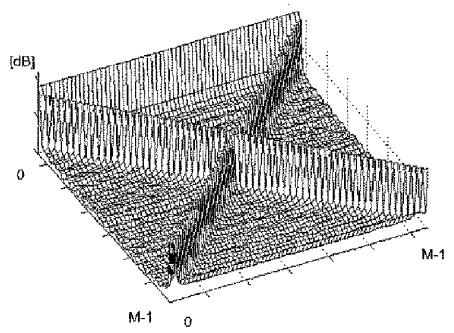


Fig. 2

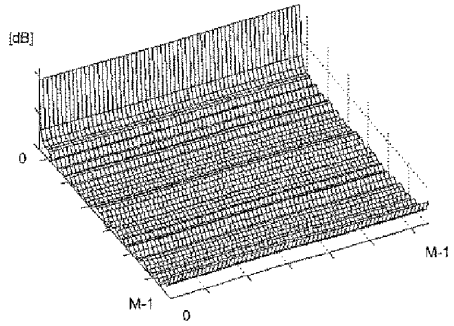


Fig. 3

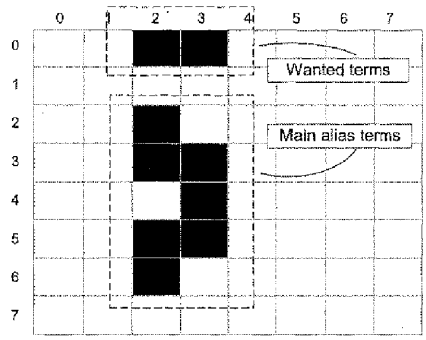

0
1
2
3
4
5
6
7
Wanted terms
Main alias terms

Fig. 4

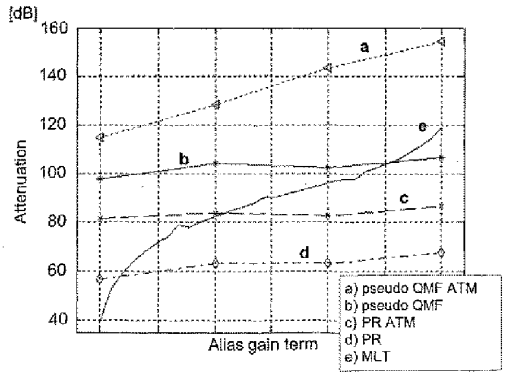

[dB]
160
140
120
100
80
60
40
Attenuation
Alias gain term
a
b
c
d
e
a) pseudo QMF ATM
b) pseudo QMF
c) PR ATM
d) PR
e) MLT

Fig. 5

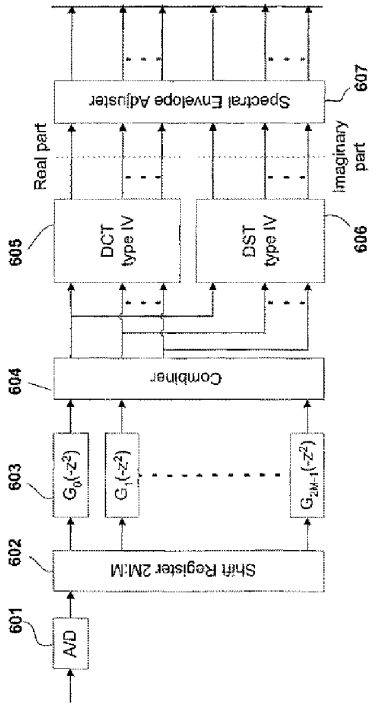

601
A/D
602
Shift Register 2M:M
603
$G_0(-z^2)$
$G_1(-z^2)$
$G_{2M-1}(-z^2)$
604
Combiner
605
DCT type IV
606
DST type IV
607
Spectral Envelope Adjuster
Real part
imaginary part

Fig. 6

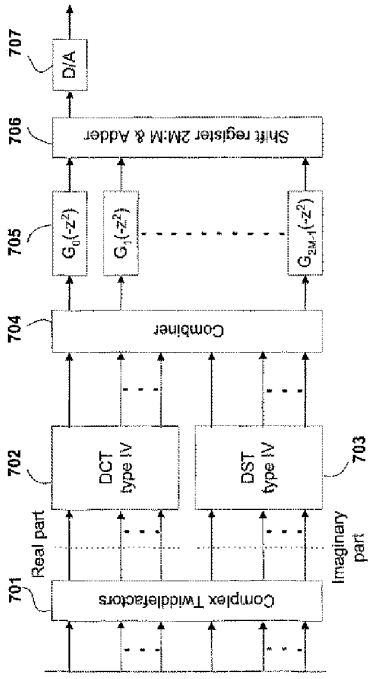

707
D/A
706
Shift register 2M:M & Adder
705
$G_0(-z^2)$
$G_1(-z^2)$
$G_{2M-1}(-z^2)$
704
Combiner
702
DCT type IV
703
DST type IV
Real part
Imaginary part
701
Complex Twiddlefactors

Fig. 7

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 02/00626

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H03H 17/00
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: H03H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5732189 A (JOHNSTON ET AL), 24 March 1998 (24.03.98), column 5, line 11 - column 9, line 63	1-15
A	US 5285498 A (JOHNSTON), 8 February 1994 (08.02.94), column 3, line 1 - column 4, line 9	1-15
A	WO 0141320 A1 (SOLANA TECHNOLOGY DEVELOPMENT CORPORATION), 7 June 2001 (07.06.01), See the whole document	1-15
A	US 5568142 A (VELAZQUEZ ET AL), 22 October 1996 (22.10.96), See the whole document	1-15

☐ Further documents are listed in the continuation of Box C.　　☒ See patent family annex.

* Special categories of cited documents
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 July 2002	19 -07- 2002
Name and mailing address of the ISA Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Stefan Hultquist/SN Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT
Information on patent family members

10/06/02

International application No.
PCT/SE 02/00526

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5732189 A	24/03/98	NONE	
US 5285498 A	08/02/94	CA 2090159 A	03/09/93
		DE 69326492 D,T	04/05/00
		EP 0559383 A,B	08/09/93
		JP 3258424 B	18/02/02
		JP 6075590 A	18/03/94
		KR 9707661 B	15/05/97
		US 5481614 A	02/01/96
WO 0141320 A1	07/06/01	AU 1837200 A	12/06/01
US 5568142 A	22/10/96	CA 2203156 A	02/05/96
		EP 0787385 A	06/08/97
		JP 10507891 T	28/07/98
		WO 9613097 A	02/05/96

Form PCT/ISA/210 (patent family annex) (July 1998)

(81)指定国　　AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZM,ZW

AUDIEN.FOR0045

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **05-172865**

(43)Date of publication of application : **13.07.1993**

(51)Int.Cl. **G01R 23/16**
G01H 3/04

(21)Application number : **04-121314**

(22)Date of filing : **15.04.1992**

(71)Applicant : **HEWLETT PACKARD CO <HP>**

(72)Inventor : **WAITE JR JAMES W**

(30)Priority

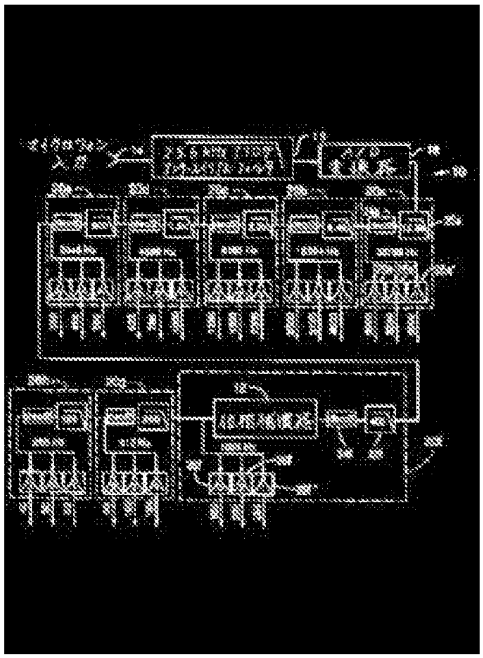
Priority number : **91 685370** Priority date : **15.04.1991** Priority country : **US**

(54) TIME REGION SPECTRUM ANALYZING METHOD, METHOD FOR DETERMINING INTENSITY OF SOUND AND REAL TIME OCTAVE ANALYSER

(57)Abstract:

PURPOSE: To simply compensate the phase error between the measuring channels of a signal analyser having a plurality of measuring channels.

CONSTITUTION: The signal inputted from a microphone 14 is digitalized by an A/D converter 18 through an anti-alias filter 16 to be analyzed by a series of analyzing stages 20a-20h. A phase compensating device 12 for compensating the phase error between this channel and other channel is inserted in at least one analyzing stage and the phase error for this phase compensating device 12 is approximated by a digital filter properly selecting pole/zero to be used in compensation.



(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号

特開平5－172865

(43)公開日　平成５年(1993)７月13日

(51)Int.Cl.[5]	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０１Ｒ　23/16	Ｂ	8803－2Ｇ		
Ｇ０１Ｈ　3/04		8117－2Ｇ		

審査請求　未請求　請求項の数３(全　８　頁)

(21)出願番号　特願平4－121314

(22)出願日　平成４年(1992)４月15日

(31)優先権主張番号　６８５３７０

(32)優先日　1991年４月15日

(33)優先権主張国　米国（ＵＳ）

(71)出願人　590000400
ヒューレット・パッカード・カンパニー
アメリカ合衆国カリフォルニア州パロアルト　ハノーバー・ストリート　3000

(72)発明者　ジェイムズ・ダブリュウ・ウェイト，ジュニア
アメリカ合衆国ワシントン州エバレット，グランド・アベニュウ　1012

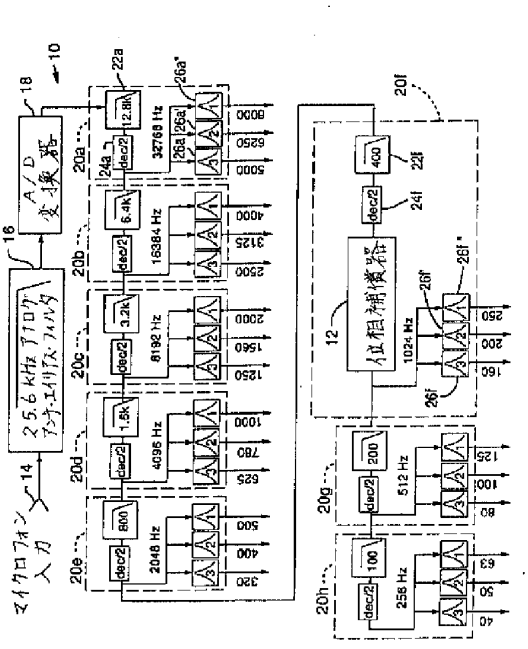
(74)代理人　弁理士　長谷川　次男

(54)【発明の名称】　時間領域スペクトル解析方法，音の強さを決定するための方法及び実時間オクターブ解析装置

(57)【要約】

【目的】複数測定チャネルを有する信号解析装置の測定チャネル間の位相エラーを簡便に補償する。

【構成】マイクロフォン１４から入力した信号はアンチ・エイリアス・フィルタ１６を経てＡ／Ｄ変換器１８でデジタル化された後一連の解析段２０ａ～２０ｈにおいて解析される。このチャネルの他のチャネルとの位相エラーを補償するための位相補償器１２を少なくとも１つの解析段に挿入した。この位相補償器のための位相エラーは極／ゼロを適切に選んだデジタル・フィルタによって近似され，補償に用いられる。



1

【特許請求の範囲】

【請求項１】後記（イ）乃至（ハ）のステップを含む時間領域スペクトル解析方法。

（イ）補償すべき伝達特性を決定するステップ，（ロ）前記伝達特性を近似する極／零伝達関数を合成するステップ，及び（ハ）前記極／零伝達関数に従って適応デジタル・フィルタをプログラムスルステップ。

【請求項２】（イ）第一，第二のマイクロフォンからの信号に関する位相エラー関数を決定するステップと，

（ロ）前記第一のマイクロフォンの出力信号のスペクトル成分を決定するために前記第一のマイクロフォンの出力信号を第一の帯域フィルタ列を通して処理するステップと，（ハ）前記第二のマイクロフォンの出力信号のスペクトル成分を決定するために前記第二のマイクロフォンの出力信号を第二の帯域フィルタ列を通して処理するステップと，（ニ）音の強さを決定するために前記第一，第二の帯域フィルタ列の出力信号を解析するステップとからなり，前記処理するステップの少なくとも一方が前記決定された位相エラー関数に応じた補正を行うための時間領域フィルタリングを含む音の強さを決定するための方法。

【請求項３】少なくとも二つの測定チャネルを有し，各測定チャネルが後記（イ）乃至（ト）から成る実時間オクターブ解析装置。

（イ）入力信号源，（ロ）前記入力信号源に結合した入力端を有するアンチ・エイリアス・フィルタ，（ハ）前記アンチ・エイリアス・フィルタの出力端に結合した入力端を有するアナログ・デジタル変換器。（ニ）それぞ＊

2

＊れの入力端，出力端及び少なくとも一つの帯域フィルタを有する複数の逐次解析段，（ホ）前記逐次解析段の第一の解析手段の入力端を前記アナログ・デジタル変換器の出力端に結合する手段，（ヘ）前記逐次解析段の各解析手段の入力端を前段の各解析手段の出力端に結合する手段，（ト）前記測定チャネルの一方に前記入力信号源に付随する位相エラーを補償するための時間領域デジタル・フィルタを挿入するための手段。

【発明の詳細な説明】

【０００１】

【発明の技術分野】本発明は、マイクロフォンなどの、１対の測定用変換器の特性を正確に整合させるための方法及び装置に関するものである。

【０００２】

【０００３】通常、音の強さは、精確に整合がとれ、近接して配置された１対のマイクロフォン（一般に音の強さの「プローブ」として知られる）を利用して、測定される。１９８８年1月の、Ｊｏｕｒｎａｌ　ｏｆ　Ｖｉｂｒａｔｉｏｎ，Ｓｔｒｅｓｓ，ａｎｄ　Ｒｅｌｉａｂｉｌｉｔｙ　ｉｎ　Ｄｅｓｉｇｎ　第１１０巻、９７頁～１０３頁におけるＰｏｐｅ，Ｊ．による“Ｔｗｏ－ＭｉｃｒｏｐｈｏｎｅＳｏｕｎｄｏ　Ｉｎｔｅｎｓｉｔｙ　Ｐｒｏｂｅ”にさらに詳細に説明されているように、音の強さは、下記の公式に従って、２つのマイクロフォンにより、周波数領域で検出されるクロス・スペクトルに関連したものである。

【０００４】

$$Ir(\omega) = -Im(G_{AB}) / \rho\omega\triangle r \qquad (1)$$

【０００５】時間領域における相当する式は、次の通り※30※である。

$$I_r = -(p_a + p_b) \int_t (p_a - p_b)\, dt / (2\rho\triangle r) \qquad (2)$$

ここで：

ρは、音響媒体の密度、

ωは、ラジアンで表示の周波数、

△ｒは、マイクロフォンの有効離隔距離、

Ｐａは、マイクロフォンａにおける音圧、

Ｐｂは、マイクロフォンｂにおける音圧、

Ｉｒは、ａからｂの方向における音の強さ、

$\int_t (x)\, dt$はｘの時間積分，

$Im(G_{AB})$は、ＰａとＰｂの間におけるクロス・スペクトルの虚部を表している。

【０００６】上記の詳細な計算を行うのに必要な測定は、２つの異なるタイプの計器、すなわち、高速フーリエ変換（ＦＦＴ）スペクトル解析器及び実時間オクターブ解析器を利用して実施することが出きる。各計器とも、それぞれ、長所と欠点を備えている。ヒューレット・パッカード社製のＨＰ３５６６０Ａ及び３５６６５ＡのようなＦＦＴ解析器は、デジタル方式でアナログ入力信号のサンプリングを行い、サンプリングしたデータの高速フーリエ変換を実施して、そのスペクトル成分を求める働きをする。フーリエ解析の結果は、一連のスペクトル係数であり、それぞれが、複数の周波数「ビン」のそれぞれに対応している。ＦＦＴ解析器技法の欠点は、フーリエ解析によるスペクトル成分の分解で得られる周波数ビンは、周波数の間隔が均等ということにある。例えば、０～１０ＫＨｚの間のスペクトルに関する４００ビンの解析によって、それぞれ、２５Ｈｚに相当するビンが得られる。こうした分解能は、高周波には十分以上であるが、低周波には、不十分である。一般に、低周波で十分な分解能を得るには、より狭い周波数範囲（０～１ＫＨｚ）に対する第２のＦＦＴ測定が必要になる。次に、２つの測定結果を組み合わせて、最終結果が得られる。しかし、この手順では、単一の計器が同時に両方の測定を実施できないので、問題がある。それどころか、組み合わせた測定値は、２つの異なる時点に行った２つの異なる測定値を反映したものになる。この非実時間操作は、受け入れることができない場合がよくある。

【０００７】対照的に、実時間オクターブ解析器は、一連の帯域フィルタ（しばしば、サンプル値データ・シス

テムが用いられる）を利用して、スペクトル成分を求める。これらのフィルタは、対数的に間隔をあけた周波数（しばしば、１／３オクターブのステップ）を中心におき、これによって、周波数が低くなるにつれて、分解能もますます精密の度を増してゆくことになる。（人間の耳は、対数式に音を知覚するので、オクターブをベースにした解析が一般的な測定モードになっている。）Ｂｒｕｅｌ＆Ｋｊａｅｒ２１３３及びＨＰ３５６６５Ａを含む、いくつかの実時間オクターブ解析器が知られている。（ＦＦＴ解説器として上述のＨＰ３５６６５Ａは、ＦＦＴ及び実時間オクターブ解析能力を組み合わせて、単一の計器としたものである。）前記Ｂ＆Ｋ２１３３の場合、音の強さは、帯域フィルタの出力を組み合わせ、上述の式（２）に記載の操作に対するサンプル値データのもう一方を実施することによって計算される。

【０００８】ＦＦＴ及び実時間オクターブ解析測定のいずれかのセット・アップに必要なクロス・チャネル・データが有効であるためには、プローブ・マイクロフォンの特性を精確に整合させなければならない。振幅の整合は、比較的容易にとれる。（さらに、音の強さの計算は、振幅の不整合にそれほど影響されない。上記Ｐｏｐｅの文献と対照されたい。）しかし、位相の整合は、困難である。

【０００９】２つの主たるメカニズムが、他の点では同じマイクロフォンの間において位相の不整合を生じさせることになる。第１のメカニズムは、ダイヤフラムの減衰における変動性に関するものである。この現象は、とりわけ、上述の数百Ｈｚの周波数において顕著である。製造公差のため、他の点では同一のマイクロフォンが、１キロヘルツで測定すると、位相応答に数度の差を示すのが一般的である。マイクロフォンの位相の不整合を生じさせる第２のメカニズムは、音響低域フィルタを有効に形成する、マイクロフォンのダイヤフラム後方のキャビティとこのキャビティから大気中への通風穴との間における相互作用である。マイクロフォン間における静圧等化の差によって、他の点では整合のとれているマイクロフォン間において、２０ヘルツで３度までの位相差異を生じる可能性がある。（一般に、２０ヘルツは、音の強さの測定にとって低周波数の限界である。）

【００１０】１対のマイクロフォンが、音の強さの測定にうまく利用されるためには、それらの位相特性が、一般に、問題となる測定スパン全体にわたって０．３度未満に納まるように整合しなければならない。（この０．３度という数字は、用途に極めて依存するものである。１度までの差異が、許容できる場合も有り得る。また、０．１度未満の差異が必要とされる場合も有り得る。上述のように、現在の製造プロセスで確実にマイクロフォンを整合させることのできるのは、この近似しきい値要件より１桁上の、約３度以内でしかない。従って、マイクロフォンの整合は、退屈な手動選択プロセスによって行わなければならないのが普通である。

【００１１】一般に、低周波における位相エラーは、音の強さを測定する上で最もクリティカルである。上述のように、音の強さは、ベクトル測定値である。測定される音の方向は、同じ音が２つのマイクロフォンに到達する位相遅延によって決まる。低周波の場合、音の波長が数メートルの長さになり、マイクロフォンは１センチメートルないし２センチメートルしか間隔をあけて配置できないので、オフ軸の音のベクトルに関連したｃｏｓφ位相遅延は、極めてわずかであり、わずかな位相エラーによってさえもマスクされてしまうことになる（ここで、φは、強さのベクトルとマイクロフォンの中心線との角度である）。低周波位相エラーの問題を解決できれば、手動で高周波応答を整合することで、マイクロフォンを選択する残りの仕事は、かなり簡単になる。実時間オクターブ解析器の場合、マイクロフォン自体を精密に整合させなければならないが、ＦＦＴ解析器の場合、解析器内部の補正率を利用してマイクロフォンの整合をトルコとができる。すなわち、１対のマイクロフォンの間の位相エラーを周波数領域で測定し、マイクロフォン間のクロス・スペクトルにこのエラー項の複素共役をかけて、補正を実施することができる。（１対のマイクロフォンの間における位相エラーを測定するため、いくつかの異なる方法を用いることができる。典型的なのは、１９７８年版、Ｊｏｕｒｎａｌ ｏｆ ｔｈｅ Ａｃｏｕｓｔｉｃａｌ Ｓｏｃｉｅｔｙ ｏｆ Ａｍｅｒｉｃａ 第６４巻、第６号、１６１３～１６１６頁のＣｈｕｎｇ，Ｊ．Ｙ．による“Ｃｒｏｓｓ－Ｓｐｅｃｔｒａｌ Ｍｅｔｈｏｄ ｏｆ ＭｅａｓｕｒｉｎｇＡｃｏｕｓｔｉｃ Ｉｎｔｅｎｓｉｔｙ Ｗｉｔｈｏｕｔ Ｅｒｒｏｒ Ｃａｕｓｅｄ ｂｙ Ｉｎｓｔｒｕｍｅｎｔ Ｐｈａｓｅ Ｍｉｓｍａｔｃｈ”、及び、１９８５年版ＮＯＩＳＥ－ＣＯＮ ８５ Ｐｒｏｃｅｅｄｉｎｇｓ ４２３～４２８頁のＳｅｙｂｅｒｔ，Ａ．Ｆ．による“Ｍｅａｓｕｒｕｍｅｎｔ ｏｆＰｈａｓｅ Ｍｉｓｍａｔｈ Ｂｅｔｗｅｅｎ Ｔｗｏ Ｍｉｃｒｏｐｈｏｎｅｓ”に開示のものである。

【００１２】実時間オクターブ解析器は、これまで、このタイプの位相補正を受け入れることができなかった。従って、精密に整合のとれたマイクロフォンを作製するという方法論で、広範囲にわたる研究が実施されてきた。この研究の典型的な例が、１９８５年版、Ｐｒｏｃｅｅｄｉｎｇｓ ｏｆ ２ｎｄ Ｉｎｔｅｒｎａｔｉｏｎａｌ Ｃｏｎｇｒｅｓｓ ｏｎ Ａｃｏｕｓｔｉｃ Ｉｎｔｅｎｓｉｔｙ，Ｓｅｎｌｉｓ ５０～５７頁のＦｒｅｄｅｒｉｋｓｅｎ，Ｅ．による“Ｐｈａｓｅ Ｃｈａｒａｃｔｅｒｉｓｔｉｃｓ ｏｆ ＭｉｃｒｏｐｈｏｎｅｓｆｏｒＩｎｔｅｎｓｉｔｙ Ｐｒｏｂｅｓ”、及び、１９８６年版、Ｂｒｕｅｌ＆Ｋｊｒ Ｔｅｃｈｎｉｃａｌ Ｒｅｖｉｅｗ、第４号、１１～２１頁のＦｒ

ｅｄｅｒｉｋｓｅｎ，Ｅ．他による“Ｐｒｅｓｓｕｒｅ Ｍｉｃｒｏｐｈｏｎｅｓｆｏｒ Ｉｎｔｅｎｓｉｔｙ Ｍｅａｓｕｒｅｍｅｎｔｓ ｗｉｔｈ Ｓｉｇｎｉｆｉｃａｎｔｌｙ Ｉｍｐｒｏｖｅｄ Ｐｈａｓｅ Ｐｒｏｐｅｒｔｉｅｓ”である。後者の論文では、機械的補償、すなわち、ダイヤフラム後方に２つのキャビティを追加して、低周波の通気キャビティにおける共振効果を減衰させ、これによって、マイクロフォン間における位相変動性を最小限に抑えるという、低周波の位相不整合に対する解決策が提案されている。また、米国特許第４，８８７，６５０号及び第４，７７７，６５０号には、位相精度にとって最適なマイクロフォンに関する開示もある。有効ではあるが、こうしたアプローチでは、製造コストが大幅に増大することになる。

【００１３】

【発明の目的】従って，本発明の目的は，複数の測定チャネル間の位相エラーを補償できる簡単で廉価な方法を提供し，さらにそれらを音声等のスペクトル解析に応用して測定精度を向上せしめることにある。

【００１４】

【発明の概要】本発明によれば、実時間オクターブ解析器には、極及びゼロが所定の対をなすマイクロフォンに関連した低周波位相エラーを相殺するように選択された、適応時間領域位相補償フィルタが設けられている。従って、複雑な機械的補償案に頼らなくても、不整合なマイクロフォンの位相エラーを精確に等化することができる。望ましい実施例の場合、適応フィルタの伝達関数は、プローブ・マイクロフォンのＦＦＴスペクトル解析によって得られたクロス・スペクトル・データに基づく極／ゼロ曲線適合技法を利用して、求められる。本発明の以上の及びそれ以外の特徴及び利点については、添付の図面に関連した下記の詳細な説明から明らかになる。

【００１５】

【発明の実施例】図１を参照すると、本発明による１／３オクターブ解析器１０には、２つの測定チャネルＡ及びＢが含まれている。図示されているのは、チャネルＡだけである。チャネルＢは、後述のフィルタ１２が省かれている点を除けば、同じである。チャネルＡには、マイクロフォン１４、アナログ・アンチ・エイリアス・フィルタ１６、アナログ・デジタル（Ａ／Ｄ）変換器１８、及び、一連の解析段２０が含まれている。各解析段２０には、デジタル・アンチ・エイリアス低域フィルタ２２、デシメーション段２４、及び、３つの１／３オクターブ・デジタル帯域フィルタ２６／２６′／２６″が含まれている。

【００１６】マイクロフォン１４は、音の強さプローブを構成する２つの内の一方である。このマイクロフォンからの出力は、２５．６ＫＨｚを超えるスペクトル成分を減衰させるため、アナログ・アンチ・エイリアス・フィルタ１６によって低域フィルタリングを施される。この帯域に制限が加えられた信号は、次に、Ａ／Ｄ変換器１８によって、６５，５３６（216）回／秒のサンプリングが施される。

【００１７】Ａ／Ｄ変換器１８からのサンプル値信号が、第１の解析段２０ａに加えられ、この解析段で、１２．８ＫＨｚを超える成分を減衰させるため、フィルタ２２ａによって、入力サンプル値データに低域フィルタリングが施される。フィルタリングを施された信号は、デシメータ２４ａによってデシメートされる。（デシメーションは、ナイキストの判定基準の達成にとって必須のものではない周期的データ・サンプルを取り上げないようにする既知のプロセスである。この場合、デシメーションによって、通常は、元の６５，５３６ＫＨｚのサンプル・レートで１オクターブの解析を行うのに必要な計算力により、低い方の全てのフル・オクターブ帯域を解析することが可能になる。

【００１８】デシメートされた出力信号（この場合、毎秒３２，７６８のサンプルで構成される）が、３つのデジタル帯域フィルタ２６ａ、２６ａ′、及び、２６ａ″（アメリカ標準規格　Ｓ１．１１　１９８６　種類３、タイプ１－Ｄフィルタ）に加えられ、サンプル値スペクトルが、それぞれ、８０００、６２５０、及び、５０００Ｈｚを中心にした１／３オクターブの帯域に分解される。第１のデシメータ２４ａからの出力は、第２の解析段２０ｂによって用いられるサンプル値データを提供する。このデータは、低域フィルタリングを施され（今回は、６．４ＫＨｚで）、デシメートされ（１６，３８４サンプル／秒）、それぞれ、４０００、３１２５、及び、２５００Ｈｚを中心にした１／３オクターブのデジタル・フィルタ２６ｂ、２６ｂ′、及び、２６ｂ″に加えられる。たまたま、デシメーションによって、サンプル・レートが半分になったので、また、解析帯域も、２で関連づけられるので（アメリカ標準規格Ｓ１．１１　１９８６によれば、デジタル１／３オクターブ・フィルタのベース２システム）、フィルタ２６ｂ、２６ｂ′、及び、２６ｂ″の極とゼロの位置は、フィルタ２６ａ、２６ａ′、及び、２６ａ″と同じになる。この結果、解析段２０ａに用いられた同じ組をなす３つのフィルタが、解析段２０ｂにおいて複製されるだけのため、帯域フィルタの実現が大幅に容易化される。

【００１９】追加解析段２０ｃ、２０ｄ、及び、２０ｅが、同様に、次々と段階的に設けられる。それぞれ、先行フィルタの１／２の周波数で入力信号に低域フィルタリングを施し、２でデシメートし、デシメートされたデータを次に低い３つの１／３オクターブ帯域に分解する。解析帯域を１／２にすることと結びつくデータのデシメーションによって、同じ組をなす３つのフィルタを各順次解析段で利用することが可能になる。解析段２０ｆには、他の解析段にも、また、他の測定チャネル（「Ｂ」）にも見られないフィルタ１２が含まれてい

る。例示のフィルタ１２は、３００Ｈｚ未満の周波数におけるプローブ・マイクロフォン１４に関連した位相エラーを相殺することを目的とした、適応形、単位利得、時間領域、無限インパルス応答（ＩＩＲ）フィルタである。このフィルタの極及びゼロを決める特定の方法の詳細については、図２～図４に関連して、以下に解説し、示すものとする。

【００２０】第１に、２つのプローブ・マイクロフォンの間の位相エラーが、測定される。適合する方法につい＊

$$\exp(j\theta_e) = (G_{AB} / G^{S*}_{AB})^{1/2} \qquad (3)$$

ここで、$\exp(j\theta_e)$は、周波数の関数としての位相エラーであり、G^{S*}_{AB}はG^{S}_{AB}の複素共役である。同様に、位相エラーは、G_{AB}とG^{S*}_{AB}の乗算を行うことによって相殺され、位相の不整合によるエラーのないクロス・スペクトルの推定値が得られる。この推定値が、次に、式（１）に挿入される。

【００２２】図３には、曲線３０によって、図２に示すＧＡＢとＧＳ ＡＢのデータに関するこの計算の共役結果が示されている。エラー位相の共役は、補正フィルタ１２に関する基準として用いられる。曲線３０に対して任意に類似した補正フィルタ伝達関数を合成することが可能である。図３の曲線３２は、３つの極と３つのゼロを利用して合成された伝達関数を示している。こうした近似は、ほとんどの用途において十分である。曲線３０によって表されるようなデータから時間領域フィルタ伝達関数を合成するのに用いられる方法論は、何年も前から知られており、例えば、その開示が参考までに本書に組み込まれている、米国特許第４，６５４，８０８号、第４，６５４，８０９号、及び、第４，６５８，３６７号に示されている。最良の態様の場合、Ｃｈｕｎｇのマイクロフォンの位相測定が、これらの特許に開示の方法論を利用して、任意の曲線に合成されたＳ領域伝達関数を当てはめる組み込み能力を備えた、ＨＰ　３５６３Ａ　ＦＦＴ解析器によって行われる。曲線当てはめプロセスの出力は、合成された伝達関数の極とゼロを表したデータである。

【００２３】場合によっては、上述の特許に記載の、ＨＰ　３５６３Ａ解析器によって実現される曲線当てはめ方法論によって、Ｓ面の右手側に極が生じ、結果得られるフィルタが実現不能になる。この結果を回避するためには、前もって極の１つを０Ｈｚ（ＤＣ）に固定するの※

$$H(z) = A\Pi^{3}_{r=1}(1-c_r z^{-1}) / \Pi^{3}_{k=1}(1-d_k z^{-1}) = \Sigma^{3}_{r=0} b_r z^{-r} / (1-\Sigma^{3}_{k=0} a_k z^{-k}) \qquad (4)$$

Ｈ（ｚ）に関する極－ゼロ式は、c_rをゼロ位置に等しくなるようにセットし、d_kを極位置にセットすることによって、書くことができる。いったんＺ領域になると、伝達関数は、多項式の形に拡張することができる。次に、ｚの多項係数を利用して、正準直接形式ＩＩと呼ばれる構成の合成補正フィルタを実現することができる。補正フィルタの非単位利得は、Ａの倍数でスケーリ

＊ては、Ｃｈｕｎｇ他による解説がある。Ｃｈｕｎｇの方法は、基本的に、クロス・スペクトルをいったん測定し（測定スペクトルＧＡＢが得られる）、それから、マイクロフォンを切り替えてもう１度測定する（測定スペクトルＧＳＡＢが得られる）、回路スイッチング技法である。このスイッチングによって、実際の強さによる位相応答を分割によって相殺し、下記の式におけるプローブ位相エラー項だけが残るようにすることが可能になる。

【００２１】

※が有効であることが分かった。（さらに、曲線当てはめプロセスから均一な利得係数を生じさせるのに、極とゼロの配置を前もって割りつけることができる。マイクロフォンの通気路とキャビティの相互作用によって生じる、音響低域フィルタのカット・オフ周波数に、追加固定極位置を選択するための物理的正当性を見いだすことができる。）極の１つがＳ面の原点に拘束されている、図示実施例の場合、極／ゼロ合成機能によって、下記の極とゼロが生じた（この場合、それぞれが、実Ｓ軸に沿って存在する）。

【００２４】

極	ゼロ
０	－０．８９６２
－１．１	－６．０
－１３５	－１３２．３３

【００２５】ＨＰ３５６３Ａは、通常、Ｓ面において合成される伝達関数を求める。ただし、ＨＰ３５６３Ａ解析器における自動操作としても実現される機能である、双線形変換を利用すれば、Ｚ領域への変換は簡単である。２００Ｈｚのプリワープ（ｐｒｅ－ｗａｒｐ）周波数を利用することによって、下記の極－ゼロ位置が、Ｚ領域で得られる。

【００２６】

極	ゼロ
１．０	．９５８６１８
．３５５５７５	．３６４２６９
．９９２２８３	．９９３７０８

【００２７】（プリワープは、Ｓ領域表現とＺ領域表現が一致する周波数を表している。）Ｚ領域のフィルタ伝達関数は、下記のように表される。

ングを施される。

【００２８】上述のフィルタに関するこうした実施例が、図５に示されており、無限インパルス応答（ＩＩＲ）トポロジとして認識される。この例の場合、係数は、実数であるが、その必要はない。また、係数は、再プログラミングが可能であって、さまざまなマイクロフォン、または、１組のマイクロフォンの時間に関するト

ラック・エラーに適応させることができることが望ましい。図４には、時間領域補正フィルタ１２による補正後のプローブ位相エラー（曲線３４）が示されており、それと補正前の位相エラー（曲線３０）の比較が行われている。明らかに、位相エラーは、大幅に抑えられている。（図２～図４の曲線は、ランダムに選択されたマイクロフォンを非無響環境に利用し、概念測定を立証することによって得られた。描かれたものを大幅に超えるエラー補正は、より制御された条件で得ることができる。）

【００２９】図１に戻ると、補正フィルタ１２からの出力が、それぞれ、２５０、２００、及び、１６０Ｈｚを中心とした、１／３オクターブ・フィルタ２６ｆ／２６ｆ′／２６ｆ″に加えられる。さらに、上述のような段階的構成が後続し、解析段２０ｇ及び２０ｈによって、最低域の２オクターブに関する出力が得られる（１２５／１００／８０Ｈｚ及び６２．５／５０／４０Ｈｚ）。音の強さを求めるため、チャネルＡの１／３オクターブ・フィルタ２６／２６′／２６″からの出力が、式（２）によって表された従来の方法で、チャネルＢの対応する出力と組み合わされる。望ましい実施例の場合、Ｍｏｔｏｒｏｌａ　ＭＣ５６００１のような汎用デジタル信号処理回路を利用して、時間領域補正フィルタ１２を含む解析段２０を実現することができる。代替実施例の場合、帯域フィルタだけが、汎用ＤＳＰ回路を用いて実現され、解析段のデシメータ２４及び低域フィルタ２２は、カスタム・ゲート・アレイによって実現される。

【００３０】出願人による最良の態様の場合、上述のマイクロフォンの較正及び補正ステップは、マルチ・チャネルＦＦＴ／実時間オクターブ解析計器にプログラムされた自動手順に組み込まれる。この手順を実行すると、該計器は、ユーザにマイクロフォンのセット・アップ命令を与え、ピンク・ノイズのソースを加えて、第１のクロス・スペクトルを捕捉し、ユーザに命じて、マイクロフォンを切り替えさせ（Ｃｈｕｎｇの手順が用いられるものと仮定する）、第２のクロス・スペクトルを捕捉し（再び、ＦＦＴモード）、式（３）の操作を行って、位相エラーを求め、特許の曲線の当てはめ手順に従って、補正フィルタの係数を合成し、図１の実時間オクターブ解析器アーキテクチャによって補正フィルタを実現し、それから、ユーザに対し、システムの較正及び補正が済み、音の強さの測定準備の整ったことを示す。その後較正されたシステムで得られるデータは、式（２）のサンプル値データの対の片方に従って処理され、実時間の音の強さの測定値が得られることになる。

【００３１】

【発明の効果】以上詳述したように，本発明の実施により，相互に位相エラーを持つ測定チャネルによる測定での位相エラーの補正が精密におこなえる。望ましい実施例に関連して、本発明の解説及び図解を行ってきたが、もちろん、こうした原理から逸脱することなく、構成及び細部において、本発明に修正を加えることが可能である。例えば、本発明の例示は、低周波にちょうど最適の時間領域補正フィルタに関連して行ったが、同じ原理によって、任意の広帯域幅にわたって補正が可能であることは明らかである。（全測定帯域の補正に適応するフィルタの場合、４つ以上の極とゼロがのぞましい。）同様に、本発明の例示は、単一の適応時間領域補正フィルタを用いた装置に関連して行ったが、他の実施例の場合、複数のこうしたフィルタを利用するのが有効な場合もある。複数の直列補正フィルタによって、位相補正曲線のより精密な適合が可能になる場合もある。１つのフィルタで、高周波の位相エラーを補正し（低周波では、効果がない）、もう１つのフィルタで、低周波の位相エラーを補正することも可能でる。代替案として、複数フィルタは、連鎖をなす解析段において直列にではなく、各解析段２０のデシメータ２４とフィルタ２６／２６′／２６″の間に１つずつ挿入して用いることも可能である。さらに、また、複数フィルタは、各測定チャネルに１つ以上配置して利用することも可能であり、この場合、有効に実現されるのは、フィルタ伝達関数間の位相差である。

【００３２】本発明の例示は、３極／ゼロ補正フィルタに関連して行ってきたが、他の状況には、他のフィルタが適合するということになるのは、明らかである。例えば、マイクロフォンがかなり精密に整合されている場合、単一項の補正フィルタで十分かもしれない。同様に、本発明の例示は、Ｚ領域トポロジを発生する後続の双線形関数を備えた、Ｓ領域曲線フィッタに関連して行ったが、他の実施例では、ＨＰ　３５６３Ａに見られるようなＺ領域曲線フィッタによって、極とゼロを直接決めることもできる。最後に、本発明の例示は、時間領域補正フィルタに関する適応ＩＩＲフィルタ・トポロジに関連して行ったが、他の位相補償フィルタ・トポロジを容易に用いることができるのも明らかである。

【図面の簡単な説明】

【図１】本発明の一実施例の実時間３分の１オクターブ解析装置の概略ブロック図である。

【図２】標準状態と切り換えた状態での整合していない一対のマイクロフォンによってサンプルされたピンク・ノイズのクロス・スペクトラムを周波数の関数としてプロットしたグラフである。

【図３】図２から決定されたマイクロフォンの位相エラーのグラフと３個の極と３個のゼロにより合成した補正カーブのグラフである。

【図４】図１の装置により補正した後のマイクロフォンの位相エラーのグラフである。

【図５】図１の装置に用いる時間領域位相補正フィルタを正準直接形式ＩＩで実現した場合の該フィルタのブロック図である。

11　　12

【符号の説明】

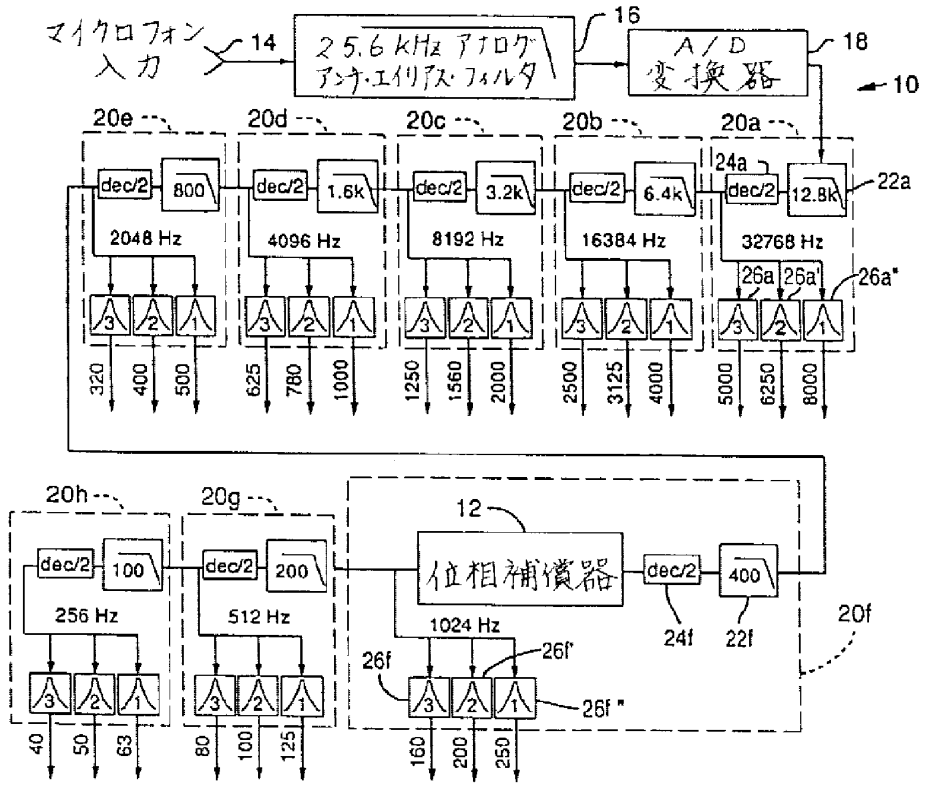
１４：マイクロフォン

１６：アナログ・アンチ・エイリアス．フィルタ　　＊

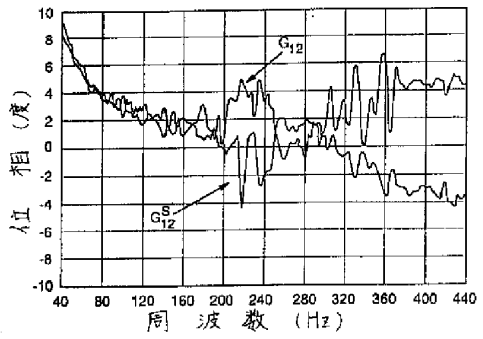
＊１８：Ａ／Ｄ変換器

２０ａ，２０ｂ，・・・，２０ｈ：解析段

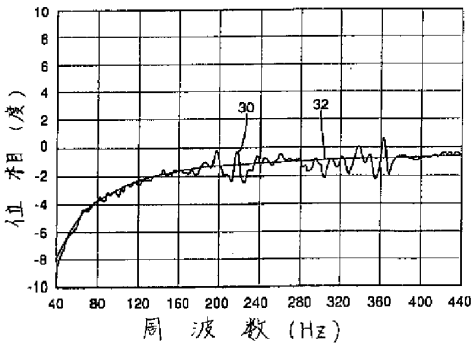
【図１】



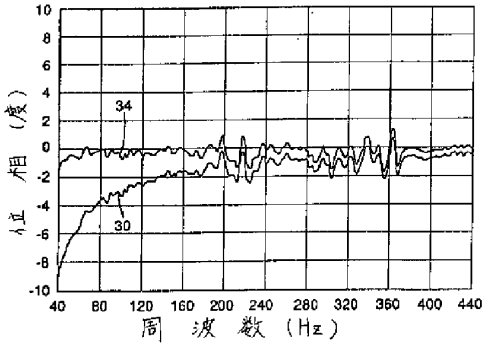
【図２】



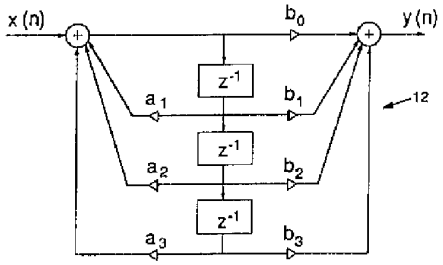
【図３】



【図４】


10
8
6
4
2
0
-2
-4
-6
-8
-10
位相（度）
34
30
40 80 120 160 200 240 280 320 360 400 440
周波数（Hz）

【図５】


x(n)
y(n)
b0
z-1
a1
b1
z-1
a2
b2
z-1
a3
b3
12

Electronic Acknowledgement Receipt	
EFS ID:	12306752
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	James William Drapinski
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	14-MAR-2012
Filing Date:	29-JAN-2007
Time Stamp:	19:52:52
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Transmittal Letter	Audience_3909US_IDS_Transmittal_Letter.pdf	78517 ef104761e89e9febf9d7bc5ecf3a81c5865e8235	no	3
Warnings:					
Information:					

2	Information Disclosure Statement (IDS) Form (SB08)	Audience_3909US_IDS_SB08_1.PDF	616698 b0bedc5e5cd6c9c5cc8663470be3210d9b8d98de	no	21
Warnings:					
Information:					
3	Foreign Reference	AUDIEN_FOR0002_JP5053587A.pdf	631496 ef4b9356e3e0a4de3f633049f9c770cbf8cf2642	no	5
Warnings:					
Information:					
4	Foreign Reference	AUDIEN_FOR0015_WO2003069499.pdf	3297041 4e895f54b5e331437eb5086ddfadee0592db19b2	no	31
Warnings:					
Information:					
5	Foreign Reference	AUDIEN_FOR0022_WO2010005493.pdf	4577978 7a06ac37e732f77e51d3bfb5f0471a7dacf8a74a	no	40
Warnings:					
Information:					
6	Foreign Reference	AUDIEN_FOR0023_WO2007140003.pdf	2953677 0a02b640200c72ba6e62af0b4083933706e52c85	no	28
Warnings:					
Information:					
7	Foreign Reference	AUDIEN_FOR0029_WO2003043374.pdf	2860538 dc41a2cbcf0b554b3acb9bf671422e8213c5efc3	no	26
Warnings:					
Information:					
8	Foreign Reference	AUDIEN_FOR0030_WO2001074118.pdf	2710354 40c6d34431aaa85986b77d3642ba9d8961372ed6	no	27
Warnings:					
Information:					
9	Foreign Reference	AUDIEN_FOR0031_WO2007081916.pdf	2649522 c4f2f81107591c7be77e692234ad322f1b4075de	no	24
Warnings:					
Information:					
10	Foreign Reference	AUDIEN_FOR0032_JP11249693.pdf	863960 dc650f86b1641d1e1ccb9805e04c15ec7e378b42	no	11
Warnings:					
Information:					

11	Foreign Reference	AUDIEN_FOR0033_JP10313497.pdf	3956921 ba4503a6ed860df11907f820f4a22f4fac3b7e1c	no	34
Warnings:					
Information:					
12	Foreign Reference	AUDIEN_FOR0034_JP2005195955.pdf	1692925 75bcfc75b573e8467cad18b059e096076bb009d8	no	19
Warnings:					
Information:					
13	Foreign Reference	AUDIEN_FOR0035_JP2005110127.pdf	331248 f5ce423273227d0f5346c28ae2b38bcb0c48323c	no	10
Warnings:					
Information:					
14	Foreign Reference	AUDIEN_FOR0036_WO2004010415.pdf	3796264 a039e8df50a92657ea90272c027fbf24843a3275	no	49
Warnings:					
Information:					
15	Foreign Reference	AUDIEN_FOR0037_JP2004053895.pdf	1703060 9f539c1ae00844806a1225d883fc6bd6db54f796	no	18
Warnings:					
Information:					
16	Foreign Reference	AUDIEN_FOR0038_JP2005518118.pdf	2142737 43e9ba664d8fbaf538874755b6295330ae1b2783	no	20
Warnings:					
Information:					
17	Foreign Reference	AUDIEN_FOR0039_WO03069499A1.pdf	3295477 4954426eb59582e356cf96d00793b1917ce065fd	no	31
Warnings:					
Information:					
18	Foreign Reference	AUDIEN_FOR0040_JP2004531767.pdf	2878572 0588cc4a0126bde8d50d0326f6155e2e805aedbc	no	52
Warnings:					
Information:					
19	Foreign Reference	AUDIEN_FOR0041_WO02103676A1.pdf	3637029 31b3f68ea6d1dc67ff5006ba0c43a0e44b042752	no	35
Warnings:					
Information:					

20	Foreign Reference	AUDIEN_FOR0042_JP2005148274.pdf	3208515 0e2c950deb55b3ef109568c7882b246adfc0351c	no	41
Warnings:					
Information:					
21	Foreign Reference	AUDIEN_FOR0043_WO02080362A1.pdf	3030915 041574859f69e7a404fe53d21da7c9bb89b88c1a	no	22
Warnings:					
Information:					
22	Foreign Reference	AUDIEN_FOR0044_JP2004533155.pdf	2735567 363dec174748078df18012be7719e0f2454f633a	no	42
Warnings:					
Information:					
23	Foreign Reference	AUDIEN_FOR0045_JP05172865.pdf	344411 e61c94150d9c5a3845561e1e1aad2b3d5bbaa43c	no	9
Warnings:					
Information:					
			Total Files Size (in bytes):	53993422	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano et al.
APPLICATION NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
GROUP ART UNIT:	2614
CONFIRMATION NO.:	7721
ATTY.DKT.NO.:	PA3909US

MAIL STOP AMENDMENT
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. § 1.56, and 1.97-1.98

SIR:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08, listing several references. The Examiner is requested to make these references of official record in the application.

The references cited may be material to examination of the application and are submitted in compliance with the Applicant's duty of disclosure as defined by 37 C.F.R. § 1.56. Additionally, the Applicant wishes to make the Examiner aware of, and invite the Examiner to consider, as is appropriate, the following office actions from the following U.S. patent applications:

- Office Actions mailed October 27, 2003, February 10, 2004, December 17, 2004, and April 20, 2005 and Notice of Allowance mailed October 26, 2005, in U.S. patent application number 09/534,682, filed March 24, 2000.
- Office Actions mailed October 6, 2005 and March 28, 2006 in U.S. patent application number 10/177,049, filed June 21, 2002.
- Office Actions mailed May 3, 2005, October 19, 2005, May 17, 2006, November 16, 2006, and June 15, 2007 in U.S. patent application number 09/993,442, filed November 13, 2001.
- Office Action mailed October 8, 2003, and Notice of Allowance mailed February 24, 2004, in U.S. patent application number 10/004,141, filed November 14, 2001.
- Office Actions mailed May 9, 2003, June 26, 2006, and February 23, 2007 and Notice of Allowance mailed June 4, 2003, in U.S. patent application number 10/074,991, filed February 13, 2002.
- Office Action mailed June 13, 2006 in U.S. patent application number 10/840,201, filed May 5, 2004.
- Office Actions mailed December 13, 2006 and June 13, 2007 in U.S. patent application number 10/613,224, filed July 3, 2003.
- Office Actions mailed March 30, 2010, September 13, 2010, and March 30, 2011 in U.S. patent application number 11/343,524, filed January 30, 2006.
- Office Actions mailed August 5, 2008, January 21, 2009, September 3, 2009, May 10, 2011, and October 24, 2011 in U.S. patent application number 11/441,675, filed May 25, 2006.
- Office Actions mailed August 18, 2010 and April 28, 2011 in U.S. patent application number 11/825,563, filed July 6, 2007.
- Office Action mailed July 28, 2011 and Notice of Allowance mailed March 1, 2012 in U.S. patent application number 12/072,931, filed February 29, 2008.
- Office Action mailed September 14, 2011 and Notice of Allowance mailed January 27, 2012 in U.S. patent application number 12/004,897, filed December 21, 2007.
- Office Action mailed September 8, 2011 in U.S. patent application number 12/004,899, filed December 21, 2007.
- Office Actions mailed March 28, 2011, October 26, 2011, and February 15, 2012 in U.S. patent application number 12/228,034, filed August 8, 2008.
- Office Action mailed October 3, 2011 and Notice of Allowance mailed February 23, 2012 in U.S. patent application number 12/004,788, filed December 21, 2007.
- Office Action mailed September 8, 2011 in U.S. patent application number 12/004,896, filed December 21, 2007.
- Office Action mailed December 28, 2011 in U.S. patent application number 12/288,228, filed October 16, 2008.
- Office Action mailed October 17, 2011 in U.S. patent application number 12/077,436, filed March 18, 2008.
- Notice of Allowance mailed March 1, 2012 in U.S. patent application number 12/080,115, filed March 31, 2008

- Office Action mailed November 14, 2011 in U.S. patent application number 12/215,980, filed June 30, 2008.
- Office Actions mailed July 13, 2011 and November 16, 2011 in U.S. patent application number 12/217,076, filed June 30, 2008.
- Office Action mailed September 1, 2011 in U.S. patent application number 12/286,909, filed October 2, 2008.
- Office Action mailed November 15, 2011 in U.S. patent application number 12/286,995, filed October 2, 2008.
- Office Action mailed December 6, 2011 in U.S. patent application number 12/319,107, filed December 31, 2008.
- Office Action mailed December 30, 2011 in U.S. patent application number 12/422,917, filed April 13, 2009.

No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

The Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600.

Respectfully submitted,
Carlos Avendano et al.

March 14, 2012

By: /James W. Drapinski/

James W. Drapinski, Reg. No. 46,242
Carr & Ferrell *LLP*
120 Constitution Drive
Menlo Park, CA 94025
TEL: (650) 812-3400
FAX: (650) 812-3444

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Patent citation information please click the Add button.						Add

U.S.PATENT APPLICATION PUBLICATIONS						Remove
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Published Application citation information please click the Add button.						Add

FOREIGN PATENT DOCUMENTS							Remove	
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button							Add	

NON-PATENT LITERATURE DOCUMENTS			Remove
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	1	BOLL, STEVEN F. "Suppression of Acoustic Noise in Speech using Spectral Subtraction", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27, No. 2, April 1979, pgs. 113-120.	☐
	2	BOLL, STEVEN F. "Suppression of Acoustic Noise in Speech Using Spectral Subtraction", Dept. of Computer Science, University of Utah Salt Lake City, Utah, April 1979, pp. 18-19.	☐
	3	DAHL, MATTIAS et al., "Simultaneous Echo Cancellation and Car Noise Suppression Employing a Microphone Array", 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, April 21-24, pgs. 239-242.	☐
	4	"ENT 172." Instructional Module. Prince George's Community College Department of Engineering Technology. Accessed: October 15, 2011. Subsection: "Polar and Rectangular Notation". <http://academic.ppgcc.edu/ent/ent172_instr_mod.html>	☐
	5	FULGHUM, D. P. et al., "LPC Voice Digitizer with Background Noise Suppression", 1979 IEEE International Conference on Acoustics, Speech, and Signal Processing, pgs. 220-223.	☐
	6	GRAUPE, DANIEL et al., "Blind Adaptive Filtering of Speech from Noise of Unknown Spectrum Using a Virtual Feedback Configuration", IEEE Transactions on Speech and Audio Processing, March 2000, vol. 8, no. 2, pgs. 146-158.	☐
	7	HAYKIN, SIMON et al. "Appendix A.2 Complex Numbers." Signals and Systems. 2nd Ed. 2003. pg. 764	☐
	8	HERMANSKY, HYNEK "Should Recognizers Have Ears?", In Proc. ESCA Tutorial and Research Workshop on Robust Speech Recognition for Unknown Communication Channels, pages 1-10, France 1997.	☐
	9	HOHMANN, V. "Frequency Analysis and Synthesis Using a Gammatone Filterbank", ACTA Acustica United with Acustica, 2002, Vol. 88, pages 433-442.	☐
	10	JEFFRESS, LLOYD A. et al. "A Place Theory of Sound Localization," Journal of Comparative and Physiological Psychology, 1948, vol. 41, p. 35-39.	☐
	11	JEONG, HYUK et al., "Implementation of a New Algorithm Using the STFT with Variable Frequency Resolution for the Time-Frequency Auditory Model", J. Audio Eng. Soc., April 1999, Vol. 47, No. 4., pages 240-251.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	12	KATES, JAMES M. "A Time-Domain Digital Cochlear Model", IEEE Transactions on Signal Processing, December 1991, Vol. 39, No. 12, pages 2573-2592.	☐
	13	LAZZARO, JOHN et al., "A Silicon Model of Auditory Localization," Neural Computation Spring 1989, vol. 1, pgs. 47-57, Massachusetts Institute of Technology.	☐
	14	LIPPMANN, RICHARD P. "Speech Recognition by Machines and Humans", Speech Communication, July 1997, vol. 22, no. 1, pgs. 1-15	☐
	15	MARTIN, RAINER "Spectral Subtraction Based on Minimum Statistics", in Proceedings Europe. Signal Processing Conf., 1994, pgs. 1182-1185	☐
	16	MITRA, SANJIT K. Digital Signal Processing: a Computer-based Approach. 2nd Ed. 2001. pgs. 131-133	☐
	17	WATTS, LLOYD Narrative of Prior Disclosure of Audio Display on Feb. 15, 2000 and May 31, 2000.	☐
	18	COSI, PIERO et al. (1996), "Lyon's Auditory Model Inversion: a Tool for Sound Separation and Speech Enhancement," Proceedings of ESCA Workshop on 'The Auditory Basis of Speech Perception,' Keele University, Keele (UK), 15-19 July, 1996, pp. 194-197.	☐
	19	RABINER, LAWRENCE R. et al. "Digital Processing of Speech Signals", (Prentice-Hall Series in Signal Processing). Upper Saddle River, NJ: Prentice Hall, 1978.	☐
	20	WEISS, RON et al., "Estimating Single-Channel Source Separation Masks: Revelance Vector Machine Classifiers vs. Pitch-Based Masking", Workshop on Statistical and Perceptual Audio Processing, 2006.	☐
	21	SCHIMMEL, STEVEN et al., "Coherent Envelope Detection for Modulation Filtering of Speech," 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1, no. 7, pgs 221-224.	☐
	22	SLANEY, MALCOM, "Lyon's Cochlear Model", Advanced Technology Group, Apple Technical Report #13, Apple Computer, Inc., 1988, pp 1-79	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	23	SLANEY, MALCOM, et al. "Auditory Model Inversion for Sound Separation," 1994 IEEE International Conference on Acoustics, Speech and Signal Processing, April 19-22, vol. 2, pgs 77-80.	☐
	24	SLANEY, MALCOM. "An Introduction to Auditory Model Inversion", Interval Technical Report IRC 1994-014, http://coweb.ecn.purdue.edu/~maclom/interval/1994-014/, September 1994, accessed on July 6, 2010	☐
	25	SOLBACH, LUDGER "An Architecture for Robust Partial Tracking and Onset Localization in Single Channel Audio Signal Mixes", Technical University Hamburg-Harburg, 1998.	☐
	26	STAHL, V. et al., "Quantile Based Noise Estimation for Spectral Subtraction and Wiener Filtering," 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, June 5-9, vol.3, pgs. 1875-1878	☐
	27	Syntrillium Software Corporation, "Cool Edit User's Manual", 1996, pp. 1-74.	☐
	28	TCHORZ, JURGEN et al., "SNR Estimation Based on Amplitude Modulation Analysis with Applications to Noise Suppression", IEEE Transactions on Speech and Audio Processing, Vol. 11, No. 3, May 2003, pgs. 184-192.	☐
	29	WATTS, LLOYD, "Robust Hearing Systems for Intelligent Machines," Applied Neurosystems Corporation, 2001, pp. 1-5.	☐
	30	YOO, HEEJONG et al., "Continuous-Time Audio Noise Suppression and Real-Time Implementation", 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, May 13-17, pgs. IV3980-IV3983.	☐
	31	International Search Report dated June 8, 2001 in Application No. PCT/US01/08372	☐
	32	International Search Report dated April 3, 2003 in Application No. PCT/US02/36946	☐
	33	International Search Report dated May 29, 2003 in Application No. PCT/US03/04124	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	34	International Search Report and Written Opinion dated October 19, 2007 in Application No. PCT/US07/00463	☐
	35	International Search Report and Written Opinion dated April 9, 2008 in Application No. PCT/US07/21654	☐
	36	International Search Report and Written Opinion dated September 16, 2008 in Application No. PCT/US07/12628	☐
	37	International Search Report and Written Opinion dated October 1, 2008 in Application No. PCT/US08/08249	☐
	38	International Search Report and Written Opinion dated May 11, 2009 in Application No. PCT/US09/01667	☐
	39	International Search Report and Written Opinion dated August 27, 2009 in Application No. PCT/US09/03813	☐
	40	International Search Report and Written Opinion dated May 20, 2010 in Application No. PCT/US09/06754	☐
	41	FAST COCHLEA TRANSFORM, US Trademark Reg. No. 2,875,755 (August 17, 2004).	☐
	42	DAHL, MATTIAS et al., "Acoustic Echo and Noise Cancelling Using Microphone Arrays", International Symposium on Signal Processing and its Applications, ISSPA, Gold coast, Australia, August 25-30, 1996, pgs. 379-382.	☐
	43	DEMOL, M. et al. "Efficient Non-Uniform Time-Scaling of Speech With WSOLA for CALL Applications", Proceedings of InSTIL/ICALL2004 - NLP and Speech Technologies in Advanced Language Learning Systems - Venice 17-19 June, 2004.	☐
	44	LAROCHE, JEAN. "TIME AND PITCH SCALE MODIFICATION OF AUDIO SIGNALS", in "Applications of Digital Signal Processing to Audio and Acoustics", The Kluwer International Series in Engineering and Computer Science, Volume 437, pgs. 279-309, 2002.	☐

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	45	MOULINES, ERIC et al., "Non-Parametric Techniques for Pitch-Scale and Time-Scale Modification of Speech", Speech Communication, Vol. 16, pgs. 175-205, 1995.	☐
	46	VERHELST, WERNER, "Overlap-Add Methods for Time-Scaling of Speech", Speech Communication Vol. 30, pgs 207-221, 2000.	☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/James W. Drapinski/	Date (YYYY-MM-DD)	2012-03-14
Name/Print	James W. Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/08372

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : H04R 25/00
US CL : 381/312, 320
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 381/312, 320

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
IEEE Explore Web site

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4,536,844 A (LYON) 20 August 1985 (20.08.1985), abstract, figure 1, column 2 lines 22-25	1-4, 12-14, 17, 21, 22
X	US 5,473,759 A (SLANEY et al) 05 December 1995 (05.12.1995), figures 2 and 4	1-4, 12-14, 17, 21, 22
X	KATES, JAMES M. "A Time-Domain Digital Cochlear Model" IEEE Transactions on Signal Processing, December 1991, Volume 39 Number 12, pages 2573-2592	1-4, 12-14, 17, 21, 22

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
18 May 2001 (18.05.2001)

Date of mailing of the international search report
08 JUN 2001

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Forester W. Isen
Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/36946

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : H04R 5/00
US CL : 381/17
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 381/17,1,309,310,311,300; 379/406.01

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,223,090 B1 (BRUNGART) 24 April 2001 (24.04.2001), figures 3, 6, 7	1-20
A	US 5,729,612 A (ABEL et al) 17 March 1998 (17.03.1998), see entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
12 February 2003 (12.02.2003)

Date of mailing of the international search report
03 APR 2003

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Forrester W. Isen

Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/04124

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G06F 17/00; H03G 5/00; G10L 19/00
US CL : 700/94; 381/98; 704/500
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 700/94; 381/98; 704/500

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	US 2002/0147595 A1 (BAUMGARTE) 10 October 2002 (10.10.2002), figures1, 2 and 7, page 3 and 5.	1,3-5, 8-12, 14-19, 22-25
A	US 6,513,004 B1 (RIGAZIO et al) 28 January 2003 (28.01.2003), entire document.	1-25
A	US 4,536,844 A (LYON) 20 August 1985 (20.08.1985), entire document.	1-25
A	US 5,473,759 A (SLANEY et al) 05 December 1995 (05.12.1995), entire document.	1-25
A	KATES, JAMES M. "A Time-Domain Digital Cochlear Model" IEEE Transactions of Siganl Processing, December 1991, Volume 39 Number 12, pages 2573-2592.	1-25

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 May 2003 (05.05.2003)	29 MAY 2003
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Forester W Isen Telephone No. (703) 305-4700

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US07/00463

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04R 25/00 (2007.01)
USPC - 381/313
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - H04R 25/00 (2007.01)
USPC - 381/313, 094.300, 91-92, 111

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatBase: ft=microphone% ft=filter*~ ft=nois*~ ft=energ*~ ft=estimat*~ ft=speech*~ or speak*~ ft=smooth*~ ; MicroPatent (Databases-USG USA EPA EPB WO JP DEG DEA DET DEU GBA FRA): Patent/Publication No. US20010016020 US20010031053 US20020009203 US20020041693 Us20020080980 US20020106092 US20020116187 US20030039369 us6430295

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 2001/0031053 A1 (FENG et al) 18 October 2001 (18.10.2001) entire document	1, 2, 5, 8, 12-17, 19, 20 --- 3, 4, 6, 7, 9-11, 18
Y	US 6,430,295 B1 (HANDEL et al) 06 August 2002 (06.08.2002) entire document	3, 4, 6, 7, 9-11, 18
A	US 2003/0072382 A1 (RALEIGH et al) 17 April 2003 (17.04.2003) entire document	1-20
A	US 6,717,991 B1 (GUSTAFSSON et al) 06 April 2004 (06.04.2004) entire document	1-20
A	US 2003/0099345 A1 (GARTNER et al) 29 May 2003 (29.05.2003) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 August 2007	19 OCT 2007
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To: DANIEL KLOKE
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CALIFORNIA 94303

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	19 OCT 2007
Applicant's or agent's file reference: PA3503PCT	FOR FURTHER ACTION: See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US07/00463	05 January 2007	05 January 2006

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - H04R 25/00 (2007.01)
USPC - 381/313

Applicant: AUDIENCE, INC.

1. This opinion contains indications relating to the following items:

- [X] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [X] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	20 August 2007	Blaine Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/00463

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. type of material

      ☐ a sequence listing

      ☐ table(s) related to the sequence listing

   b. format of material

      ☐ on paper

      ☐ in electronic form

   c. time of filing/furnishing

      ☐ contained in the international application as filed

      ☐ filed together with the international application in electronic form

      ☐ furnished subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/00463

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	3, 4, 6, 7, 9-11, 18	YES
	Claims	1, 2, 5, 8, 12-17, 19, 20	NO
Inventive step (IS)	Claims	None	YES
	Claims	1-20	NO
Industrial applicability (IA)	Claims	1-20	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1, 2, 5, 8, 12-17, 19 and 20 lack novelty under PCT Article 33(2) as being anticipated by Feng et al (US 2001/0031053 A1).

Regarding claim 1, Feng et al disclose a method for enhancing speech, comprising: receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone (paragraph 76 "acoustic sensor 22, 24 signals to microphones may be employed"); determining a filter estimate during a frame, the filter estimate based on a noise estimate of the primary acoustic signal (paragraphs 145-150 "filtering applying estimated acoustic sources of comparable levels and noise in a given time frame generally results in better intelligibility of the desired signal"), an energy estimate of the primary acoustic signal, and an inter-microphone level difference based on the primary and secondary acoustic signals (paragraphs 134-135 "energy signal difference estimation accounts for secondary relationships"); and applying the filter estimate to the primary acoustic signal to produce a speech estimate (paragraphs 7-8 "the speech system employs a windowing technique to estimate a sound source").

Regarding claim 2, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses determining an energy estimate for each of the acoustic signals during the frame (paragraphs 134-135 "each of the current frame sources reflects the energy signal acoustic estimation").

Regarding claim 5, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 2, and further discloses using the energy estimates to determine the inter-microphone level difference for the frame (paragraphs 134-135 "energy signal difference estimation accounts for secondary relationships"; paragraphs 145-150 "applying estimated levels in a given time frame generally results in better intelligibility of the desired signal").

Regarding claim 8, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses wherein the noise estimate is based on an energy estimate of the primary acoustic signal and the inter-microphone level difference (paragraphs 134-135 "energy signal difference estimation accounts for secondary relationships"; paragraphs 145-150 "applying estimated acoustic sources of comparable levels and noise in a given time frame generally results in better intelligibility of the desired signal").

Regarding claim 12, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses converting the speech estimate to a time domain (paragraph 66 "conversion to a time domain representation").

Regarding claim 13, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses outputting the speech estimate to a user (paragraph 67 "Output device 90 supplies the speaker").

Regarding claim 14, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses wherein the filter estimate is based on a Wiener filter (paragraph 7 "the system employs a Wiener filter").

Regarding claim 15, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses a system for enhancing speech on a device, comprising: a primary microphone configured to receive a primary acoustic signal; a secondary microphone located a distance away from the primary microphone and configured to receive a secondary acoustic signal; and an audio processing engine configured to enhance speech received at the primary microphone, the audio processing engine comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the primary acoustic signal and an inter-microphone level difference; and a filter module configured to determine a filter estimate to be applied to the primary acoustic signal to generate a filtered acoustic signal, the filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference (paragraph 156 "a processing system is configured for generating a desired acoustic signal from a first source and an interfering acoustic signal from a second source that is spaced apart from the first source").

Regarding claim 16, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 15, and further discloses wherein the audio processing engine further comprises an energy module configured to determine energy estimates for a frame of the primary and secondary acoustic signals (paragraphs 134-135 "energy estimation accounts for secondary relationships"; paragraphs 145-150 "applying estimated levels in a given time frame generally results in better intelligibility of the desired signal").

(Continued in Supplemental Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/00463

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box No. V
2. Citations and explanations:

Regarding claim 17, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 15, and further discloses wherein the audio processing engine further comprises an inter-microphone level difference module configured to determine the intermicrophone level difference (paragraphs 134-135 "energy difference estimation accounts for secondary relationships"; paragraphs 145-150 "applying estimated levels generally results in better intelligibility of the desired signal").

Regarding claim 19, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 15, and further discloses wherein the audio processing engine further comprises a masking module configured to determine the speech estimate (paragraphs 7-8 "the speech system employs a windowing technique to estimate a sound source").

Regarding claim 20, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, and further discloses a computer readable medium having embodied thereon a program, the program being executable by a machine to perform a method for enhancing speech on a device, the method comprising: receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone; determining an energy estimate over a frame for each of the acoustic signals; using the energy estimates to determine an inter-microphone level difference for the frame; generating a noise estimate based on the energy estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference; calculating a filter estimate based on the noise estimate and the inter-microphone level difference; and applying the filter estimate to the primary acoustic signal to produce a speech estimate (paragraph 71 "the present invention may be performed on variously arranged processing equipment configured with one or more modules combination not limited to programming memory").

Claims 3, 4, 6, 7, 9-11 and 18 lack an inventive step under PCT Article 33(3) as being obvious over Feng et al (US 2001/0031053 A1) in view of Handel et al (US 6,430,295 B1).

Regarding claim 3, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 2, and further discloses approximate filtering by applying equation sets (paragraphs 145-146 "approximate"), approximate acoustic sources (paragraphs 145-146 "acoustic sources") and energy estimated from signal processing frequency to time (paragraphs 62-66 "energy can be estimated from signal processing frequency to time"). Feng et al lacks the teaching of wherein the energy estimate of the primary acoustic signal is approximated as the E1 (t, w) equation recited in the claim. However, Handel et al teaches time-varying equations, such as the Recursive Least Squares RLS algorithm used toward an approximation to estimate prior art systems noise energy concentrated around one or more frequencies (column 6, lines 1-67, column 7, lines 1-10 "time-varying") as in the E1 (t, w) equation claimed. Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the energy estimate of the primary acoustic signal is approximated as the E1 (t, w) equation recited in the claim into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Regarding claim 4, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 2, and further discloses approximate filtering by applying equation sets (paragraphs 145-146 "approximate"), approximate acoustic sources (paragraphs 145-146 "acoustic sources") and energy estimated from signal processing frequency to time (paragraphs 62-66 "energy can be estimated from signal processing frequency to time"). Feng et al lacks the teaching of wherein the energy estimate of the secondary acoustic signal is approximated as the E2 (t, w) equation recited in the claim. However, Handel et al teaches time-varying equations, such as the Recursive Least Squares RLS algorithm used toward an approximation to estimate prior art systems noise energy concentrated around one or more frequencies (column 6, lines 1-67, column 7, lines 1-10 "time-varying") as in the E2 (t, w) equation claimed. Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the energy estimate of the secondary acoustic signal is approximated as the E2 (t, w) equation recited in the claim into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Regarding claim 6, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 5, and further discloses approximate filtering by applying equation sets (paragraphs 145-146 "approximate"), approximate acoustic sources (paragraphs 145-146 "acoustic sources") and energy estimated from signal processing frequency to time (paragraphs 62-66 "energy can be estimated from signal processing frequency to time"). Feng et al lacks the teaching of wherein the inter-microphone level difference is approximated by the ILD(t, w) equation recited in the claim. However, Handel et al teaches time-varying equations, such as the Recursive Least Squares RLS algorithm used toward an approximation to estimate prior art systems noise energy concentrated around one or more frequencies (column 6, lines 1-67, column 7, lines 1-10 "time-varying") as in the ILD(t, w) equation claimed. Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the inter-microphone level difference is approximated by the ILD(t, w) equation recited in the claim into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

(Continued in next Supplemental Box)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/00463

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Previous Supplemental Box:

Regarding claim 7, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 5, and further discloses approximate filtering by applying equation sets (paragraphs 145-146 "approximate"), approximate acoustic sources (paragraphs 145-146 "acoustic sources") and energy estimated from signal processing frequency to time (paragraphs 62-66 "energy can be estimated from signal processing frequency to time"). Feng et al lacks the teaching of wherein the inter-microphone level difference is approximated by the ILD(t, w) equation recited in the claim. However, Handel et al teaches time-varying equations, such as the Recursive Least Squares RLS algorithm used toward an approximation to estimate prior art systems noise energy concentrated around one or more frequencies (column 6, lines 1-67, column 7, lines 1-10 "time-varying") as in the ILD(t, w) equation claimed. Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the inter-microphone level difference is approximated by the ILD(t, w) equation recited in the claim into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Regarding claim 9, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 8, and further discloses approximate filtering by applying equation sets (paragraphs 145-146 "approximate"), approximate acoustic sources (paragraphs 145-146 "acoustic sources") and energy estimated from signal processing frequency to time (paragraphs 62-66 "energy can be estimated from signal processing frequency to time"). Feng et al lacks the teaching of wherein the noise estimate is approximated as the N(t,w) equation recited in the claim. However, Handel et al teaches time-varying equations, such as the Recursive Least Squares RLS algorithm used toward an approximation to estimate prior art systems noise energy concentrated around one or more frequencies (column 6, lines 1-67, column 7, lines 1-10 "time-varying") as in the N(t,w) equation claimed. Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the noise estimate is approximated as the N(t,w) equation recited in the claim into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Regarding claim 10, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 1, but lacks the teaching of smoothing the filter estimate prior to applying the filter estimate to the primary acoustic signal. However, Handel et al teaches smoothing the filter estimate prior to applying the filter estimate to the primary acoustic signal (column 9, lines 66-67, column 10, lines 1-65 "An exemplary system using signal processing components is smoothed by a first order recursion estimated using an acoustic scenario that can be applied using an arbitrary number of sources and sensors"). Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate smoothing the filter estimate prior to applying the filter estimate to the primary acoustic signal into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Regarding claim 11, Feng et al discloses the claimed invention features as discussed in the lack of inventive step of claim 10, and further discloses approximate filtering by applying equation sets (paragraphs 145-146 "approximate"), approximate acoustic sources (paragraphs 145-146 "acoustic sources") and energy estimated from signal processing frequency to time (paragraphs 62-66 "energy can be estimated from signal processing frequency to time"). Feng et al lacks the teaching of wherein the smoothing is approximated as the M(t,w) equation recited in the claim. However, Handel et al teaches time-varying equations, such as the Recursive Least Squares RLS algorithm used toward an approximation to estimate prior art systems noise energy concentrated around one or more frequencies (column 6, lines 1-67, column 7, lines 1-10 "time-varying") and smoothing the filter estimate prior to applying the filter estimate to the primary acoustic signal (column 9, lines 66-67, column 10, lines 1-65 "An exemplary system using signal processing components is smoothed by a first order recursion estimated using an acoustic scenario that can be applied using an arbitrary number of sources and sensors") as in the M(t,w) equation claimed. Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the smoothing is approximated as the M(t,w) equation recited in the claim into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Regarding claim 18, Feng et al discloses the claimed invention features as discussed in the lack of novelty of claim 15, but lacks the teaching of wherein the audio processing engine further comprises a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal. However, Handel et al teaches wherein the audio processing engine further comprises a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal (column 9, lines 66-67, column 10, lines 1-65 "An exemplary system using signal processing components is smoothed by a first order recursion estimated using an acoustic scenario that can be applied using an arbitrary number of sources and sensors"). Handel et al further shows that this is relevant for filtering (column 5, lines 35-60 "filtering"). Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to incorporate wherein the audio processing engine further comprises a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal into the invention of Feng et al as taught in Handel et al in order to provide filtering (column 5, lines 35-60 "filtering").

Claims 1-20 meet the criteria set out in PCT Article 33(4), and thus have industrial applicability because the subject matter claimed can be made or used in industry.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US07/21654

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04R 25/00 (2008.01)
USPC - 381/313
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8) - H04R 25/00 (2008.01)
USPC - 381/313, 312

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Database: MicroPatent

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2003/0147538 A1 (ELKO) 07 August 2003 (07.08.2003) entire document	1-26
Y	US 2003/0169891 A1 (RYAN et al) 11 September 2003 (11.09.2003) entire document	1-26

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 March 2008	09 APR 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2005)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
Susan Yee
2200 Geng Road
Palo Alto, California 94303-3220

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	09 APR 2008

Applicant's or agent's file reference	FOR FURTHER ACTION
PA3909PCT	See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US07/21654	09 October 2007	10 October 2006

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - H04R 25/00 (2008.01)
USPC - 381/313

Applicant AUDIENCE, INC.

1. This opinion contains indications relating to the following items:

- [x] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [x] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	06 March 2008	Blaine Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/21654

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. type of material

      ☐ a sequence listing

      ☐ table(s) related to the sequence listing

   b. format of material

      ☐ on paper

      ☐ in electronic form

   c. time of filing/furnishing

      ☐ contained in the international application as filed

      ☐ filed together with the international application in electronic form

      ☐ furnished subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/21654

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	1-26	YES
	Claims	None	NO
Inventive step (IS)	Claims	None	YES
	Claims	1-26	NO
Industrial applicability (IA)	Claims	1-26	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1-26 lack an inventive step under PCT Article 33(3) as being obvious over Elko (US 2003/0147538 A1) in view of Ryan et al. (US 2003/0169891 A1).

Referring to claims 1, 14, and 26, Elko discloses a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal (paragraphs 0008 and 0095); a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal (paragraph 0026), but is silent on an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference. However, in disclosing a low-noise directional microphone system (abstract), Ryan et al. teaches an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference (paragraphs 0038 and 0040). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Referring to claims 2 and 15, Elko (as discussed in lack of inventive step of claims 1 and 14) is silent wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level-equalized secondary electric signal. However, Ryan et al. teaches wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level-equalized secondary electric signal (paragraph 0070). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include the limitation wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between the primary electric signal and a delayed and level-equalized secondary electric signal in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Referring to claims 3 and 16, Elko (as discussed in lack of inventive step of claims 1 and 14) discloses wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary signal adjusted by the gain (paragraphs 0099 and 0102).

Referring to claims 4 and 17, Elko (as discussed in lack of inventive step of claims 3 and 16) is silent wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal. However, Ryan et al. teaches wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal (paragraph 0042). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal in the invention of Elko as taught by Ryan et al. in order to provide a well-balanced value for the gain.

Referring to claims 5 and 18, Elko (as discussed in lack of inventive step of claims 1 and 14) discloses wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the level-equalized secondary electric signal and a delayed primary electric signal (paragraphs 0099 and 0102).

Referring to claim 6, Elko (as discussed in lack of inventive step of claim 1) discloses a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal (paragraph 0063).

(Continued in Supplemental Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/21654

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box No. V
2. Citations and explanations:

Referring to claims 7 and 20, Elko (as discussed in lack of inventive step of claims 1 and 14) is silent on an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal. However, Ryan et al. teaches an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal (paragraph 0039). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Referring to claims 8 and 21, Elko (as discussed in lack of inventive step of claims 1 and 14) discloses a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference (paragraph 0051).

Referring to claim 9, Elko (as discussed in lack of inventive step of claim 1) is silent on a filter module configured to determine a filter estimate to be applied to the primary acoustic signal. However, Ryan et al. teaches a filter module configured to determine a filter estimate to be applied to the primary acoustic signal (paragraph 0028). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include a filter module configured to determine a filter estimate to be applied to the primary acoustic signal in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Referring to claim 10, Elko (as discussed in lack of inventive step of claims 9) is silent on a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal. However, Ryan et al. teaches a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal (paragraph 0028). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Referring to claim 11, Elko (as discussed in lack of inventive step of claim 1) discloses a masking module configured to determine a speech estimate (paragraph 0052).

Referring to claim 12, Elko (as discussed in lack of inventive step of claim 11) discloses a frequency synthesis module configured to convert the speech estimate into a time domain for output (paragraph 0052).

Referring to claims 13 and 25, Elko (as discussed in lack of inventive step of claims 1 and 14) discloses wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal (paragraph 0105).

Referring to claim 19, Elko (as discussed in lack of inventive step of claim 14) discloses wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal (paragraph 0029).

Referring to claim 22, Elko (as discussed in lack of inventive step of claim 21) is silent on determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference. However, Ryan et al. teaches determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference (paragraph 0039-0040). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Referring to claim 23, Elko (as discussed in lack of inventive step of claim 22) discloses producing a speech estimate by applying the filter estimate to the primary acoustic signal (paragraph 0052).

Referring to claim 24, Elko (as discussed in lack of inventive step of claim 22) is silent on smoothing the filter estimate. However, Ryan et al. teaches smoothing the filter estimate (paragraph 0028). Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to include smoothing the filter estimate in the invention of Elko as taught by Ryan et al. in order to provide a low-noise directional microphone system that is particularly well suited for use in a digital hearing instrument (paragraph 0002).

Claims 1-26 meet the criteria set out in PCT Article 33(4), and thus have industrial applicability because the subject matter claimed can be made or used in industry.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US07/12628

A. CLASSIFICATION OF SUBJECT MATTER
IPC: **H03G 5/00**(2006.01)

USPC: 381/98
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 381/98,99,97,103,94.2,94.3;333/132

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4,137,510 (IWAHARA) 30 January 1979 (30.01.1979), column 1, line 64- column 2, line 31.	1-23

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 August 2008 (14.08.2008)	16 SEP 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer Valerie Bell for Kile O. Blair Telephone No. (571) 272-2000

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
SUSAN YEE
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CA 94303

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	16 SEP 2008
Applicant's or agent's file reference PA3583PCT	**FOR FURTHER ACTION** See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US07/12628	24 May 2007 (24.05.2007)	25 May 2006 (25.05.2006)

International Patent Classification (IPC) or both national classification and IPC

IPC: **H03G 5/00**(2006.01)
USPC: 381/98

Applicant

AUDIENCE, INC

1. This opinion contains indications relating to the following items:

☒	Box No. I	Basis of the opinion
☐	Box No. II	Priority
☐	Box No. III	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
☐	Box No. IV	Lack of unity of invention
☒	Box No. V	Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
☐	Box No. VI	Certain documents cited
☐	Box No. VII	Certain defects in the international application
☐	Box No. VIII	Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis(b)* that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/ US	Date of completion of this opinion	Authorized officer
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	14 August 2008 (14.08.2008)	Valerie Bell-Harris for Kile O. Blair Telephone No. (571) 272-2000

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/12628

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:
   - ☒ the international application in the language in which it was filed
   - ☐ a translation of the international application into _____, which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).
2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))
3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. type of material
      - ☐ a sequence listing
      - ☐ table(s) related to the sequence listing

   b. format of material
      - ☐ on paper
      - ☐ in electronic form

   c. time of filing/furnishing
      - ☐ contained in the international application as filed.
      - ☐ filed together with the international application in electronic form.
      - ☐ furnished subsequently to this Authority for the purposes of search.

4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/12628

Box No. V Reasoned statement under Rule 43 *bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims 1-23	YES
	Claims NONE	NO
Inventive step (IS)	Claims NONE	YES
	Claims 1-23	NO
Industrial applicability (IA)	Claims 1-23	YES
	Claims NONE	NO

2. Citations and explanations:

Please See Continuation Sheet

Form PCT/ISA/237 (Box No. V) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/12628

Supplemental Box
In case **the space in any of the preceding boxes is not sufficient.**

V. 2. Citations and Explanations:

Claims 1-23 an inventive step under PCT Article 33(3) as being obvious over Iwahara (US Pat. No. 4,137,510).

Regarding claim 1, Iwahara teaches a method for processing audio signals, comprising: filtering an input signal with a filter of a filter cascade to
produce a first filtered signal; subtracting the first filtered signal from the input signal to derive a first sub-band signal (the input signal is sent though delay 21-1 to the adder 24-1 where the signal filtered by filter 22-1 is subtracted from the input signal as shown in Fig. 17); filtering the first filtered signal with a next complex-valued filter of the filter cascade to produce a next filtered signal; and subtracting the next filtered signal from the first filtered signal to derive a next sub-band signal (the process is repeated at adder 24-2 with low pass filter 22-2 and the signal through delay 21-2b, Fig. 17).

Although Iwahara does not explicitly teach that the filters are complex valued, it would have been obvious toone of ordinary skill in the art since such a modification would have yielded predictable results

Regarding claim 2, Iwahara teaches the method of claim 1.

Although Iwahara does not explicitly teach the feature wherein the complex-valued filter and the next complex-valued filter are single pole, complex-valued filters, the use of such filters would have been an obvious design choice to one of ordinary skill in the art.

Regarding claim 3, Iwahara teaches the method of claim 1 further comprising performing phase alignment on one or more of the sub-band signals (phase shifter 51-1 performs phase alignment on the sub-band signal of LPF 22-1, Fig. 17).

Regarding claim 4, Iwahara teaches the method of claim 3 further comprising disposing of an imaginary portion of the one or more phase aligned sub-band signals (since Iwahara does not utilize complex value, the imaginary portion of the signals is already disposed, therefore it would have been obvious to one of ordinary skill in the art to dispose of the imaginary portions if complex valued filters are used since Iwahara does not include the imaginary portions).

Regarding claim 5, Iwahara teaches the method of claim 1 further comprising performing amplitude compensation on one or more of the sub-band signals (coefficient applying operation circuit which affects the amplitude characteristics of the signal, Col. 2, lines 19-31).

Regarding claim 6, Iwahara teaches the method of claim 1 further comprising performing a time delay on one or more of the sub-band signals for cross-sub-band alignment (delay 21-1, Fig. 17).

Regarding claim 7, Iwahara teaches the method of claim 6 further comprising summing the delayed one or more sub-band signals to generate a reconstructed audio signal (output terminal 32 in Fig. 7a which is equivalent to the signal obtained by combining the out signals of the output terminals 23, 25, 26, and 27 of Fig. 6).

Regarding claim 8, Iwahara teaches the method of claim 1.

Although Iwahara does not explicitly teach the feature of pre-processing the input signal prior to filtering the input signal with

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US07/12628

Supplemental Box
In case the space in any of the preceding boxes is not sufficient.

the complex-valued filter of the filter cascade, it would have been obvious to one of ordinary skill in the art to pre-process the input signal with the motivation of providing adequate processing and control in the system.

Regarding claim 9, Iwahara teaches the method of claim 1 further comprising modifying one or more of the sub-band signals based on an analysis path from the filter cascade (the sub band signals are modified in phase and delay in their respective paths, Fig. 17).

Regarding claim 10, Iwahara teaches the method of claim 1 wherein the sub-band signals are frequency components of the input signal (frequency bands, col. 1, lines 64- col. 2, line 2).

Regarding claim 11, Iwahara teaches a system for processing an audio signal, comprising: an audio processing engine comprising a filter cascade of filters configured to derive a plurality of sub-band signals from an input signal the set of filters arranged in the filter cascade whereby an output of each filter is passed to a next filter in the filter cascade (the output of filter 22-1 inverts at adder 24-1 and it fed to filter 22-2, Fig. 17).

Although Iwahara does not explicitly teach that the filters are complex valued, it would have been obvious toone of ordinary skill in the art since such a modification would have yielded predictable results

Regarding claim 12, Iwahara teaches the system of claim 11.

Although Iwahara does not explicitly teach the feature wherein the complex-valued filter and the next complex-valued filter are single pole, complex-valued filters, the use of such filters would have been an obvious design choice to one of ordinary skill in the art.

Regarding claim 13, Iwahara teaches the system of claim 11 wherein the audio processing engine further comprises a reconstruction module configured to perform phase alignment on one or more of the sub-band signals (phase shifter 51-1 performs phase alignment on the sub-band signal of LPF 22-1, Fig. 17).

Regarding claim 14, Iwahara teaches the system of claim 11 wherein the audio processing engine further comprises a reconstruction module configured to perform amplitude compensation on one or more of the sub-band signals (coefficient applying operation circuit which affects the amplitude characteristics of the signal, Col. 2, lines 19-31).

Regarding claim 15, Iwahara teaches the system of claim 11 wherein the audio processing engine further comprises a reconstruction module configured to perform a time delay on one or more of the sub-band signals (delay 21-1, Fig. 17).

Regarding claim 16, Iwahara teaches the system of claim 11 wherein the audio processing engine further comprises a modification module configured to modify one or more of the sub- band signals based on an analysis path from the filter cascade (delay 21-2a, Fig. 17).

Regarding claim 17, Iwahara teaches the system of claim 11.

Although Iwahara does not explicitly teach the that the system of claim 11 further comprises a conditioning module configured to pre-process the input signal prior to filtering the input signal with the filter cascade, pre-processing and conditioning of an input signal prior to being filtered would have been an obvious modification to one of ordinary skill in the art with the motivation of removing noise or any other well known signal enhancement or repairing method.

Claim 18 is substantially similar to claim 12 and lacks inventive step for the same reasons.

Claim 19 is substantially similar to claim 2 and lacks inventive step for the same reasons

Claim 20 is substantially similar to claim 3 and lacks inventive step for the same reasons

Claim 21 is substantially similar to claim 5 and lacks inventive step for the same reasons

Claim 22 is substantially similar to claim 6 and lacks inventive step for the same reasons

Regarding claim 20, Iwahara teaches the machine-readable medium of claim 18.

Although Iwahara does not explicitly teach the that the system of claim 18 further comprises a conditioning module configured to pre-process the input signal prior to filtering the input signal with the filter cascade, pre-processing and conditioning of an input signal prior to being filtered would have been an obvious modification to one of ordinary skill in the art with the motivation of removing noise or any other well known signal enhancement or repairing method.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/08249

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04B 15/00 (2008.04)
USPC - 381/94.2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8): H04B 15/00 (2008.04)
USPC: 381/94.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
IPC(8): G10L 21/00 (2008.04)
USPC: 381/94.1, 381/94.2, 704/205, 704/227

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWest (USPS, USPT, USOC, EPAB, JPAB), DialogPro, GoogleScholar; adaptive noise suppression primary acoustic signal speech distortion estimate loss control adjust enhancement filter gain masks output subtract calculate spectrum power classifying inter-level secondary comfort audibility sensor generator energy instructions

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2005/0027520 A1 (Mattila et al.) 03 February 2005 (03.02.2005), para [0018]-[0019]; [0025]; [0043]; [0113]; [0124]-[0128]; [0130]; [0138]-[0141]; [0179]; [0182]-[0183]; [0197]; [0213]; [0216]-[0219]; [0233]-[0234]; [0239]	1-20
Y	US 2006/0072768 A1 (Schwartz et al.) 06 April 2006 (06.04.2006), para [0021]-[0023]; [0060]-[0063]; [0067]	1-20
Y	US 2003/0128851 A1 (Furuta) 10 July 2003 (10.07.2003), para [0023]; [0073]; [0075]; [0090]-[0093]; [0120]	2-4,13
Y	US 2003/0103632 A1 (Goubran et al.) 05 June 2003 (05.06.2003), para [0024]-[0025]; [0032]	16-17
Y	US 2006/0184363 A1 (McCree et al.) 17 August 2006 (17.08.2006), claim 9; para [0011]-[0013]	18-20
A	US 2005/0288923 A1 (Kok) 29 December 2005 (29.12.2005), entire document	1-20
A	US 6,810,273 B1 (Mattila et al.) 26 October 2004 (26.10.2004), entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 September 2008 (16.09.2008)	01 OCT 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
SUSAN YEE
2200 GENG ROAD
PALO ALTO, CA 94303

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	01 OCT 2008
Applicant's or agent's file reference: PA4190PCT	FOR FURTHER ACTION: See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US 08/08249	03 July 2008 (03.07.2008)	06 July 2007 (06.07.2007)

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - H04B 15/00 (2008.04)
USPC - 381/94.2

Applicant AUDIENCE, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	25 September 2008 (25.09.2008)	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 08/08249

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:

   a. type of material

      ☐ a sequence listing

      ☐ table(s) related to the sequence listing

   b. format of material

      ☐ on paper

      ☐ in electronic form

   c. time of filing/furnishing

      ☐ contained in the international application as filed

      ☐ filed together with the international application in electronic form

      ☐ furnished subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

Form PCT/ISA/237 (Box No. I) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 08/08249

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	1-20	YES
	Claims	None	NO
Inventive step (IS)	Claims	None	YES
	Claims	1-20	NO
Industrial applicability (IA)	Claims	1-20	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1, 5-12 and 14-15 lack an inventive step under PCT Article 33(3) as being obvious over US 2005/0027520 A1 to Mattila et al. (hereinafter 'Mattila') in view of US 2006/0072768 A1 to Schwartz et al. (hereinafter 'Schwartz').

Regarding claim 1, Mattila teaches a method for adaptively suppressing noise, comprising:
determining a speech loss distortion estimate based on the primary acoustic signal (see para [0113]; [0124]-[0125]); generating a plurality of gain masks based on the speech loss distortion estimate utilizing an enhancement filter (see para [0125]; [0182]-[0183]); applying the plurality of gain masks to the primary acoustic signal to generated a noise suppressed signal (see para [0125]; [0138]-[0141]; [0197]); and outputting the noise suppressed signal (see para [0233]-[0234]; [0239]), but does not specifically teach receiving a primary acoustic signal. Schwartz teaches a method for noise suppression comprising receiving a primary acoustic signal (see para [0021]-[0023]). It would have been obvious to one of ordinary skill in the art to utilize the method comprising receiving a primary acoustic signal of Schwartz and to combine it with the method of Mattila because it would provide acoustic signal noise reduction, thereby increasing the robustness and cost -effectiveness of the method.

Regarding claim 5, Mattila in view of Schwartz teaches the method of claim 1. Mattila further teaches a method further comprising classifying noise and speech in the primary acoustic signal (see para [0019]).

Regarding claim 6, Mattila in view of Schwartz teaches the method of claim 1. Schwartz further teaches a method further comprising determining an inter-level difference between the primary acoustic signal and a secondary acoustic signal (see para [0060]-[0063]; [0067]).

Regarding claim 7, Mattila in view of Schwartz teaches the method of claim 1. Mattila further teaches a method further comprising generating and applying a comfort noise to the noise suppressed signal prior to output (see para [0018]; [0025]; [0213]).

Regarding claim 8, Mattila in view of Schwartz teaches the method of claim 7. Mattila further teaches wherein generating the comfort noise comprises setting the comfort noise to a level that is just above a level of audibility (see para [0043]; [0179]; [0216]-[0219]).

Regarding claim 9, Mattila in view of Schwartz teaches the method of claim 1. Mattila further teaches a method further comprising deriving control signals to adjust the enhancement filter based on the speech loss distortion estimate (see para [0125]-[0128]).

Regarding claim 10, Mattila teaches a system for adaptively suppressing noise in a primary acoustic signal, comprising: an adaptive intelligent suppression generator configured to adaptively generate a plurality of gain masks to apply to the primary acoustic signal (see para [0125]; [0182]-[0183]); and a mask module configured to apply the plurality of gain masks to the primary acoustic signal to generate a noise suppressed signal (see para [0125]; [0138]-[0141]; [0197]), but does not specifically teach an acoustic sensor configured to receive the primary acoustic signal. Schwartz teaches a system for noise suppression comprising an acoustic sensor configured to receive the primary acoustic signal (see para [0021]-[0023]). It would have been obvious to one of ordinary skill in the art to utilize the system comprising an acoustic sensor configured to receive the primary acoustic signal of Schwartz and to combine it with the system of Mattila because it would provide a sensor for acoustic signal noise reception in addition to acoustic signal noise reduction, thereby increasing the robustness and cost-effectiveness of the system.

Regarding claim 11, Mattila in view of Schwartz teaches the system of claim 10. Mattila further teaches a system further comprising a comfort noise generator configured to generate a comfort noise to apply to noise suppressed signal (see para [0018]; [0025]; [0213]).

Regarding claim 12, Mattila in view of Schwartz teaches the system of claim 10. Mattila further teaches wherein the adaptive intelligent suppression generator comprises a speech distortion control module configured to determine a speech distortion estimate from the primary acoustic signal (see para [0113]; [0124]-[0125]), and configured to derive control signals to adjust computation of the gain masks based on the speech distortion estimate (see para [0125]-[0128]).

Regarding claim 14, Mattila in view of Schwartz teaches the system of claim 10. Schwartz further teaches a system further comprising an inter-level difference module configured to generate an inter-level difference used by the adaptive intelligent suppression generator to determine a speech distortion estimate (see para [0060]-[0063]; [0067]).

--- continued in Supplemental Box ---

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 08/08249

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

Box No. V

2. Citations and explanations:

Regarding claim 15, Mattila in view of Schwartz teaches the system of claim 10. Mattila further teaches wherein the adaptive intelligent suppression generator comprises a compute enhancement filter module configured to adaptively generate the gain masks based on a speech distortion estimate (see para [0125]-[0128]).

Claims 2-4 and 13 lack an inventive step under PCT Article 33(3) as being obvious over Mattila in view of Schwartz and further in view of US 2003/0128851 A1 (Furuta).

Regarding claim 2, Mattila in view of Schwartz teaches the method of claim 1, but does not specifically teach wherein determining a speech loss distortion estimate comprises subtracting a calculated noise spectrum from a power spectrum of the primary acoustic signal. Furuta teaches a method for noise suppression wherein determining a speech loss distortion estimate comprises subtracting a calculated noise spectrum from a power spectrum of the primary acoustic signal (see para [0023]; [0073]; [0075]). It would have been obvious to one of ordinary skill in the art to utilize the method comprising subtracting a calculated noise spectrum from a power spectrum of the primary acoustic signal of Furuta and to combine it with the method of Mattila in view of Schwartz because it would provide optimized noise suppression, thereby increasing the robustness and cost-effectiveness of the method.

Regarding claim 3, Mattila in view of Schwartz and further in view of Furuta teaches the method of claim 2. Furuta further teaches a method further comprising calculating the noise spectrum (see para [0090]-[0093]; [0120]).

Regarding claim 4, Mattila in view of Schwartz and further in view of Furuta teaches the method of claim 2. Mattila further teaches a method further comprising calculating the power spectrum of the primary acoustic signal (see para [0130]; [0139]).

Regarding claim 13, Mattila in view of Schwartz teaches the system of claim 10, but does not specifically teach a system further comprising a noise estimate module configured to generate a noise power spectrum used by the adaptive intelligent suppression generator to determine a speech distortion estimate. Furuta teaches a system for noise suppression comprising a noise estimate module configured to generate a noise power spectrum used by the adaptive intelligent suppression generator to determine a speech distortion estimate (see para [0090]-[0093]; [0120]). It would have been obvious to one of ordinary skill in the art to utilize the system comprising a noise estimate module configured to generate a noise power spectrum of Furuta and to combine it with the system of Mattila in view of Schwartz because it would provide a more accurate speech distortion estimate, thereby increasing the efficiency and cost-effectiveness of the system.

Claims 16-17 lack an inventive step under PCT Article 33(3) as being obvious over Mattila in view of Schwartz and further in view of US 2003/0103632 A1 to Goubran et al. (hereinafter 'Goubran').

Regarding claim 16, Mattila in view of Schwartz teaches the system of claim 10, but does not specifically teach a system further comprising an energy module configured to generate a primary spectrum for the primary acoustic signal. Goubran teaches an adaptive sound masking system comprising an energy module configured to generate a primary spectrum for the primary acoustic signal (see para [0024]). It would have been obvious to one of ordinary skill in the art to utilize the system comprising an energy module configured to generate a primary spectrum for the primary acoustic signal of Goubran and to combine it with the system of Mattila in view of Schwartz because it would provide more accurate noise suppression, thereby increasing the efficiency and cost-effectiveness of the system.

Regarding claim 17, Mattila in view of Schwartz and further in view of Goubran teaches the system of claim 16. Goubran further teaches wherein the energy module is further configured to generate a power spectrum for a second acoustic signal received by a second acoustic sensor (see para [0024]-[0025]; [0032]).

Claims 18-20 lack an inventive step under PCT Article 33(3) as being obvious over US 2006/0184363 A1 to McCree et al. (hereinafter 'McCree') in view of Mattila and further in view of Schwartz.

Regarding claim 18, McCree teaches a machine readable medium having embodied thereon a program, the program providing instructions for a method for adaptively suppressing noise (see claim 9; para [0011]-[0013]), but does not specifically teach the method comprising: receiving a primary acoustic signal; determining a speech loss distortion estimate based on the primary acoustic signal; generating a plurality of gain masks based on the speech loss distortion estimate utilizing an enhancement filter; applying the plurality of gain masks to the primary acoustic signal to generated a noise suppressed signal; and outputting the noise suppressed signal. Mattila teaches a method for noise suppression comprising determining a speech loss distortion estimate based on the primary acoustic signal (see para [0113]; [0124]-[0125]); generating a plurality of gain masks based on the speech loss distortion estimate utilizing an enhancement filter (see para [0125]; [0182]-[0183]); applying the plurality of gain masks to the primary acoustic signal to generated a noise suppressed signal (see para [0125]; [0138]-[0141]; [0197]); and outputting the noise suppressed signal (see para [0233]-[0234]; [0239]), but does not specifically teach receiving a primary acoustic signal. Schwartz teaches a method for noise suppression comprising receiving a primary acoustic signal (see para [0021]-[0023]). It would have been obvious to one of ordinary skill in the art to utilize the method comprising determining a speech loss distortion estimate, generating and applying a plurality of gain masks, and outputting the noise suppressed signal of Mattila and the method comprising receiving a primary acoustic signal of Schwartz and to combine it with the machine readable medium of McCree because it would provide more accurate noise suppression, thereby increasing the efficiency and cost-effectiveness of the machine readable medium.

--- continued in Supplemental Box ---

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 08/08249

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:

Box No. V

2. Citations and explanations:

Regarding claim 19, McCree in view of Mattila and further in view of Schwartz teaches the machine readable medium of claim 18. Mattila further teaches wherein the method further comprises deriving control signals to adjust the enhancement filter based on the speech loss distortion estimate (see para [0125]-[0128]).

Regarding claim 20, McCree in view of Mattila and further in view of Schwartz teaches the machine readable medium of claim 18. Mattila further teaches wherein the method further comprising generating and applying a comfort noise to the noise suppressed signal prior to output (see para [0018]; [0025]; [0213]).

Claims 1-20 have industrial applicability as defined by PCT Article 33(4) because the subject matter can be made or used in industry.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 09/01667

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04M 9/08 (2009.01)
USPC - 379/406.14
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC - 379/406.14

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 379/406.01, 406.14 (text search)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWest (PGPB, USPT, EPAB, JPAB); Google Scholar
Search terms used: envelop, echo, cancel, mask, suppress, subband, sub-band, noise, square, root. rms. far, end, near, floor, acoustic, gain

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,587,998 A (Velardo, JR. et al.). et al. 24 December 1996 (24.12.1996), col. 2, ln. 58 through col. 3, ln. 50, col. 6, ln. 8-19, col. 8, ln. 13-25	1-23
A	US 5,825,320 A (Miyamori et al.) 20 October 1998 (20.10.1998), entire document	1-23
A	US 2004/0057574 A1 (Faller) 25 March 2004 (25.03.2004), entire document	1-23
A	US 6,999,582 B1 (Popovic et al.) 14 February 2006 (14.02.2006), entire document	1-23

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 April 2009 (27.04.2009)	11 MAY 2009
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
STEPHEN BACHMANN
2200 GENG ROAD
PALO ALTO, CA 94303

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	11 MAY 2009

Applicant's or agent's file reference	FOR FURTHER ACTION
PA4546PCT	See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US 09/01667	16 March 2009 (16.03.2009)	18 March 2008 (18.03.2008)

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - H04M 9/08 (2009.01)
USPC - 379/406.14

Applicant AUDIENCE, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	27 April 2009 (27.04.2009)	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/01667

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:
   - [X] the international application in the language in which it was filed.
   - [ ] a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. [ ] This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:
   a. type of material
      - [ ] a sequence listing
      - [ ] table(s) related to the sequence listing

   b. format of material
      - [ ] on paper
      - [ ] in electronic form

   c. time of filing/furnishing
      - [ ] contained in the international application as filed
      - [ ] filed together with the international application in electronic form
      - [ ] furnished subsequently to this Authority for the purposes of search

4. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

Form PCT/ISA/237 (Box No. I) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/01667

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	NONE	YES
	Claims	1-23	NO
Inventive step (IS)	Claims	NONE	YES
	Claims	1-23	NO
Industrial applicability (IA)	Claims	1-23	YES
	Claims	NONE	NO

2. Citations and explanations:

Claims 1-23 lack novelty under PCT Article 33(2) as being anticipated by US 5,587,998 A to Velardo, JR. et al. (hereinafter 'Velardo, JR.').

Regarding claim 1, Velardo, JR. teaches a method for envelope-based acoustic echo cancellation, comprising: receiving a primary acoustic signal via a primary microphone and a far-end signal via a receiver (signals transmitted into the network at a near location are received and signals transmitted into the network from a far location are received by the same processing device, col. 2, ln. 58-67); performing frequency analysis on the primary acoustic signal and the far-end acoustic signal to obtain frequency sub-bands for the primary and far-end acoustic signals (the near input and the far input are each separately decomposed into plural sub-band components, col.3, ln. 3-4); applying the echo gain mask to the primary acoustic signal to generate a masked signal (a masked signal is produced, col. 3, ln. 5-50, col. 8, ln. 13-25); and outputting the masked signal (the masked signal is outputted, col. 3, ln. 5-50).

Regarding claim 2, Velardo, JR. teaches generating a noise gain mask based on at least the primary acoustic signal for each frequency sub-band (a noise mask and an echo mask are obtained by mixing the transmitted sub-band signal with a white-noise component, col. 3, ln. 23-42, col. 8, ln. 13-25).

Regarding claim 3, Velardo, JR. teaches wherein applying the echo gain mask to the primary acoustic signal to generate the masked signal comprises applying a combination of the echo gain mask and the noise gain mask to the primary acoustic signal to generate the mask signal (in order to mask the echo component, the transmitted sub-band signal is mixed with a white-noise component, col. 3, ln. 23-42, col. 8, ln. 13-25).

Regarding claim 4, Velardo, JR. teaches wherein applying the combination of the echo gain mask and noise gain mask comprises selecting a minimum between the echo gain mask and noise gain mask for each sub-band (a lower noise threshold level is selected, col. 3, ln. 23-42, col. 8, ln. 13-25).

Regarding claim 5, Velardo, JR. teaches wherein generating the noise gain mask comprises determining a substantially echo-free noise estimate (the mask is utilized to provide an echo-free noise, col. 3, ln. 23-42, col. 8, ln. 13-25).

Regarding claim 6, Velardo, JR. teaches receiving a secondary acoustic signal via a secondary microphone (signals transmitted into the network at a near location are received, col. 2, ln. 58-67).

Regarding claim 7, Velardo, JR. teaches wherein generating the noise gain mask comprises factoring in the secondary acoustic signal (the signals received at the near location are broken into sub-bands and then analyzed for echo cancellation, col. 3, ln. 5-50).

Regarding claim 8, Velardo, JR. teaches wherein generating the echo gain mask comprises taking a square root of the far-end acoustic signal and the primary acoustic signal to form the magnitude envelopes of the far-end and primary acoustic signals (a modulus signal is derived from each sub-band components, where the absolute value of each of these sub-band signals is smoothed, resulting in a waveform which is proportional to the root mean squared, i.e. rms, energy envelope of the sub-band signal, where each of these waveforms is then attenuated according to an echo-loss estimate and the resulting waveform represents the envelope of the expected echo waveform, col. 3, ln. 5-23).

Regarding claim 9, Velardo, JR. teaches smoothing the magnitude envelopes of the far-end and primary acoustic signals (the absolute value of each of these sub-band signals is smoothed, col. 3, ln. 5-23).

Regarding claim 10, Velardo, JR. teaches wherein generating the echo gain mask comprises determining an echo dominated noise estimate and using a square root of the echo dominated noise estimate to generate a noise magnitude envelope (a modulus signal is derived from each sub-band components, where the absolute value of each of these sub-band signals is smoothed, resulting in a waveform which is proportional to the root mean squared, i.e. rms, energy envelope of the sub-band signal, where each of these waveforms is then attenuated according to an echo-loss estimate and the resulting waveform represents the envelope of the expected echo waveform, col. 3, ln. 5-23).

--CONTINUED IN SUPPLEMENTAL BOX--

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/01667

Supplemental Box

In case **the space in any of the preceding boxes is not sufficient.**
Continuation of:
Box. No. V. 2. Citations and Explanations:

Regarding claim 11, Velardo, JR. teaches wherein generating the echo gain mask comprises predicting an echo envelope based on the magnitude envelopes of the far-end and primary acoustic signals (the magnitude of the near-end signal waveform is determined and the magnitude of the far-end signal waveform is determined and both are used in the smoothing operation for determining the envelope of the expected echo waveform, col. 3, ln. 5-23, col. 6, ln. 8-19).

Regarding claim 12, Velardo, JR. teaches wherein generating the echo gain mask comprises updating the gain based on one or more adaptation control cues (a center clipper transmits sub-band components which comply with a particular cue, e.g. ones that lie below a lower threshold, col. 3, ln. 36-50).

Regarding claim 13, Velardo, JR. teaches wherein the adaptation control cue comprises an adaptive far-end noise floor (a center clipper transmits sub-band components which lie below the lower threshold that represent a noise floor, col. 3, ln. 36-50).

Regarding claim 14, Velardo, JR. teaches wherein the adaptation control cue comprises a noise floor (a center clipper transmits sub-band components which lie below the lower threshold that represent a noise floor, col. 3, ln. 36-50).

Regarding claim 15, Velardo, JR. teaches wherein the adaptation control cue comprises a maximum gain (each sub-band component is subjected to a center-clipping operation that is intended to remove weak signals on the assumption that they are echoes, where the template, which represents the envelope of the expected echo waveform, is an upper threshold, col. 3, ln. 13-20).

Regarding claim 16, Velardo, JR. teaches synthesizing the masked signal into time domain for outputting (after center-clipping, the far-input sub-band components are combined, thereby to produce a synthesized, full-band, output signal, where an inverse FFT is used to transform from the frequency domain back to the time domain, col. 3, ln. 20-22, col. 8, ln. 43-60).

Regarding claim 17, Velardo, JR. teaches a system for envelope-based acoustic echo cancellation in a communication device, comprising: acoustic sensors configured to receive a primary acoustic signal and a far-end acoustic signal (signals transmitted into the network at a near location are received and signals transmitted into the network from a far location are received by the same processing device, col. 2, ln. 58-67); a frequency analysis module configured to performing frequency analysis on the primary acoustic signal and the far-end acoustic signal to obtain frequency subbands for the primary and far-end acoustic signals (the near input and the far input are each separately decomposed into plural sub-band components, col.3, ln. 3-4); an echo mask generator configured to generate an echo gain mask based on magnitude envelopes of the primary and far-end acoustic signals for each frequency sub-band (a masked signal is produced using magnitude envelopes, col. 3, ln. 5-50, col. 8, ln. 13-25); a modifier/reconstructor module configured to apply the echo gain mask to the primary acoustic signal to generate a masked signal (a noise mask and an echo mask are obtained by mixing the transmitted sub-band signal with a white-noise component, col. 3, ln. 23-42, col. 8, ln. 13-25); and an output device configured to output the masked signal (the masked signal is outputted, col. 3, ln. 5-50).

Regarding claim 18, Velardo, JR. teaches a noise mask generator configured to generate a noise gain mask based on at least the primary acoustic signal for each frequency sub-band (a noise mask and an echo mask are obtained by mixing the transmitted sub-band signal with a white-noise component, col. 3, ln. 23-42, col. 8, ln. 13-25).

Regarding claim 19, Velardo, JR. teaches wherein the modifier/reconstructor module configured to apply the echo gain mask to the primary signal to generate the masked signal comprises the modifier/reconstructor module configured to apply a combination of the echo gain mask and the noise gain mask to the primary acoustic signal to generate the masked signal (in order to mask the echo component, the transmitted sub-band signal is mixed with a white-noise component, col. 3, ln. 23-42, col. 8, ln. 13-25).

Regarding claim 20, Velardo, JR. teaches an echo and noise mask integration module configured to generate the combination of the echo gain mask and noise gain mask (after the transmitted sub-band signal is mixed with a white-noise component, the signal is generated to a post-smoothing function col. 3, ln. 23-42, col. 8, ln. 13-39).

Regarding claim 21, Velardo, JR. teaches wherein the modifier/reconstructor module is further configured to convert the masked signal into time domain for outputting (after center-clipping, the far-input sub-band components are combined, thereby to produce a synthesized, full-band, output signal, where an inverse FFT is used to transform from the frequency domain back to the time domain, col. 3, ln. 20-22, col. 8, ln. 43-60).

Regarding claim 22, Velardo, JR. teaches wherein the echo mask generator comprises an echo predictor configured to predict echo envelopes (each of the waveforms are attenuated according to an echo-loss estimate, col. 3, ln. 5-23).

--CONTINUED IN SUPPLEMENTAL BOX--

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/01667

Supplemental Box

In case **the space in any of the preceding boxes is not sufficient.**
Continuation of:
Box. No. V. 2. Citations and Explanations:

Regarding claim 23, Velardo, JR. teaches a machine readable medium having embodied thereon a program, the program providing instructions for a method for envelope-based acoustic echo cancellation in a communication device, comprising: receiving a primary acoustic signal via a primary microphone and a far-end signal via a receiver (signals transmitted into the network at a near location are received and signals transmitted into the network from a far location are received by the same processing device, col. 2, ln. 58-67); performing frequency analysis on the primary acoustic signal and the far-end acoustic signal to obtain frequency sub-bands for the primary and far-end acoustic signals (the near input and the far input are each separately decomposed into plural sub-band components, col.3, ln. 3-4); generating an echo gain mask based on magnitude envelopes of the primary and far-end acoustic signals for each frequency sub-band (a masked signal is produced using magnitude envelopes, col. 3, ln. 5-50, col. 8, ln. 13-25); generating a noise gain mask based on at least the primary acoustic signal for each frequency sub-band (a noise mask is obtained, col. 3, ln. 23-42, col. 8, ln. 13-25); applying a combination of the echo gain mask and noise gain mask to the primary acoustic signal to generate a masked signal (in order to mask the echo component using the echo and noise mask, the transmitted sub-band signal is mixed with a white-noise component, col. 3, ln. 23-42, col. 8, ln. 13-25); and outputting the masked signal (the masked signal is outputted, col. 3, ln. 5-50).

Claims 1-23 have industrial applicability as defined by PCT Article 33(4), because the subject matter can be made or used in industry.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 09/03813

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A61F 11/06 (2009.01)
USPC - 381/71.11
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 381/71.11

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 381/71.1; 381/150 (keyword limited - see search terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (PGPB, USPT, USOC, EPAB, JPAB); GOOGLE
Terms: noise, suppress, acoustic, signal, subtract, reference, predict, ratio, gain, mask, filter, coefficient, primary, secondary, microphone, mobile.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2008/0019548 A1 (Avendano) 24 January 2008 (24.01.2008), entire document, especially abstract, para [0023], [0029], [0031], [0045], [0053], [0061], [0062].	1-21
Y	US 2003/0101048 A1 (Liu) 29 May 2003 (29.05.2003), entire document, especially abstract, para [0006], [0015], [0036], [0037], [0038], [0039].	1-21

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 August 2009 (10.08.2009)	27 AUG 2009
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To:
STEPHEN BACHMANN
CARR & FERRELL LLP
2200 GENG ROAD
PALO ALTO, CA 94303

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*	27 AUG 2009

Applicant's or agent's file reference	FOR FURTHER ACTION
PA4604PCT	See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US 09/03813	26 June 2009 (26.06.2009)	30 June 2008 (30.06.2008)

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - A61F 11/06 (2009.01)
USPC - 381/71.11

Applicant AUDIENCE, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	10 August 2009 (10.08.2009)	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/03813

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:
   - [X] the international application in the language in which it was filed.
   - [ ] a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. [ ] This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of:
   a. type of material
      - [ ] a sequence listing
      - [ ] table(s) related to the sequence listing
   b. format of material
      - [ ] on paper
      - [ ] in electronic form
   c. time of filing/furnishing
      - [ ] contained in the international application as filed
      - [ ] filed together with the international application in electronic form
      - [ ] furnished subsequently to this Authority for the purposes of search

4. [ ] In addition, in the case that more than one version or copy of a sequence listing and/or table(s) relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

Form PCT/ISA/237 (Box No. I) (April 2007)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/03813

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	1-21	YES
	Claims	None	NO
Inventive step (IS)	Claims	None	YES
	Claims	1-21	NO
Industrial applicability (IA)	Claims	1-21	YES
	Claims	None	NO

2. Citations and explanations:

Claims 1-21 lack an inventive step under PCT Article 33(3) as being obvious over US 2008/0019548 A1 to Avendano in view of US 2003/0101048 A1 to Liu.

Regarding claim 1, Avendano teaches a method for suppressing noise (improved noise suppression, para [0023]), comprising: receiving at least a primary and a secondary acoustic signal (reception by the microphones 106 and 108, the acoustic signals are converted into electric signals (i.e., a primary electric signal and a secondary electric signal), para [0029]); subtracting a desired signal component from the secondary acoustic signal to obtain a noise component signal (delaying and subtracting the acoustic signals captured by the microphones 106 and 108. Signals (i.e., sounds), para [0031]); and outputting the noise subtracted signal (Sounds from the same sound source 102 are transmitted by the secondary microphone 108 as a cardioid secondary signal, para [0031]). Avendano does not expressly teach performing a first determination of at least one energy ratio related to the desired signal component and the noise component signal; performing a second determination of whether to adjust the noise component signal based on the at least one energy ratio; adjusting the noise component signal based on the second determination; subtracting the noise component signal from the primary acoustic signal to generate a noise subtracted signal. However, Liu, in an analogous art, teaches performing a first determination of at least one energy ratio related to the desired signal component and the noise component signal (reduce the energy of the background noise embedded in the voice sounds signals and enhance the voice sounds-to-noise ratio, para [0015]); performing a second determination of whether to adjust the noise component signal based on the at least one energy ratio (adjust the value of the correction coefficient ... so that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0038]); adjusting the noise component signal based on the second determination (the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0039]); subtracting the noise component signal from the primary acoustic signal to generate a noise subtracted signal (by subtracting the estimated background noise from the signal of the first microphone in the time domain, the better quality of the voice sounds signals could be obtained, para [0006]). It would have been obvious to one skilled in the art to incorporate the teachings of Liu with those of Avendano in order to reduce background noise in a spoken voice setting and enhance the voice sound to noise ratio (see Liu para [0015]).

Regarding claim 2, Liu further teaches wherein subtracting the desired signal component comprises applying a coefficient representing a source location to the primary acoustic signal to generate the desired signal component (value of the correction coefficient ... so that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0038]).

Regarding claim 3, Liu further teaches wherein the at least one energy ratio comprises a reference energy ratio and a prediction energy ratio (decides the output bit, b(k), by comparing the input speech sample ... and the prediction ... from the adaptive prediction filter 23, para [0036]).

Regarding claim 4, Liu further teaches adapting an adaptation coefficient applied to the noise component signal when the prediction energy ratio is greater than the reference energy ratio (determine the current status of the adaptive speech analysis unit 103 according to b(k) and its preceding three bits, b(k-1), b(k-2), and b(k-3), and determine one correction coefficient, para [0037]).

Regarding claim 5, Liu further teaches freezing an adaptation coefficient applied to the noise component signal when the prediction energy ratio is less than the reference energy ratio (If any two successive bits of these four bits are different, reset the correction coefficient as -1. This is because at this time the adaptive speech analysis unit 103 over-estimates the voice sounds signal and the stepsize is required to be decreased. For the other conditions, ... which represents the status that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals., para [0039]).

Regarding claim 6, Avendano further teaches determining a NP gain based on the at least one energy ratio indicating how much of the primary acoustic signal has been cancelled out of the noise subtracted signal (computing the magnitude of both signals and deriving the gain, g, as the ratio between the magnitude of the primary acoustic signal to the magnitude of the secondary acoustic signal, para [0062]).

Regarding claim 7, Avendano further teaches providing the NP gain to a multiplicative noise suppression system (the primary acoustic signal is multiplied by the smoothed Wiener filter estimate to estimate the speech, para [0053]).

Regarding claim 8, Avendano further teaches wherein the primary and secondary acoustic signals are separated into sub-band signals (filters such as short-time Fourier transform (STFT), sub-band filter banks, para [0061]).

---continued on supplemental box---

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/03813

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box No. V.2. Citations and explanations.

Regarding claim 9, Avendano further teaches wherein outputting the noise subtracted signal comprises outputting the noise subtracted signal to a multiplicative noise suppression system ((Sounds from the same sound source 102 are transmitted by the secondary microphone 108 as a cardioid secondary signal, para [0031]).

Regarding claim 10, Avendano further teaches wherein the multiplicative noise suppression system comprises generating a gain mask based at least on the noise subtracted signal (a Wiener filter is used to suppress noise/enhance speech, para [0045]).

Regarding claim 11, Avendano further teaches applying the gain mask to the noise subtracted signal to generate an audio output signal (Sounds from the same sound source 102 are transmitted by the secondary microphone 108 as a cardioid secondary signal, para [0031]).

Regarding claim 12, Avendano teaches a system for suppressing noise (improved noise suppression, para [0023]), comprising: a microphone array configured to receive at least a primary and a secondary acoustic signal (reception by the microphones 106 and 108, the acoustic signals are converted into electric signals (i.e., a primary electric signal and a secondary electric signal), para [0029]); an analysis module configured to generate a desired signal component which may be subtracted from the secondary acoustic signal to obtain a noise component signal (delaying and subtracting the acoustic signals captured by the microphones 106 and 108. Signals (i.e., sounds), para [0031]); a gain module configured to perform a first determination of at least one energy ratio related to the desired signal component and the noise component signal (computing the magnitude of both signals and deriving the gain, g, as the ratio between the magnitude of the primary acoustic signal to the magnitude of the secondary acoustic signal, para [0062]). Avendano does not expressly teach an adaptation module configured to perform a second determination of whether to adjust the noise component signal based on the at least one energy ratio, the adaption module further configured to adjust the noise component signal based on the second determination; and at least one summing module configured to subtract the desired signal component from the secondary acoustic signal and to subtract the noise component signal from the primary acoustic signal to generate a noise subtracted signal. However, Liu, in an analogou art, teaches an adaptation module configured to perform a second determination of whether to adjust the noise component signal based on the at least one energy ratio (reduce the energy of the background noise embedded in the voice sounds signals and enhance the voice sounds-to-noise ratio, para [0015]), the adaption module further configured to adjust the noise component signal based on the second determination (adjust the value of the correction coefficient ... so that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0038]); and at least one summing module configured to subtract the desired signal component from the secondary acoustic signal (the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0039]) and to subtract the noise component signal from the primary acoustic signal to generate a noise subtracted signal (by subtracting the estimated background noise from the signal of the first microphone In the time domain, the better quality of the voice sounds signals could be obtained, para [0006]). It would have been obvious to one skilled in the art to incorporate the teachings of Liu with those of Avendano in order to reduce background noise in a spoken voice setting and enhance the voice sound to noise ratio (see Liu para [0015]).

Regarding claim 13, Liu further teaches wherein the analysis module is configured to apply a coefficient representing a source location to the primary acoustic signal to generate the desired signal component value of the correction coefficient ... so that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0038]).

Regarding claim 14, Liu further teaches wherein the at least one energy ratio comprises a reference energy ratio and a prediction energy ratio (decides the output bit, b(k), by comparing the input speech sample ... end the prediction ... from the adaptive prediction filter 23, para [0036]).

Regarding claim 15, Liu further teaches wherein the adaptation module is configured to adapt an adaptation coefficient applied to the noise component signal when the prediction energy ratio is greater than the reference energy ratio (determine the current status of the adaptive speech analysis unit 103 according to b(k) and its preceding three bits, b(k-1), b(k-2), and b(k-3), and determine one correction coefficient, para [0037]).

Regarding claim 16, Liu further teaches wherein the adaptation module is configured to freeze an adaptation coefficient applied to the noise component signal when the prediction energy ratio is less than the reference energy ratio (If any two successive bits of these four bits are different, reset the correction coefficient as -1. This is because at this time the adaptive speech analysis unit 103 over-estimates the voice sounds signal and the stepsize is required to be decreased. For the other conditions, ... which represents the status that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals., para [0039]).

Regarding claim 17, Avendano further teaches wherein further comprising a gain module configured to determine a NP gain based on the at least one energy ratio indicating how much of the primary acoustic signal has been cancelled out of the noise subtracted signal (computing the magnitude of both signals and deriving the gain, g, as the ratio between the magnitude of the primary acoustic signal to the magnitude of the secondary acoustic signal, para [0062]).

---continued on supplemental box---

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/03813

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box No. V.2. Citations and explanations.

Regarding claim 18, Avendano teaches a machine readable medium having embodied thereon a program, the program providing instructions for a method for suppressing noise using noise subtraction processing (improved noise suppression, para [0023]), the method comprising: receiving at least a primary and a secondary acoustic signal (reception by the microphones 106 and 108, the acoustic signals are converted into electric signals (i.e., a primary electric signal and a secondary electric signal), para [0029]); subtracting a desired signal component from the secondary acoustic signal to obtain a noise component signal (delaying and subtracting the acoustic signals captured by the microphones 106 and 108. Signals (i.e., sounds), para [0031]); and outputting the noise subtracted signal (Sounds from the same sound source 102 are transmitted by the secondary microphone 108 as a cardioid secondary signal, para [0031]). Avendano does not expressly teach performing a first determination of at least one energy ratio related to the desired signal component and the noise component signal; performing a second determination of whether to adjust the noise component signal based on the at least one energy ratio; adjusting the noise component signal based on the second determination; subtracting the noise component signal from the primary acoustic signal to generate a noise subtracted signal. However, Liu, in an analogous art, teaches performing a first determination of at least one energy ratio related to the desired signal component and the noise component signal (reduce the energy of the background noise embedded in the voice sounds signals and enhance the voice sounds-to-noise ratio, para [0015]); performing a second determination of whether to adjust the noise component signal based on the at least one energy ratio (adjust the value of the correction coefficient ... so that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0038]); adjusting the noise component signal based on the second determination (the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals, para [0039]); subtracting the noise component signal from the primary acoustic signal to generate a noise subtracted signal (by subtracting the estimated background noise from the signal of the first microphone in the time domain, the better quality of the voice sounds signals could be obtained, para [0006]). It would have been obvious to one skilled in the art to incorporate the teachings of Liu with those of Avendano in order to reduce background noise in a spoken voice setting and enhance the voice sound to noise ratio (see Liu para [0015]).

Regarding claim 19, Liu further teaches wherein the at least one energy ratio comprises a reference energy ratio and a prediction energy ratio (decides the output bit, b(k), by comparing the input speech sample ... and the prediction ... from the adaptive prediction filter 23, para [0036]).

Regarding claim 20, Liu further teaches adapting an adaptation coefficient applied to the noise component signal when the prediction energy ratio is greater than the reference energy ratio (determine the current status of the adaptive speech analysis unit 103 according to b(k) and its preceding three bits, b(k-1), b(k-2), and b(k-3), and determine one correction coefficient, para [0037]).

Regarding claim 21, Liu further teaches freezing an adaptation coefficient applied to the noise component signal when the prediction energy ratio is less than the reference energy ratio (If any two successive bits of these four bits are different, reset the correction coefficient as -1. This is because at this time the adaptive speech analysis unit 103 over-estimates the voice sounds signal and the stepsize is required to be decreased. For the other conditions, ... which represents the status that the adaptive speech analysis unit 103 could adapt to the variation of the voice sounds signals., para [0039]).

Claims 1-21 have industrial applicability as defined by PCT Article 33(4), because the subject matter can be made or used in industry.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference PA4770PCT	FOR FURTHER ACTION	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. PCT/US 09/06754	International filing date *(day/month/year)* 30 December 2009 (30.12.2009)	(Earliest) Priority Date *(day/month/year)* 31 December 2008 (31.12.2008)
Applicant AUDIENCE, INC.		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 2 sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language**, the international search was carried out on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).

c. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (see Box No. II).

3. ☐ **Unity of invention is lacking** (see Box No. III).

4. With regard to the **title**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings**,

a. the figure of the drawings to be published with the abstract is Figure No. 1

☒ as suggested by the applicant.

☐ as selected by this Authority, because the applicant failed to suggest a figure.

☐ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figures is to be published with the abstract.

Form PCT/ISA/210 (first sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 09/06754

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - G10L 19/00 (2010.01)
USPC - 704/500
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 704/500

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 704/500

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Dialog Classic (Chinese Pat Abstr; Derwent Index, EPFT, French Pat, Jap Abstr, USPFT, WIPO/PCT PFT); Google Scholar; Terms searched: AMPLITUDE, AUDIO, AUDIO SIGNAL, COCHLEAR, COMPLEX-VALU, COMPLEX-VALUED, DECOMPOS, DECOMPOSED, DELAY, FILTER, FREQUENC, FREQUENCY, FREQUENCIES, LINEAR, PHASE, PSYCHOACOUSTIC, RECONSTRUCT...

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,956,674 A (Smyth et al.) 21 September 1999 (21.09.1999); entire document, especially: col 6, ln 15-17, 29-42; col 8, ln 2-17, 39-62; col 8, ln 39 to col 9, ln 4; col 10, ln 6-9; col 14, ln 4-7; col 47, ln 61-64	1-13, 17-20
Y		14-16
Y	US 5,682,463 A (Allen et al.) 28 October 1997 (28.10.1997); entire document, especially: col 32, ln 42-43; col 33, ln 27-30	14
Y	US 2007/0067166 A1 (Pan et al.) 22 March 2007 (22.03.2007); entire document, especially: para [0032] and [0034]	15
Y	US 2007/0100612 A1 (Ekstrand et al.) 03 May 2007 (03.05.2007); entire document, especially: para [0009], [0014]	16
A	US 5,974,380 A (Smyth et al.) 26 October 1999 (26.10.1999); entire document	1-20
A	US 6,226,616 B1 (You et al.) 01 May 2001 (01.05.2001); entire document	1-20
A	US 2006/0198542 A1 (Benjelloun Touimi et al.) 07 September 2006; entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 February 2010 (16.02.2001)	20 MAY 2010
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (July 2009)

PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To: STEPHEN BACHMANN
2200 GENG ROAD
PALO ALTO, CA 94303

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)* 20 MAY 2010

Applicant's or agent's file reference	FOR FURTHER ACTION
PA4770PCT	See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US 09/06754	30 December 2009 (30.12.2009)	31 December 2008 (31.12.2008)

International Patent Classification (IPC) or both national classification and IPC
IPC(8) - G10L 19/00 (2010.01)
USPC - 704/500

Applicant AUDIENCE, INC.

1. This opinion contains indications relating to the following items:

- ☒ Box No. I Basis of the opinion
- ☐ Box No. II Priority
- ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- ☐ Box No. IV Lack of unity of invention
- ☒ Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- ☐ Box No. VI Certain documents cited
- ☐ Box No. VII Certain defects in the international application
- ☐ Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

3. For further details, see notes to Form PCT/ISA/220.

Name and mailing address of the ISA/US	Date of completion of this opinion	Authorized officer:
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	17 February 2010 (17.02.2010)	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/237 (cover sheet) (July 2009)

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/06754

Box No. I Basis of this opinion

1. With regard to the **language**, this opinion has been established on the basis of:

   ☒ the international application in the language in which it was filed.

   ☐ a translation of the international application into _____ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing filed or furnished:

   a. (means)

   ☐ on paper

   ☐ in electronic form

   b. (time)

   ☐ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purposes of search

4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/06754

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims	14-16	YES
	Claims	1-13, 17-20	NO
Inventive step (IS)	Claims	None.	YES
	Claims	1-20	NO
Industrial applicability (IA)	Claims	1-20	YES
	Claims	None.	NO

2. Citations and explanations:

Claims 1-13, 17-20 lack novelty under PCT Article 33(2) as being anticipated by US 5,956,674 A to Smyth et al. (hereinafter "Smyth").

Regarding claim 1, Smyth discloses a method for reconstructing a decomposed audio signal, comprising: receiving a decomposed audio signal comprising a plurality of sub-band signals (col 6, ln 15-17) having successively shifted group delays as a function of frequency (col 8, ln 39-62); grouping the plurality of sub-band signals into two or more groups (col 8, ln 6-12, 44-62); applying a delay function to at least one of the two or more groups (col 8, ln 56-59); combining the groups to reconstruct the audio signal (col 8, ln 63 to col 9, ln 4); and outputting the audio signal (col 8, ln 63 to col 9, ln 4).

Regarding claim 2, Smyth discloses the method of claim 1, as applied above. Smyth further discloses further comprising adjusting one or more of a phase or amplitude of at least one of the plurality of sub-band signals (col 47, ln 61-64).

Regarding claim 3, Smyth discloses the method of claim 1, as applied above. Smyth further discloses wherein applying the delay function comprises realigning the group delays of the sub-band signals in at least one of the two or more groups (col 8, ln 12-17).

Regarding claim 4, Smyth discloses the method of claim 1, as applied above. Smyth further discloses wherein the delay function is based, at least in part, on a psychoacoustic model (col 6, ln 29-33).

Regarding claim 5, Smyth discloses the method of claim 1, as applied above. Smyth further discloses further comprising defining the delay function using a delay table (col 10, ln 6-9).

Regarding claim 6, Smyth discloses the method of claim 1, as applied above. Smyth further discloses wherein the two or more groups do not overlap (col 8, ln 2-12).

Regarding claim 7, Smyth discloses the method of claim 1, as applied above. Smyth further discloses wherein the combining comprises summing the two or more groups (col 14, ln 4-7).

Regarding claim 8, Smyth discloses a system for reconstructing a decomposed audio signal, comprising: a reconstruction module configured to receive a decomposed audio signal comprising a plurality of sub-band signals (col 6, ln 15-17) having successively shifted group delays as a function of frequency (col 8, ln 39-62), the reconstruction module comprising a grouping sub-module configured to group the plurality of sub-band signals into two or more groups (col 8, ln 6-12, 44-62), a delay sub-module configured to apply a delay function to at least one of the two or more groups (col 8, ln 56-59), and a combination sub-module configured to combine the groups to reconstruct the audio signal (col 8, ln 63 to col 9, ln 4); and a sink module configured to output the audio signal (col 8, ln 63 to col 9, ln 4).

Regarding claim 9, Smyth discloses the system of claim 8, as applied above. Smyth further discloses wherein the reconstruction module further comprises an adjustment sub-module configured to adjust one or more of a phase or amplitude of at least one of the plurality of sub-band signals (col 47, ln 61-64).

Regarding claim 10, Smyth discloses the system of claim 8, as applied above. Smyth further discloses wherein the delay sub-module is further configured to realign the group delays of the sub-band signals in at least one of the two or more groups (col 8, ln 12-17).

Regarding claim 11, Smyth discloses the system of claim 8, as applied above. Smyth further discloses wherein the delay function is based, at least in part, on a psychoacoustic model (col 6, ln 29-33).

Regarding claim 12, Smyth discloses the system of claim 8, as applied above. Smyth further discloses wherein the delay function is defined using a delay table (col 10, ln 6-9).

Regarding claim 13, Smyth discloses the system of claim 8, as applied above. Smyth further discloses wherein the combination sub-module is further configured to sum the two or more groups (col 14, ln 4-7).

–continued in Supplemental Box–

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 09/06754

Supplemental Box

In case the space in any of the preceding boxes is not sufficient.

Continuation of:
Box No. V.2. Citations and explanations:

Regarding claim 17, Smyth discloses a computer readable storage medium having embodied thereon a program (col 6, ln 33-42), the program being executable by a processor (col 6, ln 33-42) to perform a method for reconstructing a decomposed audio signal, the method comprising: receiving a decomposed audio signal comprising a plurality of sub-band signals (col 6, ln 15-17) having successively shifted group delays as a function of frequency (col 8, ln 39-62); grouping the plurality of sub-band signals into two or more groups (col 8, ln 6-12, 44-62); applying a delay function to at least one of the two or more groups (col 8, ln 56-59); combining the groups to reconstruct the audio signal (col 8, ln 63 to col 9, ln 4); and outputting the audio signal (col 8, ln 63 to col 9, ln 4).

Regarding claim 18, Smyth discloses the computer readable medium of claim 17, as applied above. Smyth further discloses further comprising adjusting one or more of a phase or amplitude of each of the plurality of sub-band signals (col 47, ln 61-64).

Regarding claim 19, Smyth discloses the computer readable medium of claim 17, as applied above. Smyth further discloses wherein applying the delay function comprises realigning the group delays of the sub-band signals in at least one of the two or more groups (col 8, ln 12-17).

Regarding claim 20, Smyth discloses the computer readable medium of claim 17, as applied above. Smyth further discloses wherein the delay function is based, at least in part, on a psychoacoustic model (col 6, ln 29-33).

Claim 14 lack an inventive step under PCT Article 33(3) as being obvious over Smyth in view of US 5,682,463 A to Allen et al. (hereinafter "Allen").

Regarding claim 14, Smyth discloses the system of claim 8, as applied above. However, Smyth does not explicitly disclose further comprising a fast cochlear transform filter bank, the fast cochlear transform filter bank providing the decomposed audio signal. However, Allen does disclose further comprising a fast cochlear transform filter bank, the fast cochlear transform filter bank providing the decomposed audio signal (col 32, ln 42-43; col 33, ln 27-30). It would have been obvious to one having ordinary skill in the art at the time of the applicant's claimed invention to modify the system of claim 8 as disclosed by Smyth to include further comprising a fast cochlear transform filter bank, the fast cochlear transform filter bank providing the decomposed audio signal as disclosed by Allen to analyze speech content contained within the the audio signal.

Claim 15 lacks an inventive step under PCT Article 33(3) as being obvious over Smyth in view of Allen and further in view of US 2007/0067166 A1 to Pan et al. (hereinafter "Pan").

Regarding claim 15, Smyth and Allen teach the system of claim 8, as applied above. However, Smyth does not explicitly disclose further comprising a linear phase filter bank, the linear phase filter bank providing the decomposed audio signal. However, Pan does disclose further comprising a linear phase filter bank, the linear phase filter bank providing the decomposed audio signal (para [0032] and [0034]). It would have been obvious to one having ordinary skill in the art at the time of the applicant's claimed invention to modify the system of claim 8 as taught by Smyth and Allen to include further comprising a linear phase filter bank, the linear phase filter bank providing the decomposed audio signal as disclosed by Pan to ensure that all frequency components in each group have equal delay times.

Claim 16 lacks an inventive step under PCT Article 33(3) as being obvious over Smyth in view of Allen and further in view of US 2007/0100612 A1 to Ekstrand et al. (hereinafter "Ekstrand").

Regarding claim 16, Smyth and Allen teach the system of claim 8, as applied above. However, Smyth does not explicitly disclose further comprising a complex-valued filter bank, the complex-valued filter bank providing the decomposed audio signal. However, Ekstrand does disclose further comprising a complex-valued filter bank, the complex-valued filter bank providing the decomposed audio signal (para [0009] and [0014]). It would have been obvious to one having ordinary skill in the art at the time of the applicant's claimed invention to modify the system of claim 8 as taught by Smyth and Allen to include further comprising a linear phase filter bank, the linear phase filter bank providing the decomposed audio signal as disclosed by Ekstrand to reduce the complexity of computing the decomposed signal.

Claims 1-20 have industrial applicability as defined by PCT Article 33(4) because the subject matter can be made or used in industry.

Electronic Acknowledgement Receipt	
EFS ID:	12306893
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	James William Drapinski
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	14-MAR-2012
Filing Date:	29-JAN-2007
Time Stamp:	20:04:40
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Form (SB08)	Audience_3909US_IDS_SB08_2.pdf	615516 5f021ecf951227f532afe67f252ba8b93d7b71c0	no	8

Warnings:
Information:

A U.S. Patent Number Citation or a U.S. Publication Number Citation is required in the Information Disclosure Statement (IDS) form for autoloading of data into USPTO systems. You may remove the form to add the required data in order to correct the Informational Message if you are citing U.S. References. If you chose not to include U.S. References, the image of the form will be processed and be made available within the Image File Wrapper (IFW) system. However, no data will be extracted from this form. Any additional data such as Foreign Patent Documents or Non Patent Literature will be manually reviewed and keyed into USPTO systems.					
2	Non Patent Literature	AUDIEN_NPL0007_BOLL_SUPPRESSION_113.pdf	1893781 f51f3616100038c095f226775941e35f4d734315	no	8
Warnings:					
Information:					
3	Non Patent Literature	AUDIEN_NPL0009_BOLL_SUPPRESSION.pdf	293675 7cb36fbd8b6692abf4e0dad3c18885e18712eeeb	no	3
Warnings:					
Information:					
4	Non Patent Literature	AUDIEN_NPL0014_DAHL_SIMULTANEOUS.pdf	674762 c83f08eed2eba8123be104cf74576c7d50f6dd38	no	4
Warnings:					
Information:					
5	Non Patent Literature	AUDIEN_NPL0017_ENT172_INSTRUCTIONAL.pdf	464397 68c1561cf0a6dbce90b81e5ace6fd41015ff3dcf	no	6
Warnings:					
Information:					
6	Non Patent Literature	AUDIEN_NPL0019_FULGUM_VOICE.pdf	871023 d97e99d4429321400e92190d2c837ac08dc59165	no	4
Warnings:					
Information:					
7	Non Patent Literature	AUDIEN_NPL0021_GRAUPE_BLIND.pdf	2719474 b04763f9e79c53607e759b5ff7ff30f341b3d054	no	13
Warnings:					
Information:					
8	Non Patent Literature	AUDIEN_NPL0022_HAYKIN_SIGNALS.pdf	372235 53dcdde12e0d2fbf8d2bd3ca0e61f26f1cbdae40	no	3
Warnings:					
Information:					
9	Non Patent Literature	AUDIEN_NPL0023_HERMANSKY_SHOULD.pdf	2529152 172ac9a99c8059386 51f94c7bd8e8bdeba7e2105	no	10
Warnings:					
Information:					

10	Non Patent Literature	AUDIEN_NPL0024_HOHMANN_FREQUENCY.pdf	2288405 e42a24190bfa3306fb3b9280797b5135af555bd2	no	10
Warnings:					
Information:					
11	Non Patent Literature	AUDIEN_NPL0025_JEFFRESS_PLACE.pdf	304937 29e94755c71ccf56c10d8cf8cd268a712a059891	no	5
Warnings:					
Information:					
12	Non Patent Literature	AUDIEN_NPL0026_JEONG_IMPLEMENTATION.pdf	2331106 f1c252d972176dbe9ea3d0b346970325394e8f2a	no	12
Warnings:					
Information:					
13	Non Patent Literature	AUDIEN_NPL0027_KATES_TIME.pdf	3887794 d61ce683fdeb20e232ba30aa1fc5af3d23e12cad	no	20
Warnings:					
Information:					
14	Non Patent Literature	AUDIEN_NPL0030_LAZZARO_SILICON.pdf	1053794 dfd0f44a0c5d268e67bcce48fe6cb4fafe8b0398	no	11
Warnings:					
Information:					
15	Non Patent Literature	AUDIEN_NPL0031_LIPPMANN_SPEECH.pdf	3100579 82ccb089c59a7d4915c56fa652c378ba92001a53	no	15
Warnings:					
Information:					
16	Non Patent Literature	AUDIEN_NPL0034_MARTIN_SPECTRAL.pdf	776718 c791ab1a1c1fc1f3296c15b245546f255a81eee1	no	4
Warnings:					
Information:					
17	Non Patent Literature	AUDIEN_NPL0035_MITRA_DIGITAL.pdf	921694 34c26166dadc0323122975ae66d5d2268a3610ef	no	5
Warnings:					
Information:					
18	Non Patent Literature	AUDIEN_NPL0038_NARRATIVE.pdf	415294 12ba547cd12a0520625ce47920e2eb96c2e97cde	no	4
Warnings:					
Information:					

19	Non Patent Literature	AUDIEN_NPL0040_COSI_LYONS.pdf	720695 cfc996352a59a4dd53ba7f6fbd67bed6d596de58	no	4
Warnings:					
Information:					
20	Non Patent Literature	AUDIEN_NPL0042_RABINER_DIGITAL.pdf	2173230 8e22faaed01492ea5f682b7944952a329f285c81	no	17
Warnings:					
Information:					
21	Non Patent Literature	AUDIEN_NPL0044_WEISS_ESTIMATING.pdf	1190613 64ec579cafa1ba838e9481c1451afde68686751f	no	6
Warnings:					
Information:					
22	Non Patent Literature	AUDIEN_NPL0045_SCHIMMEL_COHERENT.pdf	942041 2759c02d90b466ea9396b868ab9c659052a46983	no	4
Warnings:					
Information:					
23	Non Patent Literature	AUDIEN_NPL0046_SLANEY_LYONS_1988.pdf	6833847 3d7d53f25c3d3bc43333350b9f747943ece2a5f4	no	79
Warnings:					
Information:					
24	Non Patent Literature	AUDIEN_NPL0047_SLANEY_AUDITORY.pdf	930650 8637d3d12f700fda2709fa5f595e1c4f591ba7c4	no	7
Warnings:					
Information:					
25	Non Patent Literature	AUDIEN_NPL0048_SLANEY_INTRODUCTION.pdf	1241630 baa5a5510b7e07dc49e98ba534531cc3419248ee	no	7
Warnings:					
Information:					
26	Non Patent Literature	AUDIEN_NPL0050_SOLBACH_ARCHITECTURE.pdf	16482836 7ef7e2faa192a824260a5ea06f679df833c71066	no	142
Warnings:					
Information:					
27	Non Patent Literature	AUDIEN_NPL0053_STAHL_QUANTILE_BASED.pdf	890835 cbb4c5d7023be9ae8c29146b3c3e9797f1424d4f	no	4
Warnings:					
Information:					

28	Non Patent Literature	AUDIEN_NPL0057_SYNTRILLIUM_COOL_EDIT_USERS.pdf	12258515 a8ce3422f6f0be355e1a576eac94bbfc9083cf39	no	75
Warnings:					
Information:					
29	Non Patent Literature	AUDIEN_NPL0059_TCHORZ_SNR.pdf	2462032 2e88621c5ff16bd9d5d6d432ae976db3cb94bb24	no	9
Warnings:					
Information:					
30	Non Patent Literature	AUDIEN_NPL0062_WATTS_ROBUST.pdf	958023 a7b09e4edf660f2377a4363fd886dc0f0dcf0213	no	5
Warnings:					
Information:					
31	Non Patent Literature	AUDIEN_NPL0065_YOO_CONTINUOUS.pdf	827333 99f42340525b067194d3da68ddb19dea5b4d0a89	no	4
Warnings:					
Information:					
32	Non Patent Literature	AUDIEN_NPL0066_PCTUS0108372_ISR_2994PCT.pdf	154556 511b4549e0445b70fec6ed6a040113449be3b223	no	1
Warnings:					
Information:					
33	Non Patent Literature	AUDIEN_NPL0067_PCTUS0236946_ISR_2998PCT.pdf	153786 2a42789884c9affc699992bb49b4a6ba989cc815	no	1
Warnings:					
Information:					
34	Non Patent Literature	AUDIEN_NPL0068_PCTUS0304124_ISR_2999PCT.pdf	174445 ed029fdf4f8d0c28dfd3df07e50970b93483902d	no	1
Warnings:					
Information:					
35	Non Patent Literature	AUDIEN_NPL0069_PCTUS0700463_ISR_WO_3503PCT.pdf	1164438 0ed19601f1b02de84cafb603ea17da48a5f3be14	no	6
Warnings:					
Information:					
36	Non Patent Literature	AUDIEN_NPL0070_PCTUS0721654_ISR_WO_3909PCT.pdf	784078 8f6566d86c77d2a823f7d584dfe5c3684faf1fbb	no	5
Warnings:					
Information:					

37	Non Patent Literature	AUDIEN_NPL0071_PCTUS0712628_ISR_WO_3583PCT.pdf	802046 9ad59fe5d1e536fa3cf98685e8f7c75f7f677f23	no	6
Warnings:					
Information:					
38	Non Patent Literature	AUDIEN_NPL0072_PCTUS0808249_ISR_WO_4190PCT.pdf	1150441 558f535c0e055af717a9d2e05cb98e11b0c96c6a	no	6
Warnings:					
Information:					
39	Non Patent Literature	AUDIEN_NPL0073_PCTUS0901667_ISR_WO_4565PCT.pdf	900906 8f9575bcdba827ab4491feb69e2d790f1aa71fed	no	6
Warnings:					
Information:					
40	Non Patent Literature	AUDIEN_NPL0074_PCTUS0903813_ISR_WO_4604PCT.pdf	1169592 86946ecc3b7048d16662964ad8944c42972b745d	no	6
Warnings:					
Information:					
41	Non Patent Literature	AUDIEN_NPL0075_PCTUS0906754_ISR_WO_4770PCT.pdf	939989 d37c94df2f5e39f537c634dea81072910ce0967b	no	6
Warnings:					
Information:					
42	Non Patent Literature	AUDIEN_NPL0076_US_2875755.pdf	20725 abd076cb5e25214ca392d37d0ad2b5fee144bf53	no	1
Warnings:					
Information:					
43	Non Patent Literature	AUDIEN_NPL0077_DAHL_ACOUSTIC.pdf	725221 6538d2c21a980bb55c1f2893d19c43520ff8aec5	no	4
Warnings:					
Information:					
44	Non Patent Literature	AUDIEN_NPL0078_DEMOL_EFFICIENT.pdf	760088 0789123f546a3a0fb4a5027e0a9b9d909e4a83cc	no	4
Warnings:					
Information:					
45	Non Patent Literature	AUDIEN_NPL0079_LAROCHE_TIME.pdf	3006466 b398ddb52aeefd4cec72ceea9aea7a38d31c038a	no	31
Warnings:					
Information:					

46	Non Patent Literature	AUDIEN_NPL0080_MOULINES_NON-PARAMETRIC.pdf	5412239 91b078097dcc5d0af70568dd3ba92315d1537693	no	31
Warnings:					
Information:					
47	Non Patent Literature	AUDIEN_NPL0081_VERHELST_OVERLAP.pdf	2311336 ebdb8785a2d29b3a7d749d559aece24ed4b63a20	no	15
Warnings:					
Information:					
			Total Files Size (in bytes):	93056968	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

22830 7590 03/15/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER
KIM, PAUL

ART UNIT	PAPER NUMBER
2614	

DATE MAILED: 03/15/2012

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

TITLE OF INVENTION: SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$870	$300	$0	$1170	06/15/2012

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.

B. If the status above is to be removed, check box 5b on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above, or

If the SMALL ENTITY is shown as NO:

A. Pay TOTAL FEE(S) DUE shown above, or

B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check box 5a on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and 1/2 the ISSUE FEE shown above.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

22830 7590 03/15/2012
CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

TITLE OF INVENTION: SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$870	$300	$0	$1170	06/15/2012

EXAMINER	ART UNIT	CLASS-SUBCLASS
KIM, PAUL	2614	381-092000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. The following fee(s) are submitted:
- ❑ Issue Fee
- ❑ Publication Fee (No small entity discount permitted)
- ❑ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ❑ A check is enclosed.
- ❑ Payment by credit card. Form PTO-2038 is attached.
- ❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ❑ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.
- ❑ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 03/15/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	

ART UNIT	PAPER NUMBER
2614	

DATE MAILED: 03/15/2012

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment to date is 1078 day(s). If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the Patent Term Adjustment will be 1078 day(s).

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No.	Applicant(s)	
	11/699,732	AVENDANO, CARLOS	
	Examiner	Art Unit	
	PAUL KIM	2614	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *3/1/2012*.
2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.
3. ☒ The allowed claim(s) is/are *1-28*.
4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
6. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
7. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
7. ☐ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☐ Other _____.

/P. K./
Examiner, Art Unit 2614

DETAILED ACTION

Allowable Subject Matter

The following is an examiner's statement of reasons for allowance:

Regarding independent claim 1, the closest prior art of record, Elko (US Pub. 20030147538) discloses a system for enhancing speech, comprising: a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal; a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal.

However, Elko fails to teach the combination of a system for enhancing speech, comprising: a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal; a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal, the differential microphone array module being further configured to determine the cardioid primary signal based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and an inter-microphone level difference module

configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

Regarding independent claim 14, the closest prior art of record, Elko, discloses a method for enhancing speech, comprising: receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone; determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal.

However, Elko fails to teach the combination of a method for enhancing speech, comprising: receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone; determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; determining the cardioid primary signal further based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

Regarding independent claim 26, the closest prior art of record, Elko discloses a non-transitory computer readable storage medium having embodied thereon a program, the program being executable by a processor to perform a method for enhancing speech, the method comprising: receiving a primary acoustic signal at a primary

microphone and a secondary acoustic signal at a secondary microphone; determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal.

However, Elko fails to teach the combination of a non-transitory computer readable storage medium having embodied thereon a program, the program being executable by a processor to perform a method for enhancing speech, the method comprising: receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone; determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal; determining the cardioid primary signal further based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and non-linearly combining components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference. The distinct features, as disclosed in independent claims 1, 14, and 26 render the claims allowable.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled “Comments on Statement of Reasons for Allowance.”

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PAUL KIM whose telephone number is (571) 270-7697. The examiner can normally be reached on Monday - Thursday, 8 AM - 4 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Vivian Chin can be reached on (571) 272-7848. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/P. K./
Examiner, Art Unit 2614

/VIVIAN CHIN/
Supervisory Patent Examiner, Art Unit 2614

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11699732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

SEARCHED			
Class	Subclass	Date	Examiner
381	313, 312, 91, 92, 122	1/27/2011	PK
381	95, 110, 122, 94.1, 94.2, 94.3, 94.7	12/2/2011	PK
704	226, 227, 233, 275	12/2/2011	PK

SEARCH NOTES		
Search Notes	Date	Examiner
EAST - see EAST search notes	1/26/2011	PK
Inventor search - Carlos Avendano	1/26/2011	PK
Search consultation - Xu Mei (Primary, AU 2614)	8/10/2011	PK
Updated search - see EAST search notes	12/2/2011	PK
Updated search - see EAST search notes	3/10/2012	PK

INTERFERENCE SEARCH			
Class	Subclass	Date	Examiner
USPGPUBS Text Search		12/2/2011	PK

/P. K./ Examiner.Art Unit 2614	

Issue Classification

Application/Control No.	Applicant(s)/Patent Under Reexamination
11699732	AVENDANO, CARLOS
Examiner	**Art Unit**
PAUL KIM	2614

ORIGINAL		INTERNATIONAL CLASSIFICATION														
CLASS	SUBCLASS	CLAIMED						NON-CLAIMED								
381	92	H	0	4	R	3 / 00 (2006.0)										

CROSS REFERENCE(S)					
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)				
381	94.1	94.2	94.3	94.7	122
704	226	227	233	275	

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1	18	17												
2	2	19	18												
3	3	20	19												
4	4	21	20												
5	5	22	21												
6	6	23	22												
7	7	24	23												
8	8	25	24												
9	9	26	25												
10	10	28	26												
11	11	14	27												
12	12	27	28												
13	13														
15	14														
16	15														
17	16														

/PAUL KIM/ Examiner.Art Unit 2614 (Assistant Examiner)	3/10/2012 (Date)	**Total Claims Allowed:** 28	
/VIVIAN CHIN/ Supervisory Patent Examiner.Art Unit 2614 (Primary Examiner)	03/11/2012 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure 1a

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	2394	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer)))	US-PGPUB; USPAT	ADJ	OFF	2012/03/10 13:22
L2	731	L1 and (delay$5 and combin$5)	US-PGPUB; USPAT	ADJ	OFF	2012/03/10 13:22
L3	193	L1 and ((delay$5 same combin$5) and ((primary or first) same second$5))	US-PGPUB; USPAT	ADJ	OFF	2012/03/10 13:29
L4	9	l3 and (differential with microphone)	US-PGPUB; USPAT	ADJ	OFF	2012/03/10 13:29
L5	39	l1 and (differential with microphone)	US-PGPUB; USPAT	ADJ	OFF	2012/03/10 13:30
L6	5	l1 and (differential with microphone with array)	US-PGPUB; USPAT	ADJ	OFF	2012/03/10 13:30
S1	1	(11/699732).APP.	US-PGPUB; USPAT	OR	OFF	2011/01/27 05:56
S2	20	((CARLOS) near2 (AVENDANO)).INV.	US-PGPUB; USPAT	ADJ	ON	2011/01/27 06:07
S3	53	("20010016020" | "20010031053" | "20020009203" | "20020080980" | "20020106092" | "20020116187" | "20030039369" | "20030072460" | "20030099345" | "20030138116" | "20030169891" | "20050185813" | "4649505" | "4658426" | "5208864" | "5319736" | "5381473" | "5402496" | "5574824" | "5602962" | "5694474" | "5757937" | "5796819" | "6002776" | "6061456" | "6072881" | "6222927" | "6317501" | "6363345" | "6381176" | "6430295" | "6549630" | "6584203" | "6717991" | "6738482" | "6760450" | "6760805" | "6785381" | "6795558" | "6831865" | "6882736" | "6917688" | "6978159" | "7031478" | "7089349" | "7099821" | "7142677" | "7146316" | "7155019" | "7171008" | "7174022" | "7206418" | "7246058").PN.	US-PGPUB; USPAT	ADJ	ON	2011/01/27 06:09
S5	1	("20030147538").PN.	US-PGPUB; USPAT	OR	OFF	2011/01/27 07:30

S6	1	(11/343524).APP.	US-PGPUB; USPAT	OR	OFF	2011/01/27 07:33
S7	2615	(381/313,312,91,92,122).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/01/27 15:40
S8	55	("6381176" | "20020116187" | "20050185813" | "5319736" | "5402496" | "7206418" | "20030169891" | "5796819" | "6549630" | "7146316" | "6831865" | "5381473" | "6795558" | "7099821" | "7155019" | "7246058" | "20020106092" | "20030072460" | "6317501" | "6760450" | "5208864" | "5574824" | "6882736" | "20030147538" | "6760805" | "20020009203" | "20030039369" | "6430295" | "6584203" | "6717991" | "5602962" | "6917688" | "20010016020" | "20030099345" | "6002776" | "6061456" | "6363345" | "7171008" | "6978159" | "20020080980" | "5694474" | "4658426" | "5757937" | "7142677" | "7089349" | "20010031053" | "20030138116" | "4649505" | "6072881" | "6222927" | "6738482" | "6785381" | "7031478" | "7174022" | "20080260175").PN.	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 08:23
S9	1	("20060133621").PN.	US-PGPUB; USPAT; USOCR	OR	OFF	2011/12/02 08:26
S10	7369	speech with enhanc$6	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:46
S11	1	(speech with enhanc$6) same (primary with secondary with microphone$2) same (different$5 with microphone$2 with array)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:48
S12	12	(speech with enhanc$6) same (different$5 with microphone$2 with array)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:48
S13	111	(speech with enhanc$6) and (different$5 with microphone$2 with array)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:49
S14	1	(speech with enhanc$6) same (cardioid with primary with secondary with signal)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:52
S15	3	(speech with enhanc$6) same (microphone$2 with array) same (primary with secondary with signal)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 12:52
S16	6224	((381/91,92,95,110,122) or (704/226,233,275)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 12:57
S17	1508	((381/94.1,94.2,94.3,94.7) or (704/227)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 12:59
S18	1835	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer))) and (distance or separat$4)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:03

S19	2321	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer)))	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:04
S20	9	(inter-microphone with level with differenc$4)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:06
S21	21458	(inter with (level or microphone))	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:10
S22	58	S19 and S21	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:11
S25	6224	((381/91,92,95,110,122) or (704/226,233,275)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 13:54
S26	1835	(((speech or primary or mouth or voice) near2 (microphone or mic$2 or sensor or transducer)) same ((noise or reference or ambien$3 or environment$4 or secondary) near2 (microphone or mic$2 or sensor or transducer))) and (distance or separat$4)	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:54
S27	164	S26 and S25	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:54
S28	1508	((381/94.1,94.2,94.3,94.7) or (704/227)).CCLS.	US-PGPUB; USPAT	OR	OFF	2011/12/02 13:55
S29	67	S26 and S28	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
S30	21458	(inter with (level or microphone))	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
S31	80	S30 and S25	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55
S32	33	S30 and S28	US-PGPUB; USPAT	ADJ	OFF	2011/12/02 13:55

3/10/2012 1:31:04 PM
C:\Users\pkim2\Documents\EAST\Workspaces\11699732.wsp

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11699732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	01/27/2011	12/02/2011	03/10/2012						
1	1	✓	✓	=						
2	2	✓	✓	=						
3	3	✓	✓	=						
4	4	✓	✓	=						
5	5	✓	✓	=						
6	6	✓	✓	=						
7	7	✓	✓	=						
8	8	✓	✓	=						
9	9	✓	✓	=						
10	10	✓	✓	=						
11	11	✓	✓	=						
12	12	✓	✓	=						
13	13	✓	✓	=						
15	14	✓	✓	=						
16	15	✓	✓	=						
17	16	✓	✓	=						
18	17	✓	✓	=						
19	18	✓	✓	=						
20	19	✓	✓	=						
21	20	✓	✓	=						
22	21	✓	✓	=						
23	22	✓	✓	=						
24	23	✓	✓	=						
25	24	✓	✓	=						
26	25	✓	✓	=						
28	26	✓	✓	=						
14	27		✓	=						
27	28		✓	=						

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Paul Kim
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	4535473		1985-08-13	Sakata, Tomio	
	2	7949522		2011-05-24	Hetherington et al.	
If you wish to add additional U.S. Patent citation information please click the Add button.						Add

U.S.PATENT APPLICATION PUBLICATIONS						Remove
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Published Application citation information please click the Add button.						Add

FOREIGN PATENT DOCUMENTS							Remove	
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] ¡	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button							Add	

NON-PATENT LITERATURE DOCUMENTS	Remove

EFS Web 2.1.17

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Paul Kim
	Attorney Docket Number	PA3909US

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Paul Kim
	Attorney Docket Number	PA3909US

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☒ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☒ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☐ A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/James W. Drapinski/	Date (YYYY-MM-DD)	2012-04-18
Name/Print	James W. Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

IN THE
UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Carlos Avendano
APPLICATION NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
GROUP ART UNIT:	2614
CONFIRMATION NO.:	7721
ATTY.DKT.NO.:	PA3909US

MAIL STOP AMENDMENT
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT
Under 37 C.F.R. § 1.56, and 1.97-1.98
& Statement Under 37 C.F.R. § 1.97(e)(2)

SIR:

Pursuant to the provisions of 37 C.F.R. §§ 1.56 and 1.97-98 of the Rules of Practice in Patent Cases, enclosed herewith is form PTO-SB-08, listing several references. The Examiner is requested to make these references of official record in the application.

The references cited may be material to examination of the application and are submitted in compliance with the Applicant's duty of disclosure as defined by 37 C.F.R. § 1.56.

Additionally, the Applicant wishes to make the Examiner aware of, and invite the Examiner to consider, as is appropriate, the following office actions from the following U.S. patent applications:

- Notice of Allowance mailed February 13, 2012 in U.S. patent application number 11/441,675, filed May 25, 2006.
- Notice of Allowance mailed March 20, 2012 in U.S. patent application number 12/004,896, filed December 21, 2007.
- Office Action mailed March 16, 2012 in U.S. patent application number 12/077,436, filed March 18, 2008.
- Office Action mailed March 14, 2012 in U.S. patent application number 12/217,076, filed June 30, 2008.
- Notice of Allowance mailed February 28, 2012 in U.S. patent application number 12/286,909, filed October 2, 2008.
- Office Action mailed April 10, 2012 in U.S. patent application number 12/286,995, filed October 2, 2008.
- Office Action mailed April 16, 2012 in U.S. patent application number 12/319,107, filed December 31, 2008.

No representation is made or intended as to the completeness of this list, nor is the inclusion of any reference on this list an admission that it is prior art or pertinent to this application.

Applicant hereby declares, pursuant to 37 C.F.R. § 1.97(e)(2), that no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application and, to the knowledge of the person signing the certification after having made a reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 C.F.R. § 1.56(c) more than three months prior to the filing of the information disclosure statement.

The Commissioner is hereby authorized to charge any necessary fee to Account Number 06-0600.

Respectfully submitted,
Carlos Avendano

April 18, 2012

By: /James W. Drapinski/

James W. Drapinski, Reg. No. 46,242
Carr & Ferrell *LLP*
120 Constitution Drive
Menlo Park, CA 94025
TEL: (650) 812-3400
FAX: (650) 812-3444

Electronic Patent Application Fee Transmittal	
Application Number:	11699732
Filing Date:	29-Jan-2007
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Filer:	James William Drapinski/Justin Chang
Attorney Docket Number:	PA3909US
Filed as Small Entity	

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

Electronic Acknowledgement Receipt	
EFS ID:	12567955
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	James William Drapinski
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	18-APR-2012
Filing Date:	29-JAN-2007
Time Stamp:	15:10:20
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$180
RAM confirmation Number	1668
Deposit Account	060600
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Information Disclosure Statement (IDS) Form (SB08)	Audience_3909US_SB08.pdf	612319 3db0ab0a14a06f4c1db2408cd0ab0946bf2e0920	no	4
Warnings:					
Information:					
2	Transmittal Letter	Audience_3909US_SIDS_Transmittal_Letter.pdf	77959 64148b2a63933409a18fbc427c5d1e175ceb15a9	no	3
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	30597 0cf1c0b86bdfc8d968eb61f296bf4d96ca614173	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	720875		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 04/30/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	
ART UNIT	PAPER NUMBER
2614	
MAIL DATE	DELIVERY MODE
04/30/2012	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

UNITED STATES DEPARTMENT OF COMMERCE
U.S. Patent and Trademark Office
Address : COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450

APPLICATION NO./ CONTROL NO.	FILING DATE	FIRST NAMED INVENTOR / PATENT IN REEXAMINATION	ATTORNEY DOCKET NO.
11/699,732	29 January, 2007	AVENDANO, CARLOS	PA3909US

		EXAMINER	
CARR & FERRELL LLP 120 CONSTITUTION DRIVE MENLO PARK, CA 94025		PAUL KIM	
		ART UNIT	PAPER
		2614	20120329

DATE MAILED:

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner for Patents

The information disclosure statement (IDS) submitted on 3/14/2012 and 4/18/2012 is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

/P. K./
Examiner, Art Unit 2614

PTO-90C (Rev.04-03)

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99**)		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Paul Kim
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	4535473		1985-08-13	Sakata, Tomio	
	2	7949522		2011-05-24	Hetherington et al.	
If you wish to add additional U.S. Patent citation information please click the Add button.						Add

U.S.PATENT APPLICATION PUBLICATIONS						Remove
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Published Application citation information please click the Add button.						Add

FOREIGN PATENT DOCUMENTS								Remove
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button								Add

NON-PATENT LITERATURE DOCUMENTS	Remove

EFS Web 2.1.17

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 04/18/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Paul Kim
	Attorney Docket Number	PA3909US

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	04/20/2012

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 04/18/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Paul Kim	
	Attorney Docket Number	PA3909US	

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☒ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☒ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☐ A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/James W. Drapinski/	Date (YYYY-MM-DD)	2012-04-18
Name/Print	James W. Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	3976863		1976-08-24	Engel	
	2	3978287		1976-08-31	Fletcher, et al.	
	3	4137510		1979-01-30	Iwahara	
	4	4433604		1984-02-28	Ott	
	5	4516259		1985-05-07	Yato, et al.	
	6	4536844		1985-08-20	Lyon	
	7	4581758		1986-04-08	Coker, et al.	
	8	4628529		1986-12-09	Borth, et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	9	4630304		1986-12-16	Borth, et al.	
	10	4674125		1987-06-16	Carlson, et al.	
	11	4718104		1988-01-05	Anderson	
	12	4811404		1989-03-07	Vilmur, et al.	
	13	4812996		1989-03-14	Stubbs	
	14	4864620		1989-09-05	Bialick	
	15	4920508		1990-04-24	Yassaie, et al.	
	16	5027410		1991-06-25	Williamson, et al.	
	17	5054085		1991-10-01	Meisel, et al.	
	18	5058419		1991-10-22	Nordstrom, et al.	
	19	5099738		1992-03-31	Hotz	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99**)	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	20	5119711		1992-06-09	Bell, et al.	
	21	5142961		1992-09-01	Paroutaud	
	22	5150413		1992-09-22	Nakatani, et al.	
	23	5175769		1992-12-29	Hejna, Jr., et al.	
	24	5187776		1993-02-16	Yanker	
	25	5210366		1993-05-11	Sykes, Jr.	
	26	5224170		1993-06-29	Waite, Jr.	
	27	5230022		1993-07-20	Sakata	
	28	5323459		1994-06-21	Hirano	
	29	5341432		1994-08-23	Suzuki, et al.	
	30	5381512		1995-01-10	Holton, et al.	

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99**)	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	31	5400409		1995-03-21	Linhard	
	32	5402493		1995-03-28	Goldstein	
	33	5471195		1995-11-28	Rickman	
	34	5473702		1995-12-05	Yoshida, et al.	
	35	5473759		1995-12-05	Slaney, et al.	
	36	5479564		1995-12-26	Vogten, et al.	
	37	5502663		1996-03-26	Lyon	
	38	5544250		1996-08-06	Urbanski	
	39	5583784		1996-12-10	Kapust, et al.	
	40	5587998		1996-12-24	Velardo, Jr., et al.	
	41	5590241		1996-12-31	Park, et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	42	5675778		1997-10-07	Jones	
	43	5682463		1997-10-28	Allen, et al.	
	44	5706395		1998-01-06	Arslan, et al.	
	45	5717829		1998-02-10	Takagi	
	46	5729612		1998-03-17	Abel, et al.	
	47	5732189		1998-03-24	Johnston, et al.	
	48	5749064		1998-05-05	Pawate, et al.	
	49	5792971		1998-08-11	Timis, et al.	
	50	5806025		1998-09-08	Vis, et al.	
	51	5809463		1998-09-15	Gupta, et al.	
	52	5825320		1998-10-20	Miyamori, et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	53	5839101		1998-11-17	Vahatalo, et al.	
	54	5920840		1999-07-06	Satyamurti, et al.	
	55	5933495		1999-08-03	Oh	
	56	5943429		1999-08-24	Handel	
	57	5956674		1999-09-21	Smyth, et al.	
	58	5974380		1999-10-26	Smyth, et al.	
	59	5978824		1999-11-02	Ikeda	
	60	5983139		1999-11-09	Zierhofer	
	61	5990405		1999-11-23	Auten, et al.	
	62	6097820		2000-08-01	Turner	
	63	6108626		2000-08-22	Cellario, et al.	

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	64	6122610		2000-09-19	Isabelle	
	65	6134524		2000-10-17	Peters, et al.	
	66	6137349		2000-10-24	Menkhoff, et al.	
	67	6140809		2000-10-31	Doi	
	68	6173255		2001-01-09	Wilson, et al.	
	69	6180273		2001-01-30	Okamoto	
	70	6216103		2001-04-10	Wu, et al.	
	71	6223090		2001-04-24	Brungart	
	72	6226616		2001-05-01	You, et al.	
	73	6263307		2001-07-17	Arslan, et al.	
	74	6266633		2001-07-24	Higgins, et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	75	6339758		2002-01-15	Kanazawa, et al.	
	76	6355869		2002-03-12	Mitton	
	77	6381570		2002-04-30	Li, et al.	
	78	6434417		2002-08-13	Lovett	
	79	6449586		2002-09-10	Hoshuyama	
	80	6469732		2002-10-22	Chang, et al.	
	81	6487257		2002-11-26	Gustafsson, et al.	
	82	6496795		2002-12-17	Malvar	
	83	6513004		2003-01-28	Rigazio, et al.	
	84	6516066		2003-02-04	Hayashi	
	85	6529606		2003-03-04	Jackson, Jr. II, et al.	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29	
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	86	6622030		2003-09-16	Romesburg, et al.	
	87	6718309		2004-04-06	Selly	
	88	6792118		2004-09-14	Watts	
	89	6798886		2004-09-28	Smith, et al.	
	90	6810273		2004-10-26	Mattila, et al.	
	91	6915264		2005-07-05	Baumgarte	
	92	6944510		2005-09-13	Ballesty, et al.	
	93	6982377		2006-01-03	Sakurai, et al.	
	94	6999582		2006-02-14	Popovic, et al.	
	95	7016507		2006-03-21	Brennan	
	96	7020605		2006-03-28	Gao	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	97	7054452		2006-05-30	Ukita	
	98	7065485		2006-06-20	Chong-White, et al.	
	99	7076315		2006-07-11	Watts	
	100	7092529		2006-08-15	Yu, et al.	
	101	7092882		2006-08-15	Arrowood, et al.	
	102	7164620		2007-01-16	Hoshuyama	
	103	7171246		2007-01-30	Mattila, et al.	
	104	7209567		2007-04-24	Kozel, et al.	
	105	7225001		2007-05-29	Eriksson, et al.	
	106	7242762		2007-07-10	He, et al.	
	107	7254242		2007-08-07	Ise, et al.	

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29	
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	108	7359520		2008-04-15	Brennan, et al.	
	109	7412379		2008-08-12	Taori, et al.	
	110	7433907		2008-10-07	Nagai, et al.	
	111	7555434		2009-06-30	Nomura, et al.	

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20020002455		2002-01-03	Accardi, Anthony J. ; et al.	
	2	20020133334		2002-09-19	Coorman, Geert ; et al.	
	3	20020147595		2002-10-10	Baumgarte, Frank	
	4	20020184013		2002-12-05	Walker, Michael	
	5	20030014248		2003-01-16	Vetter, Rolf	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	6	20030033140		2003-02-13	Taori, Rakesh ; et al.	
	7	20030040908		2003-02-27	Yang, Feng ; et al.	
	8	20030063759		2003-04-03	Brennan, Robert L. ; et al.	
	9	20030095667		2003-05-22	Watts, Lloyd	
	10	20030101048		2003-05-29	Liu, Chia-Horng	
	11	20030103632		2003-06-05	Goubran, Rafik ; et al.	
	12	20030128851		2003-07-10	Furuta, Satoru	
	13	20030228023		2003-12-11	Burnett, Gregory C. ; et al.	
	14	20040013276		2004-01-22	Ellis, Richard Thompson ; et al.	
	15	20040047464		2004-03-11	Yu, Zhuliang ; et al.	
	16	20040057574		2004-03-25	Faller, Christof	

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	17	20040078199		2004-04-22	Kremer, Hanoh ; et al.	
	18	20040131178		2004-07-08	Shahaf, Mark ; et al.	
	19	20040165736		2004-08-26	Hetherington, Phil ; et al.	
	20	20040196989		2004-10-07	Friedman, Sol ; et al.	
	21	20040263636		2004-12-30	Cutler, Ross ; et al.	
	22	20050025263		2005-02-03	Wu, Gin-Der	
	23	20050027520		2005-02-03	Mattila, Ville-Veikko ; et al.	
	24	20050216259		2005-09-29	Watts, Lloyd	
	25	20050228518		2005-10-13	Watts, Lloyd	
	26	20050288923		2005-12-29	Kok, Chi-Wah	
	27	20060072768		2006-04-06	Schwartz; Stephen R. ; et al.	

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	28	20060098809		2006-05-11	Nongpiur; Rajeev ; et al.	
	29	20060133621		2006-06-22	Chen; Juin-Hwey ; et al.	
	30	20060149535		2006-07-06	Choi; Woo Young ; et al.	
	31	20060184363		2006-08-17	McCree; Alan ; et al.	
	32	20060198542		2006-09-07	Benjelloun Touimi; Abdellatif ; et al.	
	33	20060222184		2006-10-05	Buck; Markus ; et al.	
	34	20070067166		2007-03-22	Pan; Xingde ; et al.	
	35	20070094031		2007-04-26	Chen; Juin-Hwey	
	36	20070100612		2007-05-03	Ekstrand; Per ; et al.	
	37	20070195968		2007-08-23	Jaber; Marwan	
	38	20070230712		2007-10-04	Belt; Harm Jan Willem ; et al.	

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	39	20070276656		2007-11-29	Solbach; Ludger ; et al.	
	40	20080033723		2008-02-07	Jang; Giljin ; et al.	
	41	20080140391		2008-06-12	Yen; Ming Hsiang ; et al.	
	42	20080201138		2008-08-21	Visser; Erik ; et al.	
	43	20080228478		2008-09-18	Hetherington; Phillip A. ; et al.	
	44	20090012783		2009-01-08	Klein; David	
	45	20090012786		2009-01-08	Zhang; Xianxian ; et al.	
	46	20090129610		2009-05-21	Kim; Kyu-hong ; et al.	
	47	20090220107		2009-09-03	Every; Mark ; et al.	
	48	20090238373		2009-09-24	Klein; David	
	49	20090253418		2009-10-08	Makinen; Jorma	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29	
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	50	20090271187		2009-10-29	Yen; Kuan-Chieh ; et al.	
	51	20090323982		2009-12-31	Solbach; Ludger ; et al.	
	52	20100094643		2010-04-15	Avendano; Carlos ; et al.	
	53	20100278352		2010-11-04	Petit; Nicolas ; et al.	
	54	20110178800		2011-07-21	Watts; Lloyd	

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1	5053587	JP		1993-03-05	Nakada		☐
	2	2003/069499	WO		2003-08-21	Watts		☐
	3	2010/005493	WO		2010-01-14	Solbach et al.		☐
	4	2007/140003	WO		2007-12-06	Solbach et al.		☐

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29	
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	5	2003/043374	WO		2003-05-22	Watts		☐
	6	2001/74118	WO		2001-10-04	Watts		☐
	7	2007/081916	WO		2007-07-19	Avendano et al.		☐
	8	11-249693	JP		1999-09-17	Takano et al.		☐
	9	10-313497	JP		1998-11-24	Aoki et al.		☐
	10	2005195955	JP		2005-07-21	Amada et al.		☐
	11	2005110127	JP		2005-04-21	Takahashi		☐
	12	2004/010415	WO		2004-01-29	Nomura et al.		☐
	13	2004053895	JP		2004-02-19	Osamu et al.		☐
	14	2005518118	JP		2005-06-16	Watts		☐
	15	2003069499	WO		2003-08-21	Watts		☐

EFS Web 2.1.17

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

	16	2004531767	JP		2004-10-14	Brandman		☐
	17	2002103676	WO		2002-12-27	Brandman		☐
	18	2005148274	JP		2005-06-09	Nagai et al.		☐
	19	2002080362	WO		2002-10-10	Ekstrand		☐
	20	2004533155	JP		2004-10-28	Ekstrand		☐
	21	05-172865	JP		1993-07-13	Waite		☐

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T5
	1		☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	03/29/2012

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT **(Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlos Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlos Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/James W. Drapinski/	Date (YYYY-MM-DD)	2012-03-14
Name/Print	James W. Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (01-10)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	11699732
	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

U.S.PATENTS						Remove
Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Patent citation information please click the Add button.						Add

U.S.PATENT APPLICATION PUBLICATIONS						Remove
Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					
If you wish to add additional U.S. Published Application citation information please click the Add button.						Add

FOREIGN PATENT DOCUMENTS							Remove	
Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							☐
If you wish to add additional Foreign Patent Document citation information please click the Add button							Add	

NON-PATENT LITERATURE DOCUMENTS			Remove
Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

EFS Web 2.1.17

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	1	BOLL, STEVEN F. "Suppression of Acoustic Noise in Speech using Spectral Subtraction", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27, No. 2, April 1979, pgs. 113-120.	☐
	2	BOLL, STEVEN F. "Suppression of Acoustic Noise in Speech Using Spectral Subtraction", Dept. of Computer Science, University of Utah Salt Lake City, Utah, April 1979, pp. 18-19.	☐
	3	DAHL, MATTIAS et al., "Simultaneous Echo Cancellation and Car Noise Suppression Employing a Microphone Array", 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, April 21-24, pgs. 239-242.	☐
	4	"ENT 172." Instructional Module. Prince George's Community College Department of Engineering Technology. Accessed: October 15, 2011. Subsection: "Polar and Rectangular Notation". <http://academic.ppgcc.edu/ent/ent172_instr_mod.html>	☐
	5	FULGHUM, D. P. et al., "LPC Voice Digitizer with Background Noise Suppression", 1979 IEEE International Conference on Acoustics, Speech, and Signal Processing, pgs. 220-223.	☐
	6	GRAUPE, DANIEL et al., "Blind Adaptive Filtering of Speech from Noise of Unknown Spectrum Using a Virtual Feedback Configuration", IEEE Transactions on Speech and Audio Processing, March 2000, vol. 8, no. 2, pgs. 146-158.	☐
	7	HAYKIN, SIMON et al. "Appendix A.2 Complex Numbers." Signals and Systems. 2nd Ed. 2003. pg. 764	☐
	8	HERMANSKY, HYNEK "Should Recognizers Have Ears?", In Proc. ESCA Tutorial and Research Workshop on Robust Speech Recognition for Unknown Communication Channels, pages 1-10, France 1997.	☐
	9	HOHMANN, V. "Frequency Analysis and Synthesis Using a Gammatone Filterbank", ACTA Acustica United with Acustica, 2002, Vol. 88, pages 433-442.	☐
	10	JEFFRESS, LLOYD A. et al. "A Place Theory of Sound Localization," Journal of Comparative and Physiological Psychology, 1948, vol. 41, p. 35-39.	☐
	11	JEONG, HYUK et al., "Implementation of a New Algorithm Using the STFT with Variable Frequency Resolution for the Time-Frequency Auditory Model", J. Audio Eng. Soc., April 1999, Vol. 47, No. 4., pages 240-251.	☐

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	12	KATES, JAMES M. "A Time-Domain Digital Cochlear Model", IEEE Transactions on Signal Processing, December 1991, Vol. 39, No. 12, pages 2573-2592.	☐
	13	LAZZARO, JOHN et al., "A Silicon Model of Auditory Localization," Neural Computation Spring 1989, vol. 1, pgs. 47-57, Massachusetts Institute of Technology.	☐
	14	LIPPMANN, RICHARD P. "Speech Recognition by Machines and Humans", Speech Communication, July 1997, vol. 22, no. 1, pgs. 1-15	☐
	15	MARTIN, RAINER "Spectral Subtraction Based on Minimum Statistics", in Proceedings Europe. Signal Processing Conf., 1994, pgs. 1182-1185	☐
	16	MITRA, SANJIT K. Digital Signal Processing: a Computer-based Approach. 2nd Ed. 2001. pgs. 131-133	☐
	17	WATTS, LLOYD Narrative of Prior Disclosure of Audio Display on Feb. 15, 2000 and May 31, 2000.	☐
	18	COSI, PIERO et al. (1996), "Lyon's Auditory Model Inversion: a Tool for Sound Separation and Speech Enhancement," Proceedings of ESCA Workshop on 'The Auditory Basis of Speech Perception,' Keele University, Keele (UK), 15-19 July, 1996, pp. 194-197.	☐
	19	RABINER, LAWRENCE R. et al. "Digital Processing of Speech Signals", (Prentice-Hall Series in Signal Processing). Upper Saddle River, NJ: Prentice Hall, 1978.	☐
	20	WEISS, RON et al., "Estimating Single-Channel Source Separation Masks: Revelance Vector Machine Classifiers vs. Pitch-Based Masking", Workshop on Statistical and Perceptual Audio Processing, 2006.	☐
	21	SCHIMMEL, STEVEN et al., "Coherent Envelope Detection for Modulation Filtering of Speech," 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 1, no. 7, pgs 221-224.	☐
	22	SLANEY, MALCOM, "Lyon's Cochlear Model", Advanced Technology Group, Apple Technical Report #13, Apple Computer, Inc., 1988, pp 1-79	☐

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Filing Date	2007-01-29
(Not for submission under 37 CFR 1.99)	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	23	SLANEY, MALCOM, et al. "Auditory Model Inversion for Sound Separation," 1994 IEEE International Conference on Acoustics, Speech and Signal Processing, April 19-22, vol. 2, pgs 77-80.	☐
	24	SLANEY, MALCOM. "An Introduction to Auditory Model Inversion", Interval Technical Report IRC 1994-014, http://coweb.ecn.purdue.edu/~maclom/interval/1994-014/, September 1994, accessed on July 6, 2010	☐
	25	SOLBACH, LUDGER "An Architecture for Robust Partial Tracking and Onset Localization in Single Channel Audio Signal Mixes", Technical University Hamburg-Harburg, 1998.	☐
	26	STAHL, V. et al., "Quantile Based Noise Estimation for Spectral Subtraction and Wiener Filtering," 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, June 5-9, vol.3, pgs. 1875-1878	☐
	27	Syntrillium Software Corporation, "Cool Edit User's Manual", 1996, pp. 1-74.	☐
	28	TCHORZ, JURGEN et al., "SNR Estimation Based on Amplitude Modulation Analysis with Applications to Noise Suppression", IEEE Transactions on Speech and Audio Processing, Vol. 11, No. 3, May 2003, pgs. 184-192.	☐
	29	WATTS, LLOYD, "Robust Hearing Systems for Intelligent Machines," Applied Neurosystems Corporation, 2001, pp. 1-5.	☐
	30	YOO, HEEJONG et al., "Continuous-Time Audio Noise Suppression and Real-Time Implementation", 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, May 13-17, pgs. IV3980-IV3983.	☐
	31	International Search Report dated June 8, 2001 in Application No. PCT/US01/08372	☐
	32	International Search Report dated April 3, 2003 in Application No. PCT/US02/36946	☐
	33	International Search Report dated May 29, 2003 in Application No. PCT/US03/04124	☐

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	34	International Search Report and Written Opinion dated October 19, 2007 in Application No. PCT/US07/00463	☐
	35	International Search Report and Written Opinion dated April 9, 2008 in Application No. PCT/US07/21654	☐
	36	International Search Report and Written Opinion dated September 16, 2008 in Application No. PCT/US07/12628	☐
	37	International Search Report and Written Opinion dated October 1, 2008 in Application No. PCT/US08/08249	☐
	38	International Search Report and Written Opinion dated May 11, 2009 in Application No. PCT/US09/01667	☐
	39	International Search Report and Written Opinion dated August 27, 2009 in Application No. PCT/US09/03813	☐
	40	International Search Report and Written Opinion dated May 20, 2010 in Application No. PCT/US09/06754	☐
	41	FAST COCHLEA TRANSFORM, US Trademark Reg. No. 2,875,755 (August 17, 2004).	☐
	42	DAHL, MATTIAS et al., "Acoustic Echo and Noise Cancelling Using Microphone Arrays", International Symposium on Signal Processing and its Applications, ISSPA, Gold coast, Australia, August 25-30, 1996, pgs. 379-382.	☐
	43	DEMOL, M. et al. "Efficient Non-Uniform Time-Scaling of Speech With WSOLA for CALL Applications", Proceedings of InSTIL/ICALL2004 - NLP and Speech Technologies in Advanced Language Learning Systems - Venice 17-19 June, 2004.	☐
	44	LAROCHE, JEAN. "TIME AND PITCH SCALE MODIFICATION OF AUDIO SIGNALS", in "Applications of Digital Signal Processing to Audio and Acoustics", The Kluwer International Series in Engineering and Computer Science, Volume 437, pgs. 279-309, 2002.	☐

EFS Web 2.1.17 ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732 11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29
	First Named Inventor	Carlose Avendano
	Art Unit	2614
	Examiner Name	Kim, Paul
	Attorney Docket Number	PA3909US

	45	MOULINES, ERIC et al., "Non-Parametric Techniques for Pitch-Scale and Time-Scale Modification of Speech", Speech Communication, Vol. 16, pgs. 175-205, 1995.	☐
	46	VERHELST, WERNER, "Overlap-Add Methods for Time-Scaling of Speech", Speech Communication Vol. 30, pgs 207-221, 2000.	☐

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/Paul Kim/	Date Considered	03/29/2012

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Receipt date: 03/14/2012	Application Number	11699732	11699732 - GAU: 2614
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**	Filing Date	2007-01-29	
	First Named Inventor	Carlose Avendano	
	Art Unit	2614	
	Examiner Name	Kim, Paul	
	Attorney Docket Number	PA3909US	

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ Fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ None

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/James W. Drapinski/	Date (YYYY-MM-DD)	2012-03-14
Name/Print	James W. Drapinski	Registration Number	46,242

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /P.K./

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

22830 7590 03/15/2012
CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

TITLE OF INVENTION: SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$870	$300	$0	$1170	06/15/2012

EXAMINER	ART UNIT	CLASS-SUBCLASS
KIM, PAUL	2614	381-092000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 Carr & Ferrell LLP
2
3

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: Audience, Inc.
(B) RESIDENCE: (CITY and STATE OR COUNTRY): Mountain View, CA

Please check the appropriate assignee category or categories (will not be printed on the patent) : ☐ Individual ☒ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
☒ Issue Fee
☒ Publication Fee (No small entity discount permitted)
☐ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
☐ A check is enclosed.
☐ Payment by credit card. Form PTO-2038 is attached.
☒ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number 060600 (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
☐ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.
☐ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature /James W. Drapinski/ Date May 1, 2012
Typed or printed name James W. Drapinski Registration No. 46,242

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Carlos Avendano
SERIAL NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
GROUP ART UNIT:	2614
CONFIRMATION NO.:	7721
ATTY. DKT. NO.:	PA3909US

AMENDMENT AFTER ALLOWANCE PURSUANT TO 37 CFR § 1.312

Examiner Kim:

The Applicant submits this amendment after allowance per 37 C.F.R. § 1.312. This amendment seeks to amend the **specification** as listed beginning on **page two** and the **claims** as listed beginning on **page 10.** The Applicant's **remarks** begin on **page 17,** and the **conclusion** is provided on **page 18.**

IN THE SPECIFICATION

PLEASE REPLACE PARAGRAPH [0001] WITH THE FOLLOWING PARAGRAPH:

[0001] The present application claims the priority benefit of U.S. Provisional Patent Application No. 60/850,928, filed October 10, 2006, and entitled "Array Processing Technique for Producing Long-Range ILD Cues with Omni-Directional Microphone Pair;" the present application is also a continuation-in-part of U.S. Patent Application No. 11/343,524, and entitled "System and Method for Utilizing Inter-Microphone Level Differences for Speech Enhancement," which claims the priority benefit of U.S. Provisional Patent Application No. 60/756,826, filed January 5, 2006, and entitled "Inter-Microphone Level Difference Suppressor," ~~both~~ all of which are herein incorporated by reference.

PLEASE REPLACE PARAGRAPH [0005] WITH THE FOLLOWING PARAGRAPH:

[0005] A solution to the distance constraint may be obtained by using directional microphones. Using directional microphones ~~allow~~ allows a user to extend an effective level difference between the two microphones over a larger range with a narrow inter-level difference (ILD) beam. This may be desirable for applications such as push-to-talk (PTT) or videophones where a speech source is not in as close a proximity to the microphones, as for example, a telephone application.

PLEASE REPLACE PARAGRAPH [0026] WITH THE FOLLOWING PARAGRAPH:

[0026] Referring now to FIG. 2, the exemplary audio device 104 is shown in more detail. In exemplary embodiments, the audio device 104 is an audio receiving device that comprises a processor 202, the primary microphone 106, the secondary microphone 108, an audio processing engine 204, and an output device 206. The audio device 104 may comprise further components necessary for audio device 104 operations. The audio processing engine 204 will be discussed in more detail[[s]] in connection with FIG. 3.

PLEASE REPLACE PARAGRAPH [0029] WITH THE FOLLOWING PARAGRAPH:

[0029] FIG. 3 is a detailed block diagram of the exemplary audio processing engine 204, according to one embodiment of the present invention. In exemplary embodiments, the audio processing engine 204 is embodied within a memory device. In operation, the acoustic signals (i.e., X_1 and X_2) received from the primary and secondary microphones 106 and 108 are converted to electric signals and processed through a differential microphone array (DMA) module 302. The DMA module 302 is configured to use DMA theory to create directional patterns for the close-spaced microphones 106 and 108. The DMA module 302 may determine sounds and signals in a front and back cardioid region about the audio device 104 by delaying and subtracting the acoustic signals captured by the microphones 106 and 108. Signals (i.e., sounds) received from these cardioid regions are hereinafter referred to as cardioid signals. In one example, sounds from a ~~sound~~ audio source 102 within the cardioid region are transmitted by the primary microphone 106 as a cardioid primary signal. Sounds from the same ~~sound~~ audio source 102 are transmitted by the secondary microphone 108 as a cardioid secondary signal.

PLEASE REPLACE PARAGRAPH [0035] WITH THE FOLLOWING PARAGRAPH:

[0035] Referring to FIG. 4a, one implementation of the DMA module 302, frequency analysis module 304, energy module 306, and the ILD module 308 is provided. In this implementation, the acoustic signals received by the microphones 106 and 108 are processed by the DMA module 302. The exemplary DMA module 302 delays the primary acoustic signal, X_1, via a delay node 402, $z^{-\tau 1}$. Similarly, the DMA module 302 delays the secondary acoustic signal, X_2, via a second delay node [[40]]404, $z^{-\tau 2}$.

PLEASE REPLACE PARAGRAPH [0038] WITH THE FOLLOWING PARAGRAPH:

[0038] The energy module 306 takes the signals from the frequency analysis module 304 and calculates the power estimates associated with the cardioid primary signal (C_f) and the cardioid secondary ~~signal~~signal (C_b). In exemplary embodiments, the power estimates may be mathematically determined by squaring and integrating an absolute value of the output of the frequency analysis module 304. Power estimates of the signals from the cardioid primary signal and the cardioid secondary signal are referred to herein as components. For example, the energy level associated with the primary microphone signal may be determined by

$$E_f(t,\omega) = \int_{\text{frame}} \left| C_f(t',\omega) \right|^2 dt',$$

and the energy level associated with the secondary microphone signal may be determined by

$$E_b(t,\omega) = \int_{\text{frame}} \left| C_b(t',\omega) \right|^2 dt'.$$

PLEASE REPLACE PARAGRAPH [0039] WITH THE FOLLOWING PARAGRAPH:

[0039] Given the calculated energy levels, the ILD may be determined by the ILD module 308. In exemplary embodiments, the ILD is determined in a non-linear manner by taking a ratio of the energy levels, such as

$$ILD(t, \omega) = E_f(t,\omega) / E_b(t,\omega) \text{ .}$$

Applying the determined energy levels to this ILD equation[[s]] results in

$$ILD(t,\omega) = \frac{\int |C_f(t',\omega)|^2 dt'}{\int_{\text{frame}} |C_b(t',\omega)|^2 dt'} .$$

PLEASE REPLACE PARAGRAPH [0044] WITH THE FOLLOWING PARAGRAPH:

[0044] $\lambda_I(t,\omega)$ in the above equation is derived from the ILD approximated by the ILD module 308, as

$$\lambda_I(t,\omega) = \begin{cases} \approx 0 & \text{if} \quad ILD(t,\omega) < threshold \\ \approx 1 & \text{if} \quad ILD(t,\omega) > threshold \end{cases} .$$

That is, when ILD at the primary microphone 106 is smaller than a threshold value (e.g., *threshold* = 0.5) above which speech is expected to be, λ_I is small, and thus the noise estimator follows the noise closely. When ILD starts to rise (e.g., because speech is present within the large ILD region), λ_I increases. As a result, the noise estimate module 312 slows down the noise estimation process and the speech energy does not contribute significantly to the final noise estimate. Therefore, exemplary embodiments of the present invention may use a combination of minimum statistics and voice activity detection to determine the noise estimate.

PLEASE REPLACE PARAGRAPHS [0047]-[0048] WITH THE FOLLOWING PARAGRAPHS:

[0047] φ in the above exemplary Wiener filter equation is a factor which further limits the noise estimate. φ can be any positive value. In one embodiment, nonlinear expansion may be obtained by setting φ to 2. According to exemplary embodiments, φ is determined empirically and applied when a body of [[W=]]$\underline{W =} \left(\frac{P_s}{P_s + P_n} \right)$ falls below a prescribed value (e.g., 12 dB down from the maximum possible value of W, which is unity).

[0048] Because the Wiener filter estimation may change quickly (e.g., from one frame to the next frame) and noise and speech estimates can vary greatly between each frame, application of the Wiener filter estimate, as is, may result in artifacts (e.g., discontinuities, blips, transients, etc.). Therefore, an optional filter smoothing module 316 is provided to smooth the Wiener filter estimate applied to the acoustic signals as a function of time. In one embodiment, the filter smoothing module 316 may be mathematically approximated as

$$M(t,\omega) = \lambda_s(t,\omega)W(t,\omega) + (1-\lambda_s(t,\omega))M(t-1,\omega)_{\text{,}}$$

[[,]]where λ_s is a function of the Wiener filter estimate and the primary microphone energy, E_1.

PLEASE REPLACE PARAGRAPH [0053] WITH THE FOLLOWING PARAGRAPH:

[0053] Referring now to FIG. 4b, a practical implementation of the DMA module 302 according to one embodiment of the present invention is shown. In exemplary embodiments, microphone differences are compensated by using a filter 412, F(z), that equalizes the microphones 106 and 108. Since the filter 412 is a non-causal filter, in some embodiments, a delay is applied to the primary microphone signal with a delay node 414, D(z). The application of the delay node 414 results in an alignment of the two channels.

PLEASE REPLACE PARAGRAPH [0057] WITH THE FOLLOWING PARAGRAPH:

[0057] As a result, [[a]]sampling rate conversion (SRC) ~~node~~nodes 424 and 426 is provided. The outputs of the SRC nodes 424 and 426 are the cardioid primary and cardioid secondary signals, C_f and C_b.

PLEASE REPLACE PARAGRAPHS [0061]-[0063] WITH THE FOLLOWING PARAGRAPHS:

[0061] FIG. 6 is a polar plot of a front-to-back cardioid directivity pattern 602 and ILD diagram produced according to exemplary embodiments of the present invention. The cardioid directivity pattern 602 illustrates a range in which the acoustic signals may be received. As shown, by using the non-linear combination process and delay ~~lines~~nodes (e.g., 420 and 422), the range of the cardioid directivity pattern 602 may be extended in the forward and backward directions (i.e., along the x-axis). The extension in the forward and backward directions allows significant ILD cues to be obtained from acoustic sources further away from the microphones 106 and 108. As a result, the omni-directional microphones 106 and 108 can achieve acoustic characteristics that mimic those of directional microphones.

[0062] Referring now to FIG. 7, a flowchart 700 of an exemplary method for utilizing ILD of omni-direction microphones for noise suppression and speech enhancement is shown. In step 702, acoustic signals are received by the primary microphone 106 and the secondary microphone 108. In exemplary embodiments, the microphones are omni-directional microphones. In some embodiments, the acoustic signals are converted by the microphones to electronic signals (i.e., the primary electric signal and the secondary electric signal) for processing.

[0063] Differential array analysis is then performed in step 704 on the acoustic signals by the DMA module 302. In exemplary embodiments, the DMA module 302 is configured to determine the cardioid primary signal and the cardioid secondary signal by delaying, subtracting, and applying a gain factor to the acoustic signals captured by the microphones 106 and 108. Specifically, the DMA module 302 determines the cardioid primary signal by taking a difference between the primary electric signal and a delayed secondary electric signal. Similarly, the DMA module 302 determines

~~thecardioid~~the cardioid secondary signal by taking a difference between the secondary electric signal and a delay primary electric signal.

PLEASE REPLACE PARAGRAPH [0071] WITH THE FOLLOWING PARAGRAPH:

[0071] The above-described modules can be ~~comprises~~comprised of instructions that are stored on storage media. The instructions can be retrieved and executed by the processor 202. Some examples of instructions include software, program code, and firmware. Some examples of storage media comprise memory devices and integrated circuits. The instructions are operational when executed by the processor 202 to direct the processor 202 to operate in accordance with embodiments of the present invention. Those skilled in the art are familiar with instructions, processor(s), and storage media.

CLAIMS

PLEASE AMEND THE CLAIMS AS FOLLOWS

1. (Previously Presented) A system for enhancing speech, comprising:

a primary and secondary microphone configured to receive a primary acoustic signal and a secondary acoustic signal;

a differential microphone array (DMA) module configured to determine a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and secondary electric signal converted from the secondary acoustic signal, the differential microphone array module being further configured to determine the cardioid primary signal based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and

an inter-microphone level difference module configured to non-linearly combine components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

2. (Previously Presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by taking a difference between a delayed primary electric signal and a delayed and level-equalized secondary electric signal.

3. (Previously Presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid primary signal by determining a gain and taking a difference between a primary electric signal and a delayed secondary electric signal adjusted by the gain.

4. (Original) The system of claim 3 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

5. (Previously Presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between the secondary electric signal and a delayed primary electric signal.

6. (Original) The system of claim 1 further comprising a frequency analysis module configured to determine frequencies for the cardioid primary signal and the cardioid secondary signal.

7. (Original) The system of claim 1 further comprising an energy module configured to determine energy estimates for a frame of the cardioid primary signal and the cardioid secondary signal.

8. (Original) The system of claim 1 further comprising a noise estimate module configured to determine a noise estimate for the primary acoustic signal based on an energy estimate of the cardioid primary signal and the inter-microphone level difference.

9. (Original) The system of claim 1 further comprising a filter module configured to determine a filter estimate to be applied to the primary acoustic signal.

10. (Original) The system of claim 9 further comprising a filter smoothing module configured to smooth the filter estimate prior to applying the filter estimate to the primary acoustic signal.

11. (Original) The system of claim 1 further comprising a masking module configured to determine a speech estimate.

12. (Original) The system of claim 11 further comprising a frequency synthesis module configured to convert the speech estimate into a time domain for output.

13. (Original) The system of claim 1, wherein the DMA module determines the cardioid primary signal and a cardioid secondary signal of a sub-band of the primary electric signal.

14. (Previously Presented) A method for enhancing speech, comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal;

determining the cardioid primary signal further based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and

non-linearly combining components of the cardioid primary signal and cardioid secondary signal to obtain an inter-microphone level difference.

15. (Previously Presented) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a delayed secondary electric signal.

16. (Previously Presented) The method of claim 14 wherein determining the cardioid primary signal comprises determining a gain and taking a difference between a primary electric signal and a delayed secondary electric signal adjusted by the gain.

17. (Original) The method of claim 16 wherein the gain is the ratio between a magnitude of the primary acoustic signal and a magnitude of the secondary acoustic signal.

18. (Original) The method of claim 14 wherein determining the cardioid secondary signal comprises taking a difference between the secondary electric signal and a delayed primary electric signal.

19. (Original) The method of claim 14 wherein non-linearly combining comprises dividing the component of the cardioid primary signal by the component of the cardioid secondary signal.

20. (Original) The method of claim 14 further comprising determining an energy estimate for each of the acoustic signals during a frame.

21. (Original) The method of claim 14 further comprising determining a noise estimate based on an energy estimate of the primary acoustic signal and the inter-microphone level difference.

22. (Original) The method of claim 21 further comprising determining a filter estimate based on the noise estimate of the primary acoustic signal, the energy estimate of the primary acoustic signal, and the inter-microphone level difference.

23. (Currently Amended) The method of claim 22 further comprising producing a speech estimate by applying the filter estimate to the primary acoustic signal.

24. (Original) The method of claim 22 further comprising smoothing the filter estimate.

25. (Previously Presented) The method of claim 14 wherein the cardioid primary signal and the cardioid secondary signal are each of a sub-band of the primary electric signal.

26. (Previously Presented) A non-transitory computer readable storage medium having embodied thereon a program, the program being executable by a processor to perform a method for enhancing speech, the method comprising:

receiving a primary acoustic signal at a primary microphone and a secondary acoustic signal at a secondary microphone;

determining a cardioid primary signal and a cardioid secondary signal based on a primary electric signal converted from the primary acoustic signal and a secondary electric signal converted from the secondary acoustic signal;

determining the cardioid primary signal further based at least in part on delaying at least one of the primary electric signal and the secondary electric signal; and

non-linearly combining components of the cardioid primary signal and the cardioid secondary signal to obtain an inter-microphone level difference.

27. (Previously Presented) The system of claim 1 wherein the DMA module is configured to determine the cardioid secondary signal by taking a difference between a level-equalized secondary electric signal and a delayed primary electric signal.

28. (Previously Presented) The method of claim 14 wherein determining the cardioid primary signal comprises taking a difference between a delayed primary electric signal and a level-equalized secondary electric signal.

REMARKS

The Applicant has amended the specification to correct minor typographical errors and for consistency. The Applicant does not believe any new matter has been introduced by these amendments, and that these amendments are fully supported by the Specification and Figures as filed. For example, paragraph [0029] has been amended to conform to the Specification and Figures, and is supported at least by Fig. 1 and paragraph [0023] of the Specification ("While the microphones 106 and 108 receive sound (i.e., acoustic signals) from the audio source 102, the microphones 106 and 108 also pick up noise 110."). In light of the foregoing, the Applicant respectfully requests entry of the amendments to the Specification.

The Applicant has amended the claims to correct minor typographical errors. The Applicant does not believe any new matter has been introduced by these amendments, and that these amendments are fully supported by the Specification and Figures as filed.

This amendment does not require additional search or examination since the amendment merely addresses formal matters and does NOT affect the disclosure or the scope of any claim. The Applicant, therefore, respectfully requests entry of the aforementioned amendment under 37 C.F.R. § 1.312.

CONCLUSION

This amendment corrects minor typographical and consistency errors in the Specification and claims. Entry of the amendment is respectfully requested. The Examiner is invited to contact the undersigned with any questions pertaining to the present application.

The Commissioner is authorized to charge any underpayment or credit any overpayment to Deposit Account No. 06-0600 for any matter in connection with this response, including any fee for extension of time up to the current date, which may be required.

Respectfully submitted,
Carlos Avendano

Date: April 30, 2012

By: / James W. Drapinski /
James W. Drapinski (Reg. No. 46,242)
Carr & Ferrell LLP
120 Constitution Drive
Menlo Park, CA 94025
T: 650.812.3400
F: 650.812.3444

Electronic Patent Application Fee Transmittal

Application Number:	11699732
Filing Date:	29-Jan-2007
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Filer:	James William Drapinski/Jillian Sweetnam
Attorney Docket Number:	PA3909US

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Utility Appl issue fee	2501	1	870	870
Publ. Fee- early, voluntary, or normal	1504	1	300	300

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Miscellaneous:				
			Total in USD ($)	**1170**

Electronic Acknowledgement Receipt	
EFS ID:	12672232
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	James William Drapinski
Filer Authorized By:	
Attorney Docket Number:	PA3909US
Receipt Date:	01-MAY-2012
Filing Date:	29-JAN-2007
Time Stamp:	17:39:10
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	$1170
RAM confirmation Number	4647
Deposit Account	060600
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Issue Fee Payment (PTO-85B)	3909US_Issue_Fee_PTOL_85b.pdf	113998 9a7a8ffad033743ffaf3480cf8469b8ed687272c	no	1
Warnings:					
Information:					
2	Amendment after Notice of Allowance (Rule 312)	3909US_312_Amendment.PDF	152270 7c59114e1079e50213b45b2203498652ad21d3b7	no	18
Warnings:					
Information:					
3	Fee Worksheet (SB06)	fee-info.pdf	32013 b7ad2eb5ae39db9201362c2c981dd7bd7cae4bd6	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	298281		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

O.K. TO ENTER - 312
/P.K./ AU 2614
May 3, 2012

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Carlos Avendano
SERIAL NO.:	11/699,732
FILED:	January 29, 2007
TITLE:	System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement
EXAMINER:	Paul Kim
GROUP ART UNIT:	2614
CONFIRMATION NO.:	7721
ATTY. DKT. NO.:	PA3909US

AMENDMENT AFTER ALLOWANCE PURSUANT TO 37 CFR § 1.312

Examiner Kim:

The Applicant submits this amendment after allowance per 37 C.F.R. § 1.312. This amendment seeks to amend the **specification** as listed beginning on **page two** and the **claims** as listed beginning on **page 10.** The Applicant's **remarks** begin on **page 17,** and the **conclusion** is provided on **page 18.**

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7721

22830 7590 05/07/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

EXAMINER	
KIM, PAUL	
ART UNIT	**PAPER NUMBER**
2614	
MAIL DATE	**DELIVERY MODE**
05/07/2012	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Response to Rule 312 Communication	Application No.	Applicant(s)
	11/699,732	AVENDANO, CARLOS
	Examiner	Art Unit
	PAUL KIM	2614

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

1. ☒ The amendment filed on *01 May 2012* under 37 CFR 1.312 has been considered, and has been:

a) ☒ entered.

b) ☐ entered as directed to matters of form not affecting the scope of the invention.

c) ☐ disapproved because the amendment was filed after the payment of the issue fee.
Any amendment filed after the date the issue fee is paid must be accompanied by a petition under 37 CFR 1.313(c)(1) and the required fee to withdraw the application from issue.

d) ☐ disapproved. See explanation below.

e) ☐ entered in part. See explanation below.

/VIVIAN CHIN/ Supervisory Patent Examiner, Art Unit 2614	/P. K./ Examiner, Art Unit 2614

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/699,732	06/05/2012	8194880	PA3909US	7721

22830 7590 05/16/2012

CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 1384 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

Carlos Avendano, Campbell, CA;

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY.DOCKET NO./TITLE	REQUEST ID
11/699,732	01/29/2007	Carlos Avendano	PA3909US	7023

Acknowledgement of Loss of Entitlement to Entity Status Discount

The entity status change request below filed through Private PAIR on 11/05/2015 has been accepted.

CERTIFICATIONS:

Change of Entity Status:

X Applicant changing to regular undiscounted fee status.

NOTE: Checking this box will be taken to be notification of loss of entitlement to small or micro entity status, as applicable.

This portion must be completed by the signatory or signatories making the entity status change in accordance with 37 CFR 1.4(d)(4).

Signature:	/James W. Drapinski/
Name:	James W. Drapinski
Registration Number:	46242

PTO/SB/96 (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

STATEMENT UNDER 37 CFR 3.73(b)

Applicant/Patent Owner: Carlos AVENDANO — Attorney Docket Number: 112057-0747

Application No./Patent No.: 11/699732 — Filed/Issue Date: 1/29/2007

Titled: SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT

Knowles Electronics, LLC (Name of Assignee) — Corporation (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that it is:

1. ☒ the assignee of the entire right, title, and interest;
2. ☐ an assignee of less than the entire right, title, and interest (The extent (by percentage) of its ownership interest is %); or
3. ☐ the assignee of an undivided interest in the entirety of (a complete assignment from one of the joint inventors was made)

in the patent application/patent identified above by virtue of either:

A. ☐ An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel _____ Frame _____, or for which a copy thereof is attached.

OR

B. ☒ A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as follows:

1. From: CARLOS AVENDANO To: AUDIENCE, INC.
   The document was recorded in the United States Patent and Trademark Office at
   Reel 018860, Frame 0667, or for which a copy thereof is attached.

2. From: AUDIENCE, INC. To: AUDIENCE LLC
   The document was recorded in the United States Patent and Trademark Office at
   Reel 037927, Frame 0424, or for which a copy thereof is attached.

3. From: AUDIENCE LLC To: KNOWLES ELECTRONICS, LLC
   The document was recorded in the United States Patent and Trademark Office at
   Reel 037927, Frame 0435, or for which a copy thereof is attached.

☐ Additional documents in the chain of title are listed on a supplemental sheet.

☒ As required by 37 CFR 3.73(b)(1)(i), the documentary evidence of the chain of title from the original owner to the assignee was, or concurrently is being, submitted for recordation pursuant to 37 CFR 3.11.

[NOTE: A separate copy (i.e., a true copy of the original document(s)) must be submitted to Assignment Division in accordance with 37 CFR Part 3, if the assignment is to be recorded in the records of the USPTO. See MPEP 302.08]

The undersigned (whose title is supplied below) is authorized to act on behalf of the assignee.

Signature — Date: 11/18/16

Mark J. Danielson (Reg. No. 40,580) — Printed or Typed Name

Attorney for Applicant — Title

This collection of information is required by 37 CFR 3.73(b). The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: **Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/SB/80 (11-08)
Approved for use through 11/30/2011. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

POWER OF ATTORNEY TO PROSECUTE APPLICATIONS BEFORE THE USPTO

I hereby revoke all previous powers of attorney given in the application identified in the attached statement under 37 CFR 3.73(b).

I hereby appoint:

☑ Practitioners associated with the Customer Number: 139331

OR

☐ Practitioner(s) named below (if more than ten patent practitioners are to be named, then a customer number must be used):

Name	Registration Number	Name	Registration Number

as attorney(s) or agent(s) to represent the undersigned before the United States Patent and Trademark Office (USPTO) in connection with any and all patent applications assigned only to the undersigned according to the USPTO assignment records or assignment documents attached to this form in accordance with 37 CFR 3.73(b).

Please change the correspondence address for the application identified in the attached statement under 37 CFR 3.73(b) to:

☑ The address associated with Customer Number: 139331

OR

☐ Firm or Individual Name			
Address			
City		State	Zip
Country			
Telephone		Email	

Assignee Name and Address:

KNOWLES ELECTRONICS, LLC
1151 Maplewood Drive
Itasca, Illinois 60143

A copy of this form, together with a statement under 37 CFR 3.73(b) (Form PTO/SB/96 or equivalent) is required to be filed in each application in which this form is used. The statement under 37 CFR 3.73(b) may be completed by one of the practitioners appointed in this form if the appointed practitioner is authorized to act on behalf of the assignee, and must identify the application in which this Power of Attorney is to be filed.

SIGNATURE of Assignee of Record

The individual whose signature and title is supplied below is authorized to act on behalf of the assignee

Signature	Renee Dubroad Brown	Date	3/7/2016
Name	Renee Brown	Telephone	650-254-2681
Title	Associate GC, IP		

This collection of information is required by 37 CFR 1.31, 1.32 and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Electronic Acknowledgement Receipt

EFS ID:	27564630
Application Number:	11699732
International Application Number:	
Confirmation Number:	7721
Title of Invention:	SYSTEM AND METHOD FOR UTILIZING OMNI-DIRECTIONAL MICROPHONES FOR SPEECH ENHANCEMENT
First Named Inventor/Applicant Name:	Carlos Avendano
Customer Number:	22830
Filer:	Mark J. Danielson/Ayisha Patrick
Filer Authorized By:	Mark J. Danielson
Attorney Docket Number:	PA3909US
Receipt Date:	20-NOV-2016
Filing Date:	29-JAN-2007
Time Stamp:	10:08:17
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Assignee showing of ownership per 37 CFR 3.73	112057_0747_373b.pdf	101619 33a590f5e569bb29514150e2994afbc918693b69	no	2

Warnings:

Information:	
Total Files Size (in bytes):	101619

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/699,732	01/29/2007	Carlos Avendano	PA3909US

CONFIRMATION NO. 7721

139331
FOLEY & LARDNER LLP
3000 K STREET N.W.
SUITE 600
WASHINGTON, DC 20007-5109

POA ACCEPTANCE LETTER

OC000000090810826

Date Mailed: 04/21/2017

NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 11/20/2016.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/tmwilliams/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
11/699,732	01/29/2007	Carlos Avendano	PA3909US

CONFIRMATION NO. 7721

22830
CARR & FERRELL LLP
120 CONSTITUTION DRIVE
MENLO PARK, CA 94025

POWER OF ATTORNEY NOTICE

OC000000090810793

Date Mailed: 04/21/2017

NOTICE REGARDING CHANGE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 11/20/2016.

- The Power of Attorney to you in this application has been revoked by the assignee who has intervened as provided by 37 CFR 3.71. Future correspondence will be mailed to the new address of record(37 CFR 1.33).

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/tmwilliams/