	Multipart Description/PDF files in .zip description			
	Document Description	**Start**	**End**	
	Assertion of entitlement to small entity status	1	1	
	Reexam Certificate of Service	2	2	
Warnings:				
Information:				
2	90_019169_Comments_on_Stmt_of_Reasons_for_Confirmation.pdf	102195 8b258d2e03a8de3cfcdf8c350fba7339d570df37	yes	3

	Multipart Description/PDF files in .zip description			
	Document Description	**Start**	**End**	
	Patent Owner Comments after Action Closing Prosecution	1	2	
	Reexam Certificate of Service	3	3	
Warnings:				
Information:				
3	90_019169_Statement_of_Related_Proceedings.pdf	104798 5207ffbc4d078db0e35311fe3aa51d187bc13f64	yes	3

	Multipart Description/PDF files in .zip description			
	Document Description	**Start**	**End**	
	Notice of concurrent proceeding(s)	1	2	
	Reexam Certificate of Service	3	3	
Warnings:				
Information:				
	Total Files Size (in bytes):	300120		

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.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : **H04R 25/00**	**A1**	**(11) International Publication Number:** **WO 99/43185** **(43) International Publication Date:** 26 August 1999 (26.08.99)

(21) International Application Number: PCT/DK98/00062

(22) International Filing Date: 18 February 1998 (18.02.98)

(71) Applicant *(for all designated States except US)*: TØPHOLM & WESTERMANN APS [DK/DK]; Ny Vestergaardsvej 25, DK–3500 Værløse (DK).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: WESTERMANN, Søren, Erik [DK/DK]; Skovmosegaard, Grønholt, DK–3480 Fredensborg (DK).

(74) Agents: RAFFNSØE, Knud, Rosenstand et al.; Internationalt Patent–Bureau, Høje Taastrup Boulevard 23, DK–2630 Taastrup (DK).

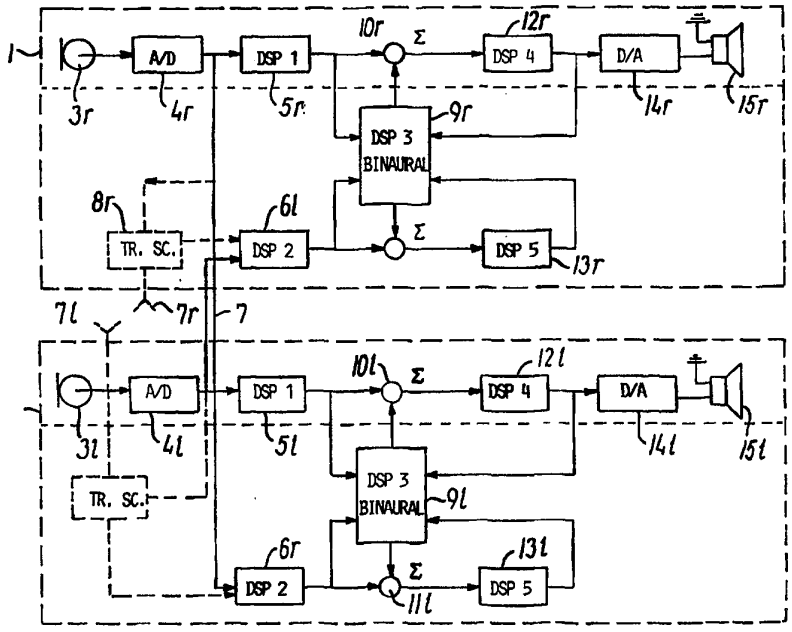
(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: A BINAURAL DIGITAL HEARING AID SYSTEM

(57) Abstract

A binaural digital hearing aid system comprises two hearing aid units (1, 2) for arrangement in a user's left and right ear, respectively. Each unit comprises input signal transducer means (3r, 3l), A/D conversion means (4r, 4l), digital signal processing means (5r–13r, 5l–13l), D/A conversion means (14r, 14l) and output signal transducer means (15r, 15l). A bi–directional communication link (7) is provided between the units. The digital signal processing means of each unit is arranged to affect a substantially full digital signal processing including individual processing of signals from the input transducer means of the actual unit and simulated processing of signals from the input transducer means of the other unit as well as binaural signal processing and includes at least a first digital signal processor part (5r, 5l) for processing said internally supplied signal, a second digital signal processor part (6l, 6r) for processing the signal supplied via said communication link (7) and a third digital signal processor part (9r, 9l) to effect common binaural digital signal processing of information derived from the signals processed in said first and second digital signal processor parts, said second digital signal processor part (6l, 6r) in each unit simulating the first digital signal processor part (5l, 5r) in the other unit with respect to adjustment parameters controlling the performance of said first signal processor part in said other unit.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

A Binaural Digital Hearing Aid System

The invention relates to a binaural digital hearing aid system comprising two hearing aid units for arrangement in a user's left and right ear, respectively, each of said units comprising input signal transducer means for conversion of a received input sound signal into an analog input signal, A/D conversion means for conversion of said analog input signal into a digital input signal, digital signal processing means for processing said digital input signal and generating a digital output signal, D/A conversion means for conversion of said digital output signal into an analog output signal and output signal transducer means for conversion of said analog output signal into an output sound signal perceivable to the user, a bidirectional communication link being provided between said units to connect a point in the signal path between the input signal transducer means and the digital signal processing means in one of said units with a corresponding point in the signal path between the input signal transducer means and the digital signal processing means of the other of said units.

For normally hearing persons the ability to localize sounds in space defined as binaural hearing ability is an important part of the sound perception. Typically the amplitude of sound received by the ipsilateral ear which is closer to the source of sound, is of greater amplitude than the sound received by the opposite contralateral ear. This difference in sound level, although often small by itself, is of great importance for a human being's perception of the direction of an incident sound.

In the human hearing system binaural sound perception results from a complicated signal processing of sounds arriving at the left and right side ears, in

which time/phase and frequency distribution of the sound plays a decisive role. Thus, time/phase differences and frequency enhancement are important for determining directions in the horizontal and vertical planes, respectively.

With conventional analog hearing aids persons suffering from a binaural hearing impairment, i.e. a hearing loss affecting both ears, the customary practice has been to use two separate hearing aids adjusted to compensate individually for the hearing loss of the respective ear for which the hearing aid is operative and compensation of the loss of binaural sound perception, although typically made even worse by the very use of a hearing aid in both ears, has in most cases by and large been ignored.

As a relatively simple compensation, it has been suggested for each of the two hearing aids of an analog system to use a microphone with a pronounced direction dependent characteristic to provide an analog signal the level of which changes, when the hearing aid is moved from a position pointing towards the sound source to other position with a minimum level, when the hearing aid points in a direction at right angles to the direction to the sound source.

In US-A-3,509,289 a different concept for compensation of binaural hearing loss in an analog hearing aid system is disclosed involving the use of cross-coupled AGC circuitry for maintaining and enhancing the interaural level difference between contralateral and ipsilateral incident sound. In this system, the gain of each of a first and second amplifying channel is varied inversely with the output of the other channel by separate AGC circuits which are cross-coupled to stabilize the system.

With the introduction of digital signal processing

in hearing aids a significant improvement of hearing aid performance has become possible and more advanced proposals for binaural hearing loss compensation have seen the light.

Thus, US-A-5,479,522 discloses a hearing enhancement system comprising in addition to two hearing aid devices for arrangement in the left and right side ears, respectively, a body-worn pack comprising a remote digital signal processor connected to each of the hearing aid device by a down-link and an up-link for interactive digital processing of the audio signals for each ear based on signals received from both hearing aid devices. The common binaural digital signal processing is predetermined and limited to attenuation of noise and narrowing of the sound field or adapting the signal level in the two channels. The signals supplied to the common binaural signal processing are not affected by the individual hearing loss compensation in the two channels.

In addition, this prior art system reduces the comfort by requiring a separate body-worn signal processor in a addition to the two hearing aid devices and the physical links between the common binaural processor and the two hearing aid devices in the form of radio communications make the system susceptible to distortion affecting the quality of sound reproduction.

In WO 97/14268, a binaural digital hearing aid system is disclosed in which the need for a separate body-worn remote control processor has been eliminated by the use of two hearing aid devices for arrangement in the left and right side ears, respectively, each of which incorporates a digital signal processor to which not only the unprocessed audio signal generated by the microphone in the same hearing aid device is supplied, but also the unprocessed audio signal generated by the

microphone in the opposite hearing aid device, the latter audio signal from each of the two devices being supplied to the respective opposite device through a bidirectional communication link.

This prior art system can be switched between distinct modes of either full binaural signal processing or performance as a conventional monaural hearing aid, which in one embodiment is done by giving the user the option of disabling the digital signal processor by either physically removing an external digital processing unit or by disabling a digital processor.

In the binaural processing mode of this prior art system no account is taken of the difference with respect to hearing loss and compensation between the two ears and, somewhat generalized, the system could be seen as an advanced digital substitute for the above-mentioned relatively simple binaural compensation using microphones with a pronounced direction dependent characteristic.

On this background, it is the object of the invention to provide an improved digital binaural hearing aid system in which the above-mentioned shortcomings of prior art systems have been eliminated to provide for a binaural signal processing, which for persons with a binaural hearing loss will restore binaural sound perception while taking into account the difference in hearing loss and compensation between the two ears.

According to the invention, a binaural digital hearing aid system as defined above is characterized in that the digital signal processing means of each unit is arranged to effect a substantially full digital signal processing including individual processing of signals from the input transducer means of the actual unit and simulated processing of signals from the input transducer means of the other unit as well as binaural

signal processing of signals supplied, on one hand, internally from the input signal transducer means of the same unit and, on the other hand, via said communication link from the input signal transducer means of the other unit, said digital signal processing means including at least a first digital signal processor part for processing said internally supplied signal, a second digital signal processor part for processing the signal supplied via said communication link and a third digital signal processor part to effect common binaural digital signal processing of information derived from the signals processed in said first and second digital signal processor parts, said second digital signal processor part in each unit simulating the first digital signal processor part in the other unit with respect to adjustment parameters controlling the performance of said first signal processor part in said other unit.

Thereby, in the binaural hearing aid system of the invention each of the hearing aid units for the left and right side ears, respectively, performs in addition to digital signal processing adapted to compensate for the hearing loss of the ear served by the unit, a simulated full digital signal processing of sound signals received by the unit for the opposite ear and adapted to compensate for the specific hearing loss of that ear, as well as a common binaural signal processing taking into account both of the normally different compensation characteristics of both units.

By the advantageous embodiments and modifications of the system set out in the dependant claims the system can be designed for user operated switching between functioning as a binaural system and a conventional monaural hearing system, and the digital signal processing means in each hearing unit may be

programmable to be switchable between different sound environments or listening situations by user operation, whereby programmed performance data for the first signal processing means of one unit is entered for programming of the second signal processing mens of the other unit, in which the simulated signal processing of signals supplied from the first unit is carried out.

By the provision of only a single bidirectional communication link between the two hearing aid units, the hearing aid system of the invention is less susceptible to signal distortion and interruption than the prior art systems described above.

In the following the invention will be further explained with reference to the accompanying drawings on which

figure 1 is a schematic block diagram representation of an embodiment of the hearing aid system of the invention;

figure 2 a further detailed representation of the embodiment shown in figure 1; and

fig. 3 is a block diagram representation of one hearing aid unit of a programmable hearing aid system of the invention.

The binaural hearing system illustrated in figure 1 comprises two hearing aid units 1 and 2 intended for arrangement in the user's right and left ears, respectively. The hearing aid units 1 and 2 are identical in structure, but as further explained in the following they will normally have been programmed or otherwise adjusted to provide different hearing loss compensation adapted to the specific hearing impairment of the ear, in which the unit is to be arranged. For the following description identical parts in the two units 1 and 2 will be designated by the same reference numeral followed by "r" and "l", respectively, to indicate the

localization of such parts in either the right ear or the left ear unit 1 or 2.

Each of units 1 and 2 comprises input signal transducer means e.g. in the form of one or more hearing aid microphones 3r, 3l which receives sound signals to be processed in the unit and transforms these sound signals into analog electrical signals which are supplied to an A/D converter 4r, 4l for conversion into digital signals.

In the embodiment shown the digital signal from A/D converter 4r, 4l in each of units 1 and 2 is supplied to a first digital signal processor 5r, 5l which is programmed or otherwise adjusted to perform signal processing functions such as filtering, band-division, amplification, gain control adjustment, compression, expansion and/or compensation for unlinearities in the microphone or the user's ear channel.

However, to the extent that some of the signal processing functions of processors 5r, 5l may be implemented in the A/D converters 4r, 4l, each of which will then supply a preprocessed digital signal, each of digital processors 5r, 5l need not be present as a separate unit.

According to the invention, each of units 1 and 2 also comprises a second digital signal processor 6l and 6r, respectively, which is structurally identical to processor 5r, 5l, but is programmed or otherwise adjusted to perform digital signal processing functions on the signals received by the opposite ear, i.e. processor 6l in unit 1 for the right ear is programmed to provide the specific signal processing intended for the left ear and will thus, in principle, provide the same signal processing as signal processor 5l in unit 2, whereas signal processor 6r in unit 2 will provide the same signal processing as processor 5r in unit 1.

In the embodiment shown the digital electrical signal from converter 4r, 4l in each of units 1 and 2 is additionally supplied via a communication link 7 to second signal processor 6r, 6l in the other unit, such that in each unit the second signal processor 6l, 6r performs a simulated processing corresponding to the processing by first signal processor 5l, 5r in the other unit. However, as an alternative the analog signals from microphone 3r, 3l in each of units 1 and 2 could be communicated directly and supplied to A/D conversion in the other unit.

The signal processors 5r, 5l and 6r, 6l will typically be state of the art digital hearing aid processors programmed to perform a relatively sophisticated signal processing with respect to sound/noise separation and user operated adaptability to a number of different sound environments or listening situations.

The communication link 7 between the right and left hearing aid units 1 and 2 is preferably a single bidirectional communication link which may be physically implemented by a cable extending between the two units. the hearing aid units 1 and 2 may be designed for arrangement in the ear (ITE) or behind the ear (BTE). In either case a cable connection between the units may extend around the neck of the user and may eventually be integrated in a necklace or similar piece of jewelry or bijouterie.

Alternatively, the bidirectional communication link 7 may be wireless and, as shown in dashed lines,comprise antennas 7r, 7l connected with appropriate transceiving means 8r, 8l in each unit. For hearing units in ITE design such an antenna may be physically implemented by a relatively short piece of wire or string which in use will project outside the ear and

may serve additionally to facilitate withdrawal of the hearing unit from its ITE position.

In each of units 1 and 2 the first and second digital signal processors 5r, 5l and 6l, 6r outputs a processed digital signal which is supplied to a third signal processor 9r, 9l which, in accordance with the invention performs a common binaural digital signal processing of the processed digital signals outputted from the first and second signal processors 5r, 5l and 6l, 6r.

The binaural signal processing in each of third signal processors 9r, 9l may make use of state of the art binaural processing techniques taking into account differences with respect to amplitude, phase-lag etc. between arrival of incoming sounds at the input transducers of the right and left ear hearing aid units. As result of this binaural signal processing which according to the invention is based on information derived from the processed digital signals in both of units 1 and 2, the third signal processor 9l, 9r in each unit outputs processed digital right and left binaural signal parts which in digital adder devices 10r, 10l and 11l, 11r is combined with the processed digital output signal from first signal processor 5r, 5l in the same unit.

In each unit the combined processed digital signals from adder device 10r, 10l may be supplied directly to a D/A converter 14r, 14l for conversion into a processed analog signal which is supplied to an output transducer device in the form of a conventional hearing aid telephone 15r, 15l. As illustrated the processed digital signals from adder devices 10r, 10l and 11r, 11l may optionally undergo a further digital signal processing in fourth signal processors 12r, 12l and fifth signal processors 13r, 13l, respectively,

which may include compensation of the specific hearing loss and automatic gain control. From processors 12r, 12l and 13r, 13l feedback signals are also supplied to the binaural processor 9r, 9l.

The processing functions of the fourth and fifth signal processors 12r, 12l and 13r, 13l in each of units 1 and 2 may be implemented in the binaural processor 9r, 9l so that processors 12r, 12l and 13r, 13l may in principle be dispensed with as separate units. The binaural signal processor 9r, 9l may then further be designed to output only the binaural digital signal part intended for the actual unit, i,e, the right ear binaural signal part for unit 1 and the left ear binaural signal part for unit 2. In both cases, the incorporation of the fourth and fifth signal processors either as separate units or in the binaural processor 9r, 9l provides an advantageous possibility, however, for an AGC function and/or hearing loss compensation in the binaural signal processor 9r, 9l by feed-back of the processed digital output signals from both of the fourth and fifth signal processors 12r, 12l and 13r, 13l to the binaural processor 9r, 9l.

Examples of digital signal processors for use in each of units 1 and 2 are disclosed e.g. in EP-B1-0 732 036, US-A-5,165,017, US-A-4,531,229 and US-A-5,144,675. An advanced signal processing method and device employing socalled dynamic AGC has been disclosed in copending international patent application PCT/DK97/00598, the disclosure of which is incorporated herein by reference.

The design and structure of the binaural hearing system of the invention, by which all information carrying signals in the separate signal channels for the right ear and the left ear sound perception are made available for processing in both of units 1 and 2

not only of the signal actually belonging to the respective side, but also, by simulated processing, of the signal belonging to the other side, opens the possibility of implementing complex and highly sophisticated binaural signal processing to restore binaural hearing ability without significant complication of the structure. In fact, both of hearing aid units 1 and 2 may be identical in structure and equipped with identical components like converters, signal processors etc.

Whereas the various signal processors in each of units 1 and 2 have been illustrated and described as separate processors they may advantageously be incorporated as separate processing parts of a common single digital processor such as a microprocessor.

The embodiment of the hearing aid system shown in figure 2 serves to illustrate the degree of complexity of binaural signal processing that can be implemented in each of the right and left ear hearing aid units 16 and 17 with a single bidirectional communication link 28 between the two units.

Using the same distinction as to reference numerals between the right and left ear units as used for the embodiment in figure 1 only the structure and function of the right ear unit 16 will be explained in the following.

The unprocessed analog signal from microphone 18r is preamplified and converted to digital form in preamplifier and A/D converter 19r and processed to compensate for unlinearity of the microphone and the sound perception in the ear in linearity control unit 20r, from which a preprocessed digital signal is supplied, on one hand to a band divider filtering unit 21r in the signal processing channel for the right ear and, on the other hand, via the bidirectional communication link 28 to a band divider filtering unit 21rs in

the processing part of the left ear unit 17 performing the simulated right ear signal processing.

In the band divider filtering unit 21r the incoming preprocessed digital signal is split into a number of frequency bands, each of which is further processed in a noise filtering unit 22r and a processing unit 23r in which the signal is amplified in accordance with the compensation characteristic adapted to compensate for the specific hearing loss of the right ear.

As for the embodiment in figure 1 each of the two hearing aids comprise in addition to the signal processing channel for the ear in which the unit is arranged a separate signal processing channel effecting a simulated signal processing corresponding to the signal processing in the other unit. In the embodiment in figure 2 this simulated processing channel comprises for the right ear hearing aid unit 16 processing units 21ls, 22ls and 23ls effecting the same processing functions as processing units 21r, 22r and 23r for the right ear compensation, but adjusted to the specific characteristics for the left ear compensation effected in the left ear hearing aid unit 17.

The left ear hearing aid unit 17 is identical in structure to the right ear hearing aid unit 16 and comprises the left ear signal processing channel with processing unit 19l and 20l and filtering and compensation units 21l, 22l and 23 l as well as the simulated right ear processing channel including units 21rs, 22rs and 23rs.

In each of hearing aid units 16 and 17 binaural signal processing may be effected in two processing units 24, 24l and 25r, 25l. Thus, in the right ear hearing aid unit 16 a first binaural processing unit 24r may receive the band divided output signals from filtering unit 21r in the right ear processing channel

as well as from filtering unit 21ls in the simulated left ear processing channel and provide correction signals affecting signal scaling in processing units 22r and 22ls and a second binaural processing unit 25r may effect further binaural signal processing on incoming signal from the first binaural processing unit 24r as well as from processing units 22r, 22ls and 23r, 23ls.

Finally, in each of hearing aid units 16 and 17 the output signal from processing unit 23r, 23l in the right and left ear processing channel, respectively, and the binaural output signal from the second binaural processing unit 25r. 25l is reconverted into analog form in an output processing unit 26r, 26l and supplied to an output transducer such as a conventional hearing aid telephone 27r, 27l.

In each of hearing aid units 16 and 17 all of processing units 22 to 25 may be designed for automatic gain control (AGC), e.g. as disclosed in the above-mentioned copending international patent application PCT/DK97/000598.

In each of hearing aid units 16 and 17 the processing units 21 to 25 are thus interconnected via a multiplicity of internal information and control signals lines, whereas the only external connection to the other hearing aid unit is via the single directional communication link 28.

According to a particular aspect of the invention the signal lines connecting the processing units 21 to 23 of the right and left ear processing channels and the simulated left and right ear processing channels to the binaural processing units 24 and 25 may be opened and closed or activated and deactivated by control of appropriate switching means, not illustrated, whereby an advantageous adjustment flexibility is obtained with

a smooth transition ranging from full binaural signal processing approximating the sound information processing of the human brain via a more simple binaural sound level control to conventional monaural sound reproduction, contrary to the separation of the processing units for normal and binaural processing in the prior art system of WO 97/14268 explained above.

The signals supplied in each of the right and left ear hearing aid units to the binaural signal processing in units 24 and 25 may as illustrated in figure 2 be microphone signals which have been preprocessed by conversion into digital form and correction of frequency and/or level distortions caused by unlinearities in the microphone circuits and/or following from the arrangement in the user's ear channel. Preferably, the incoming signals for the binaural processing have been filtered to the desired frequency band width.

Moreover as shown in figure 2 the preprocessed microphone signal supplied from each of hearing aid units 16 and 17 to the simulated processing channel of the other unit may be limited, e.g. by compression in an additional compressor unit 28r, 28l, to reduce the dynamic range, the band width and/or the number of samplings, thereby reducing the amount of data or information to be processed by the simulated processing. A similar signal limitation may also be provided, e.g. by an additional compressor unit 29r, 29l for the signals supplied in each hearing unit from the signal processing channel for the right or left ear, respectively, to binaural signal processing. In either case the processing units, to which such compressed signals are supplied, must be designed for processing these signals.

The binaural signal processing effected by processing units 24 and 25 may comprise a level correc-

tion, by which the gain in the hearing aid unit, right or left, receiving the weakest incoming sound signal is controlled on the basis of the incoming sound signal at the other hearing aid unit as represented, e.g. by the preprocessed microphone signal communicated therefrom via communication link 28 for simulated signal processing. Thereby, the sound level ratio between sounds received by the right and left ears, respectively. and the spatial information provided thereby may be maintained also for hearing aid units with automatic gain control (AGC), sin AGC control can be effected on the basis of the strongest processed signal in the right or left ear units 16 and 17

The complexity allowed for the internal signal processing in each of hearing aid units 16 and 17 would also allow a compensation for time delays introduced e.g. by the signal communication via communication link 28, if necessary.

For hearing aids which as shown in figure 2 employs sound or signal processing in a number of separate frequency bands with automatic gain control in each band the processing complexity and/or capacity further entails a data or information transfer between the real and simulated processing channels in each unit to provide for equal adaptation of the gain control of these processing channels, whereby the overall transfer function of each hearing aid unit may be adapted to take account of sound spectrum differences occurring at the right and left sides, thereby taking account of the frequency distribution in the spectra of sounds received at the right and left ears, which is very important for the localization of a sound source in space.

Since, in each of the right and left ear hearing aid units 16 and 17 in figure 2 all processing units

19r - 23r, 19l - 23l for the actual right or left ear signal processing channel and all processing units 21ls - 23ls, 21rs - 23rs for the simulated signal processing channel are programmed or otherwise adjusted to the specific processing parameters for the right and left ear signal processing, respectively, or vice versa, the binaural signal processing in each side takes fully account of the specific hearing loss characteristics of both sides up to or even beyond the output signal transducer 26r and 26l, respectively.

To accomplish this the binaural signal processing effected in the right and left ear hearing units 16 and 17 will typically be mirror images of each other to restore the actual sound level and sound spectrum differences between incoming sounds at the right and left ears, respectively.

As a special advantage of the binaural signal processing in each of the two hearing units of the system according to the invention a sophisticated noise or feedback suppression is made possible, by which tone signals deviating from the overall sound image may be effectively suppressed without suppression of tone signals present in the overall sound signal or in the right and left side at the same time. This can be accomplished by including in the binaural sound processing unit a feedback suppression system to which a residual feedback signal representing the difference feedback signals from the actual and simulated sound processing channels is supplied. By means of such a feedback suppression it is possible for the hearing aid system of the invention to distinguish between howl and information sound signals of a similar character such as a tone from a flute solo in classical music composition or alarm or signalling tones such as walk/stop beeps at traffic lights.

In each of hearing aid units 16 and 17 the performance of each of signal processing units 21r - 23r, 21l - 23l in the real signal processing channel as well as the performance of each of the processing units 21ls - 23ls, 21rs - 23rs in the simulated processing channel is controlled by adjustment parameters or data adapted to the specific compensation requirements of the right and left ears, respectively.

According to the invention such adjustment parameters may be individually programmable to compensate for the user's specific hearing impairment with respect to the right and left ears, whereby the hearing aid system may be supplied with a standard adjustment to permit individual programming to be effected by a hearing aid fitter as is customary practice in the individual user adjustment of hearing aids.

Moreover, the adjustment parameters may be organized in different programme settings to permit operation of the hearing aid system in different modes ranging from fully binaural to simple monaural operation of the hearing aid units and/or permit adaption of the hearing aid system to varying sound environments or listening situations.

Figure 3 shows for one of the hearing aid units in the system illustrated in figure 2, i.e. the right ear hearing aid unit 16 how this is accomplished by means of a performance and programme memory 30 in which all programmable adjustment parameters for a number of specific performance programmes are entered and may be selected from a selection unit 31 which may be user operated and/or operated from a sound signal analyzer 32 to effect programme selection automatically in response to occurrence of specified sound signal conditions.

Optionally, at least one of hearing aid units 16

or 17 may include means for calculation of intermediate settings between at least two consecutive performance programme settings, in which case also such intermediate settings will be selectable from the selection unit 31.

Since as a result of the structure and organiza - tion of hearing aid units 16 and 17 to effect not only the actual signal processing for the ear in which the unit is arranged, but also the simulated processing for the opposite ear memory 30 will contain all adjustment parameters needed for the signal processing for both sides, the programming of the hearing aid system may be effected by entering of adjustment parameters and user operated or automatically activated performance programmes in memory 30 of the one of hearing aid units 16 and 17 only and effecting transfer of adjustment parameters for the processing units of the other hearing aid unit via the communication link 28 in an adjustment or initiation mode activated at each change of performance programme.

Ultimately this makes possible to operate the system according to the invention by a master-slave principle, in which case one of the hearing aid units would function as a master unit and take control of the other unit functioning as a slave unit in which memory 30 would then contain the information or parameters needed for the actual function of the slave unit.

Alternatively, various kinds of intermediate or mixed organization schemes could be foreseen, e.g. by designing both hearing units with user operated as well as automatic programme selection. This could provide e.g. for consensus operation in situations where one unit would try to shift automatically to a specific programme matching prevailing sound signal conditions, by effecting an exchange of actual adjustment parameter

settings between the two units via communication link 28 to enable a decision to be made in one of the units as to whether the programme selected by one of the units should be effected for both units.

In each of the two units synchronization means 33 may further be provided for the exchange of synchronization information between the signal processing parts of the two units via the communication link 28. Such synchronizing information may be derived from the signals otherwise transferred between the two units or be generated as separate synchronizing signals.

User operability may advantageously be effected by wireless remote control from a separate control unit carried by the user. This is suitable, in particular, for embodiments in which wireless transmission is already used for the bidirectional communication link between the two hearing aid units.

P A T E N T C L A I M S

1. A binaural digital hearing aid system comprising two hearing aid units for arrangement in a user's left and right ear, respectively, each of said units comprising input signal transducer means for conversion of a received input sound signal into an analog input signal, A/D conversion means for conversion of said analog input signal into a digital input signal, digital signal processing means for processing said digital input signal and generating a digital output signal, D/A conversion means for conversion of said digital output signal into an analog output signal and output signal transducer means for conversion of said analog output signal into an output sound signal perceivable to the user, a bidirectional communication link being provided between said units to connect a point in the signal path between the input signal transducer means and the digital signal processing means in one of said units with a corresponding point in the signal path between the input signal transducer means and the digital signal processing means of the other of said units, c h a r a c t e r i z e d in that the digital signal processing means of each unit (1, 2;16, 17) is arranged to effect a substantially full digital signal processing including individual processing of signals from the input transducer means (3r, 3l; 18r, 18l) of the actual unit and simulated processing of signals from the input transducer means of the other unit as well as binaural signal processing of signals supplied, on one hand, internally from the input signal transducer means of the same unit and, on the other hand, via said communication link (7; 28) from the input signal transducer means of the other unit, said digital signal processing means including at least a first digital signal processor part (5r, 5l, 12r, 12l;

21r-23r, 21l-23l) for processing said internally supplied signal, a second digital signal processor part (6l, 6r, 13l, 13r; 21ls-23ls, 21rs-23rs) for processing the signal supplied via said communication link (28) and a third digital signal processor part (9r, 9l; 24r-25r, 24l-25l) to effect common binaural digital signal processing of information derived from the signals processed in said first and second digital signal processor parts, said second digital signal processor part (6l, 6r, 13l, 13r; 21ls-23ls, 21rs-23rs) in each unit simulating the first digital signal processor part (5l, 5r, 12l, 12r; 21l-23l, 21r-23r) in the other unit with respect to adjustment parameters controlling the performance of said first signal processor part in said other unit.

2. A hearing aid system as claimed in claim 1, c h a r a c t e r i z e d in that said bidirectional communication link (7) is a wireless transmission link, transceiving means (8r, 8l) and antenna means (7r, 7l) being provided in each of said hearing aid units.

3. A hearing aid system as claimed in claim 2, c h a r a c t e r i z e d in that said antenna means (7r, 7l) in each hearing unit comprises a short piece of antenna wire projecting externally from an ear mold arranged in ITE or CIC position and acting simultaneously as a withdrawal string.

4. A hearing aid system as claimed in claim 1, 2 or 3, c h a r a c t e r i z e d in that in at least one of said hearing aid units (16, 17) adjustment parameters for processing units in an actual signal processing channel effecting signal processing adapted to the ear in which the unit is arranged as well as further adjustment parameters for the simulated signal processing channel effecting said simulated signal processing adapted to the ear in which the other unit

is arranged are entered into a memory (30).

5. A hearing aid system as claimed in claim 4, c h a r a c t e r i z e d in that said memory (30) further includes performance programme settings to provide varying operation modes ranging from fully binaural signal processing to simple monaural signal processing in the hearing aid units and/or provide adaption of system performance to different sound environments or listening situations.

6. A hearing aid system as claimed in claim 5, c h a r a c t e r i z e d in that at least one of said hearing aid units (1, 2; 16, 17) include means for calculation of intermediate settings between at least two consecutive performance programme settings.

7. A hearing aid system as claimed in claim 5 or 6, c h a r a c t e r i z e d in that programme selection means (31) for said performance programme settings including said intermediate settings are user operated and/or automatically activated on the basis of incoming sound signal analysis.

8. A hearing aid system as claimed in any of claims 2 or 3 and claim 7, c h a r a c t e r i z e d in that said selection means (31) are remotely controlled by wireless transmission of control signals.

9. A hearing aid system as claimed in any of claims 5 to 8, c h a r a c t e r i z e d in that said memory (30) is provided in one hearing aid unit only acting as a master unit, said bidirectional communication link (28) being designed for transfer of adjustment parameter settings from said one unit to the other unit acting as a slave unit.

10. A hearing aid system as claimed in any of claims 6, 7 or 8, c h a r a c t e r i z e d in that each of said hearing aid units (16, 17) is provided with said memory (30) and said programme selection

means (31), said selection means (31) being designed in both units for user operation as well as automatic activation, said bidirectional communication link (28) being designed for exchange of parameter settings between the two units and one of the units including decision means to provide for a consensus control of the system.

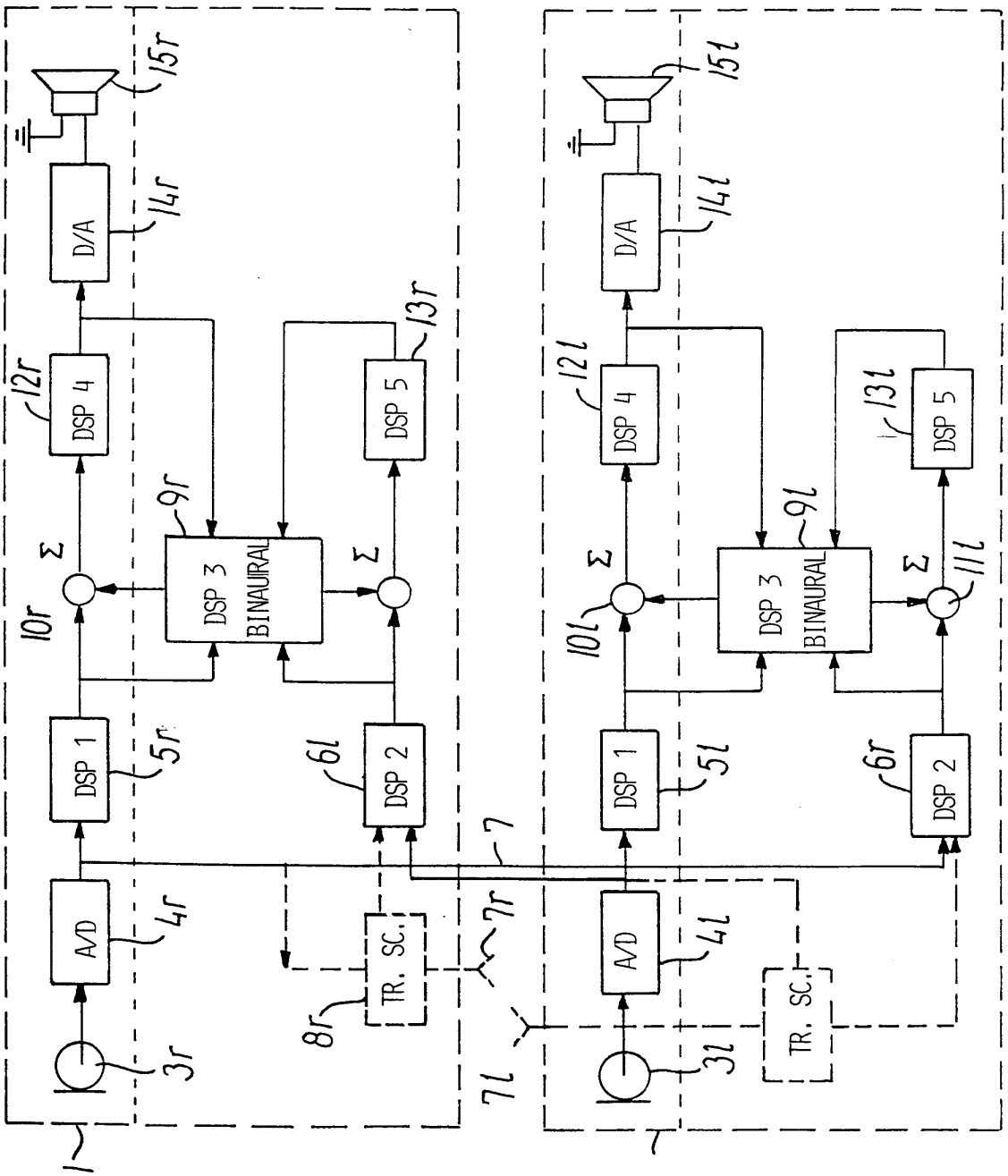
11. A hearing aid system as claimed in any of the preceding claims, c h a r a c t e r i z e d in that the binaural signal processing functions in each of the two hearing aid units are mirrored with respect to the binaural signal processing functions in the other unit to take account of sound level and/or sound spectrum differences between incoming sound signal at the right and left ear hearing aid units.

12. A hearing aid system as claimed in claim 11, c h a r a c t e r i z e d in that the binaural signal processing part of each of said hearing aid units comprises a feedback howl suppression system providing howl suppression by processing of a residual signal representing the difference between feedback signals in the actual and simulated signal processing channels.

13. A hearing aid system as claimed in any of any of the preceding claims, c h a r a c t e r i z e d in that in each of the hearing aid units means are provided for limitation, e.g. by compression of the signals supplied from said first and second processor parts to said third processor part.

14. A hearing aid system as claimed in any of the preceding claims, c h a r a c t e r i z e d in that means (33) are provided in each of the two hearing aid units for exchange of synchronization information between the signal processing parts of the two units.

FIG.1



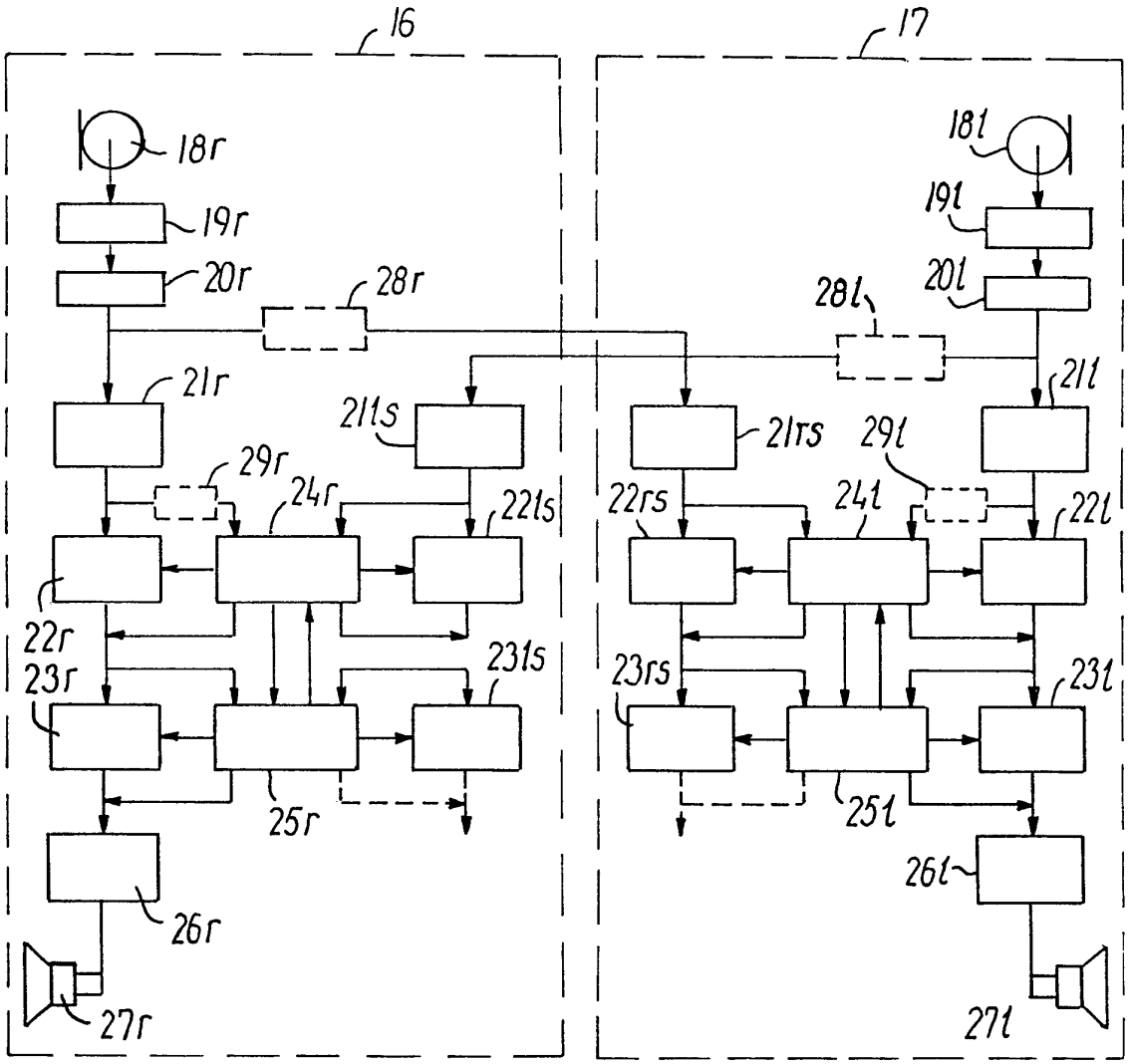

16
17
18r
19r
20r
28r
21r
211s
29r
24r
221s
22r
231s
23r
25r
26r
27r
181
191
201
28l
211
21rs
29l
22rs
241
221
23rs
231
251
261
271

FIG. 2

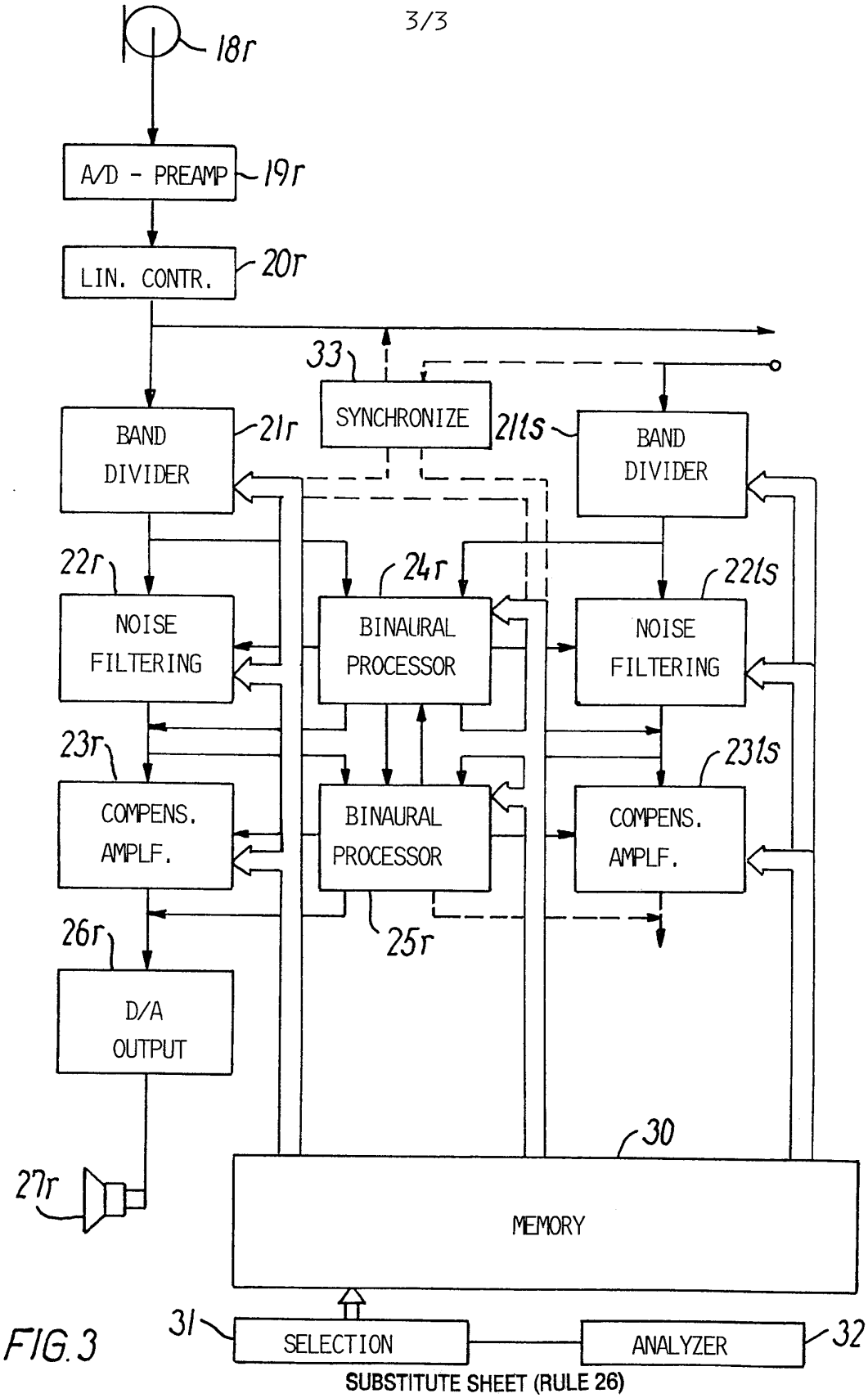

18r
A/D - PREAMP
19r
LIN. CONTR.
20r
33
SYNCHRONIZE
21r
BAND DIVIDER
21ls
BAND DIVIDER
22r
NOISE FILTERING
24r
BINAURAL PROCESSOR
22ls
NOISE FILTERING
23r
COMPENS. AMPLF.
BINAURAL PROCESSOR
25r
23ls
COMPENS. AMPLF.
26r
D/A OUTPUT
27r
30
MEMORY
31
SELECTION
ANALYZER
32

FIG. 3

INTERNATIONAL SEARCH REPORT

In itional Application No
PCT/DK 98/00062

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04R25/00

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)
IPC 6 H04R

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 practical, 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.
A	WO 97 14268 A (AUDIOLOGIC INC) 17 April 1997 cited in the application see page 3, line 5 - line 25 see page 4, line 29 - page 7, line 6; figures ---	1
A	US 5 511 128 A (LINDEMANN) 23 April 1996 see column 2, line 35 - column 4, line 19; figures ---	1
A	US 5 479 522 A (LINDEMANN ET AL.) 26 December 1995 cited in the application see column 2, line 11 - line 59 see column 3, line 50 - column 6, line 20 --- -/--	1

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in 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
3 November 1998	11/11/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gastaldi, G

2

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int tional Application No
PCT/DK 98/00062

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 608 803 A (MAGOTRA ET AL.) 4 March 1997 see column 1, line 45 - column 2, line 46	1

2

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/DK 98/00062

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9714268 A	17-04-1997	US 5757932 A	26-05-1998
		AU 7248796 A	30-04-1997
		EP 0855130 A	29-07-1998
		NO 981559 A	08-06-1998
US 5511128 A	23-04-1996	AU 1833395 A	08-08-1995
		EP 0740893 A	06-11-1996
		WO 9520305 A	27-07-1995
US 5479522 A	26-12-1995	US 5757932 A	26-05-1998
US 5608803 A	04-03-1997	NONE	

Form PCT/ISA/210 (patent family annex) (July 1992)

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 July 2006 (13.07.2006)

PCT

(10) International Publication Number
WO 2006/074082 A1

(51) International Patent Classification:
G06K 9/00 (2006.01) *A61B 5/117* (2006.01)

(21) International Application Number:
PCT/US2005/047449

(22) International Filing Date:
28 December 2005 (28.12.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/028,873 4 January 2005 (04.01.2005) US

(71) Applicant *(for all designated States except US)*: **MOTOROLA, INC.** [US/US]; 1303 East Algonquin Road, Schaumburg, Illinois 60196 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ZUREK, Robert A.,** [US/US]; 1055 Autumn Drive, Anticoh, Illinois 60002 (US). **CHARLIER, Michael L.,** [US/US]; 931 Saratoga, Palatine, Illinois 60067 (US). **JOHNSON, John,** [US/US]; 9409 Nicholas Lane, Spring Grove, Illinois 60081 (US). **KOO, Simone,** [IT/US]; 1434 Joan Drive, Palatine, Illinois 60074 (US).

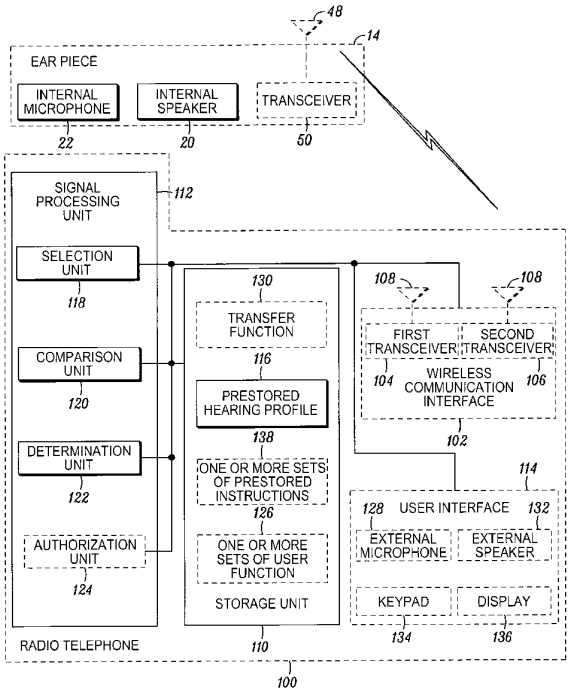
(74) Agents: **CHAPA, Lawrence J.,** et al.; 600 North Us Highway 45, Libertyville, Illinois 60048 (US).

(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, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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,

[Continued on next page]

(54) Title: A SYSTEM AND METHOD FOR DETERMINING AN IN-EAR ACOUSTIC RESPONSE FOR CONFIRMING THE IDENTITY OF A USER



(57) Abstract: A system and method is provided for confirming the identity of a user, where the confirmation is made using biometric data. The system includes an internal speaker having an audio port, which is positioned within the ear of the user, and is adapted for producing a signal having one or more frequencies. The system further includes an internal microphone having an audio port, which is positioned within the ear of the user, and is adapted for detecting a resulting signal including the signal produced by the internal speaker and any corresponding signal reflections. A determination is then made as to whether the resulting signal detected by the internal microphone matches the corresponding predetermined expected signal, based upon a prestored hearing profile.

WO 2006/074082 A1

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).

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.

Published:
— *with international search report*

A SYSTEM AND METHOD FOR DETERMINING AN IN-EAR ACOUSTIC RESPONSE FOR CONFIRMING THE IDENTITY OF A USER

FIELD OF THE INVENTION

The present invention relates generally to the confirmation of the identity of a user using biometric information. More particularly, the present invention relates to confirming the identity of the user using an in-ear acoustic response.

BACKGROUND OF THE INVENTION

Many systems have been concerned with confirming the identity of a user, and several of these have incorporated one or more of various biometric signatures, which are believed to be unique for each individual. Examples of biometric information, which has historically been used to confirm the identity of a user, include scans of finger prints, retinal scans, and voice prints.

Once a user's identity has been determined, the determination can be used to enable/disable access to a device or functions for use with a device. Furthermore, a user's identity can be used to load a set of user preferences, when the particular user is operating the device. Still further a secure manner of establishing one's identity can be used to authenticate the source information communicated by the user of a device.

Security mechanisms, generally need to have safeguards in place, which will limit their circumvention. However, it is generally understood that no security mechanism is fool proof. If one is given enough time and sufficient resources, most security mechanisms can be overcome. Correspondingly, a security minded individual will often try to make the task of overcoming the security mechanisms as difficult as possible.

At least one manner, which has been used to enhance various security mechanisms involves implementing locks which do not have a static key, i.e. a fixed solution such a password. In some instances an algorithm can be used in connection with a pseudo random set of inputs, where the response as determined by the algorithm is dependent upon the pseudo random set of inputs, which are extremely unlikely to be anticipated, as well as can similarly be set up so as to be unlikely to

ever repeat. Consequently any previously intercepted valid responses would likely be invalid and/or of limited assistance relative to a determination of future responses. As a result, one would need to be able to decode the algorithm and intercept the future pseudo-random inputs before a valid future response could be generated. Still further where the algorithm incorporates the variability associated with a user's biometric information, one would presumably need access to the individual as part of the decomposition of the algorithm. Biometric data generally is rather complex and so for many types of biometric data does not readily decode based upon casual observance, and therefore any attempt to decode would be relatively fairly intrusive, and consequently the ability to use biometric information as part of an algorithm for use as a security mechanism, would make one's attempts to decode the algorithm fairly obvious.

SUMMARY OF THE INVENTION

The present invention provides for a system for confirming the identity of a user. The system includes an internal speaker, which is positioned within the ear of the user, and is adapted for producing a signal having one or more frequencies. The system further includes an internal microphone, which is positioned within the ear of the user, and is adapted for detecting a resulting signal including the signal produced by the internal speaker and any corresponding signal reflections. The system still further includes a storage unit, which includes a prestored hearing profile, which is adapted for predicting an expected signal, resulting from a produced signal having one or more frequencies, and a signal processing unit, coupled to the internal speaker, the internal microphone, and the storage unit. The signal processing unit includes a selection unit for selecting one or more frequencies to be produced, a comparison unit for comparing the resulting signal with a predetermined expected signal, and a determination unit for determining if the resulting signal detected by the internal microphone matches the corresponding predetermined expected signal.

In at least one embodiment, the signal produced by the internal speaker has multiple frequencies, wherein the detected resulting signal includes one or more

distortion product otoacoustic emission tones generated by an inner portion of the ear having a nonlinear response.

In at least a further embodiment, the signal processing unit further includes an authorization unit adapted for providing access to one or more sets of user functions, if a match is identified by the determination unit

The present invention further provides a method for confirming the identity of a user using an audio receiving system including an internal microphone and an internal speaker proximate an ear of a user. The method includes selecting one or more frequencies. A signal is then produced by the internal speaker having the one or more frequencies. A resulting signal including the signal produced by the internal speaker and any corresponding signal reflections is then received. The received resulting signal is then compared with a predetermined expected signal, where a determination is made as to whether a substantial match exists between the received resulting signal and the predetermined expected signal.

These and other features, and advantages of this invention are evident from the following description of one or more preferred embodiments of this invention, with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE FIGURES

The present invention will be described by way of exemplary embodiments, but not limitations, illustrated in the accompanying drawings in which like references denote similar elements, and in which:

FIG. 1 is a cut away view of a user's ear illustrating an ear piece, in accordance with at least one embodiment of the present invention, inserted at least partially within the ear canal of the user;

FIG. 2 is a side plan view of the ear piece, illustrated in FIG. 1, which highlights the positional placement of some of the internal elements, in accordance with at least one embodiment of the present invention;

FIG. 3 is a block diagram of a system for confirming the identity of a user incorporating an ear piece and a radio telephone, in accordance with at least one embodiment of the present invention; and

FIG. 4 is a flow diagram of a method for confirming the identity of a user using an audio receiving system including an internal microphone and an internal speaker proximate an ear of a user, in accordance with at least one embodiment of the present invention.

DETAILED DESCRIPTION

As required, detailed embodiments of the present invention are disclosed herein, however, it is to be understood that the disclosed embodiments are merely exemplary of the invention, which can be embodied in various forms. Therefore, specific structural and functional details disclosed herein are not to be interpreted as limiting, but merely serve as a basis for the claims and as a representative basis for teaching one skilled in the art to variously employ the present invention in virtually any appropriately detailed structure. Further, the terms and phrases used herein are not intended to be limiting, but rather, to provide an understandable description of the invention.

FIG. 1 illustrates a cut away view 10 of a user's ear 12 illustrating an ear piece 14, in accordance with at least one embodiment of the present invention, inserted at least partially within the ear canal 16 of the user.

FIG. 2 illustrates an enlarged side plan view of the ear piece 14, shown in FIG. 1, which highlights the positional placement of some of the internal elements, in accordance with at least one embodiment of the present invention. The ear piece 14 includes an internal speaker 20 and an internal microphone 22, which are both coupled to an audio port 24, corresponding to an opening in the external housing, via a trunk 26 having first branch 28 and a second branch 30. The first branch 28 and the second branch 30 join at the trunk 26 via an acoustic coupling 32. The first branch 28 couples the internal microphone 22 to the audio port 24, and the second branch 30 couples the internal speaker 20 to the audio port 24.

The separate branches 28, 30 provide a fixed degree of acoustic coupling between the internal speaker 20 and the internal microphone 22. Furthermore, to the extent that the two elements continue to potentially interfere with one another, the known geometry and the static nature of the acoustic coupling 32 enables one to

determine and/or reasonably predict the amount of interference, and to take steps to filter and/or remove the interference.

In at least one embodiment, the internal microphone 22 is an omni-directional electret microphone, and the internal speaker 20 is a balanced armature loudspeaker

The output audio port 24 is arranged relative to the external housing, so as to exit the ear piece 14 into a user's ear canal 16, when the ear piece 14 is at least partially inserted. Proximate the output audio 24, the external housing is constructed and arranged as a pair of cascading (i.e. laterally offset in the direction of insertion 34) semi-spheres 36 made of a pliable material having their convex curved surface 38 which faces toward the ear canal 16 upon insertion. When placed within the ear canal 16 of the user, the semi-spheres 36 made of pliable material will flex so as to conform to the shape of the ear canal 16, thereby producing a seal, which limits the amount of sound that will traverse across the external surface of the semi-spheres 36. The sealing, generally does not affect the transmission of sound via the internal audio porting formed from the previously identified trunk 26 and the first and second branches 28, 30, which travel through the semi-spheres 36 proximate their center.

The ear piece 14 has a neck 40 located between the semi-spheres 36 and a cylindrical body 42, wherein the width of the neck 40 widens closer to the cylindrical body 42. Conversely, the neck 40 tapers toward the point of initial insertion, where the diameter of the neck 40, where the neck 40 meets the semi-spheres 36 is narrower than the diameters of the semi-spheres 36. The narrower diameter of the neck 40 proximate the semi-spheres 36 provides the outer surface of the semi-spheres 36 room to flex in connection with making a seal.

The body 42 of the ear piece 14 includes a battery 44, which can be re-charged through a port in the body 42 (not shown), as well as includes a substrate 46 upon which electronic circuit elements can be positioned or placed, such as a transceiver 50 (FIG. 3) for short range communication, e.g. a Bluetooth transceiver. A nub at the end of the body 42 houses an antenna 48 for use in transmitting and receiving signals via the transceiver 50.

The transceiver 50 provides at least one method for transmitting signals detected by the internal microphone 22 and receiving signals to be reproduced by the

internal speaker 20. Still further, the conveyance of other signals would similarly be possible. One skilled in the art will readily appreciate, that signals could be alternatively and/or additionally communicated to and from the ear piece 14, via a hard wired connection.

FIG. 3 illustrates a block diagram of a system for confirming the identity of a user incorporating a radio telephone 100 and an ear piece 14, which in at least one embodiment is consistent with the ear piece 14 illustrated in FIGS.1 and 2,.

In the illustrated embodiment, the radio telephone 100 includes a wireless communication interface 102, via which the radio telephone 100 is communicatively coupled to the ear piece 14. In the illustrated embodiment, the radio telephone includes a first transceiver 104 and a second transceiver 106, each of which is coupled to a corresponding antenna 108. In at least one embodiment, the first transceiver 104 is adapted for communicating within a cellular network, and the second transceiver 106 is adapted for short range communication, of the type incorporated as part of the ear piece 14 (e.g. Bluetooth). However one skilled in the art will recognize that the present invention is not dependent upon the exact nature of the communication capabilities, relative to the type of technology and/or the protocol employed, but are generally intended to enable signals to be communicated between the various elements.

The radio telephone 100 additionally includes a storage unit 110, a signal processing unit 112, and in accordance with at least some embodiments, a user interface 114. The storage unit 110 includes a prestored hearing profile 116 which can be used for confirming the identity of a user. In at least one instance the hearing profile includes previously measured responses including the non-linear response of the reflected signals detected by the internal microphone 22, in response to producing one or more predetermined frequencies by the internal speaker 20. The hearing profile 116 can include both frequency and amplitude information relative to the reflected signals, in response to the known stimulus pattern corresponding to the produced signal.

The signal processing unit 112 includes a selection unit 118, which is adapted for selecting one or more frequencies to be produced. The signal processing unit 112

additionally includes a comparison unit 120, which is adapted for comparing the resulting signal detected by the internal microphone 22 and comparing the resulting signal with a predetermined expected signal. The signal processing unit 112 still further includes a determination unit 122, which is adapted for determining if the resulting signal detected by the internal microphone 22 matches the corresponding predetermined expected signal.

For example, when a pair of frequencies are produced within the ear, a distortion product otoacoustic emission tone is produced, which has a frequency that is equivalent to two times the lower one of the two supplied frequencies minus the higher one of the two supplied frequencies. One would additionally expect the resulting signal would additionally include a signal component having a frequency, which matches each of the two supplied signal, as well as the additional distortion product otoacoustic emission tone. By comparing the signal characteristics of the response signal including the amplitude of each of the signal components, at least a potential partial match of the user can be determined. The original supplied frequencies could be determined using a pseudo-random number seed generator.

In at least some embodiments, the signal processing unit 112 will additionally include an authorization unit 124, which is adapted for providing access to one or more sets of user functions 126, if a match is identified. The one or more sets of user functions 126 could be stored within the storage unit 110.

In addition to the use of a prestored hearing profile for purposes of confirming the identity of the user, the system could additionally compare the user's voice received by the internal microphone 22, positioned within the user's ear, with an external microphone 128, positioned external to the user ear. In at least one embodiment, the system uses the microphone incorporated as part of the radio telephone as the external microphone 128. Alternatively, the external microphone could be incorporated within the ear piece 14 having an audio port having an opening to the external world, which is external to the sealed portion of the ear 12. Still further, the external microphone can be positioned at the end of a boom, which could extend away from the ear piece 14 and toward the user's mouth. One skilled in the art

will recognize that still further arrangements are possible without departing from the teachings of the present invention.

The corresponding signal received via the internal microphone 22 could be compared relative to the signal received via the external microphone 128, and the corresponding difference could be compared relative to a prestored transfer function 130, in order to make further determinations concerning the identity of the user.

Still further identifying characteristics could be measured and compared against previously determined base line measurements. As a further example one or more sensors could be placed along the external surface of the ear piece 14 housing at a location, which is proximate relative to the position where the ear piece 14 will produce a seal relative to the ear 12. The sensors could additionally be used to test the integrity of any seal. The sensors could be one or more of a pressure or proximity sensor.

In addition to the external microphone 128, which forms part of the user interface 114 of the radio telephone 100, in accordance with the illustrated embodiment, the user interface 114 additionally can include an external speaker 132, a keypad 134, and/or a display 136, as well as other elements which facilitate user interaction with the radio telephone 100.

In at least some embodiments, at least a portion of the signal processing unit 112 including the corresponding selection unit 118, comparison unit 120, determination unit 122, and authorization unit 124 can take the form of one or more sets 138 of prestored instructions, which are executed by one or more microprocessors. The one or more sets 138 of prestored instructions can be maintained in the storage unit 110, which in some instances can be separate from or integrated as part of any microprocessor, which might form all or part of the signal processing unit 112. Still further, one or more sets 138 of prestored instructions can include programming or software instructions and/or microcode. In the same and/or other embodiments, all or parts of the various elements including the selection unit 118, comparison unit 120, determination unit 122, and authorization unit 124 can be alternatively and/or partially implemented in hardware using discrete or integrated circuit elements, including logic gates and/or sequential state machines. However,

one skilled in the art will readily appreciate that still further approaches are possible for implementing the same without departing from the teachings of the present invention.

The storage element 110 can include both volatile and non-volatile forms of memory, many types of which are well known to one skilled in the art. One such example includes a semiconductor memory, for example a read only memory. Other examples include a fixed drive or hard drive, which commonly stores information magnetically. However any form of data storage may be suitable, as long as it supports the ability to at least initially store, and later retrieve the information contained therein.

Furthermore, while the signal processing unit 112 and the storage unit 110 were described in connection with a corresponding radio telephone 100, which is communicatively coupled to the ear piece 14, one skilled in the art will readily appreciate that the signal processing unit 112 and the storage unit 110 could alternatively be incorporated as part of a type of device, which is other than a radio telephone 100. The signal processing unit 112 and the storage unit 110 could similarly be incorporated as part of the ear piece 14, itself.

FIG. 4 illustrates a flow diagram of a method 200 for confirming the identity of a user using an audio receiving system including an internal microphone and an internal speaker proximate an ear of a user, in accordance with at least one embodiment of the present invention. The method 200 includes selecting 202 one or more frequencies, and producing 204 a signal by the internal speaker having the one or more frequencies. A resulting signal is then received 206, which includes the signal produced by the internal speaker and any corresponding signal reflections. The received resulting signal is then compared 208 with a predetermined expected signal, and a determination 210 is made whether a substantial match exists between the received resulting signal and the predetermined expected signal.

In at least some embodiments of the present invention, a confirmation of the identity will be made based upon the results of the comparison between the received resulting signal and the predetermined expected signal.

However in at least further embodiments, the results of the comparison between the received resulting signal and the predetermined expected signal will be supplemented with additional criteria from which a match can be determined, such as a comparison of the measured difference between a received signal, corresponding to the user's voice, received by each of the internal microphone and the external microphone, with a predetermined expected result, which could be derived from a prestored transfer function. In these instances the method 200 will further include receiving 214 an audio signal produced by the user via an internal microphone and an external microphone, and then comparing 216 the difference between the signal received via the internal microphone and the signal received via the external microphone with the difference predicted by a predetermined transfer function. A determination 218 is then made relative to whether a difference in the signal received via the internal microphone and the signal received via the external microphone, matches the difference predicted by the predetermined transfer function.

In this instance, if a substantial match is determined to exist between the received resulting signal and the predetermined expected signal, and a substantial match is determined to exist between the difference in the signal received via the internal microphone and the signal received via the external microphone and the difference predicted by the predetermined transfer function, then the identity of the user is confirmed 212.

While the preferred and other embodiments of the invention have been illustrated and described, it will be clear that the invention is not so limited. Numerous modifications, changes, variations, substitutions, and equivalents will occur to those of ordinary skill in the art without departing from the spirit and scope of the present invention as defined by the following claims.

WHAT IS CLAIMED IS:

1. A system for confirming the identity of a user comprising:

an internal speaker, positioned within the ear of the user, adapted for producing a signal having one or more frequencies;

an internal microphone, positioned within the ear of the user, adapted for detecting a resulting signal including the signal produced by the internal speaker and any corresponding signal reflections;

a storage unit, which includes a prestored hearing profile, which is adapted for predicting an expected signal, resulting from a produced signal having one or more frequencies; and

a signal processing unit, coupled to the internal speaker, the internal microphone, and the storage unit, the signal processing unit including:

a selection unit for selecting one or more frequencies to be produced;

a comparison unit for comparing the resulting signal with a predetermined expected signal; and

a determination unit for determining if the resulting signal detected by the internal microphone matches the corresponding predetermined expected signal.

2. A system in accordance with claim 1 wherein the signal produced by the internal speaker has multiple frequencies, wherein the detected resulting signal includes one or more distortion product otoacoustic emission tones generated by an inner portion of the ear having a nonlinear response.

3. A system in accordance with claim 1 wherein the internal microphone and internal speaker are incorporated as part of an earpiece.

4. A system in accordance with claim 3 wherein the earpiece includes an earbud, which fits at least partially within an ear canal of the user.

5. A system in accordance with claim 4 wherein the earbud includes a housing having an external surface, which provides a seal against the inner wall of the ear canal, when at least a portion of the earbud is firmly placed within the ear canal of the user.

6. A system in accordance with claim 5 wherein the external surface includes one or more pressure sensors for detecting the integrity of the seal against the inner wall of the ear canal.

7. A system in accordance with claim 6, wherein the pressure sensors are adapted to detect the geometry of the ear of the user.

8. A system in accordance with claim 5 wherein the earbud includes an audio port coupled to the internal microphone having an opening at the external surface of the housing along the portion of the external surface which is located within the ear canal of the user.

9. A system in accordance with claim 8 wherein the audio port includes a first branch and a second branch coupled to the same external surface opening, where the first branch of the audio port is coupled to the internal microphone and the second branch of the audio port is coupled to the internal speaker.

10. A system in accordance with claim 1 wherein the signal processing unit further includes an authorization unit adapted for providing access to one or more sets of user functions, if a match is identified by the determination unit.

11. A system in accordance with claim 1 further comprising an external microphone, coupled to the signal processing unit, adapted for receiving an audio signal, wherein said internal microphone is further adapted for receiving said audio signal, and wherein said storage unit additionally includes a transfer function, which represents the predetermined difference between a signal detected by each of the internal microphone and the external microphone, that is produced by the user, wherein the comparison unit is further adapted for comparing the difference of an audio signal that is produced by the user and detected by each of the internal microphone and the external microphone with the difference predicted by the transfer function, and the determination unit is further adapted for determining if the detected difference of the audio signal matches the difference predicted by the transfer function.

12. A system in accordance with claim 11 wherein the external microphone and the signal processing unit are incorporated as part of radio telephone.

13. A system in accordance with claim 1 wherein the one or more frequencies selected to be produced are selected based upon a substantially random seed value.

14. A method for confirming the identity of a user using an audio receiving system including an internal microphone and an internal speaker proximate an ear of a user, said method comprising:

selecting one or more frequencies;

producing a signal by the internal speaker having the one or more frequencies;

receiving a resulting signal including the signal produced by the internal speaker and any corresponding signal reflections;

comparing the received resulting signal with a predetermined expected signal; and

determining if a substantial match exists between the received resulting signal and the predetermined expected signal.

15. A method in accordance with claim 14 wherein, if a substantial match is determined to exist between the received resulting signal and the predetermined expected signal, then the identity of the user is confirmed.

16. A method in accordance with claim 15 wherein when the identity of the user is confirmed, access is provided to one or more sets of functions specific to the identified user.

17. A method in accordance with claim 14 wherein the predetermined expected signal is determined by sampling a resulting signal in response to producing a signal having each one of the frequencies from which the one or more frequencies are anticipated to be selected.

18. A method in accordance with claim 14 further comprising:

receiving an audio signal produced by the user via an internal microphone and an external microphone,

comparing the difference between the signal received via the internal microphone and the signal received via the external microphone with the difference predicted by a predetermined transfer function, and

determining if a substantial match exists between the difference in the signal received via the internal microphone and the signal received via the external microphone and the difference predicted by the predetermined transfer function.

19. A method in accordance with claim 18 wherein, if a substantial match is determined to exist between the received resulting signal and the predetermined expected signal, and a substantial match is determined to exist between the difference in the signal received via the internal microphone and the signal received via the external microphone and the difference predicted by the predetermined transfer function then the identity of the user is confirmed.

20. A method in accordance with claim 18 wherein the transfer function is determined by sampling previous examples of speech from the user by each of the internal microphone and the external microphone and computing the difference between the sampled examples of the same portions of speech.

21. A method in accordance with claim 20 wherein the transfer function is non-linear relative to frequency.

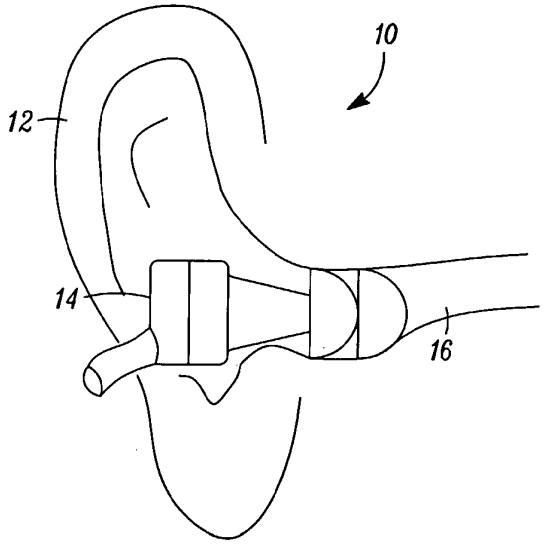


FIG. 1

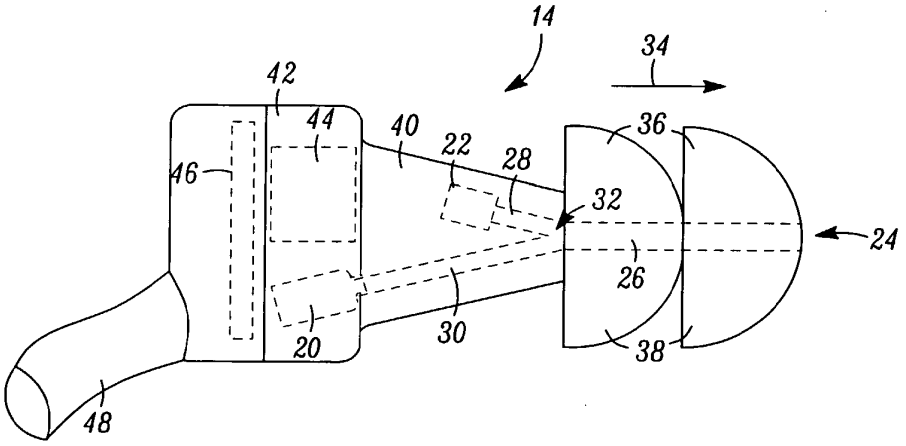


FIG.2

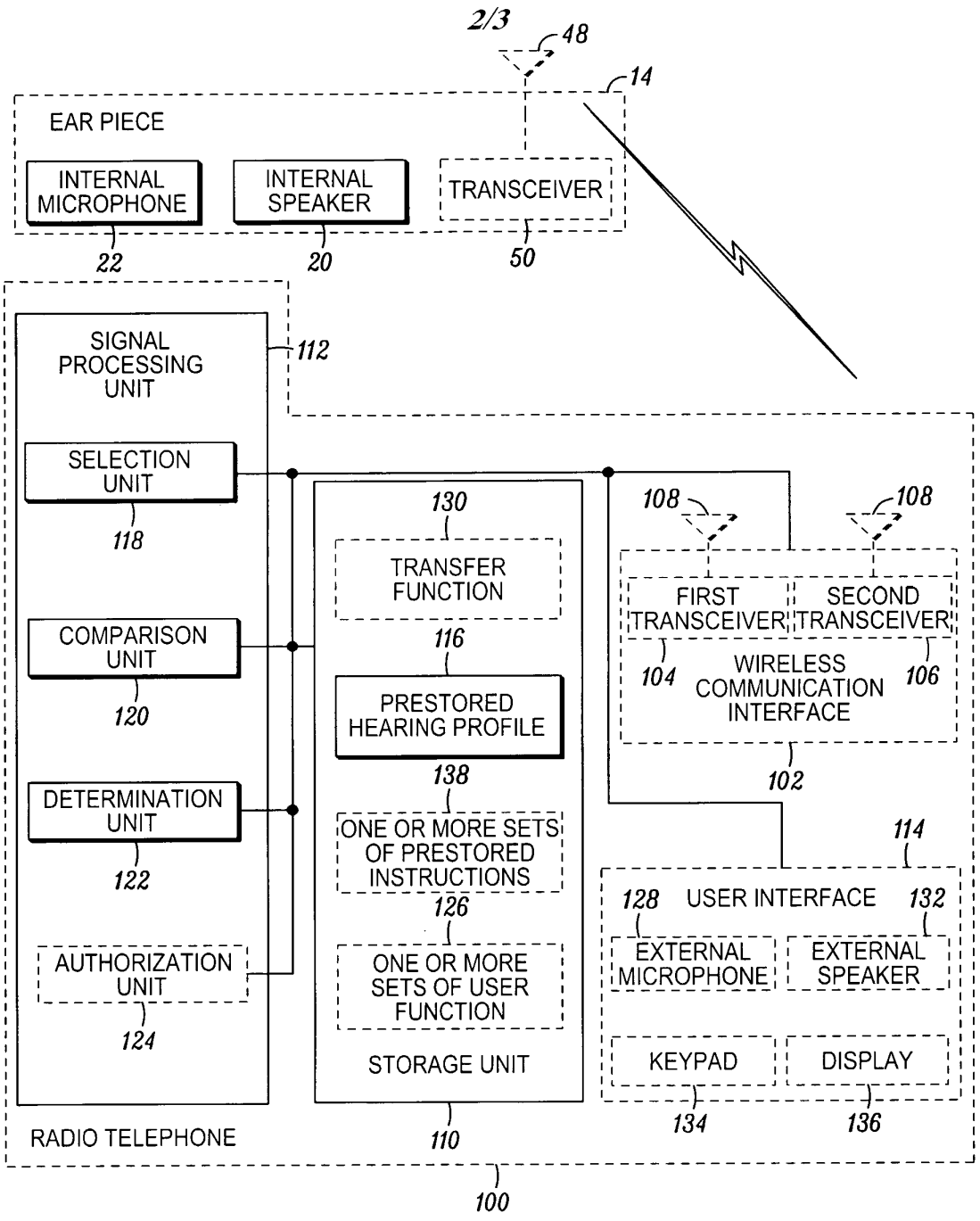

2/3
48
14
EAR PIECE
INTERNAL MICROPHONE
INTERNAL SPEAKER
TRANSCEIVER
22
20
50
SIGNAL PROCESSING UNIT
112
SELECTION UNIT
118
COMPARISON UNIT
120
DETERMINATION UNIT
122
AUTHORIZATION UNIT
124
130
TRANSFER FUNCTION
116
PRESTORED HEARING PROFILE
138
ONE OR MORE SETS OF PRESTORED INSTRUCTIONS
126
ONE OR MORE SETS OF USER FUNCTION
STORAGE UNIT
110
108
108
FIRST TRANSCEIVER
SECOND TRANSCEIVER
104
WIRELESS COMMUNICATION INTERFACE
106
102
114
128
USER INTERFACE
132
EXTERNAL MICROPHONE
EXTERNAL SPEAKER
KEYPAD
DISPLAY
134
136
RADIO TELEPHONE
100

FIG. 3

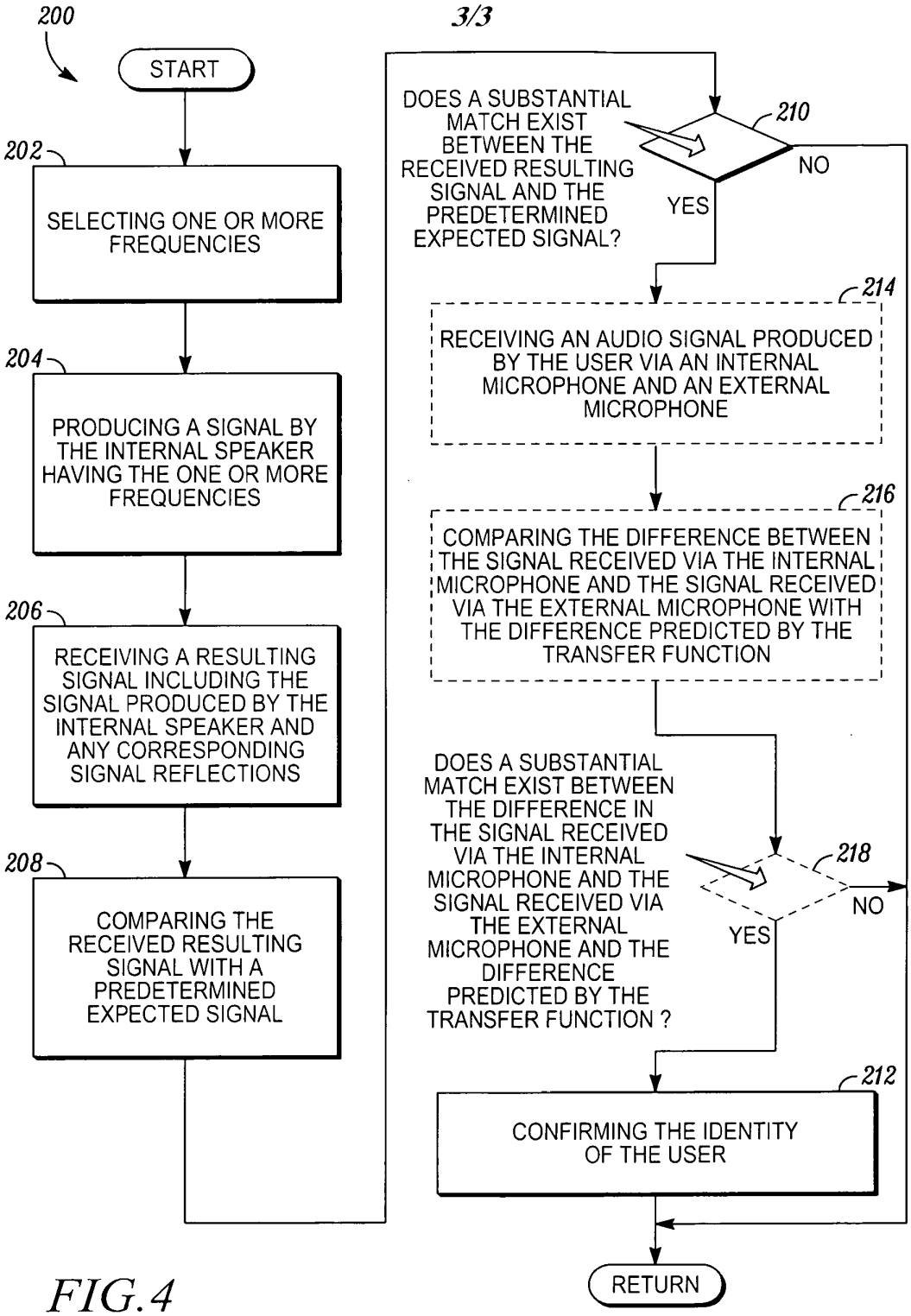

200
START
202
SELECTING ONE OR MORE FREQUENCIES
204
PRODUCING A SIGNAL BY THE INTERNAL SPEAKER HAVING THE ONE OR MORE FREQUENCIES
206
RECEIVING A RESULTING SIGNAL INCLUDING THE SIGNAL PRODUCED BY THE INTERNAL SPEAKER AND ANY CORRESPONDING SIGNAL REFLECTIONS
208
COMPARING THE RECEIVED RESULTING SIGNAL WITH A PREDETERMINED EXPECTED SIGNAL
DOES A SUBSTANTIAL MATCH EXIST BETWEEN THE RECEIVED RESULTING SIGNAL AND THE PREDETERMINED EXPECTED SIGNAL?
210
NO
YES
214
RECEIVING AN AUDIO SIGNAL PRODUCED BY THE USER VIA AN INTERNAL MICROPHONE AND AN EXTERNAL MICROPHONE
216
COMPARING THE DIFFERENCE BETWEEN THE SIGNAL RECEIVED VIA THE INTERNAL MICROPHONE AND THE SIGNAL RECEIVED VIA THE EXTERNAL MICROPHONE WITH THE DIFFERENCE PREDICTED BY THE TRANSFER FUNCTION
DOES A SUBSTANTIAL MATCH EXIST BETWEEN THE DIFFERENCE IN THE SIGNAL RECEIVED VIA THE INTERNAL MICROPHONE AND THE SIGNAL RECEIVED VIA THE EXTERNAL MICROPHONE AND THE DIFFERENCE PREDICTED BY THE TRANSFER FUNCTION ?
218
NO
YES
212
CONFIRMING THE IDENTITY OF THE USER
RETURN

FIG.4

INTERNATIONAL SEARCH REPORT

Inter al application No
PCT/US2005/047449

A. CLASSIFICATION OF SUBJECT MATTER
G06K9/00 A61B5/117

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)
G06K G07C A61B

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 practical, 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	EP 1 205 884 A (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD) 15 May 2002 (2002-05-15) paragraph [0039] - paragraph [0050]	1-5, 10-21
X	WO 02/091310 A (ORANGE PERSONAL COMMUNICATIONS SERVICES LIMITED; RODWELL, PHILIP, MAUR) 14 November 2002 (2002-11-14) page 13, line 1 - page 19, line 20	1-5, 10-21
A	US 5 787 187 A (BOUCHARD ET AL) 28 July 1998 (1998-07-28) abstract column 9, line 59 - column 10, line 9	1,14

☐ 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
22 March 2006	20/04/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Sonius, M

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

Inte il application No
PCT/US2005/047449

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1205884 A	15-05-2002	JP 2002143130 A	21-05-2002
		US 2002057805 A1	16-05-2002
WO 02091310 A	14-11-2002	EP 1388127 A1	11-02-2004
		GB 2375205 A	06-11-2002
		JP 2004532584 T	21-10-2004
		US 2004215968 A1	28-10-2004
US 5787187 A	28-07-1998	NONE	

Form PCT/ISA/210 (patent family annex) (April 2005)

Electronic Acknowledgement Receipt

EFS ID:	48591197
Application Number:	90019169
International Application Number:	
Confirmation Number:	7859
Title of Invention:	METHOD AND DEVICE FOR ACUTE SOUND DETECTION AND REPRODUCTION
First Named Inventor/Applicant Name:	11244666
Customer Number:	176134
Filer:	Jacob Andrew Snodgrass
Filer Authorized By:	
Attorney Docket Number:	0027071.00133
Receipt Date:	15-SEP-2023
Filing Date:	24-FEB-2023
Time Stamp:	07:36:02
Application Type:	Reexam (Third Party)

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	Foreign Reference	2-176_WO99-43185_Westermann.pdf	2288093 b151500334e410117ede6a671fcc3a3996e14388	no	31

Warnings:

Information:					
2	Non Patent Literature	2-181_Multi-Sensory_Microphones_for_Robust_Speech_Detection-Enhancement-and-Recognition.pdf	245624 aca526da7831216d30a57d7d5af6bfbea7a43bf4	no	4
Warnings:					
Information:					
3	Foreign Reference	2-184_WO2006-074082_Zurek082.pdf	1528609 6d1104e3c6b4e6d3bd477289c0b6948d60893803	no	22
Warnings:					
Information:					
4	Non Patent Literature	2-177_Adaptive_Feedback_Active_Noise_Control_Headset_Corrected.pdf	7128128 072df5d4c0c1c2d154e7f5eb00428834ffed7f40	no	8
Warnings:					
Information:					
		Total Files Size (in bytes):	11190454		

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/08a (01-22)
Approved for use through 05/31/2024. 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	90019169
Filing Date	2023-02-24
First Named Inventor	Goldstein et al.
Art Unit	3992
Examiner Name	Ge, Yuzhen
Attorney Docket Number	3-00051-666EPR

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	7844070	B2	2010-11-30	Abolfathi	
	2	5390254	A	1995-02-14	Adelman	
	3	6910013	B2	2005-06-21	Allegro et al.	
	4	7773759	B2	2010-08-10	Alves et al.	
	5	5251263	A	1993-10-05	Andrea et al.	
	6	5933506	A	1999-08-03	Aoki et al.	
	7	10297246	B2	2019-05-21	Asada et al.	
	8	7933423	B2	2011-04-26	Baekgaard Jensen et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	9	7158933	B2	2007-01-02	Balan et al.	
	10	4947440	A	1990-08-07	Bateman et al.	
	11	8027481	B2	2011-09-27	Beard	
	12	8189803	B2	2012-05-29	Bergeron et al.	
	13	6941161	B1	2005-09-06	Bobisuthi et al.	
	14	8270629	B2	2012-09-18	Bothra et al.	
	15	7246058	B2	2007-07-17	Burnett	
	16	9135797	B2	2015-09-15	Couper et al.	
	17	8477955	B2	2013-07-02	Engle et al.	
	18	4555677	A	1985-11-26	Beesley	
	19	7936885	B2	2011-05-03	Frank et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	20	4596902	A	1986-06-24	Gilman	
	21	5539831	A	1996-07-23	Harley	
	22	8270634	B2	2012-09-18	Harney et al.	
	23	6415034	B1	2002-07-02	Hietanen	
	24	6424721	B1	2002-07-23	Hohn	
	25	5550923	A	1996-08-27	Hotvet	
	26	5276740	A	1994-01-04	Inanaga et al.	
	27	5887070	A	1999-03-23	Iseberg et al.	
	28	6738482	B1	2004-05-18	Jaber	
	29	8150084	B2	2012-04-03	Jessen et al.	
	30	7103188	B1	2006-09-05	Jones	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	31	6526148	B1	2003-02-25	Jourjine et al.	
	32	7778434	B2	2010-08-17	Juneau et al.	
	33	5524056	A	1996-06-04	Killion et al.	
	34	5577511	A	1996-11-26	Killion	
	35	5692059	A	1997-11-25	Kruger	
	36	6728385	B2	2004-04-27	Kvaløy et al.	
	37	7082393	B2	2006-07-25	Lahr	
	38	8652040	B2	2014-02-18	LeBoeuf et al.	
	39	5479522	A	1995-12-26	Lindemann et al.	
	40	5526819	A	1996-06-18	Lonsbury-Martin et al.	
	41	7995773	B2	2011-08-09	Mao	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	42	6639987	B2	2003-10-28	McIntosh	
	43	6269161	B1	2001-07-31	McLaughlin et al.	
	44	7903825	B1	2011-03-08	Melanson	
	45	6175633	B1	2001-01-16	Morrill et al.	
	46	7107109	B1	2006-09-12	Nathan et al.	
	47	7983433	B2	2011-07-19	Nemirovski	
	48	6647368	B2	2003-11-11	Nemirovski	
	49	5204906	A	1993-04-20	Nohara et al.	
	50	6151571	A	2000-11-21	Pertrushin	
	51	7634094	B2	2009-12-15	Reber	
	52	7277722	B2	2007-10-02	Rosenzweig	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	53	7050966	B2	2006-05-23	Schneider et al.	
	54	6160758	A	2000-12-12	Spiesberger	
	55	6101256	A	2000-08-08	Steelman	
	56	6754359	B1	2004-06-22	Svean et al.	
	57	6567524	B1	2003-05-20	Svean et al.	
	58	6445799	B1	2002-09-03	Taenzer et al.	
	59	6912289	B2	2005-06-28	Vonlanthen et al.	
	60	7430300	B2	2008-09-30	Vosburgh et al.	
	61	7680465	B2	2010-03-16	Zad-Issa	
	62	7174022	B1	2007-02-06	Zhang et al.	
	63	8068627	B2	2011-11-29	Zhang et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	64	7986802	B2	2011-07-26	Ziller	

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	20050090295	A1	2005-04-28	Ali et al.	
	2	20080037801	A1	2008-02-14	Alves et al.	
	3	20040202340	A1	2004-10-14	Armstrong et al.	
	4	20050281422	A1	2005-12-22	Armstrong et al.	
	5	20080107282	A1	2008-05-08	Asada	
	6	20030055627	A1	2003-03-20	Balan et al.	
	7	20060188105	A1	2006-08-24	Baskerville et al.	
	8	20060126865	A1	2006-06-15	Blamey et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	9	20070223717	A1	2007-09-27	Boersma	
	10	20070253569	A1	2007-11-01	Bose	
	11	20070092087	A1	2007-04-26	Bothra et al.	
	12	20030228023	A1	2003-12-11	Burnett et al.	
	13	20040133421	A1	2004-07-08	Burnett et al.	
	14	20050071158	A1	2005-03-31	Byford	
	15	20060067551	A1	2006-03-30	Cartwright et al.	
	16	20060147063	A1	2006-07-06	Chen	
	17	20080130908	A1	2008-06-05	Cohen et al.	
	18	20040202333	A1	2004-10-14	Csermak et al.	
	19	20070014423	A1	2007-01-18	Darbut et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	20	20020057817	A1	2002-05-16	Darbut	
	21	20070160243	A1	2007-07-12	Dijkstra et al.	
	22	20060083387	A1	2006-04-20	Emoto	
	23	20080181419	A1	2008-07-31	Goldstein et al.	
	24	20030061032	A1	2003-03-27	Gonopolskiy	
	25	20060262935	A1	2006-11-23	Goose et al.	
	26	20080175411	A1	2008-07-24	Greve	
	27	20040008850	A1	2004-01-15	Gustavsson	
	28	20060262938	A1	2006-11-23	Gauger et al.	
	29	20050033571	A1	2005-02-10	Huang et al.	
	30	20020111798	A1	2002-08-15	Huang	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	31	20040160573	A1	2004-08-19	Jannard et al.	
	32	20020076059	A1	2002-06-20	Joynes	
	33	20050182620	A1	2005-08-18	Kabi et al.	
	34	20070185601	A1	2007-08-09	Lee et al.	
	35	20030035551	A1	2003-02-20	Light et al.	
	36	20050114124	A1	2005-05-26	Liu et al.	
	37	20050047611	A1	2005-03-03	Mao	
	38	20070100637	A1	2007-05-03	McCune	
	39	20080063228	A1	2008-03-13	Mejia et al.	
	40	20020141602	A1	2002-10-03	Nemirovski	
	41	20060053007	A1	2006-03-09	Niemisto	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	42	20020191799	A1	2002-12-19	Nordqvist et al.	
	43	20100119077	A1	2010-05-13	Platz et al.	
	44	20060262944	A1	2006-11-23	Rasmussen et al.	
	45	20010046304	A1	2001-11-29	Rast	
	46	20030026438	A1	2003-02-06	Ray et al.	
	47	20070189544	A1	2007-08-16	Rosenberg	
	48	20030198357	A1	2003-10-23	Schneider et al.	
	49	20060064037	A1	2006-03-23	Shalon et al.	
	50	20050058300	A1	2005-03-17	Suzuki et al.	
	51	20040252852	A1	2004-12-16	Taenzer	
	52	20080165988	A1	2008-07-10	Terlzzi et al.	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	53	20070206825	A1	2007-09-06	Thomasson	
	54	20050058313	A1	2005-03-17	Victorian et al.	
	55	20080201138	A1	2008-08-21	Visser et al.	
	56	20070021958	A1	2007-01-25	Visser et al.	
	57	20050123146	A1	2005-01-09	Voix et al.	
	58	20090087003	A1	2009-04-02	Zurek 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	2004289762	JP	A	2004-10-14	Amada		
	2	2006114101	WO	A1	2006-11-02	Andersen et al.		
	3	2005227511	JP	A	2005-08-25	Arimune		

EFS Web 2.1.18

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	4	1674061	EP	A1	2006-06-28	Berg		
	5	9326084	WO	A1	1993-12-23	Claybaugh	Corresponding to EP0643881	
	6	1640972	EP	A1	2006-03-29	Dijkstra et al.		
	7	1489596	EP	B1	2004-12-22	Gustavsson		
	8	10294989	JP	A	1998-11-04	Hashimoto et al.		
	9	4312155	DE	A1	1994-10-20	Hiller		
	10	9326085	WO	A1	1993-12-23	Hohman et al.		
	11	0935236	EP	A1	1999-08-11	Housni		
	12	297362	JP	A	1999-09-03	Katsumata		
	13	20030058432	KR	A	2003-07-07	Kim		
	14	20020086433	KR	A	2002-11-18	Kim		

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	15	20030069471	KR	A	2003-08-27	Kim		
	16	3085237	JP	A	2000-07-07	Kondo et al.		
	17	12878298	JP	A	2002-09-27	Kondo et al.		
	18	11331990	JP	A	1999-11-30	Kondo et al.		
	19	3556987	JP	B2	2004-05-21	Masui		
	20	2006037156	WO	A1	2006-04-13	Mejia et al.		
	21	2005168888	JP	A	2005-06-30	Murata		
	22	0157852	WO	A1	2001-08-09	Nakadai et al.		
	23	2006200446	AU	A1	2006-02-23	Nemirovski		
	24	1681903	EP	A2	2006-07-19	Platz		
	25	1320281	EP	A2	2003-06-18	Roeck		

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	26	2006054205	WO	A1	2006-05-26	Schobben et al.		
	27	1638079	EP	A2	2006-03-22	Sheikhzadeh et al.		
	28	8600133	WO	A1	1986-01-03	Spalding		
	29	1594344	EP	A2	2005-11-09	Stirnemann		
	30	0217836	WO	A1	2002-03-07	Svean et al.		
	31	1519625	EP	A2	2005-03-03	Victorian et al.		
	32	9943185	WO	A1	1999-08-26	Westermann et al.		
	33	2006074082	WO	A1	2006-07-13	Zurek et al.		

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	Aki Härmä, et al., Techniques and Applications of Wearable Augmented Reality Audio, Conference: the 114th Audio Engineering Society (AES) Convention, Amsterdam, The Netherlands, March 22-25, 2003	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	2	Alex Olwal and Steven Feiner, Interaction Techniques Using Prosodic Features of Speech and Audio Localization, Proceedings of IUI 2005 (International Conference on Intelligent User Interfaces), San Diego, CA, Jan 9-12, 2005, p. 284-286.	
	3	ATAL and LABINER, A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 3, June 1976	
	4	Boaz Rafaely; Miriam Furst, Audiometric Ear Canal Probe with Active Ambient Noise Control, IEEE Transactions on Speech and Audio Processing, Vol. 4, Issue 3, pp. 224-230, May 1996	
	5	Guentchev and Weng, Learning-Based Three Dimensional Sound Localization Using a Compact Non-Coplanar Array of Microphones, AAAI Technical Report SS-98-02, pp. 68-77, 1997	
	6	Jeffrey C. Bos; David W. Tack, Speech Input Hardware Investigation for Future Dismounted Soldier Computer Systems, DRDC Toronto CR2005-064, May 1, 2005	
	7	John Robert Franks, The Effect of Hearing Aid Microphone Location on the Intelligibility of Hearing Aid-Transduced Speech, December 1975	
	8	K.C. Zangi, A New Two-Sensor Active Noise Cancellation Algorithm, 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing, 27-30 April 1993	
	9	Kambiz C. Zangi, Optimal Feedback Control Formulation of the Active Noise Cancellation Problem: Pointwise and Distributed, RLE Technical Report No. 583, May 1994	
	10	Kyle Gann, Otto E. Laske, Colby Leider, Products of Interest, Computer Music Journal, Vol. 30, Number 3, Fall 2006, pp. 93-107 (Article)	
	11	MAKHOUL, Spectral Analysis of Speech by Linear Prediction, IEEE Transactions on Audio and Electroacoustics, Vol. AU-21, No. 3, pp. 140-148, June 1973	
	12	Mark C. Flohr, Active Noise Cancellation for Headphones Used in High Noise Environments, Using Conventional Analog Circuitry, May 1, 1987	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	13	Michael Valente, et al., Performance of Dual-Microphone In-the-Ear Hearing Aids, J Am AcadAudiol 11: 181-189, 2000	
	14	MOBLEY, HALL, YEAGER, E-3 In-flight Acoustic Exposure Studies and Mitigation Via Active Noise Reduction Headset, Dec. 2002	
	15	Nils Westerlund; Mattias Dahl; Ingvar Claesson, In-Ear Microphone Techniques for Severe Noise Situations, Research Report, Nov. 2005	
	16	Palomäki, Brown, Wang, A Binaural Processor for Missing Data Speech Recognition in the Presence of Noise and Small-Room Reverberation, Speech Communication 43 (2004) pp. 361–378, March 31, 2004	
	17	Rafael E. Delgado, et al., Adaptive Noise Cancellation in a Multimicrophone System for Distortion Product Otoacoustic Emission Acquisition, IEEE Transactions on Biomedical Engineering, Vol. 47, No. 9, pp. 1154-1164, Sep. 2000	
	18	RAFAELY and JONES, Combined feedback-feedforward active noise-reducing headset, J. Acoust. Soc. Am. 112(3), Pt. 1, Sep. 2002	
	19	Ravi Vaidyanathan, et al., A Dual-Mode Human-Machine Interface for Robotic Control Based on Acoustic Sensitivity of the Aural Cavity, The First IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, February 2006	
	20	Remzi Serdar Kurcan, Isolated word recognition from in-ear microphone data using Hidden Markov Models (HMM), March 2006	
	21	S.J. Elliott and P.A. Nelson, Active Noise Control: Low-frequency techniques for suppressing acoustic noise leap forward with signal processing, IEEE SIGNAL PROCESSING MAGAZINE, Oct. 1993	
	22	Samuel S. Job, REDUCING THE NEGATIVE EFFECTS OF EAR-CANAL OCCLUSION, 2002	
	23	Seltzer, Microphone Array Processing for Robust Speech Recognition, July 2003	

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

	24	Sullivan, Multi-Microphone Correlation-Based Processing for Robust Automatic Speech Recognition, Aug. 1996	
	25	Taddese, Sound Source Separation and Localization, May 2006	
	26	Woon S. Gan, Sohini Mitra, Sen M. Kuo, Adaptive Feedback Active Noise Control Headset: Implementation, Evaluation and Its Extensions, IEEE Transactions on Consumer Electronics, Vol. 51, No. 3, Aug. 2005	
	27	Zhengyou Zhang, et al., Multi-Sensory Microphones for Robust Speech Detection, Enhancement and Recognition, Microsoft Research, 2004	

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	90019169
	Filing Date	2023-02-24
	First Named Inventor	Goldstein et al.
	Art Unit	3992
	Examiner Name	Ge, Yuzhen
	Attorney Docket Number	3-00051-666EPR

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	/Jacob A. Snodgrass/	Date (YYYY-MM-DD)	2023-09-15
Name/Print	Jacob A. Snodgrass	Registration Number	66,032

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.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 November 2006 (02.11.2006)

PCT

(10) International Publication Number
WO 2006/114101 A1

(51) International Patent Classification:
G10L 21/02 (2006.01) *G10L 11/02* (2006.01)

(21) International Application Number:
PCT/DK2006/000221

(22) International Filing Date: 26 April 2006 (26.04.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA200500606 26 April 2005 (26.04.2005) DK
PA200500607 26 April 2005 (26.04.2005) DK

(71) Applicant *(for all designated States except US)*: **AALBORG UNIVERSITET** [DK/DK]; Fredrik Bajers Vej 5, DK-9220 Aalborg Ø (DK).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ANDERSEN, Søren, Vang** [DK/DK]; Boulevarden 44, 3th, DK-9000 Aalborg (DK). **SØRENSEN, Karsten, Vandborg** [DK/DK]; Rosenlundsgade 14, 2. tv., DK-9000 Aalborg (DK).

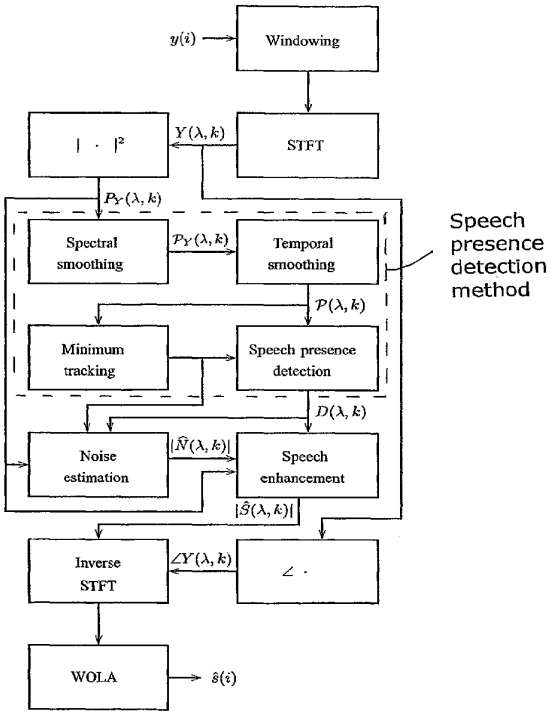
(74) Agent: **PLOUGMANN & VINGTOFT A/S**; Sundkrogsgade 9, P.O. Box 831, DK-2100 Copenhagen Ø (DK).

(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, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, 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: DETECTION OF SPEECH PRESENT IN A NOISY SIGNAL AND SPEECH ENHANCEMENT MAKING USE THEREOF



(57) Abstract: The invention provides a method for speech presence detection of speech in a noisy signal. Based on the noisy signal represented as a temporal-spectral distribution, the method includes performing a temporal-spectral smoothing on the temporal-spectral representation of the noisy signal prior to making a speech presence decision. This allows detection of connected regions in the time-frequency domain representation of the noisy signal, and thus allows improved possibilities for efficient noise suppression without musical tones as resulting artifacts. In preferred embodiments, the temporal-spectral smoothing includes performing a spectral smoothing prior to performing a temporal smoothing, and preferably the temporal smoothing includes performing a recursive smoothing with time and frequency varying smoothing parameters, these parameters being lower limited so as to ensure a minimum degree of smoothing. In preferred embodiments, the speech presence decision is based on a minimum tracking of an output of the temporal-spectral smoothing, and a noise estimation is performed based on the speech presence decision, this decision being a binary decision. The mentioned speech presence decision is advantageous for speech enhancement, especially speech enhancement including applying different attenuation functions to regions detected as noise and speech, respectively. In a preferred embodiment noise regions are attenuated by multiplying by a scalar, e.g. a scalar of 0.05-0.4. The methods are applicable within a number of devices e.g. hearing aids, headset and mobile phones.

WO 2006/114101 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.

DETECTION OF SPEECH PRESENT IN A NOISY SIGNAL AND SPEECH ENHANCEMENT MAKING USE THEREOF

Field of the invention

The invention relates to the field of signal processing, more specifically to processing aiming at detecting presence of speech in a noisy signal. Especially, the speech presence detection may serve the purpose of enhancing speech contained in the noisy signal. The invention provides a method and a device, e.g. a headset, adapted to perform the method.

Background of the invention

Signal processing with the purpose of enhancing speech contained in a noisy signal so as to improve speech intelligibility is well-known within a number of applications, e.g. hearing aids, headsets for mobile phones etc. However, most known methods with a low complexity suited for e.g. hearing aids, suffer from a poor sound quality.

In the state of the art it is known to process the noisy signal so as to represent it as a temporal-spectral distribution, e.g. a periodogram. A speech presence detection is performed in order to classify regions of the temporal-spectral distribution as either speech or speech plus noise. A spectral subtraction is then performed in order to suppress noise, e.g. an average noise spectrum is subtracted from all regions. However, by performing a spectral subtraction the noise in the regions where no speech is present is significantly changed, thus resulting in attenuation of this noise, but since it is not taken into account that the actual noise spectrum changes over the region considered, the remaining noise suffers from artificially sounding musical tones and thus an overall poor sound quality is the result.

Summary of the invention

Thus, it may be seen as an object of the present invention to provide a speech enhancement method capable of providing a low computational complexity while still providing a high sound quality without suffering from severe artefacts as musical tones etc.

In a first aspect, the invention provides a method for speech presence detection of speech in a noisy signal, the noisy signal being represented as a temporal-spectral distribution, the method including performing a temporal-spectral smoothing on the temporal-spectral representation of the noisy signal prior to making a speech presence decision.

By "represented as a temporal-spectral distribution" is understood a spectrogram type of representation of a portion of noisy signal in time-frequency domain distribution, thus a three-dimensional representation of the, i.e. time, frequency and e.g. signal intensity or signal power. E.g. the temporal dimension can be divided in discrete time frames, and the spectral dimension can be divided in discrete frequency bins. However, it is to be understood that other temporal-spectral distributions other than fixed discrete time frame and frequency bins are possible.

CONFIRMATION COPY

The method is advantageous e.g. in that a smoothing of a time-frequency domain representation of the noisy signal prior to making the speech presence decision allows a better possibility to detect connected time-frequency regions of noise and connected time-frequency regions of speech plus noise. Thus, the smoothing helps to facilitate the speech presence decision task, and it leads in addition to a better noise estimate. E.g. in a preferred embodiment, the speech presence decision is a binary decision.

The speech presence detection method is advantageous for use in connection with speech enhancement since the mentioned detection of connected regions of noise and speech plus noise, respectively, provides a basis for applying an efficient noise suppression algorithm which still, due to the connected regions, results in a natural sounding enhanced speech. E.g. it becomes possible to apply different attenuation functions to noise regions and speech plus noise regions and thus achieve a more efficient noise suppression without resulting in musical tone artifacts.

In preferred embodiments of the method, the temporal-spectral smoothing includes performing a spectral smoothing prior to performing a temporal smoothing, and especially it is preferred that the temporal smoothing includes performing a recursive smoothing with time and frequency varying smoothing parameters, and more preferably these smoothing parameters are lower limited so as to ensure a minimum degree of smoothing.

Preferably, the speech presence decision is based on a minimum tracking of an output of the temporal-spectral smoothing. Preferably, a noise estimation is performed based on the speech presence decision.

In a second aspect, the invention provides a speech enhancement method including performing the speech presence detection method described in the first aspect, and performing enhancement of the speech in the noisy signal based on the speech presence detection.

Preferably, a first attenuation function is applied to regions of the temporal-spectral distribution where presence of speech is detected, and a second attenuation function is applied to other regions of the temporal-spectral distribution, the first and second attenuation functions being different. With different attenuation of the detected noise regions and the detected speech plus noise regions, it is possible to apply more dedicated attenuation functions that result in a more efficient noise suppression still without resulting in musical noise or tone artifacts. Especially, it is preferred that the second attenuation function applied to the other regions, i.e. the regions detected as noise, of the temporal-spectral distribution is a scalar attenuation, and preferably this scalar attenuation is implemented by multiplying by a scalar in the range 0.05 to 0.4, such as in the range 0.1 to 0.3, and preferably in the range 0.15 to 0.25, such as 0.2 or approximately 0.2.

Since the speech enhancement method of the second aspect includes performing the speech presence detection method of the first aspect, it has the same advantages as

mentioned for the first aspect, and the preferred embodiments mentioned for the first aspect therefore also apply.

The speech enhancement method is suited for application where a noisy audio signal containing speech is corrupted by noise. The noise may be caused by electrical noise interfering with an electrical audio signal, or the noise may be acoustic noise such as introduced at the recording of the speech, e.g. a person speaking in a telephone at a place with traffic noise etc. The speech enhancement method can then be used to increase speech intelligibility by enhancing the speech in relation to the noise.

In a third aspect the invention provides a device including a processor adapted to perform the method of any one of the first or second aspects. Thus, the advantages and embodiments mentioned for the first and second aspects therefore apply for the third aspect as well.

Especially, the device may be: a mobile phone, a radio communication device, an internet telephony system, sound recording equipment, sound processing equipment, sound editing equipment, broadcasting sound equipment, or a monitoring system.

Alternatively, the device may be: a hearing aid, a headset, an assistive listening device, an electronic hearing protector, or a headphone with a built-in microphone (so as to allow sound from the environments to reach the listener).

In a fourth aspect, the invention provides a computer executable program code adapted to perform the method according to any one of the first, second or third aspects. Thus, the same advantages as mentioned for these aspects therefore apply.

The program code may be present on a program carrier, e.g. a memory card, a disk etc. or in a RAM or ROM memory of a device.

Brief description of the drawings

In the following the invention is described in more details with reference to the accompanying figures, of which

Fig. 1 illustrates a block diagram of a preferred speech enhancement algorithm, and

Fig. 2 illustrates a preferred device.

While the invention is susceptible to various modifications and alternative forms, specific embodiments have been shown by way of example in the drawings and will be described in detail herein. It should be understood, however, that the invention is not intended to be limited to the particular forms disclosed. Rather, the invention is to cover all modifications, equivalents, and alternatives falling within the spirit and scope of the invention as defined by the appended claims.

Description of preferred embodiments

Fig. 1 illustrates a block diagram of main parts of a preferred speech enhancement algorithm, including preferred steps of a speech enhancement detection method indicated with a dashed box. The following detailed description of a speech enhancement algorithm with natural sounding residual noise based on connected time-frequency speech presence regions is based on the principles already described. The connected time-frequency regions are used by a noise estimation method and both the speech presence decisions and the noise estimate are used in the speech enhancement method. Different attenuation rules are applied to regions with and without speech presence to achieve enhanced speech with natural sounding attenuated background noise. The preferred speech enhancement method has a low computational complexity, and thus it is suited e.g. for implementation into low processing power devices such as hearing aids. Listening tests have shown that the method provides higher scores than minimum means square error log-spectral amplitude (MMSE-LSA) and decision-directed MMSE-LSA methods.

The performance of many speech enhancement methods relies mainly on the quality of a noise Power Spectral Density (PSD) estimate. When the noise estimate differs from the true noise it will lead to artifacts in the enhanced speech. The approach taken in this paper is based on connected region speech presence detection. Our aim is to exploit spectral and temporal masking mechanisms in the human auditory system [1] to reduce the perception of these artifacts in speech presence regions and eliminate the artifacts in speech absence regions. We achieve this by leaving down-scaled natural sounding background noise in the enhanced speech in connected time-frequency regions with no speech presence. The down-scaled natural sounding background noise will spectrally and temporally mask artifacts in the speech estimate while preserving the naturalness of the background noise.

In the definition of speech presence regions we are inspired by the work by Yang [2]. Yang demonstrates high perceptual quality of a speech enhancement method where constant gain is applied in frames with no detected speech presence. Yang lets a single decision cover a full frame. Thus, musical noise is present in the full spectrum of the en-

hanced speech in frames with speech activity. We therefore extend the notion of speech presence to individual time-frequency locations. This, in our experience, significantly improves the naturalness of the residual noise. The speech enhancement method, proposed in this paper, thereby eliminates audible musical noise in the enhanced speech. However, fluctuating speech presence decisions will reduce the naturalness of the enhanced speech and the background noise. Thus, reasonably connected regions of the same speech presence decision must be established.

To achieve this, we use spectral-temporal periodogram smoothing. To this end, we make use of the spectral-temporal smoothing method by Martin and Lotter [3], which extends the original ground-breaking work by Martin [4, 5]. Martin and Lotter derive optimum smoothing coefficients for (generalized) χ^2-distributed spectrally smoothed spectrograms, which is particularly well suited for noise types with a smooth power spectrum. The underlying assumption in this approach is that the real and imaginary parts of the associated STFT coefficients for the averaged periodograms have the same means and variances. For the application of spectral-temporal smoothing to obtain connected regions of speech presence decisions, we augment Martin and Lotters smoothing method with the spectral smoothing method used by Cohen and Berdugo [6].

For minimum statistics noise estimation, Martin [5] has suggested a theoretically founded bias compensation factor, which is a function of the minimum search window length, the smoothed noisy speech and the noise PSD estimate variances. This enables a low-biased noise estimate that does not rely on a speech presence detector. However, as our proposed speech enhancement method has connected speech presence regions as an integrated component, this enables us to make use of a new, simple, yet efficient, bias compensation. To verify the performance of the new bias compensation we objectively evaluate the noise estimation method that uses this bias compensation of minimum tracks from our spectrally-temporally smoothed periodograms, prior to integrating this noise estimate in the final speech enhancement method.

In result, our proposed speech enhancement algorithm has a low computational complexity, which makes it particularly relevant for application in digital signal processors

Exhibit 1009
Page 1406 of 2198

with limited computational power, such as those found in digital hearing aids. In particular, the obtained algorithm provides a significantly higher perceived quality than our implementation of the Decision-Directed Minimum Mean-Square Error Log-Spectral Amplitude (MMSE-LSA-DD) estimator [7] when evaluated in listening tests. Furthermore, the noise PSD estimate that we use to obtain a noise magnitude spectrum estimate for the attenuation rule in connected regions of speech presence is shown to be superior to estimates from Minimum Statistics (MS) noise estimation [5] and our implementation of χ^2-based noise estimation [3] for spectrally smooth noise types.

The remainder of this paper is organized as follows. In Section II, we describe the signal model and give an overview of the proposed algorithm. In Section III, we list the necessary equations to perform the spectral-temporal periodogram smoothing. Section IV contains a description of our detector for connected speech presence regions, and in Section V, we describe how the spectrally-temporally smoothed periodograms and the speech presence regions can be used to obtain both a noise PSD estimate and a noise periodogram estimate, which both relies on the new bias compensation. In the latter noise estimation method, we estimate the squared magnitudes of the noise Short-Time Fourier Transform (STFT) coefficients. In Section VI the connected region speech presence detector is introduced in a speech enhancement method with the purposes of reducing noise and augmenting listening comfort. Section VII contains the experimental setup and all necessary initializations. Finally, Section VIII describes the experimental results and Section IX concludes the paper with a discussion of the proposed methods and obtained results.

Structure of the Algorithm

After an introduction to the signal model, we give a structural description of the algorithm to provide an algorithmic overview before the individual methods, which constitute the algorithm, are described in detail.

Signal Model

We assume that noisy speech $y(i)$ at sampling time index i consists of speech $s(i)$ and

Exhibit 1009
Page 1407 of 2198

additive noise $n(i)$. For joint time-frequency analysis of $y(i)$ we apply the K-point STFT, i.e.

$$Y(\lambda,k)=\sum_{\mu=0}^{L-1}y(\lambda R+\mu)h(\mu)\exp\left(-j2\pi k\mu/K\right), \tag{1}$$

where $\lambda\in\mathcal{Z}$ is the (sub-sampled) time index, $k\in\{0,1,\ldots,K-1\}$ is the frequency index, and L is the window length. In this paper, we have that L equals K. The quantity R is the number of samples that successive frames are shifted and $h(\mu)$ is a unit energy window function, i.e. $\sum_{\mu=0}^{L-1}h^2(\mu)=1$. From the linearity of (1) we have that

$$Y(\lambda,k)=S(\lambda,k)+N(\lambda,k), \tag{2}$$

where $S(\lambda,k)$ and $N(\lambda,k)$ are the STFT coefficients of speech $s(i)$ and additive noise $n(i)$, respectively. We further assume that $s(i)$ and $n(i)$ are zero mean and statistically independent, which leads to a power relation, where the noise is additive [8], i.e.

$$E\{|Y(\lambda,k)|^2\}=E\{|S(\lambda,k)|^2\}+E\{|N(\lambda,k)|^2\}. \tag{3}$$

Structural Algorithm Description

The structure of the proposed algorithm, and names of variables with a central role, are shown in Figure 1. After applying an analysis window to the noisy speech, we take the STFT, from which we calculate periodograms $P_Y(\lambda,k)\triangleq|Y(\lambda,k)|^2$. These periodograms are spectrally smoothed, yielding $\mathcal{P}_Y(\lambda,k)$, and then temporally smoothed to produce $\mathcal{P}(\lambda,k)$. These smoothed periodograms are temporally minimum tracked, and by comparing ratios and differences of the minimum tracked values to $\mathcal{P}(\lambda,k)$, they are used for speech presence detection. As a distinct feature of the proposed method we use speech presence detection to achieve low-biased noise PSD estimates $\widetilde{P}_{\widehat{N}}(\lambda,k)$, but also for noise periodogram estimates $P_{\widehat{N}}(\lambda,k)$, which equal $P_Y(\lambda,k)$ when $D(\lambda,k)=0$, i.e. no detected speech presence. When $D(\lambda,k)=1$, i.e. detected speech presence, the noise

Exhibit 1009
Page 1408 of 2198

periodogram estimate equals the noise PSD estimate, i.e. a recursively smoothed bias compensation factor applied on the minimum tracked values. The bias compensation factor is recursively smoothed power ratios between the noise periodogram estimates and the minimum tracks. This factor is only updated while no speech is present in the frames and kept fixed while speech is present. A noise magnitude spectrum estimate $|\widehat{N}(\lambda,k)|$ obtained from the noise PSD estimate and the speech presence decisions are used in a speech enhancement method that applies different attenuation rules for speech presence and no speech presence. For speech synthesis we take the inverse STFT of the estimated speech magnitude spectrum with the phase from the STFT of the noisy speech. The synthesized frame is used in a Weighted OverLap Add (WOLA) method, where we apply a synthesis window before overlap and add.

Spectral-Temporal Periodogram Smoothing

In this section we briefly describe the spectral-temporal periodogram smoothing method.

Spectral Smoothing

First the noisy speech periodograms $P_Y(\lambda,k)$ are spectrally smoothed by letting a spectrally smoothed periodogram bin $\mathcal{P}_Y(\lambda,k)$ consist of a weighted sum of $2D+1$ periodogram bins, spectrally centered at k [6], i.e.

$$\mathcal{P}_Y(\lambda,k) = \sum_{\nu=-D}^{D} b(\nu) P_Y(\lambda,((k-\nu))_K), \tag{4}$$

where $((m))_K$ denotes m modulus K, and K is the length of the full (mirrored) spectrum. The window function $b(\nu)$ used for spectral weighting is chosen such that it sums to 1, i.e. $\sum_{\nu=-D}^{D} b(\nu) = 1$, and therefore preserves the total power of the spectrum.

Temporal Smoothing

The spectrally smoothed periodograms $\mathcal{P}_Y(\lambda,k)$, see Figure 1, are now temporally smoothed recursively with time and frequency varying smoothing parameters $\alpha(\lambda,k)$, to produce a spectrally-temporally smoothed noisy speech periodogram $\mathcal{P}(\lambda,k)$, i.e.

Exhibit 1009
Page 1409 of 2198

$$\mathcal{P}(\lambda,k) = \alpha(\lambda,k)\mathcal{P}(\lambda-1,k) + (1-\alpha(\lambda,k))\mathcal{P}_Y(\lambda,k). \quad (5)$$

We use the optimum smoothing parameters proposed by Martin and Lotter [3]. Their method consists of optimum smoothing parameters for χ^2-distributed data with some modifications that makes it suited for practical implementation. The optimum smoothing parameters are given by

$$\alpha^*(\lambda,k) = \frac{2}{2+\widetilde{K}(\mathcal{P}(\lambda-1,k)/E\{|N(\lambda,k)|^2\}-1)^2}, \quad (6)$$

with

$$\widetilde{K} = (4D+2)\frac{\left[\sum_{\mu=0}^{L-1} b^2(\mu)\right]^2}{L\sum_{\mu=0}^{L-1} b^4(\mu)} \quad (7)$$

"equivalent" degrees of freedom of a χ^2-distribution [3]. For practical implementation the noise PSD, which is used in the calculation of the optimum smoothing parameters, is estimated as the previous noise PSD estimate, i.e.

$$E\{|\widehat{N(\lambda,k)}|^2\} = \widetilde{P}_{\widehat{N}}(\lambda-1,k). \quad (8)$$

Complete Periodogram Smoothing Algorithm

Pseudocode for the complete spectral-temporal periodogram smoothing method is provided in Algorithm 1.

A smoothing parameter correction factor $\alpha_c(\lambda,k)$, proposed by Martin [5], is multiplied on $\tilde{\alpha}(\lambda,k)$. Additionally, in this paper, we lower limit the resulting smoothing parameters to ensure a minimum degree of smoothing, i.e.

$$\alpha(\lambda,k) = \max(\alpha_c(\lambda,k)\tilde{\alpha}(\lambda,k), 0.4). \quad (9)$$

In the next section, we use temporal minimum tracking on the spectrally-temporally smoothed noisy speech periodograms in a method for detection of connected speech presence regions, which later will be used for noise estimation and speech enhancement.

Algorithm 1 Periodogram Smoothing

{Initialize as listed in Table 1 and Table 2}
for $\lambda = 0$ to $M-1$ **do**
  for $k = 0$ to $K-1$ **do**
    $P_Y(\lambda,k) \leftarrow abs\left(\sum_{\mu=0}^{L-1} y(\lambda R+\mu)h(\mu)\exp\left(-j2\pi k\mu/K\right)\right)^2$
    $\mathcal{P}_Y(\lambda,k) \leftarrow \sum_{\nu=-D}^{D} b(\nu)P_Y(\lambda, mod(k-\nu,K))$
    $\alpha^*(\lambda,k) \leftarrow 2/\left(2+\widetilde{K}(\mathcal{P}(\lambda-1,k)/\widetilde{P}_{\widehat{N}}(\lambda-1,k)-1)^2\right)$
  end for
  $R \leftarrow \left(\sum_{k=0}^{K-1}\mathcal{P}(\lambda-1,k)\right)/\left(\sum_{k=0}^{K-1}P_Y(\lambda,k)\right)$
  $\tilde{\alpha}_c \leftarrow 1/(1+(R-1)^2)$
  $\alpha_c(\lambda) \leftarrow 0.7\alpha_c(\lambda-1)+0.3\max(\tilde{\alpha}_c, 0.7)$
  for $k = 0$ to $K-1$ **do**
    $\alpha(\lambda,k) \leftarrow \max(\alpha_c(\lambda)\tilde{\alpha}(\lambda,k), 0.4)$
    $\mathcal{P}(\lambda,k) \leftarrow \alpha(\lambda,k)\mathcal{P}(\lambda-1,k)+(1-\alpha(\lambda,k))\mathcal{P}_Y(k)$
    {Obtain a noise PSD estimate $\widetilde{P}_{\widehat{N}}(\lambda,k)$, e.g as proposed in Section 0.6.}
  end for
end for

Table 1: General Setup.

Variable	Value	Description
F_s	$8\ kHz$	Sample frequency
K	256	FFT size
L	256	Frame size
R	128	Frame skip
$h(\mu)$	$G_h^{-1}\sqrt{\text{HANNING}(K)}$[i]	Analysis window
$h_s(\mu)$	$G_h\sqrt{\text{HANNING}(K)}$[ii]	Synthesis window

[i] G_h is the square root of the energy of $\sqrt{\text{HANNING}(K)}$, which scales the analysis window to unit energy. This is to avoid scaling factors throughout the paper.

[ii] G_h scales the synthesis window $h_s(\mu)$, such that the analysis window $h(\mu)$, multiplied with $h_s(\mu)$, yields a HANNING(K) window.

Table 2: Smoothing Setup and Initializations for Algorithm 1.

Variable	Value	Description
D	7	Spectral window length: $2D+1$
$b(\nu)$	$G_b\text{TRIANG}(2D+1)$[i]	Spectral smoothing window
M	154-220[ii]	Number of frames
$\widetilde{K}$	20.08 [iii]	"Equivalent" degrees of freedom of χ^2-distribution
$P_{\widehat{N}}(-1,k)$	$P_Y(0,k)$	Initial noise periodogram estimate
$\alpha_c(-1)$	1	Initial correction variable

[i] $G_b = (sum(\text{TRIANG}(2D+1)))^{-1}$ scales the window to unit sum.

[ii] Calculated at run-time as $M = round(length(y(i))/R - 1/2) - 1$.

[iii] Calculated at run-time as $\widetilde{K} = (4D+2)\left(\left[\sum_{\mu=0}^{L-1} b^2(\mu)\right]^2\right) / \left(L\sum_{\mu=0}^{L-1} b^4(\mu)\right)$ [3, 16].

Connected Speech Presence Regions

We now base a speech presence detection method on comparisons, at each frequency, between the smoothed noisy speech periodograms and temporal minimum tracks of the smoothed noisy speech periodograms.

Temporal Minimum Tracking

From the spectrally-temporally smoothed noisy speech periodograms $\mathcal{P}(\lambda, k)$, we track temporal minimum values $\mathcal{P}_{min}(\lambda, k)$, within a minimum search window of length D_{min}, i.e.

$$\mathcal{P}_{min}(\lambda, k) = \min(\mathcal{P}(\psi, k)|\lambda - D_{min} < \psi \leq \lambda), \tag{10}$$

with $\psi \in \mathcal{Z}$. D_{min} is chosen as a tradeoff between the ability to bridge over periods of speech presence [5], which is crucial for the minimum track to be robust to speech presence, and the ability to follow non-stationary noise. Typically, a window length corresponding to $0.5 - 1.5$ seconds, yields an acceptable tradeoff between these two properties [5, 6]. We now have that $\mathcal{P}_{min}(\lambda, k)$ is approximately unaffected by periods of speech presence, but on the average biased towards lower values, when no spectral smoothing is applied [5]. Memory requirements of the tracking method can be reduced at the cost of lost temporal resolution, see e.g. [5]. In the following the temporal minimum tracks $\mathcal{P}_{min}(\lambda, k)$ are used in a speech presence decision rule.

Binary Speech Presence Decision Rule

We have shown in previous work [9], that temporally smoothed periodograms and their temporal minimum tracks can be used for speech presence detection. Also shown in [9] is that including terms to compensate for bias on the minimum tracks improves the speech presence detection performance (measured as the decrease in a cost function) by less than one percent. In this paper, we therefore do not consider a bias compensation factor in the speech presence decision rule. Rather, as we show later in this paper, the speech presence decisions can be used in the estimation of a simple and very well performing bias compensation factor for noise estimation. Similar to our previous

approach for temporally smoothed periodograms [9], we now exploit the properties of spectrally-temporally smoothed periodograms $\mathcal{P}(\lambda, k)$ in a binary decision rule for the detection of speech presence. The presence of speech will cause an increase of power in $\mathcal{P}(\lambda, k)$ at a particular time-frequency location, due to (3). Thus, the ratio between $\mathcal{P}(\lambda, k)$ and a noise PSD estimate, given by a minimum track $\mathcal{P}_{min}(\lambda, k)$ with a bias reduction, yields a robust (due to the smoothing) estimate of the signal-plus-noise to noise ratio at the particular time-frequency location. Our connected region speech presence detection method is based on the smooth nature of $\mathcal{P}(\lambda, k)$ and $\mathcal{P}_{min}(\lambda, k)$. The smoothness will ensure that spurious fluctuations in the noisy speech power will not cause spurious fluctuations in our speech presence decisions. Thus, we will be able to obtain connected regions of speech presence and of no speech presence. This property is fundamental for the proposed noise estimation and speech enhancement methods. As a rule to decide between the two speech presence hypotheses, namely

$$H_0(\lambda, k): \ \textit{'No speech present'} \tag{11}$$

$$H_1(\lambda, k): \ \textit{'Speech present'}, \tag{12}$$

which can be written in terms of the STFT coefficients, i.e.

$$H_0(\lambda, k): \ Y(\lambda, k) = N(\lambda, k) \tag{13}$$

$$H_1(\lambda, k): \ Y(\lambda, k) = N(\lambda, k) + S(\lambda, k), \tag{14}$$

we use a combination of two binary initial decision rules. First, let $D(\lambda, k) = i$ be the decision to believe in hypothesis $H_i(\lambda, k)$ for $i \in \{0, 1\}$. We define two initial decision rules, which will give two initial decisions $D'(\lambda, k)$ and $D''(\lambda, k)$. The initial decision rules are given by a rule, where the smoothed noisy speech periodograms $\mathcal{P}(\lambda, k)$ are compared with the temporal minimum tracks $\mathcal{P}_{min}(\lambda, k)$, weighted with a constant γ', i.e.

$$D'(\lambda, k): \quad \mathcal{P}(\lambda, k) \underset{D'(\lambda,k)=0}{\overset{D'(\lambda,k)=1}{\gtrless}} \gamma' \mathcal{P}_{min}(\lambda, k), \tag{15}$$

Exhibit 1009
Page 1414 of 2198

and one where, at time λ, the difference is compared to the average of the minimum tracks scaled by γ'', i.e.

$$D''(\lambda,k): \quad \mathcal{P}(\lambda,k) \underset{D''(\lambda,k)=0}{\overset{D''(\lambda,k)=1}{\gtrless}} \mathcal{P}_{min}(\lambda,k) \tag{16}$$

$$+\gamma'' \frac{1}{K} \sum_{k=0}^{K-1} \mathcal{P}_{min}(\lambda,k). \tag{17}$$

For the initial decision rules we have adopted the notation used by Shanmugan and Breipohl [8]. Because the minimum tracks are representative of the noise PSD's [5], the first initial decision rule classifies time-frequency bins based on the estimated signal-plus-noise to noise power ratio. Note that this can be seen as a special case of the indicator function proposed by Cohen [10] (with $\zeta_0 = \gamma'/B_{min}$ and $\gamma_0 = \infty$). The second initial decision rule $D''(\lambda,k)$ classifies bins from the estimated power difference between the noisy speech and the noise using a threshold that adapts to the minimum track power level in each frame. Multiplication of the two binary initial decisions corresponds to the logical AND-operation, when we define true as deciding on $H_1(\lambda,k)$ and false as deciding on $H_0(\lambda,k)$. We therefore propose a decision that combines the two initial decisions from the initial decision rules above, i.e.

$$D(\lambda,k) = D'(\lambda,k) \cdot D''(\lambda,k). \tag{18}$$

In effect, the combined decision allows detection of speech in low signal-to-noise ratios (SNR's) without letting low power regions with high SNR's contaminate the decisions. Thereby we obtain connected time-frequency regions of speech presence. The constant γ' is not sensitive to the type and intensity of environmental noise [11] and it can be adjusted empirically. This is also the case for γ''. For applications where a reasonable objective performance measure can be defined, the constants γ' and γ'' can be obtained by interpreting the decision rule as artificial neural network and then conduct a supervised training of this network [9].

Speech at frequencies below 100 Hz is considered perceptually unimportant, and bins below this frequency are therefore always classified with no speech presence. Real-life

Exhibit 1009
Page 1415 of 2198

noise sources often have a large part of their power at the low frequencies, so this rule ensures, that this power does not cause the speech presence detection method to falsely classify these low-frequency bins as if speech is present. If less than 5% of the K periodogram bins are classified with speech presence, we expect that these decisions have been falsely caused by the noise characteristics, and all decisions in the current frame are reclassified to no speech presence. When the speech presence decisions are used in a speech enhancement method, as we propose in Section 0.7, this reclassification will ensure the naturalness of the background noise in periods of speaker silence.

Noise Estimation

The spectral-temporal smoothing method [3], which we use in this paper, reduces the bias between the noise PSD and the minimum track $\mathcal{P}_{min}(\lambda, k)$ if the noise is assumed to be ergodic in its PSD. That is, it reduces the bias compared to minimum tracked values from periodograms, smoothed temporally using Martins first method [5]. Martin gives a parametric description of a bias compensation factor, which depends on the minimum search window length, the smoothed noisy speech, and the noise PSD estimate variances. The spectral smoothing lowers the smoothed noisy speech periodogram variance, and as a consequence, a longer minimum search window can be applied when the noise spectrum is not changing rapidly. This give the ability to bridge over longer speech periods.

We propose to use the speech presence detection method from Section 0.5 to obtain two different noise estimates, i.e. a noise PSD estimate and a noise periodogram estimate. The PSD estimate will be used in the speech enhancement methods and the noise periodogram estimate will illustrate some of the properties of the residual noise from the speech enhancement method we propose in Section 0.7.

Noise Periodogram Estimation

The noise periodogram estimate is equal to a time-varying power scaling of the minimum tracks $\mathcal{P}_{min}(\lambda, k)$, for $D(\lambda, k) = 1$. For $D(\lambda, k) = 0$ it is equal to the noisy speech

Exhibit 1009
Page 1416 of 2198

periodogram $P_Y(\lambda, k)$, i.e.

$$P_{\widehat{N}}(\lambda, k) = \begin{cases} R_{min}(\lambda)\mathcal{P}_{min}(\lambda, k) & \text{if } D(\lambda, k) = 1 \\ P_Y(\lambda, k) & \text{if } D(\lambda, k) = 0. \end{cases} \tag{19}$$

In the above equation a bias compensation factor $R_{min}(\lambda)$ scales the minimum. The scaling factor is updated in frames where no speech presence is detected and kept fixed while speech presence is detected in the frames. We let $\widetilde{R}_{min}(\lambda)$ be given by the ratio between the sums of the previous noise periodogram estimate $P_{\widehat{N}}(\lambda - 1, k)$ and the minimum tracks $\mathcal{P}_{min}(\lambda, k)$, i.e.

$$\widetilde{R}_{min}(\lambda) = \frac{\sum_{k=0}^{K-1} P_{\widehat{N}}(\lambda - 1, k)}{\sum_{k=0}^{K-1} \mathcal{P}_{min}(\lambda, k)}, \tag{20}$$

which is recursively smoothed, when no speech is present in the frame, and fixed, when speech is present in the frame, i.e.

$$R_{min}(\lambda) = \tag{21}$$
$$\begin{cases} R_{min}(\lambda - 1) & \text{if } \sum_{k=0}^{K-1} D(\lambda, k) > 0 \\ \alpha_{min} R_{min}(\lambda - 1) + (1 - \alpha_{min})\widetilde{R}_{min}(\lambda) & \text{if } \sum_{k=0}^{K-1} D(\lambda, k) = 0, \end{cases}$$

where $0 \leq \alpha_{min} \leq 1$ is a constant recursive smoothing parameter. The magnitude spectrum, at time index λ, is obtained by taking the square root of noise periodogram estimate, i.e.

$$|\widehat{N}(\lambda, k)| = \sqrt{P_{\widehat{N}}(\lambda, k)}. \tag{22}$$

This noise periodogram estimate equals the true noise periodogram $|N(\lambda, k)|^2$ when the speech presence detection is correctly detecting no speech presence. When entering a region with speech presence, the noise periodogram estimate will take on the smooth shape of the minimum track, scaled with the bias compensation factor in (21) such that the power develops smoothly into the speech presence region.

Exhibit 1009
Page 1417 of 2198

Noise PSD Estimation

The noise PSD estimate $\widetilde{P}_{\widehat{N}}(\lambda, k)$ is obtained exactly as the noise periodogram estimate but with (19) modified such that the noise PSD estimate is obtained directly as the power-scaled minimum tracks, i.e.

$$\widetilde{P}_{\widehat{N}}(\lambda, k) = R_{min}(\lambda)\mathcal{P}_{min}(\lambda, k). \tag{23}$$

A smooth estimate of the noise magnitude spectrum can be obtained by taking the square root of the noise PSD estimates, i.e.

$$|\widehat{N}(\lambda, k)| = \sqrt{\widetilde{P}_{\widehat{N}}(\lambda, k)}. \tag{24}$$

Speech Enhancement

We now describe the speech enhancement method for which the speech presence detection method has been developed. It is well known that methods that subtract a noise PSD estimate from a noisy speech periodogram, e.g. using an attenuation rule, will introduce musical noise. This happens whenever the noisy speech periodogram exceeds the noise PSD estimate. If, on the other hand, the noise PSD estimate is too high, the attenuation will reduce more noise, but also cause the speech estimate to be distorted. To mitigate these effects, we propose to distinguish between connected regions with speech presence and no speech presence. In speech presence we will use a traditional estimation technique, by means of generalized spectral subtraction, with the noise magnitude spectrum estimate, obtained using (24) from the noise PSD estimate. In no speech presence we will use a simple noise-scaling attenuation rule to preserve the naturalness in the residual noise. Note that this approach, but with only a single speech presence decision covering all frequencies in each frame, has previously been proposed by Yang [2]. Moreover, Cohen and Berdugo [11] propose a binary detection of speech presence/absence (called the indicator function in their paper), which is similar to the one we propose in this paper. However, their decision includes noisy speech

Exhibit 1009
Page 1418 of 2198

periodogram bins without smoothing, hence some decisions will not be regionally connected. In our experience, this leads to artifacts if the decisions are used directly in a speech enhancement scheme with two different attenuation rules for speech absence and speech presence. Cohen and Berdugo smoothes their binary decisions to obtain estimated speech presence probabilities, which are used for a soft-decision between two separate attenuation functions. Our approach, as opposed to this, is to obtain adequately time-frequency smoothed spectra from which connected speech presence regions can be obtained directly in a robust manner. As a consequence, we avoid distortion in speech absence regions, and thereby obtain a natural sounding background noise.

Let the generalized spectral subtraction variant be given similar to the one proposed by Berouti et. al. [12], but with the decision, of which attenuation rule to use, given explicitly by the proposed speech presence decisions, instead of comparisons between the estimated speech power and an estimated noise floor. The immediate advantage of our approach is a higher degree of control with the properties of the enhancement algorithm. Our proposed method is given by

$$|\widehat{S}(\lambda,k)| = \begin{cases} \left(|Y(\lambda,k)|^{a_1} - \beta_1|\widehat{N}(\lambda,k)|^{a_1}\right)^{\frac{1}{a_1}} & \text{if } D(\lambda,k)=1 \\ \beta_0|Y(\lambda,k)| & \text{if } D(\lambda,k)=0, \end{cases} \tag{25}$$

where a_1 determines the power, in which the subtraction is performed, β_1 is a noise overestimation factor, that scales the estimated magnitude of the noise STFT coefficient $|\widehat{N}(\lambda,k)|$, obtained from the noise PSD estimate by (24) in Section 0.6, raised to the a_1'th power. The factor β_0 scales the noisy speech STFT coefficient magnitude, which before this scaling equals the square root of the noise periodogram estimate for bins with $D(\lambda,k)=0$. After the scaling these noisy speech STFT magnitudes lead to the noise component that will be left, after STFT synthesis, in the speech estimate as artifact masking [1] and natural sounding attenuated background noise.

For synthesis we let the STFT spectrum of the estimated speech be given by the magnitude, obtained from (25), and the noisy phase $\angle Y(\lambda,k)$, i.e.

Exhibit 1009
Page 1419 of 2198

$$\widehat{S}(\lambda, k) = |\widehat{S}(\lambda, k)| e^{j\angle Y(\lambda,k)}. \quad (26)$$

By applying the inverse STFT we synthesize a time-domain frame, which we use in a WOLA scheme, as illustrated in Figure 1, to form the synthesized signal. Depending on the analysis window, a corresponding synthesis window $h_s(\mu)$ is applied before overlap add is performed.

Experimental Setup

In the experiments we use 6 speech recordings from the TIMIT database [13]. The speech is spoken by 3 different male and 3 different female speakers - all uttering different sentences of 2-3 seconds duration. These sentences are added with zero-mean highway noise and car interior noise in 0, 5, and 10 dB overall signal-to-noise ratios to form a test set of 36 noisy speech sequences. Spectrograms of time-domain signals are shown with time-frequency axes and always with the time-domain signals. When we plot intermediate coefficients the figures are shown with axes of sub-sampled time index λ and frequency index k. For all illustrations in this paper, we use the noisy speech from one of the male speakers with additive highway noise in a 5 dB over-all SNR.

The general setup in the experiments is listed in Table 1. The analysis window $h(\mu)$ is the square-root of a Hanning window, scaled to unit energy. As the synthesis window $h_s(\mu)$, we also use the square-root of a Hanning window, but scaled, such that an unmodified frame would be windowed by a Hanning window after both the analysis and synthesis window have been applied. It will therefore be ready for overlap add with 50% overlapping frames.

For the spectral-temporal periodogram smoothing, we use the settings and algorithm initializations given in Table 2. The decision rules, that are used for speech presence detection, have the threshold values listed in Table 3. For noise estimation, we use the

Table 3: Speech Presence Detection Setup.

Variable	Value	Description
γ'	6	Constant for ratio-based decision rule
γ''	0.5	Constant for difference-based decision rule

Table 4: Noise Estimation Setup.

Variable	Value	Description
D_{min}	150[i]	Minimum tracking window length
α_{min}	0.7	Scaling factor smoothing parameter

[i]Corresponds to a time duration of $D_{min} \cdot R/F_s = 2.4\ s$.

Table 5: Speech Enhancement Setup.

Variable	Value	Description
β_0	0.1	Noise scaling factor for no speech presence
β_1	1.4	Noise over-est. factor for speech presence
a_1	0.8	Attenuation rule order for speech presence

two parameters from Table 4. The speech enhancement method uses the parameter settings that are listed in Table 5.

Experimental Results

In this section, we evaluate the performance of the proposed algorithm. We measure the performance of the algorithm by means of visual inspection of spectrograms, spectral distortion measures, and informal listening tests.

By removing the power in no speech presence regions and speech presence regions from the noisy speech periodogram, we see in Figure 4 (top) and (bottom), respectively, that most of the speech, that is detectable by visual inspection, has been detected by the proposed algorithm.

We evaluate the performance of the noise estimation methods by means of their spectral distortion, which we measure as segmental noise-to-error ratios (SegNER). We calculate the SegNER in the time-frequency domain, as the ratio (in dB) between the noise energy and the noise estimation error energy. These values are upper and lower limited by 35 and 0 dB [14], respectively, i.e.

$$SegNER(\lambda) = \min(\max(NER(\lambda), 0), 35), \quad \text{where} \tag{27}$$

$$NER(\lambda) = 10\log_{10}\left(\frac{\sum_{k=0}^{K-1}|N(\lambda,k)|^2}{\sum_{k=0}^{K-1}(|N(\lambda,k)| - |\widehat{N}(\lambda,k)|)^2}\right), \tag{28}$$

and averaged over all (M) frames, i.e.

$$SegNER = \frac{1}{M}\sum_{\lambda=0}^{M-1} SegNER(\lambda). \tag{29}$$

In Table 6 we list the average SegNER's over the same 6 speakers that are used in the informal listening test of the speech enhancement method. We list the average SegNER's for the noise periodogram estimation method, the noise PSD estimation

Exhibit 1009
Page 1422 of 2198

Table 6: Segmental Noise-to-Error Ratios in db

Noise Type	Highway Traffic			Car Interior		
Noisy Speech SNR [dB]	0	5	10	0	5	10
Noise Periodogram Estimation	**19.3**	**17.0**	**14.7**	**18.3**	**16.6**	**15.0**
Noise PSD Estimation	4.6	4.6	4.4	3.0	3.1	3.2
χ^2-based Noise Estimation [3]	3.6	3.1	2.6	2.7	2.3	2.0
MS Noise Estimation [5]	1.0	1.8	2.4	1.9	2.1	2.6

Table 7: Opinion Score Scale

Score	Description
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Table 8: Highway noise speech enhancement results

	SNR [dB]	MOS	SNR [dB]	MOS	SNR [dB]	MOS
Proposed Method	**7.7**	**3.50**	10.3	**3.56**	13.0	**3.74**
MMSE-LSA-DD	7.4	2.75	**11.1**	2.85	**15.0**	3.07
MMSE-LSA	4.6	1.63	9.3	1.92	14.0	2.04
Noisy Speech	0.0	-	5.0	-	10.0	-

Exhibit 1009
Page 1423 of 2198

method, our implementation of χ^2-based noise estimation [3], and minimum statistics (MS) noise estimation [5]. Our implementation of the χ^2-based noise estimation uses the MS noise estimate [5] in the calculation of the optimum smoothing parameters, as suggested by Martin and Lotter [3]. The spectral averaging in our implementation of the χ^2-based noise estimation is performed in sliding spectral windows, of the same size as used by the two proposed noise estimation methods.

We see that the noise PSD estimate has less spectral distortion than both our implementation of the χ^2-based noise estimation [3] and MS noise estimation [5]. This can be explained by a more accurate bias compensation factor, which uses speech presence information. Note that in many scenarios the proposed smooth and low-biased noise PSD estimate is preferable over the noise periodogram estimate.

As an objective measure of time-domain waveform similarity, we list the signal-to-noise ratios, and as a subjective measure of speech quality, we conduct an informal listening test. In this test, test subjects give scores from the scale in Table 7 ranging from 1 to 5, in steps of 0.1, to three different speech enhancement methods, with the noisy speech as a reference signal. Higher score is given to the preferred speech enhancement method. The test subjects are asked to take parameters, such as the naturalness of the enhanced speech, the quality of the speech, and the degree of noise reduction into account, when assigning a score to an estimate. The presentation order of estimates from individual methods is blinded, randomized, and vary in each test set and for each test subject.

A total of 8 listeners, all working within the field of speech signal processing, participated in the test. The proposed speech enhancement method was compared with our implementation of two reference methods:

- MMSE-LSA: Minimum mean-square error log-spectral amplitude estimation, as proposed by Ephraim and Malah [7].
- MMSE-LSA-DD: Decision-directed MMSE-LSA, which is the MMSE-LSA estimation in combination with a smoothing mechanism [7]. Constants are as proposed by Ephraim and Malah.

Table 9: Car interior noise speech enhancement results

	SNR [dB]	MOS	SNR [dB]	MOS	SNR [dB]	MOS
Proposed Method	**10.5**	**3.53**	**13.4**	**3.82**	**16.5**	**3.95**
MMSE-LSA-DD	7.3	2.54	10.9	2.99	15.4	3.29
MMSE-LSA	3.1	1.89	7.7	2.07	12.6	2.37
Noisy Speech	0.0	-	5.0	-	10.0	-

Table 10: Highway traffic noise statistics at a 99% level of confidence

SNR [dB]	Test statistic	Test result	Interval estimate
0	z=10.3	Reject H_0	0.75 ± 0.19
5	z=10.2	Reject H_0	0.72 ± 0.18
10	z=10.1	Reject H_0	0.67 ± 0.17

Table 11: Car interior noise statistics at a 99% level of confidence

SNR [dB]	Test statistic	Test result	Interval estimate
0	z=11.4	Reject H_0	1.00 ± 0.23
5	z=9.4	Reject H_0	0.83 ± 0.23
10	z=6.7	Reject H_0	0.66 ± 0.25

All three methods in the test use the proposed noise PSD estimate,

Also, they all use the analysis/synthesis setup described in Section 0.8.

SNR's and Mean Opinion Scores (MOS) from the informal subjective listening test are listed in Table 8 and Table 9. All results are averaged over both speakers and listeners. The best obtained results are emphasized using bold letters.

To identify if the proposed method is significantly better, i.e. has a higher MOS, than MMSE-LSA-DD, we use the matched sample design [15], where the absolute values of the opinion scores are eliminated as a source of variation. Let μ_d be the mean of the opinion score difference between the proposed method and the MMSE-LSA-DD. Using this formulation, we write the null and alternative hypothesis as:

$$H_0 : \mu_d \leq 0, \tag{30}$$

$$H_A : \mu_d > 0, \tag{31}$$

respectively. The null hypothesis H_0 in this context should not be mistaken for the hypothesis H_0 in the speech presence detection method. With 48 experiments at each combination of SNR and noise type, we are in the large sample case, and we therefore assume that the differences are normally distributed. The rejection rule, at 1% level of significance, is

$$\text{Reject } H_0 \text{ if } z > z_{.01}, \tag{32}$$

with $z_{.01} = 2.33$. Table 10 and 11 list the test statistic z, and the corresponding test result. Also, the two-tailed 99% confidence interval [15], of the difference between the MOS of the proposed method and MMSE-LSA-DD, for highway traffic and car interior noise, respectively, is listed. From our results we can therefore state, with a confidence level of 99% that the proposed method has a higher perceptual quality than MMSE-LSA-DD. Furthermore, the difference corresponds generally to more than 0.5

MOS, which generally change the ratings from somewhere between Poor and Fair to somewhere between Fair and Good on the MOS scale.

Discussion

We have in this paper presented new noise estimation and speech enhancement methods that utilize a proposed connected region speech presence detection method. Despite the simplicity, the proposed methods are shown to have superior performance when compared to our implementation of state-of-the-art reference methods in the case of both noise estimation and speech enhancement.

In the first proposed noise estimation method the connected speech presence regions are used to achieve noise periodogram estimates in the regions where no speech is present. In the remaining regions, where speech is present, minimum tracks of the smoothed noisy speech periodograms are bias compensated with a factor that is updated in regions with no speech presence. A second proposed noise estimation method provides a noise PSD estimate by means of the same power-scaled minimum tracks that are used by the noise periodogram estimation method when speech is present. It is shown that the noise PSD estimate has less spectral distortion than both our implementation of χ^2-based noise estimation [3] and minimum statistics noise estimation [5]. This can be explained by a more accurate bias compensation factor, which uses speech presence information. The noise periodogram estimate is by far the less spectrally distorted noise estimate of the tested noise estimation methods. This verifies the connected region speech presence principle which is fundamental for the proposed speech enhancement method.

Our proposed enhancement method uses different attenuation rules for each of the two types of speech presence regions. When no speech is present the noisy speech is down-scaled and left in the speech estimate as natural sounding masking noise, and when speech is present a noise PSD estimate is used in a traditional generalized spectral subtraction. In addition to enhancing the speech, the most distinct feature of the proposed speech enhancement method is that it leaves natural sounding background noise matching the actual surroundings of the person wearing the hearing aid. The

proposed method performs well at SNR's equal to or higher than 0 dB for noise types with slowly changing and spectrally smooth periodograms. Rapid, and speech-like, changes in the noise will be treated as speech, and will therefore be enhanced, causing a decrease in the naturalness of the background noise. At very low SNR's the detection of speech presence will begin to fail. In this case, we suggest the implementation of the proposed method in a scheme, where low SNR is detected and causes a change to an approach with only a single and very conservative attenuation rule. Strong tonal interferences will affect the speech presence decisions as well as the noise estimation and enhancement method and should be detected and removed by preprocessing of the noisy signal immediately after the STFT analysis. Otherwise, a sufficiently strong tonal interference with duration longer than the minimum search window will cause the signal to be treated as if speech is absent and the speech enhancement algorithm will down-scale the entire noisy speech by multiplication with β_0.

Our approach generalizes to other noise reduction schemes. As an example the proposed binary scheme can also be used with MMSE-LSA-DD for the speech presence regions. For such a combination we expect performance similar to, or better than, what we have shown in this paper for the generalized spectral subtraction. This is supported by the findings of Cohen and Berdugo [11] that have shown that a soft-decision approach improves MMSE-LSA-DD.

The informal listening test confirms that listeners prefer the down-scaled background noise with fully preserved naturalness over the less realistic whitened residual noise from e.g. MMSE-LSA-DD. From our experiments we can conclude, with a confidence level of 99%, that the proposed speech enhancement method receives significantly higher MOS than MMSE-LSA-DD at all tested combinations of SNR and noise type.

References

[1] T. Painter and A. Spanias, "Perceptual Coding of Digital Audio," *Proc. IEEE*, vol. 88, no. 4, pp. 451–513, Apr. 2000.

[2] J. Yang, "Frequency Domain Noise Suppression Approaches in Mobile Telephone Systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 2, Apr. 1993, pp. 363–366.

[3] R. Martin and T. Lotter, "Optimal Recursive Smoothing of Non-Stationary Periodograms," in *Proc. IEEE Int. Workshop on Acoustic Echo and Noise Control*, Sept. 2001, pp. 43–46.

[4] R. Martin, "Spectral Subtraction Based on Minimum Statistics," in *Proc. EURASIP European Signal Processing Conference*, Sept. 1994, pp. 1182–1185.

[5] ——, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," *IEEE Trans. Speech Audio Processing*, vol. 9, no. 5, pp. 504–512, July 2001.

[6] I. Cohen and B. Berdugo, "Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement," *IEEE Signal Processing Lett.*, vol. 9, no. 1, pp. 12–15, Jan. 2002.

[7] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 443–445, Apr. 1985.

[8] K. S. Shanmugan and A. M. Breipohl, *Random Signals - Detection, Estimation and Data Analysis.* John Wiley & Sons, Inc., 1988.

[9] K. V. Sørensen and S. V. Andersen, "Speech Presence Detection in the Time-Frequency Domain using Minimum Statistics," in *Proc. Nordic Signal Processing Symposium*, June 2004, pp. 340–343.

[10] I. Cohen, "Noise Spectrum Estimation in Adverse Environments: Improved Minima Controlled Recursive Averaging," *IEEE Trans. Speech Audio Processing*, vol. 11, no. 5, pp. 466–475, Sept. 2003.

[11] I. Cohen and B. Berdugo, "Speech Enhancement for Non-Stationary Noise Environments," *Signal Processing*, vol. 81, no. 11, pp. 2403–2418, Nov. 2001, Elsevier Science.

[12] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 4, Apr. 1979, pp. 208–211.

[13] *DARPA TIMIT Acoustic-Phonetic Speech Database*, National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA, CD-ROM.

[14] J. R. Deller, Jr., J. H. L. Hansen, and J. G. Proakis, *Discrete-Time Processing of Speech Signals.* Wiley-Interscience, 2000.

[15] D. R. Anderson, D. J. Sweeney, and T. A. Williams, *Statistics for Business and Economics.* South-Western, 1990.

[16] D. Brillinger, *Time Series: Data Analysis and Theory.* Holden-Day, 1981.

In the following, a detailed description of a preferred method for noise power spectral density estimation from noisy speech using in-line trained hidden Markov models. This method is used in conjunction with the described speech presence detection method that provides connected time-frequency regions of each decision type as a basis for the noise estimation. In speech presence regions hidden Markov models are trained on-line and in speech presence regions the trained models are used for (MMSE) optimum estimation. Both types of speech presence regions can be present in each frame and on-line training of the models in speech absence can be conducted while the models in speech presence are used for estimation. Experiments show that the proposed noise power spectral density estimation performs better than three state of the art reference methods. For real life noise types the special case of the hidden Markov model where it reduces to a Gaussian mixture model is shown to be nearly as good as the hidden Markov model.

Statistical noise estimation methods very often relies on an assumption of stationarity; the parameters of the noise PDF are estimated during speech absence and kept constant during speech presence. In this paper we investigate if dynamic modeling of the noise PDF using hidden Markov models (HMM's) results in better performance than when using static Gaussian mixture models (GMM's). We evaluate the performance for real-life noise types. Transition probabilities in the HMM's will, to a certain degree, capture the dynamic behavior of non-stationary noise. The general method of HMM's have already proven its worth for modeling different classes of noise. In particular, Sameti et. al. [1] and Gaunard et. al. [2] have used off-line trained noise models for noise classification. In their approach only a predefined set of noise classes are modeled and the off-line training is supervised. Similarly trained models have been suggested by Ghoreishi and Sheikhzadeh [3] for speech pause detection. A subband approach has been proposed by Hosoki et. al. [4] for noise detection. If a subband is unlikely to contain clean speech they classify it as being noise contaminated.

In previous work [5] we have proposed a connected time-frequency domain region speech presence detector and applied it to find bias compensation factors for minimum statistics based noise estimation. We have shown that this approach results in a less spectrally distorted noise estimate than the original minimum statistics based noise estimation [6]. In this paper, we apply our new approach to train subband HMM's on-line while speech is absent. When speech is present we use the most recently trained noise models for MMSE optimum noise power spectral density (PSD) estimation. This way, instead of choosing from a finite set of predefined noise class models, the proposed method models the local behavior of the noise to more accurately adapt to the actual noise environment.

The remainder of this paper is organized as follows. Section 2 describes the signal model, the speech presence hypotheses, the spectral smoothing method, and the fundamental noise PSD estimation approach. Section 3 provides the details of the applied statistical model. In Section 4 the methods for estimation of state probabilities and unknown observations are described. Section 5 contains the experiments and Section 6 provides a discussion of the proposed method and the obtained results.

Signal Model

We assume that noisy speech $y(i)$ at sampling time index i consists of speech $s(i)$ and additive noise $n(i)$. For joint time-frequency analysis of $y(i)$ we apply the L-point discrete Short-Time Fourier Transform (STFT), i.e.

$$Y(\tau,\omega)=\sum_{\mu=0}^{L-1} y(\tau R+\mu)h(\mu)\exp\left(-j2\pi\omega\mu/L\right), \tag{1}$$

where $\tau\in\mathbb{Z}$ is the (sub-sampled) time index, $\omega\in\{0,1,\ldots,L-1\}$ is the frequency index, and L is the STFT size, which in this paper equals the window length. R is the skip between frames and $h(\mu)$ is a unit energy window function, i.e. $\sum_{\mu=0}^{L-1}h^2(\mu)=1$. From the linearity of (1) we have that $Y(\tau,\omega)=S(\tau,\omega)+N(\tau,\omega)$, where $S(\tau,\omega)$ and $N(\tau,\omega)$ are the STFT coefficients of speech $s(i)$ and additive noise $n(i)$, respectively. We further assume that $s(i)$ and $n(i)$ are zero mean and statistically independent, which leads to a power relation where the noise is additive. Now, let the hypotheses

H_0 and H_1 for speech absence and speech presence, respectively, be defined by two power relations, i.e.

$$H_0: \quad E\{|Y(\tau,\omega)|^2\} = E\{|N(\tau,\omega)|^2\} \tag{2}$$

$$H_1: \quad E\{|Y(\tau,\omega)|^2\} = E\{|S(\tau,\omega)|^2\} + E\{|N(\tau,\omega)|^2\}. \tag{3}$$

The decision of which hypothesis to believe is true is done by means of the connected region speech presence detection method, which we have proposed in previous work [5]. This speech presence detection method provides individual decisions at each time-frequency location. At the same time it ensures that decisions of the same type are connected in larger time-frequency regions. In the regions where no speech is detected we can directly observe what we assume to be the realizations of the stochastic noise process. The approach taken in this paper is to exploit this property to train dynamic statistical noise models in connected regions of speech absence and use it for noise PSD estimation in regions of speech presence. Initially, we apply a spectral window of size $2D+1$, centered at ω, to reduce the fluctuations of the noisy speech periodogram bins, i.e.

$$B(\tau,\omega) = \frac{1}{2D+1} \sum_{\widetilde{\omega}=\omega-D}^{\omega+D} |Y(\tau,\widetilde{\omega})|^2. \tag{4}$$

At each time-frequency location (τ,ω) we let the spectrally averaged $B(\tau,\omega)$ constitute the noise PSD estimate if speech is absent in all the noisy speech periodogram bins within the spectral window in (4). If speech is present in any of these bins we turn to HMM based MMSE estimation.

Training The Statistical Model

We model the spectrally averaged noisy periodogram bins $B(\tau,\omega)$ at each ω using a continuous density HMM with a Gaussian mixture model in each state of the HMM modeling the observation PDF. At current time, say T, we consider the T' most recent

spectrally smoothed periodograms, i.e. $B(\tau,\omega)$ for $\mathcal{T}-T'<\tau\leq\mathcal{T}$. If no noisy speech periodogram bins with speech presence was used in (4) to calculate any of these, we denote the case $D(\mathcal{T},\omega)=0$ and otherwise we denote it $D(\mathcal{T},\omega)=1$. Binary speech presence detection methods generally needs a certain amount of speech power to detect speech presence. Therefore, the last few noisy speech periodogram bins leading up to a speech presence region will most likely contain enough speech power to contaminate the HMM training. To avoid this, we train the HMM on the training set that consist of the first T spectrally smoothed periodograms within the sliding window of length T'. We train the model on the training set only if $D(\mathcal{T},\omega)=0$. If $D(\mathcal{T},\omega)=1$, the model parameters from $\mathcal{T}-1$ are preserved except for the forward likelihoods, which are estimated by prediction from the forward likelihoods and state transition probabilities of the model at $\mathcal{T}-1$. The set of T spectrally smoothed bins $B(\tau,\omega)$ for $\mathcal{T}-T'<\tau\leq\mathcal{T}-T'+T$ that constitutes a training set of T scalar observations at ω will cause the means and variances of the GMM's to be scalars. For easy generalization to the vector case we will, however, describe the theory using vector/matrix notation. As the training procedure is the same for all sets of training vectors we use the same notation for all $\mathcal{T}$ and ω, i.e. $\mathbf{X}_1^T\triangleq[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_{T-1},\mathbf{x}_T]$. We want to find model parameters Φ that maximize the joint likelihood[1] conditioned on the training set $\mathbf{X}_1^T$ of T observation vectors by adjusting the model parameters, i.e.

$$\widehat{\Phi}_{ML}=\arg\max_{\Phi}p(\mathbf{X}_1^T|\Phi). \tag{5}$$

This optimization, however, of the (generally) non-convex objective function $p(\mathbf{X}_1^T|\Phi)$ requires knowledge of the hidden states and mixture components and is therefore not feasible. Instead we use the Baum-Welch algorithm [7], which by alternating maximization will converge to a model parameter estimate corresponding to a local maximum of the likelihood function. To ensure numerical stability we use the lower limit $\epsilon_1=10^{-6}$ on the multivariate Gaussian mixture components c_{jk}, the single element of the 1-by-1 covariance matrices Σ_{jk}, and the sampled individual Gaussians $b_{jk}(\mathbf{x}_t)$ for all mixture numbers $1\leq k\leq K$ and states $1\leq j\leq N$. Also, we use a lower limit $\epsilon_2=10^{-2}$ on

[1] We use $p(\cdot)$ to denote probability density functions and $P(\cdot)$ to denote probability mass functions.

all entries in the state transition probability matrix $\mathbf{A} = \{a_{ij}\} = P(s_t = j|s_{t-1} = i)$ [8, pp. 381-382] for the states $1 \leq i, j \leq N$.

Let $\alpha_t(i)$ denote the forward likelihoods, defined as the likelihoods for being in state i at t while having produced the observations $\mathbf{X}_1^t$, conditioned on the model, i.e.

$$\alpha_t(i) \triangleq p(\mathbf{X}_1^t, s_t = i|\Phi). \tag{6}$$

The forward likelihoods $\alpha_0(i)$ initializing the HMM training for $t = 0$ at each $\mathcal{T}$ are initialized as uniform distributions. After initialization at $t = 0$ the forward likelihoods for $1 \leq t \leq T$ in the Baum-Welch training algorithm are induced from the forward likelihoods at $t - 1$, i.e.

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) a_{ij} b_j(\mathbf{x}_t), \tag{7}$$

where the sampled observation PDF $b_j(\mathbf{x}_t)$ from state j is given by the weighted mixture of sampled Gaussians $b_{jk}(\mathbf{x}_t)$, i.e.

$$b_j(\mathbf{x}_t) = \sum_{k=1}^{M} c_{jk} b_{jk}(\mathbf{x}_t) = \sum_{k=1}^{M} c_{jk} \mathcal{N}(\mathbf{x}_t; \boldsymbol{\mu}_{jk}, \boldsymbol{\Sigma}_{jk}). \tag{8}$$

To avoid numerical instability in the implementation all forward likelihood vectors are scaled to unity l_1-norm [8], i.e.

$$\widetilde{\boldsymbol{\alpha}}_t = \boldsymbol{\alpha}_t / \|\boldsymbol{\alpha}_t\|_1. \tag{9}$$

This scaling does not affect training nor does it affect HMM based estimation.

Hidden Markov Model based Estimation

When speech presence is detected in a noisy speech periodogram bin it will cause $D(\tau, \omega) = 1$ for T' successive values of τ at all $2D + 1$ frequency indices ω where it is

located within the spectral window in (4). While $D(\tau,\omega)=1$ the forward likelihoods $\boldsymbol{\alpha}_{T'}$ at ω, which are required for HMM based estimation, are predicted from the most recently trained model at ω, i.e. the model at the most recent τ for which $D(\tau,\omega)=0$. We denote the last observable training set, on which this model has been trained, $\widetilde{\mathbf{X}}_1^T$. When $D(\mathcal{T},\omega)=1$ we set the sampled observation PDF equal to one for all states j, i.e. $b_j(\mathbf{x}_t)=1$. This corresponds to an observation $\mathbf{x}_t$ that does not affect the forward likelihoods and therefore it leads to a simplified version of the forward likelihood induction equation in (6), i.e. $\alpha_t(j)=\sum_{i=1}^N \alpha_{t-1}(i)a_{ij}$ for $j\in\{1,\ldots,N\}$. The simplified equation can be compactly written in vector/matrix notation, i.e.

$$\boldsymbol{\alpha}_t=\mathbf{A}^T\boldsymbol{\alpha}_{t-1}. \tag{10}$$

From (10) it follows that the F'th successively estimated forward likelihood vector for T' (at current time $\mathcal{T}$) is given by

$$\boldsymbol{\alpha}_{T'}=(\mathbf{A}^T)^F\boldsymbol{\alpha}_{T'-F}. \tag{11}$$

In the above $\boldsymbol{\alpha}_{T'-F}$ is relative to current time $\mathcal{T}$, i.e. it equals $\boldsymbol{\alpha}_T$ from time $\mathcal{T}-F+(T'-T)$. We now investigate whether or not this estimate will converge to a "steady state". Suppose $\mathbf{A}^T\in\mathbb{R}^{N\times N}$ has N linearly independent eigenvectors. Then, by means of the eigenvalue decomposition we have that $\mathbf{A}^T=\mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix with non-increasingly sorted eigenvalues $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_{N-1}\geq\lambda_N$ on the diagonal and $\mathbf{S}=[\mathbf{s}_1,\ldots,\mathbf{s}_N]$ is a matrix of associated eigenvectors. We then have that $(\mathbf{A}^T)^F=\mathbf{S}\boldsymbol{\Lambda}^F\mathbf{S}^{-1}$. For the strictly positive, hence primitive, Markov matrix $\mathbf{A}^T$ the Perron-Frobenius theorem for primitive matrices [9, Theorem 1.1] states that there will be a unique dominant eigenvalue $\lambda_1>|\lambda_i|$ for any eigenvalue $\lambda_i\neq\lambda_1$, which can be associated a strictly positive (or strictly negative) dominant eigenvector $\mathbf{s}_1$. For a positive Markov matrix the dominant eigenvalue will be $\lambda_1=1$ [9, p.118]. Now, if we let $F\to\infty$ we have that the dominant eigenvalue remains one while the rest goes to zero. Therefore, $\lim_{F\to\infty}\mathbf{S}\boldsymbol{\Lambda}^F\mathbf{S}^{-1}$ becomes a rank-1 projection matrix that projects onto the subspace $\mathcal{C}(\mathbf{s}_1)\in\mathbb{R}^N$, i.e. the line, spanned by a dominant eigenvector $\mathbf{s}_1$ of

$\mathbf{A}^T$. Since we have that $\alpha_{T'-F}$ is strictly positive it is not possible that $\mathbf{s}_1 \perp \alpha_{T'-F}$. In effect,

$$\lim_{P\to\infty} \mathbf{S}\Lambda^F\mathbf{S}^{-1}\alpha_{T'-F} = c\mathbf{s}_1, \tag{12}$$

where $c \in \mathbb{R} \setminus 0$ is a constant. In the practical implementation, we have that $\alpha_{T'-F}$ is scaled by (9) to unity l_1-norm. Since left multiplication with the Markov matrix $\mathbf{A}^T$ does not affect the l_1-norm the limit value remains $c\mathbf{s}_1$, but readily scaled to unity l_1-norm. Thus, $c = sign(s_1(i))/||\mathbf{s}_1||_1$ for any $i \in \{1, \ldots, N\}$. To summarize, forward likelihood vectors predicted successively by (10) converge to a dominant eigenvector of $\mathbf{A}^T$. In fact, this implies that the HMM converges to a special case where it reduces to a GMM.

Once the forward likelihoods have been estimated we obtain the noise PSD estimate $\widehat{\mathbf{x}}_T$ as the conditional MMSE optimum observation vector $\widehat{\mathbf{x}}_T^\star$. i.e.

$$\widehat{\mathbf{x}}_T^\star = \arg\min_{\widehat{\mathbf{x}}_T} E\{(\widehat{\mathbf{x}}_T - \mathbf{x}_T)^T(\widehat{\mathbf{x}}_T - \mathbf{x}_T)|\widetilde{\mathbf{X}}_1^T, \Phi\}. \tag{13}$$

We assume that the unknown observation vector $\mathbf{x}_T$ is drawn from a continuous density HMM with observation PDF's modeled by a GMM in each state, i.e.

$$\mathbf{x}_T \sim \sum_{j=1}^{N}\sum_{k=1}^{K} P(s_T = j|\widetilde{\mathbf{X}}_1^T, \Phi)c_{jk}\mathcal{N}(\mathbf{x}_T; \boldsymbol{\mu}_{jk}, \boldsymbol{\Sigma}_{jk}). \tag{14}$$

The conditional MMSE estimate of the observation parameter vector is therefore given by

$$\widehat{\mathbf{x}}_T^\star = \sum_{j=1}^{N}\sum_{k=1}^{K} P(s_T = j|\widetilde{\mathbf{X}}_1^T, \Phi)c_{jk}\boldsymbol{\mu}_{jk}, \tag{15}$$

where the j'th conditional state probability

$$P(s_T = j|\widetilde{\mathbf{X}}_1^T, \Phi) = p(s_T = j, \widetilde{\mathbf{X}}_1^T|\Phi)/p(\widetilde{\mathbf{X}}_1^T|\Phi) \tag{16}$$

is obtained as the j'th entry in $\widetilde{\alpha}_T$, given by (9).

Experiments

For the experiments we use signals sampled at 8 kHz sampling frequency. Both frame and FFT size are $L = 256$ and the frame skip is $R = 128$ samples. The analysis window is a square-root Hanning scaled to unit energy. $D = 2$ defines the size of the spectral window and the number of observations in a training set is $T = 16$. These are taken from a sliding window of length $T' = 20$. In the experiments we evaluate the performance in subbands consisting of individual frequency tracks, i.e. scalar observations for all HMM's. The following methods are compared:

- *HMM(5,1)*: The proposed HMM based noise estimation method with $M = 1$ Gaussian in each state and $N = 5$ states.
- *GMM(5)*: The proposed method with $M = 5$ and $N = 1$. With $N = 1$ the HMM reduces to a GMM with M Gaussians.
- *CR-SPD*: Connected time-frequency region speech presence detection based smooth noise estimation [5].
- *MCRA*[2]: Minima controlled recursive averaging [10].
- *MS*: Minimum statistics noise estimation [6].

The first three methods are based on the connected time-frequency region speech presence detector proposed in [5]. The last two methods, MCRA and MS, serve as well known reference methods, which both feature independence from explicit speech presence detection. The performance of these methods are evaluated by means of their spectral distortion, which we measure as average segmental noise-to-error ratios (SegNER's). We calculate the SegNER directly in the time-frequency domain, as the ratio in dB between the noise energy and the noise estimation error energy. These values are upper and lower limited by 35 and 0 dB [11, p.586], respectively, and averaged over all T' frames, i.e.

[2]MCRA is implemented for 8 kHz sampling frequency and 16 ms frame skip. Filter coefficients have time constants equal to the coefficients proposed by Cohen and Berdugo [10].

$$SegNER = \frac{1}{T'} \sum_{\tau=1}^{T'} \min(\max(NER(\tau), 0), 35), \tag{17}$$

where the noise-to-error ratio, in dB, at τ is given by

$$NER(\tau) = 10 \log_{10} \left(\frac{\sum_{\omega=0}^{L-1} |N(\tau,\omega)|^2}{\sum_{\omega=0}^{L-1} (|N(\tau,\omega)| - |\widehat{N}(\tau,\omega)|)^2} \right). \tag{18}$$

We evaluate for noisy speech with 4 different noise types, i.e. highway traffic, car interior, white, and helicopter noise, which (with zero mean) are added to the speech in -5, 0, 5, and 10 dB SNR. In each combination of SNR and noise type we average the SegNER's over speech from 3 male and 3 female speakers from the TIMIT database [12]. Initially, we illustrate the distinct dynamic modeling abilities and the "steady state" estimate convergence of HMM(5,1) with an example where the model is trained on the first 16 observations. For illustrative purposes, as well as for implementation verification, we have in this example used artificially created noise with dynamics that are well captured by the applied model. We conclude that the initialization and subsequent Baum-Welch training lead to model parameters where the dynamics of the noise have been well captured by the model.

The best average SegNER in each combination of noise type and input SNR is emphasized using bold letters. For -5 dB highway traffic noise HMM(5,1) and GMM(5) are equally good, so we emphasize the average SegNER of the method that was best before the SegNER's were rounded.

From the table we see that HMM(5,1) has the highest average SegNER's for highway traffic, car interior, and helicopter noise and GMM(5) has the highest average SegNER's for white noise. That GMM(5) is better than HMM(5,1) in white noise is to be expected since the additional degrees of freedom in the HMM(5,1) relative to the GMM(5) tend to cause over-modeling. This is particularly the case when the training sets are small. In case of the HMM(5,1) the local variations in the realizations of stationary stochastic noise processes are, generally, captured by a dynamic model and in case of the GMM(5) by a static model. In all the tested combinations of noise type and SNR the proposed method is considerably better than both MCRA [10] and MS [6]. This is the case in

regions of speech presence as well as in regions of speech absence. We note that in all cases the SegNER's for HMM(5,1) and GMM(5) have only minor differences.

Discussion

We have proposed a HMM based method for noise PSD estimation that depends on an external connected time-frequency region speech presence detector. The method is trained on-line in speech absence and applied on-line for noise PSD estimation in connected regions of speech presence. Estimates from the proposed method are consistently less spectrally distorted than estimates from any of the three reference methods.

We have shown that for the tested real-life noise types there are only minor differences in performance between GMM(5) and HMM(5,1), i.e. the static model leads to similar performance as the dynamic model. For noise environments with clearly dynamically changing noise PSD PDF's the HMM would be advantageous, though, as does have the ability to model dynamic behavior of the PDF, that is, non-stationary noise. However, in order for the HMM to be a significantly better model than the GMM the number of states and Gaussians in each GMM must be adequately chosen. Also, the size of the training sets needs to be large enough for the dynamics to be captured during the Baum-Welch training.

At the cost of increased computational complexity and memory requirements the number of states and Gaussians in the HMM could be estimated on-line during speech pauses. At each ω various models could be trained and the best choice could be made from measuring the distortion between estimated and observed noise.

We have shown that the MMSE estimate under certain conditions converges to a "steady state" estimate determined by a dominant eigenvector of the transposed transition matrix. In case that the columns of the transposed state transition probability matrix A^T after Baum-Welch training remains uniformly distributed the trace, which equals the sum of the eigenvalues, will be 1. Therefore the dominant eigenvalue of

the symmetric and positive semi-definite matrix, implying non-negative eigenvalues [13, p.269], will be the only eigenvalue different from zero. This will cause immediate convergence to the "steady state" estimate where the HMM reduces to a GMM. For stationary noise we consider this to be desired behavior. More generally, the gap in magnitude between the dominant eigenvalue and the remaining eigenvalues affects the rate of convergence. There will also be an impact from the angle between the forward likelihood vector and the individual subspaces spanned by associated eigenvectors. Experiments have shown that the rate of convergence for a number of models trained on the same training set with different initializations differs for parameters associated with each of the local minima of the likelihood function. There is no direct relationship, however, between the likelihood of a local minimum and the rate of convergence.

For environments with increasing or decreasing noise levels delta parameters will be better suited for the proposed noise estimation method. They will give the "steady state" estimate the ability to follow the level of the noise. Using non-delta representation is the most conservative approach and will unlike the delta representation ensure a stable MMSE estimate. It will, however, not be able to follow increasing nor decreasing noise during speech presence.

If the statistical models are trained with no spectral smoothing of the training sets, i.e. with $D = 0$, the method proposed in this paper could easily be modified to provide estimated noise PSD PDF's at each frequency. This makes the proposed method applicable, and very well suited, for statistical speech enhancement.

The method described in this paper can be applied on subbands of any width. For the model to benefit from any inter-frequency dependencies the vectors in each subband should be modeled using full (non-diagonal) covariance matrices. Better performance could very well be a consequence of the ability to model inter-frequency dependencies in the noise. HMM modeling employing parametric descriptions of larger time-frequency regions is a topic of current research.

References

[1] H. Sameti, H. Sheikhzadeh, L. Deng, and R. L. Brennan, "HMM-Based Strategies for Enhancement of Speech Signals Embedded in Nonstationary Noise," *IEEE Trans. Speech Audio Processing*, vol. 6, no. 5, pp. 445–455, Sept. 1998.

[2] P. Gaunard, C. G. Mubikangiey, C. Couvreur, and V. Fontaine, "Automatic Classification of Environmental Noise Events by Hidden Markov Models," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 6, May 1998, pp. 3609–3612.

[3] M. H. Ghoreishi and H. Sheikhzadeh, "A Hybrid Speech Enhancement System Based on HMM and Spectral Subtraction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 3, June 2000, pp. 1855–1858.

[4] M. Hosoki, T. Nagai, and A. Kurematsu, "Speech Signal Band Width Extension and Noise Removal using Subband HMM," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 1, May 2002, pp. 245–248.

[5] K. V. Sørensen and S. V. Andersen, "Speech Enhancement with Natural Sounding Residual Noise based on Connected Time-Frequency Speech Presence Regions," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 18, pp. 2954–2964, Oct. 2005.

[6] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," *IEEE Trans. Speech Audio Processing*, vol. 9, no. 5, pp. 504–512, July 2001.

[7] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Annals of Mathematical Statistics*, vol. 41, pp. 164–171, 1970.

[8] L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition.* Prentice-Hall, 1993.

[9] E. Seneta, *Non-negative Matrices and Markov Chains.* Springer-Verlag, 1981.

[10] I. Cohen and B. Berdugo, "Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement," *IEEE Signal Processing Lett.*, vol. 9, no. 1, pp. 12–15, Jan. 2002.

[11] J. R. Deller, Jr., J. H. L. Hansen, and J. G. Proakis, *Discrete-Time Processing of Speech Signals.* Wiley-Interscience, 2000.

[12] *DARPA TIMIT Acoustic-Phonetic Speech Database*, National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA, CD-ROM.

[13] P. Stoica and R. Moses, *Introduction to Spectral Analysis.* Prentice-Hall, 1997.

Exhibit 1009
Page 1444 of 2198

Fig. 2 illustrates a block diagram of a preferred device embodiment. The illustrated device may be such as a mobile phone, a headset or a part thereof. The device is adapted to receive a noisy signal, e.g. an electrical analog or digital signal representing an audio signal containing speech and unintended noise. The device includes a digital signal processor DSP that performs a signal processing on the noisy signal. First, a speech presence detection is performed such as described in the foregoing including spectral-temporal smoothing prior to making a speech presence decision, and such as will be described in details in the following. The speech presence detection serves as input to a speech enhancement algorithm as already mention and which will also be described in more details in the following. The output of the speech enhancement algorithm is a signal where the speech is enhanced in relation to the noise. This signal with enhanced speech is applied to a loudspeaker, preferably via an amplifier, so as to present an acoustic representation of the speech enhanced signal to a listener.

As mentioned, the device in Fig. 2 may be a hearing aid, a headset or a mobile phone or the like. In case of a headset, the DSP may either be built into the headset, or the DSP may be positioned remote from the headset, e.g. built into other equipment such as amplifier equipment. In case of a hearing aid, the noisy signal can originate from a remote audio source or from microphone built into the hearing aid.

It is to be understood that reference signs in the claims should not be construed as limiting with respect to the scope of the claims.

Claims

1. A method for speech presence detection of speech in a noisy signal, the noisy signal being represented as a temporal-spectral distribution, the method including performing a temporal-spectral smoothing on the temporal-spectral representation of the noisy signal prior to making a speech presence decision.

2. Method according to claim 1, wherein the temporal-spectral smoothing includes performing a spectral smoothing prior to performing a temporal smoothing.

3. Method according to claim 2, wherein the temporal smoothing includes performing a recursive smoothing with time and frequency varying smoothing parameters.

4. Method according to claim 3, wherein the smoothing parameters are lower limited so as to ensure a minimum degree of smoothing.

5. Method according to any of the preceding claims, wherein the speech presence decision is based on a minimum tracking of an output of the temporal-spectral smoothing.

6. Method according to any of the preceding claims, wherein a noise estimation is performed based on the speech presence decision.

7. Method according to any of the preceding claims, wherein the speech presence decision is a binary decision.

8. A speech enhancement method including performing the method according to any of the preceding claims, and performing enhancement of the speech in the noisy signal based on the speech presence detection.

9. Method according to claim 8, wherein a first attenuation function is applied to regions of the temporal-spectral distribution where presence of speech is detected, and a second attenuation function is applied to other regions of the temporal-spectral distribution, the first and second attenuation functions being different.

10. Method according to claim 9, wherein the second attenuation function applied to the other regions of the temporal-spectral distribution is a scalar attenuation.

11. Method according to claim 10, wherein the scalar attenuation is implemented by multiplying by a scalar in the range 0.05 to 0.4, preferably 0.2.

12. Device including a processor adapted to perform the method according to any of the preceding claims.

13. Device according to claim 12, the device being selected from the group consisting of:

a mobile phone, a radio communication device, an internet telephony system, sound recording equipment, sound processing equipment, sound editing equipment, broadcasting sound equipment, and a monitoring system.

14. Device according to claim 12, the device being selected from the group consisting of: a hearing aid, a headset, an assistive listening device, an electronic hearing protector, and a headphone with a built-in microphone.

15. Computer executable program code adapted to perform the method according to any of claims 1-11.

1/2

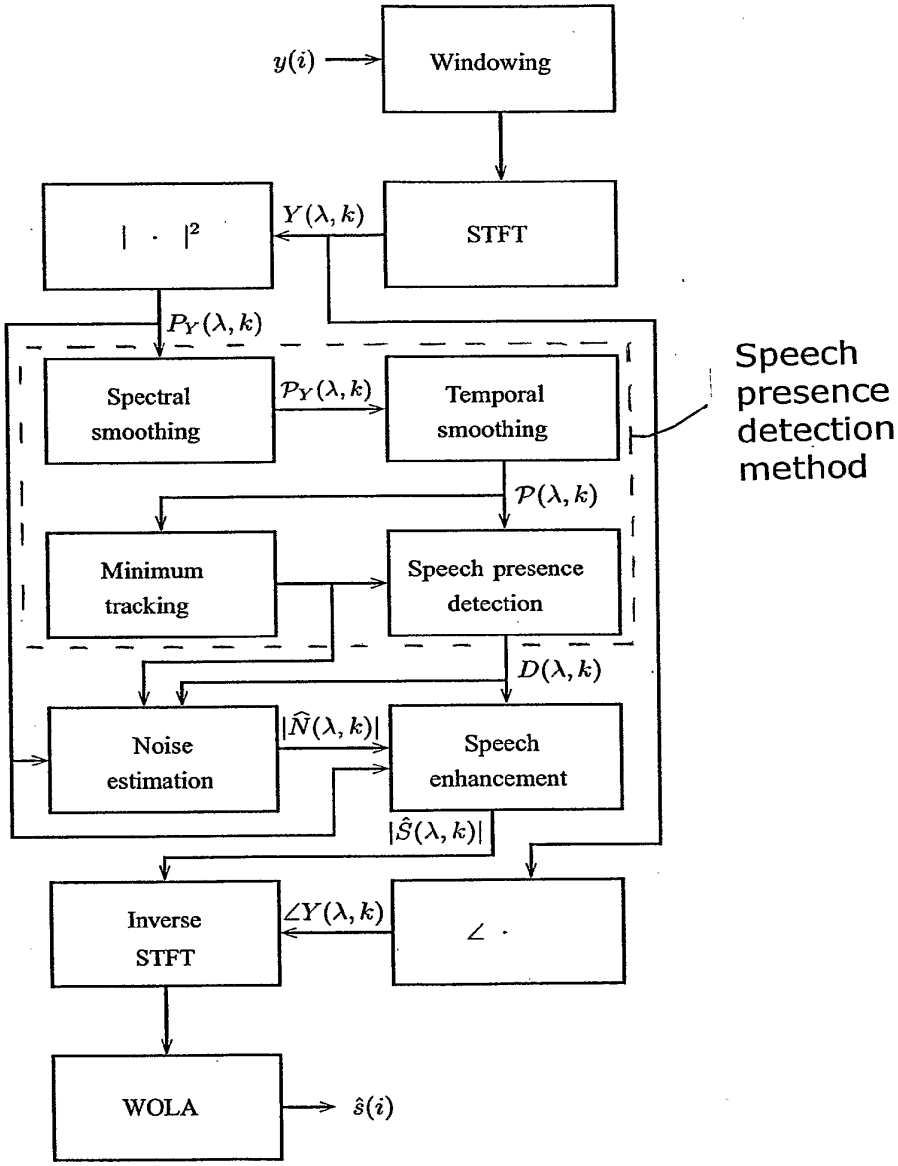

$y(i)$
Windowing
STFT
$Y(\lambda,k)$
$|\cdot|^2$
$P_Y(\lambda,k)$
Spectral smoothing
$\mathcal{P}_Y(\lambda,k)$
Temporal smoothing
$\mathcal{P}(\lambda,k)$
Minimum tracking
Speech presence detection
$D(\lambda,k)$
Noise estimation
$|\widehat{N}(\lambda,k)|$
Speech enhancement
$|\hat{S}(\lambda,k)|$
Inverse STFT
$\angle Y(\lambda,k)$
$\angle\cdot$
WOLA
$\hat{s}(i)$
Speech presence detection method

Fig. 1

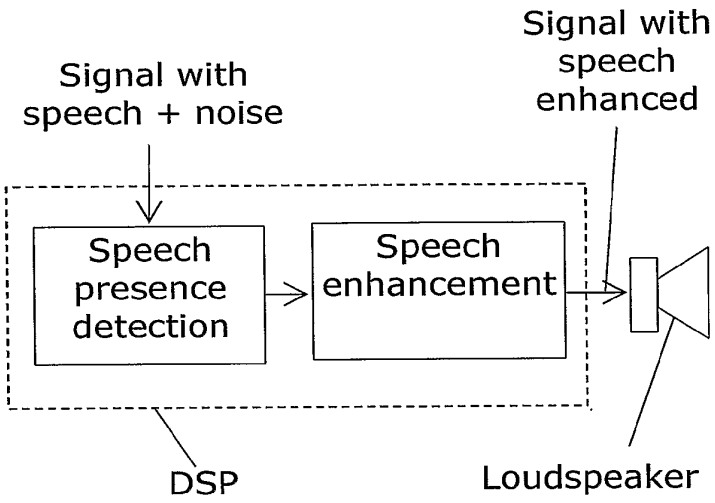

Signal with speech + noise
Signal with speech enhanced
Speech presence detection
Speech enhancement
DSP
Loudspeaker

Fig. 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/DK2006/000221

A. CLASSIFICATION OF SUBJECT MATTER
INV. G10L21/02 G10L11/02

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)
G10L

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 practical, search terms used)
EPO-Internal, WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SORENSEN K V ET AL: "Speech presence detection in the time-frequency domain using minimum statistics" PROCEEDINGS OF THE 6TH NORDIC SIGNAL PROCESSING SYMPOSIUM (IEEE CAT. NO.04EX856) IEEE PISCATAWAY, NJ, USA, 2004, pages 340-343, XP002385166 page 340, right-hand column, line 21 - page 341, right-hand column, line 21	1-3,5-9, 12-15
Y	page 341, left-hand column, line 6 - line 7	4
Y	page 340, right-hand column, line 11 - line 13	10,11

-/--

[X] Further documents are listed in the continuation of Box C.

[X] 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
13 June 2006	23/06/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Burchett, S

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MARTIN R: "Noise power spectral density estimation based on optimal smoothing and minimum statistics" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING IEEE USA, vol. 9, no. 5, July 2001 (2001-07), pages 504-512, XP002385167 ISSN: 1063-6676 cited in the application page 507, left-hand column, line 25 - line 29	4
Y	BEROUTI M ET AL: "ENHANCEMENT OF SPEECH CORRUPTED BY ACOUSTIC NOISE" INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH & SIGNAL PROCESSING. ICASSP. WASHINGTON, APRIL 2 - 4, 1979, NEW YORK, IEEE, US, vol. CONF. 4, 1979, pages 208-211, XP001079151 page 209, left-hand column, paragraph 3 - page 210, left-hand column, paragraph 2	10,11
X	US 6 363 345 B1 (MARASH JOSEPH ET AL) 26 March 2002 (2002-03-26) column 5, line 34 - column 7, line 43	1,2,5-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/DK2006/000221

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6363345 B1	26-03-2002	NONE	

Form PCT/ISA/210 (patent family annex) (April 2005)

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 674 061 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
28.06.2006 Bulletin 2006/26

(51) Int Cl.:
A61F 11/14 *(2006.01)* ***H04R 1/10*** *(2006.01)*

(21) Application number: **04030646.6**

(22) Date of filing: **23.12.2004**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR LV MK YU

(71) Applicant: **PHONAK AG**
8712 Stäfa (CH)

(72) Inventor: **Berg, Christian**
8713 Uerikon (CH)

(74) Representative: **Schwan - Schwan - Schorer**
Bauerstrasse 22
80796 München (DE)

(54) Active hearing protection system and method

(57) The invention relates to a hearing protection system comprising a first and a second hearing protection device (10), each hearing protection device comprising an active unit (16) comprising an acoustic input transducer (18, 19) for converting ambient sound into input audio signals, and an acoustic output transducer (22) for transforming filtered audio signals into sound perceivable by said user, wherein at least one of said active units comprises an audio signal processing unit (20) for processing said input audio signals into said filtered audio signals; wherein said audio signal processing unit comprises an analyzer module (34) for determining the intensity of said input audio signals separately for a plurality of spectral classes, a judgement module (36) for judging, depending on said determined spectral intensities of said input digital audio signals, which one of a plurality of predetermined criteria is presently fulfilled by the input audio signals, and a filter module (38) having adaptive frequency and time domain filter settings, depending on the judgement made by said judgement module, for producing said filtered audio signals, said judgement being adapted to detect close speech; and wherein said filter settings are selected by said judgement module such that said filter settings provide for a transparent mode if the judgement module judges that no noise is present, for a first attenuation mode if the judgement module judges that noise without close speech is present, and for a second attenuation mode if the judgement module judges that noise with close speech is present.

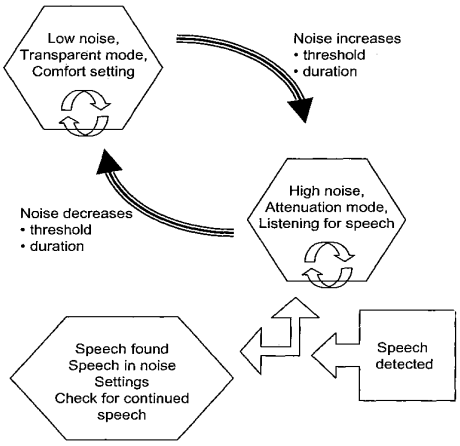


Fig. 2

EP 1 674 061 A1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The present invention relates to a system and a method for providing active hearing protection to a person.

[0002] A large part of the population is exposed to hazardous noise from time to time. This can be at work, whilst traveling, during leisure activities or at home. The exposure can lead to permanent hearing loss, distract people's attention from other hazards or simply cause stress. In order to prevent both accidents and permanent hearing damage, hearing protection devices (HPDs) have been provided in many styles and over many years. It started with the earmuff which is still very relevant and addresses very noisy environments (e.g. airports, construction, shooting) or complex working/communication situations (e.g. fighter pilots). Over the years development of biocompatible soft materials has enabled soft earplugs in different styles and colors as well as recent development of "one fits many" standard semi-soft earplugs in silicon-rubber type materials. For severe situations even the combination of an earmuff and an "in-the-ear" HPD is required to achieve desired attenuation. The physical limitation of hearing protection based on ear worn devices is defined where bone-conduction (body acoustics) becomes dominant at around 40 dB attenuation.

[0003] A common disadvantage of the above mentioned HPD styles is wearing discomfort. In case of the earmuffs, they are large which creates difficulties in combination with other head worn gear and they "close off" the ear too much for most applications. The in-the-ear styles mentioned are devices made to fit "the average" ear in one way or the other. Either the fit is provided by softness of the material which leads to undefined device insertion and undefined attenuation, or the fit is provided by standard shaped structures intended to block off the ear canal. In both cases the flat distribution of the individual shape of the outer ear and the ear canal leads to bad fit, pressure points in the ear and undefined positioning of the device.

[0004] To address this wearing comfort issue, in-the-ear hearing aid technology has been applied making customized ear molds with passive acoustical filter. These are long lasting devices with good wearing comfort. However, this customization process is traditionally a very manual process creating varying results over time, low reproducibility and the quality is very operator skill dependent.

[0005] Customized earplugs are earplugs comprising a hard shell which has an outer surface individually shaped according to the measured inner shape of the user's outer ear and ear canal. Such earplugs presently are primarily used for housing hearing aids. The inner shape of the user's outer ear and ear canal may be measured, for example, by direct laser scanning or by forming an impression. The customized hard shell may be produced by an additive process, such as layer-by-layer laser sintering of a powder material. Customized earplugs are described, for example, in WO 02/071794 A1, US 2003/0133583 A1 or US 6,533,062 B1.

[0006] On the other hand, soft earplugs are widely used, in particular also as hearing protection devices. A soft earplug has an outer surface with a standardized shape and is made of a relatively soft material so that the outer surface of the earplug is capable of adapting its shape to the individual inner shape of the user's outer ear and ear canal.

[0007] It is commonly known to design hearing protection devices as so-called active hearing protection devices wherein each device is provided with an outer microphone for converting ambient sound into input audio signals, a signal processing unit for processing the input audio signals into output audio signals, and an acoustic output transducer, i.e. speaker or receiver, which converts the audio output signals into sound perceivable by the user when wearing the hearing protection device. Thereby the hearing protection device is provided with a communication function enabling the user to perceive ambient sound signals in a controlled manner even when wearing the hearing protection device in a noisy environment. In order to provide for a selective communication function, i.e. in order to enable the user to perceive, for example, speech while suppressing undesired noise, it is known to provide the audio signal processing unit with the capability of assessing the sound picture and adapting filter and gain settings to the noise level dynamically.

[0008] An example of such a system is described in US 4,677,678 which relates to hearing protection devices with an active processing and adaptive circuitry ("Automatic Gain Control" (AGC)), with one AGC taking care of the sound in both ears and with the hearing protection devices being electrically connected for providing a binaural functionality. WO 02/17837 A1 relates to an active hearing protection device which comprises, in addition to the outer microphone, an inner microphone picking up sound at the distal end of the hearing protection device, wherein desired sound is selectively filtered and amplified based on input audio signal analysis.

[0009] US 5,027,410 relates to an adaptive digital hearing aid system which digitally classifies the sound input and selects spectral coefficients of the audio signal processing unit according to the result of the sound input classification in order to reduce background noise levels relative to speech and in order to limit the speaker output for avoiding sound signals exceeding the comfort level of the user.

[0010] US 4,750,207 relates to an analogue hearing aid system with frequency selective audio signal amplification in response to the sound input in order to provide for a background noise suppression feature.

[0011] US 2001/0046304 A1 relates to a headset device for providing external audio signals to the user's ear having a hear-through function for selected audio signals, which is achieved by analyzing the ambient sound for predetermined stored sound signals and by selectively amplifying the ambient sound in this mode over the ex-

Exhibit 1009
Page 1454 of 2198

ternal audio signals to be provided to the user's ear by the headset.

[0012] It is an object of the invention to provide for a hearing protection system which provides for sufficient hearing protection in noisy environments while it, at the same time, provides for a communication function, in particular in order to enable the user to perceive speech, both in high noise environments and in low noise environments. In addition, the hearing protection system should be easy to handle by the user. It is a further object to provide for a corresponding hearing protection method.

[0013] These objects are achieved by a hearing protection system as defined in claim 1 and a method as defined in claim 17.

[0014] The invention is beneficial in that, by providing for a classification of the actual ambient sound scenario and accordingly selecting a transparent mode, a first attenuation mode without close speech being present and a second attenuation mode with close speech being present, an automatic and appropriate adaptation of the total amplification provided by the active part of the hearing protection system according to the present ambient sound scenario is achieved, whereby both safe hearing protection and a communication function in quiet and noisy environments is achieved while handling of the hearing protection system by the user is kept particularly simple.

[0015] Preferred embodiments of the invention are defined in the dependent claims.

[0016] In the following, examples of the invention will be illustrated by reference to the attached drawings.

[0017] Fig. 1 shows a schematic representation of a hearing protection system according to the invention; and

[0018] Fig. 2 shows a schematic representation of the operation modes of the system of Fig. 1.

[0019] Fig. 1 is a schematic representation of a hearing protection system for a person comprising a hearing protection earplug 10 which is to be worn at least partly within a person's right ear canal and a hearing protection earplug 12 which is to be worn at least partly within the person's left ear canal. In a less preferred embodiment, the earplugs 10, 12 may be replaced by hearing protection devices which are to be worn at the person's right and left ear, respectively, such as an earmuff.

[0020] In general, the hearing protection devices according to the invention are adapted to provide for an acoustic attenuation of at least 10 dB averaged over the audible frequency range when worn by the user.

[0021] Each hearing protection earplug 10, 12 comprises a shell 14 which is adapted to be worn at least in part in a user's ear canal, i.e. at least a distal portion of the shell is to be inserted into the outer part of the user's ear canal, in order to provide for an acoustic attenuation of at least 10 dB averaged over the audible frequency range when the earplug is worn by the user in order to protect the user from excessive levels of ambient sound. The earplug may comprise an acoustic filter for adjusting the desired total acoustic attenuation or for adjusting the frequency dependent acoustic attenuation.

[0022] The shell preferably is a hard shell having an elasticity from shore D85 to D65 and preferably is made of polyamide. In order to achieve optimized fit of the shell within the user's outer ear and ear canal, the shell has an outer surface individually shaped according to the measured shape of the user's outer ear and ear canal, i.e. the shell has an individually customized outer shape. The shape of the user's outer ear and ear canal may be determined by direct three-dimensional scanning of the ear canal and the concha or by producing an impression of the ear canal and the concha which subsequently undergoes scanning. The scanning process may be carried out optically, preferably by laser scanning.

[0023] The digital data obtained by the scanning process is then used to create the hard shell by an additive or incremental layer-by-layer build up process. Such processes are also known as "rapid prototyping". A preferred additive build-up process is a layer-by-layer laser sintering process of powder material, preferably polyamide powder. Such processes are also known as "selective laser sintering" (SLS). The basic principle therein is the repeated deposition of a thin layer of material on a surface, with the desired sectional shape then being stabilized, i.e. hardened, by laser action. Other preferred additive layer-by-layer build-up processes are laser stereo-lithography or photo-polymerization. An overview regarding additive layer-by-layer build-up processes for producing customized shells for hearing aids can be found, for example, in US 2003/0133583 A1 or US 6,533,062 B1.

[0024] In general, the invention relates to active hearing protection devices, i.e. hearing protection devices which do not only provide for an acoustic attenuation of ambient sound in order to protect the user's hearing, but in addition comprise means for electro-acoustically bypassing this acoustic attenuation function in order to provide the user with a communication function, e.g. in order to enable the user to perceive speech signals, even when wearing the hearing protection devices in a noisy environment.

[0025] To this end, each earplug 10, 12 is provided with an active unit 16 which comprises a microphone 18 for converting ambient sound impinging on the proximal, i.e. outer, end of the earplug 10, 12 into input audio signals, an audio signal processing unit 20 for processing the input audio signals provided by the microphone 18 into filtered audio signals and an acoustic output transducer, i.e. speaker, 22 for converting the filtered audio signals provided by the audio signal processing unit 20 into sound which is perceivable by the user's ear via a sound output opening 24 provided at the distal end of each earplug 10, 12. In practice, the audio signal processing unit 20 may be realized by a digital signal processor which includes a memory 26 for storing programs and data. Consequently, the active unit 16 in addition comprises an analogue-to-digital converter between the mi-

Exhibit 1009
Page 1455 of 2198

crophone 18 and the audio signal processing unit 20 and a digital-to-analogue converter between the audio signal processing unit 20 and the speaker 22 (these converters are not shown in Fig. 1). Further, the audio signal processing unit 20 is provided with an interface 28 for temporarily connecting with a remote unit 30 via a wireless or wired data connection 32 for uploading programs and data into the audio signal processing unit 20, i.e. the program/data memory 26. Such data may include user specific audio data, such as an existing hearing loss, in order to individually configure the audio signal processing unit for optimizing performance. The active unit 16 further comprises a battery which is not shown in Fig. 1.

[0026] According to the invention, the audio signal processing unit 20 is designed to enable, on the one hand, selective perception of desired sound signals, such as speech signals, both in low noise and high noise environments while, on the other hand, protection of the user's hearing from excessive ambient sound levels is ensured. To this end, the audio signal processing unit 20 comprises an analyzer module 34 for determining the intensity of the input audio signals provided by the microphone 18 separately for a plurality of spectral classes, a judgement module 36 for judging, depending on the spectral intensities of the input audio signals determined by the analyzer module 34, which one of a plurality of predetermined criteria is presently fulfilled by the input audio signals, and a filter module 38 having adaptive frequency and time domain filter settings for producing the filtered audio signals for the speaker 22, with the filter settings of the filter module 38 depending on the judgement made by the judgement module 36. The judgement module 36 is capable of detecting close speech, i.e. speech signals originating from a person which is presently located close to the user's ear. "Close" in this respect is to be understood as a distance of less than 3 m, preferably less than 1.5 m.

[0027] The filter settings of the filter module 38 are selected by the judgement module 36 such that the filter settings provide at least for the following modes: a transparent mode if the judgement module 36 judges that no noise is present (for example, the noise levels are below predefined threshold values); a first attenuation mode if the judgement module 36 judges that noise without close speech is present (for example, the noise levels exceed the predefined threshold values); and a second attenuation mode if the judgement module 36 judges that noise with close speech is present (i.e. in addition to noise also close speech signals are detected by the judgement module 36).

[0028] The criteria used by the judgement module 36 for judging in which of these modes the filter module 38 is to be operated may be individually defined according to, for example, previous measurements of the user's hearing capability.

[0029] Preferably, the judgement module 36 is designed to determine, in order to make the filter settings mode judgement, whether for each spectral class the intensity has fallen within a given range for a given minimum time interval, with the ranges and time intervals depending on the respective filter settings mode. Further, the judgement module 36 may be designed to analyze, in order to make the filter settings mode judgement, at least one audio signal parameter selected from the group consisting of attack time, release time, signal dynamics, spectral content, spectral dynamics, and signal history. Consequently, in preferred embodiments of the invention the judgement module 36 will not only consider the present levels in the different spectral classes provided by the analyzer module 34 but in addition will also consider the variation of these levels with time in order to achieve an optimized judgement of the present audio situation.

[0030] The judgement module 36 and the filter module 38 may be designed such that in the transparent mode the frequency averaged amplification provided by the active unit 16 has a first value, in the first attenuation mode the frequency averaged amplification provided by the active unit 16 has a second value which is lower than the first value, and in the second attenuation mode the frequency averaged amplification provided by the active unit 16 has a third value lower than the first value and speech frequencies, i.e. frequencies preferably between 500 Hz and 5 kHz, are selectively amplified over other frequencies.

[0031] Preferably not only the judgement criteria of the judgement module 36 but also the filter settings of the filter module 38 in each mode are individually defined, for example, taking into account the user's hearing capability including hearing loss, in order to individually optimize the performance of the active unit 16.

[0032] In the transparent mode the first amplification value may be selected such that the total amplification, i.e. the amplification including the acoustic attenuation provided by the shell 14, is between -10 dB and 0 dB. Preferably, in the first attenuation mode the second amplification value is selected such that the total amplification, i.e. including the acoustic attenuation provided by the shell 14, is less than -10 dB.

[0033] In order to allow the judgement module 36 to detect close speech, each earplug 10, 12 is provided with a second microphone 19 which is located spaced apart from the microphone 18 in order to enable the judgement module 36 to assign angular directions to the input audio signals by analyzing phases of the input audio signals, i.e. by determining the phase difference of the input audio signal provided by the microphone 18 and the input audio signal provided by the microphone 19. Further, the judgement module 36 analyzes the power spectrum of the input audio signals for these angular directions and determines, by comparison with a predefined speech judgement condition, whether angularly directed speech signals are present in the input audio signals. Preferably, such speech judgement condition includes intensity limits for a plurality of frequency bands and a threshold value for the inhomogeneity of the angular distribution of

Exhibit 1009
Page 1456 of 2198

speech signals.

[0034] In addition or alternatively to the above described approach of determining the angular distribution of speech signals the judgement module 36 may be adapted to determine, by time, frequency and intensity dependent analysis of the input audio signals for determining the presence of spatially reflected signals and by comparison with a predefined speech judgement condition, whether reflected signals within a pre-defined speech frequency range, preferably 500 Hz to 5 kHz, and with a delay time above a pre-defined limit are present in the input audio signals, with the speech judgement condition including values for this speech frequency range and for this delay time limit.

[0035] Examples for methods and devices for achieving ambient sound classification, and in particular also the distinction between undesired noise signals and desired sound signals, are described, for example, in US 5,027,410, wherein also selective filter setting of the audio signal processing unit according to the result of the sound classification is described.

[0036] Examples and examples of methods and devices for localizing speech sources in background noise are described, for example, in EP 1 320 281 A1, WO 00/68703, EP 1 005 783 B1, US 6,522,756 B1, WO 00/33634, WO 01/60112 and EP 0 802 699 A2. In these documents it is also described how detected speech signals can be selectively amplified over background noise.

[0037] Fig. 2 is a schematic representation of an example for the operation of an active hearing protection system according to the invention. It is assumed that the active unit 16 starts in the transparent mode, i.e. the system is worn in a sound environment with low noise levels. In this mode, the main purpose of the active unit 16 is to more or less bypass the (mechanical) acoustic attenuation provided by the earplugs 10, 12 when having been inserted into the user's ear canals in order to enable the user to perceive the ambient sound, in particular speech signals, in a close-to-natural manner. In this mode, the settings of the filter module 38 are selected such that maximum hearing comfort is achieved ("comfort setting").

[0038] Once the noise increases in such a manner that the judgement module 36 judges that the "high noise" criteria are fulfilled, e.g. that threshold levels in the noise frequency bands are exceeded for a given minimum duration, the settings of the filter module 38 are changed into the first attenuation mode in order to protect the user from excessive noise levels, i.e. the amplification provided by the active unit 16 is significantly reduced or even completely eliminated in order to provide for a maximum acoustic attenuation by the earplugs 10, 12. Once the judgement module 36 finds that the noise has decreased in such a manner that the "high noise" criteria are no longer fulfilled, the settings of the filter module 38 will be changed to those for the transparent mode again.

[0039] However, in the first attenuation mode the judgement module 36 continuously searches for close speech signals. If such close speech signals are found, the system will enter the second attenuation mode and the settings of the filter module 38 will be changed in such a manner that, while noise signals are still suppressed as in the first attenuation mode, the detected speech signals are selectively amplified over the noise signals by selectively amplifying only speech frequencies and/or selectively amplifying audio signals only from angular directions in which a speech source has been detected. If no close speech signals are detected any longer, the system will return to the first attenuation mode.

[0040] In order to enhance comfort of the hearing protection system, soft fading between the various modes may be applied. In general, the change between the various modes of operation occurs fully automatically without any manipulation by the user being necessary.

[0041] The evaluation of the noise situation should occur essentially permanently, for example, at least once or twice per second, in order to ensure a sufficiently fast response to changes in the noise/communication situation.

[0042] In general, additional operation modes are possible. For example, the judgement module 36 may be adapted to detect acoustic alarm signals in order to change, once an acoustic alarm signal has been detected, the settings of the filter module 38 in such a manner that acoustic perception of this alarm signal by the user is enabled, for example, by selective amplification of the alarm signal frequencies over background noise frequencies. Such a mode could be present as a third attenuation mode.

[0043] Similar to the second attenuation mode, the active unit 16 would return to the first attenuation mode once the alarm signal is no longer detected by the judgement module 36.

[0044] Preferably, the active unit 16 of each of the earplugs 10, 12 is provided with one of the signal processing units 20.

[0045] In general, the hearing protection system may be designed as a binaural system (see elements shown in dashed lines in Fig. 1). In this case, each earplug 10, 12 may be provided with a binaural communication unit 50 which communicates bidirectionally with its counterpart in the other earplug via a communication link 52 which may be integrated into means for mechanically connecting the earplugs 10, 12, such as a cord 54, for preventing loss of the earplugs 10, 12. In addition, each binaural communication unit 50 communicates with the respective audio signal processing unit 20, in particular the respective judgement unit 36.

[0046] In such a system the audio signal processing units 20, and in particular the judgement units 36 of the right earplug 10 and the left earplug 12, are enabled to communicate with each other in order to improve the accuracy of the judgement of the speech/noise situations and the accuracy of the detection of close speech directions.

Exhibit 1009
Page 1457 of 2198

Claims

1. A hearing protection system comprising a first hearing protection device (10) to be worn at a person's right outer ear and/or at least in part in a person's right ear canal and a second hearing protection device (12) to be worn at said person's left outer ear and/or at least in part in said person's left ear canal, each hearing protection device comprising an active unit (16) comprising an acoustic input transducer (18, 19) for converting ambient sound into input audio signals, and an acoustic output transducer (22) for transforming filtered audio signals into sound perceivable by said user, wherein at least one of said active units comprises an audio signal processing unit (20) for processing said input audio signals into said filtered audio signals;
wherein each hearing protection device is adapted to provide for an acoustic attenuation of at least 10 dB averaged over the audible frequency range when worn by the person;
wherein said audio signal processing unit comprises an analyzer module (34) for determining the intensity of said input audio signals separately for a plurality of spectral classes, a judgement module (36) for judging, depending on said determined spectral intensities of said input digital audio signals, which one of a plurality of predetermined criteria is presently fulfilled by the input audio signals, and a filter module (38) having adaptive frequency and time domain filter settings, depending on the judgement made by said judgement module, for producing said filtered audio signals, said judgement being adapted to detect close speech; and
wherein said filter settings are selected by said judgement module such that said filter settings provide for a transparent mode if the judgement module judges that no noise is present, for a first attenuation mode if the judgement module judges that noise without close speech is present, and for a second attenuation mode if the judgement module judges that noise with close speech is present.

2. The hearing protection system of claim 1, wherein in said transparent mode the frequency averaged amplification provided by said active unit (16) has a first value, in said first attenuation mode the frequency averaged amplification provided by said active unit has a second value lower than the first value, and in said second attenuation mode the frequency averaged amplification provided by said active unit has a third value lower than the first value and speech frequencies are selectively amplified.

3. The hearing protection system of one of claims 1 or 2, wherein each active unit (16) comprises one of said audio signal processing units (20).

4. The hearing protection system of one of the preceding claims, wherein said filter settings in each mode are individually defined.

5. The hearing protection system of one of the preceding claims, wherein said judgement module (36), in order to make said filter settings mode judgement, is adapted to determine whether for each spectral class the intensity has fallen within a given range for a given minimum time interval, said ranges and time intervals depending on the respective filter settings mode.

6. The hearing protection system of one of the preceding claims, wherein said judgement module (36) is adapted, in order to make said filter settings mode judgement, to analyze at least one audio signal parameter selected from the group consisting of attack time, release time, signal dynamics, spectral content, spectral dynamics, and signal history.

7. The hearing protection system of claim 2, wherein said first amplification value is selected such that the total amplification, including said acoustic attenuation, provided by each hearing protection device (10, 12) is between -10 dB and 0 dB.

8. The hearing protection system of one of claims 2 or 7, wherein said second amplification value is selected such that the total amplification, including said acoustic attenuation, provided by each hearing protection device (10, 12) is less than- 10 dB.

9. The hearing protection system of claim 2, wherein said speech frequencies are between 500 Hz and 5 kHz

10. The hearing protection system of one of the preceding claims, wherein said hearing protection devices are designed as earplugs (10, 12) having a shell (14).

11. The hearing protection system of one of the preceding claims, wherein the outer surface of said shell (14) is individually shaped according to the measured inner shape of the person's outer ear and ear canal.

12. The hearing protection system of claim 11, wherein said shell (14) is a hard shell with an elasticity of from shore D85 to shore D65.

13. The hearing protection system of one of the preceding claims, wherein said judgement module (36) is adapted to assign, by analyzing phases of said input audio signals, angular directions to said input audio signals and to judge, by analyzing the power spectrum of said audio input signals for said angular directions and by comparison with a speech judge-

Exhibit 1009
Page 1458 of 2198

ment condition, whether angularly directed speech signals are present in said audio input signals.

14. The hearing protection system of claim 13, wherein said speech judgement condition includes intensity limits for a plurality of frequency bands and a threshold for the inhomogeneity of the angular distribution of speech signals, and wherein said judgement module (36) judges that close speech is present if said intensity limits and said threshold value are exceeded.

15. The hearing protection system of one of claims 13 or 14, wherein each active unit (16) includes at least one additional acoustic input transducer (19).

16. The hearing protection system of one of the preceding claims, wherein said judgement module (36) is adapted to determine, by time, frequency and intensity dependent analysis of said input audio signals for determining the presence of spatially reflected signals and by comparison with a speech judgement condition, whether reflected signals within a defined speech frequency range and with a delay time above a defined limit are present in said audio input signals, said speech judgement condition including values for said speech frequency range and for said delay time limit.

17. A method for providing active hearing protection to a person, comprising:

providing a first hearing protection device (10) at said person's right outer ear and/or at least in part in said person's right ear canal and a second hearing protection device (12) at said person's left outer ear and/or at least in part in said person's left ear canal in such a manner that an acoustic attenuation of at least 10 dB averaged over the audible frequency range is achieved;
converting, in each hearing protection device, ambient sounds into input audio signals and converting filtered audio signals into sound perceivable by said person, wherein said input audio signals are processed into said filtered audio signals in at least one of said hearing protection devices;
determining, in at least one of said hearing protection devices, the intensity of said input audio signals separately for a plurality of spectral classes; judging, depending on said determined spectral intensities of said input audio signals, which one of a plurality of predetermined criteria is presently fulfilled by the input audio signals; and specifically selecting, depending on the judgement made, frequency and time domain filter settings for producing said filtered audio signals;
wherein said filter settings are selected such that said filter settings provide for a transparent mode if the judgement made is that no noise is present, for a first attenuation mode if the judgement made is that noise without close speech is present, and for a second attenuation mode if the judgement made is that noise with close speech is present.

18. The method of claim 17, wherein in said transparent mode the frequency averaged amplification has a first value, in said first attenuation mode the frequency averaged amplification has a second value lower than the first value, and in said second attenuation mode the frequency averaged amplification has a third value lower than the first value and speech frequencies are selectively amplified.

19. The method of one of claims 17 or 18, wherein said input audio signals are processed into said filtered audio signals in both of said hearing protection devices (10, 12).

20. The method of one of claims 17 to 19, wherein said filter settings in each mode are individually defined.

21. The method of one of claims 17 to 20, further comprising: determining, in order to make said filter settings mode judgement, whether for each spectral class the intensity has fallen within a given range for a given minimum time interval, said ranges and time intervals depending on the respective filter settings mode.

22. The method of one of claims 17 to 21, further comprising: analyzing, in order to make said filter settings mode judgement, at least one audio signal parameter selected from the group consisting of attack time, release time, signal dynamics, spectral content, spectral dynamics, and signal history.

23. The method of claim 18, wherein said first amplification value is selected such that the total amplification, including said acoustic attenuation, provided by each hearing protection device (10, 12) is between -10 dB and 0 dB.

24. The method of one of claims 18 and 23, wherein said second amplification value is selected such that the total amplification, including said acoustic attenuation, provided by each hearing protection device is less than -10 dB.

25. The method of claim 18, wherein said speech frequencies are between 500 Hz and 5 kHz.

26. The method of one of claims 17 to 25, further comprising, in order to detect close speech, assigning, by analyzing phases of said input audio signals, an-

Exhibit 1009
Page 1459 of 2198

gular directions to said input audio signals and judging, by analyzing the power spectrum of said audio input signals for said angular directions and by comparison with a speech judgement condition, whether angularly directed speech signals are present in said audio input signals.

27. The method of claim 26, wherein said speech judgement condition includes intensity limits for a plurality of frequency bands and a threshold for the inhomogeneity of the angular distribution of speech signals, and wherein it is judged that close speech is present if said intensity limits and said threshold value are exceeded.

28. The method of one of claims 17 to 27, further comprising, in order to detect close speech, determining, by frequency dependent analysis of said input audio signals for determining the presence of spatially reflected signals and by comparison with a speech judgement condition, whether reflected signals within a defined speech frequency range and with a delay time above a defined limit are present in said audio input signals, said speech judgement condition including values for said speech frequency range and for said delay time limit.

Exhibit 1009
Page 1460 of 2198

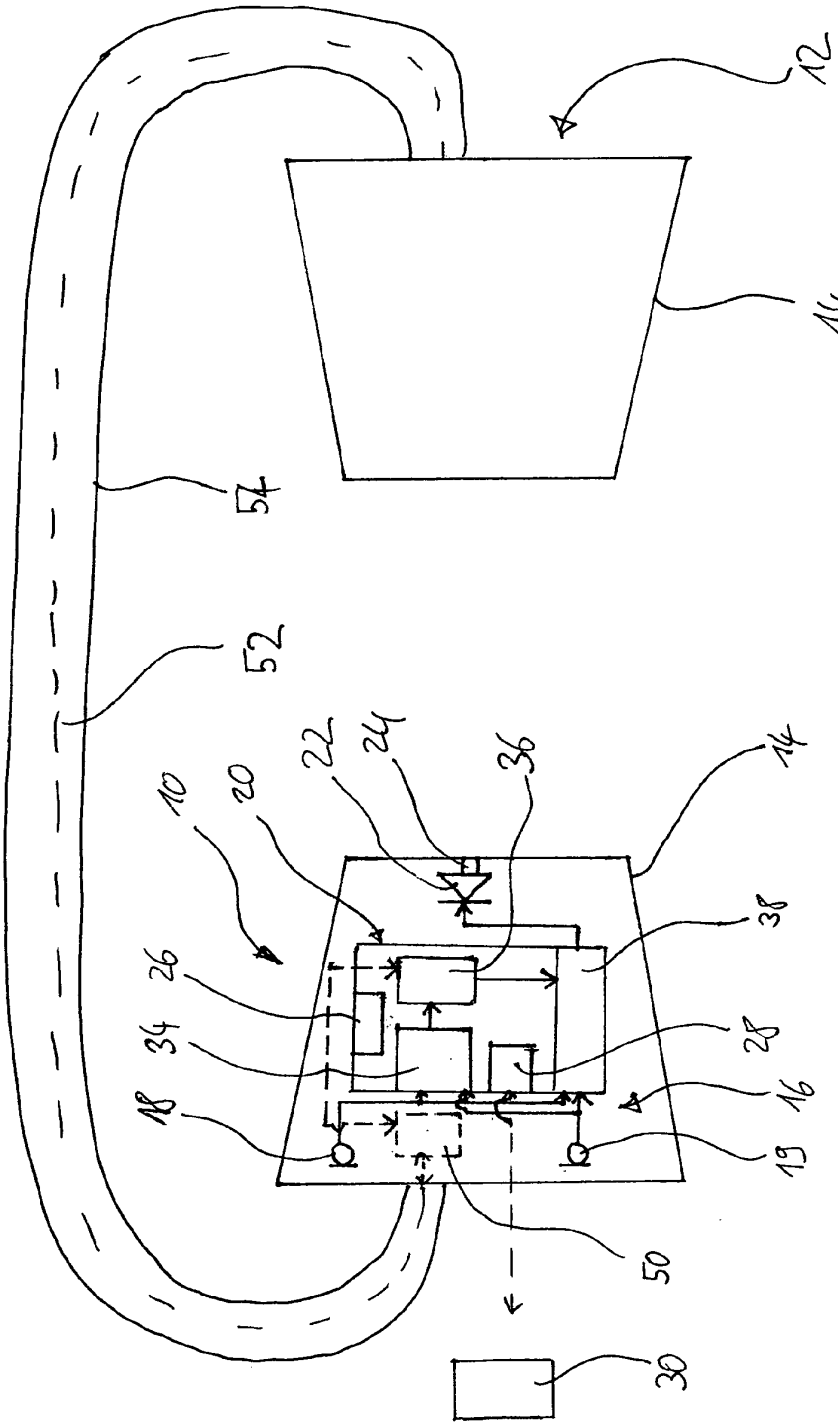


Fig. 1

Exhibit 1009
Page 1461 of 2198

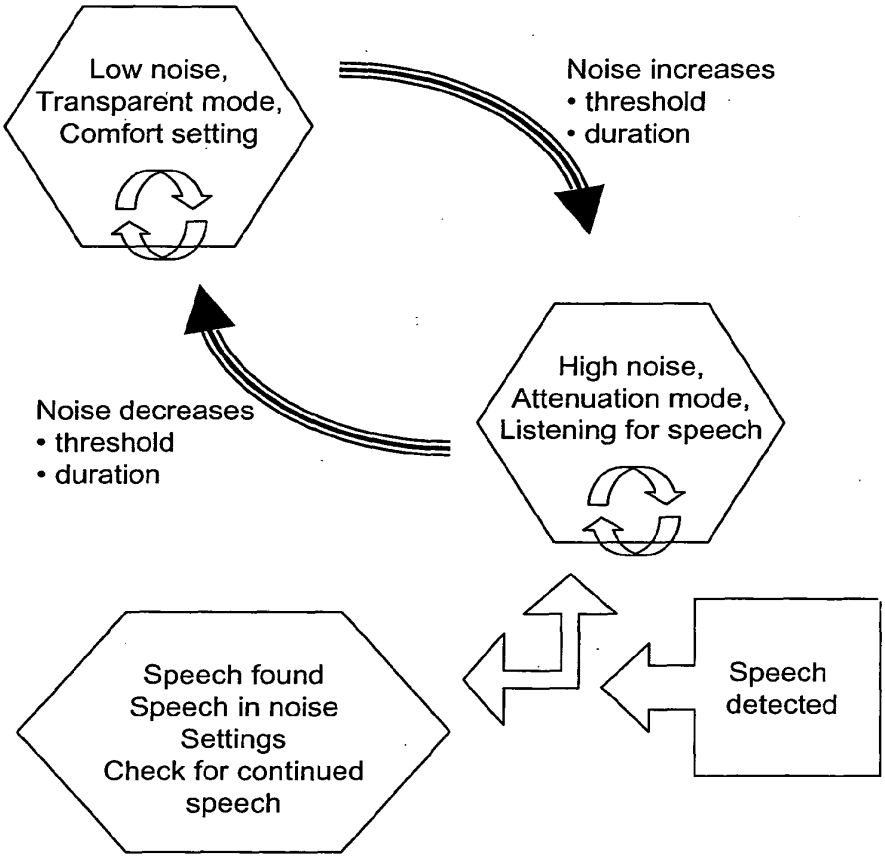


Fig. 2

Exhibit 1009
Page 1462 of 2198

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 04 03 0646

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X Y	US 2002/080979 A1 (BRIMHALL OWEN D ET AL) 27 June 2002 (2002-06-27) * paragraphs [0004], [0016], [0020], [0050] - [0057]; claims 1-3,8-11,15,16 *	1-5, 9-12,15, 18-21,25 6,22	A61F11/14 H04R1/10
Y A	DE 101 14 015 A1 (SIEMENS AUDIOLOGISCHE TECHNIK GMBH) 24 October 2002 (2002-10-24) * the whole document *	6,22 1-5, 7-21, 23-28	
A	US 2003/165246 A1 (KVALOY OLAV ET AL) 4 September 2003 (2003-09-04) * paragraphs [0025], [0036] - [0045], [0048] - [0061], [0065] - [0078], [0081] - [0087]; figures 1-8 *	1-28	
A	EP 1 448 014 A (GN RESOUND A/S) 18 August 2004 (2004-08-18) * column 2, line 41 - line 43; claim 11 *	12	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
A	US 5 631 965 A (CHANG ET AL) 20 May 1997 (1997-05-20) * column 3, line 13 - column 4, line 37 *	1-28	A61F H04R G10K
A	DE 199 35 013 C1 (SIEMENS AUDIOLOGISCHE TECHNIK GMBH) 30 November 2000 (2000-11-30) * figure 1 *	15	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	16 June 2005	Timms, O

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 04 03 0646

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

16-06-2005

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002080979 A1	27-06-2002	NONE	
DE 10114015 A1	24-10-2002	NONE	
US 2003165246 A1	04-09-2003	WO 03073790 A1	04-09-2003
		AU 2002237590 A1	09-09-2003
		CA 2474792 A1	04-09-2003
		EP 1479265 A1	24-11-2004
EP 1448014 A	18-08-2004	EP 1448014 A1	18-08-2004
		US 2005002539 A1	06-01-2005
US 5631965 A	20-05-1997	AU 4298493 A	24-01-1994
		WO 9400089 A1	06-01-1994
DE 19935013 C1	30-11-2000	CH 694481 A5	31-01-2005
		DK 200001130 A	27-01-2001

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : H03B 29/00 — A1

(11) International Publication Number: WO 93/26084

(43) International Publication Date: 23 December 1993 (23.12.93)

(21) International Application Number: PCT/US92/04567

(22) International Filing Date: 5 June 1992 (05.06.92)

(71) Applicant *(for all designated States except US):* NOISE CANCELLATION TECHNOLOGIES, INC. [US/US]; Suite 101, 1015 West Nursery Road, Linthicum, MD 21090 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* CLAYBAUGH, David [US/US]; 20522 Staffordshire Drive, Germantwon, MD 20874 (US). DENENBERG, Jeffrey [US/US]; 345 Putting Green Road, Trumbull, CT 06611 (US). BUSCH, Ralph [US/US]; 316 Boyd Avenue, Takoma Park, MD 20912 (US). HOHMAN, John [US/US]; 1545 Putty Hill Road, Towson, MD 21204 (US).

(74) Agent: HINEY, James, W.; Noise Cancellation Technologies, Inc., Suite 101, 1015 West Nursery Road, Linthicum, MD 21090 (US).

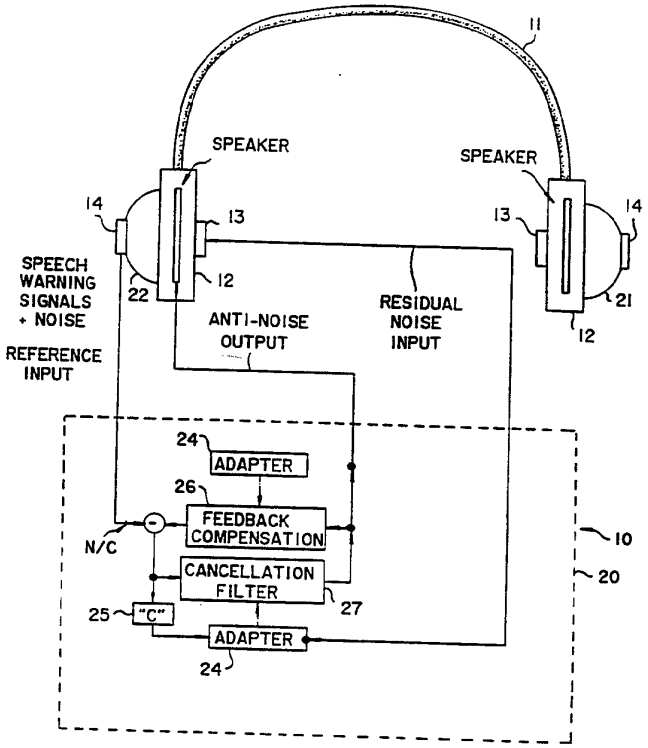
(81) Designated States: CA, JP, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE).

Published
With international search report.

(54) Title: ACTIVE PLUS SELECTIVE HEADSET

(57) Abstract

An active plus selective headset system for provision of active attenuation of broadband noise as well as speech filtering comprising a headset (11) with reference microphones (14), residual microphones (13) and speakers (12) on each of a pair of open backed muffs (21, 22) and a controller (20).



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	Côte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML	Mali	VN	Viet Nam
FI	Finland	MN	Mongolia		

ACTIVE PLUS SELECTIVE HEADSET

This invention relates to a headset for actively canceling unwanted noise while selectively allowing necessary speech to reach the user's ear.

In the past, attempts to combine the two protections i.e., high and low frequency attenuation, has resulted in not only the noise being attenuated but also the speech that the wearer needs to hear. Some systems met only limited success with fixed or "near-stationary" noise but not with the other noise of either (a) varying spectral characteristics or (b) brief duration noises with "spikes". Examples of such a system is found in U.S. Patent No. 4,025,721, to Graupe et al and U.S. Patent 4,185,168 to Graupe et al. Other systems like that found in U.S. Patent No. 4,455,675 to Bose actively attenuate all sounds at low frequencies and passively attenuate all high frequency sounds. These sounds include speech and warning signals that want to be heard by the person wearing the headset.

The instant invention solves the problem now existant, that of total attenuation of the noise and speech, by providing a solution of an active headset that can employ any of several selective algorithms such as those disclosed in U.S. Patent No. 4,654,871 to Chaplin, hereby incorporated by reference herein. Alternatively, it can employ the algorithm disclosed in U.S. Patent No. 5,105,377 to Ziegler which is also incorporated herein by reference. In addition it can employ other algorithms such as that disclosed in the application of Ziegler in U.S. Patent Application No. 07/421 759 which is hereby incorporated by reference herein.

In applications for noise canceling headsets, particularly in industrial environments, attenuation of low frequency noise as well as noise that covers the speech band (300 to 3300 Hz) passive hearing protection works extremely well at higher frequencies (typically above 1000 Hz) whereas active noise cancellation has been shown to achieve similar levels of protection at lower frequencies (50 to 1000 Hz). Passive, however, also attenuates speech and warning signals and the protectors are uncomfortable to wear. This invention provides a solution that simultaneously

provides the protection offered by a passive headset in a lightweight open back headset while using active adaptive feed forward control algorithms that attenuate all sounds in the 20 to 3300 Hz frequency band. Additionally, adaptive speech filtering or in-wire control technology separates speech from noise and passes the speech to the user.

Accordingly it is an object of this invention to provide an active noise canceling headset with selectivity.

Another object of this invention is the provision of an open back muff headset with selective filtering.

These and other objects of this invention where reference is had to the accompanying drawings in which

Fig. 1 shows a typical active/passive headset system incorporating the instant invention.

Fig. 2 shows an active plus selective headset system with an open back muff that incorporates active control and adaptive speech filtering to allow speech to pass with the "anti-noise" signal.

Fig. 3 shows a more detailed description of the active control system of Figure 2.

Fig. 4 shows a more detailed description of the adaptive speech filtering technique to be used in this headset design.

In Fig. 1 there is shown an active/passive closed back headset system 10. It consists of a typical passive headset 11, loudspeakers 12 that drive the anti-noise and residual microphones 13 to sense any remaining noise near the ear and reference microphones 14 to send advanced information for feed forward approaches and a system controller 20 which synthesizes the anti-noise signal.

The headset shown has closed backs 21, 22 for passive attenuation without the speakers, microphones and system controller, this headset would be a typical passive hearing protector.

The system is designed to use various algorithms such as that of Ziegler in U.S. Patent 5,105,377 or an adaptive feed forward approach. Both these algorithms use a

reference signal as inputs. The digital virtual earth (DVE) algorithm develops a reference signal by subtracting an equalized version of its own anti-noise signal from the residual signal. The adaptive feed forward uses the reference microphone as its input and is very effective on complicated noise environments that are broadband and random in character. The Least Means Square (LMS) adapter 24 shown in Fig. 1 are Filtered - X versions which have inherent compensation for the effects of the feedback delays around the loop. Box "C" at 25 is the impulse response of active cancellation system.

Feedback compensator 26 and cancellation filter 27 complete the component portions of the controller.

DVE is highly effective to use in simple noise environments having only a few harmonics even where the noise varies tremendously. It has also been demonstrated to be very effective doing broadband cancellation at low frequencies (50 - 700 Hz).

Speakers 12 of the headset are large enough to be capable of producing anti-noise at the same level as the noise to be canceled. They have little or no distortion and have a minimum of input-to-output delay as any delay in the feedback loop slows down the system adaptation rate.

Residual microphones 13 are typically small electret microphones mounted on the speaker frame near the ear. It must faithfully reproduce the sound that remains at the ear after cancellation so that the controller can make further adjustments to the anti-noise signal.

Reference microphones 14 are small electret microphones attached to the outside of the headset at a distance from the ear canal. This reference microphone is used to provide advanced information about the noise. The higher the frequency of the noise the more advanced information is needed to effectively cancel the noise.

Fig. 2 shows an active plus selective headset system 50 with headset 51 having open backed muff positions 52, reference microphones 53, speakers 54 and residual microphones 55. An earplug (not shown) may be substituted for the open backed muff.

The active /passive system 10 previously described can be configured to actively attenuate all sounds in the frequency band from 20 to 3300 Hz without the need for a passive muff or earplug. The approach uses an adaptive feed forward control algorithm to actively attenuate the damaging noise in this band. In order to accomplish this it is necessary to minimize the delays of the digital signal processing system, which include delays introduced by the anti-aliasing and reconstruction filters shown in Figure 3 and the acoustic delay of the speaker and residual microphone physical system, in order to effectively attenuate noise at the higher frequencies.

The controller 60 has adapters 61,62, feedback compensation 63, cancellation filter 64 and adaptive speech filter 65. Controller 60 uses a parallel adaptive speech filtering technique to pass speech to the user. Adaptive speech filtering techniques can be employed to work with the particular noisy environment. The active controller attenuates noise in the band of interest and allows speech and warning signals to pass via the adaptive speech filtering path which incorporates a warning signal filter as shown in Fig. 2. It is similar to the active/passive system except for the open backed headset design and the addition of a parallel adaptive speech filtering path and warning signal filter path as integral parts of the controller. The input to the speech filter and controller are the upstream reference microphones 53.

This reference microphone contains noise and speech. The speech is filtered from the noise and passed with the "anti-noise" generated from the adaptive feed forward controller and sent to the headset loud speaker. Both speech and warning signals, which are typically above the speech band and of known frequencies, will be heard by the user of the lightweight and open back headset.

With reference to Figure 2, the "anti-noise" and speech output signals are mixed and input to the speakers. This combined signal output sample, u_k, is given by

$$\underline{x}_k = \underline{r}_k - \underline{z}_k$$

$$y_k = \underline{A}_k \underline{s}_k$$

$$u_k = w_k + y_k$$

u_k is the output speech and anti-noise value

where $\underline{r}_k$ is a vector of the most recent examples of the residual signal

$\underline{z}_k$ is a vector of the output of the speech filter after it passed through the impulse response $\underline{C}_k$

$\underline{A}_k$ is a vector of cancellation filter coefficients

y_k is the output anti-noise value

w_k is the output speech value.

$\underline{s}_k$ is the vector of compensated inputs.

Inputs to the controller and speech filter are the reference signal, v_k, and residual signal r_k that are picked up via the reference sensor and residual sensor respectively. The adaptive feedforward controller generates an "anti-noise", y_k, and the adaptive speech filter generates a clean speech signal, w_k, that are mixed to form the output signal u_k which is sent to the speakers. Each ear piece operates independently with separate reference and residual sensors and actuator.

It is essential that the output of the speech filter, w_k, be filtered through the system in pulse response, $\underline{C}_k$, and subtracted from the residual input, r_k, so as not to interfere with the operation of the adaptive feedforward controller. Otherwise, the controller will attempt to adapt to and cancel the speech signal that is output to the speaker.

Several techniques can be used to minimize the delays of the system. First, passive material can effectively act as a low pass filter for the input reference and residual sensors. This would eliminate the need for anti-aliasing filters and thus the delays introduced by these filters would be eliminated. This technique has been shown to be quite effective in the active control of noise in ducts using the adaptive feedforward controller.

Another technique removes neither the anti-aliasing filters nor the reconstruction filters but essentially by-passes the delays introduced by these filters by inserting an analog zero'th order tap. This is achieved by placing an amplifier between

the output of the incoming gain control and the output of the reconstruction filters shown in Figure 3.

A final technique, which will be even more effective as the speed of microprocessor technology increases, is to sample at a rate of 40 kHz or greater, this eliminates the need for anti-aliasing and reconstruction filters because the cut off frequency of 20 kHz is at the limit of the loudspeaker response.

CLAIMS

1. An active plus selective headset system which provides active broadband attenuation of noise as well as speech filtering, said system comprising

a headset means adapted to be worn by a user,

reference sensing means on said headset means adapted to sense speech, noise and warning signals,

speaker means on said headset means adapted to convey speech and warning signals to the user, and

controller means adapted to provide active broadband attenuation of noise as well as speech filtering so that only the warning signals and speech reach the ear.

2. A system as in claim 1 wherein said headset means includes residual microphone means located on said headset means and including a feedback means connected to said speaker means.

3. A system as in claim 2 wherein said feedback means includes an adaptive speech filter means and a cancellation filter means.

4. A system as in claim 2 wherein said controller means is run by an adaptive feedforward algorithm.

5. A system as in claim 2 wherein said headset means includes an open backed muff.

6. A system as in claim 2 wherein said reference sensing means is an electret microphone.

7. A system as in claim 6 wherein said headset system comprises a pair of open backed muffs with the reference sensing means located on the outside of said muffs.

8. A system as in claim 1 wherein said adaptive speech filtering means adapted to filter speech from noise.

9. A system as in claim 1 wherein said adaptive speech filtering output is mixed with said controller output and resulting signal passed to speaker means.

10. A system as in claim 1 wherein said adaptive speech filtering output is filtered by the impulse response of the active cancellation system and subtracted from the residual microphone signal so as not to interfere with the operation of said controller.

11. A system as in claim 1 wherein said reference sensing means is external of speaker means on said headset means and said controller is adapted to employ a feedforward method of noise cancellation.

12. A system as in claim 1 wherein said reference sensing means is internal of said speaker means on said headset means and said controller is adapted to employ a feedback method of noise cancellation.

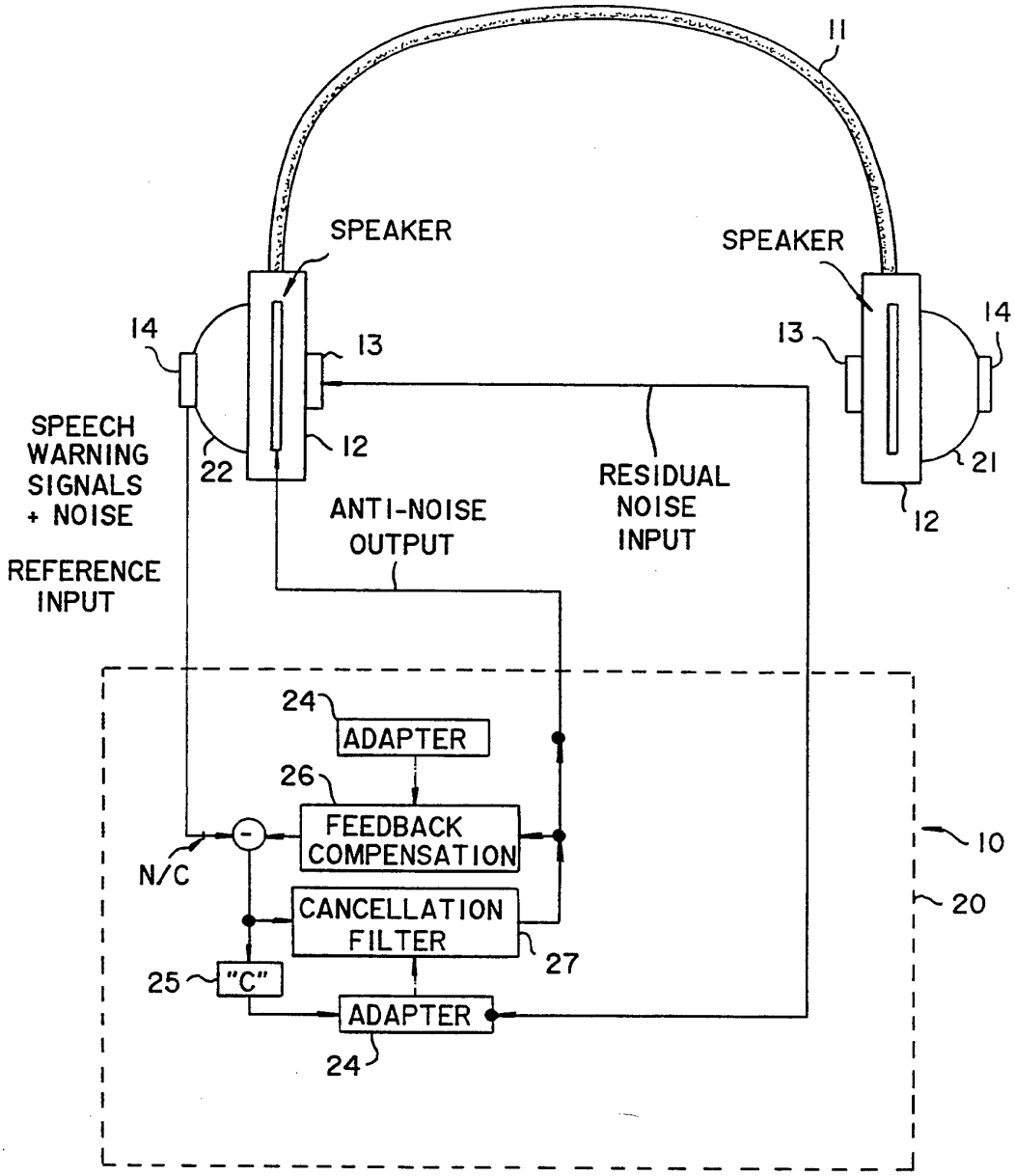

11
SPEAKER
SPEAKER
14
13
13
14
12
22
12
21
SPEECH
WARNING
SIGNALS
+ NOISE
REFERENCE
INPUT
ANTI-NOISE
OUTPUT
RESIDUAL
NOISE
INPUT
24
ADAPTER
26
FEEDBACK
COMPENSATION
N/C
CANCELLATION
FILTER
27
25
"C"
ADAPTER
24
10
20

FIG.1

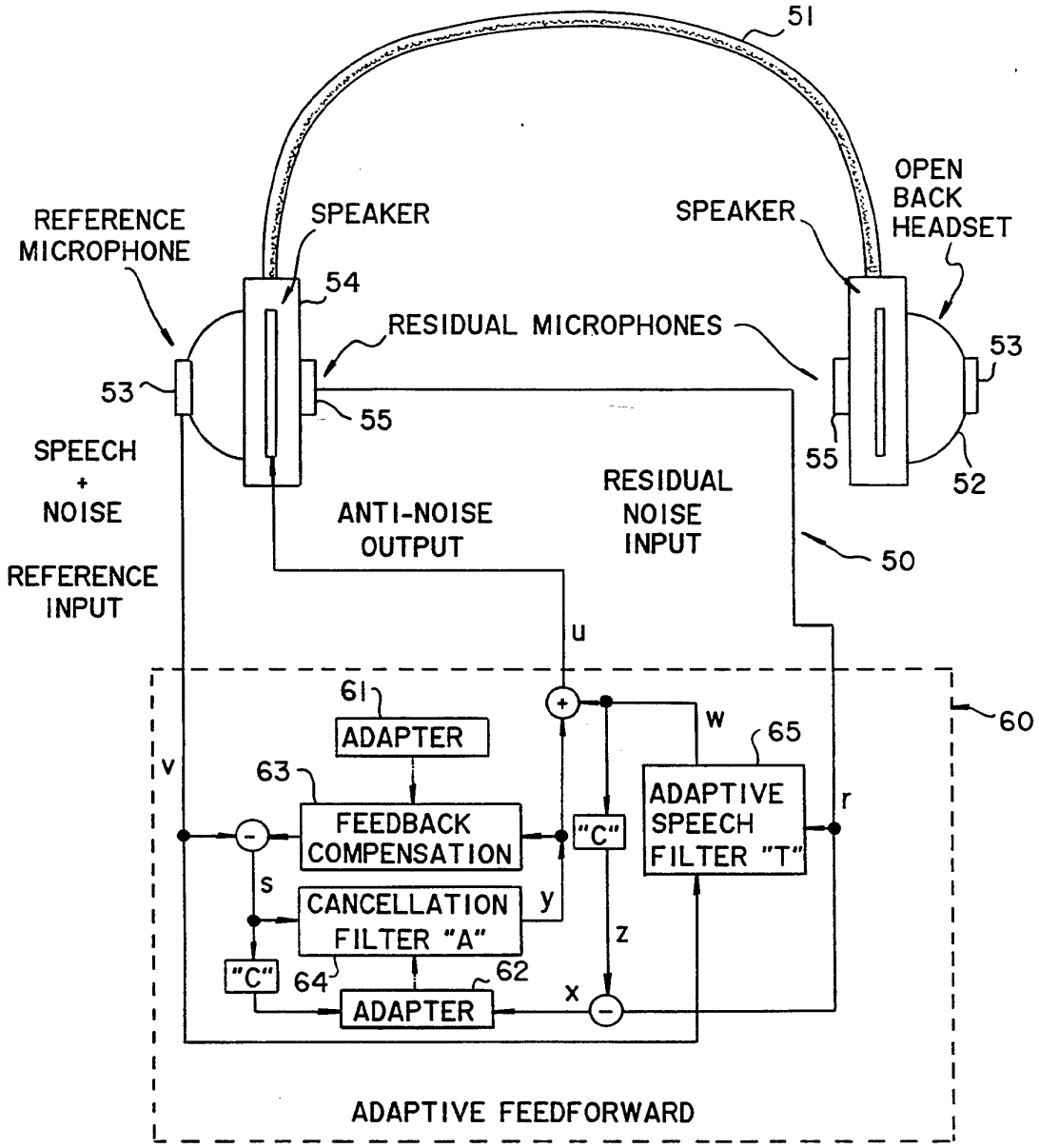

51
OPEN BACK HEADSET
REFERENCE MICROPHONE
SPEAKER
SPEAKER
54
RESIDUAL MICROPHONES
53
53
55
55
52
SPEECH + NOISE
ANTI-NOISE OUTPUT
RESIDUAL NOISE INPUT
50
REFERENCE INPUT
u
60
61
ADAPTER
w
65
v
63
FEEDBACK COMPENSATION
"C"
ADAPTIVE SPEECH FILTER "T"
r
s
CANCELLATION FILTER "A"
y
z
"C"
64
62
ADAPTER
x
ADAPTIVE FEEDFORWARD

FIG.2

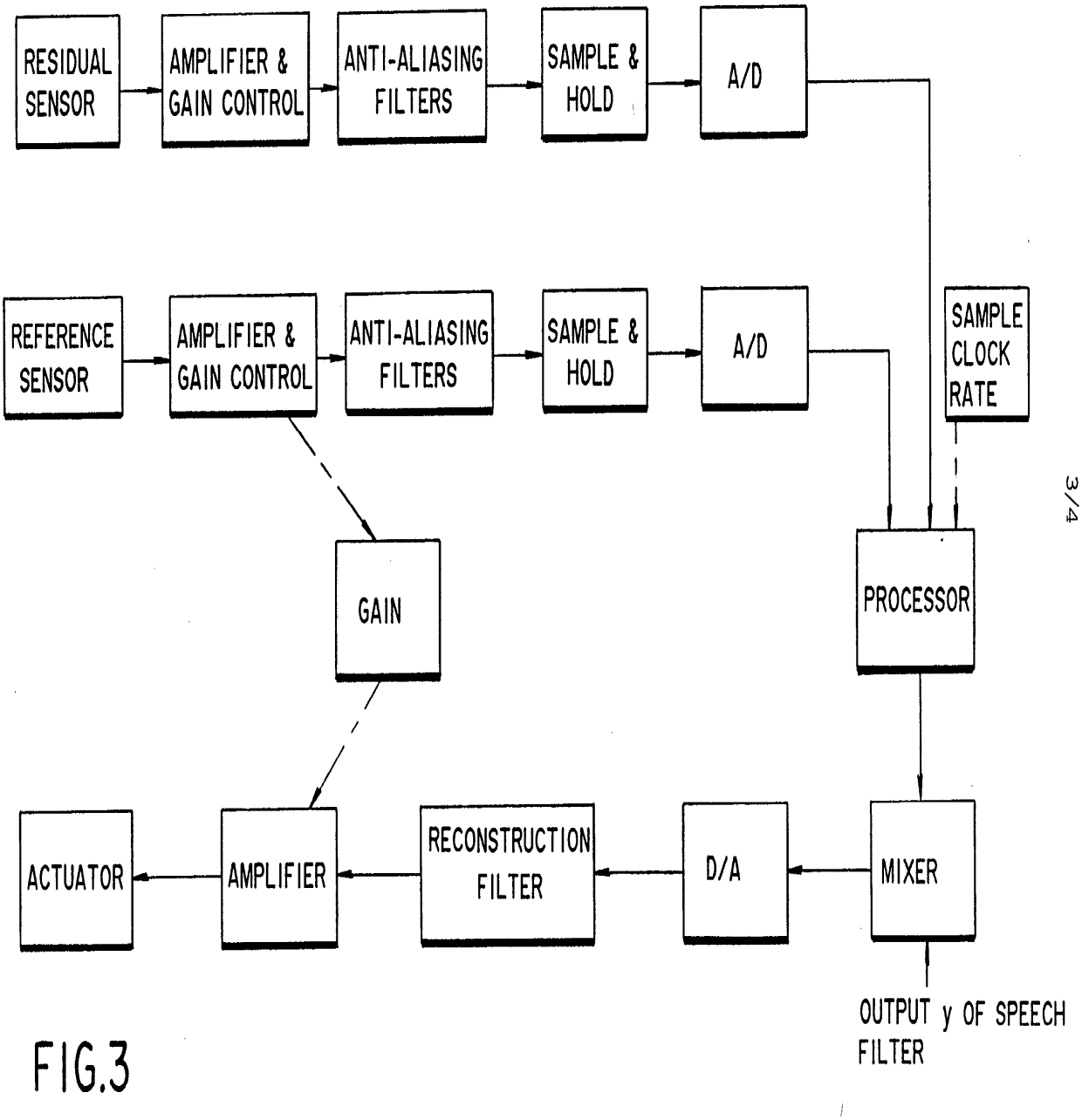


FIG.3

SUBSTITUTE SHEET

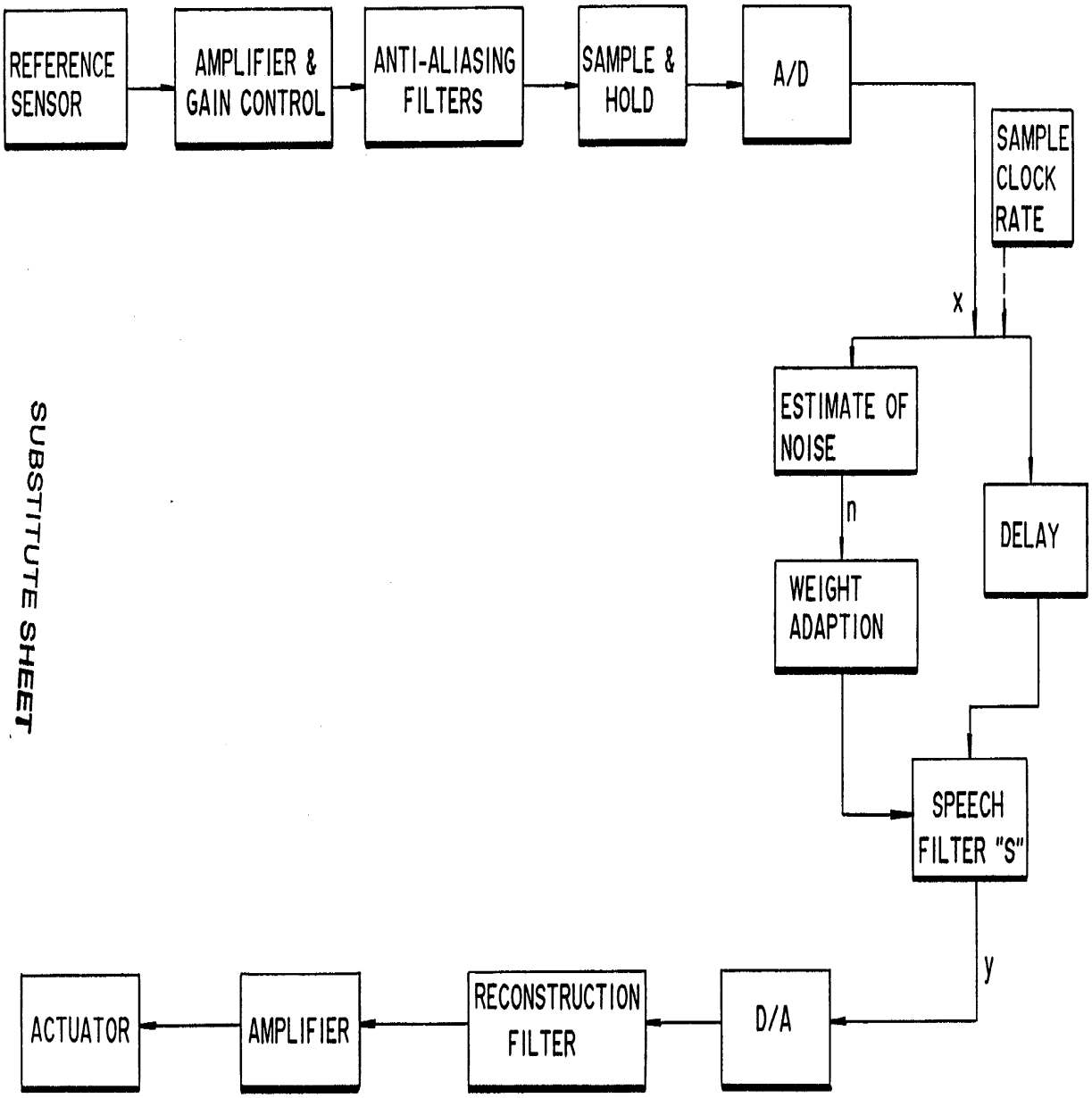


FIG.4

SUBSTITUTE SHEET

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/04567

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :H03B 29/00
US CL :381/71
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. : 381/72,74

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 Y	US,A 3,952,158 (KYLE ET AL) 20 APRIL 1976 SEE FIGS. 1-5	1,8 2-7,9-12
Y	US,A 5,046,103 (WARNAKA ET AL) 03 SEPTEMBER 1991 SEE ABSTRACT AND FIGS. 1-2	2-7,9-12
Y	US,A 5,091,953 (TRETTER) 25 FEBRUARY 1992 SEE ABSTRACT AND FIGS. 1 AND 5	2-7,9-12
X Y	US,A 4,061,158 (FREIFELD) 06 DECEMBER 1977 SEE FIGS. 2-5	1-8 2-7,9-12
Y	US,A 4,064,362 (WILLIAMS) 20 DECEMBER 1977 SEE ABSTRACT AND FIG. 1	1,8
Y	US,A 4,953,217 (TWINEY ET AL) 28 AUGUST 1990 SEE ABSTRACT AND FIG. 1	2-7,9-12
A	US,A 4,677,678 (MCCATCHEN) 30 JUNE 1987 SEE ABSTRACT AND FIGS. 1-7	1-12

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"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
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"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
"E"	earlier document published on or after the international filing date	"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
"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)	"&"	document member of the same patent family
"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		

Date of the actual completion of the international search 06 AUGUST 1992	Date of mailing of the international search report 30 SEP 1992
Name and mailing address of the ISA/ Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer EDWARD LEFKOWITZ Telephone No. (703) 305-4816

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/04567

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US,A 5,105,377 (ZIEGLER, JR.) 14 APRIL 1992 SEE ABSTRACT AND FIGS. 1-5	2-7, 9-12

Form PCT/ISA/210 (continuation of second sheet)(July 1992)★

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 640 972 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
29.03.2006 Bulletin 2006/13

(21) Application number: **05028366.2**

(22) Date of filing: **23.12.2005**

(51) Int Cl.:
G10L 21/02 *(2006.01)* ***H04R 25/00*** *(2006.01)*
H04R 3/00 *(2006.01)* ***H04R 1/10*** *(2006.01)*

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR
Designated Extension States:
AL BA HR MK YU

(71) Applicant: **Phonak AG**
8712 Stäfa (CH)

(72) Inventors:
- **Dijkstra, Evert**
  2046 Fontaines (CH)
- **Hautier, Olivier**
  2065 Savagnier (CH)
- **Destrez, Nicolas**
  1700 Fribourg (CH)

(74) Representative: **Schwan - Schwan - Schorer Patentanwälte**
Bauerstrasse 22
80796 München (DE)

(54) System and method for separation of a users voice from ambient sound

(57) The invention relates to a system for separation of a user's voice from ambient sound, comprising a device (10) to be wom at the user's ear or at least partly in the user's ear canal (12) comprising a first microphone (M1) oriented outwardly towards the environment and a second microphone (M2) oriented inwardly towards the user's ear canal, and an audio signal processing unit (24, 26) for processing audio signals from the first and second microphone by a blind source separation algorithm adapted to separate the user's voice from ambient sound.

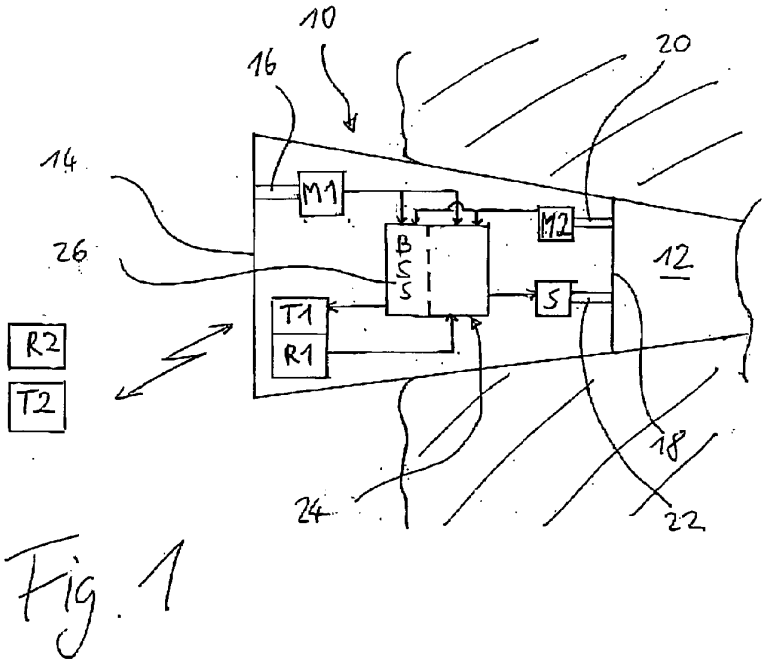


Fig. 1

EP 1 640 972 A1

Printed by Jouve, 75001 PARIS (FR)

Description

[0001] The present invention relates to a system and a method for separation of a user's voice from ambient sound by using at least one device to be worn at the user's ear or at least partly in the user's ear canal.

[0002] For communication purposes, in particular for wireless electronic communication between or from persons exposed to a noisy environment, such as workers in industrial plants, policemen, soldiers, firemen, etc., it is desirable to have a sound pick-up system which is capable at least to some extent to separate the user's voice from ambient noise, or generally ambient sound, in order to improve the intelligibility of the person's speech to the listener, who may be one of the other persons exposed to the noisy environment or who may be a remote person.

[0003] A common approach to achieve such separation of a person's voice is the use of a boom microphone, i.e. a microphone which is placed close to the mouth, carried by a headset, helmet or any other device worn by the person. Such microphone selectively emphasizes the near field around the mouth.

[0004] Other approaches are vibration pick-up devices which are in direct contact with the throat, picking up the vibrations of the vocal chord, or which are in direct contact with the meatus wall or the outer ear canal, picking up the vibrations of the head tissue (i.e. "bone conduction" microphones) or which are in direct contact with the cheek-bone.

[0005] Devices of these types are either fairly sensitive to acoustic noise masking the speech or certain speech sounds are poorly transmitted, especially the high frequency consonant sounds necessary for good intelligibility. Furthermore, for industrial applications boom microphones have the drawbacks that they limit the freedom of movement of the user and that, when combined with a hearing protection device, they will affect the stability and hence the attenuation of the hearing protection device. Bone-conduction microphones have the drawbacks that they have a very limited audio bandwidth which limits the intelligibility of the speech and that they often have to be pressed fairly hard which causes discomfort to the user.

[0006] US 6,661,901 B1 relates to an active hearing protection system comprising an earplug with an outer microphone for picking up ambient sound and an inner microphone which is sealed with respect to ambient sound but is open towards the inner part of the user's ear canal. In an operation mode in which separation of the user's voice from ambient noise is desired only the inner microphone is activated while the outer microphone is not, with the signal from the inner microphone being processed by an electronics unit integrated within the earplug in order to make the user's voice highly natural and intelligible, either for the user himself or his external communication partners.

[0007] Another approach is based on a so-called "blind source separation" (BSS) algorithm in order to separate a person's voice from background noise by corresponding audio signal processing. In this respect, US 2003/0055535 A1 relates to the use of a BSS algorithm for separating the voice of an operator of a vehicle wheel alignment system with a voice audio interface from background noise by using a microphone array in order to avoid the necessity to use a headset. US 2005/0060142 A1 relates in a more general manner to the use of two spaced-apart microphones operated with a BSS algorithm for voice separation from background noise in audio applications.

[0008] It is an object of the invention to provide for a system and a method for separation of a person's voice from ambient sound wherein good intelligibility of speech is achieved while nevertheless discomfort to the person is to be avoided as far as possible.

[0009] According to the invention this object is achieved by a system as defined in claims 1 and 25, and by corresponding methods as defined in claims 28 and 29, respectively.

[0010] The invention is beneficial in that, by using a first microphone and a second microphone, wherein according to the solution of claims 1 and 28 the first microphone is oriented outwardly towards the environment and the second microphone oriented inwardly towards the user's ear canal and according to the solution of claims 25 and 29 the first microphone is located at the right ear and the second microphone is located at the left ear, and by processing the audio signals from the first and second microphone by a blind source separation algorithm, good separation of the user's voice from ambient sound with resulting high intelligibility of the user's speech is achieved without the need for additional restrictions regarding the location of the microphones, so that in particular the need for a boom microphone or a bone-conduction microphone can be avoided. In particular, by orienting the first microphone towards the environment and orienting the second microphone towards the user's ear canal or by locating the first microphone at the right ear and the second microphone at the left ear these two microphones pick up sufficiently differently mixed signals of the ambient sound and the user's voice so that a BSS algorithm will work efficiently.

[0011] Generally, blind source separation (also referred to as "independent component analysis" (ICA)) is a technique for separating mixed source signals (components) which are presumably statistically independent from each other. In its simplified form, blind source separation applies an "un-mixing" matrix of weights to the mixed signals, for example, multiplying the matrix with the mixed signals, to produce separated signals. The weights are assigned initial values, and then adjusted to maximize joint entropy of the signals in order to minimize information redundancy. This weight-adjusting and entropy-increasing process is repeated until the information redundancy of the signals is reduced to a minimum. Because this technique does not require information on the

Exhibit 1009
Page 1482 of 2198

source of each signal, it is referred to as "blind source separation". An introduction to blind source separation is found, for example, in US 2005/0060142 A1.

[0012] In the most simple cage, BSS is applied to two different mixtures of two (acoustic) sources, wherein the two different mixtures are obtained by using two spaced apart microphones. Mixing of the two sources can be represented by a matrix A, with the BSS algorithm corresponding mathematically to finding the inverse matrix of A without knowing anything about the matrix nor about the sources, except that they are statistically independent. In the case of a person's voice mixed with background noise the latter assumption usually is valid. The mixtures of the two sources could be different with respect to amplitude and/or phase of the two sources. In other words, by picking up sound signals with two differently oriented microphones the signal of each of these microphones will correspond to a mixture which is different with regard to the difference in amplitude and/or phase of the two acoustic sources (i.e. user's voice on the one hand and ambient noise on the other hand). By orienting one of the microphones outwardly towards the environment and the other microphone inwardly to the ear canal, a particularly large difference between the two mixtures can be obtained in a simple and particularly comfortable manner, i.e. no boom microphones or bone-cunduction microphones which would cause discomfort to the user need to be used.

[0013] According to one embodiment, the two microphones are part of a hearing protection device. With such a configuration, the microphones can be arranged such that the ambient sound reaching the inwardly oriented microphone is attenuated by the hearing protection device relative to ambient sound reaching the outwardly oriented microphone. Although such hearing protection device could be an earmuff, according to a preferred embodiment the hearing protection device is an earplug comprising a shell which is to be inserted at least partially into the user's ear canal. According to one embodiment, the shell is a customized hard shell having an elasticity from shore D85 to D65 and having an outer shape according to the measured inner shape of the user's outer ear and ear canal. According to an alternative embodiment, the shell is a generic soft shell capable of adapting to the shape of the user's outer ear and ear canal.

[0014] Preferably, the inwardly oriented microphone is located at or is open to the inner part of the shell which is to be inserted into the ear canal, with the inwardly oriented microphone preferably being located at the inner end of the shell or within a channel of the shell open to the inner end of the shell.

[0015] Preferably, the outwardly oriented microphone is located at or is open to the outer part of the shell which is not to be inserted into the ear canal, with the outwardly oriented microphone preferably being located at the outer end of the shell or within a channel of the shell open to the outer end of the shell.

[0016] If the hearing protection device is an earmuff, the first microphone is located at or is open to the outer side of the earmuff and the second microphone is located at or is open to the inner side of the earmuff.

[0017] In all embodiments, the device preferably comprises a speaker adapted to provide an external audio signal to the user's ear. Thereby bidirectional communication with the user is achieved. Preferably the speaker is located at or is open to a portion of the device which is to be worn within the user's ear canal, whereby the speaker is brought acoustically close to the user's ear drum so that good intelligibility of the sound provided by the speaker is achieved even in noisy environments.

[0018] Usually the device will be binaural, i.e. it will comprise one unit for the right ear and another unit for the left ear. The speaker and the microphones may be integrated in the same unit, i.e. in at least one of the units, or the speaker may be part of the unit for one ear and the microphones may be part of the unit for the other ear.

[0019] According to one embodiment, the device may be adapted to be worn completely within the user's ear canal, whereby the device can be more or less completely hidden from the views of other persons. According to other embodiments, the device may be any other kind of a wired or wireless headset.

[0020] Preferably the audio signal processing unit is integrated within the device. In an alternative embodiment, the audio signal processing unit may be adapted to be worn behind the user's ear or somewhere at the user's body. In this case, although the device becomes more handy, more space, and hence also more power, is available for the audio signal processing unit, which may reduce the costs an/or may improve the performance of the blind source separation. In such case, the audio signal processing unit may be connected to the microphone either by wires, which is the most simple solution, or via a wireless link, for example, a radio frequency link such as Bluetooth link, an inductive link or an infrared link, which solution would result in enhanced wearing comfort for the user.

[0021] As a further alternative, the audio signal processing unit could be designed to be located remote from the user and is connected to the microphones via a radio frequency link such as Bluetooth link. In this case, there would be even less restrictions regarding the size and power consumption of the audio signal processing unit, which might result in reduced costs and/or increased performance of the blind source separation.

[0022] In all embodiments, the system preferably comprises a radio frequency transmitter for transmitting the processed audio signal output of the audio signal processing unit to a remote radio frequency receiver in order to provide the user's voice to another person. According to one embodiment, the radio frequency transmitter is integrated within the audio signal processing unit, i.e. the radio frequency transmitter may be either integrated within the device or it may be adapted to be worn behind the use's ear or at the user's body or it could be even located remote from the user. However, in an

Exhibit 1009
Page 1483 of 2198

alternative embodiment, the radio frequency transmitter could be remote from the audio signal processing unit, in which case the processed audio signal output of the audio signal processing unit would be provided to a radio frequency transmitter by wires or via a radio frequency link, an inductive link or an infrared link. In this case the location of the audio signal processing unit and the radio frequency transmitter can be optimized independently form each other.

[0023] In all embodiments, the blind source separation algorithm preferably works in the frequency domain, i.e. the algorithm is simultaneously carried out in different frequency bands/bins and the outcomes of these bands/bins are combined in an appropriate way.

[0024] In all embodiments, the blind source separation algorithm preferably works with the assumption that the sources are statistically independent, i.e. that the user's voice is independent of the ambient sound.

[0025] In the following, examples of the invention will be explained in detail by reference to the attached drawings.

Fig. 1 shows a schematic view of a first embodiment of a hearing protection earplug system according to the invention inserted into a person's ear canal with a block diagram of the most relevant electronic components;

Fig. 2 shows a schematic view of a modified embodiment of a hearing protection earplug according to the invention;

Figs. 3 and 4 show further modified embodiments of a system according to the invention;

Fig. 5 shows a modification of the embodiment of Fig. 1; and

Fig. 6 shows a modification of the embodiment of Fig. 5.

[0026] The system of Fig. 1 comprises a hearing protection earplug 10 which may have a hard shell with an elasticity from shore D85 to shore D65, which is customized, i.e. it has an outer shape according to the individual measured inner shape of the user's outer ear and ear canal. The hard shell may be manufactured by layer-by-layer laser sintering of a powder material, for example, polyamide powder, or by laser stereo-lithography or photo-polymerization. An overview regarding such additive layer-by-layer build-up processes for manufacturing customized shells of hearing devices can be found, for example, in US 2003/0133583 A1 or US 6,533,062 B1. The inner shape of the person's ear canal and outer ear can be measured, for example, by taking an impression which then undergoes laser scanning or by direct laser scanning of the ear.

[0027] Preferably the hard shell is designed such that it provides for an acoustic attenuation, averaged over the audible frequency range, of at least 10 dB when inserted into the user's ear canal.

[0028] Rather than having a customized hard shell, earplug 10 may have a generic soft shell which adapts to the shape of the users outer ear and ear canal due to its elasticity.

[0029] In Fig. 1 the earplug 10 is shown to be inserted with its inner end portion into a user's ear canal 12. The earplug 10 comprises an outer microphone M1, which is located at the outer part of the earplug 10 and which is open to the outer end 14 of the earplug 10 via a sound channel 16, and an inner microphone M2, which is located at the inner part of the earplug 10 and which is open to the inner end 18 of the earplug 10 via a sound channel 20. The earplug 10 further comprises a speaker S which is likewise located at the inner part of the earplug 10 and is open to the inner end 18 via a sound channel 22. The earplug 10 also includes a digital audio signal processing unit 24 comprising a digital signal processor as well as a radio frequency transmitter unit T1 and a radio frequency receiver unit R1. The audio signal processing unit 24 functionally includes a unit 26 for applying a blind source separation algorithm to the audio signals provided by the microphones M1 and M2. It is evident to the skilled person that the earplug 10 includes additional electrical components, such as amplifiers and analogue-to-digital converters for the audio signals provided by the microphones M1 and M2, a digital-to-analogue converter and an amplifier for the audio signal to be transformed into sound by the speaker S, program and data memory components for the audio signal processing unit 24, a battery for power supply, a demodulator, a modulator, an antenna, etc., which are not shown in Fig. 1 for the sake of simplicity.

[0030] The transmitter T1 of the earplug 10 is adapted to transmit audio signals from the earplug 10 to a remote receiver R2 via a radio link, while the receiver R1 of the earplug 10 is adapted to receive audio signals from a remote transmitter T2. The audio signals received by the receiver R1 are demodulated and then undergo signal processing in the audio signal processing unit 24 as input to the speaker S in order to provide remote audio signals to the user. Such remote audio signals could be the speech of another person picked up by a microphone whose output is sent to the remote transmitter T2 by wires or via, for example, a mobile telephone or mobile radio device.

[0031] The audio signals provided by the microphones M1 and M2 are passed as input to the blind source separation unit 26, in which a processed audio signal output for the transmitter T1 is produced, with the processed audio signal output consisting completely or at least essentially of the user's voice which has been separated from the ambient sound by action of the blind source separation algorithm carried out in the BSS unit 26. Such BSS signal processing utilizes the fact that the sound

Exhibit 1009
Page 1484 of 2198

mixtures picked up by the microphone M1 which is oriented towards the environment and M2 which is oriented towards the ear canal 12, respectively, consist - due to the different orientation of the microphones M1 and M2 - of essentially different mixtures of the ambient sound and the user's voice, which are different regarding amplitude ratio of these two signal contributions or sources (i.e. ambient sound on the one hand and user's voice on the other hand) and regarding phase difference of these two signal contributions of the mixture.

[0032] The output signal of the BSS unit 26 is transmitted via the transmitter T1 to the remote receiver R2 which usually will be connected to a remote speaker for presenting the user's voice to another person. The remote speaker and the remote microphone connected to the remote transmitter T2 could be part of an earpiece or an earplug worn by another person, which may be similar or identical to the earplug 10. Thereby in-the-ear hearing protection devices with integrated communication function can be achieved. Such a system could be used by any persons who need to communicate in a noisy environment, such as workers, soldiers, firemen, etc. However, the remote receiver R2 also might serve for communication via a mobile telephone or a mobile radio device.

[0033] In general, the remote receiver R2 / transmitter T2 could be a part of an interface of a standard wireless communication device, such as a mobile telephone device or a mobile radio device Preferably, the wireless link between the transmitter T1 / receiver R1 and the remote receiver R2 / remote transmitter T2 is a Bluetooth link. The remote receiver R2 / transmitter T2 then could be a part of the Bluetooth interface of a standard wireless communication device, such as a mobile telephone device or a mobile radio device.

[0034] According to Fig. 1, the audio signal produced by the microphones M1 and M2, which is used as input to the BSS unit 26, in addition could undergo other audio signal processing within the audio signal processing unit 24, for example, in order to provide the ambient sound picked up by the microphone M1 to the speaker S for enabling perception of desired ambient sounds during times when the noise level is low so that the hearing protection function provided by the earplug 10 is not needed, or to provide for an in-situ measurement of the actual acoustic attenuation provided by the earplug 10 when inserted into the user's ear canal 12 by comparing the sound levels measured by the microphones M1 and M2.

[0035] In Fig. 2 a modified embodiment is shown, wherein the transmitter T1 and the receiver R1 have been removed from the earplug 10 and now are located in an external unit 28 which is connected by wires 30 to the earplug 10 and which is to be worn at an appropriate location at the user's body, for example behind the user's ear or somewhere within or below the user's clothing, for example, within a helmet. The external unit 28 may include further electrical components removed from the earplug 10, for example, the power supply battery.

[0036] In Fig. 3 a further modified embodiment is shown, wherein not only the transmitter T1 and the receiver R1 (and the power supply battery) but in addition also the audio signal processing unit 24 including the BSS unit 26 has been removed from the earplug 10 and is connected via corresponding wires to the microphones M1 and M2 and the speaker S, respectively.

[0037] In Fig. 4 a still further modified embodiment is shown, wherein the wire connection 30 between the earplug 10 and the external unit 28 has been replaced by a bidirectional wireless link 32, which is established by a transmitter T3 and a receiver R3 located in the earplug 10 and a corresponding receiver R4 and transmitter T4 located in the external unit 28. The wireless link 32 could be a radio frequency link, an inductive link or an infrared link. If it is an inductive link, the external unit 28 would have to be located relatively close to the earplug 10, for example at a distance of not more than 20 cm. With an RF link, the distance could be larger; in this case it would be even be possible to locate the external unit 28 at a position remote from the user's body, for example, within a distance of a few meters. Preferably the wireless link 32 of the embodiment of Fig. 4 is a Bluetooth link.

[0038] Although in the embodiments shown is Figs. 1 to 4 the speaker S is shown as being integrated in the same earplug 10 as the microphones M1 and M2, in other embodiments the speaker S may be provided at an earplug other than the earplug at which the microphones M1 and M2 are provided.

[0039] Fig. 5 shows a modified version of the embodiment of Fig. 1, wherein the inwardly oriented second microphone M2 is not part of the earplug 10 — as shown in Fig. 1 — but rather is provided at the earplug 11 for the other ear. The second microphone M2 is located at the inner part of the earplug 11 and is open to the inner end 18 of the earplug 11 via a sound channel 20. In addition, the earplug 11 further comprises a speaker S which is likewise located at the inner part of the earplug 11 and is open to the inner end 18 via a sound channel 22. The microphone M2 and the speaker S of the earplug 11 are connected to the audio signal processing unit 24 located in the earplug 10 via a wire connection 40 which be integrated in a mechanical connection (not shown) between the earplugs 10 and 11 which may be provided for preventing loss of the earplugs 10, 11.

[0040] According to a modification of the embodiment of Fig. 5, the outwardly oriented first microphone M1 may be provided at the earplug 11, while the inwardly oriented second microphone M2 may be provided at the earplug 10 as in Fig. 1.

[0041] An alternative modification of the embodiment of Fig. 5 is shown in Fig. 6, wherein both the first microphone M1 of the earplug 10 and the second microphone M2 of the earplug 11 are outwardly oriented. In this case the differences in the sound mixtures picked up by the microphone M1 of the earplug 10 and the microphone M2 of the earplug 11 are due to the different location of

Exhibit 1009
Page 1485 of 2198

the microphone M1 at one ear and the microphone M2 at the other ear (for example, a disturbing noise source usually will be located closer to one of the two ears and hence will contribute more strongly and with different phase to the signal picked up by the microphone located at that ear than to that located at the other ear).

[0042] According to a modification of the embodiment of Fig. 6 both microphones M1 and M2 could be inwardly oriented rather than outwardly oriented.

[0043] While Figs. 5 and 6 and the above-discussed variants thereof are modifications of the embodiment of Fig. 1, according to which both the audio signal processing unit 24 and the transmitter T1 / receiver R1 are part of the earplug 10, it is to be understood that the embodiments of Figs. 2 to 4 may be analogously modified, i.e. by providing the second microphone M2 at the earplug 11 rather than at the earplug 10.

[0044] While the examples discussed so far relate to an application of the communication system to hearing protection earplugs, the invention also may be applied to earmuffs serving as a hearing protection device or generally to any kind of headset, i.e. also to hearing devices which do not provide for a hearing protection function. For example, the headset may consist of a device with a customized or a generic shell which is designed such that it can be worn completely in the ear canal 12 (i.e. as a "CIC device") and which serves exclusively communication purposes, for example, for security persons, policemen, firemen, etc., i.e. for persons who are exposed to a noisy environment, with the noise level, however, being below a threshold value which would require the use of hearing protection devices. Such CIC device generally could have the construction of the earplug of Fig. 1, however, without considerable acoustic attenuation provided due to the material and the design of the shell, and would be reduced in size so that it can be worn completely within the ear canal 12.

[0045] In general, the device need not be designed as an earplug but it may rather be any kind of headset.

Claims

1. A system for separation of a user's voice from ambient sound, comprising a device (10) to be worn at the user's ear or at least partly in the user's ear canal (12) comprising a first microphone (M1) oriented outwardly towards the environment and a second microphone (M2) oriented inwardly towards the user's ear canal, and an audio signal processing unit (24, 26) for processing audio signals from the first and second microphone by a blind source separation algorithm adapted to separate the user's voice from ambient sound.

2. The system of claim 1, wherein the device (10) is designed as a hearing protection device.

3. The system of claim 2, wherein the hearing protection device (10) is designed to provide for an acoustic attenuation, averaged over the audible frequency range, of at least 10 dB.

4. The system of claim 3, wherein the microphones (M1, M2) are arranged such that ambient sound reaching the second microphone (M2) is attenuated by the hearing protection device (10) relative to ambient sound reaching the first microphone (M1).

5. The system of one of claims 2 to 4, wherein the hearing protection device is an earplug (10) comprising a shell which is to be inserted at least partially into the user's ear canal (12).

6. The system of claim 5, wherein the shell is a hard shell having an elasticity from Shore D 85 to Shore D 65 and having an outer shape according to the measured inner shape of the user's outer ear and ear canal (12).

7. The system of one of claims 5 and 6, wherein the second microphone (M2) is located at or is open to the inner part of the shell which is to be inserted into the ear canal (12).

8. The system of claim 7, wherein the second microphone (M2) is located at the inner end of the shell or within a channel (20) of the shell open to the inner end of the shell.

9. The system of one of claims 5 to 8, wherein the first microphone (M1) is located at or is open to the outer part of the shell which is not to be inserted into the ear canal (12).

10. The system of claim 9, wherein the second microphone (M2) is located at the outer end of the shell or within a channel (16) of the shell open to the outer end of the shell.

11. The system of one of claims 1 to 4, wherein the hearing protection device is an earmuff and wherein the first microphone is located at or is open to the outer side of the earmuff and the second microphone is located at or is open to the inner side of the earmuff.

12. The system of one of the preceding claims, wherein the device comprises a speaker (S) adapted to provide an external audio signal to the user's ear.

13. The system of claim 12, wherein the speaker (S) is located at or within a channel open to that end of the device which is to be worn within the user's ear canal.

14. The system of one of the preceding claims, wherein the device is adapted to be worn completely within

Exhibit 1009
Page 1486 of 2198

the user's ear canal.

15. The system of one of the preceding claims, wherein the audio signal processing unit (24, 26) is integrated within the device (10).

16. The system of one of claims 1 to 14, wherein the audio signal processing unit (24, 26) is adapted to be worn behind the user' s ear or at the user's body.

17. The system of claim 16, wherein the audio signal processing unit (24, 26) is connected to the microphones (M1, M2) by wires (30) or via a radio frequency link (32) such as a Bluetooth link, an inductive link or an infrared link.

18. The system of one of claims 1 to 14, wherein the audio signal processing unit (24, 26) is designed to be located remote from the user and is connected to the microphones (M1, M2) via a radio frequency link (32) such as a Bluetooth link.

19. The system of one of the preceding claims, wherein the system comprises a radio frequency transmitter (T1) for transmitting the processed audio signal output of the audio signal processing unit (24, 26) to a remote radio frequency receiver (R2) in order to provide the user's voice to another person.

20. The system of claim 19, wherein the remote radio frequency receiver (R2) is part of a Bluetooth interface of a mobile telephone or a mobile radio system.

21. The system of claim 19 or 20, wherein the radio frequency transmitter (T1) is integrated with the audio signal processing unit (24, 26).

22. The system of claim 19 or 20, wherein the radio frequency transmitter (T1) is remote from the audio signal processing unit (24, 26) and wherein the processed audio signal output of the audio signal processing unit is provided to the radio frequency transmitter by wires (30) or via an inductive link.

23. The system of claim 12, wherein the system comprises a radio frequency receiver (R1) for receiving the external audio signal from a remote radio frequency transmitter (T2).

24. The system of claim 23, wherein the remote radio frequency transmitter (T2) is part of a Bluetooth interface of a mobile telephone or a mobile radio system.

25. A system for separation of a user's voice from ambient sound, comprising: a first device (10) to be worn at the user's right ear or at least partly in the user's right ear canal (12), a second device (11) to be worn at the user's left ear or at least partly in the user's left ear canal (12), the first device comprising a first microphone (M1) and the second device comprising a second microphone (M2), and an audio signal processing unit (24, 26) for processing audio signals from the first and second microphone by a blind source separation algorithm adapted to separate the user's voice from ambient sound.

26. The system of claim 25, wherein one of the first microphone (M1) and second microphone (M2) is oriented outwardly towards the environment and the other one of the first microphone and second microphone is oriented inwardly towards the user's ear canal (12).

27. The system of one of claims 25 and 26, wherein the audio signal processing unit (24, 26) is located in one of the first device (10) and the second device (11) or is adapted to be worn behind the user's ear or at the user's body or is designed to be located remote from the user and is connected to the microphones via a radio frequency link (32) such as a Bluetooth link.

28. A method for separation of a user's voice from ambient sound, comprising:

providing the user with a device (10) to be worn at a user's ear or at least partly in the user's ear canal (12) and comprising a first microphone (M1) oriented towards the environment and a second microphone (M2) oriented towards the user's ear canal (12),
picking up sound by the first microphone to create a first audio signal and by the second microphone to create a second audio signal, and
processing the first and second audio signals by a blind source separation algorithm in order to produce a processed audio signal wherein the user's voice is separated from ambient sound.

29. A method for separation of a user's voice from ambient sound, comprising:

providing the user with a first device (10) to be worn at the user's right ear or at least partly in the user's right ear canal (12) and a second device (11) to be worn at the user's left ear or at least partly in the user's left ear canal (12), the first device comprising a first microphone (M1) and the second device comprising a second microphone (M2),
picking up sound by the first microphone to create a first audio signal and by the second microphone to create a second audio signal, and
processing the first and second audio signals by a blind source separation algorithm in order to

Exhibit 1009
Page 1487 of 2198

produce a processed audio signal wherein the user's voice is separated from ambient sound.

30. The method of one of claims 28 and 29, wherein the blind source separation algorithm works in the frequency domain.

31. The method of one of claims 28 to 30, wherein the processed audio signal is provided to another person via a wireless link such as a Bluetooth link.

Exhibit 1009
Page 1488 of 2198

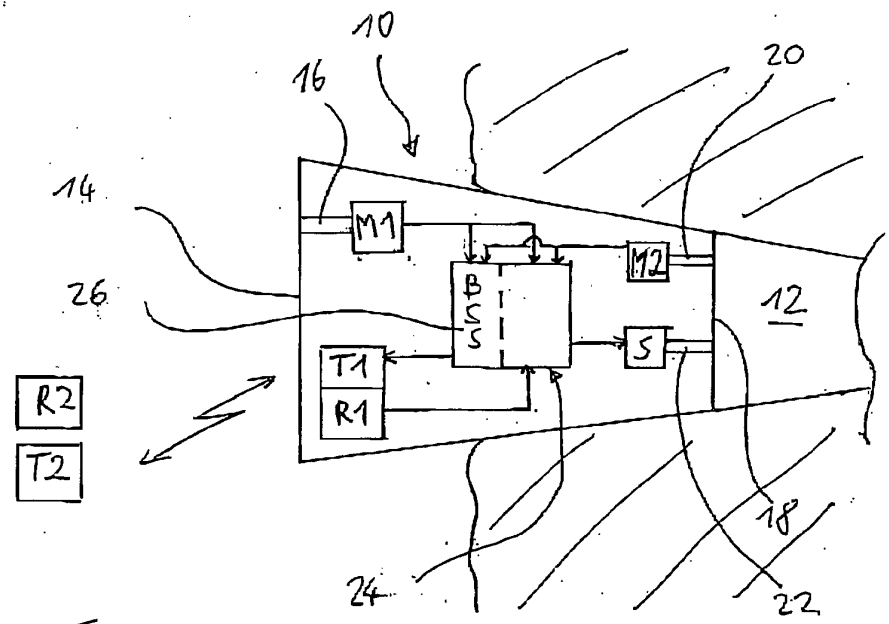


Fig. 1

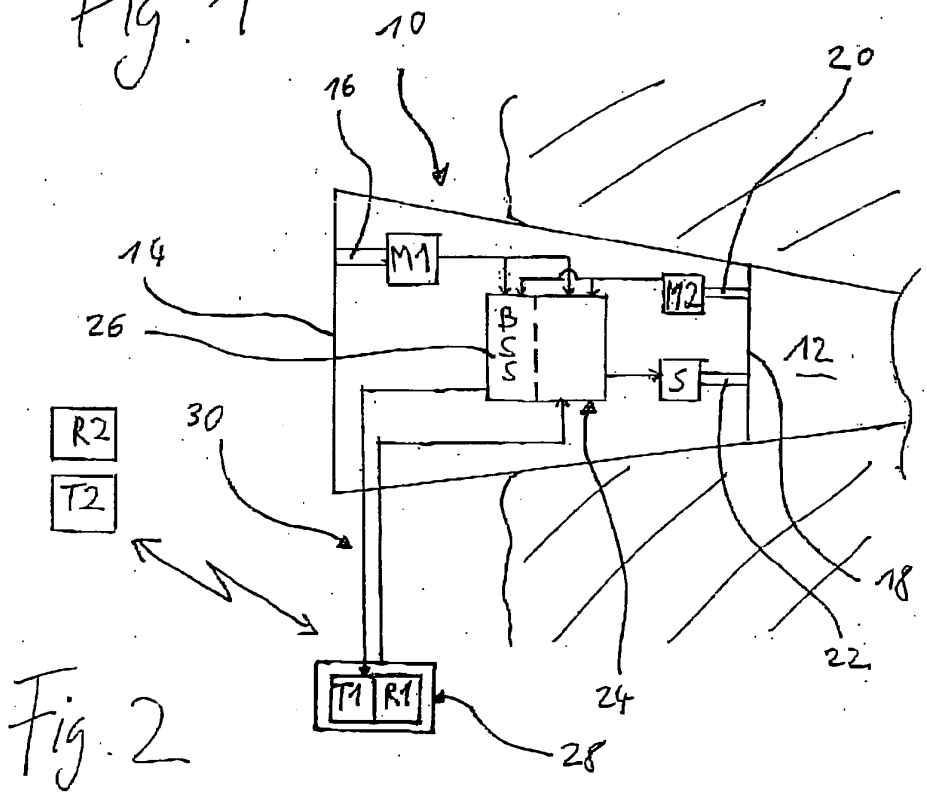


Fig. 2

Exhibit 1009
Page 1489 of 2198

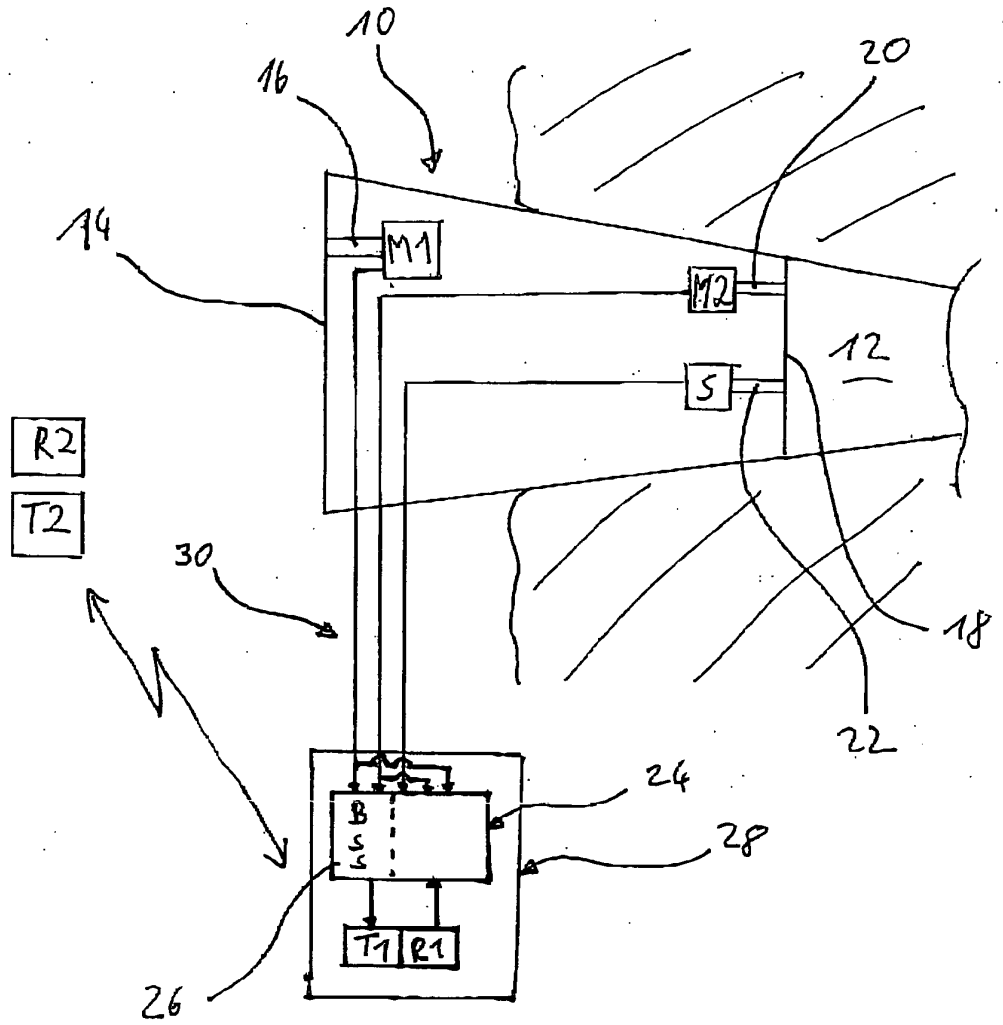


Fig. 3

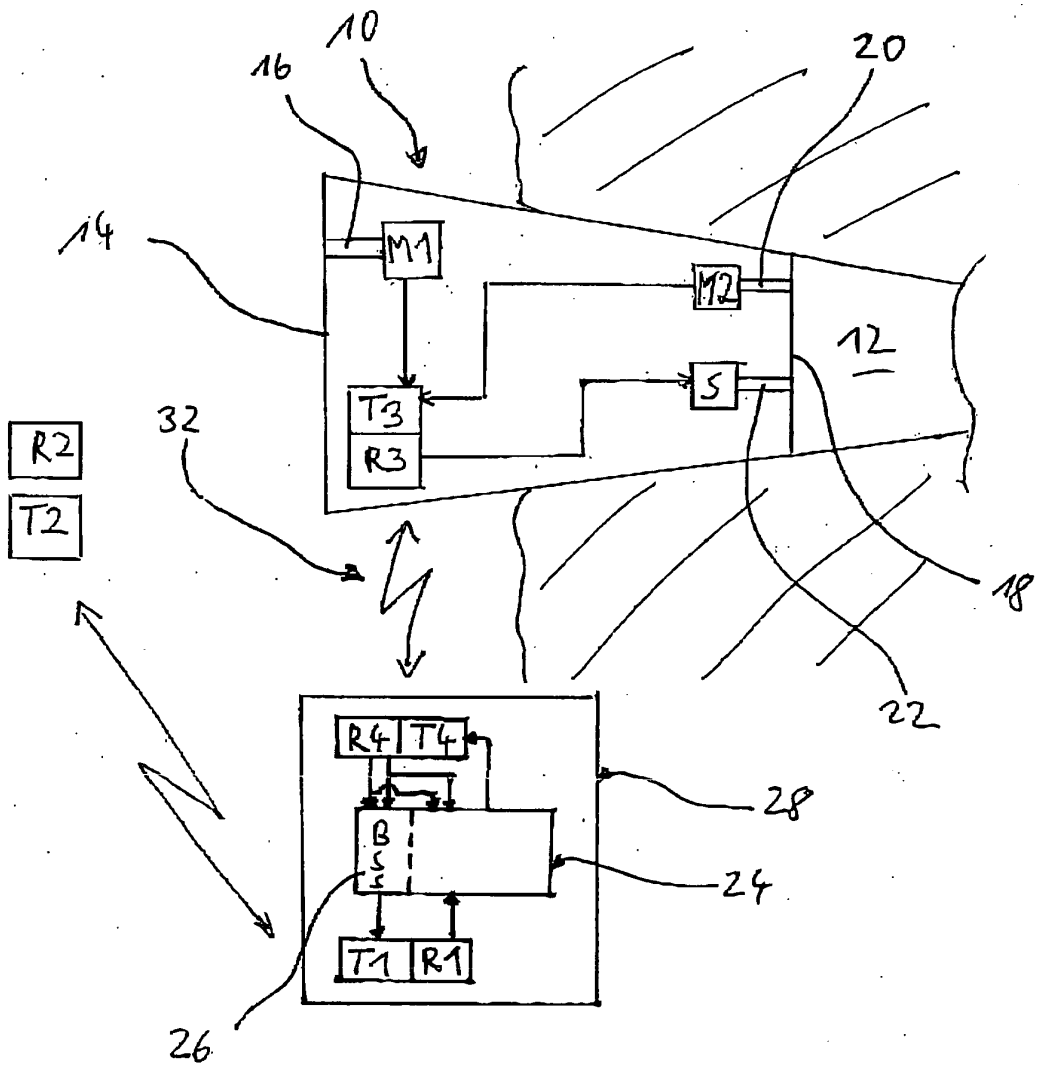


Fig. 4

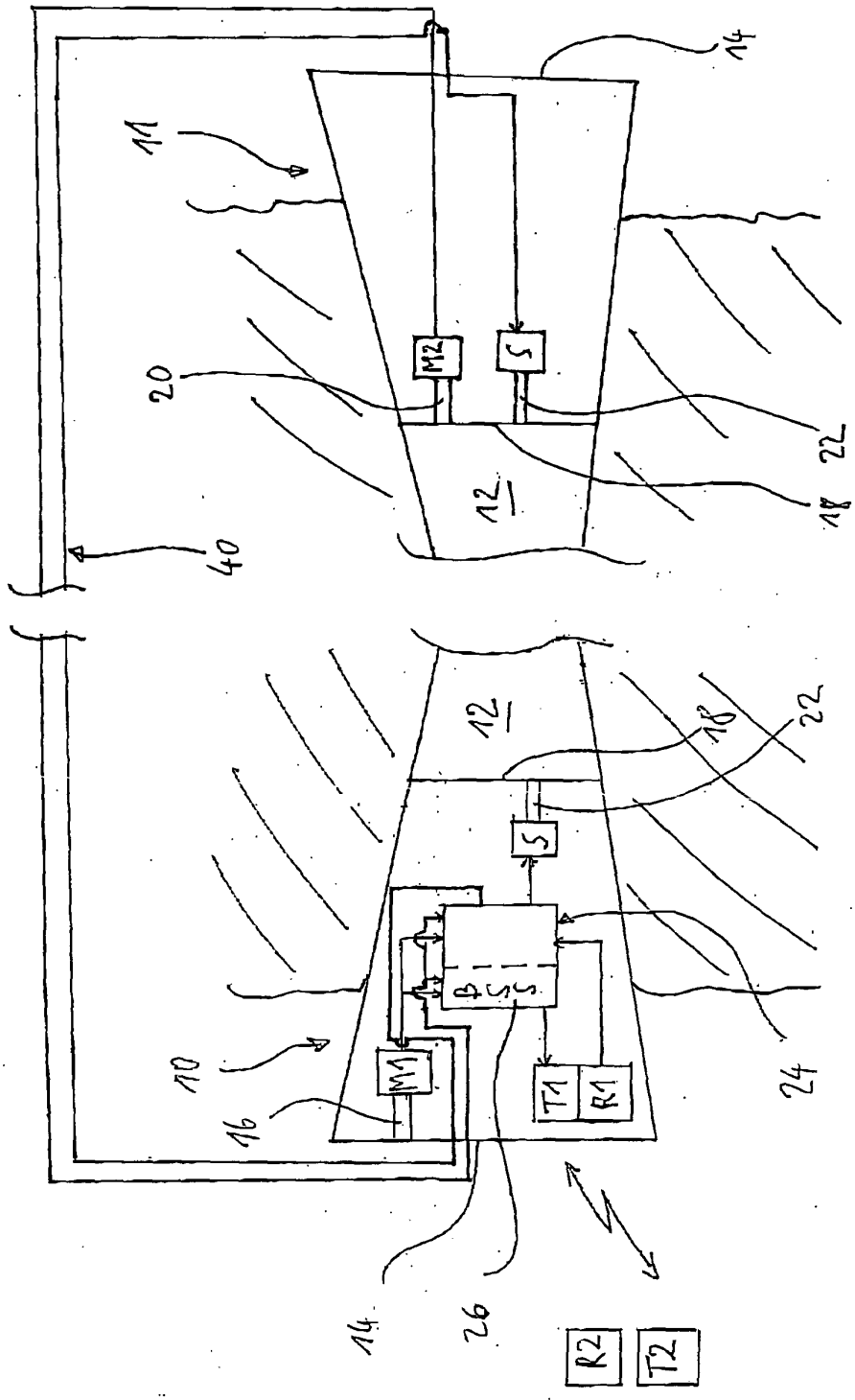


Fig. 5

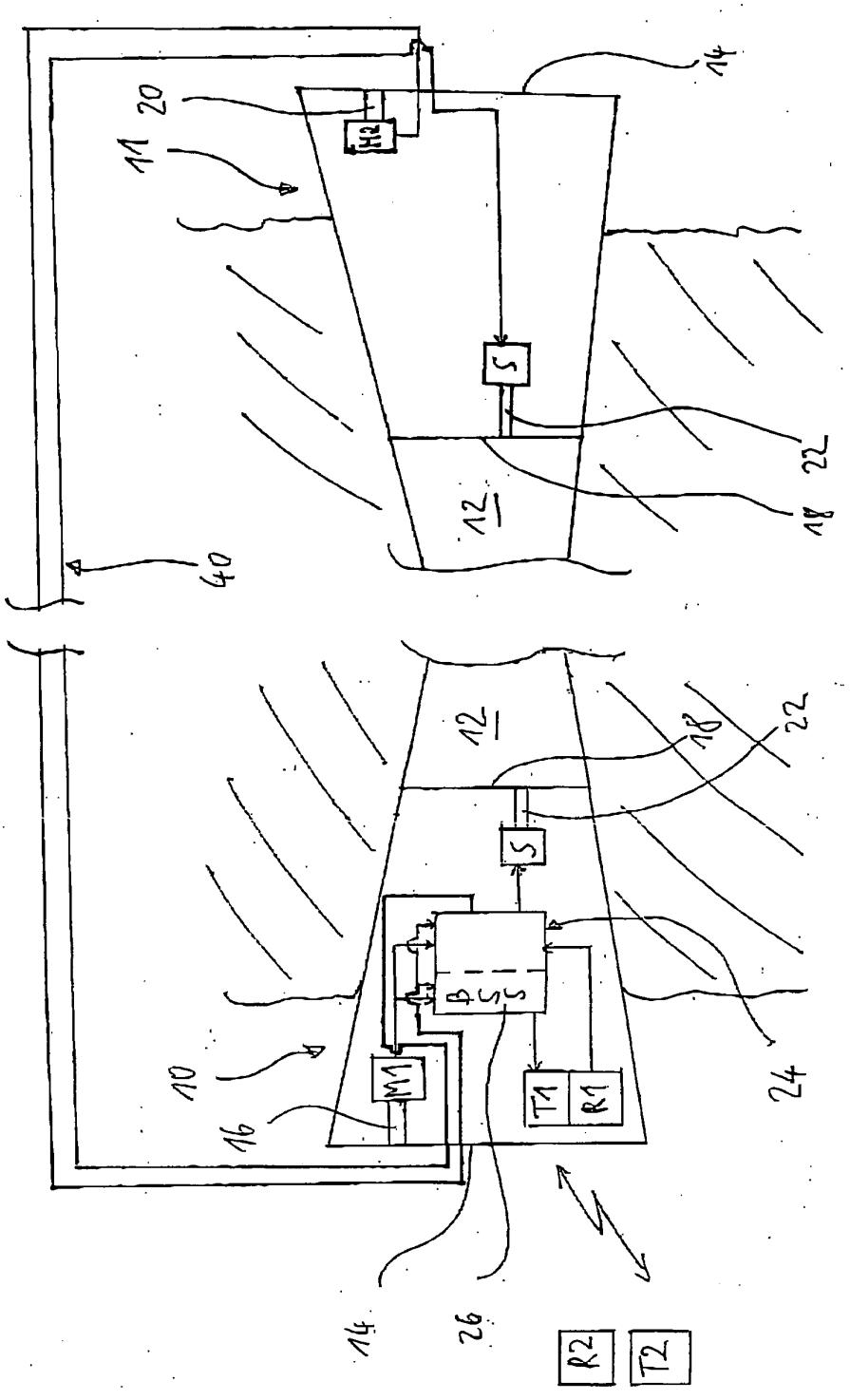


Fig. 6

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 05 02 8366

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X Y	EP 1 509 065 A (BERNAFON AG) 23 February 2005 (2005-02-23) * page 2, paragraph 1 - page 3, paragraph 14 * * page 5, paragraphs 15,16 * * page 6, paragraphs 23,27 * * claim 1 * -----	25,29,30 26,27,31	G10L21/02 H04R25/00 H04R3/00 H04R1/10
Y,D	US 6 661 901 B1 (SVEAN JARLE ET AL) 9 December 2003 (2003-12-09) * column 1, line 5 - column 2, line 33 * * column 4, line 21 - line 32 * * column 5, line 26 - column 7, line 60 * * column 10, line 52 - column 11, line 40 * * claims 10,11 * -----	1-24, 26-28,31	
Y	US 2005/060142 A1 (VISSER ERIK ET AL) 17 March 2005 (2005-03-17) * page 1, paragraph 5-7 * * page 3, paragraph 19 - page 5, paragraph 37 * * page 8, paragraph 61 * -----	1-24,28	TECHNICAL FIELDS SEARCHED (IPC) G10L H04R
A	SANG MIN LEE ET AL: "New idea of hearing aid algorithm to enhance speech discrimination in a noisy environment and its experimental results" SAN FRANCISCO, CA, USA 1-5 SEPT. 2004, PISCATAWAY, NJ, USA,IEEE, US, vol. 3, 1 September 2004 (2004-09-01), page 978Vol2, XP010774935 ISBN: 0-7803-8439-3 * the whole document * -----	1-31	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	3 February 2006	Coda, R

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

Exhibit 1009
Page 1494 of 2198

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 05 02 8366

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

03-02-2006

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1509065 A	23-02-2005	WO 2005020633 A1	03-03-2005
US 6661901 B1	09-12-2003	NONE	
US 2005060142 A1	17-03-2005	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 489 596 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
13.09.2006 Bulletin 2006/37

(51) Int Cl.:
G10L 11/02 *(2006.01)* ***H04R 3/00*** *(2006.01)*
H04R 1/40 *(2006.01)*

(21) Application number: **03445076.7**

(22) Date of filing: **17.06.2003**

(54) **Device and method for voice activity detection**

Verfahren und Vorrichtung zur Sprachaktivitätsdetektion

Procédé et dispositif de détection de l'activité vocale

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PT RO SE SI SK TR

(43) Date of publication of application:
22.12.2004 Bulletin 2004/52

(73) Proprietor: **Sony Ericsson Mobile Communications AB**
221 88 Lund (SE)

(72) Inventor: **Gustavsson, Stefan**
254 43 Helsingborg (SE)

(74) Representative: **Akerman, Marten Lennart et al**
Albihns Malmö AB
Box 4289
203 14 Malmö (SE)

(56) References cited:
EP-A- 1 206 161
US-A1- 2002 009 203
US-A1- 2003 027 600

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

Printed by Jouve, 75001 PARIS (FR)

EP 1 489 596 B1

Description

Field of the invention

[0001] The present invention relates to a device, a mobile apparatus incorporating the device, an accessory therefor and a method for voice activity detection, particularly in a mobile telephone, using the directional sensitivity of a microphone system and exploiting the knowledge about the voice source's orientation in space. The device assists the existing voice activity detection to achieve higher sensitivity and requiring less processor power.

State of the art

[0002] Voice activity detectors are used e.g. in mobile phones to enhance the performance in certain situations. The most common way to construct a voice activity detector is to look at the levels of the sub-bands of the incoming signal. Then the background noise level and the speech level are estimated and compared with a threshold to determine whether speech is present or not. An example of a voice activity detector is disclosed in U.S. patent 6,427,134.

[0003] For instance in noisy environments it is hard to make a uniform parameter set-up for the voice activity detector. Therefore several voice activity detectors are needed, trimmed to the specific cases. For example in some modules you need to be sure that if there is speech it should be detected (echo canceller), but in other cases it is better to indicate no speech if the signal to noise ratio level is too low. The plurality of voice activity detectors put a load on the digital signal processors that have to take care of performing the various voice activity detection algorithms.

[0004] The document US2003/027600 A1 discloses a device using a voice activity detector operating on the signal output by a microphone array. Microphone arrays are known to discriminate the sounds originating from a specific direction by steering the array in order to enhance the signal-to-noise ratio of a given source, thereby focusing on that source.

Summary of the invention

[0005] An object of the present invention is to complement existing voice activity detection taking into account the direction of the source of the sound.

[0006] In a first aspect, the invention provides a device for voice activity detection comprising a sound signal analyser arranged to determine whether a sound signal comprises speech, and
a microphone system arranged to discriminate sounds emanating from sources located in different directions from the microphone system.

[0007] According to the invention, the device is adapted to determine the direction of a sound source causing sound signals;
and is adapted to further analyse the sound to determine whether the sound signal comprises speech, if the sounds emanate from a first range of directions;
but to decide that the sound signal does not comprise speech, if the sounds emanate from a second, different range of directions.

[0008] Suitably, the first range of directions is directed in the direction of an intended user's mouth.

[0009] In one embodiment, the microphone system comprises two microphone elements separated a distance and located on a line directed in the direction of an intended user's mouth.

[0010] The range of directions may be defined as all sounds falling inside a cone with a cone angle α, wherein $10° < \alpha < 30°$, and preferably, α is approximately 25°.

[0011] In another embodiment, the microphone system comprises three microphone elements separated a distance and located in a plane directed in the direction of an intended user's mouth.

[0012] Suitably, two of said three microphone elements are separated a distance and located on a line directed perpendicular to the direction of an intended user's mouth.

[0013] In another embodiment, the microphone system comprises four microphone elements located such that the fourth microphone is not located in the same plane as the three others.

[0014] The microphone elements may be directional with a pattern having maximal sensitivity in the direction of an intended user's mouth.

[0015] In still a further embodiment, the microphone system comprises one directional microphone element together with one or more other microphone elements to remove the uncertainty in the direction of the sound source. The directional microphone element may be used to measure the sound pressure level relative to the other microphone element.

[0016] In a second aspect, the invention provides a mobile apparatus comprising a device as mentioned above.

[0017] Suitably, the microphone elements are located at the lower edge of the apparatus.

[0018] In one embodiment, a plurality of microphone elements are located at the lower edge of the apparatus and at

Exhibit 1009
Page 1497 of 2198

least one further microphone element is located at a distance from the lower edge.

[0019] The mobile apparatus may be a mobile radio terminal, e.g. a mobile telephone, a pager, a communicator, an electric organiser or a smartphone.

[0020] In a third aspect, the invention provides an accessory for a mobile apparatus comprising a microphone system as mentioned above.

[0021] Suitably, the first direction of the range of directions is adjustable.

[0022] The accessory may be a hands-free kit or a telephone conference microphone.

[0023] In a fourth aspect, the invention provides a method for voice activity detection, including the steps of:

receiving sound signals from a microphone system arranged to discriminate sounds emanating from sources located in different directions from the microphone system; determining the direction of the sound source causing the sound signals;
if the sounds emanate from a first range of directions, further analyse the sound to determine whether the sound signal comprises speech;
but if the sounds emanate from a second, different range of directions decide that the sound signal does not comprise speech.

[0024] Suitably, the first range of directions is directed in the direction of an intended user's mouth.

[0025] The first range of directions may be defined as all sounds falling inside cone with a cone angle α, wherein $10° < \alpha < 30°$, and preferably α is approximately 25°.

[0026] In one embodiment, the microphone system comprises at least two microphone elements located at a distance from each other and located on a line directed in the direction of an intended user's mouth, said two microphone elements being separated a distance d, wherein the direction to the sound source θ is calculated as

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

where
Δt is the time difference between the sounds from the two microphone elements,
v is the velocity of sound.

[0027] In another embodiment, one directional microphone element is used together with one or more other microphone elements to remove the uncertainty in the direction of the sound source.

[0028] The directional microphone element may be used to measure the sound pressure level relative to the other microphone element.

[0029] The invention is defined in the attached independent claims 1 and 20, while preferred embodiments are set forth in the dependent claims.

Brief description of the drawings

[0030] The invention will be described below in greater detail with reference to the accompanying drawings, in which:

fig. 1 is a perspective view of a mobile phone incorporating the present invention, and
fig. 2 is a schematic drawing of the receiving angle of an embodiment of the present invention.

Detailed description of preferred embodiments

[0031] As mentioned briefly in the introduction, many signal processing algorithms, such as echo cancellation and background noise synthesis, used in phones and hands-free kits are based on the fact that the user is speaking or not. For example the speech codec is active when the near-end user is speaking and the background synthesis is active when the near-end user is silent. All these algorithms need good voice activity detectors (VAD) to perform well. An error in the detection can result in artefacts or malfunctions caused by divergence of the algorithms or other problems.

[0032] Existing voice activity detectors are directed to determine whether speech is present or not in a sound signal. However, in fact not all speech is interesting or relevant, but only the user's speech. All other speech, e.g. in a noisy environment with several persons speaking, could be ignored and regarded as just noise.

[0033] The present inventor has realised that a microphone system having some kind of directional sensitivity could be used to discriminate sound emanating from different sources located in different directions. Sound not emanating

Exhibit 1009
Page 1498 of 2198

from the user can be declared as non-speech, and those signals do not have to be analysed with the conventional voice activity detectors.

[0034] The existing voice activity detectors may be conventional and are only referred to as a sound signal analyser in this application.

[0035] Generally, a microphone system having some kind of directional sensitivity can be used. Fig. 1 shows an example with at least two separate microphone elements.

[0036] A general mobile telephone is indicated at 1. The invention is equally applicable to other devices such as mobile radio terminals, pagers, communicators, electric organisers or smartphones. The common feature is that voice activity detection is employed, e.g. in connection with communicating speech or receiving voice commands by means of speech recognition.

[0037] In the simplest version, the microphone system comprises two microphones 2a and 2b. Suitably, they are located on a line directed in the calculated direction of an intended user's mouth. Suitably, the microphone elements are located at the lower edge of the mobile apparatus 1.

[0038] Fig. 2 shows a schematic diagram of the calculation of the direction of the sound source, typically the user's mouth 3. In the case of two microphones, only the angle to the line on which the microphone elements are located can be determined. In other words, the direction of the sound source is on a cone with a cone angle θ. To calculate the angle θ, first a cross-correlation between the two signals from the microphones 2a and 2b is made. The maximum indicates the time difference Δt between the two microphones 2a and 2b. The distance between the two microphones 2a and 2b is e.g. 20 millimetres. The angle θ is calculated as

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

[0039] Note that arccos is only defined for arguments between -1 and 1. If the time difference is negative, this means that the angle is greater than 90° and the sound emanates from behind the apparatus.

[0040] Suitably, the device is adapted to determine that all sounds with an angle θ less than a fixed angle α are emanating from the user. The threshold angle α may be set within a range of e.g. 10° to 30°, suitably at 25°.

[0041] In the case of three microphones, the direction of the sound source can be further determined to be at two points (e.g. on the above cone). The three microphone elements are suitably located in a plane directed in the general direction of the user's mouth. In fig. 1 microphone elements 2b, 2c and 2d are a possible set-up. The two microphone elements 2c and 2d at the front are located on a line perpendicular to the direction of the user's mouth, while the third microphone element 2b is located at the rear side.

[0042] In the case of four microphones (or more) detection of all direction angles may be calculated, provided that four microphone elements are located such that the fourth microphone is not located in the same plane as the three others, e.g. on a tetrahedron. A possible set-up is two microphone elements 2c and 2d at the front on the lower edge, while a third microphone element 2b is located at the rear side, and a fourth microphone element 2e is located at the front at a distance from the lower edge.

[0043] A similar microphone arrangement may be used in an accessory to a mobile apparatus, such as a hands-free kit or a telephone conference microphone system intended to be placed on a table. Apart from the microphone elements the logic circuitry may be located in the main/mobile apparatus. In this case the reception angle of the microphone system can be adjustable. This is useful e.g. when the microphone system is placed in a car, where the user can be seated either in the driver's seat or in the passenger's seat or even both the driver and the passenger may be speakers during the same call. The adjustment of the reception angle can be achieved mechanically or electronically, for example by beam forming or adaptation of the directional sensitivity of the microphone system.

[0044] To further enhance the sensitivity of the microphone system, directional microphone elements with a pattern having a maximum sensitivity in the direction of the user's mouth could be used.

[0045] In a further embodiment, one directional microphone element is used together with one or two other microphone elements (that may be non-directional). The directional microphone element is used to measure the sound pressure level relative to the other(s), thus removing the uncertainty in the direction of the sound source. Various combinations of directional microphone elements and non-directional microphone elements are possible.

[0046] The present invention leads to a voice activity detector having enhanced performance. With the present invention only one voice activity detector may be necessary throughout the whole signal path. This will in turn reduce the computational complexity, decreasing the load on the digital signal processors as well as improving the performance. It is especially favourable in environments with high background noise and noise with similar spectral properties as speech.

[0047] A person skilled in the art will realise that the invention may be realised with various combinations of hardware

Exhibit 1009
Page 1499 of 2198

and software. The scope of the invention is only limited by the claims below.

Claims

1. A device for voice activity detection comprising a sound signal analyser arranged to determine whether a sound signal comprises speech, comprising a microphone system (2a, 2b, 2c, 2d, 2e) arranged to discriminate sounds emanating from sources located in different directions from the microphone system, **characterised in that** the device is adapted to determine the direction of a sound source causing sound signals;
and is adapted to further analyse the sound to determine whether the sound signal comprises speech, if the sounds emanate from a first range of directions; but to decide that the sound signal does not comprise speech, if the sounds emanate from a second, different range of directions.

2. A device according to claim 1, **characterised in that** the first range of directions is directed in the direction of an intended user's mouth (3).

3. A device according to claim 2, **characterised in that** the microphone system comprises two microphone elements (2a, 2b) separated a distance and located on a line directed in the direction of an intended user's mouth (3).

4. A device according to claim 3, **characterised in that** the first range of directions is defined as all sounds falling inside a cone with a cone angle α, wherein $10° < \alpha < 30°$.

5. A device according to claim 4, **characterised in that** α is approximately 25°.

6. A device according to claim 2, **characterised in that** the microphone system comprises three microphone elements (2b, 2c, 2d) separated a distance and located in a plane directed in the direction of an intended user's mouth (3).

7. A device according to claim 6, **characterised in that** two (2c, 2d) of said three microphone elements are separated a distance and located on a line directed perpendicular to the direction of an intended user's mouth (3).

8. A device according to claim 2, **characterised in that** the microphone system comprises four microphone elements (2b, 2c, 2d, 2e), located such that the fourth microphone (2e) is not located in the same plane as the three others (2b, 2c, 2d).

9. A device according to any one of claims 1 to 8, **characterised in that** the microphone elements (2a, 2b, 2c, 2d, 2e) are directional with a pattern having maximal sensitivity in the direction of an intended user's mouth (3).

10. A device according to claim 1, **characterised in that** the microphone system comprises one directional microphone element together with one or more other microphone elements adapted to remove the uncertainty in the direction of the sound source.

11. A device according to claims 10, **characterised in that** the directional microphone element is adapted to measure the sound pressure level relative to the other microphone element.

12. A mobile apparatus, **characterised in that** it comprises a device as defined in any one of claims 1 to 11.

13. A mobile apparatus according to claim 12, **characterised in that** the microphone elements (2a, 2b, 2c, 2d) are located at the lower edge of the apparatus.

14. A mobile apparatus according to claim 12, **characterised in that** a plurality of microphone elements (2a, 2b, 2c, 2d) are located at the lower edge of the apparatus and at least one further microphone element (2e) is located at a distance from the lower edge.

15. A mobile apparatus according to any one of claims 12 to 14, **characterised in that** it is a mobile radio terminal, e.g. a mobile telephone (1), a pager, a communicator, an electric organiser or a smartphone.

16. An accessory for a mobile apparatus, **characterised in that** it comprises a device as defined in any one of claims 1 to 11.

Exhibit 1009
Page 1500 of 2198

17. An accessory according to claim 16, **characterised by** comprising adjustment means to adjust the first range of directions.

18. An accessory according to claim 16 or 17, **characterised in that** it is a hands-free kit.

19. An accessory according to claim 16 or 17, **characterised in that** it is a telephone conference microphone.

20. A method for voice activity detection, by the steps of:

receiving sound signals from a microphone system (2a, 2b, 2c, 2d, 2e) arranged to discriminate sounds emanating from sources located in different directions from the microphone system; and **characterized by**
determining the direction of the sound source causing the sound signals;
if the sounds emanate from a first range of directions, further analyse the sound to determine whether the sound signal comprises speech;
but if the sounds emanate from a second, different range of directions decide that the sound signal does not comprise speech.

21. A method according to claim 20, **characterised in that** the first range of directions is directed in the direction of an intended user's mouth (3).

22. A method according to claims 21, **characterised in that** the first range of directions is defined as all sounds falling inside a cone with a cone angle α, wherein $10° < \alpha < 30°$.

23. A method according to claims 22, **characterised in that** α is approximately 25°.

24. A method according to any one of claims 22 or 23, **characterised in that** the microphone system comprises at least two microphone elements (2a, 2b) located at a distance from each other and located on a line directed in the direction of an intended user's mouth (3), said two microphone elements being separated a distance d, wherein the direction to the sound source θ is calculated as

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

where
Δt is the time difference between the sounds from the two microphone elements, v is the velocity of sound.

25. A method according to claims 20, **characterised in that** one directional microphone element is used together with one or more other microphone elements to remove the uncertainty in the direction of the sound source.

26. A method according to claims 25, **characterised in that** the directional microphone element is used to measure the sound pressure level relative to the other microphone element.

Patentansprüche

1. Einrichtung zur Sprachaktivitätsdetektion, aufweisend einen Tonsignalanalysierer, der so ausgelegt ist, dass er feststellen kann, ob ein Tonsignal Sprache enthält, wobei die Einrichtung ein Mikrofonsystem (2a, 2b, 2c, 2d, 2e) aufweist, welches so ausgerichtet ist, dass es Töne, die von Quellen, welche sich in verschiedenen Richtungen vom Mikrofonsystem befinden, stammen unterscheidet, **dadurch gekennzeichnet, dass** die Einrichtung so ausgestaltet ist, dass sie die Richtung einer Tonquelle, welche Tonsignale aussendet, feststellen kann;
und so ausgestaltet ist, dass sie das Tonsignal weiter analysieren kann um festzustellen ob das Tonsignal Sprache enthält, falls Töne aus einem ersten Bereich von Richtungen stammen; und jedoch festzustellen, dass das Tonsignal keine Sprache enthält, falls die Töne aus einem zweiten anderen Bereich von Richtungen stammen.

2. Einrichtung nach Anspruch 1, **dadurch gekennzeichnet, dass** der erste Bereich von Richtungen auf die mutmaßliche Richtung des Mundes eines Nutzers (3) ausgerichtet ist.

Exhibit 1009
Page 1501 of 2198

3. Einrichtung nach Anspruch 2, **dadurch gekennzeichnet, dass** das Mikrofonsystem zwei Mikrofonelemente (2a, 2b) aufweist, die beabstandet sind und sich auf einer Linie befinden, die in Richtung der mutmaßlichen Position des Mundes des Nutzers (3) zeigt.

4. Einrichtung nach Anspruch 3, **dadurch gekennzeichnet, dass** der erste Bereich von Richtungen als alle Töne, die in einen Kegel mit Öffnungswinkel α fallen, definiert ist, wobei $10° < \alpha < 30°$.

5. Einrichtung nach Anspruch 4, **dadurch gekennzeichnet, dass** α ungefähr 25° ist.

6. Einrichtung nach Anspruch 2, **dadurch gekennzeichnet, dass** das Mikrofonsystem drei Mikrofonelemente (2b, 2c, 2d) aufweist, die beabstandet sind und sich in einer Ebene befinden, die in diejenige Richtung gerichtet ist, in der sich mutmaßlich der Mund des Nutzers befindet.

7. Einrichtung nach Anspruch 6, **dadurch gekennzeichnet, dass** zwei (2c, 2d) der drei Mikrofonelemente beabstandet sind und sich auf einer Linie befinden, die zu der Richtung der mutmaßlichen Position des Mundes des Nutzers (3) senkrecht steht.

8. Einrichtung nach Anspruch 2, **dadurch gekennzeichnet, dass** das Mikrofonsystem vier Mikrofonelemente (2b, 2c, 2d, 2e) aufweist, die so positioniert sind, dass das vierte Mikrofon (2e) sich nicht in der gleichen Ebene wie die drei anderen (2b, 2c, 2d) befindet.

9. Einrichtung nach einem der Ansprüche 1 bis 8, **dadurch gekennzeichnet, dass** die Mikrofonelemente (2a, 2d, 2c, 2d, 2e) gerichtet sind mit einem Muster, maximale Empfindlichkeit in Richtung der mutmaßlichen Position des Mundes des Nutzers hat.

10. Einrichtung nach Anspruch 1, **dadurch gekennzeichnet, dass** das Mikrofonsystem ein Richtmikrofonelement und zugleich ein oder mehrere andere Mikrofonelemente aufweist, die geeignet sind die Ungenauigkeit in der Richtung der Tonquelle zu beseitigen.

11. Einrichtung nach Anspruch 10, **dadurch gekennzeichnet, dass** das Richtmikrofonelement geeignet ist, den Schalldruckpegel relativ zum anderen Mikrofonelement zu messen.

12. Ein mobiles Gerät, **dadurch gekennzeichnet, dass** es eine in den Anspruches 1 bis 11 definierte Einrichtung aufweist.

13. Ein mobiles Gerät nach Anspruch 12, **dadurch gekennzeichnet, dass** sich die Mikrofonelemente (2a, ab, 2c, 2d) am unteren Rand des Gerätes befinden.

14. Mobiles Gerät nach Anspruch 12, **dadurch gekennzeichnet, dass** sich eine Mehrzahl von Mikrofonelementen (2a, 2b, 2c, 2d) am unteren Rand des Gerätes befindet und mindestens ein weiteres Mikrofonelement (2e) vom unteren Rand beabstandet angebracht ist.

15. Mobiles Gerät nach einem der Ansprüche 12 bis 14, **dadurch gekennzeichnet, dass** es ein mobiles Funk-Endgerät, beispielsweise ein Mobiltelefon (1), ein Pager, ein Communicator, ein elektronischer Kalender oder ein Smartphone ist.

16. Zubehör für ein mobiles Gerät, **dadurch gekennzeichnet, dass** es eine Einrichtung nach einem der Ansprüche 1 bis 11 aufweist.

17. Zubehör nach Anspruch 16, **dadurch gekennzeichnet, dass** es Einstellungselemente zur Einstellung des ersten Bereiches von Richtungen aufweist.

18. Zubehör nach Anspruch 16 oder 17, **dadurch gekennzeichnet, dass** es eine Freisprecheinrichtung ist.

19. Zubehör nach Anspruch 16 oder 17, **dadurch gekennzeichnet, dass** es ein Mikrofon für eine Telefonkonferenz ist.

20. Verfahren zur Erkennung von Sprachaktivität durch die Schritte: Empfangen eines Tonsignals von einem System von Mikrofonen (2a, 2b, 2c, 2d, 2e), das ausgerichtet ist, Töne die von Quellen stammen, welche sich in verschie-

Exhibit 1009
Page 1502 of 2198

denen Richtungen befinden, zu unterscheiden, **dadurch gekennzeichnet, dass** die Richtung der Tonquelle, die die Tonsignale ausstrahlt, festgestellt wird und, falls die Töne einem ersten Bereich vou Richtungen entspringen, ferner festgestellt wird, ob das Tonsignal Sprache aufweist; jedoch entschieden wird, dass das Tonsignal keine Sprache enthält wenn die Töne jedoch aus einem zweiten anderen Bereich von Richtungen entstammen.

21. Verfahren nach Anspruch 20, **dadurch gekennzeichnet, dass** der erste Bereich von Richtungen in Richtung der mutmaßlichen Position des Mundes des Nutzers ausgerichtet ist.

22. Verfahren nach Anspruch 21, **dadurch gekennzeichnet, dass** der erste Bereich von Richtungen definiert wird durch alle Töne, die in einen Kegel mit Öffnungswinkel α fallen, wobei gilt $10° < \alpha < 30°$.

23. Verfahren nach Anspruch 22, **dadurch gekennzeichnet, dass** α ungefähr 25° ist.

24. Verfahren nach einem der Ansprüche 22 und 23, **dadurch gekennzeichnet, dass** das Mikrofonsystem mindestens zwei Mikrofonelemente (2a, 2b) aufweist, die beabstandet sind und sich auf einer Linie befinden welche in die Richtung der mutmaßlichen Position des Mundes des Nutzers (3) zeigt, wobei die beiden Mikrofonelemente den Abstand d haben, wobei die Richtung zur Tonquelle θ berechnet wird als

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d},$$

wobei Δt die Zeitdifferenz zwischen den Tönen der beiden Mikrofone und v die Schallgeschwindigkeit ist.

25. Verfahren nach Anspruch 20, **dadurch gekennzeichnet, dass** ein Richtmikrofonelement und zugleich ein oder mehrere andere Mikrofonelemente benutzt werden, um die Ungenauigkeit in der Richtung der Tonquelle zu beseitigen.

26. Verfahren nach Anspruch 25, **dadurch gekennzeichnet, dass** das Richtmikrofonelement genutzt wird, um den Schalldruckpegel relativ zum anderen Mikrofonelement zu messen.

Revendications

1. Dispositif de détection de l'activité vocale, comprenant un analyseur de signal sonore agencé pour déterminer si un signal sonore comprend do la parole, comprenant un système de microphones (2a, 2b, 2c, 2d, 2e), agencé pour discriminer des sons provenant de sources situées dans différentes directions à partir du système de **caractérisé en ce que** le dispositif est adapté pour déterminer la direction d'une source sonore produisant des signaux sonores ; et
est en outre adapté pour analyser le son afin de déterminer si le signal sonore comprend de la parole, si les sons proviennent d'une première plage de directions ; mais pour décider que le signal sonore ne comprend pas de parole, si les sons proviennent d'une seconde plage, différente, de directions.

2. Dispositif selon la revendication 1, **caractérisé en ce que** la première plage de directions est dirigée dans la direction de la bouche (3) d'un utilisateur potentiel.

3. Dispositif selon la revendication 2, **caractérisé en ce que** le système de microphones comprend deux éléments microphoniques (2a, 2b) séparés d'une certaine distance et situés sur une droite dirigée dans la direction de la bouche (3) d'un utilisateur potentiel.

4. Dispositif selon la revendication 3, **caractérisé en ce que** la première plage de directions est définie comme regroupant tous les sons entrant dans un angle de cône α, avec $10° < \alpha < 30°$.

5. Dispositif selon la revendication 4, **caractérisé en ce que** α vaut approximativement 25 °.

6. Dispositif selon la revendication 2, **caractérisé en ce que** le système de microphones comprend trois éléments microphoniques (2b, 2c, 2d) séparés d'une certaine distance et situés dans un plan tourné dans la direction de la

Exhibit 1009
Page 1503 of 2198

bouche (3) d'un utilisateur potentiel.

7. Dispositif selon la revendication 6, **caractérisé en ce que** deux (2c, 2d) desdits trois éléments microphoniques sont séparés d'une certaine distance et situés sur une droite dirigée perpendiculairement à la direction de la bouche (3) d'un utilisateur potentiel.

8. Dispositif selon la revendication 2, **caractérisé en ce que** le système de microphones comprend quatre éléments microphoniques (2b, 2c, 2d, 2e) situés de telle manière que le quatrième microphone (2e) ne soit pas situé dans le même plan que les trois autres (2b, 2c, 2d).

9. Dispositif selon l'une quelconque des revendications 1 à 8, **caractérisé en ce que** les éléments microphoniques (2a, 2b, 2c, 2d, 2e) sont directionnels, avec un diagramme possédant une sensibilité maximale dans la direction de la bouche (3) d'un utilisateur potentiel.

10. Dispositif selon la revendication 1, **caractérisé en ce que** le système de microphones comprend un élément microphonique directionnel ainsi qu'un ou plusieurs autres éléments microphoniques adaptés pour lever l'incertitude sur la direction de la source sonore.

11. Dispositif selon la revendication 10, **caractérisé en ce que** l'élément microphonique directionnel est adapté pour mesurer le niveau de pression acoustique par rapport à l'autre élément microphonique.

12. Appareil mobile, **caractérisé en ce qu'**il comprend un dispositif selon l'une quelconque des revendications 1 à 11.

13. Appareil mobile selon la revendication 12, **caractérisé en ce que** les éléments microphoniques (2a, 2b, 2c, 2d) sont placés sur le bord intérieur de l'appareil.

14. Appareil mobile selon la revendication 12, **caractérise en ce qu'**une pluralité d'éléments microphoniques (2a, 2b, 2c, 2d) sont placés sur le bord inférieur de l'appareil, et **en ce qu'**au moins un autre élément microphonique (2e) est situé à distance du bord inférieur.

15. Appareil mobile selon l'une quelconque des revendications 12 à 14, **caractérisé en ce qu'**il s'agit d'un terminal mobile radio, par exemple un téléphone mobile (1), un boîtier de recherche de personnes, un communicateur, un organisateur électronique ou un téléphone multifonction.

16. Accessoire destiné à un appareil mobile, **caractérisé en ce qu'**il comprend un dispositif selon l'une quelconque des revendications 1 à 11.

17. Accessoire selon la revendication 16, **caractérisé en ce qu'**il comprend un moyen d'ajustement destiné à ajuster la première plage de directions.

18. Accessoire selon la revendication 16 ou 17, **caractérisé en ce qu'**il s'agit d'un kit mains libres.

19. Accessoire selon la revendication 16 on 17, **caractérisé en ce qu'**il s'agit d'un microphone de conférence téléphonique.

20. Procédé de détection de l'activité vocale, comprenant l'étape consistant à recevoir des signaux sonores d'un système de microphones (2a, 2b, 2c, 2d, 2e) agencé pour :

discriminer des sons provenant de sources situées dans différentes directions par rapport au système de microphones ;
et **caractérise par** les étapes consistant à :

déterminer la direction de la source sonore produisant les signaux sonores ;
si les sons proviennent d'une première plage de directions, analyser en outre le son pour déterminer si le signal sonore contient de la parole ;
mais si les sons proviennent d'une seconde plage, différente, de directions, décider que le signal sonore ne comprend pas de parole.

Exhibit 1009
Page 1504 of 2198

21. Procédé selon la revendication 20, **caractérisé en ce que** la première plage de directions est dirigée dans la direction de la bouche (3) d'un utilisateur potentiel.

22. Procédé selon la revendication 21, **caractérisé en** α, avec $10° < \alpha < 30°$.

23. Procédé selon la revendication 22, **caractérisé en ce que** α vaut approximativement 25°.

24. Procédé selon l'une quelconque des revendications 22 et 23, **caractérisé en ce que** le système de microphones comprend an moins deux éléments microphoniques (2a, 2b) situés à une certaine distance l'un de l'autre et situés sur une droite tournée dans la direction de la bouche (3) d'un utilisateur potentiel, lesdits deux éléments microphoniques étant séparés d'une distance d, la direction de la source sonore θ étant calculée à l'aide de l'expression :

$$\theta = \arccos \frac{\Delta t \cdot v}{2 \cdot d}$$

dans laquelle Δt est la différence de temps entre les sons des deux éléments microphoniques et v est la vitesse du son.

25. Procédé selon la revendication 20, **caractérisé en ce qu'**un élément microphonique directionnel est utilisé avec un ou plusieurs autres éléments microphoniques pour lever l'incertitude sur la direction de la source sonore.

26. Procédé selon la revendication 2 **caractérisé en ce que** l'élément microphonique directionnel est utilisé pour mesurer le niveau de pression acoustique par rapport à l'autre élément microphonique.

Exhibit 1009
Page 1505 of 2198

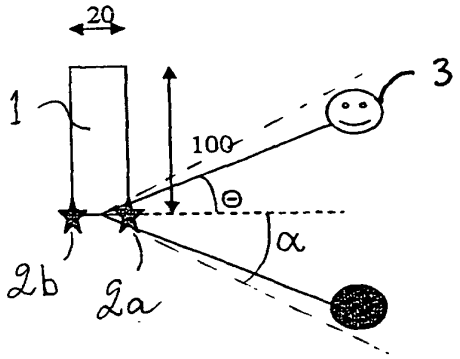


FIG 2

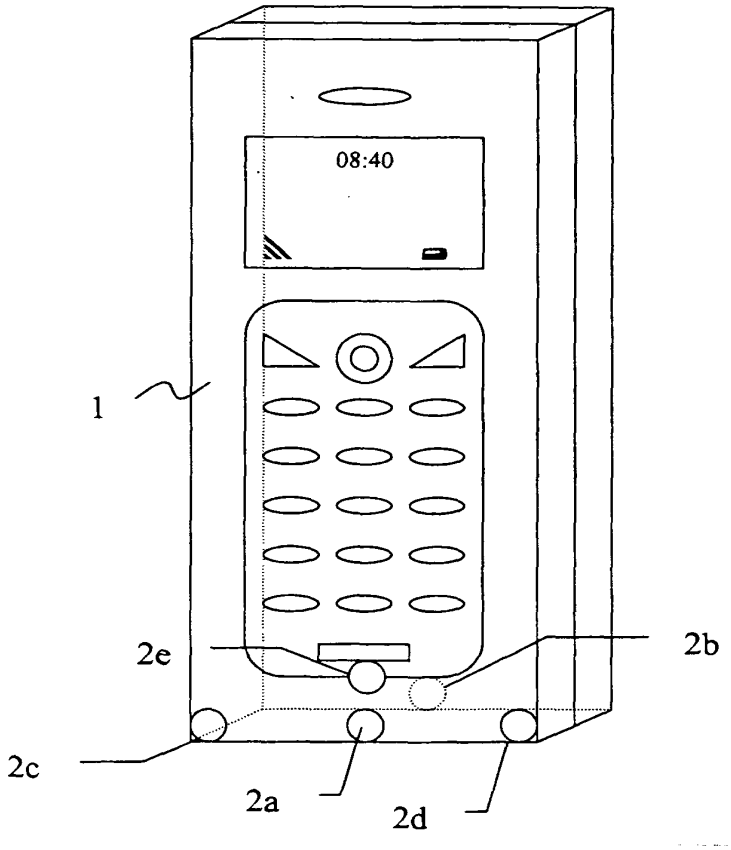


FIG 1

(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10－294989

(43)公開日　平成10年(1998)11月４日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｈ０４Ｒ　3/02		Ｈ０４Ｒ　3/02	
Ｇ１０Ｋ　11/178		Ｈ０３Ｈ　17/00	６０１Ｍ
Ｈ０３Ｈ　17/00	６０１	17/02	６０１Ｍ
17/02	６０１	21/00	
21/00		Ｈ０４Ｒ　1/10	１０１Ｚ

審査請求　未請求　請求項の数65　ＦＤ　（全 41 頁）　最終頁に続く

(21)出願番号　特願平9－116444

(22)出願日　平成９年(1997)４月18日

(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

(72)発明者　橋本　裕之
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　寺井　賢一
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　角張　勲
大阪府門真市大字門真1006番地　松下電器産業株式会社内

(74)代理人　弁理士　東島　隆治　（外１名）

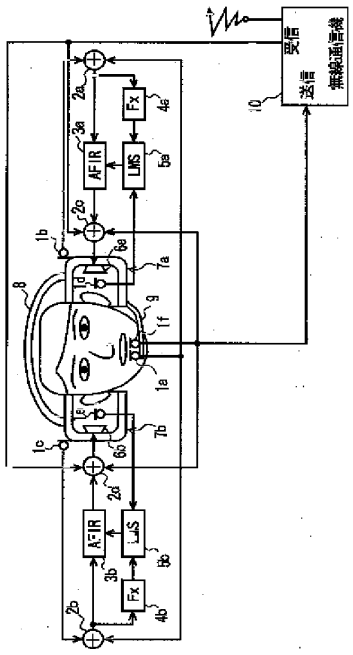
最終頁に続く

(54)【発明の名称】　騒音制御ヘッドセット

(57)【要約】

【課題】　ヘッドホンにおいて、外部の騒音を能動的に制御することにより騒音環境下での会話を可能とする騒音制御ヘッドセットを提供する。

【解決手段】　外部騒音を検出するマイクロホン１ａ～１ｃと、ヘッドホン７ａ～７ｂ内部の騒音を検出するマイクロホン１ｄ～１ｅと、自己音声を検出するマイクロホン１ｆと、マイクロホン１ａ～１ｃの検出信号をそれぞれ加算する加算器２ａ～２ｂと、その出力をリファレンス信号とする適応フィルタ３ａ～３ｂと、適応フィルタ３ａ～３ｂの出力とマイクロホン１ｆの出力と通信機１０の受信信号を加算する加算器２ｃ～２ｄと、加算器２ｃ～２ｄの出力を再生するヘッドホン７ａ～７ｂ内に設置されたスピーカ６ａ～６ｂと、スピーカ６ａ～６ｂとマイクロホン１ｄ～１ｅ間の伝達関数を補正するＦｘフィルタ４ａ～４ｂと、Ｆｘフィルタ４ａ～４ｂの出力とマイクロホン１ｄ～１ｅの出力により適応フィルタ３ａ～３ｂの係数を演算して更新するＬＭＳ演算器５ａ～５ｂとから構成されている。



【特許請求の範囲】
【請求項１】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、
前記音声検出器の近傍に取り付けられ、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記音声検出器の出力と、他の人との通信を行う通信手段からの受信信号とを加算する第１の出力信号加算器と、
前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、
前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、
を有し、但し、
前記音声検出器は使用者の口の方向に強い指向性を有する単一指向性あるいは双指向性を有し、その出力信号を前記通信手段の送信信号とし、
前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされている、よう構成したことを特徴とする騒音制御ヘッドセット。
【請求項２】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項１記載の騒音制御ヘッドセット。
【請求項３】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられている、
ことを特徴とする請求項１記載の騒音制御ヘッドセット。
【請求項４】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、
前記音声検出器の出力により音声の有無を判断する音声判別器と、
他の人との通信を行なう通信手段の受信信号の有無を判断する受信信号判別器と、
前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号とを加算する第１の出力信号加算器と、
前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、

前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器とを有し、但し前記音声検出器は使用者の口の方向に強い指向性を有する単一指向性あるいは双指向性を有し、前記音声判別器により音声が発生していると判断した間だけその出力信号を前記通信手段の送信信号および前記第１、第２の出力信号加算器への入力信号として用いており、
前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は、右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、
前記第１および第２の係数更新器は前記音声判別器により音声が発生していると判断した場合、あるいは前記受信信号判別器により受信信号が存在すると判断した場合の少なくともいずれかの場合の間だけ、前記第１および第２の適応フィルタの係数更新を停止するよう構成したことを特徴とする騒音制御ヘッドセット。

【請求項５】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項４記載の騒音制御ヘッドセット。

【請求項６】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられている、
ことを特徴とする請求項４記載の騒音制御ヘッドセット。

【請求項７】　音声判別器が音声を検出している期間、通信手段の送信電源が立ち上がっているように制御するよう構成したことを特徴とする請求項４記載の騒音制御ヘッドセット。

【請求項８】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、
前記音声検出器の出力により音声の有無を判断する音声判別器と、
他の人との通信を行なう通信手段の受信信号の有無を判断する受信信号判別器と、
前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、
前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、
前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力とから前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、
前記音声検出器の出力を遅延させる遅延器と、
前記第１の騒音検出器の出力信号を処理する第３の適応

フィルタと、
前記遅延器の出力から前記第３の適応フィルタの出力を減算する減算器と、
前記第１の騒音検出器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、
前記減算器の出力を通過させるかさせないかを制御するミュート回路とを有し、但し、
前記音声判別器は音声が発生していると判断した間だけ前記ミュート回路を導通状態にしてその出力を前記第１および第２の出力信号加算器に入力し、第１の出力信号加算器において第１の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、また第２の出力信号加算器において第２の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、さらにミュート回路の出力を前記通信手段の送信信号としており、
前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は、右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、
前記第１および第２の係数更新器は前記音声判別器により音声が発生していると判断された場合、あるいは前記受信信号判別器により受信信号が存在すると判断した場合の少なくともいずれかの場合の間だけ、前記第１および第２の適応フィルタの係数更新を停止し、
前記第３の係数更新器は前記音声判別器により音声が発生していると判断した間だけ、前記第３の適応フィルタの係数更新を停止するよう構成したことを特徴とする騒音制御ヘッドセット。
【請求項９】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項８記載の騒音制御ヘッドセット。
【請求項１０】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられている、
ことを特徴とする請求項８記載の騒音制御ヘッドセット。
【請求項１１】　前記第１の騒音検出器は無指向性であり、前記音声検出器は指向性を有する、
ことを特徴とする請求項８記載の騒音制御ヘッドセット。
【請求項１２】　前記音声検出器の振動板直径が前記第１の騒音検出器の振動板直径の２倍以上である、
ことを特徴とする請求項８記載の騒音制御ヘッドセット。
【請求項１３】　前記音声検出器は使用者の口の方向へ指向性を有するように取り付けられており、
前記第１の騒音検出器は無指向性であり、かつ前記音声検出器の指向性を有する方向とは正反対の方向に音孔が向くように取り付けられており、
前記音声検出器の指向性を有する方向に垂直である軸上で前記音声検出器の真横に前記第１の検出器を設置したことを特徴とする請求項８記載の騒音制御ヘッドセット。
【請求項１４】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔を吸音材で覆った、
ことを特徴とする請求項１３記載の騒音制御ヘッドセット。
【請求項１５】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔と、前記正反対の面と同じ面にある前記第１の騒音検出器の音孔を吸音材で覆った、
ことを特徴とする請求項１３記載の騒音制御ヘッドセット。
【請求項１６】　前記音声判別器が音声を検出している期間、通信手段の送信電源が立ち上がっているように制御することを特徴とする請求項８記載の騒音制御ヘッドセット。
【請求項１７】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、
前記音声検出器の出力により音声の有無を判断する音声判別器と、
他の人との通信を行なう通信手段の受信信号の有無を判断する受信信号判別器と、
前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、

前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、
前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、
前記音声検出器の出力を遅延させる第１の遅延器と、
前記第１の騒音検出器の出力を遅延させる第２の遅延器と、
前記第２の遅延器の出力信号を処理する第３の適応フィルタと、
前記第１の遅延器の出力から前記第３の適応フィルタの出力を減算する減算器と、
前記第２の遅延器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、
前記減算器の出力を通過させるかさせないかを制御するミュート回路と、
前記ミュート回路の動作を制御するミュート制御回路と、
を有し、但し、
前記音声判別器は音声が発生していると判断した間だけ前記ミュート制御回路に音声発生信号を入力し、ミュート制御回路はその信号を入力している間と音声発生信号がなくなってから一定時間内は、前記ミュート回路を導通状態にするミュート制御信号を発生させ、その信号をミュート回路に入力している間は導通状態にしてミュート回路の出力を前記第１および第２の出力信号加算器に入力し、第１の出力信号加算器において第１の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、また第２の出力信号加算器において第２の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、さらにミュート回路の出力を前記通信手段の送信信号とし、
前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は、右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、
前記第１および第２の係数更新器は前記音声判別器により音声が発生していると判断された場合、あるいは前記受信信号判別器により受信信号が存在すると判断された場合の少なくともいずれかの間だけ、前記第１および第２の適応フィルタの係数更新を停止し、
前記第３の係数更新器は前記音声判別器により音声が発生していると判断された間だけ、前記第３の適応フィルタの係数更新を停止する、
よう構成したことを特徴とする騒音制御ヘッドセット。

【請求項１８】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、
前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項１７記載の騒音制御ヘッドセット。

【請求項１９】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、
前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられている、
ことを特徴とする請求項１７記載の騒音制御ヘッドセット。

【請求項２０】　前記第１の騒音検出器は無指向性であり、前記音声検出器は指向性を有することを特徴とする請求項１７記載の騒音制御ヘッドセット。

【請求項２１】　前記音声検出器の振動板直径が、前記第１の騒音検出器の振動板直径の２倍以上である、
ことを特徴とする請求項１７記載の騒音制御ヘッドセット。

【請求項２２】　前記音声検出器は使用者の口の方向へ指向性を有するように取り付けられており、
前記第１の騒音検出器は無指向性であり、かつ前記音声検出器の指向性を有する方向とは正反対の方向に音孔が向くように取り付けられており、
前記音声検出器の指向性を有する方向に垂直である軸上でしかも前記音声検出器の真横に前記第１の検出器を設置した、

ことを特徴とする請求項１７記載の騒音制御ヘッドセット。

【請求項２３】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔を吸音材で覆った、

ことを特徴とする請求項２２記載の騒音制御ヘッドセット。

【請求項２４】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔と、前記正反対の面と同じ面にある前記第１の騒音検出器の音孔を吸音材で覆った、

ことを特徴とする請求項２２記載の騒音制御ヘッドセット。

【請求項２５】　前記ミュート制御回路からのミュート制御信号が発生している間だけ通信手段の送信電源が立ち上がっているように制御するよう構成したことを特徴とする請求項１７記載の騒音制御ヘッドセット。

【請求項２６】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、

前記音声検出器の出力により音声の有無を判断する音声判別器と、

他の人との通信を行なう通信手段の受信信号の有無を判断する受信信号判別器と、

前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、

前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、

前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、

前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、

前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、

前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、

前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、

前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、

前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、

前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、

前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、

前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、

前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、

前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、

前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、

前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、

前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、

前記音声検出器の出力を遅延させる遅延器と、

前記第１の騒音検出器の出力信号を処理する第３の適応フィルタと、

前記第２の騒音検出器の出力信号を処理する第４の適応フィルタと、

前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、

前記遅延器の出力から前記第３の適応フィルタの出力と前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する減算器と、

前記第１の騒音検出器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、

前記第２の騒音検出器の出力と前記減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、

前記第３の騒音検出器の出力と前記減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、

前記減算器の出力を通過させるかさせないかを制御するミュート回路と、　を有し、但し、

前記音声判別器は音声が発生していると判断した間だけ前記ミュート回路を導通状態にしてその出力を前記第１および第２の出力信号加算器に入力し、第１の出力信号加算器において第１の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、また第２の出力信号加算器において第２の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、さらにミュート回路の出力を前記通信手段の送信信号としており、

前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、

前記第２の伝達関数補正器の係数は、右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、

前記第１および第２の係数更新器は前記音声判別器により音声が発生していると判断した場合、あるいは前記受信信号判別器により受信信号が存在すると判断した場合の少なくともいずれかの場合の間だけ、前記第１および第２の適応フィルタの係数更新を停止し、

前記第３の係数更新器は前記音声判別器により音声が発生していると判断した間だけ、前記第３の適応フィルタの係数更新を停止するよう構成したことを特徴とする騒音制御ヘッドセット。

【請求項２７】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられていることを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項２８】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられている、
ことを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項２９】　前記第４および第５の係数更新器は、前記音声判別器により音声が発生していると判断した間だけ、前記第４および第５の適応フィルタの係数更新を停止するよう構成したことを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項３０】　前記第１の騒音検出器は無指向性であり、前記音声検出器は指向性を有することを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項３１】　前記音声検出器の振動板直径が、前記第１の騒音検出器の振動板直径の２倍以上であることを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項３２】　前記音声検出器は使用者の口の方向へ指向性を有するように取り付けられており、
前記第１の騒音検出器は無指向性であり、かつ前記音声検出器の指向性を有する方向とは正反対の方向に音孔が向くように取り付けられており、
前記音声検出器の指向性を有する方向に垂直である軸上でしかも前記音声検出器の真横に前記第１の検出器を設置したことを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項３３】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔を吸音材で覆ったことを特徴とする請求項３２記載の騒音制御ヘッドセット。

【請求項３４】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔と、前記正反対の面と同じ面にある第１の騒音検出器の音孔を吸音材で覆ったことを特徴とする請求項３２記載の騒音制御ヘッドセット。

【請求項３５】　前記音声判別器が音声を検出している期間、前記通信手段の送信電源が立ち上がっているように制御するよう構成したことを特徴とする請求項２６記載の騒音制御ヘッドセット。

【請求項３６】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、
前記音声検出器の出力により音声の有無を判断する音声判別器と、
前記通信手段の受信信号の有無を判断する受信信号判別器と、
前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、
前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、
前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、
前記音声検出器の出力を遅延させる第１の遅延器と、
前記第１の騒音検出器の出力信号を処理する第３の適応フィルタと、
前記第２の騒音検出器の出力信号を処理する第４の適応

フィルタと、
前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、
前記第１の遅延器の出力から前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する第１の減算器と、
前記第１の減算器の出力を遅延させる第２の遅延器と、
前記第２の遅延器の出力から前記第３の適応フィルタの出力を減算する第２の減算器と、
前記第１の騒音検出器の出力と前記第２の減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、
前記第２の騒音検出器の出力と前記第１の減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、
前記第３の騒音検出器の出力と前記第１の減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、
前記第２の減算器の出力を通過させるかさせないかを制御するミュート回路と、
を有し、但し、
前記音声判別器は音声が発生していると判断した間だけ前記ミュート回路を導通状態にしてその出力を前記第１および第２の出力信号加算器に入力し、第１の出力信号加算器において第１の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、また第２の出力信号加算器において第２の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、さらにミュート回路の出力を前記通信手段の送信信号としており、
前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、
前記第１および第２の係数更新器は前記音声判別器により音声が発生していると判断した場合、あるいは前記受信信号判別器により受信信号が存在すると判断した場合の少なくともいずれかの場合の間だけ、前記第１および第２の適応フィルタの係数更新を停止し、
前記第３の係数更新器は、前記音声判別器により音声が発生していると判断した間だけ、前記第３の適応フィルタの係数更新を停止するよう構成したことを特徴とする騒音制御ヘッドセット。
【請求項３７】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項３８】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられていることを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項３９】　前記第４および第５の係数更新器は、前記音声判別器により音声が発生していると判断した間だけ、前記第４および第５の適応フィルタの係数更新を停止する、ことを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項４０】　前記第１の騒音検出器は無指向性であり、前記音声検出器は指向性を有する、　ことを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項４１】　前記音声検出器の振動板直径が、前記第１の騒音検出器の振動板直径の２倍以上である、
ことを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項４２】　前記音声検出器は使用者の口の方向へ指向性を有するように取り付けられており、
前記第１の騒音検出器は無指向性であり、かつ前記音声検出器の指向を有する方向とは正反対の方向に音孔が向くように取り付けられており、
前記音声検出器の指向性を有する方向に垂直である軸上で前記音声検出器の真横に前記第１の検出器を設置した、
ことを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項４３】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔を吸音材で覆った、
ことを特徴とする請求項４２記載の騒音制御ヘッドセット。
【請求項４４】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔と、前記正反対の面と同じ面にある前記第１の騒音検出器の音孔を吸音材で覆った、　ことを特徴とする請求項４２記載の騒音制御ヘッドセット。
【請求項４５】　前記音声判別器が音声を検出している期間、通信手段の送信電源が立ち上がっているように制御する、
ことを特徴とする請求項３６記載の騒音制御ヘッドセット。
【請求項４６】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近

傍に取り付けられている音声検出器と、
音声検出器の出力により音声の有無を判断する音声判別器と、
他の人との通信を行う通信手段の受信信号の有無を判断する受信信号判別器と、
前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側に取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、
前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、
前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、
前記音声検出器の出力を遅延させる第１の遅延器と、
前記第１の騒音検出器の出力を遅延させる第２の遅延器と、
前記第２の遅延器の出力信号を処理する第３の適応フィルタと、
前記第２の騒音検出器の出力信号を処理する第４の適応フィルタと、
前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、
前記第１の遅延器の出力から前記第３の適応フィルタの出力と前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する減算器と、
前記第２の遅延器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、
前記第２の騒音検出器の出力と前記減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、
前記第３の騒音検出器の出力と前記減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、
前記減算器の出力を通過させるかさせないかを制御するミュート回路と、
前記ミュート回路の動作を制御するミュート制御回路と、
を有し、但し、
前記音声判別器は音声が発生していると判断した間だけ前記ミュート制御回路に音声発生信号を入力し、ミュート制御回路はその信号を入力している間と音声発生信号がなくなってから一定時間内は、前記ミュート回路を導通状態にするミュート制御信号を発生させ、その信号をミュート回路に入力している間はミュート回路を導通状態にしてミュート回路の出力を前記第１および第２の出力信号加算器に入力し、第１の出力信号加算器において第１の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、また第２の出力信号加算器において第２の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、さらにミュート回路の出力を前記通信手段の送信信号とし、
前記第１の伝達関数補正器の係数は左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、
前記第１および第２の係数更新器は前記音声判別器により音声が発生していると判断した場合、あるいは前記受信信号判別器により受信信号が存在すると判断した場合の少なくともいずれかの場合の間だけ、前記第１および第２の適応フィルタの係数更新を停止し、
前記第３の係数更新器は、前記音声判別器により音声が発生していると判断した間だけ、前記第３の適応フィルタの係数更新を停止するよう構成したことを特徴とする騒音制御ヘッドセット。
【請求項４７】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着し状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネット外側を、使用の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上

部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項４８】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着し状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割し左上部の領域に取り付けられている、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項４９】　第４および第５の係数更新器は、前記音声判別器により音声が発生していると判断した間だけ、前記第４および第５の適応フィルタの係数更新を停止する、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項５０】　前記第１の騒音検出器は無指向性であり、前記音声検出器は指向性を有する、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項５１】　前記音声検出器の振動板直径が、前記第１の騒音検出器の振動板直径の２倍以上である、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項５２】　前記音声検出器は使用者の口の方向へ指向性を有するように取り付けられており、
前記第１の騒音検出器は無指向性であり、かつ前記音声検出器の指向性を有する方向とは正反対の方向に音孔が向くように取り付けられており、
前記音声検出器の指向性を有する方向に垂直である軸上でしかも前記音声検出器の真横に前記第１の騒音検出器を設置した、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項５３】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔を吸音材で覆った、
ことを特徴とする請求項５２記載の騒音制御ヘッドセット。
【請求項５４】　前記音声検出器の使用者の口に対向する面とは正反対の面にある音孔と、同じ面にある第１の騒音検出器の音孔を吸音材で覆った、
ことを特徴とする請求項５２記載の騒音制御ヘッドセット。
【請求項５５】　前記ミュート制御回路からのミュート制御信号が発生している間だけ通信手段の送信電源が立ち上がっているように制御する、
ことを特徴とする請求項４６記載の騒音制御ヘッドセット。
【請求項５６】　左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、
前記音声検出器の出力により音声の有無を判断する音声判別器と、
他の人との通信を行なう通信手段の受信信号の有無を判断する受信信号判別器と、
前記音声検出器の近傍に取り付けられており、主として周囲の騒音を検出する第１の騒音検出器と、
前記ヘッドホンの左側キャビネットの外側取り付けられた第２の騒音検出器と、
前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、
前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、
前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、
前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、
前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、
前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、
前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、
前記ヘッドホンの右側キャビネットの外側に取り付けられた第３の騒音検出器と、
前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、
前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、
前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、
前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、
前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、
前記右側ヘッドホン内部に取り付けられた第２の誤差検出器と、
前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、
前記音声検出器の出力を遅延させる第１の遅延器と、
前記第１の騒音検出器の出力を遅延させる第２の遅延器と、
前記第２の遅延器の出力信号を処理する第３の適応フィルタと、
前記第２の騒音検出器の出力信号を処理する第４の適応

フィルタと、
前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、
前記第１の遅延器の出力から前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する第１の減算器と、
前記第１の減算器の出力を遅延させる第３の遅延器と、
前記第３の遅延器の出力から前記第３の適応フィルタの出力を減算する第２の減算器と、
前記第２の遅延器の出力と前記第２の減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、
前記第２の騒音検出器の出力と前記第１の減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、
前記第３の騒音検出器の出力と前記第１の減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、
前記第２の減算器の出力を通過させるかさせないかを制御するミュート回路と、
前記ミュート回路の動作を制御するミュート制御回路とを有し、但し、
前記音声判別器は、音声が発生していると判断した間だけ前記ミュート制御回路に音声発生信号を入力し、ミュート制御回路はその信号を入力している間と音声発生信号がなくなってから一定時間内は、前記ミュート回路を導通状態にするミュート制御信号を発生させ、その信号をミュート回路に入力している間はミュート回路を導通状態にしてミュート回路の出力を前記第１および第２の出力信号加算器に入力し、第１の出力信号加算器において第１の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、また第２の出力信号加算器において第２の適応フィルタの出力と通信手段からの受信信号とミュート回路の出力が加算され、さらにミュート回路の出力を前記通信手段の送信信号とし、
前記第１の伝達関数補正器の係数は、左側ヘッドホンから第１の誤差検出器までの伝達関数に近似するようになされており、
前記第２の伝達関数補正器の係数は、右側ヘッドホンから第２の誤差検出器までの伝達関数に近似するようになされており、
前記第１および第２の係数更新器は、前記音声判別器により音声が発生していると判断した場合、あるいは前記受信信号判別器により受信信号が存在すると判断した場合の少なくともいずれかの間だけ、前記第１および第２の適応フィルタの係数更新を停止し、
前記第３の係数更新器は、前記音声判別器により音声が発生していると判断した間だけ、前記第３の適応フィルタの係数更新を停止するよう構成したことを特徴とする騒音制御ヘッドセット。

【請求項５７】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第２の騒音検出器は、左側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、左側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した右上部の領域に取り付けられている、
ことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【請求項５８】　前記音声検出器が使用者の口元に来るように騒音制御ヘッドセットを装着した状態において、前記第３の騒音検出器は、右側ヘッドホンのキャビネットの外側を、使用者の側方から見た場合に、右側ヘッドホンの中心点から水平軸と垂直軸により左上部と左下部と右上部および右下部に４分割した左上部の領域に取り付けられている、
ことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【請求項５９】　前記第４および第５の係数更新器は、前記音声判別器により音声が発生していると判断された間だけ、前記第４および第５の適応フィルタの係数更新を停止する、
ことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【請求項６０】　前記第１の騒音検出器は無指向性であり、前記音声検出器は指向性を有する、
ことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【請求項６１】　前記音声検出器の振動板直径が、前記第１の騒音検出器の振動板直径の２倍以上である、
ことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【請求項６２】　前記音声検出器は使用者の口の方向へ指向性を有するように取り付けられており、
前記第１の騒音検出器は無指向性であり、かつ前記音声検出器の指向性を有する方向とは正反対の方向に音孔が向くように取り付けられており、
前記音声検出器の指向性の方向に垂直である軸上でしかも前記音声検出器の真横に前記第１の検出器を設置した、
ことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【請求項６３】　前記音声検出器の使用者の口に対向する方向面とは正反対の面にある音孔を吸音材で覆った、
ことを特徴とする請求項６２記載の騒音制御ヘッドセット。
【請求項６４】　前記音声検出器の使用者の口に対向する方向面とは正反対の面にある音孔と、前記正反対の面と同じ面にある前記第１の騒音検出器の音孔を吸音材で覆ったことを特徴とする請求項６２記載の騒音制御ヘッ

ドセット。
【請求項６５】　前記ミュート制御回路からのミュート制御信号が発生している間だけ通信手段の送信電源が立ち上がっているように制御するよう構成したことを特徴とする請求項５６記載の騒音制御ヘッドセット。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は能動騒音制御に関するものである。
【０００２】
【従来の技術】近年、スピーカを用いて、例えば空調用のダクト内の騒音を能動的に消音する制御法が実用化されつつある。また工場などで用いられる、作業者の聴覚保護などを目的とした騒音制御型のヘッドホン（以後、騒音制御ヘッドホンと称する）が種種提案されている。
【０００３】以下、図面を参照しながら従来の騒音制御ヘッドホンについて説明する。従来例としては、特開平３－９６１９９号公報に示されている。図２５は、この従来例の騒音制御ヘッドホンのブロック図を示すものである。図２５において、イヤホン２５の外筐にマイクロホン２３が取り付けられている。また、イヤホン２５の外耳道２８Ａに面する内側部にはマイクロホン２４が取り付けられている。イヤホン２５が耳２８に装着された状態では、マイクロホン２３では矢印Ｎで示す外部騒音が収音され、マイクロホン２４では外耳道２８Ａ内の音が収音される。
【０００４】マイクロホン２３及び２４の出力は適応処理システム２６に供給される。適応処理システム２６では、マイクロホン２３からイヤホン２５までの伝達特性が、マイクロホン２４で検出される外部騒音Ｎが外耳道２８Ａに到達するまでの伝達特性と同様な特性となるように同定される。適応処理システム２６の出力はアンプ２７を介してイヤホン２５に供給される。イヤホン２５により、アンプ２７の出力信号が音に変換され、この音が外耳道２８Ａに供給される。外耳道２８Ａ内には、外部から騒音Ｎが入ってくるが、適応処理システム２６の働きにより、この騒音成分が最小になるように制御される。
【０００５】すなわち、図２６に示すような閉ループ適応システムが図２５の装置により実現されて、騒音成分を除去するものである。図２６において、適応処理システム２６は、信号処理システム３０の特性と同様の特性で動作するように働く。すなわち、信号Ｓが適応処理システム２６と信号処理システム３０に供給される。減算器２９で、適応処理システム２６の出力ｙと信号処理システム３０の出力ｄとの誤差ε（ε＝ｄ－ｙ）が求められる。適応処理システム２６にはこの誤差成分εが入力され、この誤差εが最小となるように制御される。適応処理が完了した後は、信号処理システム３０の伝達関数が適応処理システム２６の伝達関数と等しいという意味で、両システムは同定される。
【０００６】図２６の適応処理システムと図２５の騒音制御ヘッドホン装置とを対比して説明すると、信号Ｓが外部騒音Ｎと対応し、誤差成分εがマイクロホン２４で検出される信号と対応し、信号処理システム３０がマイクロホン２３からマイクロホン２４までの伝達系と対応したものである。誤差分εに対応する外耳道２８Ａの音圧は、マイクロホン２４で検出され、このマイクロホン２４の出力に基づいて適応処理システム２６が制御される。これにより、外耳道２８Ａ内における騒音成分の音圧が最小となる。
【０００７】以上のように、適応処理システム２６により、マイクロホン２３からヘッドホン２５までの伝達特性が、外部騒音が外耳道２８Ａに入るときの遮音特性と同様な特性となるように常に同定されるため、外耳道２８Ａ内の騒音成分を打ち消すことができる。よって、イヤホン２５の装着の仕方が変わったり、イヤホン２５やマイクロホン２３が経年変化を起こしても遮音特性が悪化しない。
【０００８】次に特開平６－６７６７９号公報に示されている他の従来の騒音制御ヘッドホンについて説明する。図２７は、上記従来の騒音制御ヘッドホンのブロック図を示すものである。図２７において、左右一対で構成される個人携帯用のヘッドホン３１の左右の筐体の内部には、それぞれ誤差検出用のマイクロホン３２と制御用のスピーカ３３が設けられている。これらマイクロホン３２とスピーカ３３は、左右のそれぞれＬチャンネル及びＲチャンネルの誤差信号線３２ａと制御信号線３３ａを介して、ヘッドホン３１と一体構成となっている信号処理装置３４に接続されている。振動源または音源となる機械３８の近傍には、機械３８から発生する振動または音を検出するための検出センサ３５と、この検出センサ３５の出力をそれぞれ所定のレベルに増幅する増幅器３５ａが一組、または、必要に応じて複数組が、機械３８の振動または発生音と関連度の高い参照信号を生成するため、予めよく調査された位置に設けられている。
【０００９】検出センサ３５が複数個設けられる場合は、機械３８を代表する参照信号とするため、それぞれの増幅器３５ａの出力をミキサ３６において混合して、混合の結果生じたミックス信号を機械３８の参照信号とする。この参照信号は機械３８近傍に設けられた複数の出力ジャック３６ａに出力されている。ヘッドホン３１を機械３８の近くで使用する場合は、信号処理装置３４の参照信号入力用の接続ケーブル３４ａのプラグ３４ｂを出力ジャック３６ａに装着して、参照信号を信号処理装置３４内に取り込む。この参照信号に基づいてマイクロホン３２の出力が最小となるような信号をスピーカ３３に与えて、能動的消音を行う構成になっている。なお、ここでは、参照信号はＬ、Ｒの両チャンネルに共用されるようになっている。

【００１０】以上のように、検出センサ３５を音源である機械３８の近傍に複数個適切に配置して設け、さらに、これらの出力を適正なレベルでミキシング（混合すること）しているので、耳に入る音と関連性が高い参照信号が得られ、この参照信号を使用することで高い消音効果が得られる。また、検出センサ３５は音源である機械３８の近傍に設置されているので、参照信号が信号処理装置３４内に取り込まれる時間と、機械３８の音が直接耳に入る時間との差が十分大きい。したがって、信号処理装置３４が消音用の打ち消し信号を算出して、スピーカ３３に与えるために必要な処理時間を確保できる。このことによりランダム音に対しても適確に消音できる。

【００１１】図２８は図２７のシステムにおける参照信号を無線で伝送する方式を示しており、人が携帯する信号処理装置３４側に無線受信機３７ａを設け、ミキサ３６の出力側に無線送信機３７ｂを設けている。これによって機械３８が発する振動や音を代表する参照信号は無線送信機３７ｂを介して、所定の強さの電界内にあるすべての無線受信機３７ａに送信されて利用可能となる。このため、図２７のような信号処理装置３４を接続ケーブル３４ａを介してミキサ３６の出力に接続する手間が省け、しかも人の行動の自由が確保できる。

【００１２】さらに図２９は音源となる機械が複数存在する場合を示している。それぞれの機械３８－１～３８－ｎに対応する同一周波数の無線送信機３７ｂを設け、各機械の参照信号をそれぞれの無線送信機３７ｂから所定の強さの電界内にあるすべての無線受信機３７ａに送信するものである。人が携帯する信号処理装置３４に接続される無線受信機３７ａは、各機械３８－１～３８－ｎから送られる複数の参照信号をミキシングして受信して、このミキシングされた受信信号に基づいて、スピーカ３３に送る制御信号を各参照信号を含んだ形で生成し、消音制御を行う。従って複数の音源の機械３８－１～３８－ｎから音が聞こえてくるような場合に、非常に効果がある。

【００１３】

【発明が解決しようとする課題】しかしながら図２５に示す従来例では、マイクロホン２３とマイクロホン２４の距離が非常に近いため、マイクロホン２３が外部騒音を検出してからその騒音がマイクロホン２４に到達するまでの時間が短くなる。また使用者はヘッドホン２５を装着した状態で移動するために騒音源とマイクロホン２３とマイクロホン２４の位置関係が必ずしも一直線上にあるとは限らない。このため、適応処理システム２６が正常に動作するために必要な、当技術分野ではよく知られているので詳細な説明を省略するが、収音した外部騒音とアンプからイヤホンに供給されるべき信号との間の因果律が成立しない場合が生じる。従って広帯域ランダム騒音の制御は困難であった。また外部騒音が複数存在する場合や、残響時間の長い反射性の部屋あるいは大きな部屋の場合には、１個のマイクロホンでは複数の音源からの音を平等には検出できず、マイクロホン２３とマイクロホン２４で検出される騒音信号の関連度が低下して消音効果が十分に得られないという問題があった。さらに、使用者の音声を検出するマイクロホンが付属していないために、複数の使用者間での会話ができないという問題点もあった。

【００１４】次に図２７、図２８及び図２９に示す従来例では、騒音源である機械３８の近傍に検出センサ３５を設置したことにより、使用者がヘッドホン３１を装着した状態で移動すると信号処理装置３４で実現すべき特性が大きく変化し、その動作が不安定なものとなってしまう。また信号処理装置３４を構成するフィルタが有限長のＦＩＲフィルタ（有限インパルス応答フィルタ、Finite impulse response filter)であったり、次数の小さなＩＩＲフィルタ(無限インパルス応答フィルタ、Infinite impulse response filter)である場合、検出センサ３５からマイクロホン３２までの伝送路のインパルスに対するレスポンスを十分な程度に実現できない。そのため、消音効果が劣化する問題がある。また残響時間の長い反射性の部屋あるいは大きな部屋の場合にも、有限長のＦＩＲフィルタや、次数の小さなＩＩＲフィルタを用いたのでは、検出センサ３５からマイクロホン３２までの伝送路のインパルスに対するレスポンスを十分な程度に実現できず、消音効果が劣化する問題がある。さらに、音声を検出するマイクロホンが付属していないために、複数の使用者間での会話ができないという問題点もあった。

【００１５】本発明の第１の目的は、ヘッドホンを音声検出器が一体に構成された騒音制御ヘッドセットを装着した状態で移動することによる騒音源と騒音制御ヘッドセットの位置を示す制御点の相対位置に変化があった場合や、外部騒音が複数存在する場合、あるいは残響時間の長い反射性の部屋あるいは大きな部屋の場合にも、広帯域ランダム騒音の制御、あるいは周期性騒音の制御を安定に、しかも消音効果が十分に得られるようにし、さらに自己の音声と通信機を介した相手の音声のモニタを実現することである。

【００１６】本発明の第２の目的は、第１の目的に加えて、自己の音声および相手の音声のモニタによって生じる消音制御時の誤動作を防止し、また自己の音声発生時以外は通信機への音声を送信せず、さらに通信機の省電力化を可能とすることである。本発明の第３の目的は、第２の目的に加えて、自己の音声信号に含まれる周囲騒音を適応フィルタを用いて能動的に低減し、自己の音声発生時の適応フィルタの誤動作を防止することである。

【００１７】本発明の第４の目的は、第３の目的に加えて、自己の音声の音声判別に要する時間を補償することにより、自己の音声のモニタによる消音制御時の誤動作

および自己の音声信号の騒音制御時の誤動作の防止を確実なものとし、さらに自己の音声のモニタおよび通信機への送信時における音声信号の欠落を防止することである。本発明の第５の目的は、第３の目的をさらに精度よく実現することである。本発明の第６の目的は、第４の目的をさらに精度よく実現することである。

【００１８】

【課題を解決するための手段】第１の目的を達成するために、ヘッドホンと音声検出器が一体に構成された本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンの筐体であるキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の近傍に取り付けられた第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力を信号処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力を信号処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号とを加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力とから前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器とを有している。

【００１９】また、第２の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられた第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号とを加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記音声検出器の出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器とを有している。

【００２０】また、第３の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられた第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記通信手段からの受信信号とを加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フ

ィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力とから前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、前記音声検出器の出力を遅延させる遅延器と、前記第１の騒音検出器の出力信号を処理する第３の適応フィルタと、前記遅延器の出力から前記第３の適応フィルタの出力を減算する減算器と、前記第１の騒音検出器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、前記減算器の出力を通過させるかさせないかを制御するミュート回路とを有している。

【００２１】また、第４の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられており、主に周囲の騒音を検出する第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、前記音声検出器の出力を遅延させる第１の遅延器と、前記第１の騒音検出器の出力を遅延させる第２の遅延器と、前記第２の遅延器の出力信号を処理する第３の適応フィルタと、前記第１の遅延器の出力から前記第３の適応フィルタの出力を減算する減算器と、前記第２の遅延器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、前記減算器の出力を通過させるかさせないかを選択するミュート回路と、前記ミュート回路の動作を制御するミュート制御回路とを有している。

【００２２】また、第５の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられた第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、前記音声検出器の出力を遅延させる遅延器と、前記第１の騒音検出器の出力信号を処理する第３の適応フィルタと、前記第２の騒音検出器の出力信号を処理する第４の適応フィルタと、前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、前記遅延器の出力から前記第３の適応フィルタの出力と前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する減算器と、前記第１の騒音検出器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の

係数更新器と、前記第２の騒音検出器の出力と前記減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、前記第３の騒音検出器の出力と前記減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、前記減算器の出力を通過させるかさせないかを制御するミュート回路とを有している。

【００２３】また同様に、第５の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、前記通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられており、主に周囲の騒音を検出する第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、前記音声検出器の出力を遅延させる第１の遅延器と、前記第１の騒音検出器の出力信号を処理する第３の適応フィルタと、前記第２の騒音検出器の出力信号を処理する第４の適応フィルタと、前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、前記第１の遅延器の出力から前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する第１の減算器と、前記第１の減算器の出力を遅延させる第２の遅延器と、前記第２の遅延器の出力から前記第３の適応フィルタの出力を減算する第２の減算器と、前記第１の騒音検出器の出力と前記第２の減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、前記第２の騒音検出器の出力と前記第１の減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、前記第３の騒音検出器の出力と前記第１の減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、前記第２の減算器の出力を通過させるかさせないかを制御するミュート回路とを有している。

【００２４】また、第６の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられた第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホンのキャビネットの内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、前記音声検出器の出力を遅延させる第１の遅延器と、前記第１の騒音検出器の出力を遅延させる第２の遅延器と、前記第２の遅延器の出力信号を処理する第３の適応フィルタと、前記第２の騒音検出器の出力信号を処理する第４の適応フィルタと、前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、前記第１の遅延器の出力から前記第３の適応フィルタの出力と前記第４の適応フ

ィルタの出力と前記第５の適応フィルタの出力とを減算する減算器と、前記第２の遅延器の出力と前記減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、前記第２の騒音検出器の出力と前記減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、前記第３の騒音検出器の出力と前記減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、前記減算器の出力を通過させるかさせないかを制御するミュート回路と、前記ミュート回路の動作を制御するミュート制御回路とを有している。

【００２５】また同様に、第６の目的を達成するために本発明の騒音制御ヘッドセットは、左側あるいは右側のヘッドホンのキャビネットに設けられた支持アームにより使用者の口元近傍に取り付けられている音声検出器と、前記音声検出器の出力により音声の有無を判断する音声判別器と、通信手段の受信信号の有無を判断する受信信号判別器と、前記音声検出器の近傍に取り付けられており、主に周囲の騒音を検出する第１の騒音検出器と、前記ヘッドホンの左側キャビネットに取り付けられた第２の騒音検出器と、前記第１の騒音検出器と前記第２の騒音検出器の出力を加算する第１の騒音信号加算器と、前記第１の騒音信号加算器の出力信号を処理する第１の適応フィルタと、前記第１の騒音信号加算器の出力信号を処理する第１の伝達関数補正器と、前記第１の適応フィルタの出力と前記通信手段からの受信信号を加算する第１の出力信号加算器と、前記第１の出力信号加算器の出力を再生する左側ヘッドホンと、前記左側ヘッドホンのキャビネットの内部に取り付けられた第１の誤差検出器と、前記第１の誤差検出器からの出力と前記第１の伝達関数補正器の出力から前記第１の適応フィルタの係数を演算して更新する第１の係数更新器と、前記ヘッドホンの右側キャビネットに取り付けられた第３の騒音検出器と、前記第１の騒音検出器と前記第３の騒音検出器の出力を加算する第２の騒音信号加算器と、前記第２の騒音信号加算器の出力信号を処理する第２の適応フィルタと、前記第２の騒音信号加算器の出力信号を処理する第２の伝達関数補正器と、前記第２の適応フィルタの出力と前記通信手段からの受信信号を加算する第２の出力信号加算器と、前記第２の出力信号加算器の出力を再生する右側ヘッドホンと、前記右側ヘッドホン内部に取り付けられた第２の誤差検出器と、前記第２の誤差検出器からの出力と前記第２の伝達関数補正器の出力から前記第２の適応フィルタの係数を演算して更新する第２の係数更新器と、前記音声検出器の出力を遅延させる第１の遅延器と、前記第１の騒音検出器の出力を遅延させる第２の遅延器と、前記第２の遅延器の出力信号を処理する第３の適応フィルタと、前記第２の騒音検出器の出力信号を処理する第４の適応フィルタと、前記第３の騒音検出器の出力信号を処理する第５の適応フィルタと、前記第１の遅延器の出力から前記第４の適応フィルタの出力と前記第５の適応フィルタの出力とを減算する第１の減算器と、前記第１の減算器の出力を遅延させる第３の遅延器と、前記第３の遅延器の出力から前記第３の適応フィルタの出力を減算する第２の減算器と、前記第２の遅延器の出力と前記第２の減算器の出力から前記第３の適応フィルタの係数を演算して更新する第３の係数更新器と、前記第２の騒音検出器の出力と前記第１の減算器の出力から前記第４の適応フィルタの係数を演算して更新する第４の係数更新器と、前記第３の騒音検出器の出力と前記第１の減算器の出力から前記第５の適応フィルタの係数を演算して更新する第５の係数更新器と、前記第２の減算器の出力を通過させるかさせないかを制御するミュート回路と、前記ミュート回路の動作を制御するミュート制御回路とを有している。

【００２６】

【発明の実施の形態】以下本発明の実施の形態について、図１から図２４を用いて説明する。

【００２７】《実施例１》実施例１を図１ないし図８を参照して説明する。図１は本発明の実施例１における騒音制御ヘッドセットのブロック図を示すものである。図１において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。マイクロホン１ｆの出力と無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィ

ルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。、以上のように構成された騒音制御ヘッドヘットについて、以下その動作について説明する。以後の説明において、騒音を抑制してヘッドセットのスピーカ６ａ、６ｂから出る音の耳元で音圧レベルを低減させることを「騒音制御」と称する。

【００２８】まずマイクロホン１ａ、１ｂ及び１ｃで検出した周囲騒音は、加算器２ａ及び２ｂにおいてそれぞれ所定のレベル比で加算され、適応フィルタ３ａ及び３ｂおよびＦｘフィルタ４ａ及び４ｂにそれぞれ入力される。適応フィルタ３ａ及び３ｂから出力される制御信号は、それぞれの加算器２ｃ及び２ｄにおいて、マイクロホン１ｆからの音声信号及び通信機１０からの受信信号とそれぞれ所定のレベル比で加算され、それぞれのヘッドホン７ａ及び７ｂ内に設置されているスピーカ６ａ及び６ｂに印加され制御音として再生される。ヘッドホン７ａ及び７ｂ内にそれぞれ設置されているマイクロホン１ｄ及び１ｅは、スピーカ６ａ及び６ｂからの制御音と、ヘッドホン７ａ及び７ｂを伝わって内部に進入してくる騒音が検出され、この制御音と騒音が干渉した結果のそれぞれの誤差信号がＬＭＳ演算器５ａ及び５ｂに入力される。

【００２９】一方、加算器２ａ及び２ｂからの騒音信号は、それぞれのＦｘフィルタ４ａ及び４ｂにおいてその係数との畳み込み演算が行われ、それぞれＬＭＳ演算器５ａ及び５ｂに入力される。そしてＬＭＳ演算器５ａ及び５ｂは、それぞれＦｘフィルタ４ａ及び４ｂからの信号とマイクロホン１ｄ及び１ｅからの信号により、それぞれのマイクロホン１ｄ及び１ｅからの出力信号を最小とするように、適応フィルタ３ａ及び３ｂの係数を演算しその値を更新する。ここで、Ｆｘフィルタ４ａはスピーカ６ａ（正確には加算器２ｃ）からマイクロホン１ｄまでの伝達関数に係数として近似するようになされており、Ｆｘフィルタ４ｂはスピーカ６ｂ（正確には加算器２ｄ）からマイクロホン１ｅまでの伝達関数に係数として近似するようになされている。

【００３０】この方法をFiltered-x LMSアルゴリズム（例えば参考文献として、B. Widrowand S. Stearns, Adaptive Signal Processing,(Prentice-Hall,Englewood Cliffs,NJ,1985)）という。そして、Filtered-x LMSアルゴリズムの係数更新演算は、以下の式で表される。

【００３１】

$$w(n+1)=w(n)+\mu r(n)e(n)\cdots(1)$$

【００３２】ただし、$r(n)=x(n)^{T}c(n)$

【００３３】ｗ（ｎ）；適応フィルタ３ａあるいは３ｂの係数

【００３４】ｃ（ｎ）；Ｆｘフィルタ４ａあるいは４ｂの係数

【００３５】ｘ（ｎ）；加算器２ｃあるいは２ｄからの騒音信号

【００３６】ｅ（ｎ）；マイクロホン１ｄあるいは１ｅからの誤差信号

【００３７】μ　　；ステップパラメータ

【００３８】Ｔは転置を示す。

【００３９】以上の動作を繰り返すことにより、ヘッドホン７ａ及び７ｂの空間内（つまり耳元）の騒音が低減される。

【００４０】ところで、各マイクロホン１ａ、１ｂ、１ｃ、１ｅ及び～１ｆの位置関係は図２に示すように、マイクロホン１ａはマイクロホン１ｆと共に使用者の口元近傍に設置される。マイクロホン１ｂはヘッドホン７ａの外側の後方上部、すなわち図２の（ｃ）に示すように、左側ヘッドホン７ａをキャビネット外側を側方から見た場合に、左側ヘッドホン７ａの中心点Ｐ１を通る水平軸と垂直軸により左上部、左下部、右上部および右下部の４領域に分割した右上部領域に設置される。マイクロホン１ｃはヘッドホン７ｂの外側の後方上部すなわち図２の（ａ）に示すように、右側ヘッドホン７ｂのキャビネット外側を側方から見た場合に、右側ヘッドホン７ｂの中心点Ｐ２を通る水平軸と垂直軸により左上部、左下部、右上部および右下部の４領域に分割した左上部領域に設置される。マイクロホン１ｄは、図１に示すように、ヘッドホン７ａ内のスピーカ６ａの前面で外耳穴近くに設置され、マイクロホン１ｅはヘッドホン７ｂ内のスピーカ６ｂの前面で外耳穴近くに設置されている。つまり、図２の（ｃ）に示すように、マイクロホン１ａ、１ｄ及び１ｂが側方から見ておよそ一本の直線Ｌ１上に位置するように設置されている。同様に、図２の（ａ）に示すように、マイクロホン１ａ、１ｅ及び１ｃが側方から見ておよそ一本の直線Ｌ２上に位置するように設置されている。

【００４１】このように設置した構成とすることにより、図３に示すように、使用者の前方から騒音が来る場合はマイクロホン１ａが主に騒音を検出し、左方から騒音が来る場合はマイクロホン１ｂが主に騒音検出をする。また使用者の右方から騒音が来る場合はマイクロホン１ｃが主に騒音を検出し、後方から騒音が来る場合はマイクロホン１ｂ及び１ｃが主に騒音を検出をする。これによって、どのような方向から騒音が到来しても、つまりどこに騒音源があっても、また騒音源が複数であっても、残響時間の長い空間であっても、さらに使用者が

どのように移動しても、騒音を適確に検出することができる。

【００４２】図４は拡散音場である残響室において実測した複数の騒音源が存在する場合の騒音制御効果を示すグラフであり、横軸は周波数、縦軸は耳元での音圧レベルを表している。曲線Ａはヘッドセットを装着しないときの音圧レベルである。曲線Ｂはヘッドセットを装着し、騒音制御を行わないときのヘッドホン７ａ及び７ｂの遮音特性により減衰した音圧レベルを示す。曲線Ｃは騒音制御を行うヘッドセットを装着したときの音圧レベルであり、曲線ＢとＣの差が適応フィルタ３ａ及び３ｂによる能動的騒音制御の効果を示す。残響室で複数の騒音を発生させると、図３に示すように全ての方向から騒音が到来していると考えられるが、それでもヘッドホン７ａ及び７ｂの遮音特性があまりよくない低周波では騒音が低減されている。

【００４３】このように、高周波についてはヘッドホン７ａ及び７ｂの遮音により遮断し、遮音しきれない低周波についてはその制御帯域において騒音制御の因果律を満たすようにシステムの動作速度を上げれば、図４のように安定に消音できる。また、騒音検出をマイクロホン１ａだけでなく、マイクロホン１ｂ及び１ｃでも行い、加算器２ａ及び２ｂでマイクロホン１ａの出力と加算することにより、ある一定方向だけでなく、周囲の騒音をくまなく検出でき、それによって因果律を満たすようにある程度補正することができる。

【００４４】図５は、騒音源の直前での騒音制御効果で、つまり図３における前方騒音源だけが存在する場合の騒音制御効果を示している。図５の曲線Ａ，曲線Ｂ及び曲線Ｃはそれぞれ図４と同じ測定条件のとき音圧レベルを示す。この場合は、マイクロホン１ａからマイクロホン１ｄ及び１ｅまでの距離がある程度確保できるので、図４では制御されなかった高い周波数まで制御できるようになる。これは騒音制御の因果律を満足する周波数帯域が拡大したことによる。つまり、マイクロホン１ｄあるいはマイクロホン１ｅにおいて、騒音源と各マイクロホン１ｄ及び１ｅの間に、騒音検出用のマイクロホン１ａ、１ｂ及び１ｃが存在するとき、騒音制御に要する時間が騒音の伝達時間に近づくことになる。換言すれば、騒音の伝達時間が一番大きくなるとき、騒音制御に対する時間的余裕が生まれることになる。

【００４５】図６は、１ＫＨｚの音の制御効果を示している。図において、曲線Ａ，Ｂ、Ｃは、それぞれヘッドセットを装着しない場合、騒音制御をしないヘッドセットを装着した場合、騒音制御をするヘッドセットを装着した場合の音圧レベルである。このように適応フィルタ３ａ及び３ｂを用いれば、従来のアナログ式騒音制御ヘッドホンでは不可能であった高周波音の制御が可能となる。以上のように、本実施例の騒音制御ヘッドセットは、どのような騒音源、あるいは音場であっても、安定かつ十分に消音効果を得ることができる。また音声検出用のマイクロホン１ｆを使用者の口の方向に指向性のある双指向性、あるいは単一指向性のマイクロホンを用いることにより、音声以外の余計な周囲騒音を極力検出しないようにすることができる。そしてこの音声信号を加算器２ｃ及び２ｄにおいて制御信号および通信機１０からの受信信号と加算することにより、騒音を低減しながら自己の音声と相手方の音声をモニタすることができる。当然、この音声信号は通信機１０を介して相手方に送信されるので、例えばこの騒音制御ヘッドセットを双方が使用していれば、通常会話できないような騒音環境であっても十分に会話ができ意志疎通を図ることができる。

【００４６】なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結ぶ構成としているが、図７に示すように支持部１２を介してヘルメット１１に取り付ける構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。また、マイクロホン１ａ、１ｂ、及び１ｃの出力のＡＤ変換は、加算器２ａ及び２ｂの前で行ってもよく、また後で行ってもよい。加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータの数が削減できる。

【００４７】さらにＬＭＳ演算器５ａ及び５ｂは、それぞれのＦｘフィルタ４ａ及び４ｂの出力によりマイクロホン１ｄ及び１ｅからの誤差信号を正規化する、よく知られたいわゆる学習同定法を用いるものであってもよい。通信機１０は、騒音制御ヘッドセット内に内蔵していてもよく、あるいは外部の通信機を使用してもよい。さらに、図８に示すように、加算器２ａ及び２ｂがマイクロホン１ａ、１ｂ及び１ｃの検出信号を全て加算する構成でもよい。

【００４８】《実施例２》実施例２を図９を参照して説明する。図９は本発明の実施例２における騒音制御ヘッドセットのブロック図を示すものである。図９において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の

受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３に入力されかつスイッチ１５ｃを介して無線通信機１０に入力される。また前記の受信信号は受信信号判別器１４にも入力される。本実施例の騒音制御ヘッドセットは、図１における実施例１の騒音制御ヘッドセットの構成要素にさらに音声判別器１３と、受信信号判別器１４と、スイッチ１５ａ、１５ｂ及び１５ｃとを加えたものであり、以下にその動作について説明する。

【００４９】使用者が音声を発すると、その音声はマイクロホン１ｆで検出されて音声信号が音声判別器１３に印加され、音声判別器１３で音声が発生されたことを検知する。音声判別器１３の出力の検知信号によってスイッチ１５ｃが導通し、これによって音声信号が加算器２ｃ及び２ｄに加えられてスピーカ６ａ及び６ｂによって自己の音声が再生される。また音声信号は通信機１０にも入力される。このとき音声判別器１３からの検知信号により通信機１０の送信用電源を立ち上げることにより、自己の音声信号が相手方に送信される。同時に、この検知信号によってスイッチ１５ａ及び１５ｂが開放され、これによってＬＭＳ演算器５ａ及び５ｂの誤差信号（式１におけるｅ（ｎ））が０となるので、適応フィルタ３ａ及び３ｂの係数更新動作が停止する。

【００５０】使用者が話すのを止めて音声が無くなると、音声判別器１３が音声の無いことを検出し、これによってスイッチ１５ｃが開放されるので加算器２ｃ及び２ｄおよび通信機１０にはマイクロホン１ｆからの出力信号が入力されなくなる。このとき音声判別器１３からの検知信号により通信機１０の送信用電源を切ることにより、通信機１０の消費電力を削減することができる。そして同時に、スイッチ１５ａ及び１５ｂが導通し、ＬＭＳ演算器５ａ～５ｂの係数更新動作が再び開始される。

【００５１】これによって、音声発生時にスピーカ６ａ及び６ｂで再生される自己の音声がマイクロホン１ｄ及び１ｅで検出され、ＬＭＳ演算器５ａ及び５ｂの係数更新動作に影響を及ぼし、適応フィルタ３ａ及び３ｂを不安定にすることを防止できる。また、スイッチ１５ｃの開放によって音声の無い無言時には周囲騒音を相手に送信しないようにすることが可能となる。また、受信信号判別器１４は通信機１０から印加される受信信号の有無を検出し、受信信号がある場合には受信信号を加算器２ｃ及び２ｄを経てスピーカ６ａ及び６ｂで再生する。またスイッチ１５ａ及び１５ｂは開放されて適応フィルタ３ａ～３ｂの係数更新動作が停止する。

【００５２】これによって、無線信号の受信時に適応フィルタ３ａ及び３ｂが不安定になることを防止できる。以上より、無言時あるいは通信機１０からの受信信号が無い場合に、本実施例の騒音制御ヘッドセットは、使用者の耳元での騒音を低減する。また音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減するとともに、適応フィルタ３ａ及び３ｂの動作が不安定になることを防止できる。また、無言時には周囲騒音を相手に送信しないとともに通信機１０の消費電力を低減することが可能となる。

【００５３】なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結合した構成としているが、図７に示すようにヘルメット１１に取り付けた構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。また、マイクロホン１ａ、１ｂ及び１ｃの出力のＡＤ変換は加算機２ａ及び２ｂの前で行ってもよく、また後で行ってもよい。加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータの数が削減できる。

【００５４】さらに、音声判別器１３および受信信号判別器１４は、マイクロホン１ｆからの音声信号と通信機１０からの受信信号をそれぞれデジタル化して所定の音声検出アルゴリズムにより音声の有無の判別をしてもよいが、例えば音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。

【００５５】さらにＬＭＳ演算器５ａ及び５ｂは、Ｆｘフィルタ４ａ及び４ｂの出力によりマイクロホン１ｄ及び１ｅからの誤差信号を正規化する、いわゆる学習同定法であってもよい。通信機１０は、騒音制御ヘッドセットに内蔵しても、あるいは外部の通信機を使用してもよい。さらに、図８で説明したように、図９において加算器２ａ及び２ｂでマイクロホン１ａ、１ｂ及び１ｃの検

出信号を全て加算する構成でもよい。

【００５６】《実施例３》図１０は本発明の実施例３における騒音制御ヘッドセットのブロック図を示すものである。図１０において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３と遅延器１６に入力される。また前記の受信信号は受信信号判別器１４にも入力される。遅延器１６の出力は減算器１７に入力される。マイクロホン１ａの出力は第３の適応フィルタ３ｃと第３のＬＭＳ演算器５ｃにも入力される。第３のＬＭＳ演算器５ｃの出力は第３の適応フィルタ３ｃに入力され、第３の適応フィルタ３ｃの出力は減算器１７に入力される。減算器１７の出力はスイッチ１５ｃを介してＬＭＳ演算器５ｃに入力され、ミュート回路であるところのスイッチ１５ｄを介して無線通信機１０に入力される。本実施例の騒音制御ヘッドセットは、図９における騒音制御ヘッドセットの構成要素にさらに適応フィルタ３ｃと、ＬＭＳ演算器５ｃと、スイッチ１５ｃ及び１５ｄと、遅延器１６と減算器１７とを加えたものであり、以下にその動作について説明する。

【００５７】マイクロホン１ｆで検出された周囲騒音の信号は、遅延器１６で所定の、適応フィルタ３ｃが因果律を満たす時間だけ遅延され、減算器１７に入力される。一方、マイクロホン１ａで検出された騒音信号は、適応フィルタ３ｃで信号処理され、その出力が減算器１７に入力される。減算器１７では、遅延器１６の出力から適応フィルタ３ｃの出力が減算され、その結果の出力信号がスイッチ１５ｃ及び１６ｄに入力される。そしてＬＭＳ演算器５ｃは、マイクロホン１ａからの騒音信号とスイッチ１５ｃを経て入力される誤差信号とにより、スイッチ１５ｃを経て入力される誤差信号を最小にするように適応フィルタ３ｃの係数を演算して更新する。これによって、減算器１７の出力信号に含まれる騒音成分が低減されることになる。

【００５８】マイクロホン１ａ及び１ｆは、図１１および図１２に示すように、マイクロホン１ｆが使用者の口の方向に点線の長円Ｄ１で示す指向性を有する単一指向性型であり、マイクロホン１ａはマイクロホン１ｆとは正反対の方向の面に音孔を有し、マイクロホン１ｆの真横に設置されている点線の円Ｄ２で示す無指向性のマイクロホンとなっている。

【００５９】図１３は拡散音場である残響室で複数の騒音源から音を発生させたときのマイクロホン１ｆの検出信号における騒音制御効果を示している。曲線Ｂは騒音制御を行っていない場合の減算器１７の出力レベルであり、曲線Ｃは騒音制御を行っている場合の減算器１７の出力における騒音レベルである。また図１４は、前記の拡散音場で音声を発した場合の音声信号の変化を示している。曲線Ｂ及びＣはそれぞれ騒音制御を行っていない場合と行っている場合の減算器１７の出力における音声レベルを示す。

【００６０】図１３に示すように、マイクロホン１ｆで検出される音声信号に含まれる騒音成分は十分に減衰されるが、音声成分のレベル変化は図１４に示すようにほとんど無いことがわかる。図１５は、騒音源の直前での騒音制御効果で、つまり図３における前方騒音源だけが存在する場合の制御効果を示している。曲線ＢとＣとの比較からわかるように、曲線Ｃで示す騒音制御を行う場合には、騒音の出力レベルが大幅に低減される。

【００６１】このように、騒音源の直前で制御すれば、騒音低減量が増加するのがわかる。ところで、使用者が音声を発すると、音声判別器１３では音声が発生したことを検知し、検知信号を出力する。この検知信号によってスイッチ１５ｄが導通し、これによって音声信号が加算器２ｃ及び２ｄに加えられてそれぞれのスピーカ６ａ

及び６ｂで自己音声が再生される。検知信号は通信機１０にも入力される。このとき音声判別器１３からの検知信号により通信機１０の送信用電源が立ち上げられ、自己の音声信号が相手側に送信される。そして同時に、この検知信号によってスイッチ１５ａ～１５ｃが開放され、これによってＬＭＳ演算器５ａ～５ｃの誤差信号が０となるので、適応フィルタ３ａ～３ｃの係数更新動作が停止する。

【００６２】次に話すのを止めて音声が無くなると、音声判別器１３が音声の無いことを検出し、これによってスイッチ１５ｄが開放されるので加算器２ｃ及び２ｄおよび通信機１０にはマイクロホン１ｆからの出力信号が入力されなくなる。このとき音声判別器１３からの検知信号により通信機１０の送信用電源を切ることにより、通信機１０の消費電力を削減することができる。このとき、スイッチ１５ａ～１５ｃが導通し、これによってＬＭＳ演算器５ａ～５ｃの係数更新動作が再び開始される。

【００６３】その結果音声発生時にスピーカ６ａ及び６ｂで再生される自己の音声がマイクロホン１ｄ及び１ｅで検出されて、ＬＭＳ演算器５ａ及び５ｂの係数更新動作に影響を及ぼし、適応フィルタ３ａ及び３ｂの動作を不安定にするのを防止できる。また音声発生時にＬＭＳ演算器５ｃの誤差信号に残る音声信号が、ＬＭＳ演算器５ｃの係数更新動作に影響を及ぼし、適応フィルタ３ｃの動作を不安定にするのを防止できる。さらにマイクロホン１ａに音声が混入した場合、適応フィルタ３ｃの出力によって減算器１７において音声信号が低減されることを防止できる。そして、スイッチ１５ｄによって無言時には余計な周囲騒音を相手に送信しないようにすることが可能となる。

【００６４】また、受信信号判別器１４は通信機１０からの受信信号の有無を検出し、受信信号がある場合には、その受信信号は加算器２ｃ及び２ｄを経てスピーカ６ａ及び６ｂで再生され、その間はスイッチ１５ａ及び１５ｂが開放されて適応フィルタ３ａ及び３ｂの係数更新動作が停止する。これによって、無線受信時に適応フィルタ３ａ及び３ｂの動作を不安定にすることを防止できる。

【００６５】以上より、無言時あるいは通信機１０からの受信信号が無い場合、本実施例の騒音制御ヘッドセットは、図１で説明したように使用者の耳元での騒音を低減し、同時に無言時には、適応フィルタ３ｃによりマイクロホン１ｆからの検出信号に含まれる周囲騒音信号を低減できる。音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減しながら適応フィルタ３ａ及び３ｂの動作が不安定にならないように制御でき、同時に音声発生時には、音声信号に混入する周囲騒音成分を低減しながら適応フィルタ３ｃが不安定にならないように、あるいは音声信号を低減しないように制御できる。また、無言時には余計な周囲騒音を相手に送信しないことと通信機１０の消費電力を低減することが可能となる。

【００６６】なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結ぶ構成としているが、図７に示すようにヘルメット１１に取り付けた構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。また、加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータが削減できる。さらに、音声判別器１３および受信信号判別器１４は、マイクロホン１ｆからの音声信号と通信機１０からの受信信号をデジタル化して所定の音声検出アルゴリズムにより音声判別してもよいが、例えば音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。

【００６７】また、マイクロホン１ａとマイクロホン１ｆの振動板直径は、同じでもよいが、発明者による実験の結果マイクロホン１ｆの振動板直径がマイクロホン１ａのものの２倍以上あれば、より好ましい。これはマイクロホン１ｆの感度がマイクロホン１ａより高い方が好ましいからである。さらに、図１６に示すように、マイクロホン１ａとマイクロホン１ｆの騒音源方向の面（口の方向とは正反対の面）に、マイクロホン１ａ及び１ｆの音孔を覆うように、例えば不織布などの吸音材１８をかぶせてもよい。

【００６８】さらにＬＭＳ演算器５ａ、５ｂ及び５ｃは、Ｆｘフィルタ４ａ及び４ｂの出力およびマイクロホン１ａの出力により、マイクロホン１ｄ及び１ｅと減算器１７からの誤差信号を正規化する、いわゆる学習同定法であってもよい。通信機１０は、騒音制御ヘッドセットに内蔵していてもよく、またあるいは外部の通信機を使用してもよい。さらに、図８で説明したように、図１０において加算器２ａ及び２ｂがマイクロホン１ａ、１ｂ及び１ｃの検出信号を全て加算する構成でもよい。

【００６９】《実施例４》実施例４を図１７を参照して説明する。図１７は本発明の実施例４における騒音制御ヘッドセットのブロック図を示すものである。図１７において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加

算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３と遅延器１６ａに入力される。また前記の受信信号は受信信号判別器１４にも入力される。遅延器１６ａの出力は減算器１７に入力される。マイクロホン１ａの出力は遅延器１６ｂを経て第３の適応フィルタ３ｃと第３のＬＭＳ演算器５ｃにも入力される。第３のＬＭＳ演算器５ｃの出力は第３の適応フィルタ３ｃに入力され、第３の適応フィルタ３ｃの出力は減算器１７に入力される。減算器１７の出力はスイッチ１５ｃを介してＬＭＳ演算器５ｃに入力され、ミュート回路であるところのスイッチ１５ｄを介して無線通信機１０に入力される。音声判別器１３の出力はミュート制御回路１９に入力され、ミュート制御回路１９の出力は無線通信機１０に入力される。本実施例の騒音制御ヘッドセットは、図１０における騒音制御ヘッドセットの構成要素に遅延器１６ｂとミュート制御回路１９とを加えたものであり、以下にその動作について説明する。

【００７０】マイクロホン１ｆで検出された周囲騒音の信号は、遅延器１６ａで所定の、適応フィルタ３ｃが因果律を満たす時間だけ遅延され、減算器１７に入力される。一方、マイクロホン１ａで検出された騒音信号は、遅延器１６ｂで所定の、音声判別器１３が音声の発生を確実に判別するのに要する時間だけ遅延され、適応フィルタ３ｃで信号処理され、その出力が減算器１７に入力される。減算器１７では、遅延器１６ａの出力から適応フィルタ３ｃの出力が減算され、その結果の誤差信号がスイッチ１５ｃ及び１５ｄに入力される。ＬＭＳ演算器５ｃは、遅延器１６ｂからの騒音信号とスイッチ１５ｃからの誤差信号により、スイッチ１５ｃからの誤差信号を最小にするように適応フィルタ３ｃの係数を演算し更新する。これによって、減算器１７の出力信号に含まれる騒音成分が低減されることになる。

【００７１】ここで、マイクロホン１ａ及び１ｆは、実施例３の場合と同様に、図１１あるいは図１６に示されるように構成されており、そのときの制御効果は、図１３、図１４及び図１５に示すグラフのようになる。図１７において、音声を発すると、音声判別器１３で音声が発生したことを検知し検知信号を出力する。この検知信号はミュート制御回路１９に入力される。ミュート制御回路１９は音声判別器１３からの検知信号を受けるとスイッチ１５ｄを導通させる。これによって音声信号が加算器２ｃ及び２ｄに加えられてスピーカ６ａ及び６ｂで自己の音声が再生されるとともに、通信機１０にも入力される。このとき音声判別器１３からの検知信号により通信機１０の送信用電源を立ち上げることにより、自己の音声信号が相手方に送信される。また同時に、音声判別器１３の検知信号によってスイッチ１５ａ、１５ｂ及び１５ｃが開放され、これによってＬＭＳ演算器５ａ、５ｂ及び５ｃの誤差信号が０となるので、適応フィルタ３ａ、３ｂ及び３ｃの係数更新が停止する。

【００７２】音声が無くなると、音声判別器１３が音声の無いことを検出し、その検知信号がミュート制御回路１９に入力される。ミュート制御回路１９はこの検出信号を受けてから一定時間後にスイッチ１５ｄを開放し、これによって加算器２ｃ及び２ｄ、および通信機１０にはマイクロホン１ｆからの出力信号が入力されなくなる。このときミュート制御回路１９からの信号により通信機１０の送信用電源を切ることにより、通信機１０の消費電力を節減することができる。また同時に、スイッチ１５ａ～１５ｃが導通し、これによってＬＭＳ演算器５ａ、５ｂ及び５ｃの係数更新動作が再び開始される。

【００７３】ここで音声判別器１３は、例えば図１８に示すように、バンドパスフィルタ２０とパワー演算器２１ａ及び２１ｂと比較器２２ａ及び２２ｂとから構成されている。バンドパスフィルタ２０はマイクロホン１ｆで検出した信号のうち、音声帯域成分のみを通過させ、パワー演算器２１ａ及び２１ｂに入力する。パワー演算器２１ａ及び２１ｂは、ある一定期間のパワー（二乗和）を求めるが、パワー演算器２１ａとパワー演算器２１ｂは、位相がずれて動作するようになされる。次にパワー演算器２１ａ及び２１ｂの出力としきい値設定器２２で設定されている所定のしきい値とをそれぞれの比較器２２ａ及び２２ｂで比較し、パワー演算器２１ａ及び２１ｂの両方の出力が、このしきい値より小さければスイッチ１５ｃを導通させ、いずれかが大きければスイッ

チ１５ｃを開放する。また、パワー演算器２１ａ及び２１ｂのいづれかの出力がしきい値より大きければ、ミュート制御回路１９はスイッチ１５ｄを導通させ、どちらも小さくなれば一定時間後にスイッチ１５ｄを開放する。

【００７４】ところで、音声判別器１３をこのように構成すると、音声を発してからパワー演算器２１ａ及び２１ｂの出力がしきい値以上になるまでにある一定の時間が必要となる。つまり音声を発してから、スイッチ１５ｃとミュート制御回路１９に音声判別器１３より検知信号が入力されるまで、遅延時間があることになる。そこでこの遅延時間を補正するために、マイクロホン１ａの出力を遅延器１６ｂで遅延させ、さらにマイクロホン１ｆの出力を遅延器１６ａで音声検出に要する遅延時間（遅延器１６ｂの遅延時間に等しい）を含めて適応フィルタ３ｃが因果律を満たす時間だけ遅延させる。これによって、音声発声初期における音声判別器１３の音声を検出するのに要する時間中、適応フィルタ３ｃはその音声信号を制御しない。

【００７５】また、音声判別器１３が音声がないと判断してからミュート制御回路１９が一定時間後にスイッチ１５ｄを開放することにより、音声レベルが小さくなったときや、言葉（会話）の最後の部分を途切れることなく通信機１０および加算器２ｃ及び２ｄに入力することができる。これによって、音声発生時にスピーカ６ａ及び６ｂで再生される自己の音声がマイクロホン１ｄ及び１ｅで検出されるためにＬＭＳ演算器５ａ及び５ｂの係数更新動作に影響を及ぼし、適応フィルタ３ａ及び３ｂの動作を不安定にすることを防止できる。また音声発生時にＬＭＳ演算器５ｃの誤差信号に残る音声信号によってＬＭＳ演算器５ｃの係数更新動作が影響を受け、適応フィルタ３ｃの動作を不安定にすることを防止できる。さらにマイクロホン１ａに音声が混入した場合、適応フィルタ３ｃの出力によって減算器１７において音声信号が低減されることを防止できる。そして、スイッチ１５ｄによって無言時には余計な周囲騒音を相手に送信しないようにすることが可能となる。

【００７６】また、受信信号判別器１４は通信機１０からの受信信号の有無を検出し、受信信号がある場合には、加算器２ｃ及び２ｄによりスピーカ６ａ及び６ｂでそれぞれ再生されるとともに、スイッチ１５ａ及び１５ｂが開放されて適応フィルタ３ａ及び３ｂの係数更新動作が停止する。これによって、無線受信時に適応フィルタ３ａ及び３ｂの動作が不安定になるのを防止できる。

【００７７】以上より、無言時あるいは通信機１０からの受信信号が無い場合、本実施例の騒音制御ヘッドセットは、図１の実施例１で説明したように耳元での騒音を低減する。そして無言時に適応フィルタ３ｃによりマイクロホン１ｆからの検出信号に含まれる周囲騒音信号を低減できる。音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減しながら適応フィルタ３ａ及び３ｂの動作が不安定にならないようにできる。そして同時に音声発生時に音声信号に混入する周囲騒音成分を低減しながら適応フィルタ３ｃが不安定にならないように、あるいは音声信号を低減してしまわないようにできる。また、無言時に余計な周囲騒音を相手に送信しないことと通信機１０の消費電力を低減することが可能となる。さらに、ミュート制御回路１９の出力によって通信機１０の送信用電源を制御することにより、スイッチ１５ｄが導通状態の期間と送信用電源が入っている期間が同じになるので、音声信号を途切れることなく相手に送信できる。

【００７８】なお、図１７に示す本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結合する構成としているが、図７に示すように、ヘルメット１１に取り付ける構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。また、加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータが削減できる。

【００７９】さらに、受信信号判別器１４は、通信機１０からの受信信号をデジタル化して所定の音声検出アルゴリズムにより音声判別してもよいが、受信音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。同様に、音声判別器１３も図１８のような構成ではなく、音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別してもよい。また、マイクロホン１ａとマイクロホン１ｆの振動板直径は、同じでもよいが、マイクロホン１ｆの振動板直径の方がマイクロホン１ａの２倍以上あれば、より好ましい。

【００８０】さらにＬＭＳ演算器５ａ、５ｂ及び５ｃは、Ｆｘフィルタ４ａ及び４ｂの出力および遅延器１６ｂの出力により、マイクロホン１ｄ及び１ｅと減算器１７からの誤差信号を正規化する、いわゆる学習同定法であってもよい。通信機１０は、騒音制御ヘッドセットとして内蔵してよく、また外部の通信機を使用してもよい。さらに、実施例１の図８で説明したように、図１７において加算器２ａ及び２ｂがマイクロホン１ａ、１ｂ及び１ｃの検出信号を全て加算する構成でもよい。

【００８１】《実施例５》実施例５を図１９を参照して説明する。図１９は本発明の実施例５における騒音制御ヘッドセットのブロック図を示すものである。図１９において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネッ

ト内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３と遅延器１６に入力される。また前記の受信信号は受信信号判別器１４にも入力される。遅延器１６の出力は減算器１７に入力される。マイクロホン１ａの出力は第３の適応フィルタ３ｃと第３のＬＭＳ演算器５ｃにも入力される。第３のＬＭＳ演算器５ｃの出力は第３の適応フィルタ３ｃに入力され、第３の適応フィルタ３ｃの出力は減算器１７に入力される。減算器１７の出力はスイッチ１５ｃを介してＬＭＳ演算器５ｃに入力され、ミュート回路であるところのスイッチ１５ｄを介して通信機１０と加算器２ｃ、２ｄに入力される。マイクロホン１ｂの出力は第４の適応フィルタ３ｄと第４のＬＭＳ演算器５ｄにも入力される。またマイクロホン１ｃの出力は第５の適応フィルタ３ｅと第５のＬＭＳ演算器５ｅにも入力される。適応フィルタ３ｄ、３ｅの出力は減算器１７に入力される。ＬＭＳ演算器５ｄの出力は適応フィルタ３ｄに入力され、ＬＭＳ演算器５ｅの出力は適応フィルタ３ｅに入力される。減算器１７の出力はＬＭＳ演算器５ｄ、５ｅに入力される。本実施例の騒音制御ヘッドセットは、図１０における騒音制御ヘッドセットの構成要素に第４及び第５の適応フィルタ３ｄ及び３ｅと、第４及び第５のＬＭＳ演算器５ｄ及び５ｅとを加えたものであり、以下にその動作について説明する。

【００８２】マイクロホン１ｆで検出された周囲騒音の信号は、遅延器１６で所定の、適応フィルタ３ｃが因果律を満たす時間だけ遅延され、減算器１７に入力される。一方、マイクロホン１ａで検出された騒音信号は、適応フィルタ３ｃで信号処理され、その出力が減算器１７に入力される。また、マイクロホン１ｂで検出された騒音信号は、適応フィルタ３ｄで信号処理され、マイクロホン１ｃで検出された騒音信号は、適応フィルタ３ｅで信号処理される。適応フィルタ３ｄ及び３ｅの出力は減算器１７に入力される。減算器１７では、遅延器１６の出力から適応フィルタ３ｃ、３ｄ及び３ｅの出力が減算され、その結果の誤差信号はスイッチ１５ｃ及び１５ｄとＬＭＳ演算器５ｄ及び５ｅに入力される。ＬＭＳ演算器５ｃは、マイクロホン１ａからの騒音信号とスイッチ１５ｃからの誤差信号により、スイッチ１５ｃからの誤差信号を最小にするように適応フィルタ３ｃの係数を演算し更新する。同様に、ＬＭＳ演算器５ｄは、マイクロホン１ｂからの騒音信号と減算器１７からの誤差信号により、減算器１７からの誤差信号を最小にするように適応フィルタ３ｄの係数を更新する。ＬＭＳ演算器５ｅは、マイクロホン１ｃからの騒音信号と減算器１７からの誤差信号により、減算器１７からの誤差信号を最小にするように適応フィルタ３ｅの係数を更新する。これによって、減算器１７の出力信号に含まれる騒音成分が低減されることになる。

【００８３】ここで、マイクロホン１ａ、１ｆは実施例３の場合と同様に、図１１あるいは図１６に示されるように構成されており、そのときの制御効果は、図２０に示すグラフのようになる。図２０を図１３と比べれば、騒音制御を行わないときの曲線Ｂと騒音制御を行うときの曲線Ｃの出力レベルの差が増加して騒音制御効果が向上しているのがわかる。また、音声判別器１３および受信信号判別器１４による適応フィルタ３ａ、３ｂ及び３ｃの係数更新動作の制御と、スイッチ１５ｄの制御、および通信機１０の送信用電源の制御についての動作と効果は図１０に示す実施例３の場合と同様である。

【００８４】ただし、適応フィルタ３ｄ及び３ｅの係数更新動作において、それぞれの騒音信号は図２で示すように口元から遠いために音声を検出しにくいマイクロホン１ｂ及び１ｃからの検出信号なので、音声信号が含まれない。従って減算器１７の出力における音声信号成分は図１４に示すように低減しないので、適応フィルタ３ｄ及び３ｅの係数更新動作は停止しないようになっている。

【００８５】以上より、無言時あるいは通信機１０からの受信信号が無い場合、本実施例の騒音制御ヘッドセットは、図１の実施例１で説明したように耳元での騒音を

低減する。そして無言時に適応フィルタ３ｃ～３ｅによりマイクロホン１ｆからの検出信号に含まれる周囲騒音信号を低減できる。音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減しながら適応フィルタ３ａ及び３ｂが不安定にならないようにできる。そして音声発生時に音声信号に混入する周囲騒音成分を低減しながら適応フィルタ３ｃが不安定にならないように、あるいは音声信号を低減してしまわないようにできる。また、無言時に余計な周囲騒音を相手に送信しないことと通信機１０の消費電力を低減することが可能となる。

【００８６】なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結合する構成としているが、図７に示すようにヘルメット１１に取り付ける構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。また、加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータが削減できる。

【００８７】さらに、音声判別器１３および受信信号判別器１４は、マイクロホン１ｆからの音声信号と通信機１０からの受信信号をデジタル化して所定の音声検出アルゴリズムにより音声判別してもよいが、例えば音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。また、マイクロホン１ａとマイクロホン１ｆの振動板直径は、同じでもよいが、マイクロホン１ｆの振動板直径の方がマイクロホン１ａの２倍以上あれば、より好ましい。

【００８８】さらにＬＭＳ演算器５ａ～５ｃは、Ｆｘフィルタ４ａ及び４ｂの出力およびマイクロホン１ａの出力により、マイクロホン１ｄ及び１ｅと減算器１７からの誤差信号を正規化する、いわゆる学習同定法であってもよい。加えて、図２１に示すように、ＬＭＳ演算器５ｄ及び５ｅの誤差信号をスイッチ１５ｅ及び１５ｆによって制御するようにしてもよい。この場合、音声判別器１３で音声があると判断すればスイッチ１５ｅ及び１５ｆを開放し、無言と判断すればスイッチ１５ｅ及び１５ｆを導通することにより、減算器１７に含まれる音声レベルが大きい場合でも、適応フィルタ３ｄ及び３ｅが不安定にならないようにできる。

【００８９】通信機１０は、騒音制御ヘッドセットとして内蔵してよく、あるいは外部の通信機を使用してもよい。さらに、図８で説明したように、図１９において加算器２ａ及び２ｂがマイクロホン１ａ、１ｂ及び１ｃの検出信号を全て加算する構成でもよい。

【００９０】《実施例６》実施例６を図２２を参照して説明する。図２２は本発明の実施例６における騒音制御ヘッドセットのブロック図を示すものである。図２２において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３と遅延器１６ａに入力される。また前記の受信信号は受信信号判別器１４にも入力される。遅延器１６ａの出力は減算器１７ａに入力される。マイクロホン１ａの出力は第３の適応フィルタ３ｃと第３のＬＭＳ演算器５ｃにも入力される。第３のＬＭＳ演算器５ｃの出力は第３の適応フィルタ３ｃに入力され、第３の適応フィルタ３ｃの出力は減算器１７ｂに入力される。減算器１７ｂの出力はスイッチ１５ｃを介してＬＭＳ演算器５ｃに入力され、ミュート回路であるところのスイッチ１５ｄを介して通信機１０と加算器２ｃ、２ｄに入力される。マイクロホン１ｂの出力は第４の適応フィルタ３ｄと第４のＬＭＳ演算器５ｄにも入力される。またマイクロホン１ｃの出力は第５の適応フィルタ３ｅと第５のＬＭＳ演算器５ｅにも入力される。適応フィルタ３ｄ、３ｅの出力は減算器１７ａに入

力される。ＬＭＳ演算器５ｄの出力は適応フィルタ３ｄに入力され、ＬＭＳ演算器５ｅの出力は適応フィルタ３ｅに入力される。減算器１７ａの出力はＬＭＳ演算器５ｄ、５ｅに入力され、かつ遅延器１６ｂを経て減算器１７ｂに入力される。本実施例の騒音制御ヘッドセットは、図１９における騒音制御ヘッドセットの構成要素に遅延器１６ａと、減算器１７ａを加えたものであり、以下にその動作について説明する。

【００９１】マイクロホン１ｆで検出された周囲騒音の信号は、遅延器１６ａで所定時間の、適応フィルタ３ｄ及び３ｅが因果律を満たす時間だけ遅延され、減算器１７ａに入力される。一方、マイクロホン１ｂで検出された騒音信号は適応フィルタ３ｄで信号処理される。またマイクロホン１ｃで検出された騒音信号は適応フィルタ３ｅで信号処理される。適応フィルタ３ｄ及び３ｅの出力は減算器１７ａに入力される。減算器１７ａでは、遅延器１６ａの出力から適応フィルタ３ｄ及び３ｅの出力が減算される。ＬＭＳ演算器５ｄは、マイクロホン１ｂからの騒音信号と減算器１７ａからの誤差信号により、減算器１７ａからの誤差信号を最小にするように適応フィルタ３ｄの係数を更新する。またＬＭＳ演算器５ｅは、マイクロホン１ｃからの騒音信号と減算器１７ａからの誤差信号により、減算器１７ａからの誤差信号を最小にするように適応フィルタ３ｅの係数を更新する。これによって、減算器１７ａの出力信号に含まれる騒音成分が低減されることになる。

【００９２】次に、減算器１７ａの出力は遅延器１６ｂで所定の、適応フィルタ３ｃが因果律を満たす時間だけ遅延され、減算器１７ｂに入力される。一方マイクロホン１ａで検出された騒音信号は、適応フィルタ３ｃで信号処理され、その出力が減算器１７ｂに入力される。減算器１７ｂでは、遅延器１６ｂの出力から適応フィルタ３ｃの出力が減算され、その結果の出力信号がスイッチ１５ｃ及び１５ｄに入力される。ＬＭＳ演算器５ｃは、マイクロホン１ａからの騒音信号とスイッチ１５ｃからの誤差信号により、スイッチ１５ｃからの誤差信号を最小にするように適応フィルタ３ｃの係数を演算し更新する。これによって、減算器１７ｂの出力信号に含まれる騒音成分が低減されることになる。

【００９３】ここで、マイクロホン１ａ、１ｆは実施例３の場合と同様に、図１１あるいは図１６に示されるように構成されており、そのときの制御効果は、図２０のグラフのようになる。図２０において、曲線Ｂは騒音制御を行わないときの減算器１７の出力レベルであり、曲線Ｃは騒音制御を行うときの同出力レベルである。また、音声判別器１３および受信信号判別器１４による適応フィルタ３ａ、３ｂ及び３ｃの係数更新動作の制御と、スイッチ１５ｄの制御、および通信機１０の送信用電源の制御については、その動作と効果は実施例５の場合と同様である。

【００９４】以上より、無言時あるいは通信機１０からの受信信号が無い場合、本実施例の騒音制御ヘッドセットは、図１の実施例１で説明したように耳元での騒音を低減する。そして無言時に適応フィルタ３ｃ、３ｄ及び３ｅによりマイクロホン１ｆからの検出信号に含まれる周囲騒音信号を低減できる。音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減しながら適応フィルタ３ａ及び３ｂが不安定にならないようにできる。そして音声発生時に音声信号に混入する周囲騒音成分を低減しながら適応フィルタ３ｃが不安定にならないように、あるいは音声信号を低減してしまわないようにできる。また、無言時に余計な周囲騒音を相手に送信しないことと通信機１０の消費電力を低減することが可能となる。

【００９５】また、遅延器１６ａ及び１６ｂと減算器１７ａ及び１７ｂを設けたことにより、適応フィルタ３ｃ、３ｄ及び３ｅに実現すべき係数のインパルス応答を最適に調整することができる。なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結ぶ構成としているが、図７に示すようにヘルメット１１に取り付ける構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。

【００９６】また、加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータが削減できる。さらに、音声判別器１３および受信信号判別器１４は、マイクロホン１ｆからの音声信号と通信機１０からの受信信号をデジタル化して所定の音声検出アルゴリズムにより音声判別してもよいが、例えば音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。

【００９７】また、マイクロホン１ａとマイクロホン１ｆの振動板直径は、同じでもよいが、マイクロホン１ｆの振動板直径の方がマイクロホン１ａの２倍以上あれば、より好ましい。さらにＬＭＳ演算器５ａ、５ｂ及び５ｃは、Ｆｘフィルタ４ａ及び４ｂの出力およびマイクロホン１ａの出力により、マイクロホン１ｄ及び１ｅと減算器１７からの誤差信号を正規化する、いわゆる学習同定法であってもよい。

【００９８】加えて、図２１に示す実施例５で説明したように、図２２においてＬＭＳ演算器５ｄ及び５ｅの誤差信号をスイッチ１５ｅ及び１５ｆを設けて制御するようにしてもよい。さらに、図８で説明したように、図２２において、加算器２ａ及び２ｂがマイクロホン１ａ、１ｂ及び１ｃの検出信号を全て加算する構成でもよい。通信機１０は、騒音制御ヘッドセットとして内蔵しても、あるいは外部の通信機を使用してもよい。

【００９９】《実施例７》実施例７を図２３を参照して説明する。図２３は本発明の実施例７における騒音制御ヘッドセットのブロック図を示すものである。図２３に

おいて、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット２７ａ内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット２７ｂ内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３と遅延器１６ａに入力される。また前記の受信信号は受信信号判別器１４にも入力される。遅延器１６ａの出力は減算器１７に入力される。マイクロホン１ａの出力は遅延器１６ｂを経て第３の適応フィルタ３ｃと第３のＬＭＳ演算器５ｃにも入力される。第３のＬＭＳ演算器５ｃの出力は第３の適応フィルタ３ｃに入力され、第３の適応フィルタ３ｃの出力は減算器１７に入力される。減算器１７の出力はスイッチ１５ｃを介してＬＭＳ演算器５ｃに入力され、ミュート回路であるところのスイッチ１５ｄを介して無線通信機１０に入力される。音声判別器１３の出力はミュート制御回路１９に入力され、ミュート制御回路１９の出力は無線通信機１０に入力される。マイクロホン１ｂの出力は第４の適応フィルタ３ｄと第４のＬＭＳ演算器５ｄにも入力される。またマイクロホン１ｃの出力は第５の適応フィルタ３ｅと第５のＬＭＳ演算器５ｅにも入力される。ＬＭＳ演算器５ｄ、５ｅの出力はそれぞれ適応フィルタ３ｄ、３ｅに入力される。適応フィルタ３ｄ、３ｅの出力は減算器１７に入力される。本実施例の騒音制御ヘッドセットは、図１７における騒音制御ヘッドセットの構成要素に第４及び第５の適応フィルタ３ｄ及び３ｅと、第４及び第５のＬＭＳ演算器５ｄ及び５ｅを加えたものであり、以下にその動作について説明する。

【０１００】マイクロホン１ｆで検出された周囲騒音の信号は、遅延器１６ａで所定の、適応フィルタ３ｃが因果律を満たす時間だけ遅延され、減算器１７に入力される。一方、マイクロホン１ａで検出された騒音信号は、遅延器１６ｂで所定の、音声判別器１３が音声を確実に判別するのに要する時間だけ遅延され、適応フィルタ３ｃで信号処理され、その出力が減算器１７に入力される。マイクロホン１ｂで検出された騒音信号は、適応フィルタ３ｄで信号処理され、マイクロホン１ｃで検出された騒音信号は、適応フィルタ３ｅで信号処理され、それぞれの出力が減算器１７に入力される。減算器１７では、遅延器１６ａの出力から適応フィルタ３ｃ、３ｄ及び３ｅの出力が減算され、その結果の出力信号はスイッチ１５ｃ及び１５ｄとＬＭＳ演算器５ｄ及び５ｅに入力される。ＬＭＳ演算器５ｃは、遅延器１６ｂからの騒音信号とスイッチ１５ｃからの誤差信号により、スイッチ１５ｃからの誤差信号を最小にするように適応フィルタ３ｃの係数を演算し更新する。同様に、ＬＭＳ演算器５ｄは、マイクロホン１ｂからの騒音信号と減算器１７からの誤差信号により、減算器１７からの誤差信号を最小にするように適応フィルタ３ｄの係数を更新する。ＬＭＳ演算器５ｅは、マイクロホン１ｃからの騒音信号と減算器１７からの誤差信号により、減算器１７からの誤差信号を最小にするように適応フィルタ３ｅの係数を更新する。これによって、減算器１７の出力信号に含まれる騒音成分が低減されることになる。

【０１０１】ここで、マイクロホン１ａ及び１ｆは図１１あるいは図１６に示されるように構成されており、そのときの制御効果は、図２０のグラフに示すようになる。また、音声判別器１３および受信信号判別器１４による適応フィルタ３ａ、３ｂ及び３ｃの係数更新動作の制御と、ミュート制御回路１９とスイッチ１５ｄの制御、および通信機１０の送信用電源の制御については、その動作と効果は図１７に示す実施例４の場合と同様である。

【０１０２】以上より、無言時あるいは通信機１０からの受信信号が無い場合、本実施例の騒音制御ヘッドセットは、実施例１で説明したように耳元での騒音を低減する。そして無言時に適応フィルタ３ｃ、３ｂ及び３ｅに

よりマイクロホン１ｆからの検出信号に含まれる周囲騒音信号を低減できる。音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減しながら適応フィルタ３ａ及び３ｂが不安定にならないようにできる。そして音声発生時に音声信号に混入する周囲騒音成分を低減しながら適応フィルタ３ｃが不安定にならないように、あるいは音声信号を低減してしまわないようにできる。また、無言時に余計な周囲騒音を相手に送信しないことと通信機１０の消費電力を低減することが可能となる。さらに、ミュート制御回路１９の出力により通信機１０の送信用電源を制御することにより、スイッチ１５ｄが導通状態の期間と送信用電源が入っている期間が同じになるので、音声信号を途切れることなく相手に送信できる。

【０１０３】なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結合する構成としているが、図７に示すようにヘルメット１１に取り付ける構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。また、加算器２ａ及び２ｂの後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータが削減できる。

【０１０４】さらに、音声判別器１３と受信信号判別器１４は、マイクロホン１ｆからの音声信号と通信機１０からの受信信号をデジタル化して所定の音声検出アルゴリズムにより音声判別してもよいが、例えば音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。また、マイクロホン１ａとマイクロホン１ｆの振動板直径は、同じでもよいが、マイクロホン１ｆの振動板直径の方がマイクロホン１ａの２倍以上あれば、より好ましい。

【０１０５】さらにＬＭＳ演算器５ａ、５ｂ及び５ｃは、Ｆｘフィルタ４ａ及び４ｂの出力および遅延器１６ｂの出力により、マイクロホン１ｄ及び１ｅと減算器１７からの誤差信号を正規化する、いわゆる学習同定法であってもよい。通信機１０は、騒音制御ヘッドセットとして内蔵しても、あるいは外部の通信機を使用してもよい。

【０１０６】加えて、図２１で説明したように、図２３においてＬＭＳ演算器５ｄ及び５ｅの誤差信号をスイッチ１５ｅ及び１５ｆを設けて制御するようにしてもよい。さらに、実施例１の図８で説明したように、図２３において加算器２ａ及び２ｂがマイクロホン１ａ及び１ｃの検出信号を全て加算する構成でもよい。

【０１０７】《実施例８》実施例８を図２４を参照して説明する。図２４は本発明の実施例８における騒音制御ヘッドセットのブロック図を示すものである。図２４において、ヘッドホンと音声検出器が一体に構成された騒音制御ヘッドセットは、使用者の口元に配置された第１の騒音検出器であるところのマイクロホン１ａ、左側ヘッドホン７ａのキャビネットの外側に配置された第２の騒音検出器であるところのマイクロホン１ｂ、右側ヘッドホン７ｂのキャビネットの外側に配置された第３の騒音検出器であるところのマイクロホン１ｃ、キャビネット内に配置された第１の誤差検出器であるところのマイクロホン１ｄ、キャビネット内に配置された第２の誤差検出器であるところのマイクロホン１ｅ、口元に配置された音声検出器であるところのマイクロホン１ｆを有する。マイクロホン１ａ、１ｂの出力は第１の騒音信号加算器であるところの加算器２ａに入力され、マイクロホン１ａ、１ｃの出力は第２の騒音信号加算器であるところの加算器２ｂに入力される。無線又は有線の通信機１０の受信信号は第１の出力信号加算器であるところの加算器２ｃ、第２の出力信号加算器であるところの加算器２ｄに入力される。加算器２ａ、２ｂの出力は、第１の適応フィルタ３ａ、第２の適応フィルタ３ｂ（図中ではＡＦＩＲと表示する）、にそれぞれ入力される。また加算器２ａ、２ｂの出力は第１の伝達関数補正器であるところのＦｘフィルタ４ａ、第２の伝達関数補正器であるところのＦｘフィルタ４ｂにそれぞれ入力される。Ｆｘフィルタ４ａ、４ｂの出力は第１の係数更新器であるところのＬＭＳ（最小二乗平均）演算器５ａ、第２の係数更新器であるところのＬＭＳ演算器５ｂにそれぞれ入力される。マイクロホン１ｄ、１ｅの出力はそれぞれのスイッチ１５ａ、１５ｂを介してＬＭＳ演算器５ａ、５ｂに入力される。ＬＭＳ演算器５ａ、５ｂの出力は適応フィルタ３ａ、３ｂにそれぞれ入力される。適応フィルタ３ａ、３ｂの出力は加算器２ｃ、２ｄにそれぞれ入力され、加算器２ｃ、２ｄの出力は左側ヘッドホン７ａのキャビネット内にあるスピーカ６ａ、右側ヘッドホン７ｂのキャビネット内にあるスピーカ６ｂにそれぞれ入力される。左側ヘッドホン７ａと右側ヘッドホン７ｂはヘッドバンド８で結合されている。支持アーム９はマイクロホン１ａ、１ｆを口元に近づけるように支持する。マイクロホン１ｆの出力は音声判別器１３と遅延器１６ａに入力される。また前記の受信信号は受信信号判別器１４にも入力される。遅延器１６ａの出力は減算器１７ａに入力される。マイクロホン１ａの出力は遅延器１６ｂを経て第３の適応フィルタ３ｃと第３のＬＭＳ演算器５ｃにも入力される。第３のＬＭＳ演算器５ｃの出力は第３の適応フィルタ３ｃに入力され、第３の適応フィルタ３ｃの出力は減算器１７ｂに入力される。減算器１７ｂの出力はスイッチ１５ｃを介してＬＭＳ演算器５ｃに入力され、ミュート回路であるところのスイッチ１５ｄを介して通信機１０に入力される。音声判別器１３の出力はミュート制御回路１９に入力され、ミュート制御回路１９の出力は通信機１０に入力される。マイクロホン１ｂの出力は第４の適応フィルタ３ｄと第４のＬＭＳ演算器５ｄにも入力される。またマイクロホン１ｃの出力は第５の適応フィルタ３ｅと第５のＬＭＳ演算器５ｅにも入

力される。ＬＭＳ演算器５ｄ、５ｅの出力はそれぞれ適応フィルタ３ｄ、３ｅに入力される。適応フィルタ３ｄ、３ｅの出力は減算器１７ａに入力される。減算器１７ａの出力は遅延器１６ｃを経て減算器１７ｂに入力される。本実施例の騒音制御ヘッドセットは、図２３における騒音制御ヘッドセットの構成要素に遅延器１６ａと、減算器１７ａを加えたものであり、以下にその動作について説明する。

【０１０８】マイクロホン１ｆで検出された周囲騒音の信号は、遅延器１６ａで所定時間の、適応フィルタ３ｄ及び３ｅが因果律を満たす時間だけ遅延され、減算器１７ａに入力される。一方、マイクロホン１ｂで検出された騒音信号は、適応フィルタ３ｄで信号処理され、またマイクロホン１ｃで検出された騒音信号は、適応フィルタ３ｅで信号処理され、それぞれの出力が減算器１７ａに入力される。減算器１７ａでは、遅延器１６ａの出力から適応フィルタ３ｄ及び３ｅの出力が減算され、結果の出力はＬＭＳ演算器５ｄ及び５ｅに入力される。そしてＬＭＳ演算器５ｄは、マイクロホン１ｂからの騒音信号と減算器１７ａからの誤差信号により、減算器１７ａからの誤差信号を最小にするように適応フィルタ３ｄの係数を更新する。またＬＭＳ演算器５ｅは、マイクロホン１ｃからの騒音信号と減算器１７ａからの誤差信号により、減算器１７ａからの誤差信号を最小にするように適応フィルタ３ｅの係数を更新する。これによって、減算器１７ａの出力信号に含まれる騒音成分が低減されることになる。

【０１０９】次に、減算器１７ａの出力は、遅延器１６ｃで所定時間の、適応フィルタ３ｃが因果律を満たす時間だけ遅延され、減算器１７ｂに入力される。一方、マイクロホン１ａで検出された騒音信号は、遅延器１６ｂで所定時間の、音声判別器１３が音声を確実に判別するのに要する時間だけ遅延され、適応フィルタ３ｃで信号処理され、その出力が減算器１７ｂ入力される。減算器１７ｂでは、遅延器１６ｃの出力から適応フィルタ３ｃの出力が減算され、その結果の信号をスイッチ１５ｃ及び１５ｄに入力する。これによって、減算器１７ｂの出力信号に含まれる騒音成分が低減されることになる。

【０１１０】ここで、マイクロホン１ａ、１ｆは実施例７の場合と同様に、図１１あるいは図１６に示されるように構成されており、そのときの制御効果は、図２０のグラフに示すようにとなる。また、音声判別器１３および受信信号判別器１４による適応フィルタ３ａ、３ｂ及び３ｃの係数更新動作の制御と、ミュート制御回路１９とスイッチ１５ｄの制御、および通信機１０の送信用電源の制御については、その動作と効果は図２３に示す実施例７の場合と同様である。

【０１１１】以上により、無言時あるいは通信機１０からの受信信号が無い場合、本実施例の騒音制御ヘッドセットは、図１の実施例１で説明したように耳元での騒音を低減する。そして無言時に適応フィルタ３ｃ、３ｄ及び３ｅによりマイクロホン１ｆからの検出信号に含まれる周囲騒音信号を低減できる。音声発生時あるいは通信機１０からの受信信号がある場合には、騒音を低減しながら適応フィルタ３ａ及び３ｂが不安定にならないようにできる。そして音声発生時に音声信号に混入する周囲騒音成分を低減しながら適応フィルタ３ｃが不安定にならないように、あるいは音声信号を低減してしまわないようにできる。また、無言時に余計な周囲騒音を相手に送信しないことと通信機１０の消費電力を低減することが可能となる。さらに、ミュート制御回路１９の出力により通信機１０の送信用電源を制御することにより、スイッチ１５ｄが導通状態の期間と送信用電源が入っている期間が同じになるので、音声信号を途切れることなく相手に送信できる。

【０１１２】また、遅延器１６ａ及び１６ｃと減算器１７ａ及び１７ｂを設けたことにより、適応フィルタ３ｃ、３ｄ及び３ｅに実現するべき係数のインパルス応答を最適に調整することができる。なお、本実施例ではヘッドホン７ａ及び７ｂをヘッドバンド８で結合する構成としているが、図７に示すようにヘルメット１１に取り付ける構成でもよい。また、ヘッドホン７ａ及び７ｂについては、密閉型でも開放型でもよく、さらにイヤホン型でもよい。

【０１１３】また、加算器２ａ及び２ｂの処理の後でアナログの加算信号をデジタル信号に変換すれば、ＡＤコンバータが削減できる。さらに、音声判別器１３と受信信号判別器１４は、マイクロホン１ｆからの音声信号と通信機１０からの受信信号をデジタル化して所定の音声検出アルゴリズムにより音声判別してもよいが、例えば音声信号をアナログ信号のまま、コンパレータによって基準しきい値と比較して判別する構成でもよい。

【０１１４】また、マイクロホン１ａとマイクロホン１ｆの振動板直径は、同じでもよいが、マイクロホン１ｆの振動板直径の方がマイクロホン１ａの２倍以上あれば、より好ましい。さらにＬＭＳ演算器５ａ、５ｂ及び５ｃは、Ｆｘフィルタ４ａ及び４ｂの出力および遅延器１６ｂの出力により、マイクロホン１ｄ及び１ｅと減算器１７からの誤差信号を正規化する、いわゆる学習同定法であってもよい。

【０１１５】通信機１０は、騒音制御ヘッドセットとして内蔵していても、あるいは外部の通信機を使用してもよい。加えて、図２１で説明したように、図２４においてＬＭＳ演算器５ｄ及び５ｅの誤差信号をスイッチ１５ｅ及び１５ｆを設けて制御するようにしてもよい。さらに、実施例１の図８で説明したように、図２４において加算器２ａ及び２ｂがマイクロホン１ａ及び１ｃの検出信号を全て加算する構成でもよい。

【０１１６】

【発明の効果】以上のように図１の実施例について説明

した第１の発明によれば、 口元近傍の第１の騒音検出器と左右のヘッドホンに設置された第２および第３の騒音検出器からの騒音信号を、それぞれ第１および第２の騒音信号加算器で加算し、その出力を第１および第２の適応フィルタのリファレンス信号とする。これによって、使用者が騒音制御ヘッドセットを装着した状態で移動することによる騒音源と制御点の相対位置の変化があった場合や、外部騒音が複数存在する場合や、残響時間の長い反射性の部屋あるいは大きな部屋の場合にも、広帯域ランダム騒音の制御、あるいは周期性騒音の制御を安定に行うことができ、しかも消音効果を十分に得ることができる。さらに音声検出器の出力と第１および第２の適応フィルタの出力と通信機からの受信信号を第１および第２の出力信号加算器で加算して、左側ヘッドホンと右側ヘッドホンより再生する。これによって、自己の音声と通信機を介した相手の音声のモニタを実現できる。

【０１１７】また図９の実施例について説明した第２の発明によれば、第１の発明の効果に加えて、音声判別器と受信信号判別器により自己の音声発生時あるいは受信信号検出時には第１および第２の適応フィルタのそれぞれの第１および第２の係数更新器の係数更新動作を停止することにより、自己音声および相手音声のモニタによる消音制御時の誤動作を防止する。また自己の音声発生時以外は通信機へ自己音声を送信せず、さらに通信機の省電力化を可能とする。さらに図１０の実施例について説明した第３の発明によれば、第２の発明の効果に加えて、第３の適応フィルタにより自己の音声信号に含まれる周囲騒音を能動的に低減し、さらに音声判別器により自己の音声発生時の第３の適応フィルタの誤動作を防止することができる。

【０１１８】さらに図１７の実施例について説明した第４の発明によれば、第３の発明の効果に加えて、第１および第２の遅延器を用いて自己の音声の音声判別に有する時間を補償する。これによって、自己の音声のモニタよる消音制御時の誤動作および自己の音声信号の騒音制御時の誤動作の防止を確実なものにすることができる。さらにミュート制御回路で音声判別器からの音声検出信号がなくなってから一定時間後に自己の音声をミュートすることにより、自己の音声のモニタおよび通信機への送信時における音声信号の欠落を防止することである。

【０１１９】また図１９の実施例について説明した第５の発明によれば、音声信号に含まれる周囲騒音を低減するために、第１、第２および第３の騒音検出器の信号をそれぞれ第３、第４および第５の適応フィルタのリファレンス信号として用い音声検出器からの出力を制御することにより、第３の発明の効果をさらに精度よく実現できる。本発明の第６の発明によれば、第４の発明の効果に加えて、音声信号に含まれる周囲騒音の信号を低減するために、第１、第２および第３の騒音検出器の信号をそれぞれ第３、第４および第５の適応フィルタのリファレンス信号として用いて音声検出器からの出力を制御することにより、第４の発明の効果をさらに精度よく実現することができる。

【図面の簡単な説明】

【図１】本発明の実施例１の騒音制御ヘッドセットを示すブロック図である。

【図２】（ａ）は各マイクロホンの設置場所を示す右側面図である。（ｂ）は各マイクロホンの設置場所を示す平面図である。（ｃ）は各マイクロホンの設置場所を示す左側面図である。

【図３】各マイクロホンと騒音到来方向の位置関係を示す図である。

【図４】拡散音場でのヘッドホンの騒音制御効果を示すグラフである。

【図５】直接音場でのヘッドホンの騒音制御効果を示すグラフである。

【図６】１ＫＨｚの周期騒音のある音場でのヘッドホンの騒音制御効果を示すグラフである。

【図７】（ａ）は本発明のヘルメットへの応用を示す平面図である。（ｂ）は本発明のヘルメットへの応用を示す左側面図である。

【図８】実施例１の他の例を示すブロック図である。

【図９】本発明の実施例２の騒音制御ヘッドセットを示すブロック図である。

【図１０】本発明の実施例３の騒音制御ヘッドセットを示すブロック図である。

【図１１】（ａ）は実施例３におけるマイクロホン１ａとマイクロホン１ｆの位置関係を示す正面図である。（ｂ）は実施例３におけるマイクロホン１ａとマイクロホン１ｆの位置関係を示す側面図である。（ｃ）は実施例３におけるマイクロホン１ａとマイクロホン１ｆの位置関係を示す背面図である。

【図１２】実施例３におけるマイクロホン１ａとマイクロホン１ｆの指向性を示す図である。

【図１３】拡散音場での音声マイクの騒音制御効果を示すグラフである。

【図１４】騒音制御の影響による音声マイクの特性変化を示すグラフである。

【図１５】直接音場での音声マイクの騒音制御効果を示すグラフである。

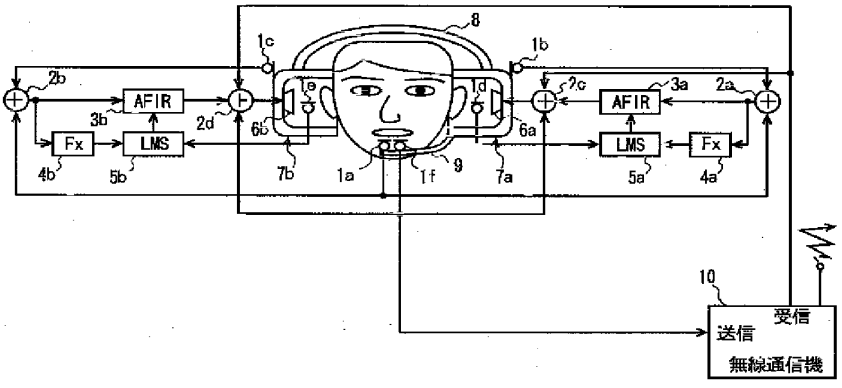
【図１６】（ａ）は実施例３の他の例におけるマイクロホン１ａとマイクロホン１ｆの位置関係を示す正面図である。（ｂ）は実施例３の他の例におけるマイクロホン１ａとマイクロホン１ｆの位置関係を示す側面図である。（ｃ）は実施例３の他の例におけるマイクロホン１ａとマイクロホン１ｆの位置関係を示す背面図である。

【図１７】本発明の実施例４の騒音制御ヘッドセットを示すブロック図である。

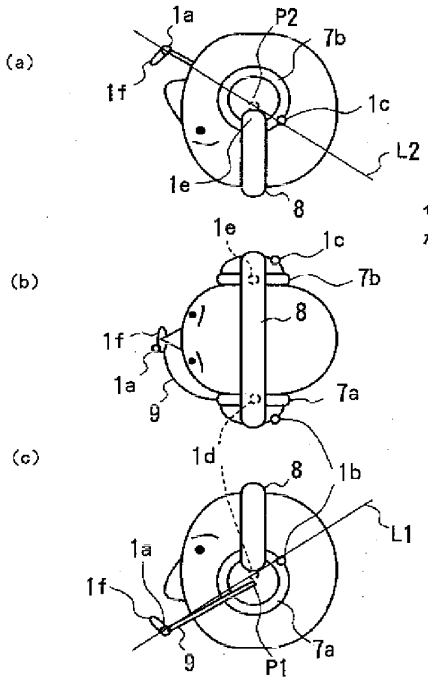
【図１８】実施例４における音声判別回路の１例を示す

ブロック図である。
【図１９】本発明の実施例５の騒音制御ヘッドセットを示すブロック図である。
【図２０】拡散音場での音声マイクの騒音制御効果を示すグラフである。
【図２１】実施例５の他の例を示すブロック図である。
【図２２】本発明の実施例６の騒音制御ヘッドセットを示すブロック図である。
【図２３】本発明の実施例７の騒音制御ヘッドセットを示すブロック図である。
【図２４】本発明の実施例８の騒音制御ヘッドセットを示すブロック図である。
【図２５】従来の騒音制御ヘッドホンを示すブロック図である。
【図２６】従来の適応フィルタを示すブロック図である。
【図２７】従来の騒音制御ヘッドホンの他の例を示すブロック図である。
【図２８】従来の騒音制御ヘッドホンの更に他の例を示すブロック図である。
【図２９】従来の騒音制御ヘッドホンの更に他の例を示すブロック図である。
【符号の説明】
１ａ、１ｂ、１ｃ、１ｄ、１ｅ、１ｆ　マイクロホン
２ａ、２ｂ、２ｃ、２ｄ　加算器
３ａ、３ｂ、３ｃ、３ｄ、３ｅ　適応フィルタ
４ａ、４ｂ　Ｆｘフィルタ
５ａ、５ｂ、５ｃ、５ｄ、５ｅ　ＬＭＳ演算器
６ａ、６ｂ　スピーカ
７ａ、７ｂ　ヘッドホン
８　ヘッドバンド
９　支持アーム
１０　通信機
１１　ヘルメット
１３　音声判別器
１４　受信信号判別器
１５ａ、１５ｂ、１５ｃ、１５ｄ、１５ｅ、１５ｆ　スイッチ
１６、１６ａ、１６ｂ、１６ｃ　遅延器
１７、１７ａ、１７ｂ　減算器
１８　吸音材
１９　ミュート制御回路
２０　バンドパスフィルタ
２１ａ、２１ｂ　パワー演算器
２２ａ、２２ｂ　比較器
２３　マイクロホン
２４　マイクロホン
２５　ヘッドホン
２６　適応処理システム
２７　アンプ
２８　耳
２８Ａ　外耳道
２９　減算器
３０　システム
３１　ヘッドホン
３２　マイクロホン
３２ａ　誤差信号線
３３　スピーカ
３３ａ　制御信号線
３４　信号処理装置
３４ａ　接続ケーブル
３４ｂ　プラグ
３５　検出センサ
３５ａ　増幅器
３６　ミキサ
３６ａ　出力ジャック
３７ａ　無線受信機
３７ｂ　無線送信機
３８－１、３８－ｎ　機械

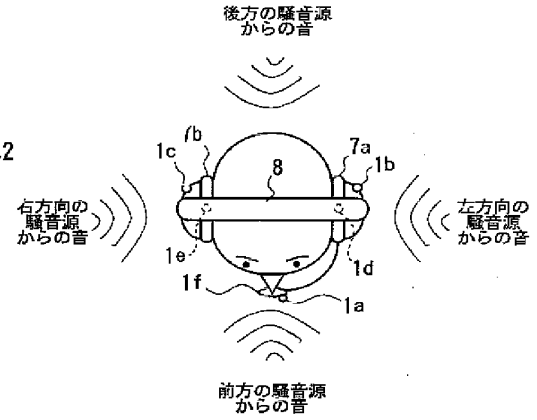
【図１】



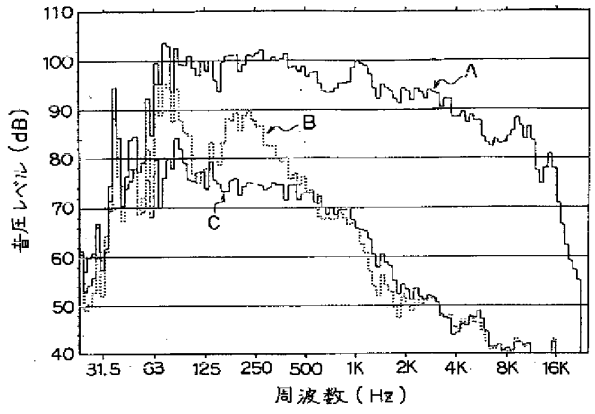
【図２】


(a)
1a
P2
7b
1f
1c
L2
1e
8
(b)
1e
1c
7b
8
1f
1a
9
7a
(c)
1d
8
1b
L1
1a
1f
9
P1
7a

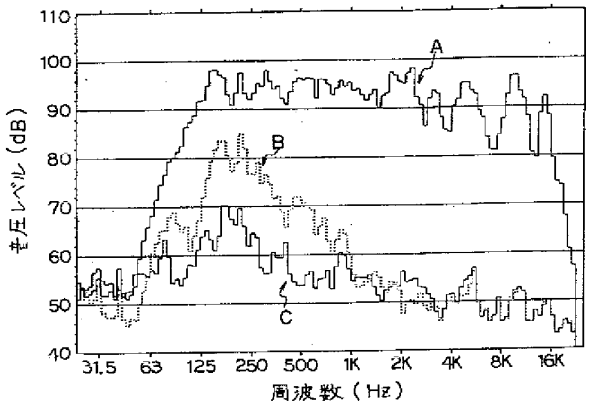
【図３】


後方の騒音源
からの音
右方向の
騒音源
からの音
左方向の
騒音源
からの音
前方の騒音源
からの音
1c
7b
8
7a
1b
1e
1d
1f
1a

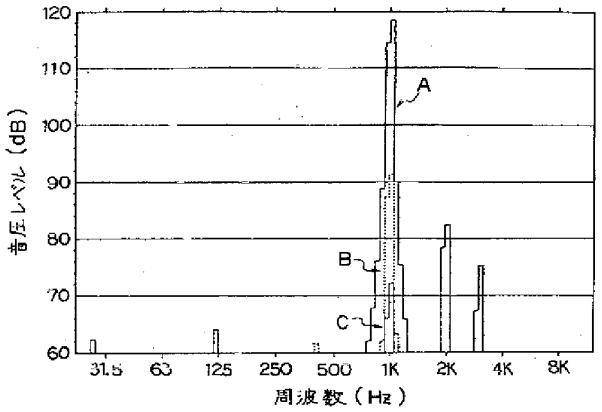
【図４】


110
100
90
80
70
60
50
40
音圧レベル（dB）
A
B
C
31.5
63
125
250
500
1K
2K
4K
8K
16K
周波数（Hz）

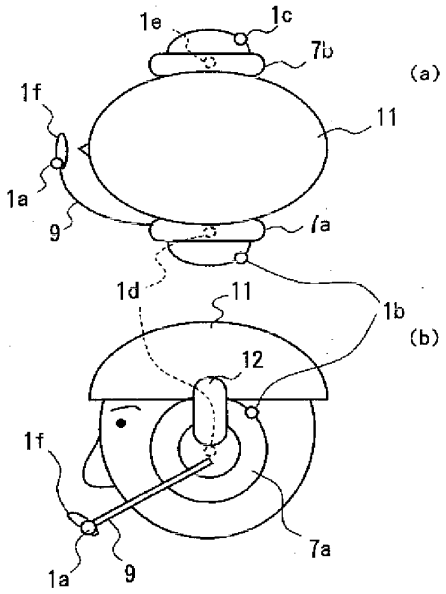
【図５】


110
100
90
80
70
60
50
40
音圧レベル（dB）
A
B
C
31.5
63
125
250
500
1K
2K
4K
8K
16K
周波数（Hz）

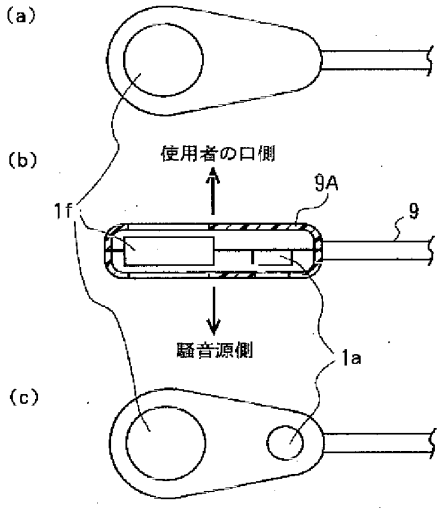
【図６】


120
110
100
90
80
70
60
音圧レベル（dB）
A
B
C
31.5
63
125
250
500
1K
2K
4K
8K
周波数（Hz）

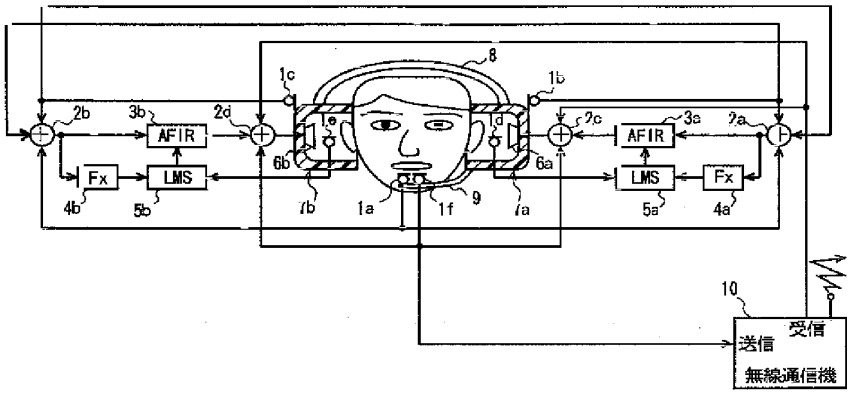
【図７】


1e
1c
7b
(a)
1f
11
1a
9
1a
1d
11
1b
(b)
12
1f
7a
1a
9

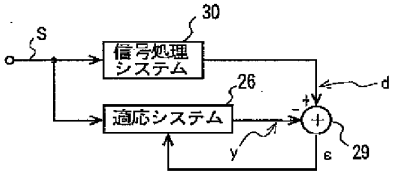
【図１１】


(a)
(b)
使用者の口側
9A
1f
9
騒音源側
1a
(c)

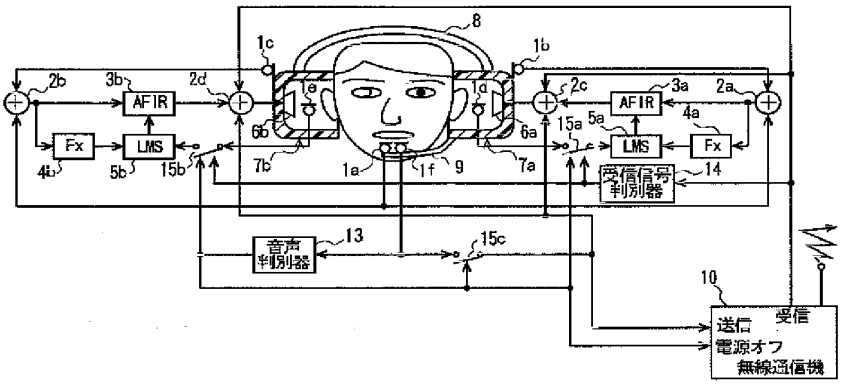
【図８】


8
1c
1b
2b
3b
AFIR
2d
2c
3a
2a
Fx
LMS
6b
6a
LMS
Fx
4b
5b
1b
1a
1f
9
1a
5a
4a
10
送信
受信
無線通信機

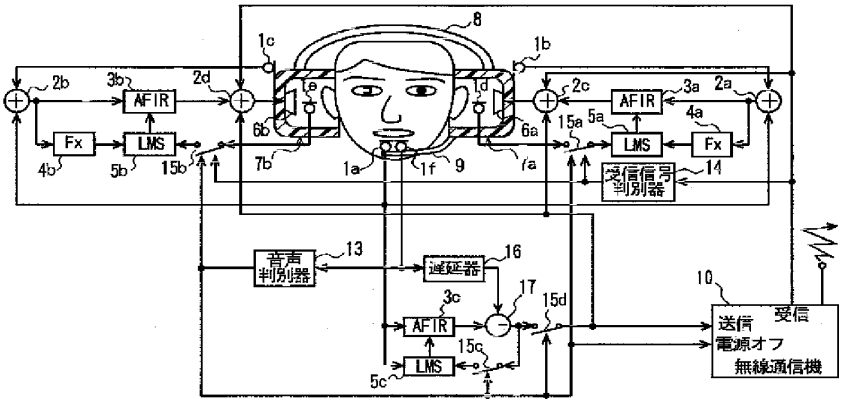
【図２６】


S
30
信号処理システム
26
適応システム
d
y
ε
29

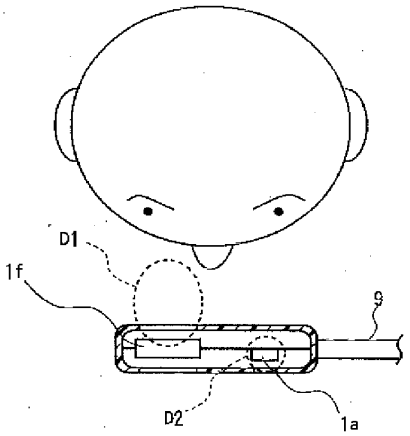
【図９】



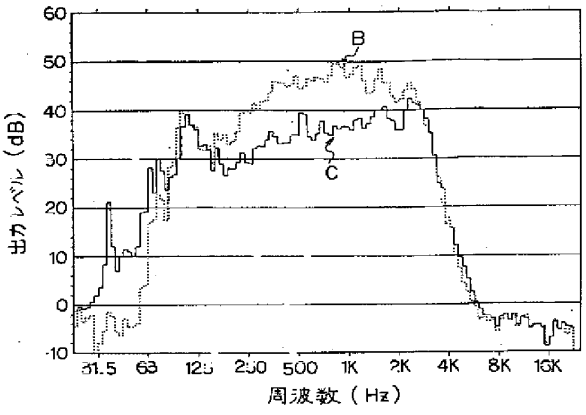
【図１０】



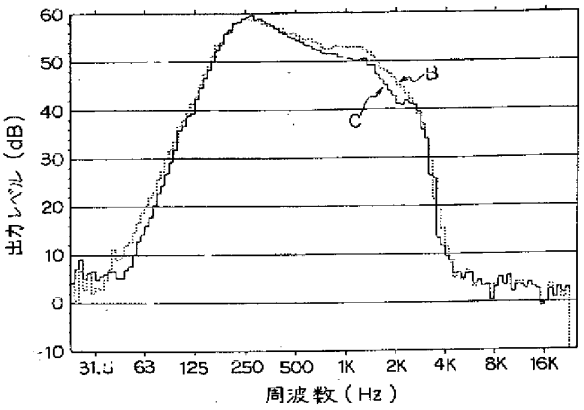
【図１２】



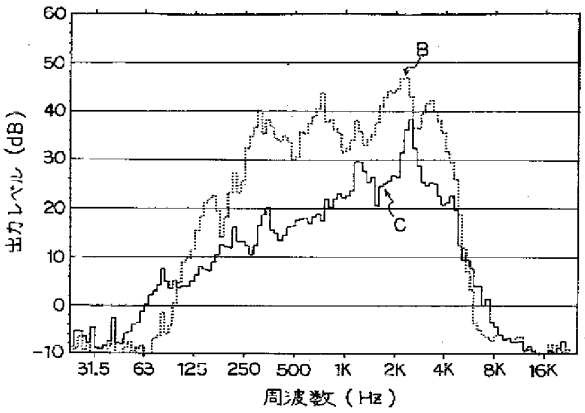
【図１３】



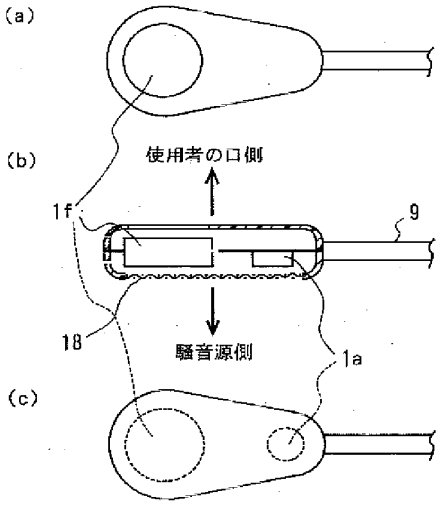
【図１４】



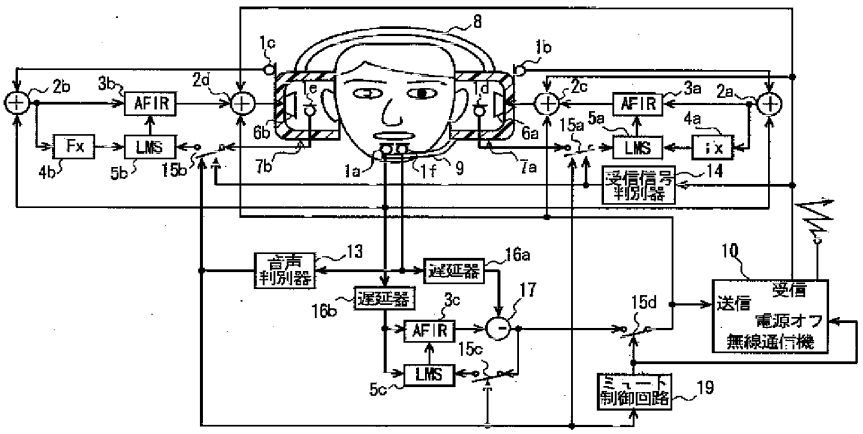
【図１５】



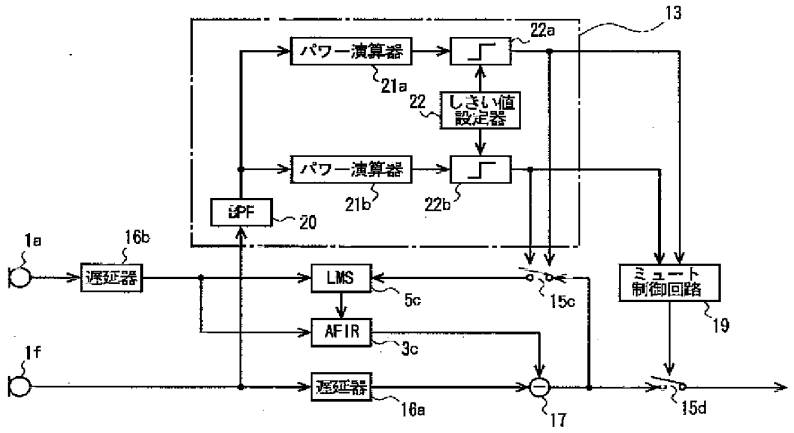
【図１６】



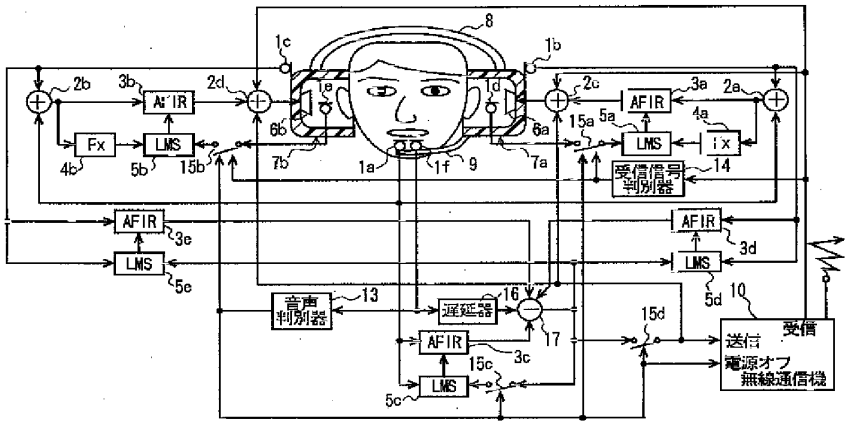
【図１７】



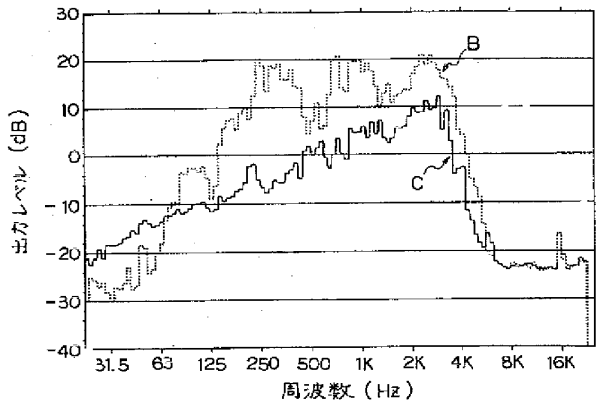
【図１８】



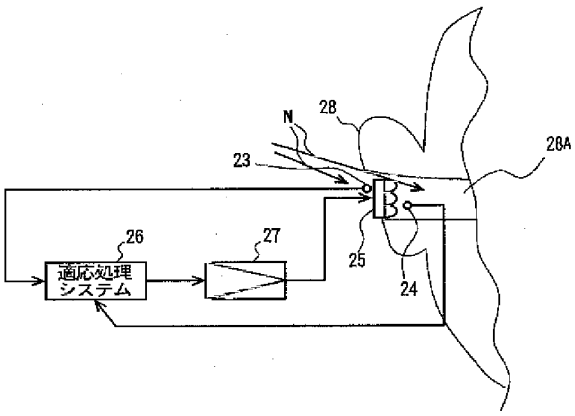
【図１９】



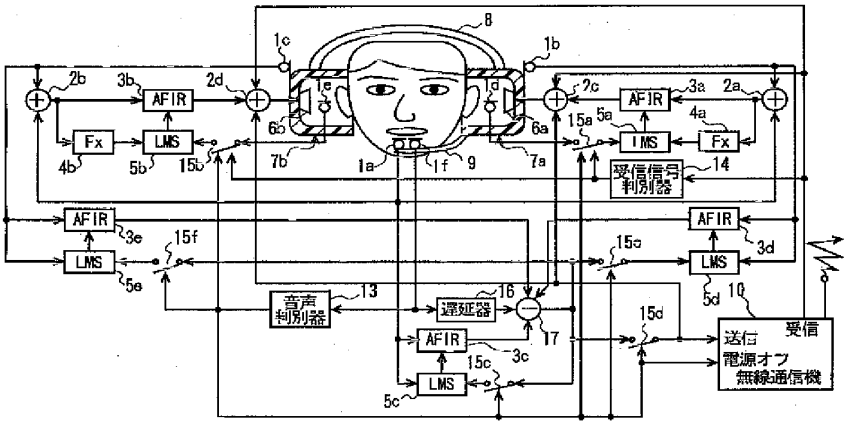
【図２０】



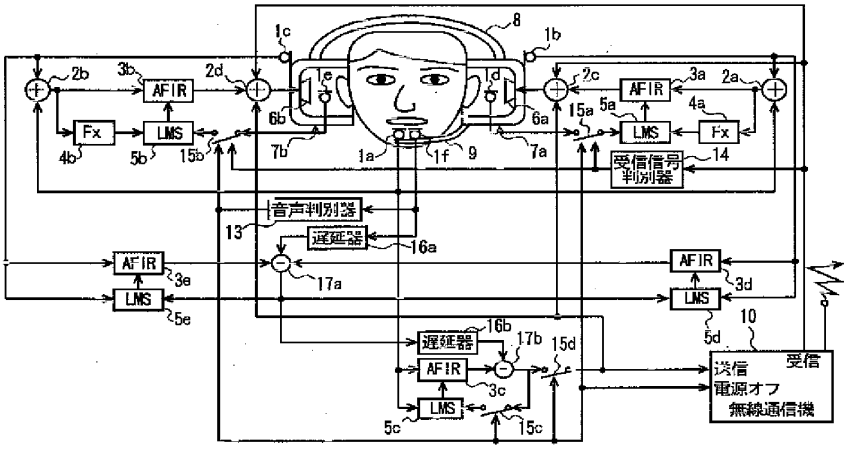
【図２５】



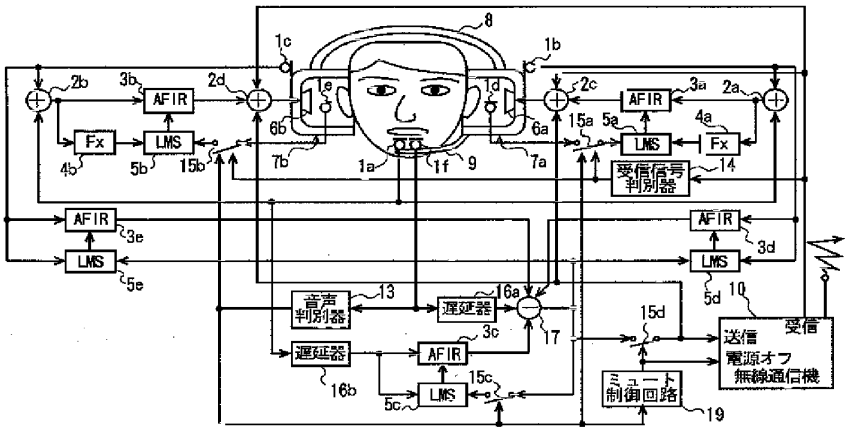
【図２１】


8
1c
1b
2b
3b
AFIR
2d
1e
1d
2c
3a
2a
6b
6a
15a
5a
4a
Fx
LMS
4b
5b
15b
7b
1a
1f
9
7a
受信信号判別器
14
AFIR
3e
15f
LMS
5e
15c
AFIR
3d
LMS
5d
10
音声判別器
13
遅延器
16
17
AFIR
3c
15c
15d
送信
受信
電源オフ
無線通信機
LMS
5c

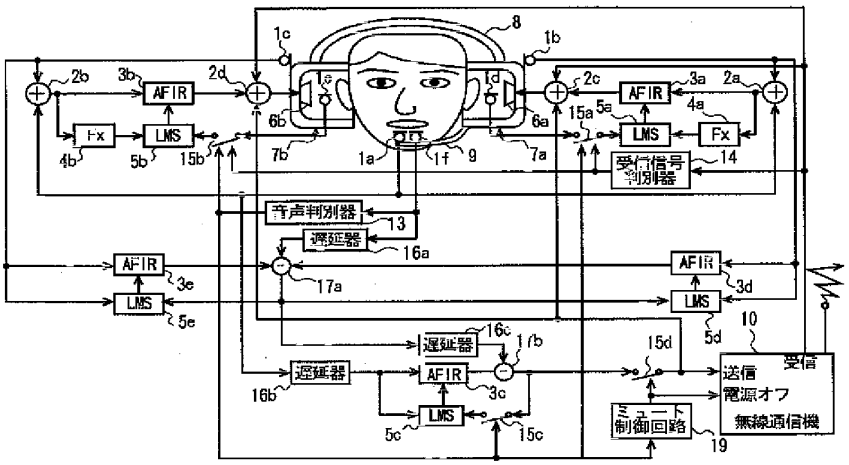
【図２２】


8
1c
1b
2b
3b
AFIR
2d
1e
1d
2c
3a
2a
6b
6a
15a
5a
4a
Fx
LMS
4b
5b
15b
7b
1a
1f
9
7a
受信信号判別器
14
音声判別器
13
遅延器
16a
17a
AFIR
3e
AFIR
3d
LMS
5e
LMS
5d
10
遅延器
16b
17b
15d
AFIR
3c
送信
受信
電源オフ
無線通信機
LMS
5c
15c

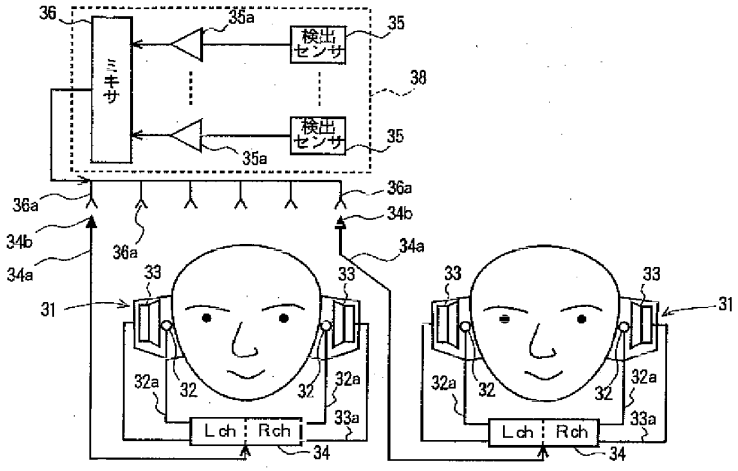
【図２３】


8
1c
1b
2b
3b
AFIR
2d
1e
1d
2c
3a
2a
AFIR
6b
6a
15a
5a
4a
Fx
LMS
Fx
LMS
15b
4b
5b
7b
1a
1f
9
7a
受信信号判別器
14
AFIR
3e
AFIR
3d
LMS
5e
LMS
5d
10
音声判別器
13
遅延器
16a
15d
受信
送信
電源オフ
無線通信機
遅延器
AFIR
3c
17
16b
5c
LMS
15c
ミュート制御回路
19

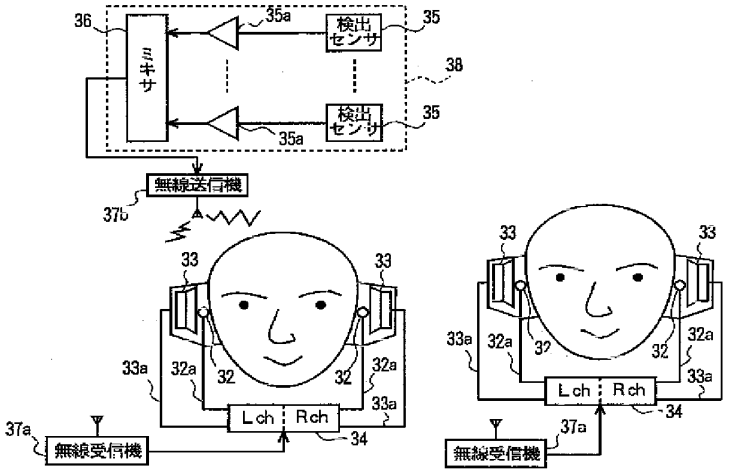
【図２４】


8
1c
1b
2b
3b
AFIR
2d
1e
1d
2c
3a
2a
AFIR
6b
6a
15a
5a
4a
Fx
LMS
Fx
LMS
15b
4b
5b
7b
1a
1f
9
7a
受信信号判別器
14
音声判別器
13
遅延器
16a
AFIR
3e
17a
AFIR
3d
LMS
5e
LMS
5d
10
16c
遅延器
1fb
15d
受信
送信
遅延器
AFIR
3c
電源オフ
16b
LMS
5c
15c
ミュート制御回路
19
無線通信機

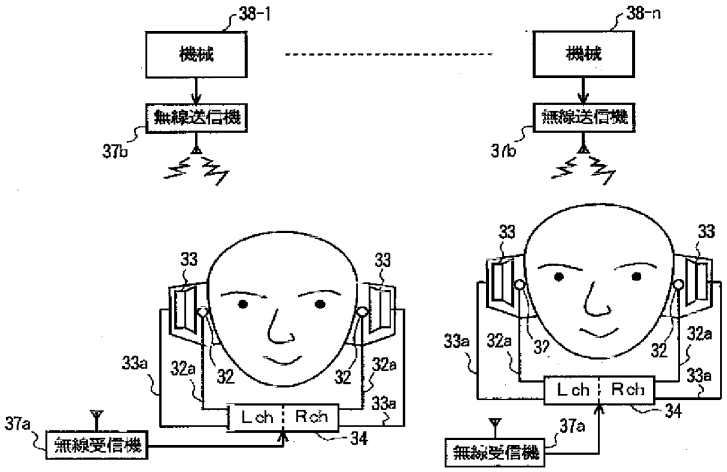
【図２７】


36
35a
検出センサ
35
38
ミキサ
36a
34b
34a
33
31
32a
32
Lch
Rch
34
33a

【図２８】


36
35a
検出センサ
35
38
ミキサ
無線送信機
37b
33
32a
32
33a
Lch
Rch
34
37a
無線受信機

【図２９】



フロントページの続き

(51)Int.Cl.[6]		識別記号	FI	
Ｈ０４Ｒ	1/10	１０１	Ｈ０４Ｒ　5/033	Ｚ
	5/033		Ｇ１０Ｋ　11/16	Ｈ

(72)発明者　毛利　哲也
大阪府門真市大字門真1006番地　松下電器
産業株式会社内

(19) BUNDESREPUBLIK DEUTSCHLAND

DEUTSCHES PATENTAMT

(12) **Offenlegungsschrift**

(10) **DE 43 12 155 A1**

(51) Int. Cl.[5]: G 10 L 5/00

(21) Aktenzeichen: P 43 12 155.1
(22) Anmeldetag: 14. 4. 93
(43) Offenlegungstag: 20. 10. 94

DE 43 12 155 A1

(71) Anmelder:
Hiller, Friedrich, Dipl.-Ing., 4400 Münster, DE

(72) Erfinder:
gleich Anmelder

(54) Verfahren und Vorrichtung zur Verbesserung der Erkennungsleistung und Erhöhung der Zuverlässigkeit bei der automatischen Spracherkennung in geräuschvoller Umgebung

(57) Anordnung von zwei Luftschallmikrofonen in Verbindung mit einem Verfahren der Subtraktion spektraler Energiekomponenten der beiden Mikrofonsignale zur Verbesserung der Erkennungsleistung bei der automatischen Spracherkennung in geräuschvoller Umgebung.
Durch die Anordnung zweier baugleicher Mikrofone an einem am Kopf getragenen Geschirr wird erreicht, daß beliebiger, tonfrequenter Störschall aus dem umgebenden Raum in seiner nachteiligen Wirkung auf die Arbeitsweise des Spracherkenners deutlich abgeschwächt wird.
Durch die Detektion von Wortgrenzen und Bildung der Merkmale aus den spektralen Differenzkomponenten wird sowohl die Zuverlässigkeit von Sprachsteuerungssystemen als auch die Erkennungsleistung von allgemeinen Spracheingabesystemen erhöht.
Die Beherrschung auch sehr starker, pulsartiger Geräusche, wie sie in industrieller Umgebung oder in Fahrzeugen auftreten, und die kostengünstige Realisierung kennzeichnen die Verbesserungen gegenüber gattungsgemäßen Verfahren.

DE 43 12 155 A1

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

BUNDESDRUCKEREI 08. 94 408 042/127 6/31

Beschreibung

Die Erfindung bezieht sich auf ein Verfahren nach dem Oberbegriff des Verfahrensanspruchs 1 und auf eine Vorrichtung zur Durchführung des erfindungsgemäßen Verfahrens, welche die Merkmale des Oberbegriffs des Vorrichtungsanspruchs 5 aufweist.

Die sichere automatische Erkennung gesprochener Sprache bereitet bei gleichzeitig vorhandenen Umgebungsgeräuschen Schwierigkeiten. Hierbei wird unter Sprache eine beliebige lautliche Äußerung (Wort, Silbe, Pfeifton ...) verstanden, die von der nutzenden Person abgegeben wird, um Funktionselemente zu steuern (sog. Sprachsteuerungssystem) oder um Text und/oder Daten in einen Rechner einzugeben ("hörende Schreibmaschine"). Sprachsteuerungssysteme arbeiten mit Einzelworten, deren Beginn und Ende durch Sprechpausen zwischen den Worten festgestellt werden kann (Isoliertworterkenner). Bei der Texteingabe wird auch bereits Fließsprache verwendet.

Tritt in den Sprechpausen Störschall auf, kann dieser als Laut (fehl-)interpretiert werden und bei genügender Ähnlichkeit zu einem verwendeten Wort unerwünscht eine Funktion bewirken — die Zuverlässigkeit der Sprachsteuerung ist dann ungenügend. Tritt das Geräusch gleichzeitig beim Sprechen eines Wortes auf, so kann es durch Überlagerung beider Schalle zu einer Falschinterpretation kommen. Durch Wortverwechslungen oder Rückweisungen wird die Erkennungsleistung dabei unerwünscht herabgesetzt.

Es sind einkanalige Störunterdrückungssysteme bekannt, wie sie an Telefonleitungen vorzugsweise Verwendung finden. Dabei wird unterstellt, daß Eigenschaften später auftretender Geräusche vorab bereits bekannt sind. In der DE 36 42 591 C2 wird mit max. 12 verschiedenen Geräuschmustern gearbeitet. Bei anderen Verfahren wird das in den Sprechpausen analysierte Geräusch als im nachfolgenden Wort weiterhin wirksam unterstellt. Diese Verfahren versagen jedoch, wenn unbekannte Geräusche gleichzeitig beim Sprechen auftreten.

Es werden deshalb auch mehrkanalige Verfahren (mindestens 2 Mikrofone im Raum) angewendet. In der DE 39 29 481 C2 wird mit mindestens drei Luftschallmikrofonen und einem neuronalen Netzwerk ein sehr aufwendiges Verfahren berichtet. Die Anordnung nach DE 36 10 797 A1 benutzt zwei ortsfeste Luftschallmikrofone in erheblichem räumlichen Abstand. Das "Sprach"-Mikrofon nahe am Mund wandelt Laute und das Geräusch — das abgelegene "Geräusch"-Mikrofon Geräusche samt dem stark geschwächten Sprachschall. Ein vorhandenes Sprachsignal wird eindeutig aus der positiven Energiedifferenz beider Kanäle abgeleitet und schaltet die Signale vom "Sprach"-Mikrofon zum Erkenner durch. Nachteile dieser Anordnung:

— Geräusche während des Sprechens wirken unvermindert,
— inhomogener Störschall im Raum kann Durchschalten bewirken,
— Bewegungen der Person (gegenüber dem "Sprach"-Mikrofon) verändern Signalamplitude und -spektrum nachteilig.

In der DE 37 42 929 C1 befindet sich ein Luftschallmikrofon an einem Geschirr in konstanter Nähe zum Mund. Ein weiteres Körperschallmikrofon am Körper der Person wird mit beim Sprechen erzeugten Körperschall beaufschlagt und schaltet Sprache samt Geräusch des Luftschallmikrofons zum Erkenner durch. Hierbei wird zwar eine Variation des Sprachschalls bei Bewegungen der Person vermieden und ein reines Geräusch kann keine Funktion auslösen, Störschall beim Sprechen wird beim Erkenner jedoch auch hier die Erkennungsrate vermindern.

Durch die gewählte Mikrofonposition und dessen Richtwirkung ist der Signal-/Geräuschabstand an diesem Wandler jedoch schon optimiert.

Weiterhin ist bekannt, daß ein identisch in zwei Kanälen vorhandenes Störsignal durch Differenzbildung der entsprechenden Energiekomponenten einer Spektralrepräsentation des jeweiligen Kanals entfernt werden kann (spektrale Subtraktion phasenunabhängiger Größen). Sowohl die gleichgerichteten und tiefpaßgefilterten Ausgangssignale einer Filterbankanordnung als auch das Betragsspektrum der Fourier-Transformierten ist hierfür geeignet.

Der Erfindung liegt die Aufgabe zugrunde, eine kostengünstige Anordnung mit zwei Luftschallmikrofonen zu schaffen, die in Verbindung mit dem Verfahren der spektralen Subtraktion die zuverlässige Erkennung von Sprache auch bei gleichzeitig auftretendem beliebigen, auch sehr lautem pulsartigem Störgeräusch ermöglicht.

Die Erfindung löst diese Aufgabe durch die kennzeichnenden Merkmale des Verfahrensanspruchs 1 bzw. des Vorrichtungsanspruchs 5.

Durch die Befestigung von zwei baugleichen Luftschallmikrofonen an einer Haltevorrichtung am Kopf der nutzenden Person gemäß dem kennzeichnenden Teil des Anspruchs 5 werden die Voraussetzungen für eine wirksame Störgeräuschverminderung durch spektrale Subtraktion optimiert. Das Mikrofon in Mundnähe optimiert den Signal/Geräuschabstand im erzeugten Signal und vermeidet Signalvariationen bei Bewegungen der Person. Das Mikrofon am Hinterkopf ist räumlich hinreichend nahe am Ort des ersten Mikrofons um sicherzustellen, daß es in guter Näherung durch dasselbe Geräusch aus dem umgebenden Raum beaufschlagt wird. Andererseits wird das Sprachsignal durch den Kopf sehr gut abgeschattet, so daß dieses bei der Subtraktion nur geringfügig gedämpft wird. Dabei könnte die Haltevorrichtung z. B. als Helm oder Hut realisiert sein, um speziellen Einsatzfällen gerecht zu werden. Die leichte Handhabung der Mikrofonanordnung stellt dabei eine wichtige konstruktive Forderung dar.

Vorteilhafte Ausbildungen der Vorrichtung gemäß Anspruch 6 bis Anspruch 14 zeigt **Fig.** 2.

In der **Fig.** 1 ist das Verfahren nach Anspruch 1 dargestellt, wobei vorteilhafte Ausbildungen nach Anspruch 2, 3, 11, 12 und 14 skizziert sind:

Die Signale der Mikrofone (**1, 2**) werden über eine drahtlose Übertragungseinrichtung (**4**) an den Erkenner (**5**) geleitet. Im Eingangsteil (**5.1**) werden die elektrischen Signale den Vorverstärkern (**6, 9**) zugeführt, um anschließend in den Spektralanalysatoren (**8, 10**) verarbeitet zu werden. Die unterbrochenen Doppellinien kennzeichnen die Vorgabe, daß Parameter (Verstärkung, Analyseintervall ...) nur synchron in beiden Kanälen verändert werden. Das Laufzeitglied (**7**) wurde in einer realisierten Anordnung nach **Fig.** 2 mit einer Verzögerungszeit von $\tau = 0{,}7$ ms ausgestattet. Dadurch wird bei der Subtraktion der spektralen Energiekomponenten sichergestellt, daß die Anteile des Sprachsignals zeitrichtig zugeordnet sind. Dabei wurde die Laufzeit des Sprachsignals vom Mund → Mikrofon (**1**) von der Laufzeit Mund → Mikrofon (**2**) an einer Kopfnachbil-

dung bei Raumtemperatur subtrahiert.

Die Spektralanalyse wurde als Filterbank mit 16 Bandpässen mit einem Signalprozessor realisiert. Nach einer linearen Quantisierung mit 14 Bit an äquidistanten Zeitpunkten (50 µs) erfolgte eine Höhenanhebung, Hanningfensterung und eine 512-Punkte FFT. Die Fenster überlappten sich mit 50%. Die Amplituden der erhaltenen Spektrallinien wurden in Frequenzgruppen zusammengefaßt, die der Frequenzauflösung des menschlichen Ohres näherungsweise entsprach.

Da bei der nachfolgenden Subtraktion, besonders in Sprechpausen, negative Ergebnisse entstehen können, werden diese definitiv zu Null gesetzt. Ist der Taster (3) geöffnet, so werden keine (Differenz-) Signale zur weiteren Verarbeitung an den eigentlichen Erkenner weitergegeben ("Räuspertaste") Parallel zu den spektralen Energiewerten aus der Differenzbildung können Merkmale aus dem Zeitsignal erzeugt werden.

Im unteren Teil von **Fig.** 1 ist eine zweikanalige Signalisierung von tonfrequenten Schwingungen an Lautsprecher dargestellt. Hier werden vorteilhaft Hinweise aus der Sprachausgabeeinheit des Erkenners (in 5.2) drahtlos übertragen, um den Dialog mit dem Benutzer zu führen.

In der **Fig.** 2 wird die bevorzugte Ausformung der Mikrofonanordnung schräg von vorne (**Fig.** 2.1), von der Seite (**Fig.** 2.2) und von oben (**Fig.** 2.3) gezeigt. Dabei ist die Hauptrichtung der Mikrofonempfindlichkeit durch einen Pfeil mit einem "R" veranschaulicht. Die Platzierung des Sprachmikrofons (1) im Mundwinkel vermeidet Geräusche beim Ausatmen und sorgt für maximale Signalamplitude. Das Störgeräuschmikrofon (2) am Hinterkopf ist dort vom Sprachschall gut abgeschattet und wird mit typisch weniger als einem Zehntel des Schalldrucks an (1) beaufschlagt.

Da bei vielen handelsüblichen Kopfbügelmikrofonen bzw. Hör-/Sprechgarnituren das Mikrofon samt Aufnahme (15) lösbar mit dem Geschirr (13) verbunden ist, kann die dargestellte Anordnung leicht durch Hinzufügen des zweiten Mikrofons auf der anderen Seite des Geschirrs erhalten werden. Hiermit wird die wirtschaftliche Herstellung gewährleistet.

Die in den Ohrpolstern integrierten Hörer (11, 12) gestatten die akustische Dialogführung. Es kann jedoch z. B. auch Unterhaltungsmusik eingespeist werden. Auf eine Kopfnachbildung gesetzt, kann die Anordnung auch als Stereomikrofon für allgemeine Verwendung eingesetzt werden.

Eine Skala dient der Markierung (17) und eine Feststellvorrichtung (18) der Fixierung der Mikrofonposition, wenn diese verstellbar ist. Dies gewährleistet optimal reproduzierbare Signale. Die Markierung des Sprachmikrofons (14) vermeidet, daß das Geschirr falsch herum (Geräuschmikrofon (2) am Mund) getragen wird. Beim Gespräch mit Arbeitskollegen wird der Erkenner mit der Taste (3) inaktiviert. Das flexible Anschlußkabel (16) wird mit dem Stecker an den Erkenner oder die sendeseitige drahtlose Übertragungseinheit (4.1) gesteckt. Sie wird batteriegespeist am Körper getragen.

Patentansprüche

1. Verfahren zur Verbesserung der Erkennungsleistung und Erhöhung der Zuverlässigkeit bei der automatischen Erkennung gesprochener menschlicher Sprache in geräuschvoller Umgebung, durch Verringerung der Störgeräuschanteile im als Frequenzmuster dargestellten Sprachsignal, wobei ein erstes Mikrofon bestimmungsgemäß vornehmlich das Sprachsignal aufnimmt, wobei ein zweites Mikrofon bestimmungsgemäß vornehmlich mit dem Umgebungsgeräusch beaufschlagt wird, wobei die konditionierten Ausgangssignale beider Mikrofone verknüpft werden, um sowohl die fälschliche Interpretation eines reinen Umgebungsgeräusches als lautliche Äußerung zu vermeiden, als auch die Wahrscheinlichkeit einer Falschinterpretation einer lautlichen Äußerung bei gleichzeitig vorhandenem Umgebungsgeräusch zu vermindern, **dadurch gekennzeichnet,**
daß sich das erste der beiden baugleichen Luftschallmikrofone nahe am Mund befindet und mit dem Sprachsignal samt den Umgebungsgeräuschen beaufschlagt wird, während das zweite Mikrofon näherungsweise dasselbe Geräusch samt dem stark abgeschwächten Sprachsignal aufnimmt, wobei das Verhältnis der Energieanteile des Sprachsignals der beiden Mikrofone im Freifeld näherungsweise auch bei Bewegungen der nutzenden Person konstant bleibt,
daß die elektrischen Signale der Mikrofone einer synchronen spektralen Kurzzeitanalyse der Signalenergie zugeführt werden,
daß anschließend die entsprechenden spektralen Energiekomponenten des zweiten Kanals von denen des ersten subtrahiert werden, wobei negative Werte vermieden werden und diese Werte anschließend zur Wortgrenzendetektion und/oder Bildung von Merkmalsvektoren dem eigentlichen Erkenner zugeführt werden.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß sich im Signalpfad des ersten Mikrofons eine Verzögerungsschaltung (Laufzeitglied) befindet, dessen Verzögerungszeit die längere Schallaufzeit vom Mund zum zweiten Mikrofon kompensiert.

3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß bei der Subtraktion der entsprechenden spektralen Komponenten zunächst gebildete negative Werte identisch Null gesetzt werden.

4. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß bei der Subtraktion der entsprechenden spektralen Komponenten der Absolutbetrag des Ergebnisses weiterverarbeitet wird.

5. Vorrichtung zur Durchführung des Verfahrens nach Anspruch 1,
mit einem ersten, vornehmlich mit beim Sprechen erzeugten Luftschallwellen beaufschlagbaren Mikrofon,
mit einem zweiten, wenigstens mit Umgebungsgeräuschen beaufschlagbaren Mikrofon,
mit einer elektrischen Schaltung, welche die von den Mikrofonen abgegebenen elektrischen Signale geeignet verarbeitet und für die Merkmalsgewinnung im Spracherkennungsgerät bereitstellt, dadurch gekennzeichnet,
daß sich das erste der baugleichen Luftschallmikrofone nahe am Mund, das zweite Mikrofon am Hinterkopf der benutzenden Person befindet,
daß die räumliche Position der Mikrofone zueinander und zum Mund durch eine geeignete Haltevorrichtung am Kopf der Person weitgehend konstant gehalten wird,
daß die Signale der Mikrofone je einem baugleichen Vorverstärker und anschließend je einer bau-

gleichen und synchron arbeitenden Vorrichtung zur spektralen Kurzzeitanalyse der Signalenergie zugeführt werden,
daß die positiven Werte aus der spektralen Subtraktion zusammen mit dem Zeitsignal vom ersten Mikrofon an den Teil des Spracherkennungsgerätes geführt sind, der Wortgrenzen detektiert und/oder Merkmalsvektoren generiert.

6. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, daß das erste Mikrofon (1), wie auch das zweite Mikrofon (2), mit einer baugleichen Mechanik an einem Geschirr (13) befestigt ist, welches, wie in **Fig.** 2.2 schematisch dargestellt,von der sprechenden Person am Kopf getragen wird, daß die von den Mikrofonen abgehenden Anschlußkabel gebündelt (16) an einem geeigneten Stecker enden, der in das nicht räumlich dargestellte Spracherkennungsgerät (5) gesteckt werden kann.

7. Vorrichtung nach Anspruch 6, dadurch gekennzeichnet, daß die Mikrofone (1, 2) eine Richtcharakteristik aufweisen, (in **Fig.** 2 ist die Hauptachse der Richtwirkung mit "R" gekennzeichnet), daß durch eine verstellbare Mechanik in der Aufnahme der Mikrofonhalter am Geschirr (15) die Position von (1) zum Mundwinkel und auch die Position von (2) am Hinterkopf (mögl. wie in **Fig.** 2.2 gegenüber dem Mund) individuell optimal eingestellt werden kann.

8. Vorrichtung nach Anspruch 7, dadurch gekennzeichnet, daß die eingestellte Position der Mikrofone durch z. B. eine Skala (17) eindeutig gemerkt und reproduziert werden kann.

9. Vorrichtung nach Anspruch 7, dadurch gekennzeichnet, daß die gewählte Position der Mikrofone mit einer Feststellvorrichtung an der Aufnahme (15) fixiert werden kann (in **Fig.** 2.1 als (18) Schraubenkopf dargestellt).

10. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, daß das erste Mikrofon (1) gut sichtbar eindeutig markiert ist (14).

11. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, daß durch einen Schalter oder Taster (3), welcher als Öffner wirkt, der Erkenner manuell temporär inaktiviert werden kann.

12. Vorrichtung nach Anspruch 6, dadurch gekennzeichnet, daß in die Ohrpolster des Geschirrs elektrisch/akustische Wandler (11, 12) integriert sind, die elektrisch mit dem Erkenner verbunden, eine Dialogführung gestatten.

13. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, daß eine drahtlose Übertragungseinrichtung (4) die Mikrofonsignale von der Person zum Erkenner (5) überträgt, daß der batteriegespeiste Sendeteil (4.1) mittels einer Halterung am Körper getragen werden kann und die Mikrofonanordnung an diesem Sender angeschlossen wird.

14. Vorrichtung nach Anspruch 13, dadurch gekennzeichnet, daß für die Ansteuerung der Hörer (11, 12), die drahtlose Übertragungseinrichtung bidirektional ausgeführt ist.

Hierzu 1 Seite(n) Zeichnungen

– Leerseite –

Nummer: DE 43 12 155 A1
Int. Cl.[5]: G 10 L 5/00
Offenlegungstag: 20. Oktober 1994

Fig. 1

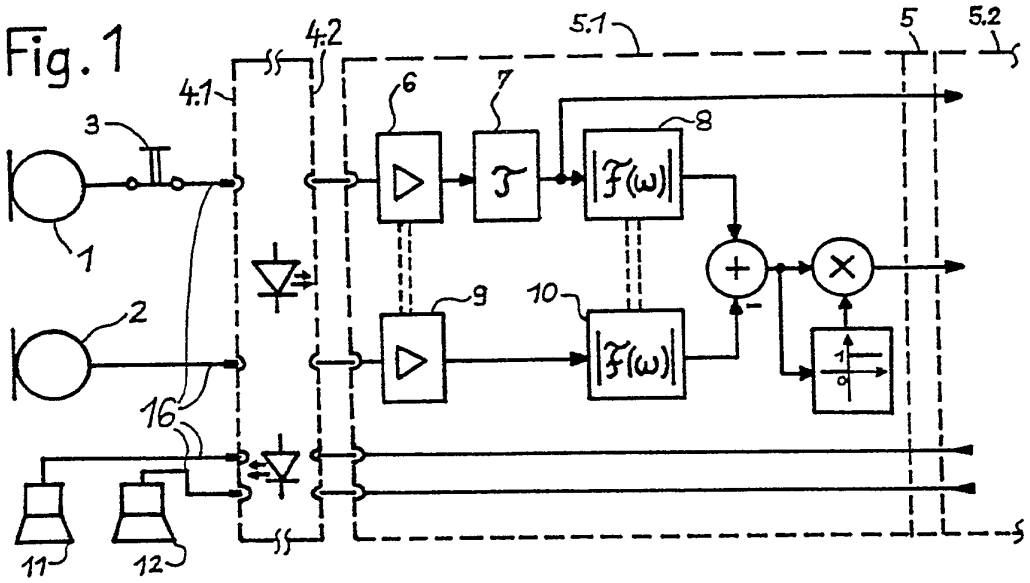


Fig. 2.1

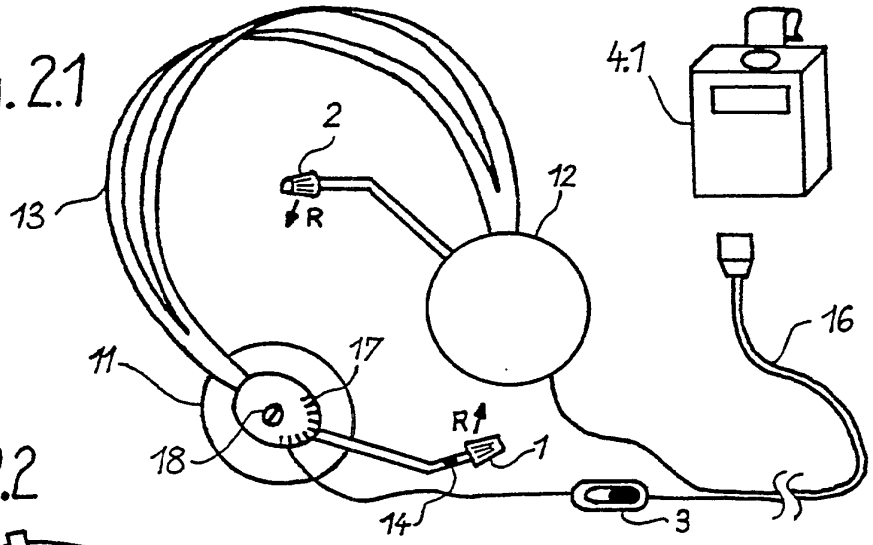


Fig. 2.2

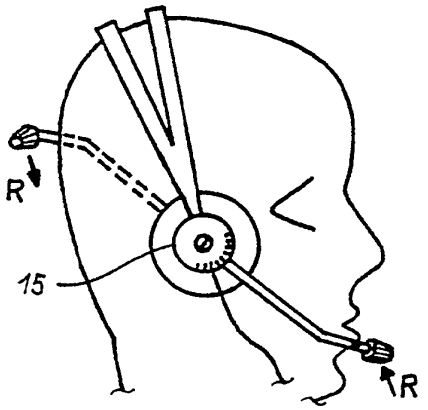
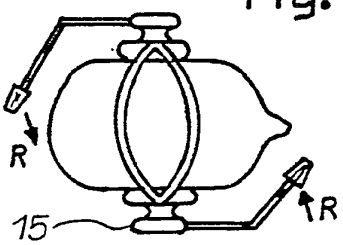


Fig. 2.3



PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : H03B 29/00

A1

(11) International Publication Number: WO 93/26085

(43) International Publication Date: 23 December 1993 (23.12.93)

(21) International Application Number: PCT/US92/04568

(22) International Filing Date: 5 June 1992 (05.06.92)

(71) Applicant *(for all designated States except US):* NOISE CANCELLATION TECHNOLOGIES [US/US]; Suite 101, 1015 West Nursery Road, Linthicum, MD 21090 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* HOHMAN, John [US/US]; 1545 Putty Hill Road, Towson, MD 21204 (US). EVANS, Robert [US/US]; 14 Medici Court, Baltimore, MD 21234 (US). BUSCH, Ralph [US/US]; 316 Boyd Avenue, Takoma Park, MD 20912 (US). SCOTT, Roy [US/US]; 10536 Cross Fox Lane E2, Columbia, MD 21044 (US).

(74) Agent: HINEY, James, W.; Noise Cancellation Technologies, Suite 101, 1015 West Nursery Road, Linthicum, MD 21090 (US).

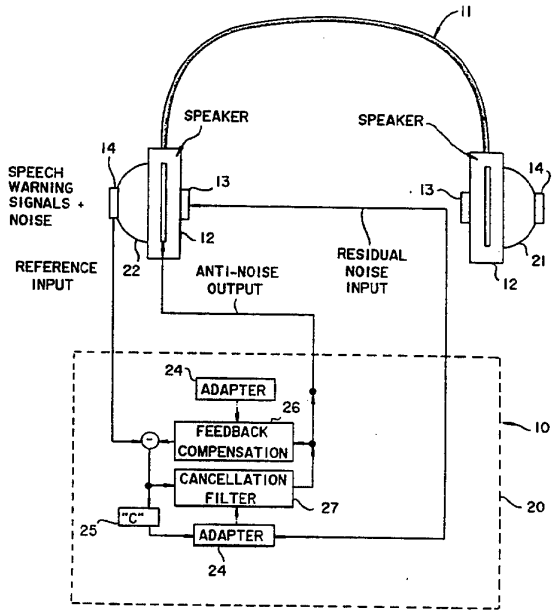
(81) Designated States: CA, JP, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LU, MC, NL, SE).

Published
With international search report.

(54) Title: ACTIVE/PASSIVE HEADSET WITH SPEECH FILTER

(57) Abstract

An active/passive headset with a speech filter characterized by a controller (20) processing speech, noise and warning signals input from microphones (14) on closed back muffs (20) and providing an anti-noise output by a digital virtual earth or adaptive feedforward algorithm to speakers (12) adjacent a user's ear.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**MR**	Mauritania
AU	Australia	**GA**	Gabon	**MW**	Malawi
BB	Barbados	**GB**	United Kingdom	**NL**	Netherlands
BE	Belgium	**GN**	Guinea	**NO**	Norway
BF	Burkina Faso	**GR**	Greece	**NZ**	New Zealand
BG	Bulgaria	**HU**	Hungary	**PL**	Poland
BJ	Benin	**IE**	Ireland	**PT**	Portugal
BR	Brazil	**IT**	Italy	**RO**	Romania
CA	Canada	**JP**	Japan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SK**	Slovak Republic
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SN**	Senegal
CM	Cameroon	**LI**	Liechtenstein	**SU**	Soviet Union
CS	Czechoslovakia	**LK**	Sri Lanka	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**MC**	Monaco	**UA**	Ukraine
DK	Denmark	**MG**	Madagascar	**US**	United States of America
ES	Spain	**ML**	Mali	**VN**	Viet Nam
FI	Finland	**MN**	Mongolia		

ACTIVE/PASSIVE HEADSET WITH SPEECH FILTER

This invention relates to the area of headsets worn by a user in specific environments and, in particular, those areas where the background noise must be abated while simultaneously allowing speech through.

Generally, environmental noise can be broken into several areas such as noises of short duration with varying spectral characteristics such as conversations, other long duration noise with fixed or relatively fixed spectral characteristics such as operating pumps, fans, engines etc., and noises of very brief duration with "spike" characteristics such as something rattling, etc.

The instant invention is intended for use in environments found in industry, such as those requiring attenuation of low frequency noise as well as noise covering the speech band (300 - 3300 Hz).

Passive headsets have been employed and work extremely well at higher frequencies above 1000 Hz but don't perform at all in the lower frequency ranges. In these lower ranges active cancellation has been shown to work extremely well and the provision of high levels of protection are afforded users in the 50 to 1000 Hz range.

The instant headset provides protection at both the high frequency range (above 1000 Hz) as well as the low range (50 to 1000 Hz).

In the past, attempts to combine the two protections, i.e., high and low frequency attenuation, has resulted in not only the noise being attenuated but also the speech that the wearer needs to hear. Some systems met only limited success with fixed or "near-stationary" noise but not with the other noise of either (a) varying spectral characteristics or (b) brief duration noises with "spikes". Examples of such a system is found in U.S. Patent No. 4,025,721, to Graupe et al and U.S. Patent 4,185,168 to Graupe et al.

The instant invention solves the problem now existant, that of <u>total</u> attenuation of the noise and speech, by providing a solution of an active/passive headset that can employ any of several selective algorithms such as that disclosed in U.S. Patent No.

5,091,953 to Tretter, hereby incorporated by reference herein. Alternatively, it can employ the algorithm disclosed in U.S. Patent No. 5,105,377 to Ziegler which is also incorporated herein by reference. In addition, it can employ other algorithms such as that disclosed in the application of Ziegler in U.S. Patent Application No. 07/421 759 which is hereby incorporated by reference herein.

The technique of actively controlling noise is disclosed in various patents to Barry Chaplin especially U.S. Patent 4,654,871 which is hereby incorporated by reference herein.

Accordingly it is an object of this invention to provide an active/passive headset that will attenuate a wide range of environmental noise while allowing speech to be heard.

It is another object of this invention to provide a headset that blocks unwanted noise but allows speech to pass through a speech filter or "in-wire" canceller and be combined with the output of an active noise controller.

Still another object of this invention is to combine passive quieting with active noise cancellation in a headset and allow for passage of speech with simultaneous quieting of environmental noise.

These and other objects of the invention will become readily apparent when reference is had to the accompanying drawings in which

Fig. 1 shows a typical active/passive headset system incorporating the instant invention.

Fig. 2 shows a typical active/passive headset system using a digital virtual earth controller.

Fig. 3 is a more detailed block diagram of the digital virtual earth and adaptive feed forward systems shown in Figures 1 and 2.

Fig. 4 is a detailed description of an adaptive speech filter system.

Fig. 5 shows an active/passive closed back headset that combines the cancellation system with the speech filter system.

In Fig. 1 there is shown an active/passive closed back headset system 10. It consists of a typical passive headset 11, loudspeakers 12 that drive the anti-noise and residual microphones 13 to sense any remaining noise near the ear and reference microphones 14 to send advanced information for feed forward approaches and a system controller 20 which synthesizes the anti-noise signal.

The headset shown has closed backs 21, 22 for passive attenuation. Without the speakers, microphones and system controller, this headset would be a typical passive hearing protector.

The system is designed to use various algorithms such as that of Ziegler in U.S. Patent 5,105,377 or an adaptive feed forward approach by using the reference sensor as an input to the cancellation algorithm. Both these algorithms use a reference signal as inputs. The digital virtual earth (DVE) algorithm develops a reference signal by subtracting an equalized version of its own anti-noise signal from the residual signal. The adaptive feed forward uses the reference microphone as its input and is very effective on complicated noise environments that are broadband and random in character. The Least Means Square (LMS) adapter 24 shown in Fig 1 are Filtered - X versions which have inherent compensation for the effects of the feedback delays around the loop. Box "C" at 25 is the impulse response of the entire cancellation system.

Feedback compensator 26 and cancellation filter 27 complete the component portions of the controller.

DVE is highly effective to use in simple noise environments having only a few harmonics even where the noise varies tremendously.

Actuators 12 of the headset are large enough to be capable of producing anti-noise at the same level as the noise to be canceled. They have little or no distortion and have a minimum of input-to-output delay as any delay in the feedback loop slows down the system adaptation rate.

Residual sensors 13 are typically small electret microphones mounted on the speaker frame near the ear. It must faithfully reproduce the sound that remains at the

ear after cancellation so that the controller can make further adjustments to the anti-noise signal.

Reference microphones 14 are small electret microphones attached to the outside of the headset at a distance from the ear canal. This reference microphone is used to provide advanced information about the noise. The higher the frequency of the noise the more advanced information is needed to effectively cancel the noise, however, the passive muff limits the need for cancellation above 1000 Hz.

Figure 2 shows a basic block diagram of a digital virtual earth system, phenomena (such as noise) is detected by a residual sensor which sends out a sensor signal. The sensor signal is affected by filters and other factors shown in Figure 3 before reaching the processor. The DVE processor produces an output cancellation signal y. This output signal is affected by filters, and other factors. The output from the actuator combines with the original phenomena and the residual is detected at a location by the residual sensor of Figure 2.

The processor output for a specific sample k are given by

$$\underline{x}_k = \underline{r}_k - \underline{y}_{k-1} . \underline{C}_k$$

$$Y_k = -\underline{A}_k r_k$$

where

y_k is cancellation output value

$\underline{y}_{k-1}$ is a vector of previous cancellation output values

$\underline{C}_k$ is a vector of filter coefficients approximating S x E where S is the impulse response of the system after the processor and E is the impulse response of the system prior to the processor where a x b is the convolution of a and b.

$\underline{x}_k$ is a vector of estimates of the noise

$\underline{r}_k$ is a vector of most recent values of the residual signal

$\underline{A}_k$ is a vector of cancellation filter coefficients and a.b is the product of the two vectors.

When an LMS algorithm is used the cancellation filter coefficients are adapted as

$$\underline{g}_k = \underline{C}_k \underline{x}_k$$

$$A_{k+1} = A_k + \alpha_k g_k$$

where

$\underline{g}_k$ is the result of convolving $\underline{C}_k$ and $\underline{x}_k$

A_{k+1} is a vector of the most recent values of g_k and α is set for convergence.

Figure 3 shows in more detail the active cancellation system in either a DVE or AFF configuration. The active cancellation system has an input that receives residual noise signals (and reference signals in an AFF configuration) from a residual sensor (and reference sensor in the AFF configuration) and provides an output signal to an actuator. The actuator produces canceling phenomena from an electronic waveform (the output signal) into the area to be controlled (the ear canal). The canceling phenomena combines with the original phenomena so that the two sum algebraically.

Figure 4 shows in detail the adaptive speech filtering system. The system has as an input a reference signal that is composed of noise and speech from a reference sensor. From this input the adaptive speech filter separates the speech from noise and provides an output signal composed of speech only to an actuator.

Fig. 5 shows the instant invention with an adaptive speech filter 30 and the impulse "C" using a DVE cancellation approach. Reference microphones 14 not only pick up damaging noise but also pick up speech and warning signals. The embodiment of Fig. 2 is designed to provide protection against damaging noise which typically interferes with the speech band. It also attenuates noise in the audible range of the human ear effectively attenuating any warning siren. Using reference microphones 14 as inputs to a speech filter 30 that is then mixed with the anti-noise signal to be sent to the speaker is a means of allowing speech to pass through to the worker. If warning signals are outside the speech band and outside the frequency range of the active controller 20, these can be passed to the worker as well. Fig. 5 is a diagrammatic view

of the integration of active/passive hearing protection with speech (and warning signal) filtering so that local communication can be maintained for the safety of the worker.

The speech filter 30 separates speech from noise and output speech to the speaker with a minimal effect on intelligibility. The specific in-wire approach to use is dependent upon the noise but the DVE algorithm could be used. Various approaches have specific advantages in selectively filtering speech. Even the approach used in the Graupe patents can be used and since the output of filter 30 cannot interfere with the working of controller 20, it needs to be filtered from the control system as shown in Fig. 5 and, using the equations developed for the DVE algorithm, the output signal u_k is a specific sample k given by

$$w_k = \underline{T}_k \underline{v}_k$$

$$u_k = y_k + w_k$$

where

u_k is the cancellation and speech output value.

w_k is the filtered speech value.

$\underline{T}_k$ is a vector of filter coefficients determined by the adaptive speech filtering algorithm.

y_k is the cancellation output value as before.

$\underline{v}_k$ is the reference signal (speech and noise) input vector.

The residual sensor used to control the cancellation system also picks up the speech that is passed via the actuator. Therefore, the residual signal r_k must remove the speech signal w_k before input to the cancellation system so as not to interfere with the operation of the controller. It does so by modifying the residual signal r_k by

$$z_k = r_k - \underline{C}_k w_k$$

where

z_k is the filtered residual input to the controller.

$\underline{r}_k$ is the original residual input.

$\underline{C}_k$ is a vector of filter coefficients that is the overall impulse response of the cancellation system, as before.

w_k is the filtered output speech signal that is output to the actuator with the cancellation signal y_k.

The need to remove $\underline{C}_k w_k$ from the residual signal may not be required for all applications, for example, a siren canceling headset system will not try and cancel speech while the siren is in operation.

While the preferred embodiment of the invention is described it will be obvious to those of ordinary skill in the art that many changes and modifications can be made without departing from the scope of the appended claims.

CLAIMS

What is claimed is:

1. An active/passive noise canceling headset for attenuating high and low frequency noise, said headset comprising

a headset means having closed back earpiece means adapted to prevent external noise from entering the ear through passive attenuation,

speaker means associated with said earpiece means and adapted to produce anti-noise of the same level as the noise to be canceled,

reference sensing means also associated with said earpiece means and adapted to produce advanced information about the noise to be canceled, and

controller means adapted to secure information from said reference sensing means and cause emission of anti-noise by said speaker means.

2. A headset as in claim 1 and including residual sensing means also associated with said earpiece means and adapted to produce a signal characterizing the sound remaining at the ear.

3. A headset as in claim 2 and including a feedback loop means connecting said residual sensing means to said controller means and whereby said controller means makes adjustments to the anti-noise in response to the sound sensed by said residual sensing means.

4. A headset as in claim 3 wherein said reference sensing means are electret microphones or other type of microphones.

5. A headset as in claim 3 wherein said residual sensing means are electret microphones.

6. A headset as in claim 1 and wherein said controller means contains an adaptive speech filter means adapted to attenuate disturbing noise in the audible range of the human ear yet leave the speech to the headset wearer intelligible.

7. The method of controlling the amount of noise interfering with speech, said method comprising sensing signal containing speech and noise and converting it to electronic signals, using "in-wire" control approach producing "anti-noise" to cancel said noise component of the signal.

8. The method of claim 7 and including the step of passively quieting the higher frequency portions of the noise to be canceled simultaneous with the production of anti-noise.

9. The method of claim 8 and including the further step of sensing the noise remaining adjacent the ear after cancellation by said anti-noise and changing the anti-noise production in response to the amount of remaining noise.

10. The method of claim 9 and including the step of mixing the signals from sensing the noise containing noise and speech with the signal producing the anti-noise thus allowing speech to pass to the ear.

11. The method of filtering speech from noise said method comprising sensing signal containing speech and noise and converting it to electronic signals, adaptively filtering the signal to separate the noise component from the speech component and producing a signal that contains only the speech component.

12. The method of manual selection of either the method of claim 7 or the method of claim 8 or both based upon characteristics of the noise environment wherein

said manual selection is a switch that allows the user to select a preferred approach for allowing speech to be most intelligible.

13. The method of automatic selection of either the method of claim 7 or the method of claim 8 or both based upon the spectral and temporal characteristics of the noise environment wherein said automatic selection is achieved using a detection algorithm that continuously monitors the spectral and temporal characteristics of the noisy signal.

14. The method as in claim 10 wherein said selection of adaptive speech filtering occurs when the noise is near-stationary and of limited bandwidth.

15. The method as in claim 10 wherein said selection of in-wire control occurs when the noise is varying and tonal or broadband.

16. The method of manual selection of claim 12 wherein a setting to automatic selection can be made.

17. The method of manual selection of claim 12 wherein a setting to an off position can be made.

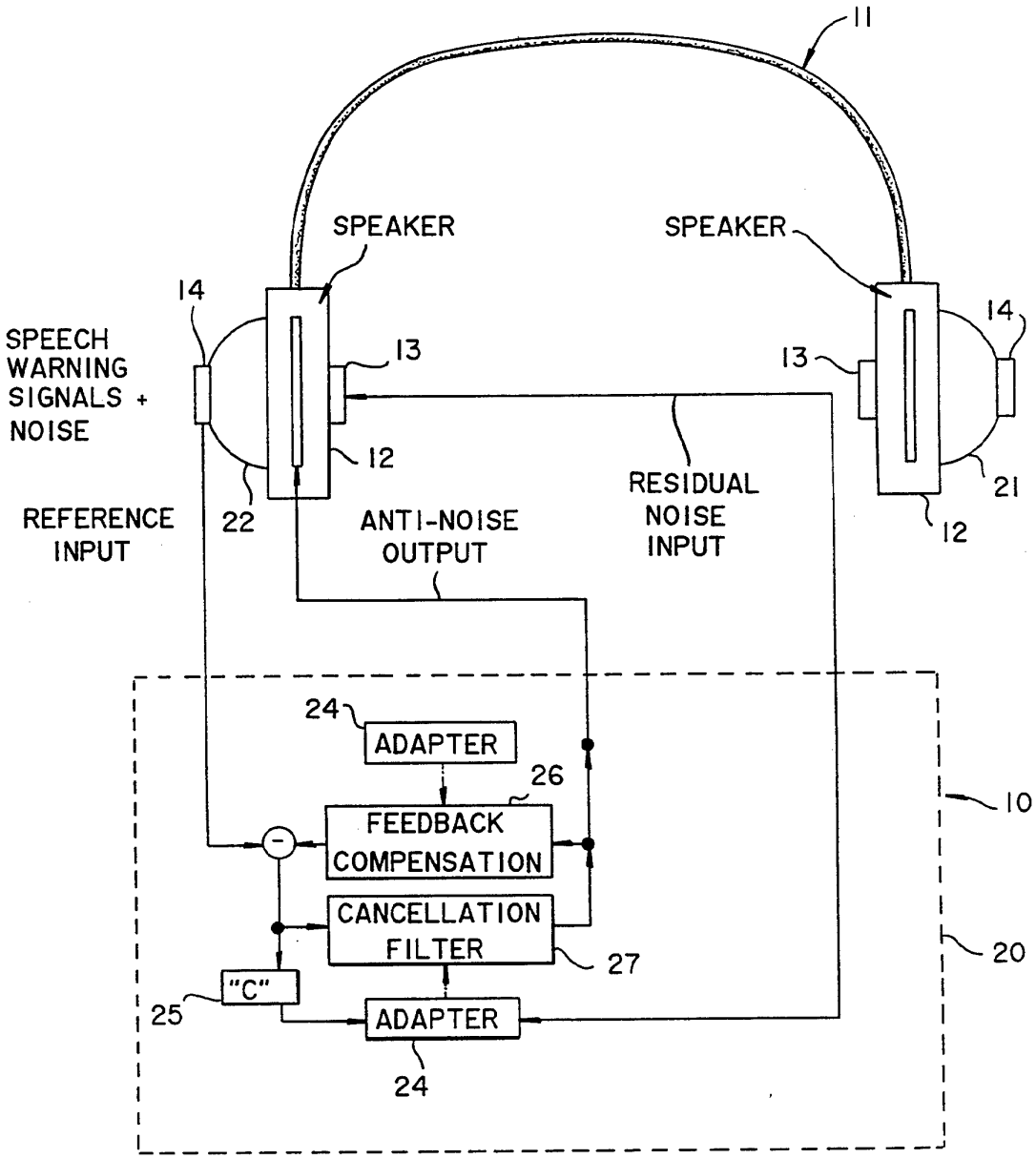

11
SPEAKER
SPEAKER
14
14
SPEECH
WARNING
SIGNALS +
NOISE
13
13
12
21
REFERENCE
INPUT
22
ANTI-NOISE
OUTPUT
RESIDUAL
NOISE
INPUT
12
24
ADAPTER
26
10
FEEDBACK
COMPENSATION
CANCELLATION
FILTER
27
20
"C"
25
ADAPTER
24

FIG.1

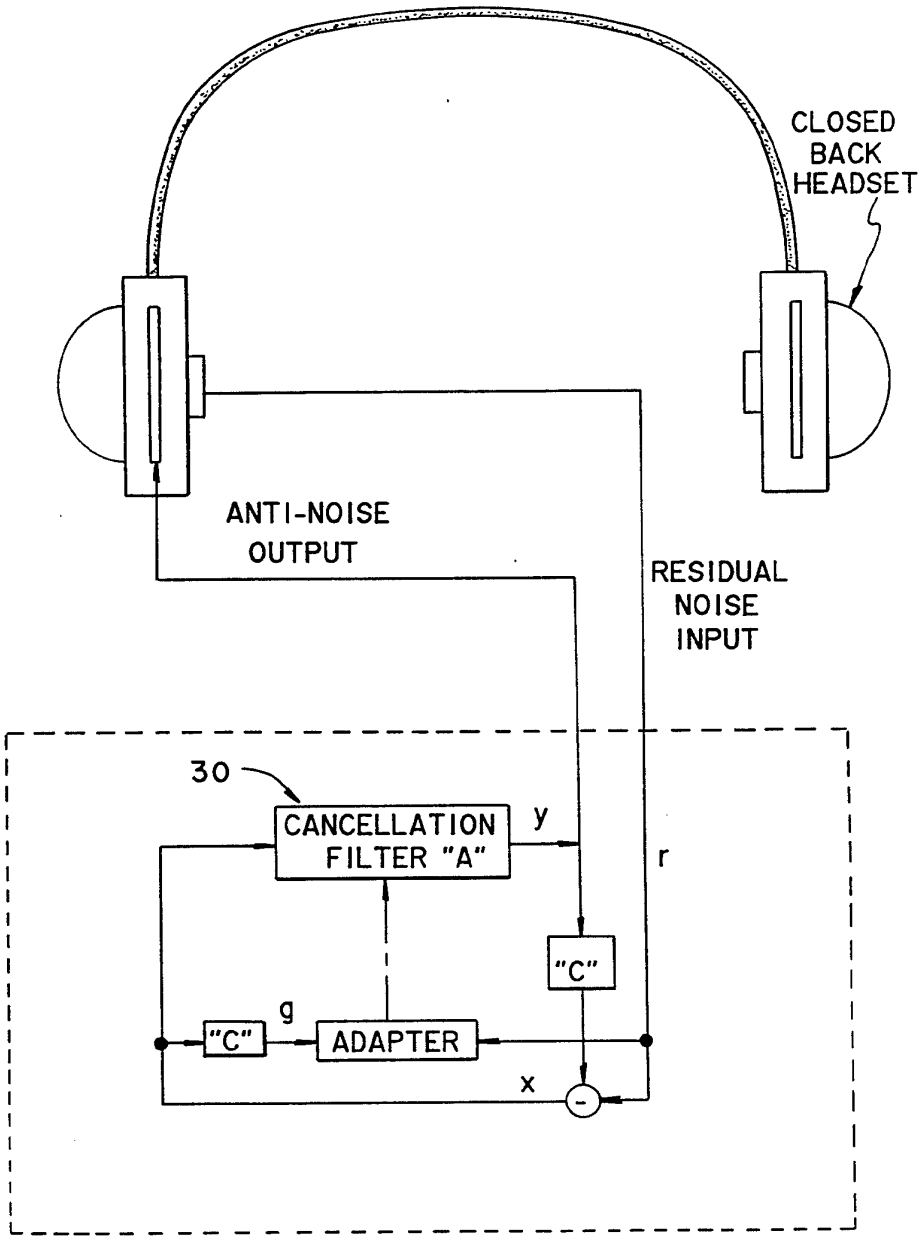

CLOSED
BACK
HEADSET
ANTI-NOISE
OUTPUT
RESIDUAL
NOISE
INPUT
30
CANCELLATION
FILTER "A"
y
r
"C"
g
"C"
ADAPTER
x

FIG.2

SUBSTITUTE SHEET

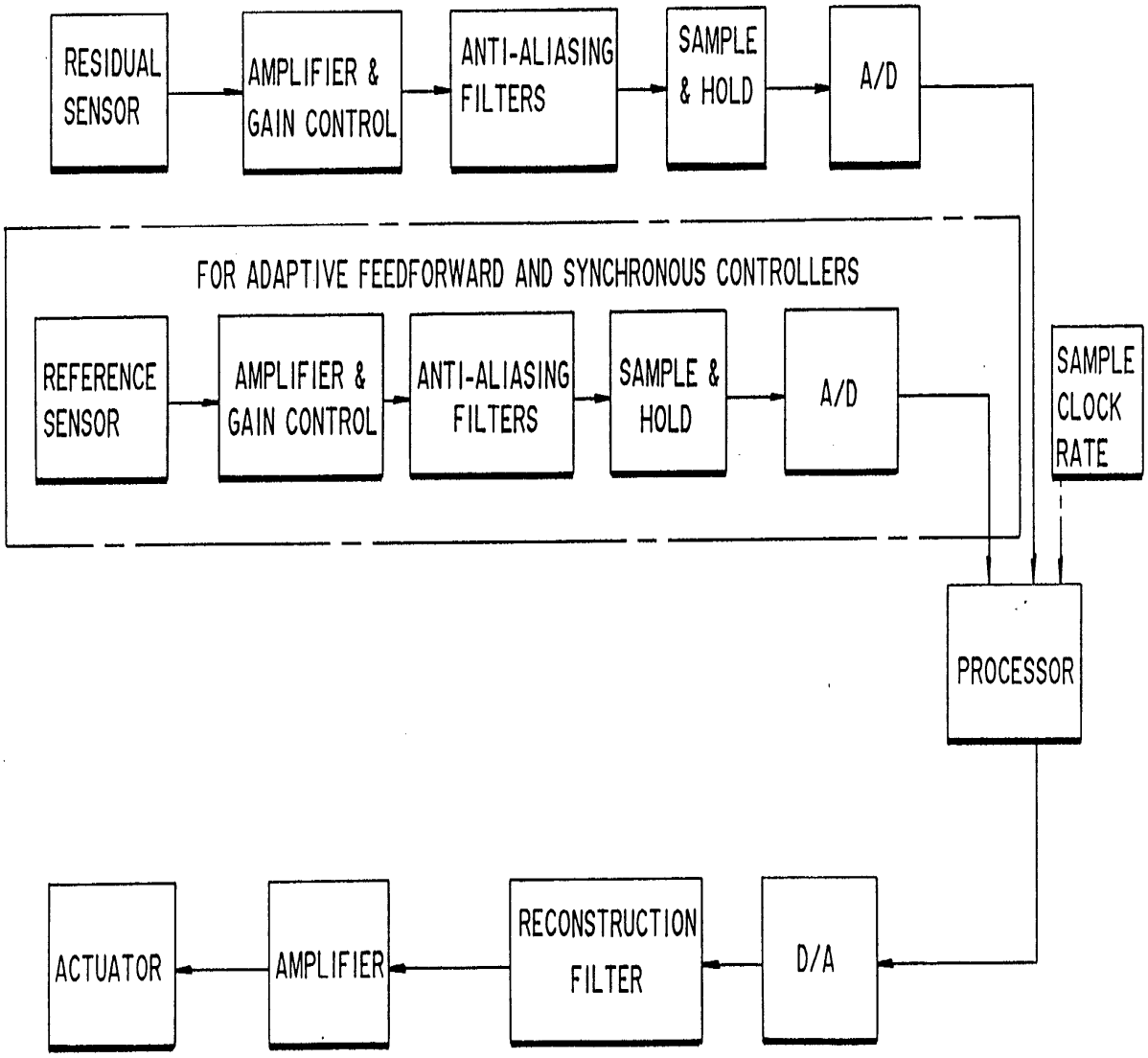


FIG.3

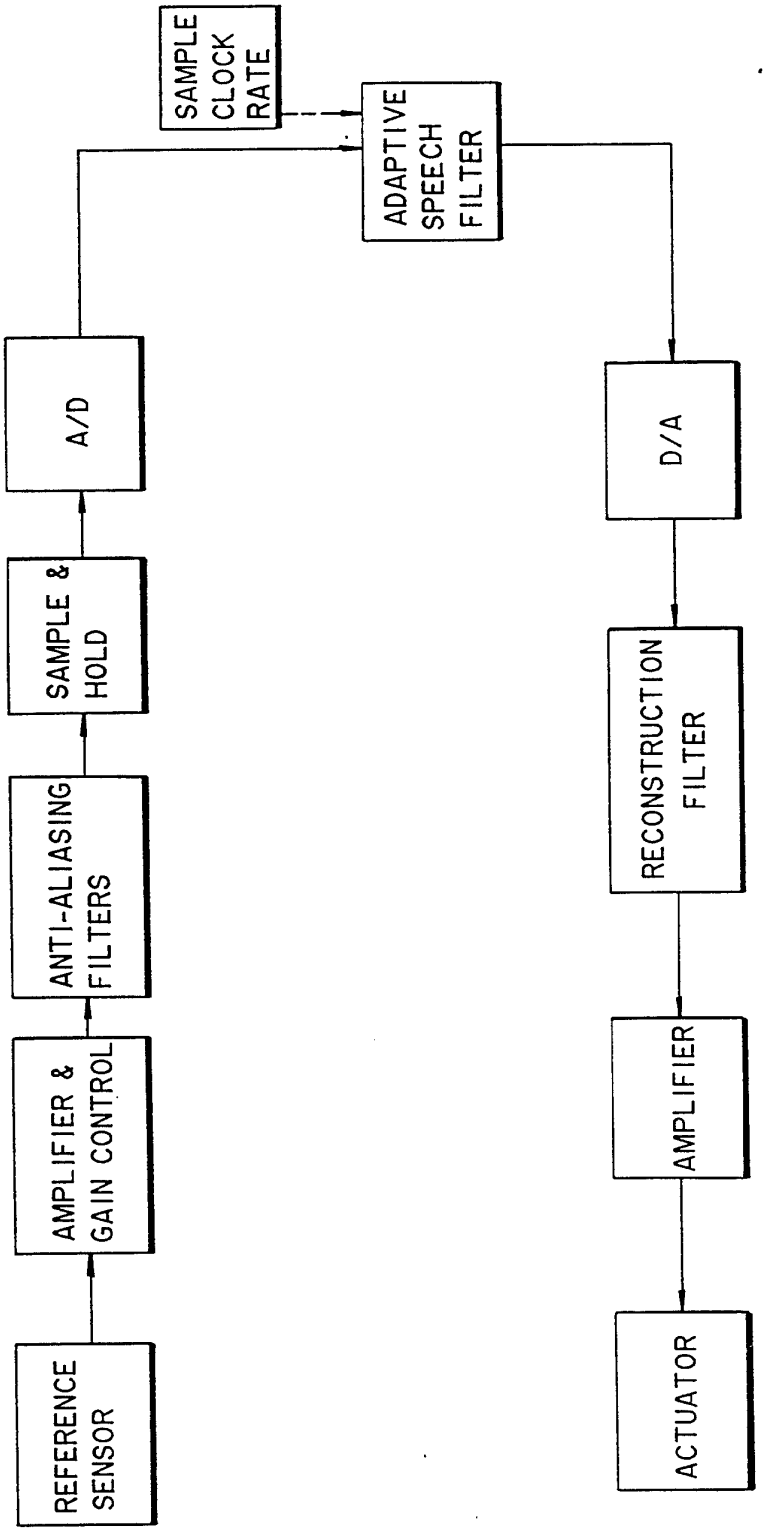

SAMPLE CLOCK RATE
ADAPTIVE SPEECH FILTER
A/D
SAMPLE & HOLD
ANTI-ALIASING FILTERS
AMPLIFIER & GAIN CONTROL
REFERENCE SENSOR
D/A
RECONSTRUCTION FILTER
AMPLIFIER
ACTUATOR

FIG.4

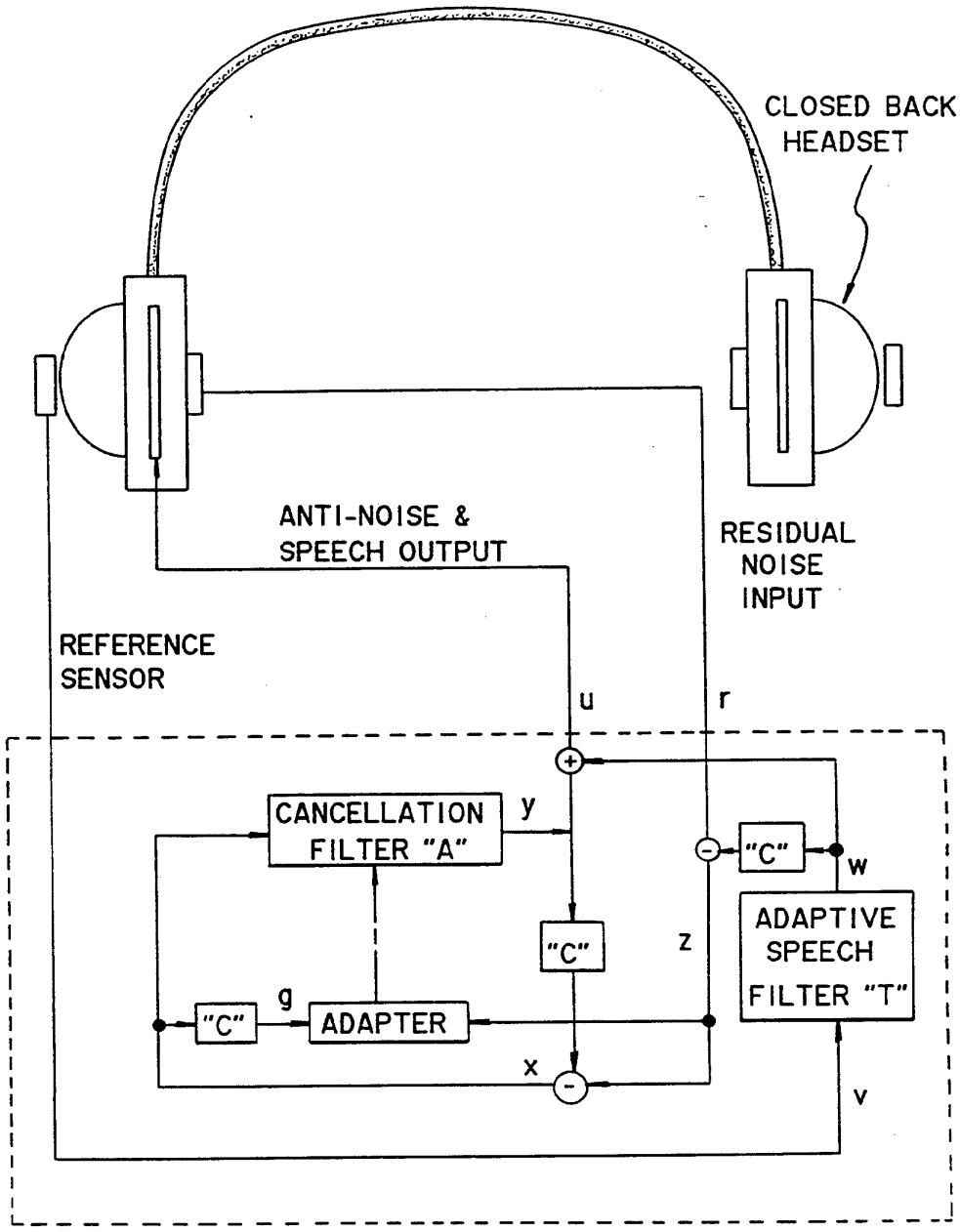


FIG.5

SUBSTITUTE SHEET

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/04568

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :H03B 29/00
US CL :381/71

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. : 381/72, 74

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 Y	US, A, 3,952,158 (Kyle et al.) 20 April 1976 See Figs. 1-5	1 1-6
X Y	US, A, 4,061,079 (Freifeld) 06 December 1977 See Figs. 2-5	1 2-6
Y,E	US, A, 5,126,681 (Ziegler, Jr. et al.) 30 June 1992 See Abstract and Figs. 1-3	7-17
Y	US, A, 5,105,377 (Ziegler, Jr.) 14 April 1992 See Abstract and Figs. 1-5	1-17
Y	US, A, 4,953,217 (Twiney et al.) 28 August 1990 See Abstract Fig. 1	2-6
Y	US, A, 5,091,953 (Tretter) 25 February 1992 See Abstract and Figs. 1 and 5	7-17
Y	US, A, 4,064,362 (Williams) 20 December 1977 See Abstract and Fig. 1	1
A	US, A, 4,677,678 (McCutchen) 30 June 1987 See Abstract and Figs. 1-7	1-6
A	US, A, 4,455,675 (Bose et al.) 19 June 1984 See Abstract and Figs. 1 and 5-6	1-6

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"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
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"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
"E"	earlier document published on or after the international filing date	"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
"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)	"&"	document member of the same patent family
"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		

Date of the actual completion of the international search 06 AUGUST 1992	Date of mailing of the international search report 09 NOV 1992
Name and mailing address of the ISA/ Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer EDWARD LEFKOWITZ Telephone No. (703) 305-4816

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/04568

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,046,103 (Warnaka et al.) 03 September 1991 See Abstract and Figs. 1-2	1-6

Form PCT/ISA/210 (continuation of second sheet)(July 1992)*

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 935 236 A1

(12) DEMANDE DE BREVET EUROPEEN

(43) Date de publication:
11.08.1999 Bulletin 1999/32

(51) Int Cl.[6]: **G10K 11/178**

(21) Numéro de dépôt: **99400221.0**

(22) Date de dépôt: **02.02.1999**

(84) Etats contractants désignés:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Etats d'extension désignés:
AL LT LV MK RO SI

(30) Priorité: **06.02.1998 FR 9801411**

(71) Demandeur: **SAGEM SA**
75116 Paris (FR)

(72) Inventeur: **Housni, Jamal**
92400 Courbevoie (FR)

(74) Mandataire: **Fruchard, Guy et al**
CABINET BOETTCHER,
22, rue du Général Foy
75008 Paris (FR)

(54) Dispositif anti-bruit à détection de seuil

(57) Dispositif anti-bruit comportant au moins une oreillette (13) ayant une cavité (2) délimitée par des moyens d'atténuation passive (3), un circuit d'atténuation active (4) ayant un domaine spatial d'efficacité et comportant au moins une unité de traitement (6) électronique reliée à des organes d'atténuation active (5, 7) comprenant au moins un microphone (5) captant des ondes sonores dans la cavité et un haut-parleur (7) disposé dans les cavités pour émettre des ondes sonores d'atténuation active élaborées par l'unité de traitement (6), au moins un organe interrupteur (8) sélectivement actionnable entre un état de service du circuit d'atténuation active et un état de repos du circuit d'atténuation active, un microphone de détection de seuil (10) captant des bruits extérieurs au domaine spatial d'efficacité du circuit d'atténuation active, des moyens de filtrage (3) associés au microphone de détection de seuil pour éliminer des bruits correspondant aux bruits absorbés par les moyens d'atténuation passive, et un détecteur (9) de dépassement d'un seuil de bruit pour commander l'organe interrupteur (8) dans l'état de service ou de repos du circuit d'atténuation active en fonction d'un niveau de puissance des bruits détectés en aval des moyens de filtrage.

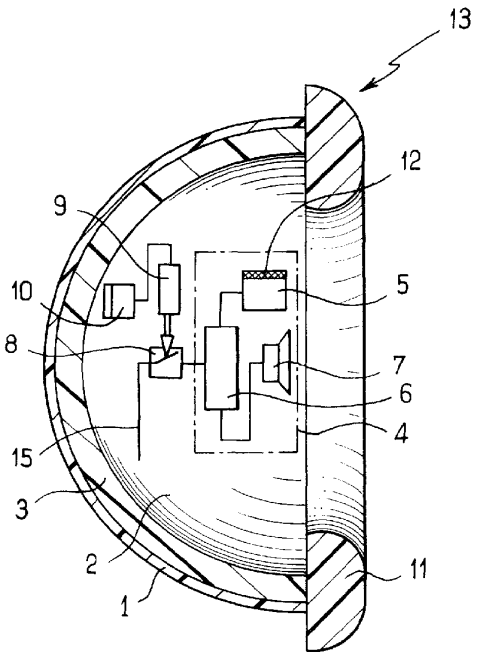


FIG. 2

EP 0 935 236 A1

Description

[0001] L'invention concerne un dispositif anti-bruit destiné à limiter la gêne qu'occasionnent les bruits à un utilisateur évoluant dans un milieu bruyant, que cet utilisateur soit ou non en communication avec une tierce personne.

[0002] On connaît des dispositifs anti-bruits réalisés sous la forme de casques comportant deux oreillettes reliées l'une à l'autre par un organe support permettant de maintenir chaque oreillette sur une oreille de l'utilisateur. Chaque oreillette comporte une paroi définissant une cavité. La paroi est tapissée, du côté interne de la cavité, d'un matériau d'isolation phonique. La paroi ainsi pourvue réalise une atténuation passive en absorbant les ondes sonores de moyennes et hautes fréquences mais demeure perméable aux ondes sonores de basses fréquences. Pour cette raison, le casque est de préférence équipé d'un circuit d'atténuation active de bruit éliminant les ondes sonores de basses fréquences. Le circuit d'atténuation active engendre, à partir d'une onde sonore incidente captée dans chaque cavité, une onde sonore d'atténuation, d'amplitude identique mais en opposition de phase avec l'onde sonore incidente. L'onde se propageant dans le conduit auditif de l'utilisateur est l'onde résultant de la somme de l'onde incidente et de l'onde engendrée par le circuit d'atténuation active.

[0003] Le circuit d'atténuation active comporte une unité de traitement électronique reliée dans chaque cavité à des organes d'atténuation active comprenant un microphone captant une onde sonore dans la cavité et un haut-parleur émettant l'onde sonore d'atténuation élaborée par l'unité de traitement. Un tel circuit doit être relié à une source d'alimentation électrique. De nombreux circuits d'atténuation active sont alimentés de façon continue de sorte qu'il en résulte une consommation électrique importante particulièrement préjudiciable à la durée d'utilisation des casques autonomes dans lesquels la source d'alimentation est constituée par des accumulateurs ne fournissant qu'une quantité limitée d'électricité.

[0004] Pour remédier à cet inconvénient, on a pensé à pourvoir le circuit d'atténuation active d'un organe interrupteur à commande manuelle, à charge pour l'utilisateur d'actionner cet organe afin de mettre en fonctionnement le circuit d'atténuation active lorsqu'il évolue dans un environnement bruyant et d'en arrêter le fonctionnement lorsque le bruit environnant diminue. Toutefois, un utilisateur soumis de façon régulière au bruit subit une accoutumance. Malgré la présence de ce bruit, l'utilisateur retarde alors la mise en fonctionnement du circuit d'atténuation active de sorte qu'il tolère le bruit de manière croissante même si la puissance de celuici est suffisante pour entraîner une fatigue importante.

[0005] Un but de l'invention est d'obvier aux inconvénients précités.

[0006] On propose selon l'invention, un dispositif anti-bruit comportant au moins une oreillette ayant une cavité délimitée par des moyens d'atténuation passive, un circuit d'atténuation active ayant un domaine spatial d'efficacité et comportant au moins une unité de traitement électronique reliée à des organes d'atténuation active comprenant au moins un microphone captant des ondes sonores dans la cavité et un haut-parleur disposé dans la cavité pour émettre des ondes sonores d'atténuation active élaborées par l'unité de traitement, au moins un organe interrupteur sélectivement actionnable entre un état de service du circuit d'atténuation active et un état de repos du circuit d'atténuation active, un microphone de détection de seuil captant des bruits extérieurs au domaine spatial d'efficacité du circuit d'atténuation active, des moyens de filtrage associés au microphone de détection de seuil pour éliminer des bruits correspondant aux bruits absorbés par les moyens d'atténuation passive, et un détecteur de dépassement d'un seuil de bruit pour commander l'organe interrupteur dans l'état de service ou de repos du circuit d'atténuation active en fonction d'un niveau de puissance des bruits détectés en aval des moyens de filtrage.

[0007] Ainsi, le déclenchement du circuit d'atténuation active est fonction du niveau de puissance des bruits filtrés, c'est-à-dire des bruits qui correspondent exactement à ceux que doit éliminer le circuit d'atténuation active. Une protection optimale est ainsi assurée pour une consommation réduite.

[0008] Selon un mode de réalisation préféré, le microphone de détection de seuil est porté par une paroi de la cavité de l'oreillette pour capter des bruits extérieurs à la cavité, les moyens de filtrage comprenant un organe de filtrage monté entre le microphone de détection de seuil et le détecteur de dépassement de seuil.

[0009] Par ailleurs, dans les dispositifs anti-bruit comportant deux oreillettes pour former un casque, les moyens d'atténuation passive sont couramment constitués d'un matériau d'isolation phonique fixé à l'intérieur des oreillettes. A moins de retirer son casque, l'utilisateur ne peut donc entendre les paroles des personnes qui l'entourent même lorsque le bruit environnant est faible. Il en résulte un isolement permanent de l'utilisateur, même lorsque cela n'est pas nécessaire.

[0010] Selon une variante du mode de réalisation préféré de l'invention, le dispositif comprend un microphone d'écoute porté par une paroi de la cavité de la seconde oreillette pour capter les bruits extérieurs à la cavité, des moyens de commutation reliés au microphone de détection de seuil et au microphone d'écoute et sélectivement actionnables par un détecteur supplémentaire de dépassement de seuil en fonction d'un niveau de puissance des bruits captés par le microphone de détection de seuil entre un premier état dans lequel le microphone de détection de seuil et le microphone d'écoute sont connectés à des haut-parleurs respectifs et un second état dans lequel le microphone de détection de seuil est connecté à l'organe de filtrage et le microphone d'écoute est coupé. Ainsi, l'utilisateur peut écouter les bruits extérieurs lorsque ces bruits ont un niveau de

Exhibit 1009
Page 1574 of 2198

puissance inférieur au seuil de bruit du détecteur supplémentaire de dépassement de seuil.

[0011] Selon un autre aspect de l'invention en liaison avec un dispositif anti-bruit selon la variante précitée et comportant un circuit de communication, lesdits haut-parleurs respectifs sont associés au circuit de communication et le dispositif comporte des moyens pour interrompre toute liaison entre les microphones et lesdits haut-parleurs respectifs lors de la réception d'une communication. Ainsi, lors de la réception d'un signal de communication, l'utilisateur n'est pas gêné par les bruits extérieurs quel que soit le niveau de puissance de ceux-ci.

[0012] D'autres caractéristiques et avantages de l'invention ressortiront à la lecture de la description qui suit de modes de réalisation particuliers non limitatifs de l'invention en relation avec les figures annexées parmi lesquelles :

- la figure 1 est une vue en perspective d'un dispositif anti-bruit selon l'invention ;
- la figure 2 est une vue schématique, en coupe par un plan vertical II de la figure 1, d'une oreillette du dispositif anti-bruit selon un premier mode de réalisation de l'invention ;
- la figure 3 est une vue schématique d'un deuxième mode de réalisation ;
- la figure 4 est une vue schématique d'une variante du deuxième mode de réalisation.

[0013] En référence aux figures, le dispositif anti-bruit selon l'invention est ici réalisé sous forme d'un casque et comporte deux oreillettes 13 reliées l'une à l'autre par un arceau 14 constituant un organe support des oreillettes 13.

[0014] Une oreillette 13 comprend une paroi 1, ici représentée sous la forme d'une calotte sphérique, définissant une cavité 2. Un matériau d'isolation phonique 3 constituant un moyen d'atténuation passive de bruits ayant une fréquence comprise dans une plage déterminée de fréquences est disposé, de façon connue en soi, dans la cavité 2. Un bourrelet 11 est fixé le long du bord de la paroi 1.

[0015] Le casque anti-bruit comprend, également de façon connue en soi, un circuit d'atténuation active de bruit généralement désigné en 4 et relié à une source d'alimentai tion en énergie électrique non représentée.

[0016] Le circuit d'atténuation active 4 comprend, pour chaque oreillette 13, une unité de traitement 6 électronique reliée à un microphone 5 et à un haut-parleur 7. L'unité de traitement 6 est en outre reliée à la source d'alimentation par une ligne d'alimentation 15.

[0017] En fonctionnement classique, le microphone 5 capte une onde sonore dans la cavité et transmet un signal électrique incident représentatif de l'onde sonore à l'unité de traitement 6. L'unité de traitement 6 élabore alors un signal électrique d'atténuation représentatif d'une onde sonore d'amplitude identique mais en opposition de phase avec l'onde sonore captée par le microphone 5. Le signal électrique d'atténuation est transmis au haut-parleur 7 qui émet l'onde sonore d'atténuation correspondante dans une portion de la cavité constituant le domaine d'efficacité du circuit d'atténuation active.

[0018] Selon l'invention, un organe interrupteur 8 est disposé dans chaque oreillette 13 sur la ligne d'alimentation 15 entre la source d'alimentation et l'unité de traitement 6. L'organe interrupteur 8 est soit dans un état passant dans lequel l'organe interrupteur permet l'alimentation de l'unité de traitement 6, soit dans un état bloquant (représenté aux figures 2, 3 et 4) dans lequel l'organe interrupteur 8 interrompt l'alimentation de l'unité de traitement 6.

[0019] Selon le premier mode de réalisation de l'invention et en référence à la figure 2, un détecteur de dépassement de seuil 9, disposé dans une des deux oreillettes 13, est relié aux organes interrupteurs 8 pour les commander. Le détecteur de dépassement de seuil 9 est relié à un microphone de détection de seuil 10 logé dans la cavité de cette oreillette pour capter les bruits dans la portion de cette cavité extérieure au domaine spatial d'efficacité du circuit d'atténuation active. Ces bruits ont été filtrés par le matériau d'isolation phonique et ont des fréquences comprises dans la plage de fréquences des bruits traités par le circuit d'atténuation active de sorte que le détecteur de dépassement de seuil 9 analyse des bruits qui sont en majeure partie susceptibles d'être traités par le circuit d'atténuation active.

[0020] En fonctionnement, le bruit capté dans la cavité 2 par le microphone de détection de seuil 10 est transmis au détecteur de dépassement de seuil 9.

[0021] Lorsque ce bruit est d'un niveau de puissance supérieure à un seuil prédéterminé du détecteur de dépassement de seuil 9, celui-ci commande la mise à l'état passant des organes interrupteur 8, mettant ainsi en service le circuit d'atténuation active.

[0022] Lorsque le bruit filtré est d'un niveau de puissance inférieure au seuil prédéterminé, le détecteur de dépassement de seuil 9 commande la mise à l'état bloquant de chaque organe interrupteur 8. L'unité de traitement 6 n'étant plus alimentée, n'émet alors aucun signal électrique d'atténuation. Le circuit d'atténuation active est alors dans un état de repos et l'unité de traitement ne consomme aucune énergie.

[0023] Pour éviter des interférences entre le microphone de détection de seuil et le circuit d'atténuation active il est essentiel que le microphone de détection de seuil soit en dehors du domaine spatial d'efficacité du circuit d'atténuation active. En pratique ce mode de réalisation pourra être utilisé dans le cas où le circuit d'atténuation active est monté dans un bouchon d'oreille.

[0024] Afin d'augmenter la capacité du microphone 5 à capter des ondes sonores de forte puissance, une résistance acoustique 12, constituée par exemple d'un tissu à mailles serrées, est de préférence disposée sur le microphone 5. Par exemple, pour un microphone 5 sen-

Exhibit 1009
Page 1575 of 2198

sible à des bruits d'un niveau de puissance allant de 20dB à 120dB, la disposition d'une résistance acoustique de 20dB permet de porter la plage de fonctionnement de 40dB à 140dB. Le seuil de mise en service du circuit d'atténuation active étant par exemple de l'ordre de 60dB, on obtient ainsi une plage d'atténuation efficace de 60dB à 140dB soit 80dB, au lieu d'une plage de 60dB à 120dB soit 60dB seulement avec le même microphone sans résistance acoustique. La protection apportée par le casque est ainsi étendue vers des bruits de plus forte puissance sans saturation des microphones du circuit d'atténuation active.

[0025] Les éléments identiques ou analogues a ceux précédemment décrits porteront dans la suite de la description la même référence numérique.

[0026] Selon le deuxième mode de réalisation, représenté à la figure 3, le casque anti-bruit comprend comme précédemment un détecteur de dépassement de seuil 9 relié aux organes interrupteurs 8 pour les commander. Le microphone de détection de seuil 10 est cette fois porté par la paroi d'une oreillette pour capter l'ensemble des bruits extérieurs à celle-ci et le détecteur de dépassement de seuil 9 est relié à un organe de filtrage 19 adapté à assurer une élimination des bruits dans la gamme de fréquences des moyens d'atténuation passive.

[0027] En outre l'organe de filtrage 19 et le haut-parleur 7 de l'oreillette supportant le détecteur de dépassement de seuil 9 sont reliés au microphone de détection de seuil 10 par l'intermédiaire d'un organe de commutation 20.

[0028] Un microphone d'écoute 22 est fixé à la paroi de l'autre oreillette pour capter les bruits extérieurs à cette oreillette et est relié au haut-parleur 7 correspondant par l'intermédiaire de l'organe de commutation 20.

[0029] Un détecteur supplémentaire de dépassement de seuil 21 est relié au microphone de détection de seuil 10 et à l'organe de commutation 20 pour commander l'organe de commutation entre un premier état (représenté à la figure 2) dans lequel le microphone de détection de seuil 10 et le microphone d'écoute 22 sont connectés aux haut-parleurs 7 et un second état dans lequel le microphone de détection de seuil 10 est connecté à l'organe de filtrage 19 et le microphone d'écoute 22 est coupé.

[0030] En fonctionnement, le bruit capté à l'extérieur des cavités par le microphone de détection de seuil 10 est transmis au détecteur supplémentaire de dépassement de seuil 21.

[0031] Lorsque le bruit capté par le microphone de détection de seuil 10 est d'un niveau de puissance inférieur à un seuil prédéterminé du détecteur supplémentaire de dépassement de seuil 21, par exemple 80 dB, le détecteur supplémentaire de dépassement de seuil 21 commande la mise de l'organe de commutation 20 dans le premier état. Les bruits extérieurs captés par les microphones 10 et 22 sont alors transmis à l'utilisateur par les haut-parleurs 7.

[0032] Lorsque ce bruit est d'un niveau de puissance supérieur au seuil prédéterminé du détecteur supplémentaire de dépassement de seuil 21, ce dernier commande la mise de l'organe de commutation 20 dans le second état. Les haut-parleurs 7 sont donc coupés de l'extérieur. Simultanément, le bruit capté par le microphone de détection de seuil 10 filtré par l'organe de filtrage 19, est transmis au détecteur de dépassement de seuil 9. Lorsque le bruit filtré a un niveau de puissance supérieur à un seuil prédéterminé du détecteur de dépassement de seuil 9, par exemple 120 dB, le détecteur de dépassement de seuil 9 commande la mise à l'état passant des organes interrupteurs 8, mettant ainsi en service le circuit d'atténuation active. Lorsque le bruit filtré a un niveau de puissance inférieur au seuil prédéterminé du détecteur de dépassement de seuil 9, le détecteur de dépassement de seuil 9 commande la mise à l'état bloquant des organes interrupteurs 8 de sorte que seule l'atténuation passive des bruits est assurée.

[0033] En référence à la figure 4, et selon une variante du deuxième mode de réalisation, le casque comporte un circuit de communication extérieure comprenant une antenne de réception associée à un élément de traitement 30 du signal de communication, et des haut-parleurs 31 disposés dans chaque cavité et reliés à l'élément de traitement 30 par des organes de basculement 32. Les organes de basculement 32 sont également reliés à l'organe de commutation 20.

[0034] Dans cette variante, on prévoit, lorsque cela est possible, de transmettre les bruits extérieurs aux haut-parleurs de communication 31. A cet effet, l'organe de commutation 20 est sélectivement actionnable entre un premier état dans lequel le microphone de détection de seuil 10 et le microphone d'écoute 22 sont connectés aux organes de basculement 32 et un second état dans lequel le microphone de détection de seuil 10 est relié à l'organe de filtrage 19 et le microphone d'écoute 22 est coupé.

[0035] Les organes de basculement 32 sont sélectivement actionnables entre un premier état dans lequel le signal de communication provenant de l'élément de traitement 30 est transmis aux haut-parleurs 31 et un second état dans lequel l'organe de commutation 20 est connecté aux haut-parleurs 31. Les organes de basculement 32 sont amenés dans leur premier état lorsqu'un signal de communication envoyé par l'élément de traitement 30 leur parvient et basculent dans le second état en l'absence d'un tel signal. Ainsi, en présence d'un signal de communication, les organes de basculement 32 passent dans le premier état pour transmettre prioritairement ce signal aux haut-parleurs 31, et ce, quel que soit le niveau de puissance des bruits extérieurs ou des bruits filtrés.

[0036] Bien entendu, l'invention n'est pas limitée aux modes de réalisation décrits et on peut y apporter des variantes de réalisation sans sortir du cadre de l'invention tel que défini par les revendications.

[0037] Bien que le support des oreillettes ait été illus-

Exhibit 1009
Page 1576 of 2198

tré sous forme d'un arceau auquel celles-ci sont reliées, le support pourrait être constitué par un casque ou une coiffe auquel sont associées des oreillettes formées par des moyens d'atténuation passive délimitant les cavités.

[0038] Bien que la mise en service ou l'arrêt des organes d'atténuation active ait été décrite en relation avec un interrupteur 8 disposé entre l'unité de traitement 6 et la source d'alimentation, on peut disposer cet interrupteur entre l'unité de traitement 6 et le haut-parleur 7 ou même entre l'unité de traitement 6 et le microphone 5. Dans ce dernier cas, l'unité de traitement ne reçoit aucun signal du microphone 5 lorsque l'organe interrupteur est dans son état bloquant. L'unité de traitement élabore alors un signal électrique d'atténuation correspondant à une onde sonore d'atténuation d'amplitude nulle.

[0039] La disposition décrite concernant les différents organes peut être modifiée. En particulier, on peut prévoir une unité de traitement commune aux deux oreillettes ou encore associer le microphone 10 et le détecteur 9 sur le support des oreillettes.

[0040] Dans le deuxième mode de réalisation, on pourrait prévoir que lorsque l'organe de commutation est dans le premier état, les microphones de détection de seuil et d'écoute soient reliés à des haut-parleurs distincts des haut-parleurs du circuit d'atténuation active.

[0041] Bien que la mise en oeuvre du second mode de réalisation ait été décrite avec un détecteur de dépassement de seuil 21 réglé à un seuil inférieur au détecteur de dépassement de seuil 9, on peut régler ces détecteurs à des niveaux quelconques l'un par rapport à l'autre en fonction de la régulation souhaitée. En effet lorsque les bruits sont essentiellement dans la gamme des fréquences de l'atténuation passive il peut être utile de couper la liaison entre le micro extérieur et les haut-parleurs sans qu'il soit nécessaire de mettre le circuit d'atténuation active en service.

[0042] Bien que dans les modes de réalisation décrits, le dispositif anti-bruit selon l'invention soit réalisé sous la forme d'un casque, le dispositif anti-bruit pourrait par exemple être intégré à l'écouteur d'un téléphone portable. Dans ce cas, le microphone de détection de seuil sera de préférence le microphone de communication du téléphone associé à un organe de filtrage.

Revendications

1. Dispositif anti-bruit comportant au moins une oreillette (13) ayant une cavité (2) délimitée par des moyens d'atténuation passive (3), un circuit d'atténuation active (4) ayant un domaine spatial d'efficacité et comportant au moins une unité de traitement (6) électronique reliée à des organes d'atténuation active (5, 7) comprenant au moins un microphone (5) captant des ondes sonores dans la cavité et un haut-parleur (7) disposé dans la cavité pour émettre des ondes sonores d'atténuation active élaborées par l'unité de traitement (6), et au moins un organe interrupteur (8) sélectivement actionnable entre un état de service du circuit d'atténuation active et un état de repos du circuit d'atténuation active, caractérisé en ce que le dispositif comporte un microphone de détection de seuil (10) captant des bruits extérieurs au domaine spatial d'efficacité du circuit d'atténuation active, des moyens de filtrage (3, 19) associés au microphone de détection de seuil pour éliminer des bruits correspondant aux bruits absorbés par les moyens d'atténuation passive, et un détecteur (9) de dépassement d'un seuil de bruit pour commander l'organe interrupteur (8) dans l'état de service ou de repos du circuit d'atténuation active en fonction d'un niveau de puissance des bruits détectés en aval des moyens de filtrage.

2. Dispositif anti-bruit selon la revendication 1, caractérisé en ce que le microphone de détection de seuil (10) est disposé à l'intérieur de la cavité (2) de l'oreillette (13).

3. Dispositif anti-bruit selon la revendication 1, caractérisé en ce que le microphone de détection de seuil (10) est porté par une paroi (1) de la cavité (2) de l'oreillette (13) pour capter des bruits extérieurs à la cavité, les moyens de filtrage comprenant un organe de filtrage (19) monté entre le microphone de détection de seuil (10) et le détecteur de dépassement de seuil (9).

4. Dispositif anti-bruit selon la revendication 3, caractérisé en ce qu'il comprend deux oreillettes, un microphone d'écoute (22) porté par une paroi (1) de la cavité (2) de la seconde oreillette (13) pour capter les bruits extérieurs à la cavité, des moyens de commutation (20) reliés au microphone de détection de seuil (10) et au microphone d'écoute (22) et sélectivement actionnables par un détecteur supplémentaire (21) de dépassement de seuil en fonction d'un niveau de puissance des bruits captés par le microphone de détection de seuil (10) entre un premier état dans lequel le microphone de détection de seuil (10) et le microphone d'écoute (22) sont connectés à des haut-parleurs respectifs (7) et un second état dans lequel le microphone de détection de seuil est connecté à l'organe de filtrage (19) et le microphone d'écoute est coupé.

5. Dispositif anti-bruit selon la revendication 4, caractérisé en ce que lesdits haut-parleurs respectifs sont les haut-parleurs (7) du circuit d'atténuation active (4).

6. Dispositif anti-bruit selon la revendication 4, caractérisé en ce qu'il comporte un circuit de communi-

Exhibit 1009
Page 1577 of 2198

cation, en ce que lesdits haut-parleurs respectifs sont associés au circuit de communication et en ce que le dispositif comporte des moyens (32) pour interrompre toute liaison entre les microphones et lesdits haut-parleurs respectifs lors de la réception d'une communication.

7. Dispositif anti-bruit selon l'une des revendications précédentes, caractérisé en ce qu'une résistance acoustique (12) est disposée sur les microphones (5) du circuit d'atténuation active.

Exhibit 1009
Page 1578 of 2198

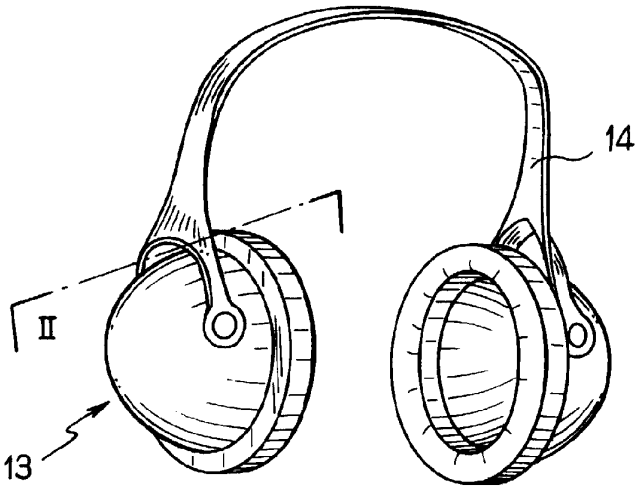


FIG_1

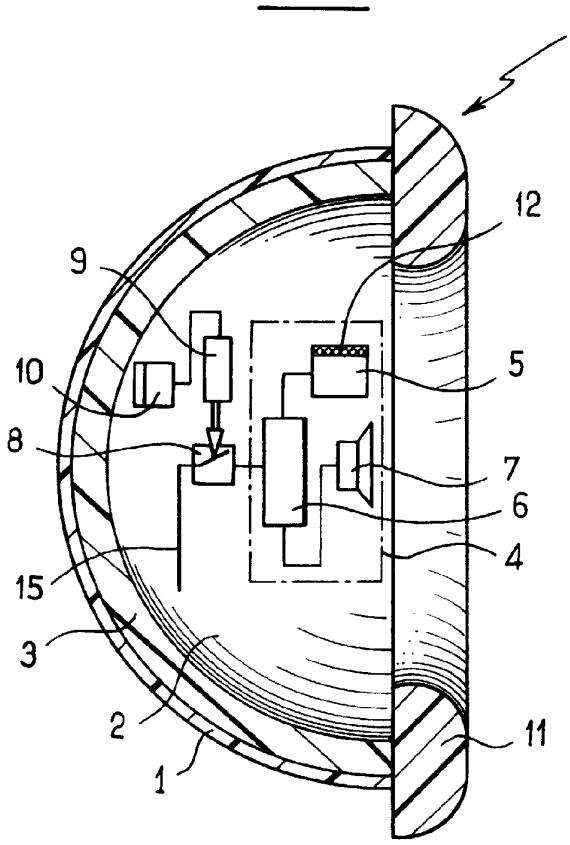


FIG_2

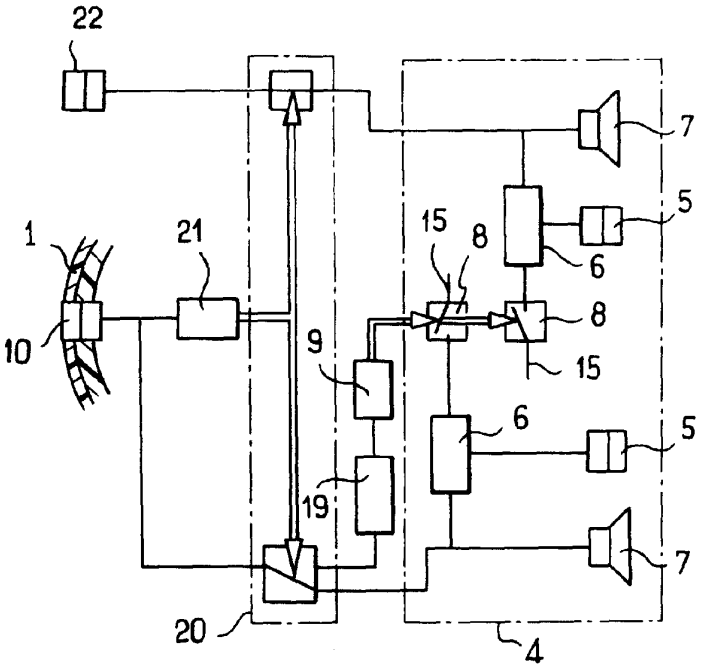

22
7
5
1
21
15
8
6
8
10
9
15
6
5
19
7
20
4

FIG.3

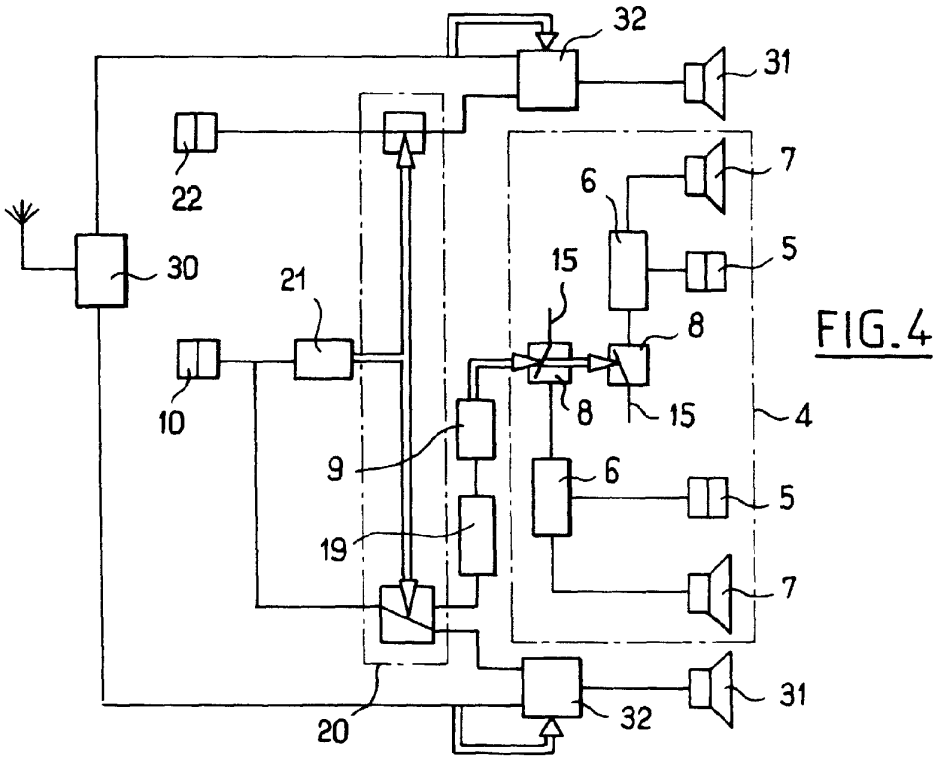

32
31
22
6
7
5
30
21
15
8
10
8
15
4
9
6
5
19
7
20
32
31

FIG.4

Office européen des brevets

RAPPORT DE RECHERCHE EUROPEENNE

Numéro de la demande
EP 99 40 0221

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.6)
X	WO 95 00946 A (ACTIVE NOISE & VIBRATION TECH) 5 janvier 1995 * abrégé * * page 1, ligne 23 - ligne 24 * * page 3, ligne 5 - ligne 30 * * page 4, ligne 33 - ligne 46 * * figures 5,12-14 * * page 15, ligne 18 - ligne 53 *	1,3,5,6	G10K11/178
A	GB 2 188 210 A (UNIV SOUTHAMPTON) 23 septembre 1987 * abrégé * * figures 8,9 *	1,7	
A	EP 0 425 129 A (BOSE CORP) 2 mai 1991 * abrégé * * figures 1,2,9 *	7	
A	FR 2 739 214 A (TECHNOFIRST) 28 mars 1997		
A	US 5 251 263 A (ANDREA DOUGLAS ET AL) 5 octobre 1993		DOMAINES TECHNIQUES RECHERCHES (Int.Cl.6)
A	EP 0 583 900 A (SONY CORP) 23 février 1994		G10K

Le présent rapport a été établi pour toutes les revendications

Lieu de la recherche	Date d'achèvement de la recherche	Examinateur
LA HAYE	2 juin 1999	de Heering, Ph.

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : arrière-plan technologique
O : divulgation non-écrite
P : document intercalaire

T : théorie ou principe à la base de l'invention
E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date
D : cité dans la demande
L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

EPO FORM 1503 03.82 (P04C02)

ANNEXE AU RAPPORT DE RECHERCHE EUROPEENNE RELATIF A LA DEMANDE DE BREVET EUROPEEN NO. EP 99 40 0221

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche européenne visé ci-dessus.
Lesdits members sont contenus au fichier informatique de l'Office européen des brevets à la date du
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets.

02-06-1999

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9500946 A	05-01-1995	AU 7355594 A	17-01-1995
		EP 0705472 A	10-04-1996
		JP 8510565 T	05-11-1996
GB 2188210 A	23-09-1987	AUCUN	
EP 0425129 A	02-05-1991	US 5305387 A	19-04-1994
		AT 135130 T	15-03-1996
		CA 2027340 A	28-04-1991
		DE 69025698 D	11-04-1996
		DE 69025698 T	25-07-1996
		JP 3215000 A	20-09-1991
FR 2739214 A	28-03-1997	AU 7136096 A	17-04-1997
		CA 2231071 A	03-04-1997
		EP 0852793 A	15-07-1998
		WO 9712359 A	03-04-1997
US 5251263 A	05-10-1993	AUCUN	
EP 0583900 A	23-02-1994	JP 6070391 A	11-03-1994
		DE 69317827 D	14-05-1998
		DE 69317827 T	13-08-1998
		US 5402497 A	28-03-1995

EPO FORM P0460

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No.12/82

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Control No.	90/019,169
U.S. Patent No.	11,244,666
Reexamination Filing Date	February 24, 2023
Inventors	Goldstein et al.
Title	METHOD AND DEVICE FOR ACUTE SOUND DETECTION AND REPRODUCTION
Examiner	Ge, Yuzhen
Art Unit	3992
Confirmation No.	7859

Submitted via EFS-Web
Mail Stop Ex Parte Reexamination
Attn: Central Reexamination Unit
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TRANSMITTAL LETTER FOR INFORMATION DISCLOSURE STATEMENT

Patent Owner herewith submits an Information Disclosure Statement (form PTO/SB/08a).

Patent Owner submits herewith the IDS fee. Should any additional fees be due in connection with this paper or at any other time during this proceeding, the Commissioner is authorized to charge such fees to Deposit Account No. 60-3815.

Dated: September 15, 2023

Respectfully submitted,

/s/ Jacob A. Snodgrass
Jacob A. Snodgrass
Registration No. 66,032

PV LAW LLP
Customer No. 176134
5335 Wisconsin Avenue, N.W., Suite 440
Washington, D.C. 20015
T: (202) 371-6861
F: (202) 888-3163

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. § 1.550(f), the undersigned hereby certifies that on September 15, 2023, a copy of this letter as well as the Information Disclosure Statement (form PTO/SB/08a) and cited non-U.S. patent documents and non-patent literature documents (on flash drive) were served on Third Party Requester at the following address.

Ali R. Sharifahmadian
Arnold & Porter Kaye Scholer LLP
601 Massachusetts Ave., N.W.
Washington, DC 20001

Dated: September 15, 2023

Respectfully submitted,

/s/ Jacob A. Snodgrass
Jacob A. Snodgrass
Registration No. 66,032

PV LAW LLP
Customer No. 176134
5335 Wisconsin Avenue, N.W., Suite 440
Washington, D.C. 20015
T: (202) 371-6861
F: (202) 888-3163

(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2005-227511
(P2005-227511A)
(43)公開日　平成17年8月25日(2005.8.25)

(51)Int.Cl.[7]	F I			テーマコード（参考）
G10L 15/20	G10L	3/02	301E	5D015
G10L 11/02	G10L	3/00	511	
G10L 15/02	G10L	3/00	513B	
G10L 15/04	G10L	9/08	301A	
G10L 15/28				

審査請求 未請求　請求項の数 15　ＯＬ　（全 22 頁）　最終頁に続く

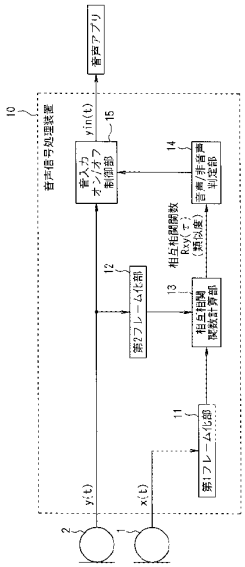
(21)出願番号　特願2004-35618(P2004-35618)
(22)出願日　平成16年2月12日(2004.2.12)

(71)出願人　000010076
ヤマハ発動機株式会社
静岡県磐田市新貝２５００番地
(74)代理人　100066980
弁理士　森　哲也
(74)代理人　100075579
弁理士　内藤　嘉昭
(74)代理人　100103850
弁理士　崔　秀▲てつ▼
(72)発明者　有宗　伸泰
静岡県磐田市新貝２５００番地　ヤマハ発動機株式会社内
(72)発明者　赤坂　貴志
静岡県磐田市新貝２５００番地　ヤマハ発動機株式会社内

Ｆターム(参考)　5D015 DD02　DD03　EE04

(54)【発明の名称】対象音検出方法、音信号処理装置、音声認識装置及びプログラム

(57)【要約】
【課題】マイクの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系の構築を可能にする。
【解決手段】音声信号処理装置は、発話音又は雑音のいずれか一方を受音するように配置されている単一指向性マイク１と、発話音及び雑音を受音するように配置されている無指向性マイク２と、単一指向性マイク１に入力された音声信号ｘ（ｔ）と無指向性マイク２に入力された音声信号ｙ（ｔ）との相互相関関数Ｒｘｙ（τ）を算出するためのフレーム化部１１，１２及び相互相関関数計算部１３と、相互相関関数計算部１３が算出した相互相関関数Ｒｘｙ（τ）に基づいて、発話音の発話区間を検出する音声／非音声判定部１４とを備える。
【選択図】図３



【特許請求の範囲】
【請求項１】
無指向性マイクで検出対象音及び雑音を受音し、単一指向性マイクで前記検出対象音又は前記雑音のいずれか一方を受音し、前記無指向性マイクに入力された声信号と単一指向性マイクに入力された音信号とを比較し、その比較結果に基づいて、前記検出対象音を検出することを特徴とする対象音検出方法。
【請求項２】
前記無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号との比較により相関度を得て、その相関度に基づいて、前記検出対象音を検出することを特徴とする請求項１記載の対象音検出方法。
【請求項３】
前記無指向性マイクに入力された音信号のパワースペクトルと、単一指向性マイクに入力された音信号のパワースペクトルとを比較して、その比較結果に基づいて、前記検出対象音を検出することを特徴とする請求項１又は２記載の対象音検出方法。
【請求項４】
前記無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号との比較により得た相関度、及び前記無指向性マイクに入力された音信号のパワースペクトルと、単一指向性マイクに入力された音信号のパワースペクトルとの比較結果に基づいて、前記検出対象音を検出することを特徴とする請求項１乃至３のいずれか１に記載の対象音検出方法。
【請求項５】
検出対象音及び雑音を受音するように配置されている無指向性マイクと、
前記検出対象音又は前記雑音のいずれか一方を受音するように配置されている単一指向性マイクと、
前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較する比較手段と、
前記比較手段の比較結果に基づいて、前記検出対象音を検出する対象音検出手段と、
を備えることを特徴とする音信号処理装置。
【請求項６】
前記比較手段は、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記対象音検出手段は、前記比較手段が算出した相関度と所定の第１しきい値とを比較して、前記検出対象音を検出することを特徴とする請求項５記載の音信号処理装置。
【請求項７】
前記比較手段は、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、
前記対象音検出手段は、前記パワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値とを比較して、前記検出対象音を検出することを特徴とする請求項５又は６記載の音信号処理装置。
【請求項８】
前記比較手段は、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、
前記比較手段は、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記対象音検出手段は、前記比較手段が算出した相関度と所定の第１しきい値との比較結果と、前記比較手段のパワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値との比較結果とに基づいて、

前記検出対象音を検出することを特徴とする請求項５乃至７のいずれか１に記載の音信号処理装置。

【請求項９】

前記無指向性マイクに入力された音信号及び単一指向性マイクに入力された音信号を時分割してフレーム化するフレーム化手段を備えており、

前記比較手段は、前記フレーム化手段から出力されるフレーム単位で、前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較し、前記対象音検出手段は、前記比較手段の比較結果に基づいて、前記フレーム化手段から出力されるフレーム単位で前記検出対象音を検出することを特徴とする請求項５乃至８のいずれか１に記載の音信号処理装置。

【請求項１０】

発話音及び雑音を受音するように配置されている無指向性マイクと、

前記発話音又は前記雑音のいずれか一方を受音するように配置されている単一指向性マイクと、

前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較する比較手段と、

前記比較手段の比較結果に基づいて、前記発話音の発話区間を検出する発話区間検出手段と、

前記発話区間検出手段が検出した前記発話音の発話区間について、音声認識処理を行う音声認識処理手段と、

を備えることを特徴とする音声認識装置。

【請求項１１】

前記比較手段は、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記発話区間検出手段は、前記比較手段が算出した相関度と所定の第１しきい値とを比較して、前記発話音の発話区間を検出することを特徴とする請求項１０記載の音声認識装置。

【請求項１２】

前記比較手段は、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、

前記発話区間検出手段は、前記パワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値とを比較して、前記発話音の発話区間を検出することを特徴とする請求項１０又は１１記載の音声認識装置。

【請求項１３】

前記比較手段は、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、

前記比較手段は、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記発話区間検出手段は、前記比較手段が算出した相関度と所定の第１しきい値との比較結果と、前記比較手段のパワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値との比較結果とに基づいて、前記発話音の発話区間を検出することを特徴とする請求項１０乃至１２のいずれか１に記載の音声認識装置。

【請求項１４】

前記無指向性マイクに入力された音信号及び単一指向性マイクに入力された音信号を時分割してフレーム化するフレーム化手段を備えており、

前記比較手段は、前記フレーム化手段から出力されるフレーム単位で、前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較し、前記発話

区間検出手段は、前記比較手段の比較結果に基づいて、前記フレーム化手段から出力されるフレーム単位で前記発話音の発話区間を検出し、前記音声認識処理手段は、前記発話区間検出手段が検出したフレーム単位の前記発話音の発話区間について、音声認識処理を行うことを特徴とする請求項１０乃至１３のいずれか１に記載の音声認識装置。

【請求項１５】

無指向性マイクで受音した検出対象音及び雑音の音信号と単一指向性マイクで受音した前記検出対象音又は前記雑音のいずれか一方の音信号とを比較し、その比較結果に基づいて、前記検出対象音を検出する処理をコンピュータに実行させることを特徴とするプログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、入力音中から検出対象音を検出する対象音検出方法及びこれを実現するプログラム、入力された音信号を処理する音信号処理装置、並びに入力された発話音について音声認識処理を行う音声認識装置に関する。

【背景技術】

【０００２】

音声は、人間の用いる種々の通信の形態の中でも最も根源的なものであると同時に、他のどの情報送出方法よりも高速度に情報を送り出すことのできる優れた通信手段である。このようなことから、音声は、古くから現在に至るまで人間の通信手段の根幹を担っている。

音声認識技術は、そのような音声を認識するための技術である。音声認識とは、その音声に含まれる情報の中で、最も基本的な意味内容に関する情報、つまり音韻情報をコンピュータなどにより抽出し、その抽出内容を判定することである。近年では、計算機プロセッサ技術の飛躍的な発達と、インターネットに代表される高度な情報ネットワークの構築により、様々な分野においてマン・マシンインタフェースとしての音声認識技術の適用が試みられている。

現在の音声認識システムの認識性能は、確率・統計的手法により格段に向上しており、理想的な環境下での音声や接話マイクロホンで収録された近距離音声などでは、非常に高い認識率が得られるようになっている。

【発明の開示】

【発明が解決しようとする課題】

【０００３】

実環境下の音声認識は、学習データと観測データとの間の環境や発話内容のミスマッチ等により、その認識率が劣化する。また、受音系となる接話マイクヘッドセットの装着によりユーザが受ける負担や不快感は大きく、音声認識システム実用化の大きな障害のひとつになっている。

また、Ｓ／Ｎ比の低下や背景雑音、室内残響の影響などにより認識が困難な遠隔音声に関し、複数の遠隔マイクロホンを用いた音声認識手法の研究が多くなされている。その代表的なものとして、マイクロホンアレーを用いる手法が挙げられる。この手法では、音源位置検出処理、目的音強調処理、雑音抑制処理、の３つの空間的な信号処理を行なうことができる。このような手法により遠隔音声の音声認識が盛んに研究されている。

【０００４】

しかし、この手法は、正確な話者方向同定処理のために複数のマイクロホンを一定間隔にて固定配置する必要があり、小型化、携帯化が困難であるため、様々な環境・状況下での音声入力への応用が難しく、用途が限定されるという問題がある。

本発明は、前述の問題に鑑みてなされたものであり、マイクロホンの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系の構築を可能にする対象音検出方法、音信号処理装置、音声認識装置及びプログラムの提供を目的とする。

【課題を解決するための手段】

【０００５】

請求項１記載の対象音検出方法は、無指向性マイクで検出対象音及び雑音を受音し、単一指向性マイクで前記検出対象音又は前記雑音のいずれか一方を受音し、前記無指向性マイクに入力された声信号と単一指向性マイクに入力された音信号とを比較し、その比較結果に基づいて、前記検出対象音を検出することを特徴とする。

また、請求項２記載の対象音検出方法は、請求項１記載の対象音検出方法において、前記無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号との比較により相関度を得て、その相関度に基づいて、前記検出対象音を検出することを特徴とする。

【０００６】

また、請求項３記載の対象音検出方法は、請求項１又は２記載の対象音検出方法において、前記無指向性マイクに入力された音信号のパワースペクトルと、単一指向性マイクに入力された音信号のパワースペクトルとを比較して、その比較結果に基づいて、前記検出対象音を検出することを特徴とする。

また、請求項４記載の対象音検出方法は、請求項１乃至３のいずれか１に記載の対象音検出方法において、前記無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号との比較により得た相関度、及び前記無指向性マイクに入力された音信号のパワースペクトルと単一指向性マイクに入力された音信号のパワースペクトルとの比較結果に基づいて、前記検出対象音を検出することを特徴とする。

【０００７】

また、請求項５記載の音信号処理装置は、検出対象音及び雑音を受音するように配置されている無指向性マイクと、前記検出対象音又は前記雑音のいずれか一方を受音するように配置されている単一指向性マイクと、前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較する比較手段と、前記比較手段の比較結果に基づいて、前記検出対象音を検出する対象音検出手段と、を備えることを特徴とする。

【０００８】

また、請求項６記載の音信号処理装置は、請求項５記載の音信号処理装置において、前記比較手段が、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記対象音検出手段が、前記比較手段が算出した相関度と所定の第１しきい値とを比較して、前記検出対象音を検出することを特徴とする。

また、請求項７記載の音信号処理装置は、請求項５又は６記載の音信号処理装置において、前記比較手段が、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、前記対象音検出手段が、前記パワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値とを比較して、前記検出対象音を検出することを特徴とする。

【０００９】

また、請求項８記載の音信号処理装置は、請求項５乃至７のいずれか１に記載の音信号処理装置において、前記比較手段が、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、前記比較手段が、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記対象音検出手段は、前記比較手段が算出した相関度と所定の第１しきい値との比較結果と、前記比較手段のパワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値との比較結果とに基づいて、前記検出対象音を検出することを特徴とする。

【００１０】

また、請求項９記載の音信号処理装置は、請求項５乃至８のいずれか１に記載の音信号

処理装置において、前記無指向性マイクに入力された音信号及び単一指向性マイクに入力された音信号を時分割してフレーム化するフレーム化手段を備えており、前記比較手段が、前記フレーム化手段から出力されるフレーム単位で、前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較し、前記対象音検出手段は、前記比較手段の比較結果に基づいて、前記フレーム化手段から出力されるフレーム単位で前記検出対象音を検出することを特徴とする。

【００１１】

また、請求項１０記載の音声認識装置は、発話音及び雑音を受音するように配置されている無指向性マイクと、前記発話音又は前記雑音のいずれか一方を受音するように配置されている単一指向性マイクと、前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較する比較手段と、前記比較手段の比較結果に基づいて、前記発話音の発話区間を検出する発話区間検出手段と、前記発話区間検出手段が検出した前記発話音の発話区間について、音声認識処理を行う音声認識処理手段と、を備えることを特徴とする。

【００１２】

また、請求項１１記載の音声認識装置は、請求項１０記載の音声認識装置において、前記比較手段が、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記発話区間検出手段が、前記比較手段が算出した相関度と所定の第１しきい値とを比較して、前記発話音の発話区間を検出することを特徴とする。

【００１３】

また、請求項１２記載の音声認識装置は、請求項１０又は１１記載の音声認識装置において、前記比較手段が、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、前記発話区間検出手段が、前記パワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値とを比較して、前記発話音の発話区間を検出することを特徴とする。

【００１４】

また、請求項１３記載の音声認識装置は、請求項１０乃至１２のいずれか１に記載の音声認識装置において、前記比較手段が、前記無指向性マイク及び単一指向性マイクに入力された各音信号のパワースペクトルをそれぞれ算出するパワースペクトル算出手段と、前記パワースペクトル算出手段が算出した２つのパワースペクトルの比を算出するパワースペクトル比算出手段とを備えており、前記比較手段が、前記無指向性マイクに入力された音信号と前記単一指向性マイクに入力された音信号との相関度を算出し、前記発話区間検出手段は、前記比較手段が算出した相関度と所定の第１しきい値との比較結果と、前記比較手段のパワースペクトル比算出手段が算出した前記パワースペクトルの比と所定の第２しきい値との比較結果とに基づいて、前記発話音の発話区間を検出することを特徴とする。

【００１５】

また、請求項１４記載の音声認識装置は、請求項１０乃至１３のいずれか１に記載の音声認識装置において、前記無指向性マイクに入力された音信号及び単一指向性マイクに入力された音信号を時分割してフレーム化するフレーム化手段を備えており、前記比較手段が、前記フレーム化手段から出力されるフレーム単位で、前記無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較し、前記発話区間検出手段が、前記比較手段の比較結果に基づいて、前記フレーム化手段から出力されるフレーム単位で前記発話音の発話区間を検出し、前記音声認識処理手段は、前記発話区間検出手段が検出したフレーム単位の前記発話音の発話区間について、音声認識処理を行うことを特徴とする。

【００１６】

また、請求項１５記載のプログラムは、無指向性マイクで受音した検出対象音及び雑音の音信号と単一指向性マイクで受音した前記検出対象音又は前記雑音のいずれか一方の音信号とを比較し、その比較結果に基づいて、前記検出対象音を検出する処理をコンピュータに実行させることを特徴とする。

無指向性マイクで検出対象音及び雑音を受音し、単一指向性マイクで検出対象音又は雑音のいずれか一方を受音するようにした場合において、無指向性マイク及び単一指向性マイクが配置されている環境で雑音だけを発したときと検出対象音及び雑音を発したときとで、無指向性マイク及び単一指向性マイクに入力された音信号が異なってくる。なお、検出対象音には、人間が発する発話音の他、物体が発する音も含まれる。

【００１７】

そこで、請求項１、５、１５記載の発明では、無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号とを比較することで、検出対象音又は発話音を検出している。また、請求項１０記載の発明では、検出対象音が人間が発する発話音であり、検出対象音の検出として、発話音の音声区間の検出を行っている。

また、請求項２、６、１１記載の発明では、無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号との相関度により、検出対象音の検出又は発話音の発話区間の検出を行っている。

【００１８】

また、請求項３、７、１２記載の発明では、無指向性マイクに入力された音信号のパワースペクトルと単一指向性マイクに入力された音信号のパワースペクトルとを比較することで、検出対象音の検出又は発話音の発話区間の検出を行っている。

また、請求項４、８、１３記載の発明では、無指向性マイクに入力された音信号と単一指向性マイクに入力された音信号との相関度と、無指向性マイクに入力された音信号のパワースペクトルと単一指向性マイクに入力された音信号のパワースペクトルとの比較結果とに基づいて、検出対象音の検出又は発話音の発話区間の検出を行っている。

また、請求項９、１４記載の発明では、前記無指向性マイクに入力された音信号及び単一指向性マイクに入力された音信号を時分割してフレーム化し、フレーム単位でその後の処理を行う。

【発明の効果】

【００１９】

本発明によれば、無指向性マイクで検出対象音（又は発話音）及び雑音を受音し、単一指向性マイクで前記検出対象音（又は発話音）又は前記雑音のいずれか一方を受音するように、無指向性マイク及び単一指向性マイクを配置する限り、検出対象音（又は発話音の音声区間）を検出することができる。これにより、マイクロホンの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系の構築が可能になる。

【発明を実施するための最良の形態】

【００２０】

本発明を実施するための最良の形態（以下、実施形態という。）を図面を参照しながら詳細に説明する。

第１の実施形態は、図１に示すように、第１及び第２マイク１，２に入力された音声信号を処理する音声信号処理装置１０である。

第１マイク１は単一指向性マイクであり、第２マイク２は無指向性マイクであり、第１及び第２マイク１，２は例えば装着型マイクである。第１及び第２マイク１，２は、図２に示すように、第１及び第２マイク１，２をできるだけ近づけて配置するとともに、単一指向性マイクである第１マイク１をその指向方向が音源（ユーザ）の位置に対して反対側となるように配置している。また、第１マイク１の指向方向に、雑音源が存在している。なお、図２に示す点線は、雑音源を基準にした第１マイク１の指向特性を示し、図２に示す一点鎖線は、第２マイク２の指向特性を示す。

このように第１及び第２マイク１，２を配置すると、雑音源からの音は、第１及び第２

マイク１，２で受音でき、音源（ユーザ）からの音は第２マイク２だけが受音できるようになる。

【００２１】

図３は、第１の実施形態の音声信号処理装置１０の構成を示す。

図３に示すように、音声信号処理装置１０は、第１及び第２フレーム化部１１，１２、相互相関関数計算部１３、音声／非音声判定部１４並びに音入力オン／オフ制御部１５を備えている。

第１及び第２マイク１，２から入力された２ｃｈの音声信号はそれぞれ、第１及び第２フレーム化部１１，１２に入力される。また、第２マイク２から入力された音声信号は、音入力オン／オフ制御部１５に入力される。ここで、第１マイク１に入力された音声信号をｘ（ｔ）とし、第２マイク２に入力された音声信号をｙ（ｔ）とする。

【００２２】

第１フレーム化部１１では、第１マイク１から入力された音声信号ｘ（ｔ）を時分割でフレーム化（或いはフレーム分割）して、複数フレームにした音声信号ｘ（ｔ）を相互相関関数計算部１３に出力する。また、第２フレームか部１２では、第２マイク２から入力される音声信号ｙ（ｔ）を時分割でフレーム化（或いはフレーム分割）して、複数フレームにした音声信号ｙ（ｔ）を相互相関関数計算部１３に出力する。ここで、第１及び第２フレーム化部１１，１２は、入力されてくる音声信号ｘ（ｔ），ｙ（ｔ）を所定時間間隔でサンプリングしていき、所定のサンプル数を１フレームとして次々にフレーム化していく。

【００２３】

相互相関関数計算部１３は、第１フレーム化部１１から出力されるフレームと、第２フレーム化部１２から出力されるフレームとを比較する。すなわち、第１マイク１に入力された音声信号ｘ（ｔ）と、第２マイク２に入力された音声信号ｙ（ｔ）とをフレーム単位で比較する。その比較結果として、相互相関関数計算部１３は、下記（１）式により、相互相関関数Ｒｘｙ（τ）を算出する。

【００２４】

【数１】

$$R_{xy}(\tau)=\frac{1}{T}\int_{-T/2}^{T/2}x(t)y(t-\tau)dt \quad \cdots (1)$$

【００２５】

ここで、τは第１マイク１と第２マイク２との間の距離によって決まる遅延時間である。また、Ｔはフレーム長である。

前述したように第１及び第２マイク１，２をできるだけ近づけて配置している場合には、遅延時間τを近似的に０とおくことができる。しかし、後述するような本発明の効果を満たす限り、第１マイク１と第２マイク２とを離して配置することは可能であり、この場合、遅延時間τを適切に与える必要がある。すなわち例えば、第１マイク１と第２マイク２との間の距離を１０ｃｍにしている場合には、その１０ｃｍ相当分の遅延時間τを与えて、相互相関関数Ｒｘｙ（τ）を算出する。このようにすれば、第１マイク１と第２マイク２との間の距離を考慮して、相互相関関数Ｒｘｙ（τ）を得ることができ、精度よく相互相関関数Ｒｘｙ（τ）を得ることができる。

【００２６】

このように算出された相互相関関数Ｒｘｙ（τ）はフレーム単位で各音声信号ｘ（ｔ），ｙ（ｔ）の波形形状の類似度を示す値となる。具体的には、相関関係を求める２つの音声信号ｘ（ｔ），ｙ（ｔ）が似ているほど、相互相関関数Ｒｘｙ（τ）は大きい値となり、相関関係を求める２つの音声信号ｘ（ｔ），ｙ（ｔ）が異なっているほど、相互相関関数Ｒｘｙ（τ）は０に近くなる。相互相関関数計算部１３は、このような相互相関関数Ｒｘｙ（τ）を音声／非音声判定部１４に出力する。

【００２７】
音声／非音声判定部１４は、相互相関関数Ｒｘｙ（τ）に基づいて音声区間（発話区間）と非音声区間（非発話区間）とを判定する。具体的には、次のように音声区間と非音声区間とを判定する。
前述したように、音源（ユーザ）と雑音源に対して図２のように第１及び第２マイク１，２を配置することで、雑音源からの音を第１及び第２マイク１，２で受音し、音源（ユーザ）からの音を第２マイク２だけで受音している。
【００２８】
一方、相互相関関数Ｒｘｙ（τ）は、前述したように、相関関係を求める２つの音声信号ｘ（ｔ），ｙ（ｔ）が似ているほど大きい値となり、相関関係を求める２つの音声信号ｘ（ｔ），ｙ（ｔ）が異なっているほど０に近くなる。
このようなことから、雑音源からの音だけを第１及び第２マイク１，２で受音している場合には、同じ音声信号が第１及び第２マイク１，２に入力されているので、すなわち、第１及び第２マイク１，２の入力音声信号のＳ／Ｎ比が同程度になるので、相互相関関数Ｒｘｙ（τ）は大きい値になる。一方、音源（ユーザ）から発話があった場合には、その発話を第２マイク２だけが受音するので、第１及び第２マイク１，２それぞれに異なる音声信号が入力されるようになり、すなわち第２マイク２の入力音声信号のＳ／Ｎ比の方が大きくなるので、相互相関関数Ｒｘｙ（τ）は０に向かって減少する。
【００２９】
このように、音源（ユーザ）から発話があった場合には相互相関関数Ｒｘｙ（τ）は０に向かって減少することから、音声／非音声判定部１４は、相互相関関数Ｒｘｙ（τ）と判定用しきい値（類似度を示すしきい値）δｒ１とを比較して、音声区間を判定する。すなわち、音声／非音声判定部１４は、相互相関関数Ｒｘｙ（τ）が判定用しきい値δｒ１未満の場合（Ｒｘｙ（τ）＜δｒ１）、音声区間と判定し、それ以外の場合（Ｒｘｙ（τ）≧δｒ１）、非音声区間と判定する。ここで、判定用しきい値δｒ１は例えば実験により得る。そして、音声／非音声判定部１４は、このような判定をフレーム単位で行う。音声／非音声判定部１４は、その判定結果をフレーム単位で音入力オン／オフ制御部１５に出力する。
【００３０】
音入力オン／オフ制御部１５には、第２マイク２からの音声信号ｙ（ｔ）が入力されており、音入力オン／オフ制御部１５は、音声／非音声判定部１４の判定結果に基づいて、第２マイク２からの音声信号ｙ（ｔ）の後段への出力のオンとオフとを切り換える。具体的には、音声／非音声判定部１４が音声区間と判定した場合、音入力オン／オフ制御部１５は、オン制御として当該音声区間に対応する音声信号ｙ（ｔ）の区間を後段に出力して、音声／非音声判定部１４が非音声区間と判定した場合、音入力オン／オフ制御部１５は、オフ制御として当該非音声区間に対応する音声信号ｙ（ｔ）の区間を後段に出力しないようにする。
【００３１】
以上のように音声信号処理装置１０が構成されている。この音声信号処理装置１０における一連の動作は次のようになる。
先ず、第１及び第２フレーム化部１１，１２が、第１及び第２マイク１，２から入力された２ｃｈの音声信号ｘ（ｔ），ｙ（ｔ）をそれぞれフレーム化し、フレーム単位で音声信号ｘ（ｔ），ｙ（ｔ）を相互相関関数計算部１３に出力する。
【００３２】
相互相関関数計算部１３では、第１及び第２フレーム化部１１，１２それぞれから出力されるフレーム単位の音声信号ｘ（ｔ），ｙ（ｔ）について相互相関関数Ｒｘｙ（τ）を算出して、算出した相互相関関数Ｒｘｙ（τ）を音声／非音声判定部１４に出力する。
音声／非音声判定部１４では、相互相関関数Ｒｘｙ（τ）と判定用しきい値δｒ１とを比較し、相互相関関数Ｒｘｙ（τ）に対応するフレームが音声区間のものか、非音声区間のものかを判定する。そして、音声／非音声判定部１４は、その判定結果を音入力オン／

オフ制御部１５に出力する。

【００３３】

音入力オン／オフ制御部１５は、音声／非音声判定部１４が音声区間と判定した場合、オン制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力して、音声／非音声判定部１４が非音声区間と判定した場合、オフ制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力しないようにする。このとき、音入力オン／オフ制御部１５から出力される音声信号ｙ（ｔ）は、音源（ユーザ）からの音と雑音源からの音とからなる音声信号となる。

【００３４】

このように、音声信号処理装置１０は、第２マイク２への入力音中の発話区間（音声区間）を検出することができる。

例えば、第１マイク１，２と音声アプリケーションとの間にこのような音声信号処理装置１０を備えることで、音声アプリケーションは、確実に発話区間についての処理を行うことができる。ここで、音声アプリケーションとしては、音声認識システム、放送システム、携帯電話及びトランシーバが挙げられる。例えば、音声アプリケーションが音声認識システムであるとすれば、音声認識システムは、音声信号処理装置１０が出力する発話区間の音声信号に基づいて音声認識できるようになる。

【００３５】

次に第１の実施形態における効果を説明する。

前述したように、無指向性マイクである第２のマイク２で発話音及び雑音を受音し、単一指向性マイクである第１マイク１で雑音を受音し、第１マイク１で受音した雑音の音声信号と第２マイク２で受音した発話音及び雑音からなる音声信号との比較により相関度を得て、その相関度に基づいて、発話音の発話区間を特定している。

【００３６】

これにより、第２のマイク２で発話音及び雑音を受音し、かつ第１マイク１で雑音を受音するように第１及び第２のマイク１，２を配置する限り、マイクの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系を構築することができる。

また、精度よく発話区間を検出することができる。そして、このように音声信号処理装置１０が検出した発話区間の音声信号を利用することにより、音声認識システムでは、高認識率、低誤認識率の音声認識が可能になり、また、携帯電話やトランシーバでは、信頼性の高いハンズフリー半二重通信が可能になり、放送システムでは、通信システムの送信電力低減が可能になる。

【００３７】

次に第２の実施形態を説明する。

この第２の実施形態も前述の第１の実施形態と同様、前記図１に示したように２つのマイク１，２に入力された音声信号を処理する音声信号処理装置１０である。そして、音声信号処理装置１０の構成は、前述の第１の実施形態と同様、図３に示したような構成になる。しかし、第２の実施形態では、第１及び第２マイク１，２の配置を前述の第１の実施形態における配置と異ならせている。

前述の第１の実施形態では、第１マイク１に単一指向性マイクを用い、第２マイク２に無指向性マイクを用い、前記図２に示したように、雑音源からの音を第１及び第２マイク１，２で受音し、音源（ユーザ）からの音を第１マイク１だけで受音するように、第１及び第２マイク１，２を配置している。

【００３８】

一方、この第２の実施形態では、音源（ユーザ）からの音を第１及び第２マイク１，２で受音し、雑音源からの音を第１マイク１だけで受音するようにしている。具体的には、第１マイク１に無指向性マイクを用い、第２マイク２に単一指向性マイクを用いる。そして、図４に示すように、第１及び第２マイク１，２をできるだけ近づけて配置するとともに、単一指向性マイクである第２マイク２を、その指向方向が音源（ユーザ）に向かい、

かつその指向方向外に雑音源が位置されるように、配置する。なお、図４に示す点線は、第１マイク１の指向特定を示し、図４に示す一点鎖線は、音源（ユーザ）を基準にした第２マイク２の指向特性を示す。

【００３９】

このように第１及び第２マイク１，２を配置した場合、前述の第１の実施形態と比較し、特に相互相関関数計算部１３で算出される相互相関関数Ｒｘｙ（τ）が異なる傾向を示すようになる。

すなわち、音源（ユーザ）からの音を第１及び第２マイク１，２で受音し、雑音源からの音を第２マイク２だけが受音しているので、雑音源からの音だけを第１マイク１で受音している場合には、第１及び第２マイク１，２それぞれに異なる音声信号が入力されるようになり、このとき相互相関関数Ｒｘｙ（τ）は０に近い値になる。一方、音源（ユーザ）から発話があった場合には、その発話を第１及び第２マイク１，２で受音するので、ほぼ同じ音声信号が第１及び第２マイク１，２に入力されるようになり、これにより、相互相関関数Ｒｘｙ（τ）は大きい値になる。このとき、第２マイク２の入力音声信号のＳ／Ｎ比は高くなり、第１マイク１の入力音声信号のＳ／Ｎ比は、第２マイク２ほどではないが、高くなる。

【００４０】

このように、音源（ユーザ）から発話があった場合には、相互相関関数Ｒｘｙ（τ）が大きくなり、第２の実施形態で得る相互相関関数Ｒｘｙ（τ）は、前述の第１の実施形態とは反対の傾向を示すようになる。

このようなことから、第２の実施形態では、音声／非音声判定部１４は、相互相関関数Ｒｘｙ（τ）と判定用しきい値（類似度を示すしきい値）δｒ２とを比較して、相互相関関数Ｒｘｙ（τ）が判定用しきい値δｒ２より大きい場合（Ｒｘｙ（τ）＞δｒ２）、音声区間と判定し、それ以外の場合（Ｒｘｙ（τ）≦δｒ２）、非音声区間と判定する。ここで、判定用しきい値δｒ２は例えば実験により得る。

【００４１】

そして、前述の第１の実施形態と同様に、音声／非音声判定部１４が音声区間と判定した場合、音入力オン／オフ制御部１５は、オン制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力して、音声／非音声判定部３４が非音声区間と判定した場合、音入力オン／オフ制御部１５は、オフ制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力しないようにする。このとき、音入力オン／オフ制御部１５から出力される音声信号ｙ（ｔ）は、音源（ユーザ）からの音のみからなる音声信号となる。

このように、第２の実施形態の音声信号処理装置１０は、第２マイク２への入力音中の発話区間（音声区間）を検出することができる。

【００４２】

次に第２の実施形態における効果を説明する。

前述したように、無指向性マイクである第１のマイク１で発話音及び雑音を受音し、単一指向性マイクである第２マイク２で発話音を受音し、第１マイク１で受音した発話音及び雑音からなる音声信号と第２マイク２で受音した発話音の音声信号との比較により相関度を得て、その相関度に基づいて、発話音の発話区間を特定している。

【００４３】

これにより、第１のマイク１で発話音及び雑音を受音し、かつ第２マイク２で発話音を受音するように第１及び第２のマイク１，２を配置する限り、マイクの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系を構築することができる。

また、精度よく発話区間を検出することができる。そして、このように音声信号処理装置１０が検出した発話区間の音声信号を利用することにより、音声認識システムでは、高認識率、低誤認識率の音声認識が可能になり、また、携帯電話やトランシーバでは、信頼性の高いハンズフリー半二重通信が可能になり、放送システムでは、通信システムの送信電力低減が可能になる。

【００４４】
次に第３の実施形態を説明する。
この第３の実施形態も前述の第１の実施形態と同様、前記図１に示したように２つのマイク１，２に入力された音声信号を処理する音声信号処理装置１０である。
前述の第１及び第２の実施形態では、相互相関関数計算部１３により音声信号ｘ（ｔ），ｙ（ｔ）の相互相関関数Ｒｘｙ（τ）を算出し、この相互相関関数Ｒｘｙ（τ）に基づいて音声区間と非音声区間とを判定している。
【００４５】
これに対して、第３の実施形態の音声信号処理装置１０は、音声信号ｘ（ｔ），ｙ（ｔ）それぞれのパワースペクトルを算出し、このパワースペクトルに基づいて音声区間と非音声区間とを判定するように構成されている。図５は、その第３の実施形態の音声信号処理装置１０の構成を示す。
図５に示すように、音声信号処理装置１０は、第１及び第２フレーム化部１１，１２、第１及び第２パワースペクトラム計算部２１，２２、パワー比計算部２３、音声／非音声判定部２４並びに音入力オン／オフ制御部１５を備えている。各部の処理内容は次のようになる。
【００４６】
なお、第１及び第２フレーム化部１１，１２及び音入力オン／オフ制御部１５については、前述の第１の実施形態のものと同様な処理を行うので、その説明を省略する。また、前述の第１の実施形態と同様に、第１マイク１は単一指向性マイクであり、第２マイク２は無指向性マイクである。さらに、第１及び第２マイク１，２の配置についても、前記図２に示したような配置にしている。これにより、雑音源からの音を第１及び第２マイク１，２で受音し、音源（ユーザ）からの音を第２マイク２だけで受音している。
【００４７】
第１及び第２フレーム化部１１，１２それぞれで複数フレームにされた音声信号ｘ（ｔ），ｙ（ｔ）は第１及び第２パワースペクトラム計算部２１，２２に入力される。
第１パワースペクトラム計算部２１は、フレーム単位で音声信号ｘ（ｔ）の第１パワースペクトル値Ｐｘ（ω）を算出し、その算出した第１パワースペクトル値Ｐｘ（ω）をパワー比計算部２３に出力する。また、第２パワースペクトラム計算部２２は、フレーム単位で音声信号ｙ（ｔ）の第２パワースペクトル値Ｐｙ（ω）を算出し、その算出した第２パワースペクトル値Ｐｙ（ω）をパワー比計算部２３に出力する。
パワー比計算部２３は、下記（２）式により、第１パワースペクトラム計算部２１からの第１パワースペクトル値Ｐｘ（ω）と、第２パワースペクトラム計算部２２からの第２パワースペクトル値Ｐｙ（ω）との比（以下、パワー比という。）Ｐｘｙ（ω）を算出する。
【００４８】
【数２】

$$P_{xy}(\omega)=G_{xy}\frac{P_x(\omega)}{P_y(\omega)} \qquad \cdots (2)$$

【００４９】
ここで、Ｇｘｙは、第１及び第２マイク１，２の感度によって決まる補正係数である。このように算出されたパワー比Ｐｘｙ（ω）はフレーム単位で各音声信号ｘ（ｔ），ｙ（ｔ）の波形形状の類似度を示す値となる。パワー比計算部２３は、このようなパワー比Ｐｘｙ（ω）を音声／非音声判定部２４に出力する。
音声／非音声判定部２４は、パワー比Ｐｘｙ（ω）に基づいて音声区間と非音声区間とを判定する。具体的には、次のように音声区間と非音声区間とを判定する。
前述したように、音源（ユーザ）と雑音源に対して前記図２のように第１及び第２マイク１，２を配置することで、雑音源からの音を第１及び第２マイク１，２で受音し、音源

（ユーザ）からの音を第２マイク２だけで受音している。

【００５０】

これにより、雑音源からの音だけを第１及び第２マイク１，２で受音している場合には、同じ音声信号が第１及び第２マイク１，２に入力されているので、すなわち第１及び第２マイク１，２の受音感度が同程度であるので、このときに第１及び２パワースペクトラム計算部２１，２２で算出される第１及び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）は同程度になる。一方、音源（ユーザ）から発話があった場合には、その発話を第２マイク２だけが受音するので、すなわち第２マイク２の受音感度の方が大きくなるので、このときに第１パワースペクトル値Ｐｘ（ω）よりも第２パワースペクトル値Ｐｙ（ω）の方が大きくなる。このとき、パワー比計算部２３が算出するパワー比Ｐｘｙ（ω）は小さくなる。

【００５１】

なお、このとき、雑音源や音源（ユーザ）の特性に応じて、所定の周波数域のパワースペクトル値Ｐｘ（ω），Ｐｙ（ω）が特に変化する。

このように、音源（ユーザ）から発話があった場合にはパワー比Ｐｘｙ（ω）は小さくなることから、音声／非音声判定部２４は、パワー比Ｐｘｙ（ω）と判定用しきい値（類似度を示すしきい値）δｐ１とを比較して、音声区間を判定する。

【００５２】

ここで、第１及び第２パワースペクトラム計算部２１，２２では、パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）を所定の周波数域を対象として得ている。よって、パワー比Ｐｘｙ（ω）は、各周波数帯について得ることができる。

このようなことから、パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）について各周波数で得ているパワー比Ｐｘｙ（ω）の総和平均値を算出し、判定では、その総和平均値と判定用しきい値δｐ１とを比較する。ここで、判定用しきい値δｐ１は例えば実験により得る。

【００５３】

なお、判定対象としてパワースペクトル値Ｐｘ（ω），Ｐｙ（ω）の全周波数域の総和平均値を用いることに限定されるものではない。例えば、音源（ユーザ）の特性を示す特定の周波数帯のパワー比Ｐｘｙ（ω）の総和平均値と判定用しきい値δｐ１とを比較したり、雑音源の特性を示す特定の周波数帯のパワー比Ｐｘｙ（ω）の平均値と判定用しきい値δｐ１とを比較したり、又は音源（ユーザ）の特性を示す特定の周波数帯のパワー比Ｐｘｙ（ω）と雑音源の特性を示す特定の周波数帯のパワー比Ｐｘｙ（ω）との平均値と判定用しきい値δｐ１とを比較したりしてもよい。この場合、それに応じて、判定用しきい値δｐ１を設定する。

【００５４】

そして、音声／非音声判定部２４は、パワー比Ｐｘｙ（ω）が判定用しきい値δｐ１未満の場合（Ｐｘｙ（ω）＜δｐ１）、音声区間と判定し、それ以外の場合（Ｐｘｙ（ω）≧δｐ１）、非音声区間と判定する。ここで、音声／非音声判定部２４は、このような判定をフレーム単位で行う。そして、音声／非音声判定部２４は、その判定結果を音入力オン／オフ制御部１５に出力する。

【００５５】

以上のように第３の実施形態の音声信号処理装置１０が構成されている。この音声信号処理装置１０における一連の動作は次のようになる。

先ず、第１及び第２フレーム化部１１，１２が、第１及び第２マイク１，２から入力された２ｃｈの音声信号ｘ（ｔ），ｙ（ｔ）をそれぞれフレーム化し、フレーム単位で音声信号ｘ（ｔ），ｙ（ｔ）を第１及び第２パワースペクトラム計算部２１，２２に出力する。

【００５６】

パワースペクトラム計算部２１，２２ではそれぞれ、第１及び第２フレーム化部１１，１２それぞれから出力されるフレーム単位の音声信号ｘ（ｔ），ｙ（ｔ）について第１及

び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）を算出して、算出した第１及び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）をパワー比計算部２３に出力する。

パワー比計算部２３では、パワースペクトラム計算部２１，２２それぞれから出力される第１及び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）について、フレーム単位でパワー比Ｐｘｙ（ω）を算出して、算出したパワー比Ｐｘｙ（ω）を音声／非音声判定部２４に出力する。

【００５７】

音声／非音声判定部２４では、パワー比Ｐｘｙ（ω）と判定用しきい値δｐ１とを比較し、パワー比Ｐｘｙ（ω）に対応するフレームが音声区間のものか、非音声区間のものかを判定する。そして、音声／非音声判定部２４は、その判定結果を音入力オン／オフ制御部１５に出力する。

音入力オン／オフ制御部１５では、第２マイク２からの音声信号ｙ（ｔ）の後段への出力のオンとオフとを切り換える。具体的には、音声／非音声判定部２４が音声区間と判定した場合、音入力オン／オフ制御部１５は、オン制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力して、音声／非音声判定部２４が非音声区間と判定した場合、音入力オン／オフ制御部１５は、オフ制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力しないようにする。このとき、音入力オン／オフ制御部１５から出力される音声信号ｙ（ｔ）は、音源（ユーザ）からの音と雑音源からの音とからなる音声信号となる。

このように、第３の実施形態の音声信号処理装置１０は、第２マイク２への入力音中の発話区間（音声区間）を検出することができる。

【００５８】

次に第３の実施形態における効果を説明する。

前述したように、無指向性マイクである第２のマイク２で発話音及び雑音を受音し、単一指向性マイクである第１マイク１で雑音を受音し、第２マイク２で受音した発話音及び雑音からなる音声信号のパワースペクトルと、第１マイク１で受音した雑音の音声信号のパワースペクトルとを比較して、その比較結果に基づいて、前記発話音の発話区間を特定している。

【００５９】

これにより、第２のマイク２で発話音及び雑音を受音し、かつ第１マイク１で雑音を受音するように第１及び第２のマイク１，２を配置する限り、マイクの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系を構築することができる。

また、精度よく発話区間を検出することができる。そして、このように音声信号処理装置１０が検出した発話区間の音声信号を利用することにより、音声認識システムでは、高認識率、低誤認識率の音声認識が可能になり、また、携帯電話やトランシーバでは、信頼性の高いハンズフリー半二重通信が可能になり、放送システムでは、通信システムの送信電力低減が可能になる。

【００６０】

次に第４の実施形態を説明する。

この第４の実施形態も前述の第３の実施形態と同様、音声信号ｘ（ｔ），ｙ（ｔ）それぞれのパワースペクトルを算出し、このパワースペクトルに基づいて音声区間と非音声区間とを判定する音声信号処理装置１０である。音声信号処理装置１０の構成は、前述の第３の実施形態と同様、前記図５に示したような構成になる。そして、第４の実施形態では、第１及び第２マイク１，２の配置を前述の第３の実施形態における配置と異ならせている。

【００６１】

前述の第３の実施形態では、前述の第１の実施形態と同様、第１マイク１に単一指向性マイクを用い、第２マイク２に無指向性マイクを用い、前記図２に示したように、雑音源からの音を第１及び第２マイク１，２で受音し、音源（ユーザ）からの音を第２マイク２だけで受音するように、第１及び第２マイク１，２を配置している。

一方、この第４の実施形態では、前述の第２の実施形態と同様、音源（ユーザ）からの音を第１及び第２マイク１，２で受音し、雑音源からの音を第１マイク１だけで受音している。具体的には、第１マイク１に無指向性マイクを用い、第２マイク２に単一指向性マイクを用いる。そして、前記図４に示したように、第１及び第２マイク１，２をできるだけ近づけて配置するとともに、単一指向性マイクである第２マイク２を、その指向方向が音源（ユーザ）に向かい、かつその指向方向外に雑音源が位置されるように、配置する。

【００６２】

このように第１及び第２マイク１，２を配置した場合、パワー比計算部２３で算出されるパワー比Ｐｘｙ（ω）は次のような傾向を示す。

音源（ユーザ）からの音を第１及び第２マイク１，２で受音し、雑音源からの音を第１マイク１だけで受音するようにしているので、雑音源からの音だけを第１マイク１で受音している場合には、第１マイク１の受音感度の方が大きくなるので、第１パワースペクトル値Ｐｘ（ω）が大きくなり、これにより、パワー比Ｐｘｙ（ω）が大きくなる。一方、音源（ユーザ）から発話があった場合には、その発話を第１及び第２マイク１，２で受音するので、これにより、第２パワースペクトル値Ｐｙ（ω）も大きくなり、パワー比Ｐｘｙ（ω）が小さくなる。

【００６３】

このように、音源（ユーザ）から発話があった場合には、パワー比Ｐｘｙ（ω）が小さくなる傾向を示すようになる。

このようなことから、第４の実施形態では、音声／非音声判定部２４は、パワー比Ｐｘｙ（ω）と判定用しきい値（類似度を示すしきい値）δｐ２とを比較して、パワー比Ｐｘｙ（ω）が判定用しきい値δｐ２未満の場合（Ｐｘｙ（ω）＜δｐ２）、音声区間と判定し、それ以外の場合（Ｐｘｙ（ω）≧δｐ２）、非音声区間と判定する。ここで、判定用しきい値δｐ２は例えば実験により得る。

【００６４】

そして、前述の第３の実施形態と同様に、音入力オン／オフ制御部１５は、音声／非音声判定部２４が音声区間と判定した場合、オン制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力して、音声／非音声判定部２４が非音声区間と判定した場合、オフ制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力しないようにする。このとき、音入力オン／オフ制御部１５から出力される音声信号ｙ（ｔ）は、音源（ユーザ）からの音のみからなる音声信号となる。

【００６５】

このように、第４の実施形態の音声信号処理装置１０は、第２マイク２への入力音中の発話区間（音声区間）を検出することができる。

次に第４の実施形態における効果を説明する。

前述したように、無指向性マイクである第１のマイク１で発話音及び雑音を受音し、単一指向性マイクである第２マイク２で発話音を受音し、第１マイク１で受音した発話音及び雑音からなる音声信号のパワースペクトルと、第２マイク２で受音した発話音の音声信号のパワースペクトルとを比較して、その比較結果に基づいて、前記発話音の発話区間を特定している。

【００６６】

これにより、第１のマイク１で発話音及び雑音を受音し、かつ第２マイク２で発話音を受音するように第１及び第２のマイク１，２を配置する限り、マイクの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系を構築することができる。

また、精度よく発話区間を検出することができる。そして、このように音声信号処理装置１０が検出した発話区間の音声信号を利用することにより、音声認識システムでは、高認識率、低誤認識率の音声認識が可能になり、また、携帯電話やトランシーバでは、信頼性の高いハンズフリー半二重通信が可能になり、放送システムでは、通信システムの送信電力低減が可能になる。

【００６７】
次に第５の実施形態を説明する。
前述の第１及び第２の実施形態では、相互相関関数計算部１３により音声信号ｘ（ｔ），ｙ（ｔ）の相互相関関数Ｒｘｙ（τ）を算出し、この相互相関関数Ｒｘｙ（τ）に基づいて音声区間と非音声区間とを判定し、また、前述の第３及び第４の実施形態では、音声信号ｘ（ｔ），ｙ（ｔ）それぞれのパワースペクトルＰｘ（ω），Ｐｙ（ω）を算出し、このパワースペクトルＰｘ（ω），Ｐｙ（ω）（具体的にはパワー比Ｐｘｙ（ω））に基づいて音声区間と非音声区間とを判定している。第５の実施形態では、第１の実施形態（第２の実施形態）の処理と、第３の実施形態（第４の実施形態）の処理とを組み合わせた処理により、音声区間と非音声区間とを判定している。すなわち、第５の実施形態では、音声信号ｘ（ｔ），ｙ（ｔ）の相互相関関数Ｒｘｙ（τ）を算出するとともに、音声信号ｘ（ｔ），ｙ（ｔ）それぞれのパワースペクトルＰｘ（ω），Ｐｙ（ω）を算出し、相互相関関数Ｒｘｙ（τ）とパワースペクトルＰｘ（ω），Ｐｙ（ω）（具体的にはパワー比Ｐｘｙ（ω））との両面から音声区間と非音声区間とを判定している。図６は、それを実現する音声信号処理装置１０の構成を示す。
【００６８】
図６に示すように、音声信号処理装置１０は、第１及び第２フレーム化部１１，１２、相互相関関数計算部１３、音声／非音声判定部１４、音入力オン／オフ制御部１５、第１及び第２パワースペクトラム計算部２１，２２、パワー比計算部２３、並びに音声／非音声判定部３１を備えている。
このような構成において、第１及び第２フレーム化部１１，１２、相互相関関数計算部１３、音声／非音声判定部１４、音入力オン／オフ制御部１５、第１及び第２パワースペクトラム計算部２１，２２、並びにパワー比計算部２３は、前述の第１及び第２の実施形態と同様な処理を行う。
【００６９】
すなわち、第１及び第２フレーム化部１１，１２は、第１及び第２マイク１，２から入力された２ｃｈの音声信号ｘ（ｔ），ｙ（ｔ）をそれぞれフレーム化し、フレーム単位で音声信号ｘ（ｔ），ｙ（ｔ）を相互相関関数計算部１３に出力する。また、第１及び第２フレーム化部１１，１２はそれぞれ、フレーム単位で各音声信号ｘ（ｔ），ｙ（ｔ）を第１及び第２パワースペクトラム計算部２１，２２それぞれに出力する。
【００７０】
相互相関関数計算部１３は、第１及び第２フレーム化部１１，１２それぞれから出力されるフレーム単位の音声信号ｘ（ｔ），ｙ（ｔ）について相互相関関数Ｒｘｙ（τ）を算出して、算出した相互相関関数Ｒｘｙ（τ）を音声／非音声判定部３１に出力する。
また、第１及び第２パワースペクトラム計算部２１，２２は、第１及び第２フレーム化部１１，１２それぞれから出力されるフレーム単位の音声信号ｘ（ｔ），ｙ（ｔ）について第１及び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）を算出して、算出した第１及び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）をパワー比計算部２３に出力する。パワー比計算部２３は、パワースペクトラム計算部２１，２２それぞれから出力される第１及び第２パワースペクトル値Ｐｘ（ω），Ｐｙ（ω）について、フレーム単位でパワー比Ｐｘｙ（ω）を算出して、算出したパワー比Ｐｘｙ（ω）を音声／非音声判定部３１に出力する。
【００７１】
音声／非音声判定部３１では、次のような判定処理を行う。この音声／非音声判定部３１で行う処理が第５の実施形態において特有の処理になる。ここで、前記図２に示したように、雑音源からの音を第１及び第２マイク１，２で受音し、音源（ユーザ）からの音を第２マイク２だけで受音するように、第１及び第２マイク１，２を配置した場合（第１又は第３の実施形態の場合）と、前記図４に示したように、雑音源からの音を第１マイク１だけで受音し、音源（ユーザ）からの音を第１及び第２マイク２で受音するように、第１及び第２マイク１，２を配置した場合（第２又は第４の実施形態の場合）とを分けして説

明する。

【００７２】

先ず、前記図２に示したように、雑音源からの音を第１及び第２マイク１，２で受音し、音源（ユーザ）からの音を第２マイク２だけで受音するように、第１及び第２マイク１，２を配置した場合について説明する。

この場合、音声／非音声判定部３１は、音源（ユーザ）から発話があった場合に相互相関関数Ｒｘｙ（τ）が０に向かって減少することから、相互相関関数Ｒｘｙ（τ）が判定用しきい値δｒ１未満の場合（Ｒｘｙ（τ）＜δｒ１）、音声区間とし、それ以外の場合（Ｒｘｙ（τ）≧δｒ１）、非音声区間とする第１判定結果を得る。また、音声／非音声判定部３１は、音源（ユーザ）から発話があった場合にはパワー比Ｐｘｙ（ω）が小さくなることから、パワー比Ｐｘｙ（ω）が判定用しきい値δｐ１未満の場合（Ｐｘｙ（ω）＜δｐ１）、音声区間とし、それ以外の場合（Ｐｘｙ（ω）≧δｐ１）、非音声区間とする第２判定結果を得る。

【００７３】

そして、音声／非音声判定部３１は、前記第１及び第２判定結果に基づいて、音声区間の最終的な判定結果を得る。例えば、音声／非音声判定部３１は、第１及び第２判定結果が共に音声区間である判定結果となった場合、最終的な判定結果を音声区間とする。または、音声／非音声判定部３１は、第１判定結果又は第２判定結果の少なくとも一方が音声区間である判定結果となった場合、最終的な判定結果を音声区間とする。そして、音声／非音声判定部３１は、それ以外の場合、最終的な判定結果を非音声区間とする。

【００７４】

一方、前記図４に示したように、雑音源からの音を第１マイク１だけで受音し、音源（ユーザ）からの音を第１及び第２マイク２で受音するように、第１及び第２マイク１，２を配置した場合には、次のような判定を行う。

音声／非音声判定部３１は、音源（ユーザ）から発話があった場合に相互相関関数Ｒｘｙ（τ）が大きくなることから、相互相関関数Ｒｘｙ（τ）が判定用しきい値δｒ２より大きい場合（Ｒｘｙ（τ）≧δｒ２）、音声区間とし、それ以外の場合（Ｒｘｙ（τ）＜δｒ２）、非音声区間とする第１判定結果を得る。また、音声／非音声判定部３１は、音源（ユーザ）からの音の出力（発話）があった場合にはパワー比Ｐｘｙ（ω）が小さくなることから、パワー比Ｐｘｙ（ω）が判定用しきい値δｐ２未満の場合（Ｐｘｙ（ω）＜δｐ２）、音声区間とし、それ以外の場合（Ｐｘｙ（ω）≧δｐ２）、非音声区間とする第２判定結果を得る。

【００７５】

そして、音声／非音声判定部３１は、前記第１及び第２判定結果に基づいて、音声区間の最終的な判定結果を得る。例えば、音声／非音声判定部３１は、第１及び第２判定結果が共に音声区間である判定結果となった場合、最終的な判定結果を音声区間とする。または、音声／非音声判定部３１は、第１判定結果又は第２判定結果の少なくとも一方が音声区間である判定結果となった場合、最終的な判定結果を音声区間とする。そして、音声／非音声判定部３１は、それ以外の場合、最終的な判定結果を非音声区間とする。

【００７６】

以上のようにして音声／非音声判定部３１は、音声区間又は非音声区間を判定している。そして、音声／非音声判定部３１は、その判定結果を音入力オン／オフ制御部１５に出力する。

音入力オン／オフ制御部１５は、音声／非音声判定部３１が音声区間と判定した場合、オン制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力して、音声／非音声判定部３１が非音声区間と判定した場合、オフ制御として第２マイク２からの音声信号ｙ（ｔ）を後段に出力しないようにする。このとき、第１及び第２のマイク１，２の配置が前記図２に示した配置であれば、音入力オン／オフ制御部１５から出力される音声信号ｙ（ｔ）は、音源（ユーザ）からの音と雑音源からの音とからなる音声信号となる。また、第１及び第２のマイク１，２の配置が前記図４に示した配置であれば、音入力オン／オフ制

御部１５から出力される音声信号ｙ（ｔ）は、音源（ユーザ）からの音のみからなる音声信号となる。

【００７７】

次に第５の実施形態における効果を説明する。

前述したように、無指向性マイクで発話音及び雑音を受音し、単一指向性マイクで発話音又は雑音のいずれか一方を受音し、無指向性マイクで受音した発話音及び雑音からなる音声信号と単一指向性マイクで受音した発話音又は雑音のいずれか一方の音声信号の比較により相関度を得ている。その一方で、無指向性マイクで受音した発話音及び雑音からなる音声信号のパワースペクトルと、単一指向性マイクで受音した発話音又は雑音のいずれか一方の音声信号のパワースペクトルとを比較して、その比較結果としてパワー比を得ている。そして、前記相関度とパワー比との両方に基づいて、最終的に前記発話音の発話区間を特定している。

【００７８】

このように、相関度とパワー比との両方に基づいて最終的に前記発話音の発話区間を特定することで、その特定を精度よく行うことができる。そして、このように音声信号処理装置１０が検出した発話区間の音声信号を利用することにより、音声認識システムでは、高認識率、低誤認識率の音声認識が可能になり、また、携帯電話やトランシーバでは、信頼性の高いハンズフリー半二重通信が可能になり、放送システムでは、通信システムの送信電力低減が可能になる。

【００７９】

また、前述の第１乃至第４の実施形態と同様に、無指向性マイクで発話音及び雑音を受音し、かつ単一指向性マイクで発話音又は雑音のいずれか一方を受音するように第１及び第２のマイク１，２を配置する限り、マイクの取り付け位置の変化等による環境の変化、話者の移動や姿勢の変化等による音源の移動に対してロバストな受音系を構築することができる。

【００８０】

なお、前述の実施形態では、第１及び第２マイク１，２から入力された音声信号ｘ１（ｔ），ｘ２（ｔ）を、直接第１及び第２フレーム化部１１，１２にそれぞれ入力しているが、具体的には、第１及び第２マイク１，２から入力された音声信号ｘ１（ｔ），ｘ２（ｔ）を、ＡＤ（アナログ／デジタル）変換した後、第１及び第２フレーム化部１１，１２に入力するようにする。また、前述の実施形態では、第２マイク２に入力された音声信号ｘ１（ｔ）を、音入力オン／オフ制御部１５に入力しているが、第２マイク２に入力され、フレーム化した音声信号ｘ１（ｔ）を音入力オン／オフ制御部１５に入力する。これらの仕様を、例えば前述の第１の実施形態の音声信号処理装置１０の構成に適用すると、図７に示すような構成になる。

【００８１】

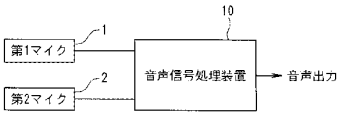
この図７に示すように、第１及び第２マイク１，２から入力された音声信号ｘ１（ｔ），ｘ２（ｔ）をそれぞれ、第１及び第２ＡＤ変換部４１，４２でＡＤ変換した後、第１及び第２フレーム化部１１，１２に入力する。また、第２ＡＤ変換部３２でＡＤ変換された信号は、第２フレーム化部１２でフレーム化されてから音入力オン／オフ制御部１５に入力される。ここで、第１及び第２ＡＤ変換部４１，４２でＡＤ変換されたデータ形式は、例えば１１０２５Ｈｚ、１６ｂｉｔ、リニアＰＣＭである。また、第１及び第２フレーム化部１１，１２でフレーム化された信号のフレーム長は、例えば５１２サンプルフレーム長である。

【００８２】

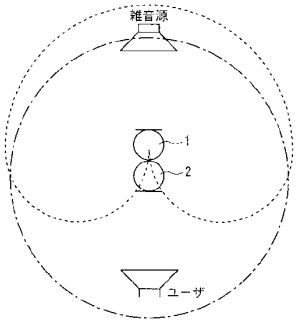
例えば、音声信号ｘ２（ｔ）を第２フレーム化部１２でフレーム化してから音入力オン／オフ制御部１５に出力することで、結果的に、音声信号処理装置１０から出力される音声信号ｘ２（ｔ）もフレーム化されているものとなり、これにより、音声信号処理装置１０から出力される音声信号ｘ２（ｔ）を利用する音声アプリケーションでは、解りやすいフレーム化された音声信号ｘ２（ｔ）で処理をすることができるようになる。

【００８３】
また、前述の実施形態では、検出対象音が人間が発する発話音である場合を説明したが、検出対象音は、人間以外の物体が発する音でもよい。
また、前述の実施形態の説明において、相互相関関数計算部１３又はパワースペクトラム計算部２１，２２及びパワー比計算部２３は、無指向性マイクに入力された音信号と、単一指向性マイクに入力された音信号とを比較する比較手段を実現しており、音声／非音声判定部１４，２４，３１は、比較手段の比較結果に基づいて、検出対象音を検出する検出対象音検出手段又は発話音の発話区間を検出する発話区間検出手段を実現している。
【００８４】
また、前述の実施形態の音声信号処理装置１０を音声認識装置に適用することができる。この場合、音声認識装置は、前述したような音声信号処理装置１０の構成に加えて、音声信号処理装置１０が検出した発話区間の音声信号について音声認識処理をする音声認識処理手段を備える。
ここで、音声認識技術としては、例えば、旭化成株式会社が提供する音声認識技術「ＶＯＲＥＲＯ」（商標）（http://www.asahi-kasei.co.jp/vorero/jp/vorero/feature.html参照）等があり、このような音声認識技術の用いた音声認識装置に適用することもできる。
【００８５】
また、前述の実施形態の音声信号処理装置１０をコンピュータで実現することができる。そして、前述したような音声信号処理装置１０の処理内容をコンピュータが所定のプログラムにより実現する。この場合、プログラムは、無指向性マイクで受音した発話音及び雑音の音声信号と単一指向性マイクで受音した前記発話音又は前記雑音のいずれか一方の音声信号とを比較し、その比較結果に基づいて、前記発話音の発話区間を検出する処理をコンピュータに実行させるプログラムになる。
【図面の簡単な説明】
【００８６】
【図１】本発明の実施形態の音声信号処理装置を含むシステム全体の構成を示すブロック図である。
【図２】本発明の第１の実施形態におけるマイクの配置を示す図である。
【図３】本発明の第１の実施形態の音声信号処理装置の構成を示すブロック図である。
【図４】本発明の第２の実施形態におけるマイクの配置を示す図である。
【図５】本発明の第３の実施形態の音声信号処理装置の構成を示すブロック図である。
【図６】本発明の第５の実施形態の音声信号処理装置の構成を示すブロック図である。
【図７】前記第１の実施形態の他の構成例を示すブロック図である。
【符号の説明】
【００８７】
１，２　マイク
１０　音声信号処理装置
１１，１２　フレーム化部
１３　相互相関関数計算部
１４，２４，３１　音声／非音声判定部
１５　音入力オン／オフ制御部
２１，２２　パワースペクトラム計算部
２３　パワー比計算部

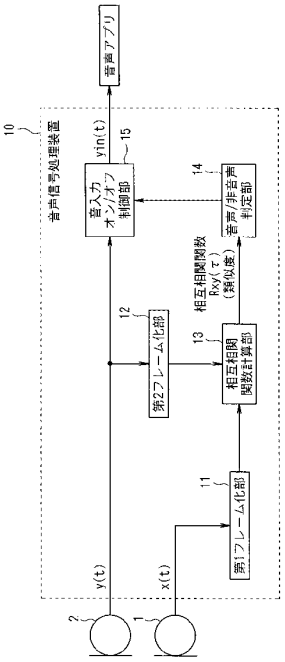
【図１】


第1マイク
第2マイク
音声信号処理装置
音声出力
1
2
10

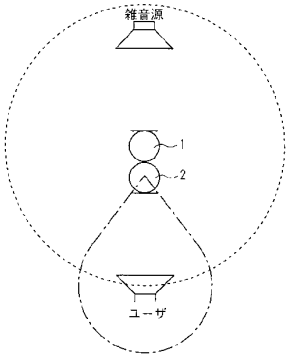
【図２】


雑音源
ユーザ
1
2

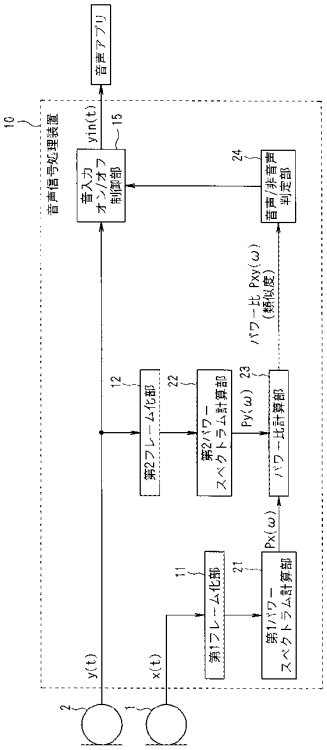
【図３】


音声アプリ
音声信号処理装置
音入力オンオフ制御部
yin(t)
15
14
音声/非音声判定部
相互相関関数
Rxy(τ)
(類似度)
12
第2フレーム化部
13
相互相関関数計算部
11
第1フレーム化部
y(t)
x(t)
2
1

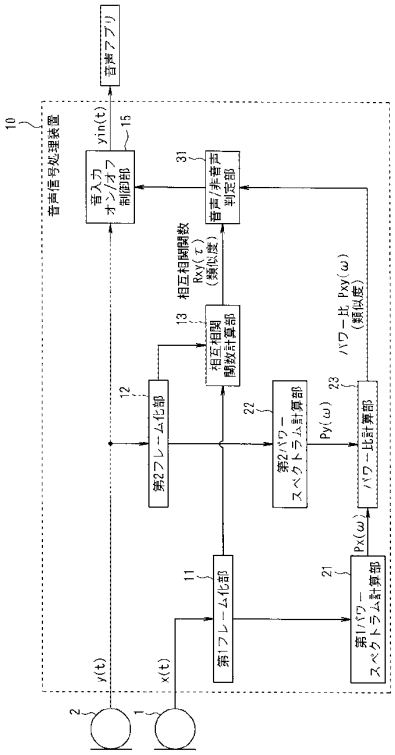
【図４】


雑音源
ユーザ
1
2

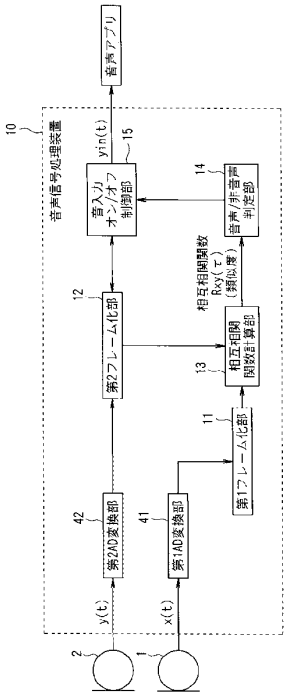
【図５】


音声アプリ
音声信号処理装置
音入力オンオフ制御部
yin(t)
15
24
音声/非音声判定部
パワー比 Pxy(ω)
(類似度)
12
第2フレーム化部
22
第2パワースペクトラム計算部
Py(ω)
23
パワー比計算部
Px(ω)
11
第1フレーム化部
21
第1パワースペクトラム計算部
y(t)
x(t)
2
1

【図６】


音声アプリ
音声信号処理装置
yin(t)
15
音入力
オン/オフ
制御部
31
音声/非音声
判定部
相互相関関数
Rxy(τ)
(類似度)
13
相互相関
関数計算部
パワー比 Pxy(ω)
(類似度)
12
第2フレーム化部
22
第2パワー
スペクトラム計算部
Py(ω)
23
パワー比計算部
11
第1フレーム化部
21
第1パワー
スペクトラム計算部
Px(ω)
y(t)
x(t)
2
1

【図７】


音声アプリ
音声信号処理装置
yin(t)
15
音入力
オン/オフ
制御部
14
音声/非音声
判定部
相互相関関数
Rxy(τ)
(類似度)
12
第2フレーム化部
13
相互相関
関数計算部
11
第1フレーム化部
42
第2AD変換部
41
第1AD変換部
y(t)
x(t)
2
1

フロントページの続き

(51)Int.Cl.[7]	F I	テーマコード（参考）
G10L 21/02		

(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2004-289762
(P2004-289762A)
(43)公開日　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）
H04R 3/00	H04R	3/00	320	5D015
G10L 11/00	H04R	1/40	320Z	5D020
G10L 15/20	G10L	3/02	301E	
G10L 15/28	G10L	3/02	301F	
G10L 21/02	G10L	3/02	301G	

審査請求 未請求　請求項の数 39　OL　（全 50 頁）　最終頁に続く

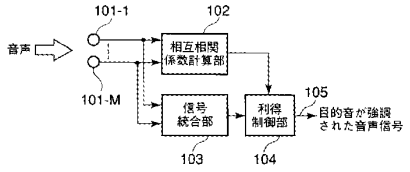
(21)出願番号　特願2003-119116(P2003-119116)
(22)出願日　平成15年4月23日(2003.4.23)
(31)優先権主張番号　特願2003-20996(P2003-20996)
(32)優先日　平成15年1月29日(2003.1.29)
(33)優先権主張国　日本国(JP)

(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
(71)出願人　503041281
永田　仁史
岩手県盛岡市青山４－１－２－４０３
(74)代理人　100058479
弁理士　鈴江　武彦
(74)代理人　100091351
弁理士　河野　哲
(74)代理人　100088683
弁理士　中村　誠
(74)代理人　100108855
弁理士　蔵田　昌俊

最終頁に続く

(54)【発明の名称】音声信号処理方法と装置及びプログラム

(57)【要約】
【課題】突発雑音を含む実環境雑音下で雑音を抑圧して目的音の成分を強調することができる音声信号処理装置を提供する。
【解決手段】空間的に離れて配置された複数のマイクロホン１０１－１～１０１－Ｍから出力される複数チャネルの入力音声信号間の相互相関係数を相互相関係数計算部１０２により算出し、信号統合部１０３により複数チャネルの入力音声信号を１チャネルに統合して得られる統合音声信号の大きさを利得制御部１０４によって相互相関係数に従って調整することにより、目的音の成分が出力音声信号１０５を生成する。
【選択図】　図１



【特許請求の範囲】
【請求項１】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号のチャネル間の相互相関係数を求めるステップと、
前記入力音声信号を１チャネルに統合して統合音声信号を出力する統合ステップと、
前記統合音声信号の大きさを前記相互相関係数に従って調整することにより出力音声信号を生成するステップとを具備する音声信号処理方法。
【請求項２】
前記統合ステップは、時間領域で動作する適応ビームフォーマを用いて前記入力音声信号を１チャネルに統合する請求項１に記載の音声信号処理方法。
【請求項３】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成するステップと、
前記複数チャネルのスペクトル情報のチャネル間の相互相関係数を求めるステップと、
前記スペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合ステップと、
前記統合スペクトル信号の大きさを前記相互相関係数に従って調整するステップとを具備する音声信号処理方法。
【請求項４】
前記統合ステップは、周波数領域で動作する適応ビームフォーマを用いて前記入力音声信号を１チャネルに統合する請求項３に記載の音声信号処理方法。
【請求項５】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成するステップと、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求めるステップと、
前記パワースペクトル及びクロススペクトルから計算される重み関数を用いて周波数毎に重みを掛けて前記コヒーレンス関数を用いて前記パワースペクトル及びクロススペクトルを修正するステップと、
修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を求めるステップとを具備する音声信号処理方法。
【請求項６】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成するステップと、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求めるステップと、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を求めるステップと、
前記コヒーレンス関数を用いて前記パワースペクトル及びクロススペクトルを修正するステップと、
修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を求めるステップとを具備する音声信号処理方法。
【請求項７】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成するステップと、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求めるステップと、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を求めるステップと、
前記パワースペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパ

ワー情報を求めるステップと、
前記コヒーレンス関数及びパワー情報を用いて前記パワースペクトル及びクロススペクトルを修正するステップと、
修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を求めるステップとを具備する音声信号処理方法。
【請求項８】
前記相互相関係数に対して、予め定めた閾値を用いて閾値処理を行うことにより前記マイクロホンに目的音が到来しているか否かを判定するステップをさらに具備する請求項５、６または７のいずれか１項に記載の音声信号処理方法。
【請求項９】
前記スペクトル情報を１チャネルに統合して統合スペクトル信号を求めるステップと、前記統合スペクトル信号の大きさを前記相互相関係数に従って調整するステップとをさらに具備する請求項５、６または７のいずれか１項に記載の信号処理方法。
【請求項１０】
前記周波数分析を高速フーリエ変換処理及び帯域フィルタバンク処理の少なくとも一方により行う請求項３、５、６または７のいずれか１項に記載の音声信号処理方法。
【請求項１１】
前記コヒーレンス関数及び前記パワー情報の少なくとも一方に従って前記統合スペクトル信号の各周波数成分に対して重み付けを行うステップをさらに具備する請求項６または７に記載の音声信号処理方法。
【請求項１２】
前記相互相関係数に従って前記複数チャネルのスペクトル情報の位相及び振幅の少なくとも一方をチャネル間で一致するように補正するステップをさらに具備する請求項６または７のいずれか１項に記載の音声信号処理方法。
【請求項１３】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号のチャネル間の相互相関係数を算出する相互相関係数計算部と、
前記入力音声信号を１チャネルに統合して統合音声信号を出力する統合部と、
前記統合音声信号の大きさを前記相互相関係数に従って調整することにより出力音声信号を生成する調整部とを具備する音声信号処理装置。
【請求項１４】
前記統合部は，時間領域で動作する適応ビームフォーマを含む請求項１３に記載の音声信号処理装置。
【請求項１５】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記複数チャネルのスペクトル情報のチャネル間の相互相関係数を算出する相互相関係数計算部と、
前記スペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合部と、
前記統合スペクトル信号の大きさを前記相互相関係数に従って調整する調整部とを具備する音声信号処理装置。
【請求項１６】
前記統合部は、周波数領域で動作する適応ビームフォーマを用いて前記入力音声信号を１チャネルに統合する請求項１３に記載の音声信号処理装置。
【請求項１７】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチ

ャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、
前記コヒーレンス関数を用いて前記パワースペクトル及びクロススペクトルを修正する修正スペクトル計算部と、
修正された前記パワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を算出する重み付き相互相関係数計算部とを具備する音声信号処理装置。
【請求項１８】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、
前記パワースペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出するパワー情報計算部と、
前記コヒーレンス関数及びパワー情報を用いて前記パワースペクトル及びクロススペクトルを修正する修正スペクトル計算部と、
修正された前記パワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を算出する重み付き相互相関係数計算部とを具備する音声信号処理装置。
【請求項１９】
前記相互相関係数に対して、予め定めた閾値を用いて閾値処理を行うことにより前記マイクロホンに目的音が到来しているか否かを判定する判定部をさらに具備する請求項１７または１８に記載の音声信号処理装置。
【請求項２０】
前記スペクトル情報を１チャネルに統合して統合スペクトル信号を求める統合部と、前記統合スペクトル信号の大きさを前記相互相関係数に従って調整する調整部とをさらに具備する請求項１７または１８に記載の音声信号処理装置。
【請求項２１】
前記周波数分析部は、高速フーリエ変換器及び帯域フィルタバンクの少なくとも一方である請求項１５、１７または１８のいずれか１項に記載の音声信号処理装置。
【請求項２２】
前記コヒーレンス関数及び前記パワー情報の少なくとも一方に従って前記統合スペクトル信号の各周波数成分に対して重み付けを行う手段をさらに具備する請求項１７または１８に記載の信号処理装置。
【請求項２３】
前記相互相関係数に従って前記複数チャネルのスペクトル情報の位相及び振幅の少なくとも一方をチャネル間で一致するように補正する手段をさらに具備する請求項１５、１７または１８のいずれか１項に記載の音声信号処理装置。
【請求項２４】
前記複数のマイクロホンは、少なくとも一つの無指向性マイクロホンと少なくとも一つの指向性マイクロホンを含む請求項１５、１７または１８のいずれか１項に記載の音声信号処理装置。
【請求項２５】
前記複数のマイクロホンは、指向性の軸の向きを異ならせた少なくとも二つの指向性マイクロホンを含む請求項１５、１７または１８のいずれか１項に記載の音声信号処理装置。
【請求項２６】
前記少なくとも二つの指向性マイクロホンは、前記指向性の軸が同一平面内に存在せず、且つ前記指向性の軸と目的音の到来方向とのなす角が一致するように配置される請求項２５に記載の音声信号処理装置。

【請求項２７】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号のチャネル間の相互相関係数を求める処理と、
前記入力音声信号を１チャネルに統合して統合音声信号を出力する処理と、
前記統合音声信号の大きさを前記相互相関係数に従って調整することにより出力音声信号を生成する処理とをコンピュータに行わせるプログラム。
【請求項２８】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、
前記複数チャネルのスペクトル情報のチャネル間の相互相関係数を求める処理と、
前記スペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合処理と、
前記統合スペクトル信号の大きさを前記相互相関係数に従って調整する処理とをコンピュータに行わせるためのプログラム。
【請求項２９】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求める処理と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を求める処理と、
前記コヒーレンス関数を用いて前記パワースペクトル及びクロススペクトルを修正する処理と、
修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を求める処理とをコンピュータに行わせるためのプログラム。
【請求項３０】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求める処理と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を求める処理と、
前記パワースペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパワー情報を求める処理と、
前記コヒーレンス関数及びパワー情報を用いて前記パワースペクトル及びクロススペクトルを修正する処理と、
修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、前記入力音声信号のチャネル間の相互相関係数を求める処理とをコンピュータに行わせるためのプログラム。
【請求項３１】
空間的に離れて配置された複数のマイクロホンに入力される音声に応答して該マイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、
音声の複数の仮想到来方向からなる仮想到来方向群に対応して、該仮想到来方向から到来する音声が複数のチャネル間で一致するように補正するための補正係数を発生する補正係数発生部と、

前記補正係数に基づいて前記パワースペクトル及びクロススペクトルを補正し、補正パワースペクトル及び補正クロススペクトルを生成するスペクトル補正部と、
前記補正パワースペクトル及び補正クロススペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出するパワー情報計算部と、
前記補正パワースペクトル及び補正クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けし、前記仮想到来方向群に対応した前記入力音声信号のチャネル間の相互相関係数を前記仮想到来方向毎に算出する相関係数計算部と、
前記相互相関係数に基づいて前記マイクロホンに入力される音声の音源方向を検出すると共に、検出した該音源方向における前記相互相関係数の値を音源相関係数として出力する音源方向検出部とを具備する音声信号処理装置。

【請求項３２】
前記音源方向に基づいて前記複数チャネルのスペクトル情報を補正するスペクトル情報補正部と、
補正された複数チャネルのスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合部と、
前記コヒーレンス関数を用いて前記統合スペクトル信号をフィルタリングするコヒーレンスフィルタ演算部と、
フィルタリングされた統合スペクトル信号の大きさを前記音源相関係数に基づいて調整することにより出力音声信号を生成する調整部とをさらに具備する請求項３１記載の音声信号処理装置。

【請求項３３】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、
前記パワースペクトルから前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出するパワー情報計算部と、
前記複数のスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合部と、
前記統合スペクトル信号のパワースペクトルを計算する統合信号パワースペクトル計算部と、
前記クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けし、重み付けたクロススペクトルをさらに前記統合スペクトル信号のパワースペクトルに基づいて正規化して利得係数を計算する利得係数計算部とを具備する音声信号処理装置。

【請求項３４】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、
前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、
前記複数のスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合部と、
前記統合したスペクトル情報のパワースペクトルを計算する統合信号パワースペクトル計算部と、
前記クロススペクトルを前記コヒーレンス関数に基づいて重み付けし、重み付けたクロススペクトルをさらに前記統合信号パワースペクトルに基づいて正規化して利得係数を計算する利得係数計算部とを具備する音声信号処理装置。

【請求項３５】
前記利得係数に基づき前記統合スペクトル信号の大きさを調整する調整部と、前記調整部から出力される統合スペクトル信号を前記コヒーレンス関数に基づいてフィルタリングして出力音声信号を出力するコヒーレンスフィルタ演算部とをさらに具備する請求項３４または３５に記載の音声信号処理装置。
【請求項３６】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、
前記複数チャネル間のクロススペクトルと各チャネルのパワースペクトルから前記複数チャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、
音声の複数の仮想到来方向からなる仮想到来方向群に対応して、該仮想到来方向から到来する音声が複数のチャネル間で一致するように補正するための補正係数を発生する補正係数発生部と、
前記補正係数に基づいて前記パワースペクトル及びクロススペクトルを補正し、補正パワースペクトル及び補正クロススペクトルを生成するスペクトル補正部と、
前記補正パワースペクトル及び補正クロススペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出するパワー情報計算部と、
前記複数チャネルのスペクトル情報を前記補正係数により補正してから統合して得られる統合スペクトル情報に対するパワースペクトルを前記補正パワースペクトル及び補正クロススペクトルに基づいて計算する仮想統合パワースペクトル計算部と、
前記補正クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けし、さらに仮想統合パワースペクトルに基づいて正規化することにより、前記仮想到来方向に対応した利得係数を求める利得係数計算部と、
前記利得係数に基づいて前記マイクロホンに入力される音声の音源方向を検出すると共に、検出した該音源方向に対応する利得係数の値を音源利得係数として出力する音源方向検出部とを具備する音声信号処理装置。
【請求項３７】
前記補正係数に基づいて前記複数チャネルのスペクトル情報を補正するスペクトル情報補正部と、
前記補正された複数チャネルのスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する信号統合部と、
前記コヒーレンス関数を用いて前記統合スペクトル信号をフィルタリングするコヒーレンスフィルタ演算部と、
フィルタリングされた統合スペクトル信号の大きさを前記音源相関係数に基づいて調整する調整部とをさらに具備する請求項３６記載の音声信号処理装置。
【請求項３８】
空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部と、
前記複数チャネルのスペクトル情報を入力として、前記複数チャネルの入力音声信号のチャネル間の第１の修正相互相関係数を計算する第１の修正相互相関係数計算部と、
前記第１の修正相互相関係数に基づいて前記複数チャネルのスペクトル情報のチャネル間の差を適応的に補正して補正スペクトル情報を生成する適応スペクトル補正部と、
前記補正スペクトル情報から第２の修正相互相関係数を計算する第２の修正相互相関係数計算部とを具備し、
前記第１及び第２の修正相互相関係数計算部は、（ａ）前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、（ｂ）前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出するコヒーレンス関数計

算部と、（ｃ）前記パワースペクトルから前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出するパワー情報計算部と、（ｄ）前記パワースペクトル及びクロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けして前記入力音声信号のチャネル間の相互相関係数を算出し、前記第１または第２の修正相互相関関数を出力する相関係数計算部とを有する音声信号処理装置。

【請求項３９】

空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルの第１スペクトル情報を生成する周波数分析部と、

前記第１スペクトル情報から第１の修正利得を計算する第１の修正利得係数計算部と、

前記第１の利得係数に基づいて前記第１スペクトル情報のチャネル間の差を適応的に補正して第２スペクトル情報を生成する適応スペクトル補正部と、

前記第２スペクトル情報から第２の修正利得を計算する第２の修正利得係数計算部とを具備し、

前記第１及び第２の修正利得係数計算部は，（ａ）前記第１または第２スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出するスペクトル計算部と、（ｂ）前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出するコヒーレンス関数計算部と、（ｃ）前記パワースペクトルから前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出するパワー情報計算部と、（ｄ）前記複数のスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合部と、（ｅ）前記統合スペクトル信号のパワースペクトルを計算する統合信号パワースペクトル計算部と、（ｆ）前記クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けし、重み付けたクロススペクトルをさらに前記統合スペクトル信号のパワースペクトルに基づいて正規化して前記第１または第２の利得係数を計算する利得係数計算部とを有する音声信号処理装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、複数のマイクロホンによって得られる入力音声信号を処理する音声信号処理方法と装置及びプログラムに関する。より詳しくは、本発明は例えばハンズフリー通話や音声認識等において用いられる雑音抑圧技術の一つとして入力音声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。

【０００２】

【従来の技術】

音声信号処理の分野では、音声認識や携帯電話の実用化に伴い、雑音対策が重要な問題となってきている。雑音抑圧技術としては、一つのマイクロホンを使う場合に用いられる，例えば雑音の定常性を仮定したスペクトルサブトラクション処理と、複数のマイクロホンを用いるマイクロホンアレイ処理がある。マイクロホンアレイ処理には、少数のマイクロホンでも高い雑音抑圧能力を発揮する適応型マイクロホンアレイがコスト面から有望である。適応マイクロホンアレイは，雑音方向に受音感度の低い死角を自動的に向けることにより雑音を抑圧するものであり、適応ビームフォーマ（適応ＢＦ）と呼ばれることもある。

【０００３】

適応ビームフォーマは，方向性の強い雑音に対しては効果的であるが，その他の雑音、例えば（１）　車で走行中に発生する雑音のような高レベルの拡散性雑音、（２）　高速で移動する車からの放射音のように音響伝達系の変化が速い雑音、あるいは（３）　突発雑音のような継続時間が非常に短い雑音、等に関しては抑圧性能が十分ではない。これらのような雑音は実環境ではごく普通に存在するため、対処が必要である。

【０００４】

非特許文献１には、複数のマイクロホンからの入力音声信号の２チャネル間のコヒーレン

ス関数に基づいたフィルタリングを行って雑音を抑圧する技術が示されている。
一方、非特許文献２では、相関の大きい雑音に対処するため、目的音のない区間でチャネル間の雑音のクロススペクトルを推定しておき、目的音のある区間で雑音の重畳した目的音のクロススペクトルから雑音のクロススペクトルを引き去る技術が開示されている。
【０００５】
非特許文献３には，例えば複数チャネルの信号間の相互相関を用いて信号検出処理を行うために、コヒーレンス関数を閾値処理することによって目的信号の存在を判別する方法が示されている。
非特許文献４には、複数のマイクロホンから出力される複数チャネルの音声信号間の相互相関係数を閾値処理することにより、目的音を検出する方法が開示されている。
非特許文献５には、適応ビームフォーマを用いて２以上のチャネルの音声信号を１チャネルに統合する方法が記載されている。
非特許文献６には、重み関数を用いて複数チャネルの音声信号のチャネル間の一般化相互相関関数（ｇｅｎｅｒａｌｉｚｅｄ ｃｒｏｓｓ ｃｏｒｒｅｌａｔｉｏｎ ｆｕｎｃｔｉｏｎ）を最尤推定する方法が開示されている。
【０００６】
【非特許文献１】
“Ｕｓｉｎｇ ｔｈｅ ｃｏｈｅｒｅｎｃｅ ｆｕｎｃｔｉｏｎ ｆｏｒ ｎｏｉｓｅ ｒｅｄｕｃｔｉｏｎ”，ＩＥＥ Ｐｒｏｃｅｅｄｉｎｇｓ－Ｉ Ｖｏｌ．１３９，Ｎｏ．３，１９９２
【０００７】
【非特許文献２】
“Ｅｎｈａｎｃｅｍｅｎｔ ｏｆ ｓｐｅｅｃｈ ｄｅｇｒａｄｅｄ ｂｙ ｃｏｈｅｒｅｎｔ ａｎｄ ｉｎｃｏｈｅｒｅｎｔ ｎｏｉｓｅ ｕｓｉｎｇ ａ ｃｒｏｓｓ－ｓｐｅｃｔｒａｌ ｅｓｔｉｍａｔｏｒ”，ＩＥＥＥ Ｔｒａｎｓ．ｏｎ Ｓｅｅｃｈ ａｎｄ Ａｕｄｉｏ ｐｒｏｃｅｓｓｉｎｇ，Ｖｏｌ．５，Ｎｏ．５，１９９７
【０００８】
【非特許文献３】
“Ｋｎｏｗｉｎｇ ｔｈｅ Ｗｈｅａｔ ｆｒｏｍ ｔｈｅ Ｗｅｅｄｓ ｉｎ Ｎｏｉｓｙ Ｓｐｅｅｃｈ“，Ｈ．Ａｇａｉｂｙ ａｎｄ Ｔ．Ｊ．Ｍｏｉｒ 著，Ｐｒｏｃ．ｏｆ ＥＵＲＯＳＰＥＥＣＨ’９７，ｖｏｌ．３，ｐｐ．１１１－１１２，１９９７
【０００９】
【非特許文献４】「２つの指向性マイクロホンを用いた目的音検出に関する検討」、永田他、電子情報通信学会誌 Ｖｏｌ．Ｊ８３－Ａ Ｎｏ．２ （２０００））
【００１０】
【非特許文献５】
“Ｔｈｅ ａｄａｐｔｉｖｅ ｆｉｌｔｅｒ ｔｈｅｏｒｙ”，Ｈｙａｋｉｎ 著，ＰＲＥＮＴＩＣＥ ＨＡＬＬ 出版
【００１１】
【非特許文献６】
“Ｔｈｅ Ｇｅｎｅｒａｌｉｚｅｄ Ｃｏｒｒｅｌａｔｉｏｎ Ｍｅｔｈｏｄ ｆｏｒ Ｅｓｔｉｍａｔｉｏｎ ｏｆ Ｔｉｍｅ Ｄｅｌａｙ”，Ｃ．Ｈ．Ｋｎａｐｐ ａｎｄ Ｇ．Ｃ．Ｃａｒｔｅｒ 著，ＩＥＥＥ Ｔｒａｎｓ，Ａｃｏｕｓｔ．，Ｓｐｅｅｃｈ，Ｓｉｇｎａｌ Ｐｒｏｃｅｓｓｉｎｇ，Ｖｏｌ．ＡＳＳＰ－２４，Ｎｏ．４，ｐｐ．３２０－３２７，１９７６
【００１２】
【発明が解決しようとする課題】
非特許文献１に記載された技術は，（１）の拡散性雑音のようにチャネル間で無相関であると仮定できる雑音に対しては有効である。しかし、（３）の突発性雑音や、ビームフォ

ーマによって抑圧できた方向性のある雑音は、チャネル間の相関が大きくなるため、抑圧することができない。非特許文献２に記載の技術によると、このようなチャネル間の相関が大きい雑音を抑圧できる。しかし、この方法が有効なのは、雑音に方向性があり、かつ、雑音の定常性が仮定できる場合に限られる。このような雑音環境では、むしろビームフォーマのように指向性の死角を雑音源に向ける手法の方がよりよく対処できる。

【００１３】

本発明は、突発雑音を含む実環境雑音下で雑音を抑圧して目的音の成分を強調することができる音声信号処理方法と装置及びプログラムを提供することを目的とする。

【００１４】

本発明の他の目的は、目的音が到来しているか否かの検出を高精度で行うことを目的とする。

【００１５】

【課題を解決するための手段】

上記の課題を解決するため、本発明の第１の観点によると、空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号間の相互相関係数を求める。入力音声信号を１チャネルに統合して得られる統合音声信号の大きさを相互相関係数に従って調整することにより、目的音の成分が強調された出力音声信号を生成する。

【００１６】

本発明の第２の観点では、各マイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成し、複数チャネルのスペクトル情報間の相互相関係数を求める。スペクトル情報を１チャネルに統合して得られる統合スペクトル信号の大きさを相互相関係数に従って調整することにより、目的音の成分が強調されたスペクトル信号を得る。

【００１７】

本発明の第３の観点では、各マイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成し、これらのスペクトル情報から入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求める。さらに、パワースペクトル及びクロススペクトルから各チャネルのスペクトル情報間のコヒーレンス関数を求める。次に、コヒーレンス関数を用いてパワースペクトル及びクロススペクトルを修正し、修正後のパワースペクトル及びクロススペクトルに基づいて重み付けられた、入力音声信号のチャネル間の相互相関係数を求める。

【００１８】

本発明の第４の観点では、各マイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成し、これらのスペクトル情報から入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求める。さらに、パワースペクトル及びクロススペクトルから各チャネルのスペクトル情報間のコヒーレンス関数を求め、またスペクトル情報から入力音声信号のチャネル間の信号パワーに関するパワー情報を求める。次に、コヒーレンス関数及びパワー情報を用いてパワースペクトル及びクロススペクトルを修正し、修正後のパワースペクトル及びクロススペクトルに基づいて重み付けられた、入力音声信号のチャネル間の相互相関係数を求める。

【００１９】

第３または第４の観点において、相互相関係数に対して予め定めた閾値を用いて閾値処理を行うことによりマイクロホンに目的音が到来しているか否かを判定してもよい。スペクトル情報を１チャネルに統合して統合スペクトル信号を求め、この統合スペクトル信号の大きさを相互相関係数に従って調整してもよい。コヒーレンス関数に従って、統合スペクトル信号の各周波数成分に対して重み付けを行ってもよい。相互相関係数に従って、複数チャネルのスペクトル情報の位相及び振幅の少なくとも一方をチャネル間で一致するように補正してもよい。

【００２０】

第３及び第４の観点において、複数のマイクロホンは、少なくとも一つの無指向性マイク

ロホンと少なくとも一つの指向性マイクロホンを含んでもよいし、指向性の軸の向きを異ならせた少なくとも二つの指向性マイクロホンを含んでもよい。後者の場合、少なくとも二つの指向性マイクロホンは、指向性の軸が同一平面内に存在せず、且つ指向性の軸と目的音の到来方向とのなす角が一致するように配置されることが好ましい。

【００２１】

さらに本発明の別の観点によると、上述した音声信号処理をコンピュータで実行するための以下のようなプログラムあるいは該プログラムを記憶した記憶媒体を提供する。

【００２２】

（１）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号間の相互相関係数を求める処理と、入力音声信号を１チャネルに統合して統合音声信号を出力する処理と、統合音声信号の大きさを相互相関係数に従って調整することにより出力音声信号を生成する処理とをコンピュータに行わせるプログラム。

【００２３】

（２）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、複数チャネルのスペクトル情報間の相互相関係数を求める処理と、スペクトル情報を１チャネルに統合して統合スペクトル信号を生成する統合処理と、統合スペクトル信号の大きさを相互相関係数に従って調整する処理とをコンピュータに行わせるためのプログラム。

【００２４】

（３）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、スペクトル情報から入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求める処理と、パワースペクトル及びクロススペクトルから複数チャネルのスペクトル情報間のコヒーレンス関数を求める処理と、コヒーレンス関数を用いてパワースペクトル及びクロススペクトルを修正する処理と、修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、入力音声信号のチャネル間の相互相関係数を求める処理とをコンピュータに行わせるためのプログラム。

【００２５】

（４）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、スペクトル情報から入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを求める処理と、パワースペクトル及びクロススペクトルから複数チャネルのスペクトル情報間のコヒーレンス関数を求める処理と、スペクトル情報に基づいて入力音声信号のチャネル間の信号パワーに関するパワー情報を求める処理と、コヒーレンス関数及びパワー情報を用いてパワースペクトル及びクロススペクトルを修正する処理と、修正されたパワースペクトル及びクロススペクトルに基づいて重み付けられた、入力音声信号のチャネル間の相互相関係数を求める処理とをコンピュータに行わせるためのプログラム。

【００２６】

（５）空間的に離れて配置された複数のマイクロホンに入力される音声に応答して該マイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する処理と、前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出する処理と、音声の複数の仮想到来方向からなる仮想到来方向群に対応して、該仮想到来方向から到来する音声が複数のチャネル間で一致するように補正するための補正係数を発生する処理と、前記補正係数に基づいて前記パワースペクトル及びクロススペクトルを補正し、補正パワースペクトル及び補正クロススペクトルを生成する処理と、前記補正パワースペクトル及び補正クロススペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出する処理と、前記補正パワースペクトル及び補正クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み

付け し、前記仮想到来方向群に対応した前記入力音声信号のチャネル間の相互相関係数を前記仮想到来方向毎に算出する処理と、前記相互相関係数に基づいて前記マイクロホンに入力される音声の音源方向を検出すると共に、検出した該音源方向における前記相互相関係数の値を音源相関係数として出力する処理とをコンピュータに行わせるためのプログラム。

【００２７】

（６）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する処理と、前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出する処理と、前記複数のスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する処理と、前記統合スペクトル信号のパワースペクトルを計算する処理と、前記クロススペクトルを前記コヒーレンス関数に基づいて重み付けし、重み付けたクロススペクトルをさらに前記統合信号パワースペクトルに基づいて正規化して利得係数を計算する処理とをコンピュータに行わせるためのプログラム。

【００２８】

（７）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する処理と、前記複数チャネル間のクロススペクトルと各チャネルのパワースペクトルから前記複数チャネル間のコヒーレンス関数を算出する処理と、音声の複数の仮想到来方向からなる仮想到来方向群に対応して、該仮想到来方向から到来する音声が複数のチャネル間で一致するように補正するための補正係数を発生する処理と、前記補正係数に基づいて前記パワースペクトル及びクロススペクトルを補正し、補正パワースペクトル及び補正クロススペクトルを生成する処理と、前記補正パワースペクトル及び補正クロススペクトルに基づいて前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出する処理と、前記複数チャネルのスペクトル情報を前記補正係数により補正してから統合して得られる統合スペクトル情報に対するパワースペクトルを前記補正パワースペクトル及び補正クロススペクトルに基づいて計算する処理と、前記補正クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けし、さらに仮想統合パワースペクトルに基づいて正規化することにより、前記仮想到来方向に対応した利得係数を求める処理と、前記利得係数に基づいて前記マイクロホンに入力される音声の音源方向を検出すると共に、検出した該音源方向に対応する利得係数の値を音源利得係数として出力する処理とをコンピュータに行わせるためのプログラム。

【００２９】

（８）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する処理と、前記複数チャネルのスペクトル情報を入力として、前記複数チャネルの入力音声信号のチャネル間の第１の修正相互相関係数を計算する処理と、前記第１の修正相互相関係数に基づいて前記複数チャネルのスペクトル情報のチャネル間の差を適応的に補正して補正スペクトル情報を生成する処理と、前記補正スペクトル情報から第２の修正相互相関係数を計算する処理とをコンピュータに実行させるためのプログラムであって、前記第１及び第２の修正相互相関係数の計算処理は、（ａ）前記スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する処理と、（ｂ）前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出する処理と、（ｃ）前記パワースペクトルから前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出する処理と、（ｄ）前記パワースペクトル及びクロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けして前記入力音声信号のチャネル間の相互相関係数を算出し、前記第１または第２の修

正相互相関関数を出力する処理とを含む。

【００３０】

（９）空間的に離れて配置された複数のマイクロホンから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルの第１スペクトル情報を生成する処理と、前記第１スペクトル情報から第１の修正利得を計算する処理と、前記第１の利得係数に基づいて前記第１スペクトル情報のチャネル間の差を適応的に補正して第２スペクトル情報を生成する処理と、前記第２スペクトル情報から第２の修正利得を計算する処理とをコンピュータに行わせるためのプログラムであって、前記第１及び第２の修正利得係数の計算処理は、（ａ）前記第１または第２スペクトル情報から前記入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する処理と、（ｂ）前記パワースペクトル及びクロススペクトルから前記複数チャネルのスペクトル情報のチャネル間のコヒーレンス関数を算出する処理と、（ｃ）前記パワースペクトルから前記入力音声信号のチャネル間の信号パワーに関するパワー情報を算出する処理と、（ｄ）前記複数のスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する処理と、（ｅ）前記統合スペクトル信号のパワースペクトルを計算する処理と、（ｆ）前記クロススペクトルを前記コヒーレンス関数及びパワー情報に基づいて重み付けし、重み付けたクロススペクトルをさらに前記統合スペクトル信号のパワースペクトルに基づいて正規化して前記第１または第２の利得係数を計算する処理とを含む。

【００３１】

【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。以下に説明する各実施形態における音声信号処理は、コンピュータ上で実行されるソフトウェア（ファームウェアを含む）として実現することが可能であり、またハードウェアによって実現することも可能である。

【００３２】

（第１の実施形態）

図１に、本発明の第１の実施形態に係る信号処理装置の構成を示す。複数のマイクロホン１０１－１～１０１－Ｍによって、例えば話者の入力音声のような目的音を含む音響信号が検出され、複数（Ｍ）チャネルの入力音声信号が出力される。ここで、雑音を抑圧して入力音声のうち最終的に出力音声として取り出したい成分を目的音という。マイクロホン１０１－１～１０１－Ｍからの入力音声信号は、図示しないＡ／Ｄ変換器によりディジタル信号に変換された後、相互相関計算部１０２と信号統合部１０３に入力される。

【００３３】

相互相関計算部１０２では、Ｍチャネルの入力音声信号間の相互相関係数が計算される。信号統合部１０３では、Ｍチャネルの入力音声信号が１チャネルに統合される。信号統合部１０３から出力される信号を統合音声信号という。統合音声信号は、相互相関係数に従って利得が制御される利得制御部１０４に入力され、その大きさが調整される。これにより、利得制御部１０４から目的音の成分が強調された出力音声信号１０５が出力される。

【００３４】

一般に、複数チャネルの観測信号に対して計算される相互相関係数は、雑音下の目的信号の検出尺度としてソナーやレーダの処理において古くから使われている。本実施形態では、音声信号処理において目的音の検出だけでなく、目的音の強調に用いる方法を提案する。この方法により、チャネル間で無相関な雑音のある環境下でも、雑音を効果的に抑圧できる。

【００３５】

本実施形態でいう相互相関係数とは、入力音声信号がｘ（ｎ），ｙ（ｎ）の２チャネルの場合、次式で計算される値ρである。

【００３６】

【数１】

$$\rho = \frac{\overline{x(n)y(n)}}{\sqrt{\overline{x(n)x(n)}}\sqrt{\overline{y(n)y(n)}}} \qquad (1-1)$$

【００３７】
ここで、上線が付された値は、期待値または時間平均値を表す（以後、同様とする）。
【００３８】
入力音声信号がＭチャネルの場合（２チャネルに限定されない場合）には、相互相関係数ρは例えば次式で計算される。
【００３９】
【数２】

$$\rho = \frac{\sum_{p=1,q=1,p<q}^{p\le M,q\le M} \overline{xp(n)xq(n)}}{\sqrt[2K]{\prod_{p=1,q=1,p<q}^{p\le M,q\le M} \overline{xp(n)xp(n)xq(n)xq(n)}}} \qquad (1-2)$$

【００４０】
ここで、ｘｐ（ｎ），ｘｑ（ｎ）はそれぞれ第ｐチャネル，第ｑチャネルの入力音声信号であり、またＫ＝Ｍ（Ｍ－１）／２である。
【００４１】
従来、複数チャネルの信号のチャネル間相互相関は信号検出処理に用いられ、例えば、コヒーレンス関数を閾値処理することによって目的信号の存在を判別する方法が例えば、非特許文献３：”Ｋｎｏｗｉｎｇ ｔｈｅ Ｗｈｅａｔ ｆｒｏｍ ｔｈｅ Ｗｅｅｄｓ ｉｎ Ｎｏｉｓｙ Ｓｐｅｅｃｈ”，Ｈ．Ａｇａｉｂｙ ａｎｄ Ｔ．Ｊ．Ｍｏｉｒ 著，Ｐｒｏｃ．ｏｆ ＥＵＲＯＳＰＥＥＣＨ’９７，ｖｏｌ．３，ｐｐ．１１１－１１２，１９９７に開示されている。
【００４２】
相互相関係数は音声検出にも使われており、この値を閾値処理して目的音を検出する方法は、例えば非特許文献４：「２つの指向性マイクロホンを用いた目的音検出に関する検討」、永田他、電子情報通信学会誌 Ｖｏｌ．Ｊ８３－Ａ Ｎｏ．２ （２０００））に開示されている。本実施形態は、閾値処理による目的音の検出ではなく、相互相関を目的音の強調に用いる点が特徴である。
【００４３】
相互相関係数ρは、入力音声に目的音が存在する場合は１に近い値をとり、雑音だけなら０に近い値となるので、音声強調に用いるには相互相関係数の大きさに従って統合音声信号に与える利得を制御すればよい。すなわち、マイクロホン１０１－１～１０１－Ｍから得られる複数チャネルの入力音声信号について、相互相関係数計算部１０２により式（１－１）または（１－２）に従ってチャネル間の相互相関係数を計算する。この相互相関係数に基づき利得制御部１０４の利得を制御し，信号統合部１０３からの統合音声信号の振幅を利得制御部１０４で調整することによって出力音声信号１０５を生成する。
【００４４】
相互相関係数ρは、－１から＋１の範囲をとる。従って、利得制御部１０４では相互相関係数の絶対値をとってから用いるか、あるいは相互相関係数が負の場合は０と置くようにする。利得制御部１０４での利得制御は、こうして計算された相互相関係数を例えば統合音声信号の振幅に乗算することにより行われる。この場合、相互相関係数と利得の関係を

図２に示す直線（Ａ）のような比例関係に設定してもよいし、例えば図２の折れ線（Ｂ）や曲線（Ｃ）のような関係にしてもよい。

【００４５】

次に、図３を用いて本実施形態における処理の流れを説明する。

まず、マイクロホン１０１－１～１０１－Ｍから音声信号を入力する（ステップＳ１１）。マイクロホンが二つの場合を例にとると、例えば図４に示すように二つのマイクロホン１０１－１～１０１－２を１０ｃｍ程度の距離を置いて、目的音源が各マイクロホン１０１－１～１０１－２から等距離となるように設置する。マイクロホン１０１－１～１０１－２の各々は指向性があってもよいし、無指向性でもよい。入力音声信号をディジタル化するＡ／Ｄ変換器のサンプリング周波数は例えば１１ｋＨｚとするが、他の周波数でもかまわない。

【００４６】

次に、相互相関係数ρを式（１－１）または式（１－２）によって計算する。このとき、相互相関係数ρの時間変化を考慮して、適当な時間間隔、例えば、Ｎ＝１２８点おきに相互相関係数ρを求めることとし、時間平均を例えば対象時点の前後Ｌ点、計２Ｌ点の波形に対して式（１－１）を適用すると、相互相関係数ρを求める数式は以下となる。

【００４７】

【数３】

$$\rho(k)=\frac{\frac{1}{2L}\sum_{i=-L}^{L-1}x(kN+i)y(kN+i)}{\sqrt{\frac{1}{2L}\sum_{i=-L}^{L-1}x(kN+i)x(kN+i)\frac{1}{2L}\sum_{i=-L}^{L-1}y(kN+i)y(kN+i)}} \tag{1-3}$$

【００４８】

ここで、ｋは相互相関係数の番号であり、ρの値は入力音声信号波形のＮサンプル毎にｉ個求まる。

【００４９】

式（１－２）を用いた場合も同様に、次式によって相関係数ρが求まる。

【００５０】

【数４】

$$\rho=\frac{\sum_{p=1,q=1,p<q}^{p\le M,q\le M}\frac{1}{2L}\sum_{i=-L}^{L-1}xp(kN+i)xp(kN+i)}{\sqrt[2K]{\prod_{p=1,q=1,p<q}^{p\le M,q\le M}\frac{1}{2L}\sum_{i=-L}^{L-1}xp(kN+i)^2\frac{1}{2L}\sum_{i=-L}^{L-1}xq(kN+i)^2}} \tag{1-4}$$

【００５１】

ここで、Ｋ＝Ｍ（Ｍ－１）／２である。

【００５２】

次に、信号統合部１０３によって複数チャネルの入力音声信号を１チャネルに統合する。信号統合部１０３の処理は、例えば単純な加算であってもよいし、図５に示すように雑音抑圧の機能を持つ、時間領域で動作する適応ビームフォーマ１０６による処理であってもよい。信号統合部１０３が単純な加算を行うとすると、統合音声信号ｚ（ｎ）は、次式のように求まる。

【００５３】

【数５】

$$z(n)=\sum_{m=1}^{M}xm(n) \qquad (1-5)$$

【００５４】

信号統合部１０３に図５のように適応ビームフォーマ１０６、例えば、よく知られているＬＭＳ適応フィルタによる２チャネルのＪｉｍ－Ｇｒｉｆｆｉｔｈビームフォーマを用いた場合、次式のように統合音声信号ｚ（ｎ）が求まる。

【００５５】

【数６】

$$\begin{aligned}U(n)&=[x(n)-y(n),x(x(n-1)-y(n-1),\ldots,x(n-T+1)-y(n-T+1)]\\ d(n)&=x(n)+y(n)\\ z(n)&=d(n-D)-U(n)*W(n-1)\\ W(n)&=W(n-1)-2\mu z(n)U(n)/|U(n)\bullet U(n)|\end{aligned} \qquad (1-6)$$

【００５６】

ここで、Ｕ（ｎ）は入力音声信号ｘ，ｙの差の値をＴ個並べたベクトル、Ｗ（ｎ）＝［ｗ１（ｎ），ｗ２（ｎ），．．．，ｗＴ（ｎ）］はｎ回更新を行った後のＬＭＳ適応フィルタの係数、ｄ（ｎ）は入力音声信号ｘ，ｙの和信号、（・）は内積である。Ｄは遅延量であり、例えばＴ／２を用いる。μはステップサイズであり、例えば０．１を用いればよい。Ｍチャネルの場合への拡張も容易であり、Ｍ－１個の適応ビームフォーマを用いて１チャネルに統合された音声信号を得る方法が、例えば非特許文献５：“Ｔｈｅ　ａｄａｐｔｉｖｅ　ｆｉｌｔｅｒ　ｔｈｅｏｒｙ”，Ｈｙａｋｉｎ　著，ＰＲＥＮＴＩＣＥ　ＨＡＬＬ　出版に詳述されているが、ここでは詳しい説明を省略する。

【００５７】

最後に、統合音声信号ｚ（ｎ）に相互相関係数ρに基づく利得を乗じて統合音声信号ｚ（ｎ）の大きさを調整することにより、出力音声信号１０５を出力する。ステップＳ１１～Ｓ１４の処理は、ディジタル化された音声信号がステップＳ１１においてフレーム単位で入力される毎に繰り返し行われる。

【００５８】

このように本実施形態によれば、複数チャネルの入力音声信号が１チャネルに統合された統合音声信号の大きさを各チャネルの入力音声信号間の相互相関関数に従って調整することにより、相関の少ない雑音が抑圧され、相関の大きい目的音の成分が強調された出力音声信号を得ることが可能となる。

【００５９】

（第２の実施形態）

図６に、本発明の第２の実施形態に係る音声信号処理装置の構成を示す。本実施形態では、第１の実施形態で説明した時間領域での音声信号処理と等価な音声信号処理を周波数領域で実現する。図６において、複数のマイクロホン１０１－１～１０１－Ｍからの入力音声信号は図示しないＡ／Ｄ変換器によりディジタル信号に変換された後、周波数分析部２０１により周波数成分が分析され、周波数スペクトルを表すスペクトル情報が生成される。周波数分析部２０１は、例えば公知のＦＦＴ（高速フーリエ変換）、ＤＦＴ（離散フーリエ変換）、あるいは通過帯域の異なる複数の帯域フィルタを並列に配置した帯域フィルタバンクによって実現される。周波数分析部２０１から出力されるスペクトル情報は、相関係数計算部２０２と信号統合部２０３に入力される。

【００６０】

相互相関計算部２０２では、Ｍチャネルのスペクトル情報間の相互相関係数、すなわち周波数領域の相互相関係数が計算される。言い換えれば、本実施形態ではスペクトル情報を用いてＭチャネルの入力音声信号のチャネル間の相互相関係数が求められる。信号統合部２０３では、Ｍチャネルのスペクトル情報が１チャネルに統合される。信号統合部２０３の処理は、第１の実施形態で説明したと同様、例えば単純な加算であってもよいし、周波数領域で動作する適応フィルタを用いたＪｉｍ－Ｇｒｉｆｆｉｔｈの適応ビームフォーマによる処理であってもよい。信号統合部２０３から出力される信号を統合スペクトル信号という。

【００６１】

信号統合部２０３から出力される統合スペクトル信号は、相互相関係数に従って利得が制御される利得制御部２０４に入力され、その大きさが調整される。これにより、利得制御部２０４から目的音の成分が強調されたスペクトル信号２０５が出力される。第１の実施形態と同様に、相互相関係数計算部２０２により得られる周波数領域の相互相関係数も、目的音が存在する場合は１に近い値をとり、雑音だけなら０に近い値となるので、目的音の強調に用いるには相互相関係数の大きさに従って統合スペクトル信号に与える利得を制御すればよい。

【００６２】

目的音の成分が強調されたスペクトル信号２０５は、必要に応じて逆変換部２０６によって周波数分析部２０１とは逆の変換、すなわち周波数領域から時間領域への変換が施されることにより、目的音の成分が強調された出力音声信号２０７が生成される。逆変換部２０６は、周波数分析部２０１が例えばＦＦＴの場合、その逆変換である逆ＦＦＴによって実現される。

【００６３】

相互相関係数計算部２０２では、入力音声信号がｘ（ｎ），ｙ（ｎ）の２チャネルの場合、周波数領域での相互相関係数として、次式で表されるρが計算される。

【００６４】

【数７】

$$\rho = \frac{\sum_{f=1}^{L} Wxy(f)}{\sqrt{\sum_{f=1}^{L} Wxx(f) \sum_{f=1}^{L} Wyy(f)}} \qquad (2\text{-}1)$$

【００６５】

ここで、Ｗｘｙ（ｆ）は入力音声信号ｘ（ｎ），ｙ（ｎ）間のクロススペクトルであり，Ｗｘｘ（ｆ），Ｗｙｙ（ｆ）は入力音声信号ｘ（ｎ），ｙ（ｎ）のパワースペクトル、Ｌは離散フーリエ変換（ＤＦＴ）における周波数成分の数である。

【００６６】

クロススペクトルとパワースペクトルは、よく知られているように、ｘ（ｎ）の離散フーリエ変換をＸ（ｆ）とし、ｙ（ｎ）の離散フーリエ変換をＹ（ｆ）とすると、

【数８】

$$\begin{aligned} Wxy(f) &= \overline{X(f)Y(f)*} \\ Wxx(f) &= \overline{X(f)X(f)*} \\ Wyy(f) &= \overline{Y(f)Y(f)*} \end{aligned} \qquad (2\text{-}2)$$

のように計算できる。ここで，上線を付した値は時間平均値、＊は複素共役である。ＤＦＴの長さは例えば２５６点を使うことができ、この場合Ｌ＝２５６である。Ｌ＝１２８と

して、得られた複素数の相互相関係数の実部をとっても等価な結果が得られる。

【００６７】

入力音声信号がＭチャネルの場合（２チャネルに限定されない場合）にも、同様に相互相関係数ρは例えば次式で計算される。

【数９】

$$\rho = \frac{\sum_{f=0}^{L-1} \sum_{i=1,j=1,i<j}^{i\leq M,j\leq M} Wij(f)}{2K\sqrt{\prod_{i=1,j=1,i<j}^{i\leq M,j\leq M} \sum_{f=0}^{L-1} Wii(f) \sum_{f=0}^{L-1} Wjj(f)}} \qquad (2-3)$$

【００６８】

ここで、Ｗｉｊ（ｆ）は入力音声信号ｘｉ（ｎ），ｘｊ（ｎ）間のクロススペクトル、Ｗｉｉ（ｆ）、Ｗｊｊ（ｆ）は入力音声信号ｘｉ（ｎ），ｘｊ（ｎ）のパワースペクトルである。

【００６９】

このようにマイクロホン１０１－１～１０ｉ－Ｍから得られる複数チャネルの入力音声信号を周波数分析部２０１でスペクトル情報に変換した後、相互相関係数計算部２０２により式（２－１）または（２－２）に従ってチャネル間の相互相関係数ρを計算する。

【００７０】

一方、周波数分析部２０１で得られる複数チャネルのスペクトル情報を信号統合部２０３で１チャネルに統合して統合スペクトル信号Ｚ（ｆ）を求める。信号統合部２０３で単純な加算を用いる場合は、

【数１０】

$$Z(f) = X(f) + Y(f) \qquad (2-4)$$

として、統合スペクトル信号Ｚ（ｆ）を得ることができる。

【００７１】

適応ビームフォーマを用いる場合は、例えば、よく知られている２チャネルのＪｉｍ－Ｇｒｉｆｆｉｔｈビームフォーマを使った場合、次式のように統合スペクトル信号Ｚ（ｆ）が求まる。

【数１１】

$$\begin{aligned} U(k,f) &= X(k,f) - Y(k,f) \\ D(k,f) &= X(k,f) + Y(k,f) \\ Z(k,f) &= D(k,f) - U(k,f)W(k-1,F) \\ W(k,f) &= X(k-1,f) - 2\mu Z(k,f)U(k,f)*/\left|U(k,f)U(k,f)*\right| \end{aligned} \qquad (2-5)$$

【００７２】

ここで、ｋはフレーム番号、Ｕはチャネル間の差分スペクトル、Ｄは加算スペクトル、Ｚは出力スペクトル、Ｗは複素数のフィルタ係数、μはステップサイズ、（＊）は複素共役である。

【００７３】

次に、相互相関係数ρに基づき利得制御部２０４の利得を制御し、信号統合部２０３からの統合スペクトル信号の大きさ（振幅）を利得制御部２０４で調整することによって、目的音の成分が強調されたスペクトル信号２０５を生成する。利得制御部２０４での利得制

御に関しては，例えば相互相関係数ρを統合スペクトル信号の振幅に乗算することにより行うことができるが、第１の実施形態と同様に例えば図２（Ａ）（Ｂ）（Ｃ）に示すような関数を用いて行うことも可能である。相互相関係数ρは負になる場合もあるが、その場合は、絶対値かまたは０と置いて利得制御に用いることも可能である。

【００７４】

図７に、本実施形態における処理の流れを示す。音声信号入力ステップＳ２１の後に周波数分析ステップＳ２２が加わったこと以外、処理の流れは第１の実施形態と基本的に同様である。すなわち、ステップＳ２２で周波数分析（例えば，ＦＦＴ）を行った後に，相互相関係数の計算（ステップＳ２３）、スペクトル情報の統合（ステップＳ２４）及び相関係数による統合スペクトル信号に対する利得制御（ステップＳ２５）を順次行って、目的音の成分が強調されたスペクトル信号を生成し、最後に必要に応じてステップＳ２６で逆変換（例えば、逆ＦＦＴ）を行って目的音の成分が強調された出力音声信号を得る。ステップＳ２１～Ｓ２６の処理は、ディジタル化された音声信号がステップＳ２１においてフレーム単位で入力される毎に繰り返し行われる。

【００７５】

このように本実施形態によれば、相関の少ない雑音が抑圧され、相関の大きい目的音の音声が強調されたスペクトル信号あるいは出力音声信号を得ることが可能となる他、相関係数の計算と信号統合の処理を周波数領域で行うことにより、相関係数の計算と信号統合の処理を時間領域で行う第１の実施形態に比較して、演算量を少なくできるという利点がある。

【００７６】

（第３の実施形態）

図８に、本発明の第３の実施形態に係る音声信号処理装置の構成を示す。本実施形態は、重み付き相互相関係数を用いて目的信号（目的音の信号）の活性度を算出する手法を提供する。こうして算出される目的信号活性度は、例えば目的音の検出や目的音の強調に有効に用いられる。

【００７７】

本実施形態では、第１の実施形態と同様に、まず複数のマイクロホン１０１－１～１０１－Ｍからの複数チャネルの入力音声信号が周波数分析部２０１により周波数領域の信号、すなわち複数の周波数成分を含むスペクトル情報に変換された後、目的信号活性度計算部３００に入力される。目的信号活性度計算部３００は、クロス・パワースペクトル計算部３０１、コヒーレンス関数計算部３０２、パワー情報計算部３０３、修正スペクトル計算部３０４及び重み付き相互相関関数計算部３０５を有する。

【００７８】

クロス・パワースペクトル計算部３０１では、複数チャネルの周波数成分から各チャネルのパワースペクトルとチャネル間のクロススペクトルが算出される。コヒーレンス関数計算部３０２では、パワースペクトルとクロススペクトルからコヒーレンス関数が算出される。パワー計算部３０３では、パワースペクトルから入力音声信号のチャネル間の信号パワーに関するパワー情報が算出される。修正スペクトル計算部３０４では、パワースペクトルとクロススペクトルに対してコヒーレンス関数とパワー情報を用いて修正が加えられる。重み付き相互相関関数計算部３０４では、修正スペクトル計算部３０４で修正されたスペクトルに従って重み付けられた相互相関係数が目的信号活性度として計算される。

【００７９】

次に、図９を用いて本実施形態における処理の流れを説明する。音声信号入力ステップＳ３１から周波数分析ステップＳ３２までは第２の実施形態と同様であり、複数チャネルの入力音声信号をフレーム単位で周波数領域の信号（スペクトル情報）に変換する。

【００８０】

次に、周波数分析で得られたスペクトル情報から各チャネルのパワースペクトルとチャネル間のクロススペクトルを計算する（ステップＳ３３）。次に、パワースペクトルとチャネル間のクロススペクトルを用いてコヒーレンス関数とパワー情報を計算する（ステップ

Ｓ３４～Ｓ３５）次に、コヒーレンス関数とパワー情報に基づいて修正されたスペクトルを計算する（ステップＳ３６）。この修正後のスペクトルに基づいて重み付き相互相関係数を計算し、これを目的信号活性度として出力する（ステップＳ３７）。ステップＳ３１～Ｓ３７の処理は、ディジタル化された音声信号がステップＳ３１においてフレーム単位で入力される毎に繰り返し行われる。

【００８１】

本実施形態は、耐雑音性を高めるように相互相関係数に修正を加える点が特徴である。一般的な相互相関係数は、雑音がチャネル間で無相関な場合の目的音検出には高い性能を示すものの、チャネル間で相関のある雑音が到来している場合と目的音が到来している場合を区別する性能は低い。本実施形態によると、相関のある雑音が到来する場合においても、目的音と雑音とを区別する性能を大幅に高めることができる。

【００８２】

通常、耳障りな大振幅の雑音はチャネル間で高い相関があるので、本実施形態で示す方法は、これを抑圧するのに好適である。出力である目的信号活性度は、入力音声に目的音が存在するか否かの尺度を示すものであり、これは以降の実施形態の音声検出や音声強調などで必要となる必須の要素である。

【００８３】

次に、クロス・パワースペクトル計算部３０１、コヒーレンス関数計算部３０２、パワー情報計算部３０３、修正スペクトル計算部３０４及び重み付き相互相関係数計算部３０４での具体的な計算方法について述べる。まず、クロス・パワースペクトル計算部３０１では、式（２－２）に従ってチャネル間のクロススペクトルとチャネル毎のパワースペクトルが計算される。次に、コヒーレンス関数計算部３０２では、入力音声信号がx，ｙの２チャネルの場合、次式に従ってコヒーレンス関数γ（ｆ）が計算される。

【数１２】

$$\gamma(f)=\frac{|Wxy(f)|}{\sqrt{Wxx(f)Wyy(f)}} \qquad (3\text{-}1)$$

【００８４】

ここで、Ｗｘｙ（ｆ）は二つのチャネル間のクロススペクトル、Ｗｘｘ（ｆ）及びＷｙｙ（ｆ）は各チャネルのパワースペクトルである。

【００８５】

入力音声信号がＭチャネルの場合（２チャネルに限定されない場合）、第ｉチャネルと第ｊチャネル間のコヒーレンス関数γｉｊ（ｆ）は、同様に次式に従って計算される。

【数１３】

$$\gamma ij(f)=\frac{|Wij(f)|}{\sqrt{Wii(f)Wjj(f)}} \qquad (3\text{-}2)$$

【００８６】

ここで、Ｗｉｊ（ｆ）は第ｉチャネルと第ｊチャネル間のクロススペクトル、Ｗｉｉ（ｆ）及びＷｊｊ（ｆ）は、第ｉチャネル及び第ｊチャネルのパワースペクトルである。

【００８７】

Ｍチャネルの場合のトータルのコヒーレンス関数γｍ（ｆ）は、例えば次式のように計算される。

【数１４】

$$\gamma m(f)=\frac{1}{K}\sum_{i=1,j=1,i<j}^{i\leq M,j\leq M}|\gamma ij(f)| \qquad (3\text{-}3)$$

【００８８】
パワー情報計算部３０３では、入力音声信号がｘ，ｙの２チャネルの場合、次式に従ってパワー情報ｐ（ｆ）が計算される。
【数１５】

$$p(f)=\frac{\min\left[\left|Wxx(f)\right|,\ \left|Wyy(f)\right|\right]}{\max\left[\left|Wxx(f)\right|,\ \left|Wyy(f)\right|\right]} \qquad (3-4)$$

【００８９】
ここで、ｍｉｎ［ａ，ｂ］はａ，ｂのうちで小さい方を選択することを意味し、ｍａｘ［ａ，ｂ］は、ａ，ｂのうちで大きい方を選択することを意味する。
【００９０】
一方、入力音声信号がＭチャネルの場合（２チャネルに限定されない場合）、第ｉチャネルと第ｊチャネル間のパワー情報はｐｉｊ（ｆ）は、次式に従って計算される。
【数１６】

$$pij(f)=\frac{\min\left[\left|Wii(f)\right|,\ \left|Wjj(f)\right|\right]}{\max\left[\left|Wii(f)\right|,\ \left|Wjj(f)\right|\right]} \qquad (3-5)$$

【００９１】
このようにして計算されるパワー情報ｐ（ｆ），ｐｉｊ（ｆ）に対して、次式のように適当な関数を用いて実際のチャネル間のパワー比に対する鋭敏性を調整することも可能である。
【数１７】

$$p(f)=\mathrm{pow}\left\{\frac{\min\left[\left|Wxx(f)\right|,\ \left|Wyy(f)\right|\right]}{\max\left[\left|Wxx(f)\right|,\ \left|Wyy(f)\right|\right]},\beta\right\} \qquad (3-6)$$

$$pij(f)=\mathrm{pow}\left\{\frac{\min\left[\left|Wii(f)\right|,\ \left|Wjj(f)\right|\right]}{\max\left[\left|Wii(f)\right|,\ \left|Wjj(f)\right|\right]},\beta\right\} \qquad (3-7)$$

【００９２】
ここで、ｐｏｗ｛ａ，ｂ｝はａのｂ乗を表す指数関数である。$\beta=1$のとき、式（３－６），（３－７）はそれぞれ式（３－４），（３－５）と同じであり、βを１より大きい値とすることで、パワー比に対する鋭敏性を増すことが可能である。
【００９３】
修正スペクトル計算部３０４では、入力音声信号が２チャネルの場合、各チャネルのパワースペクトルとチャネル間のクロススペクトルに対して、先に算出されたコヒーレンス関数ｙ（ｆ）を２乗した値である２乗コヒーレンス関数y^2（ｆ）とパワー情報ｐ（ｆ）を用いて修正が加えられたクロススペクトル及びパワースペクトルが計算される。さらに、重み付き相互相関係数計算部３０５では、修正後のクロススペクトル及びパワースペクトルに従って重み付けられた重み付き相互相関係数ρ（目的信号活性度）が計算される。
【００９４】
これら修正スペクトル計算部３０４及び重み付き相互相関係数計算部３０５での計算は、次式で示される。
【数１８】

$$\Psi a(f) = \frac{\gamma^2(f)}{(1-\gamma^2(f))|Wxy(f)|} \qquad (3-8)$$

$$\Psi b(f) = \frac{\gamma^2(f)p(f)}{(1-\gamma^2(f))|Wxy(f)|} \qquad (3-9)$$

$$\rho = \frac{\sum_{f=1}^{L} Wxy(f)\,\Psi b(f)}{\sqrt{\sum_{f=1}^{L} Wxx(f)\Psi a(f)\sum_{f=1}^{L} Wyy(f)\Psi a(f)}} \qquad (3-10)$$

【００９５】
ここで、Ψａ（ｆ），Ψｂ（ｆ）は、各々相互相関係数の計算式（３－１０）の分母、分子に用いる重み関数であり、Ｗｘｙ（ｆ）Ψｂ（ｆ）は修正後のクロススペクトル、Ｗｘｘ（ｆ）Ψａ（ｆ），Ｗｙｙ（ｆ）Ψａ（ｆ）は修正後のパワースペクトルである。
【００９６】
なお、コヒーレンス関数を用いた式（３－８）または（３－９）の重み関数以外にも、単純なクロススペクトル白色化の重み１／｜Ｗｘｙ（ｆ）｜を用いて
【数１９】

$$\Psi a(f) = \Psi b(f) = \frac{1}{|Wxy(f)|} \qquad (3-11)$$

などとすることも可能であるが、性能としては式（３－８）または（３－９）の重み関数を使うことが望ましい。
【００９７】
一方、入力音声信号がＭチャネルの場合（２チャネルに限定されない場合）、同様に各チャネルのパワースペクトルとチャネル間のクロススペクトルに対して、先に算出された第ｉチャネルと第ｊチャネル間のコヒーレンス関数ｙｉｊ（ｆ）を２乗した値である２乗コヒーレンス関数y_{ij}^2（ｆ）とパワー情報ｐｉｊ（ｆ）を用いて修正が加えられたクロススペクトル及びパワースペクトルが計算される。
【００９８】
さらに、重み付き相互相関係数計算部３０５では、修正後のクロススペクトル及びパワースペクトルに従って重み付けられた重み付き相互相関係数ρ（目的信号活性度）が計算される。この場合の修正スペクトル計算部３０４及び重み付き相互相関係数計算部３０５での計算は、次式で示される。
【数２０】

$$\Psi aij(f) = \frac{\gamma^2 ij(f)}{(1-\gamma^2 ij(f))\left|Wij(f)\right|} \qquad (3\text{-}11)$$

$$\Psi bij(f) = \frac{\gamma^2 ij(f)pij(f)}{(1-\gamma^2 ij(f))\left|Wij(f)\right|} \qquad (3\text{-}12)$$

$$\rho = \frac{\sum_{f=1}^{L} \sum_{i=1,j=1,i<j}^{i\leq M,j\leq M} Wij(f)\ \Psi bij(f)}{2N\sqrt{\prod_{i=1,j=1,i<j}^{i\leq M,j\leq M} \sum_{f=0}^{L-1} Wii(f)\Psi' aij(f) \sum_{f=0}^{L-1} Wjj(f)\Psi' aij(f)}} \qquad (3\text{-}13)$$

【００９９】
ここで、Ψａｉｊ（ｆ），Ψｂｉｊ（ｆ）は、各々相互相関係数の計算式（３－１３）の分母、分子に用いる重み関数であり、ｉ，ｊはチャネルの番号を表す。また，ｐｉｊ（ｆ）は式（３－５）または式（３－７）のパワー情報である。また、Ｋ＝Ｍ（ｍ－１）／２である。

【０１００】
Ψａ（ｆ）は、一般化相互相関関数（ｇｅｎｅｒａｌｉｚｅｄ　ｃｒｏｓｓ　ｃｏｒｒｅｌａｔｉｏｎ　ｆｕｎｃｔｉｏｎ）を最尤推定する際に使われる重み関数として知られており、チャネル間で無相関な雑音の影響を抑圧するのに効果がある。これに関しては、例えば非特許文献６：“Ｔｈｅ　Ｇｅｎｅｒａｌｉｚｅｄ　Ｃｏｒｒｅｌａｉｉｏｎ　Ｍｅｔｈｏｄ　ｆｏｒ　Ｅｓｔｉｍａｔｉｏｎ　ｏｆ　Ｔｉｍｅ　Ｄｅｌａｙ，　Ｃ．　Ｈ．　Ｋｎａｐｐ　ａｎｄ　Ｇ．　Ｃ．　Ｃａｒｔｅｒ，　ＩＥＥＥ　Ｔｒａｎｓ，　Ａｃｏｕｓｔ．，　Ｓｐｅｅｃｈ，　Ｓｉｇｎａｌ　Ｐｒｏｃｅｓｓｉｎｇ”，　Ｖｏｌ．ＡＳＳＰ－２４，　Ｎｏ．４，ｐｐ．３２０－３２７（１９７６）に詳述されている。なお、文献６は相互相関関数を求める方法を開示するものであり、相互相関係数については触れていない。

これに対して、本実施形態では重み付き相互相関係数として、上述の重み関数Ψａ（ｆ）にさらにチャネル間のパワーの比に基づいた重みを与える式（３－６）または式（３－７）によって修正したΨｂ（ｆ）を用いる点が大きく異なる。

【０１０１】
上記の処理では、チャネル間で無相関な雑音に加え、目的方向以外から到来する相関のある雑音までも効果的に抑圧するため、得られた重み付き相互相関係数は、目的信号が存在するか否かの程度を精度よく反映している。このため、重み付き相互相関係数の値を目的信号活性度として用いることができる。この目的信号活性度は、音声検出や音声強調など、種々の応用でその性能を向上するキーコンポーネントとして使用できる。

【０１０２】
本実施形態における目的信号活性度の測定において、活性度を帯域毎に分けて出力してもよい、例えば、ＤＦＴの１から１２８点を周波数上で等間隔に８帯域、つまり、１２８／８＝１６点ずつに分け、８個の目的信号活性度を出力するようにする。分割の仕方は必要に応じて変えて差し支えない。このことは以下の実施形態においても同様である。

【０１０３】
上述の説明では，コヒーレンス関数とパワー情報の両方を使って目的信号活性度を計算しているが、パワー情報を使わず、コヒーレンス関数のみを用いて目的信号活性度を計算しても、ある程度の効果がある。その場合、式（３－４）～（３－７）によって計算されるパワー情報ｐ（ｆ）またはｐｉｊ（ｆ）を１と置けばよい。

【０１０４】

（第４の実施形態）
図１０に、本発明の第４の実施形態に係る音声信号処理装置の構成を示す。本実施形態では、第３の実施形態を音声検出に適用し、目的信号活性度に対して閾値処理を行うことにより、入力音声信号から目的音の成分を検出する。
【０１０５】
複数のマイクロホン１０１－１～１０１－Ｍからの入力音声信号が周波数分析部２０１により周波数領域の信号、すなわち複数チャネルの周波数成分を含むスペクトル情報に変換された後、目的信号活性度計算部３００に入力される。目的信号活性度計算部３００の構成は、第３の実施形態で説明した通りである。
【０１０６】
目的信号活性度計算部３００から出力される目的信号活性度信号３０６は検出処理部４０１に入力され、ここで閾値処理が行われることにより、入力音声信号に目的音が存在しているかどうかを示す目的音検出ステータス信号４０２が出力される。具体的には、検出処理部４０１は入力音声信号に目的音の成分が存在すると判定した場合には“１”を、存在しないと判定した場合には“０”を目的音検出ステータス信号４０２として出力する。
【０１０７】
図１１を用いて本実施形態における処理の流れを説明すると、まずステップＳ４１で入力された入力音声信号を周波数分析し（ステップＳ４２）、得られたスペクトル情報から第３の実施形態で述べた手順により目的信号活性度を計算する（ステップＳ４３）。最後に、目的信号活性度に対して、目的に応じて予め定められた閾値を用いて閾値処理を行うことにより、入力音声信号に目的音の成分が存在しているかどうかの検出処理を行う（ステップＳ４４）。ステップＳ４１～Ｓ４４の処理は、ディジタル化された音声信号がステップＳ４１においてフレーム単位で入力される毎に繰り返し行われる。
【０１０８】
次に、図１２を用いて検出処理部４０１における閾値処理の手順を説明する。ここでは、目的音のない区間の目的信号活性度のバイアスと分散から、検出のための閾値を設定する例について示す。
まず、初期設定を行い（ステップＳ４００）、次いで音声信号の入力（ステップＳ４０１）、周波数分析（ステップＳ４０２）及び目的信号活性度の計算（ステップＳ４０３）をフレーム毎に順次行う。
【０１０９】
第ｋフレームの目的信号活性度を$\rho(k)$とすると、$\rho(k)$について目的音の無い区間（無音区間という）のバイアスと分散を推定する。無音区間か否かの暫定的な判定を$|\rho(k)-b(k-1)|$とκとの比較により行う（ステップＳ４０４）。ここで、$b(k)$は$\rho(k)$のバイアスの推定値、κは判定のための閾値である。
【０１１０】
ここで、$|\rho(k)-b(k-1)|<\kappa$の場合は、無音の可能性が高いと判断し、１次のローパスフィルタを用いて次式に示されるようにバイアス$b(k)$と分散$v(k)$の推定値を更新する（ステップＳ４０５）。
【数２１】

$$b(k)=(1-\eta)b(k-1)+\eta\rho(k) \qquad (4\text{-}1)$$

$$v(k)=(1-\eta')v(k-1)+\eta'(\rho(k)-b(k))*(\rho(k)-b(k)) \qquad (4\text{-}2)$$

【０１１１】
一方、$|\rho(k)-b(k-1)|>\kappa$の場合は、目的音が存在する可能性が高いと判断し、次式に示されるように、バイアス$b(k)$と分散$v(k)$の推定値を更新しない（ステップＳ４０６）。
【数２２】

$$b(k) = b(k-1) \qquad (4-3)$$

$$v(k) = v(k-1) \qquad (4-4)$$

【０１１２】
次に、次式によって検出のための閾値ｈ（ｋ）を設定する（ステップＳ４０７）。
【数２３】

$$h(k) = b(k) + \xi\sqrt{v(k)} \qquad (4-5)$$

【０１１３】
ここで、ξは検出閾値ｈ（ｋ）を設定するための定数である。この結果、ｈ（ｋ）＜ρ（ｋ）ならば目的信号が存在するとして”１”を、そうでなければ”０”をそれぞれ目的ステータス信号として出力する（ステップＳ４０８）。
初期設定に必要なκ，η，η’，ξの値の例は、初期設定ステップＳ４００の枠内に示した通りである。
【０１１４】
図１３に、検出処理の具体的な例を示す。図１３（Ａ）に示す曲線ρから、図１３（Ｂ）に示す検出ステータス信号の時系列が出力される。目的信号活性度の計算は、第３の実施形態で述べたようにチャネル間で相関のない雑音と相関があっても目的音とは違う方向から到来する雑音とを抑圧し、目的音だけに精度よく反応する。従って、算出される目的信号活性度を本実施形態のような音声検出のパラメータとして用いた場合、高い検出性能を達成できる。
【０１１５】
（第５の実施形態）
図１４に、本発明の第５の実施形態に係る音声信号処理装置の構成を示す。本実施形態は、第３の実施形態を音声強調に適用したものである。複数のマイクロホン１０１－１～１０１－Ｍからの入力音声信号が周波数分析部２０１により周波数領域の信号、すなわち複数チャネルの周波数成分を含むスペクトル情報に変換された後、目的信号活性度計算部３００に入力される。目的信号活性度計算部３００の構成は、第３の実施形態で説明した通りである。
【０１１６】
一方、第２の実施形態と同様に周波数分析部２０１からのスペクトル情報は信号統合部２０３にも入力され、ここで１チャネルの統合されることにより、統合スペクトル信号が生成される。信号統合部２０３から出力される統合スペクトル信号は、目的信号活性度計算部３００から出力される目的信号活性度信号（相互相関係数）３０６に従って利得が制御される利得制御部５０１に入力され、その大きさが調整される。これにより、利得制御部５０１から目的音の成分が強調されたスペクトル信号５０２が出力される。
【０１１７】
目的音の成分が強調されたスペクトル信号５０２は、必要に応じて逆変換部５０３によって周波数分析部２０１とは逆の変換、すなわち周波数領域から時間領域への変換が施され、目的音の成分が強調された出力音声信号５０４が生成される。逆変換部５０２は、周波数分析部２０１が例えばＦＦＴの場合、逆ＦＦＴによって実現される。
このように本実施形態に係る音声信号処理装置は、図６に示した第２の実施形態における相互相関係数計算部２０２が重み付き相互相関係数を計算する目的信号活性度計算部３００に変更された構成となっている。
【０１１８】
次に、図１１を用いて本実施形態における処理の流れを説明すると、まずステップＳ５１からステップＳ５３までの処理は、第４の実施形態で説明した図１１中に示すステップＳ

４１からステップＳ４３までの処理と同様である。ステップＳ５２の周波数分析の後、ステップＳ５３の目的信号活性度の計算と平行して、複数チャネルのスペクトル情報を１チャネルに統合して統合スペクトル信号を生成する処理を行う（ステップＳ５４）
次に、統合スペクトル信号に対して、ステップＳ５３で得られた目的信号活性度に応じた利得制御を行って振幅を調整することにより、目的音の成分が強調されたスペクトル信号を生成し（ステップＳ５５）、最後に必要に応じてステップＳ５６で逆変換（例えば、逆ＦＦＴ）を行って目的音の成分が強調された出力音声信号を得る。ステップＳ５１～Ｓ５６の処理は、ディジタル化された音声信号がステップＳ５１においてフレーム単位で入力される毎に繰り返し行われる。

【０１１９】
本実施形態によると、第３の実施形態で説明したように、目的信号活性度が入力音声に目的音があるか否かを高精度に反映するので、これを用いて目的音を強調する音声強調を行うことによって、種々の雑音環境において非常に高い性能の処理を実現できる。

【０１２０】
なお、第３の実施形態の中で、目的信号活性度を複数の周波数帯域に分けて求めてもよいと述べたが、本実施形態の利得制御の処理において、このような複数の周波数帯域に関する目的信号活性度を用いて、帯域毎に利得を制御することも可能である。すなわち、目的信号活性度計算の際に用いた帯域毎に統合信号、例えばスペクトル情報の算出にＬ点のＤＦＴを用い、帯域分割数をＢとする場合、Ｌ／２／Ｂ＝Ｎ点ずつを用いて以下のように目的活性度を計算する。

【０１２１】
【数２４】

$$\rho(b)=\frac{\sum_{f=s(b)}^{e(b)}Wxy(f)\psi b(f)+\sum_{L-e(b)-2}^{L-s(b)+2}Wxy(f)\psi b(f)}{2*\sum_{f=s(b)}^{e(b)}Wxx(f)\psi a(f)+2*\sum_{L-e(b)-2}^{L-s(b)+2}Wxx(f)\psi a(f)} \qquad (5\text{-}1)$$

【０１２２】
ここで、ρ（ｂ）は帯域番号ｂに関する目的信号活性度であり、帯域ｂの計算で用いる周波数成分の範囲をｓ（ｂ），ｅ（ｂ）と置いている。この値は、例えば以下のようにとる。

【数２５】

ｂ＝１のとき
ｓ（１）＝２
ｅ（１）＝Ｎ

ｂ＝１以外のとき
ｓ（ｂ）＝（ｂ－１）＊Ｎ＋１
ｅ（ｂ）＝ｂ＊Ｎ　　（５－２）

【０１２３】
これは、ＤＦＴにおいて周波数成分番号ｆが２からＬ／２の正の周波数に相当する成分と、ｆがＬ／２＋１からＬの負の周波数に相当する成分の番号の一般的な規則性を用いて求められる。ここで、ｆ＝１は直流成分に相当し、一般的な波形信号の場合、その成分は０と置いてよいので、上の計算式では除いてある。また、ｆ＝Ｌ／２の成分は利用可能な周

波数の上限であり、その大きさはやはり０に近いので、除いてある。勿論、これらを計算に含めることにしても、何ら問題はない。

【０１２４】

このようにして求められた目的信号活性度ρ（ｂ）を用いると、統合信号に対する利得制御は、以下のようにして行うことができる。

【数２６】

$$z'(f) = |\rho(b)| * z(f)$$
$$(f = (b-1)*N+1 \text{ から } b*N,\ b=1,\cdots,B)$$

(5-3)

【０１２５】

先に述べたと同様、上式のように目的信号活性度ρ（ｂ）の絶対値を用いてもよいし、ρ（ｂ）の実数部をとって負の場合は０と置いた値を用いて、以下のようにしてもよい。

【０１２６】

【数２７】

$\mathrm{Real}(\rho(b)) > 1$ のとき

$$z'(f) = \mathrm{Real}(\rho(b)) * z(f)$$
$$(f = (b-1)*N+1 \text{ から } b*N,\ b=1,\cdots,B)$$

それ以外のとき

$$z'(f) = 0 \qquad (5-4)$$

【０１２７】

以上の方法により、目的音の成分を強調する際の利得制御を帯域毎に行うことができる。これにより、ある帯域に偏って雑音が存在する場合などに、その帯域のみ抑圧することが可能となるので、目的音成分強調の性能を向上させることができる。

【０１２８】

（第６の実施形態）

図１６に、本発明の第６の実施形態に係る音声信号処理装置の構成を示す。本実施形態は、第５の実施形態にコヒーレンスとパワー情報に基づいたフィルタ演算を行うコヒーレンスフィルタ演算部６０１が加わった構成である。

【０１２９】

次に、図１７を用いて本実施形態における処理の流れについて述べる。まずステップＳ６１からステップＳ６４までの処理は、第５の実施形態の図１１中に示すステップＳ５１からステップ５４までの処理と同様である。本実施形態では、ステップＳ５４で得られた統合スペクトル信号に対して、ステップＳ６４の目的信号活性度計算の仮定で生成されるコヒーレンス関数とパワー情報を用いたフィルタ演算を行う。

【０１３０】

こうしてコヒーレンスフィルタ演算が施された統合スペクトル信号に対して、ステップＳ６３で得られた目的信号活性度に応じた利得制御を行って振幅を調整することにより目的音の成分が強調されたスペクトル信号を生成し（ステップＳ６５）、最後に必要に応じてステップＳ６６で逆変換（例えば、逆ＦＦＴ）を行って、目的音の成分が強調された出力音声信号を得る。ステップＳ６１～Ｓ６６の処理は、ディジタル化された音声信号がステップＳ６１においてフレーム単位で入力される毎に繰り返し行われる。

【０１３１】

次に、コヒーレンスフィルタ演算部６０１について詳しく述べる。コヒーレンスフィルタ演算部６０１では、目的信号活性度計算部３００で計算されるコヒーレンス関数を用いて対象のスペクトル情報をフィルタリングする。コヒーレンス関数は、式（３－１）または式（３－２）を用いて計算される。このとき、目的信号活性度計算部３００で内部的に得

られた式（３－４）～（３－７）のいずれかのパワー情報に従って、次式のようにコヒーレンス関数を修正して用いると、さらに効果的である。

【０１３２】

入力音声信号がｘ（ｆ），ｙ（ｆ）の２チャネルの場合の修正コヒーレンス関数γ（ｆ）は、次式に示される。

【数２８】

$$\gamma(f)=\frac{|Wxy(f)p(f)|}{\sqrt{Wxx(f)Wyy(f)}} \qquad (6\text{-}1)$$

【０１３３】

一方、Ｍチャネルの場合（２チャネルに限定されない場合）の修正コヒーレンス関数γ（ｆ）は、次式に示される。

【数２９】

$$\gamma(f)=\frac{\left|\sum_{i=1,j=1,i<j}^{i\leq M,j\leq M} Wij(f)pij(f)\right|}{\sqrt{\sum_{i=1,j=1,i<j}^{i\leq M,j\leq M} Wii(f)Wjj(f)}} \qquad (6\text{-}2)$$

【０１３４】

ここで、第３の実施形態と同様にｉ，ｊはチャネル番号、Ｗｉｊ（ｆ）は第ｉチャネルと第ｊチャネル間のクロススペクトル、Ｗｉｉ（ｆ），Ｗｊｊ（ｆ）は第ｉチャネル及び第ｊチャネルのパワースペクトルである。

【０１３５】

式（６－１）または式（６－２）に示される修正コヒーレンス関数γ（ｆ）を用いたフィルタ演算は、次式に従って行われる。

【数３０】

$$Zo(f)=Z(f)\gamma(f) \qquad (6\text{-}3)$$

【０１３６】

ここで、ＺＯ（ｆ）はフィルタ演算の出力、Ｚ（ｆ）は信号統合部２０３で得られる統合スペクトル信号である。

【０１３７】

このとき、例えば次式のようにコヒーレンス関数γ（ｆ）を適当な関数を用いて修正してから、フィルタ演算を行ってもよい。

【数３１】

$$Zo(f)=Z(f)pow(\gamma(f),\alpha) \qquad (6\text{-}4)$$

【０１３８】

ここで、ｐｏｗ（ａ，ｂ）は、ａのｂ乗を表す指数関数であり、例えばα＝２などを用いることがある。この場合、式（６－３）（α＝１に相当する）よりもコヒーレンス関数γ（ｆ）の値が強調され、雑音抑圧量が増加するが、代わりに目的音声の歪みも大きくなるので、状況に合わせて設定するのがよい。

【０１３９】

このように本実施形態によれば、目的信号活性度を用いた目的音の強調に際して、コヒーレンス関数に対応したスペクトルの重み付けを行うことにより、チャネル間で無相関な雑

音に対する音声強調性能をさらに改善することができる。

【０１４０】

（マイクロホンの配置について）

次に、これまでに述べたマイクロホンの好ましい配置方法について述べる。音声信号処理装置は、複数のマイクロホンに対して目的音に関しては同一の成分が入射し、雑音に関しては位相と振幅の少なくとも一方の異なった成分が入射することを想定している。このようなマイクロホンの受音状況を実現するためには、マイクロホン１０１－１～１０１－Ｍを以下に述べるように配置することが望ましい。

【０１４１】

第３の実施形態では、重み付き相互相関係数を計算する過程で、チャネル間のパワー比に関する情報を用いており、目的音に対してはチャネル間で等パワー、雑音に対してはチャネル間で異なるパワーとなるようにマイクロホン１０１－１～１０１－Ｍを配置したときに、高い性能が得られる。マイクロホン１０１－１～１０１－Ｍに全て無指向性のマイクロホンを用いた場合でも、ある程度の性能は発揮できる。これは受音位置によって反射などの条件が異なるため、無指向性マイクロホンでも到来音のパワーが異なる場合があるからである。

【０１４２】

しかしながら、高い性能を安定して発揮させるには、マイクロホン１０１－１～１０１－Ｍの少なくとも一つを指向性マイクロホンとする方がよい。これにより、目的音の到来方向以外の方向に対してチャネル間で感度差を作り出し、雑音抑圧性能を向上することができる。

【０１４３】

ここでは、マイクロホンの数Ｍが２個，すなわち２チャネルの場合について述べるが，３以上の多チャネルの場合にも容易に拡張可能である。図１８に示すように、２個のマイクロホンの一方が無指向性マイクロホン７０１で、他方が指向性マイクロホン７０２の場合と、図１９に示すように２個のマイクロホン７１１，７１２がいずれも指向性マイクロホンの場合について述べる。各々特徴のある使い分けが可能である。指向性マイクロホンとしては、通常の単一指向性マイクロホンを想定する。単一指向性以外のもっと鋭い指向性のものを使う場合は，さらに性能が高くなる可能性があるが、配置方法は単一指向性マイクロホンを用いた場合と同様である。

【０１４４】

図１８に示すように無指向性マイクロホン７０１と指向性マイクロホン７０２を用いた場合、指向性マイクロホン７０２は，目的音の方向に指向性の頂点（感度最大方向）が向くようにする。マイクロホン７０１，７０２間の距離は，例えば５ｃｍから２０ｃｍ程度が適当である。この配置においては、無指向性マイクロホン７０１の感度と指向性マイクロホン７０２の頂点方向の感度を同程度に調整しておくことが望ましい。

【０１４５】

このような配置により、指向性マイクロホン７０２における低感度の方向、例えば，図１８に示すように高感度の方向と１８０°逆の方向に関しては，チャネル間，すなわちマイクロホン７０１，７０２間の感度差が非常に大きいので，低感度の方向からの到来音の抑圧量は非常に大きくなる。一見、これは指向性マイクロホンの元々の指向性を表したにすぎないように見えるが、チャネル間のパワー比に対する鋭敏性を式（３－６）または式（３－７）のβの値によって調整できるため、指向性マイクロホン７０２の元々の指向性よりも鋭い指向性に調整することが可能である。

【０１４６】

すなわち、例えばβ＝２とすることにより、実際のパワー比の２乗の重みが目的信号活性度の計算に使われることになる。実際のパワー比は、目的音方向に関しては１であるが、目的音の到来方向以外の方向では１以下であるため、これを２乗することによって目的音以外の成分に関する重みはさらに小さくなる。このため，低感度方向と目的音方向の間の横方向などの感度もさらに小さくできる。

【０１４７】
一方、図１９に示すように２つのマイクロホンに指向性マイクロホン７１１，７１２を用いた場合は、例えば図１９（Ａ１）～（Ａ４）に示す配置が有効である。これは、同一平面上に二つのマイクロホン７１１，７１２の指向性の軸が含まれるような配置であり、図で上から見たときの指向性の軸の向きがθ＝－９０°～９０°程度の範囲内にあるのが望ましい。θ＞０の場合は指向性の軸が２つのマイクロホン７１１，７１２の中点から外に開くような形となるが、θ＜０としても同様な性能であり、この場合は指向性の軸が中点に向かう形となる。
【０１４８】
図１９（Ｂ１）～（Ｂ４）は、２つの指向性マイクロホン７１１，７１２のもう一つの好ましい配置の例である。指向性の軸は同一平面内に含まれていない。正確さを期すため、図２０に図１９（Ｂ１）～（Ｂ４）の配置における指向性の軸の向きを方位角θと仰角ϕで表した図を示す。ここで、Ｒチャネルのマイクロホン７１２の指向性の軸の向きを（θ，ϕ）とすると，Ｌチャネルのマイクロホン７１１の指向性の軸の向きは（－θ，－ϕ）となるようにするのが望ましい。すなわち、２つのマイクロホンの位置と軸方向は１８０°の回転対称をなす。マイクロホンの数がＭならば、３６０°／Ｍの回転対称となる配置が望ましい。θとϕの範囲は、１０°＜θ＜８０°，１０°＜ϕ＜８０°となるようにするのが望ましい。指向性の軸の向きを上のように設定した後、２つのマイクロホン７１１，７１２の位置を目的音の到来方向を軸として回転させた場合も全く同じ特性を有するので、必要に応じて回転させて用いてもよい。
【０１４９】
図１９（Ａ１）～（Ａ４）の配置の場合，前述した音声信号処理により最終的な指向性は、目的音の到来方向に対しては感度最大となり、指向性マイクロホン７１１，７１２から等距離の方向，すなわち、２つのマイクロホン７１１，７１２を結ぶ直線に垂直な方向に対しては感度が極大となるため，真上や真下からの到来音に対してもある程度の感度を持つようになる。
【０１５０】
これに対し、図１９（Ｂ１）～（Ｂ４）の配置では、２つの指向性マイクロホン７１１，７１２の位相が一致する方向は、図１９（Ａ１）～（Ａ４）の場合と同様に、マイクロホン７１１，７１２から等距離の方向、すなわち２つのマイクロホン７１１，７１２を結ぶ直線に垂直な面（図２１の面ａ）に含まれる方向となる。一方、２つのマイクロホン７１１，７１２の感度が一致するような到来方向は、マイクロホン７１１，７１２の軸の向きを表す２つのベクトルを１つの平面上に平行移動したとき、その２つのベクトルの差ベクトル（図２１のベクトルＣ）と垂直の平面（図２１の面ｂ）に含まれる。
【０１５１】
本実施形態における目的信号活性度は、位相と振幅がチャネル間で共に一致した場合に大きな値となるので、図２１に示す面ａと面ｂが交わる方向、すなわち正面方向（図２０または図２１において矢印で示した目的音の到来方向）とその１８０°逆の方向にのみ、指向性の大きな極大ができる。正面の逆方向に関しては、指向性マイクロホン７１１，７１２の低感度方向が向いているため，その方向からの入射音のレベルは低い。従って、実質的に正面方向のみに極大のメインローブを持つような指向性を得ることができるので、真上や真下からの到来音も抑圧したい場合には、この配置が有効である。
【０１５２】
（第８の実施形態）
図２２に、本発明の第８の実施形態に係る音声信号処理装置の構成を示す。本実施形態は、第３の実施形態における周波数分析部２０１と目的信号活性度計算部３００との間にスペクトル補正部８００を挿入した構成になっている。図２３に示されるように、スペクトル補正部８００は適応フィルタ８０１と補正フィルタ８０２を有する。
【０１５３】
前述したように、本発明の実施形態に係る音声信号処理装置は、目的音に関しては同一の

成分が複数のマイクロホン１０１－１～１０１－Ｍに入射することを想定している。従って、マイクロホン１０１－１～１０１－Ｍの感度が経年変化やバイアス設定用のバッテリの消耗などによって変化した場合、処理精度が低下する可能性がある。目的音の到来方向が想定している方向とずれた場合にも、処理精度が低下する可能性がある。

【０１５４】

本実施形態では、マイクロホン１０１－１～１０１－Ｍ毎の感度の違いや目的音の到来方向のずれを補正して、本来の性能を発揮させるために、スペクトル補正部８００により、周波数分析部２０１で得られたスペクトル情報に対して、目的信号活性度計算部３００で得られた目的信号活性度とスペクトル情報とに基づく修正を施す。

【０１５５】

次に、図２３を用いてスペクトル補正部８００での処理の詳細を述べる。ここでは、入力音声信号が２チャネルの場合について述べるが、Ｍチャネルへの拡張も同様である。スペクトルの補正は、チャネル間の差を適応フィルタ８０１により同定し、一方のチャネルのスペクトルに対して補正フィルタ８０２を用いて、適応フィルタ８０１により同定した差分を補正することにより行う。適応フィルタ８０１による差分同定の際、目的信号活性度信号３０６に従ってフィルタ更新の速さを制御するようにするようにしてもよい。

【０１５６】

適応フィルタ８０１としては、例えば、周波数領域のＬＭＳ適応フィルタを使用することが可能である。この場合、周波数領域ＬＭＳ適応フィルタの計算は、以下のように行われる。

【数３２】

$$E(k,f) = X(k,f) - Y(k,f)W(k-1,f) \qquad (8\text{-}1)$$

$$W(k,f) = \frac{W(k-1,f) - 2\mu E(k,f)X(k,f)*}{\left|X(k,f)X(k,f)*\right|} \qquad (8\text{-}2)$$

【０１５７】

ここで、ｋはフレーム番号、Ｘは第１チャネルのスペクトル、Ｙは第２チャネルのスペクトル、Ｅは誤差スペクトル、Ｗは複素数のフィルタ係数、μはステップサイズ、（＊）は複素共役である。

【０１５８】

補正フィルタ８０２の演算は、この場合、第１チャネルのスペクトルＸ（ｋ，ｆ）に対し、Ｘ′（ｋ，ｆ）＝Ｗ（ｋ，ｆ）Ｘ（ｋ，ｆ）により行う。Ｘ′（ｋ，ｆ）は、補正後の第１チャネルのスペクトルである。この演算は、適応フィルタ８０１の演算の式（８－１）で既に行われているので、新たに補正フィルタ８０２を用意せず、適応フィルタ８０１からＷ（ｋ，ｆ）Ｘ（ｋ，ｆ）の信号を取り出すだけでもよい。

【０１５９】

目的信号活性度ρ（ｋ）を用いて適応フィルタ８０１による差分同定の際のフィルタ更新の速さを制御することも可能であり、その場合は例えば、次式のように適応フィルタ８０１の更新式（８－２）を修正する。

【０１６０】

【数３３】

$\rho(k)$が閾値hより大きいとき

$$W(k,f) = \frac{W(k-1,f) - 2\mu E(k,f)X(k,f)*}{|X(k,f)X(k,f)*|}$$

それ以外のとき

$$W(k,f) = W(k-1,f) \qquad (8-3)$$

【０１６１】

ここで、閾値hとしては例えば０．５を使うことができる。これは、ρ（ｋ）の大きさが閾値より大きいときだけチャネル間の差分を求めることになるので、目的音が到来している可能性の大きいときのみフィルタ更新が行われ、雑音に適応してしまう心配がない。このような閾値を用いた適応の更新／停止の制御の他、次式のようにρ（ｋ）に比例させることも可能である。

【数３４】

$$W(k,f) = \frac{W(k-1,f) - \rho(k)2\mu E(k,f)X(k,f)*}{|X(k,f)X(k,f)*|} \qquad (8-4)$$

【０１６２】

式（８－３）を使ってチャネル間の差分を推定した場合、例えば，感度差が始めから大きく異なっている場合などは、ρ（ｋ）の値が閾値を上回らないため、適応フィルタ８０１の更新が行われず、差分が全く求まらないこともある。しかしながら、前述したようにマイクロホンの感度が経年変化やバイアス設定用のバッテリの消耗などによって変化したことを想定した場合、感度差が急に大きくなることは少なくこのような不都合はあまり問題にならない。本実施形態は、例えば第３～第６の実施形態で説明した音声信号処理における目的信号活性度を求める際の補正方法として用いることで、チャネル間の感度の差に影響を受けない動作が可能となる。

【０１６３】

（第９の実施形態）

図２４に、本発明の第９の実施形態に係る音声信号処理装置の構成を示す。第８の実施形態と同様に、スペクトル補正部９００が設けられ、さらに補正フィルタ学習指示部９１０が追加されている。

【０１６４】

第８の実施形態で示した感度補正は、マイクロホン１０１－１～１０１－Ｍの感度が大きくは違わない場合に効果があった。第９の実施形態では、目的音の振幅または位相が各マイクロホンで同一であると想定できないような場合に、学習モードの処理を設け，第８の実施形態とは別の補正フィルタの学習を行ってチャネル間の差を補正する。

【０１６５】

学習後の経年変化による感度ずれや、目的話者位置の小さなずれによる位相差などを補正する場合は、学習モードを経て学習したフィルタによる補正の後、第８の実施形態で述べたような自動的な補正を行う。本実施形態は、このような二つの補正ができる構成になっている。

【０１６６】

目的音方向が想定している方向と異なる場合や、各マイクロホン１０１－１～１０１－Ｍと目的音源との距離が異なるようなマイクロホン配置にした場合などでも、本実施形態の音声処理方法を利用可能になる。学習モードは、利用者の指示をトリガとして開始したり、装置の起動後などに装置側が自動的に学習モードに入る場合などがある。

【０１６７】

補正フィルタ学習指示部９１０は、学習モードであるか否かを表す信号を出力する。例えば、学習モードは“１”、学習モードでなければ”０”を出力する。学習モードの終了は、装置側が自動的に行ってもよいし、利用者が指示するようにしてもよい。学習モードにおいては、入力したい目的音の位置からテスト音を発生させる。利用者が発声してもよいし、スピーカなどのテスト音発生装置を目的音位置に置いて使用してもよい。テスト音は使用目的に応じて選択してよい。音声入力が目的なら音声や白色雑音を使うのが望ましい。

【０１６８】

図２５に示されるように、補正フィルタ学習指示部９１０はスイッチ９１１により利用者の指示が入力されると、一定期間を学習モードとするように、指示入力後からの経過時間をタイマ９１２で測定して、補正フィルタ学習指示信号Ｓを出力する。タイマ９１２は補正フィルタ学習指示信号Ｓとして、スイッチ９１１による指示入力時点から予め定めた時間までは例えば”１”を出力し、その他の期間は“０”を出力する。タイマ９１２は、大抵のマイクロプロセッサに備わっている機能であるので、それを使えばよい。学習モードの終了は、このように装置側がタイマ９１２を用いて自動的に行ってもよいし、利用者が指示するようにしてもよい。

【０１６９】

スペクトル補正部９００は、補正フィルタ学習指示部９１０からの指示に従って一定時間長の期間、例えば３秒間にわたって学習を行う。この期間を学習モードと呼ぶことにする。学習モードにおいては、入力したい目的音の位置からテスト音を発生させる。利用者が発声してもよいし、スピーカなどのテスト音発生装置を目的音位置に置いて使用してもよい。テスト音は使用目的に応じて選択してよい。音声入力が目的なら音声や白色雑音を使うのが望ましい。学習モードの終了後は、続けて第８の実施形態までに述べたような音声信号処理を行う。

【０１７０】

スペクトル補正部９００の構成は、第８の実施形態におけるスペクトル補正部８００の図２３に示した構成と若干異なり、図２６に示されるように図２３の補正フィルタ８０２に相当する補正フィルタ９０２に加えて、補正フィルタ９０２の前段にもう一つの補正フィルタ９０１が追加されている。補正フィルタ９０２は、第８の実施形態で説明したと同様の働きをする。すなわち、チャネル間の小さなずれを補正する。

【０１７１】

一方、追加された補正フィルタ９０１は、チャネル間の大きな差を補正する。補正フィルタ９０１は、学習モード以外は固定される。補正フィルタ学習指示部９１０からの学習フィルタ指示信号Ｓが“１”の場合、適応フィルタ９０４は補正フィルタ９０１を学習させ、学習フィルタ指示信号Ｓが“０”の場合は補正フィルタ９０２を学習させる。

【０１７２】

例えば、ＬＭＳを用いた補正フィルタ９０２の学習は次式により行われる。

【数３５】

$$E0(k,f) = X1(k,f) - Y(k,f)W0(k-1,f) \qquad (9\text{-}1)$$

$$W0(k,f) = \frac{W0(k-1,f) - 2\mu 0 E0(k,f)X(k,f)*}{|X1(k,f)X1(k,f)*|} \qquad (9\text{-}2)$$

【０１７３】

一方、補正フィルタ９０１の学習は、次式により行われる。

【数３６】

$$E1(k,f) = X(k,f) - Y(k,f)W1(k-1,f) \quad (9-3)$$

$$W1(k,f) = \frac{W1(k-1,f) - 2\mu 1E1(k,f)X(k,f)*}{|X(k,f)X(k,f)*|} \quad (9-4)$$

【０１７４】
ここで、ｋはフレーム番号、Ｘは第１チャネルのスペクトル、Ｙは第２チャネルのスペクトル、Ｘ１はＸに補正フィルタ９０１をかけた後のスペクトル、Ｗ０は補正フィルタ９０２のフィルタ係数、Ｅ０は補正フィルタ９０２の学習の際の誤差スペクトル、μ０は補正フィルタ９０２の学習の際のステップサイズ、Ｗ１は補正フィルタ１のフィルタ係数、Ｅ１は補正フィルタ９０１の学習の際の誤差スペクトル、μ１は補正フィルタ９０１の学習の際のステップサイズ，（＊）は複素共役である。ステップサイズμ０，μ１には、例えば０．１を使う。
【０１７５】
式（９－１），（９－２）の補正フィルタ９０２の学習を行う際、第８の実施形態のように目的信号活性度を用いて適応の速度を制御してよい。補正フィルタ９０１のフィルタリングは，
【数３７】

$$X1(k,f) = W1(k,f)X(k,f) \quad (9-5)$$

により行われ、補正フィルタ９０２のフィルタリングは
【数３８】

$$X'(k,f) = W0(k,f)X1(k,f) \quad (9-6)$$

ここで、Ｘ′（ｋ，ｆ）は、スペクトル補正部９００の出力となる第１チャネルのスペクトルである。
【０１７６】
次に、図２７を用いて本実施形態の処理の流れを説明する。
まず、初期設定として補正フィルタ９０１，９０２の係数の初期値を設定する（ステップＳ９０）。補正フィルタ９０１を補正フィルタ１、補正フィルタ９０２を補正フィルタ０とすると，補正フィルタ１，０の係数の初期値を全ての周波数（ｆ）で（１，０）としておけば，学習をしない場合でも音声信号の入力が可能となるので扱いやすい。ここで、（１，０）は複素数の１＋ｊ０を表す。ただし，補正フィルタ１，０の係数の初期値を全ての周波数（ｆ）で（０，０）とした場合でも、学習さえ進めば動作するようになるので，初期値をどのように選ぶかは本質的な違いはない。
【０１７７】
次に、補正フィルタ学習指示信号Ｓが“１”か否か（“０”）を調べ（ステップＳ９１）、Ｓ＝“１”であれば補正フィルタ１の学習を式（９－３）（９－４）に従って行う（ステップＳ９３）。一方、Ｓ＝“０“であれば補正フィルタ１によるフィルタリングを式（９－５）に従って行い（ステップＳ９４）、次に補正フィルタ０の学習を式（９－１）（９－２）に従って行った後、補正フィルタ０によるフィルタリングを行い（ステップＳ９３～Ｓ９４）、この後に目的信号活性度を測定する（ステップＳ９６）。ステップＳ９１からステップＳ９６までの処理は、ディジタル化された音声信号がステップＳ９１においてフレーム単位で入力される毎に繰り返し行われる。
【０１７８】
本実施形態によれば、例えばマイクロホン１０１－１～１０１－Ｍを目的音源の位置に対して距離が異なるように並べた場合にも、目的信号活性度の計算、目的音の検出及び目的

音の強調などの処理を有効に行うことが可能となる。

【０１７９】

自動車内で観測される走行雑音の環境下で用いる場合、走行雑音は拡散性が強いため、マイクロホンを異なった位置や向きに置いた場合でもチャネル間の振幅の差があまりない。各マイクロホンと目的音位置との距離が異なるように並べた場合は、本実施形態のスペクトル補正によって、チャネル間で目的音が同振幅、同位相となるように補正される。一方で、等振幅だった雑音成分は補正により異なった振幅となり、目的信号活性度における雑音区間の区別が容易になり、活性度測定の精度が向上する。このように、マイクロホンを目的音から等距離に並べない場合は、拡散性雑音下での性能向上を図ることができる。

【０１８０】

（第１０の実施形態）

図２８に、本発明の第１０の実施形態に係る音声信号処理装置の構成を示す。本実施形態は、修正相互相関係数に基づいて音源の到来方向を推定する技術に関する。音源の到来方向の推定は、音声強調や雑音源の同定など音声処理における種々の応用において重要である。特に、本実施形態に係る修正相互相関係数に基づく方法は、適応ビームフォーマなどの死角制御に基づく方法に比べて雑音源の信号や伝播状況に関する制約が少なく、広範囲の雑音環境で使用可能であるという利点がある。

【０１８１】

本実施形態に係る音声信号処理装置は、図２８に示すようにマイクロホン１０１－１～１０１－Ｍからの複数（Ｍ）チャネルの入力音声信号を周波数分析して周波数成分であるスペクトル情報に変換する周波数分析部２０１と、該スペクトル情報から音源方向を推定する音源方向推定部１０００からなる。音声分析部２０１の処理は、第２の実施形態（図６）で説明した通りである。

【０１８２】

音源方向推定部１０００は、クロス・パワースペクトル計算部１００１、コヒーレンス関数計算部１００２、補正係数発生部１００３、クロス・パワースペクトル補正部１００４、パワー情報計算部１００５、仮想方向相関係数計算部１００６及び音源方向検出部１００７を有する。以下、音源方向推定部１０００の各構成要素について説明する。

【０１８３】

クロス・パワースペクトル計算部１００１は、周波数分析部２０１により得られたスペクトル情報から各チャネルのパワースペクトルとチャネル間のクロススペクトルを計算する。

【０１８４】

コヒーレンス関数計算部１００２は、クロス・パワースペクトル計算部１００１で得られたクロススペクトルと各チャネルのパワースペクトルから入力音声信号のチャネル間のコヒーレンス関数を算出する。

【０１８５】

補正係数発生部１００３は、予め設定した信号の到来方向範囲の中に、信号の仮想的な到来方向である仮想方向を定め、この仮想方向から信号が到来したと仮定した場合に、入力音声信号のスペクトル情報中の当該信号成分がチャネル間で一致するようにスペクトル情報を補正するための補正係数を発生する。

【０１８６】

クロス・パワースペクトル補正部１００４は、発生した補正係数を用いてクロススペクトルとパワースペクトルを補正し、補正クロススペクトルと補正パワースペクトルを生成する。

【０１８７】

パワー情報計算部１００５は、補正クロススペクトルと補正パワースペクトルに基づいて入力音声信号のチャネル間の周波数毎の信号パワー比であるパワー情報を算出する。

【０１８８】

仮想方向相関係数計算部１００６は、補正パワースペクトルと補正クロススペクトルを先

のコヒーレンス関数とパワー情報に基づいて重み付けし、予め設定した１組の仮想方向に対応した相互相関係数を仮想方向毎に算出する。

【０１８９】

音源方向検出部１００７は、仮想方向相関係数計算部１００６によって計算された仮想方向毎の相互相関係数に基づき音源方向を検出して出力すると同時に、検出した音源方向における相互相関係数の値を音源相関係数として、また音源方向に対応した補正係数を音源方向補正係数として出力する。

【０１９０】

次に、各部の処理についてさらに詳しく説明する。クロス・パワースペクトル計算部１００１、コヒーレンス関数計算部１００２及びパワー情報計算１００５での計算には、例えば入力音声信号のチャネル数Ｍが２チャネルの場合は式（３－８），（３－９），（３－１０）を用い、３チャネル以上の場合は式（３－１２），（３－１３），（３－１４）を用いる。

【０１９１】

補正係数発生部１００３は、予め信号が到来する範囲を例えば図２９に示すように設定する。到来方向は、水平方向の角度である方位角θと垂直方向の角度である仰角φの組（θ，φ）で表すものとし、例えば到来範囲の中の格子点上の方向を仮想方向とするものとする。図２９の場合、到来範囲は方位角、仰角共に－４０°～４０°、格子点は方位角、仰角共に５°おきであり、全ての格子点上の方向を仮想方向の組とする。図２９では、作図の都合上格子点の間隔を５°にしてあるが、実際はもっと小さく、２°以下にすることが望ましい。

【０１９２】

格子点上の仮想方向は、ｄｈ，ｇ＝（θｈ，φｇ）で表すことにする。ここでｈは格子点の方位角に関する番号、ｇは仰角の番号である。補正係数発生部１００３は、仮想方向に対応する補正係数を次式に従って生成する。

【０１９３】

【数３９】

$$Hi(f,\theta,\phi)=\frac{\exp\left\{\dfrac{j2\pi fF\,\tau i(\theta,\phi)}{L}\right\}}{Di(\theta,\phi)} \qquad (10\text{-}1)$$

【０１９４】

ここで、ｉはチャネル番号、Ｈｉ（ｆ，θ，φ）は（θ，φ）方向に関するｉ番目のチャネルの補正係数、τｉ（θ，φ）は、ｉ番目のマイクロホンに（θ，φ）方向からの到来信号が到達するときの基準マイクロホンでの受音信号に対する伝播遅れ時間、Ｄｉ（θ，φ）は、ｉ番目のマイクロホンにおける（θ，φ）方向の感度の指向性、ｆは周波数番号、Ｆはサンプリング周波数、ＬはＦＦＴの点数である。基準マイクロホンは、例えば１番目のマイクロホンとする。

【０１９５】

伝搬遅延の値は、例えば図３０に示すようなマイクロホン配置において、到来音の方向がｄ＝（θ，φ）の場合、基準位置を座標の原点にとると、原点に対する時間遅れは、極座標と直交座標の関係を用いて以下のように計算できる。

【０１９６】

【数４０】

$$\tau 1(\theta,\phi) = -\frac{A1 \bullet d}{c} = -\frac{x1\cos\theta\cos\phi + y1\sin\theta\cos\phi + z1\sin\phi}{c} \quad (10\text{-}2)$$

【０１９７】
ここで、・は内積、ｃは音速である。マイクロホンｉの位置が　Ａｉ＝（ｘｉ，　ｙｉ，　ｚｉ）のときは、次式となる。
【０１９８】
【数４１】

$$\tau i(\theta,\phi) = -\frac{Ai \bullet d}{c} = -\frac{xi\cos\theta\cos\phi + yi\sin\theta\cos\phi + zi\sin\phi}{c} \quad (10\text{-}3)$$

【０１９９】
Ｄｉ（θ，φ）は、マイクロホン固有の特性であるので、製品情報から得るか、または測定により得る。マイクロホン感度の指向性の測定は、例えばマイクロホンへの音の入射角度を変えながら出力を測定すればよく、一般的な方法を用いればよいので、ここでは省略する。
【０２００】
補正係数発生部１００３で発生する補正係数は、音源方向探索の範囲とマイクロホン１０１－１～１０１－Ｍの指向性が変化しなければ変化しないので、最初に係数を発生した後はテーブルに記憶しておき、格子点の番号でテーブルを参照して係数の値を読み出すようにする。
【０２０１】
クロス・パワースペクトル補正部１００４では、補正係数発生部１００３で発生した補正係数を対応するチャネルのクロススペクトルとパワースペクトルに乗じて補正クロススペクトルと補正パワースペクトルを求める。計算は、次式のように行う。
【０２０２】
【数４２】

$$W'ij(f,\theta,\phi) = Wij(f)Hi(f,\theta,\phi)Hj^{*}(f,\theta,\phi) \quad (10\text{-}4)$$

【０２０３】
ここでＷ′は補正後のスペクトル、＊は複素共役、ｉ，ｊはチャネルの番号であり、ｉ≠ｊのときはクロススペクトル、ｉ＝ｊのときはパワースペクトルを意味する。
【０２０４】
式（１０－４）の補正は、スペクトル情報Ｘｉ（ｆ）をＨｉ（ｆ，θ，φ）で補正してからクロス・パワースペクトルを計算することと等価であり、上線を付した処理を時間平均化処理として、Ｈｉは時間に対して変化しないことを使うと、以下のようになることに基づいている。
【数４３】

$$\begin{aligned}&\overline{\{Xi(f)Hi(f,\theta,\phi)\}\{Xj(f)Hj(f,\theta,\phi)\}^*}\\&=\overline{Xi(f)Xj^*(f)Hi(f,\theta,\phi)Hj^*(f,\theta,\phi)\}^*}\\&=\overline{Wij(f,\theta,\phi)Hi(f,\theta,\phi)Hj^*(f,\theta,\phi)\}^*}\end{aligned} \qquad (10-5)$$

【０２０５】
パワー情報計算部１００５では、クロス・パワースペクトル補正部１００４で補正したパワースペクトルからチャネル間のパワー比を求めることにより行う。パワー比の計算は、式（３－７）において元々のパワースペクトルＷｉｉ（ｆ）の代わりに、次式のように補正したものを使う。
【０２０６】
【数４４】

$$pij(f,\theta,\phi)=pow\left\{\frac{\min\left[\left|W'ii(f,\theta,\phi)\right|,\ \left|W'jj(f,\theta,\phi)\right|\right]}{\max\left[\left|Wii'(f,\theta,\phi)\right|,\ \left|W'jj(f,\theta,\phi)\right|\right]},\beta\right\} \qquad (10-6)$$

【０２０７】
仮想方向相互相関係数計算部１００６では、補正したクロス・パワースペクトルとパワー情報を用い、仮想方向（θ，φ）に関する相互相関係数を計算する。相互相関係数の計算は、式（３－１１），（３－１２），（３－１３）において、元々のクロス・パワースペクトルとパワー情報を次式のように各々の補正したものに置き換えればよい。
【０２０８】
【数４５】

$$\Psi aij(f,\theta,\phi)=\frac{\gamma' ij(f)}{(1-\gamma' ij(f))\left|Wij(f,\theta,\phi)\right|} \qquad (10-7)$$

$$\Psi bij(f,\theta,\phi)=\frac{\gamma' ij(f)pij(f,\theta,\phi)}{(1-\gamma\ ij(f))\left|Wij(f,\theta,\phi)\right|} \qquad (10-8)$$

$$p(\theta,\phi)=\frac{\sum_{i<M,j\le M}^{i\le M,j\le M}\ \sum_{i=Li,j=Lj,i<j}^{L2} W'ij(f,\theta,\phi)\Psi' bij(f,\theta,\phi)}{\sqrt[K]{\prod_{i<M,j=Li,i<j}^{i\le M,j\le M}\ \sum_{f=L1}^{L2} W'ii(f,\theta,\phi)\Psi' aij(f,\theta,\phi)\sum_{f=L1}^{L2} W'jj(f,\theta,\phi)\Psi' aij(f,\theta,\phi)}} \qquad (10-9)$$

【０２０９】
ここで、Ｋは
【数４６】

$$K=\sum_{i=1,j=1,i<j}^{i\le M,j\le M}1=M(M-1)/2 \qquad (10\text{-}10)$$

であり、和における周波数ｆの範囲Ｌ１，Ｌ２は目的音の帯域に相当する範囲に相当する番号になるようにする。例えば、目的音の帯域を２６０Ｈｚから４ｋＨｚであると定めた場合には、ＦＦＴ長２５６、サンプリング１１ｋＨｚの場合には、Ｌ１＝６，Ｌ２＝９２とするのがよい。

【０２１０】

式（１０－６）～（１０－１０）を用い、θ＝θｈｇ，φ＝φｈｇとし、設定した到来範囲の仮想方向ｄ（θｈｇ，φｈｇ）（ｈ＝１～Ｎｈ，　ｇ＝１～Ｎｇ）に対して仮想方向相関係数を求める。

【０２１１】

音源方向検出部１００７は、仮想方向相互相関係数計算部１００６によって計算された仮想方向毎の相関係数から、そのピークを検出して音源方向として出力する。このとき、例えば次式のように仮想方向相関係数の時間的な平均化によって安定化を図ることができる。

【０２１２】

【数４７】

$$\rho'=k(\theta,\phi)=\rho k(\theta,\phi)\eta+\rho' k-1(\theta,\phi)(1-\eta) \qquad (10\text{-}11)$$

【０２１３】

ここで、ρ’ｋ　はｋ番目のフレームの処理において平均化された仮想方向相関係数、ρｋ　はｋ番目のフレームの処理において求められた仮想方向相関係数、ηは学習定数であり、例えばη＝０．０５などを用いる。ピークの検出は、ρ’ｋ（θ，φ）　から最大値を求めればよい。

【０２１４】

音源方向検出部１００７は、音源方向の他、音源方向のピークの値である音源相関係数と、音源方向に相当する補正係数である音源方向補正係数を出力する。このために、補正係数発生部１００３の内部の補正係数のテーブルから、音源方向の格子点の番号に基づいて補正係数を取り出すようにする。

【０２１５】

次に、図３１を用いて本実施形態における処理の流れを説明する。

まず，初期設定として音源方向の範囲を設定する（ステップＳ１００）。次に補正係数の生成（ステップＳ１０１）、マイクロホン１０１－１～１０１－Ｍからの音声信号の入力（ステップＳ１０２）、周波数分析（ステップＳ１０３）、クロススペクトルとパワースペクトルの計算（ステップＳ１０４）及びコヒーレンス関数の計算（ステップＳ１０５）を順次行う。次に、スペクトル補正（ステップＳ１０６），パワー情報の計算（ステップＳ１０７）及び仮想方向相互相関関数の計算（ステップＳ１０８）を全ての仮想方向について繰り返し行い、最後に音源方向の検出を行う（ステップＳ１０９）。ステップＳ１０２～Ｓ１０９の処理は、ディジタル化された音声信号がステップＳ１０２においてフレーム単位で入力される毎に繰り返し行われる。

【０２１６】

（第１１の実施形態）

本発明の音声強調処理は、マイクロホン配列の正面から目的とする音である目的音が到来すると仮定しているので、目的音の方向が仮定とずれた場合は、性能が低下する可能性がある。第８の実施形態で述べた適応処理に基づいた補正により、目的音の方向ずれへの対処がある程度は可能であるが，目的音の方向が大きくずれた場合には適応処理だけでは対処が困難である。そこで，本実施形態では第１０の実施形態で説明した音源方向推定処理

の結果を用いて目的音の方向を追尾することにより、目的音が想定している方向とずれた場合に対する音声強調処理の安定度を向上させる。

【０２１７】

図３２に、本実施形態に係る音声信号処理装置の構成を示す。本実施形態は、第１０の実施形態で説明した音源方向推定処理で音源方向を推定し、音源方向に対応する補正係数を用いて入力のスペクトル情報の補正を行い、補正したスペクトル情報を統合し、統合スペクトル情報に対して利得制御を行って音声強調を行う。

【０２１８】

このような処理を実現するため、本実施形態に係る音声信号処理装置は第１０の実施形態で説明した音源方向推定部１０００、周波数分析部２０１からの複数チャネルのスペクトル情報を音源方向補正係数に基づいて補正するスペクトル情報補正部１１００、補正したスペクトル情報を統合する信号統合部１１０１、統合スペクトル情報をコヒーレンス関数に基づいてフィルタリングするコヒーレンスフィルタ演算部１１０２、及びコヒーレンスでフィルタリングしたスペクトル情報をさらに音源相関係数に基づいて利得制御することにより雑音を抑圧する利得制御部１１０３を有する。

【０２１９】

周波数分析部２０１と音源方向推定部１０００は、第１０の実施形態で述べた通りである。スペクトル補正部１１００では、音源方向推定部から出力される音源方向補正係数を用いて複数チャネルのスペクトル情報を補正する。このスペクトル情報の補正は、音源方向からの到来音に対して相関係数を最大化する働きがある。音源方向を（θｏ，φｏ）、音源相関係数をρ（θｏ，φｏ）、音源方向補正係数をＨｉ（ｋ，θｏ，φｏ）とすれば、スペクトル情報の補正は

【数４８】

$$X'i(k) = Xi(k)Hi(k,\theta o,\phi o) \qquad (11-1)$$

に従って行われる。ここでｉはチャネル番号、Ｘ’ｉ（ｋ）は補正後のスペクトル情報、Ｘｉ（ｋ）は補正前のスペクトル情報である。

【０２２０】

以降は、補正スペクトル情報Ｘ’ｉ（ｋ）　を用いて信号統合部１１０１でｉチャネルのスペクトル情報に統合し、この統合スペクトル情報に対してコヒーレンスフィルタ演算と利得制御を行えばよい。利得制御のための利得としては、前述したようにρ（θｏ，φｏ）を使う。これ以降の処理は、第１０の実施形態と同様であるので省略する。

【０２２１】

次に、図３３を用いて本実施形態における処理の流れを説明する。

まず、初期設定として音源方向範囲を設定し、かつ第１０の実施形態で説明したように補正係数を発生する（ステップＳ２００）。次に、マイクロホン１０１－１～１０１－Ｍからの音声信号の入力（ステップＳ２０１）、周波数分析（ステップＳ２０２）、音源方向の推定（ステップＳ２０３）、スペクトル情報の補正（ステップＳ２０４）、スペクトル情報の統合（ステップＳ２０５）、コヒーレンス関数の演算（ステップＳ２０６）及び利得制御の処理（ステップＳ２０７）をディジタル化された音声信号がステップＳ２０１においてフレーム単位で入力される毎に繰り返し行う。

【０２２２】

（第１２の実施形態）

次に、本発明の第１２の実施形態について説明する。これまでに述べてきた修正相互相関係数の計算においては、式（３－１３）に示したように、相互相関の正規化の際に入力スペクトル情報のパワーの幾何平均を用いていたが、本実施形態では幾何平均の代わりに入力スペクトル情報を統合して得られる統合スペクトル情報のパワーを用いる場合について述べる。

【０２２３】

複数チャネルの信号をビームフォーマなどによって統合する際には、ビームフォーマの働

きによって方向性の雑音などが抑圧されている場合がある。このような場合、相互相関または修正相互相関係数による利得制御においては、既に抑圧されている分を考慮して軽めに利得制御した方がよい。本実施形態で説明する利得係数を用いると、抑圧された分が考慮されて利得制御が適正化できる。

【０２２４】

本実施形態に係る音声信号処理装置は、図３４に示すように空間的に離れて配置された複数のマイクロホン１０１－０～１０１－Ｍから出力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部２０１と，複数のスペクトル情報から目的音の活性度に相当する値である利得係数を計算する修正利得係数計算部２０００Ａとからなる。

【０２２５】

修正利得係数計算部２０００Ａは、クロス・パワースペクトル計算部２００１、コヒーレンス関数計算部２００２、パワー情報計算部２００３、信号統合部２００４、統合信号パワースペクトル計算部２００５及び利得係数計算部２００６からなる。

【０２２６】

クロス・パワースペクトル計算部２００１は、スペクトル情報から入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する。

【０２２７】

コヒーレンス関数計算部２００２は、複数チャネル間のクロススペクトルと各チャネルのパワースペクトルからコヒーレンス関数を算出する。

【０２２８】

パワー情報計算部２００３は、複数チャネルのパワースペクトルから入力音声信号のチャネル間の信号パワーに関するパワー情報を算出する。

【０２２９】

信号統合部２００４は、複数のスペクトル情報を統合して１チャネルの統合スペクトル情報を生成する。

【０２３０】

統合信号パワースペクトル計算部２００５は、統合スペクトル情報のパワースペクトルを計算する。

【０２３１】

利得係数計算部２００６は、クロススペクトルをコヒーレンス関数とパワー情報に基づいて重み付けし，重み付けたクロススペクトルをさらに統合信号パワースペクトルに基づいて正規化して得られる利得係数を計算する。

【０２３２】

周波数分析部２０１、クロス・パワースペクトル計算部２００１、コヒーレンス関数計算部２００２、パワー情報計算部２００３及び信号統合部２００４は第１０の実施形態と同様であるため、説明を省略する。

【０２３３】

統合信号パワースペクトル計算部２００５では、統合スペクトル情報のパワースペクトルが計算される。例えば、統合スペクトル情報をＺ（ｆ）として、統合処理が２チャネルの信号の加算平均Ｚ（ｆ）＝｛Ｘ１（ｆ）＋Ｘ２（ｆ）｝／２であったとすると、Ｚ（ｆ）のパワースペクトルは、

【数４９】

$$Wzz(f) = \overline{|Z(f)|^2} \qquad (12\text{-}1)$$

で求められる。Ｚ（ｆ）が異なる係数を持つビームフォーマから得られる統合信号であったとしても，同様である。

【０２３４】

利得係数計算部２００６で計算される利得係数αは、相互相関係数の代わりに利得制御に

用いる係数であり、Ｍ＝２の場合は次式により計算できる。

【数５０】

$$\Psi aij(f) = \frac{\gamma ij(f)2}{(1-\gamma ij(f)2)\left|Wij(f)\right|} \qquad (12\text{-}2)$$

$$\Psi bij(f) = \frac{\gamma ij(f)2\, pij(f)}{(1-\gamma ij(f)2)\left|Wij(f)\right|} \qquad (12\text{-}3)$$

$$\sigma = \frac{\sum_{f=L_1}^{L_2} Wij(f)\left|\Psi bij(f)\right|}{\sum_{f=L_1}^{L_2} Wzz(f)\left|\Psi aij(f)\right|} \qquad (12\text{-}4)$$

【０２３５】

なお、式（１２－２），（１２－３）はそれぞれ式（３－１２），（３－１３）と同じである。以上の計算により得られる利得係数σは、Ｗｚｚのパワーにおいて既に抑圧された雑音の分が除かれているので、利得を過小に計算する可能性が低くなり、性能を改善できる可能性がある。利得係数計算部２００６は、パワー比とコヒーレンス関数で重み付けた利得係数であるという意味の修正利得係数αを出力する。

【０２３６】

次に、図３５を用いて本実施形態における処理の流れについて説明する。マイクロホン１０１－１～１０１－Ｍからの音声信号の入力（ステップＳ３０１）及び周波数分析（ステップＳ３０２）の後、修正理作係数計算部２０００Ａにおいてクロススペクトルとパワースペクトルの計算（ステップＳ３０３）、、パワー情報の計算（ステップＳ３０４）、コヒーレンス関数の計算（ステップＳ３０５）、信号統合（スペクトル情報の統合）（ステップＳ３０６）、統合スペクトル情報（統合信号）のパワースペクトルの計算（ステップＳ３０７）及び修正利得係数の計算（Ｓ３０８）をディジタル化された音声信号がステップＳ３０１においてフレーム単位で入力される毎に繰り返し行う。

【０２３７】

（第１３の実施形態）

図３６に、本発明の第１３の実施形態に係る音声信号処理装置の構成を示す。本実施形態は式（１２－３）においてパワー情報ｐｉｊ（ｆ）を全て１とおいて、パワー情報を用いないようにした例であり、修正利得係数計算部２０００Ｂでは図３４中に示したパワー情報計算部２００３が除去されている。

【０２３８】

（第１４の実施形態）

次に、本発明の第１４の実施形態として、第１２の実施形態で求めた利得係数に基づいて雑音を抑圧し、目的音声を強調する音声強調処理装置について説明する。

【０２３９】

本実施形態に係る音声信号処理装置は、図３６に示すように空間的に離れて配置された複数のマイクロホン１０１－０～１０１－Ｍから出力される複数チャネルの入力音声信号を周波数分析してＭチャネルのスペクトル情報を生成する周波数分析部２０１及びスペクトル情報から目的音の活性度に相当する値である利得係数を計算する図３４に示した修正利得係数計算部２０００Ａに加えて、利得制御部２１０１及びコヒーレンスフィルタ演算部２１０２を有する。

【０２４０】

利得制御部２１０１は、修正利得係数計算部２０００Ａで計算した利得係数に基づいて、

修正利得係数計算部２０００Ａ内の信号統合部２００４で得られた統合スペクトル情報に対して利得の制御を行う。コヒーレンスフィルタ演算部２１０２は、修正利得係数計算部２０００Ａ内のコヒーレンス関数計算部２００２で得られたコヒーレンス関数に基づいて、利得制御部２１０１から出力されるスペクトル情報をフィルタリングする。

【０２４１】

次に、図３８を用いて本実施形態における処理の流れを説明する。

マイクロホン１０１－１～１０１－Ｍからの音声信号の入力（ステップＳ４０１）及び周波数分析（ステップＳ４０２）の後、修正利得係数計算部２０００Ａにおいてクロススペクトル及びパワースペクトルの計算（ステップＳ４０３）、パワー情報の計算（ステップＳ４０４）、コヒーレンス関数の計算（ステップＳ４０５）、スペクトル情報の統合（ステップＳ４０６）、統合スペクトル情報のパワースペクトルの計算（ステップＳ４０７）及び利得係数の計算（ステップＳ４０８）を行う。次に、計算された利得係数に基づく利得制御処理（ステップＳ４０９）とコヒーレンスフィルタ演算の処理（ステップＳ４１０）を行う。以上のステップＳ４０１～Ｓ４１０の処理をディジタル化された音声信号がステップＳ４０１においてフレーム単位で入力される毎に繰り返し行う。

【０２４２】

（第１５の実施形態）

図３９に、本発明の第１５の実施形態に係る音声信号処理装置の構成を示す。本実施形態は式（１０－６）のパワー情報ｐｉｊ（ｆ）を１とおいて、パワー情報を用いないようにした例であり、修正利得係数計算部２０００Ｂでは図３７中に示したパワー情報計算部２００３が除去されている。

【０２４３】

（第１６の実施形態）

次に、第１２の実施形態で説明した利得係数を用いて音源方向を推定する本発明の第１６の実施形態について説明する。本実施形態に係る音声信号処理装置は、図４０に示すようにマイクロホン１０１－１～１０１－Ｍからの複数（Ｍ）チャネルの入力音声信号を周波数分析して周波数成分であるスペクトル情報に変換する周波数分析部２０１と、該スペクトル情報から音源方向を推定する音源方向推定部３０００からなる。音声分析部２０１の処理は、第２の実施形態（図６）で説明した通りである。

【０２４４】

音源方向推定部３０００は、クロス・パワースペクトル計算部３００１、コヒーレンス関数計算部３００２、補正係数発生部３００３、クロス・パワースペクトル補正部３００４、パワー情報計算部３００５、仮想統合パワースペクトル計算部３００６、仮想方向利得係数計算部３００７及び音源方向検出部３００８を有する。以下、音源方向推定部３０００の各部について説明する。

【０２４５】

クロス・パワースペクトル計算部３００１は、周波数分析部２０１により得られたスペクトル情報から各チャネルの入力音声信号のチャネル毎のパワースペクトル及びチャネル間のクロススペクトルを算出する。

【０２４６】

コヒーレンス関数計算部３００２は、複数チャネル間のクロススペクトルと各チャネルのパワースペクトルから入力音声信号の複数チャネル間のコヒーレンス関数を算出する。

【０２４７】

補正係数発生部３００３は、信号の仮想的な到来方向である仮想方向から到来する信号がチャネル間で一致するように補正するための係数を複数の仮想方向から成る１組の仮想方向群に対応して発生する。

【０２４８】

クロス・パワースペクトル補正部３００４は、補正係数発生部３００３で発生された補正係数に基づいてクロススペクトルとパワースペクトルを補正し、補正クロススペクトルと補正パワースペクトルを生成する。

【０２４９】
パワー情報計算部３００５は、補正クロススペクトルと補正パワースペクトルに基づいて入力音声信号のチャネル間の信号パワーに関するパワー情報を算出する。
【０２５０】
仮想統合パワースペクトル計算部３００６は、周波数分析部２０１で得られた複数チャネルのスペクトル情報を補正係数発生部３００３で発生された補正係数により補正してから統合して得られる統合スペクトル情報に対するパワースペクトルを、クロス・パワースペクトル補正部３００４で得られた補正クロススペクトルと補正パワースペクトルに基づいて計算する。
【０２５１】
仮想方向利得係数計算部３００７は、クロス・パワースペクトル補正部で得られた補正クロススペクトルに対し、コヒーレンス関数とパワー情報に基づいて重み付けを行い、さらに仮想統合パワースペクトルに基づいて正規化を行った後、１組の仮想方向に対応した利得係数を求める。
【０２５２】
音源方向検出部３００８は、仮想方向利得係数計算部３００７において計算された仮想方向毎の利得係数に基づいて音源方向を検出し出力すると同時に、検出した音源方向に対応した利得係数の値を音源利得係数として、また音源方向に対応した補正係数を音源方向補正係数として出力する。
【０２５３】
ここで、周波数分析部２０１、クロス・パワースペクトル計算部３００１、コヒーレンス関数計算部３００２、補正係数発生部３００３、クロス・パワースペクトル補正部３００４及びパワー情報計算部３００５の処理については、第１０の実施形態に係る相関係数に基づく音源方向推定と同一であるので、詳細な説明を省略する。
【０２５４】
第１２～第１４の実施形態における利得係数の計算においては、利得係数αの式の分母の値を求める際、複数チャネルのスペクトル情報を統合してそのパワースペクトルを求めている。これに対し、本実施形態ではスペクトル情報の段階での統合は行わず、パワースペクトルとクロススペクトルを補正して、統合信号のパワーを直接求める。これは実際に信号を統合してからパワーを求めるよりも、計算量と記憶領域の点で有利である。すなわち、スペクトル情報を統合してからパワーを求めると、仮想方向毎にパワースペクトル推定のための時間平均化が必要となるが、本実施形態によればこれを避けることが可能である。
【０２５５】
まず、各チャネルのスペクトル情報に補正係数発生部３００３で発生した補正係数を乗じてから信号を統合したと仮定し、その処理式をここでは加算平均とする。このときの統合信号Ｚ（ｆ）は、
【数５１】

$$Z(f,\theta,\phi) = \frac{X1(f)H1(f,\theta,\phi) + X2(f)H2(f,\theta,\phi)}{2} \quad (16\text{-}1)$$

と表せる。もちろん他の統合方法でもよい。
【０２５６】
このとき、統合信号Ｚ（ｆ）のパワースペクトルは、
【数５２】

$$W_{zz}(f,\theta,\phi) = \overline{|Z(f,\theta,\phi)|^2}$$
$$= \frac{\overline{|X1H1|^2} + \overline{|X2H2|^2} + 2\overline{X2^*H2^*}}{4} \qquad (16\text{-}2)$$
$$= \frac{W11|H1|^2 + W22|H2|^2 + 2W12H1H2^*}{4}$$

となる。ここで、式（１６－２）では添え字は省略してある。また、上線は時間平均を表す。従って、クロススペクトルとパワースペクトルを一度求めておけば、後は補正係数を式（１６－２）に従って乗じるだけで、仮想方向（θ，φ）に対応した利得係数α（θ，φ）の分母の値が求まる。

【０２５７】

仮想方向利得係数計算部３００７では，まずクロス・パワースペクトル補正部３００４で求められた仮想方向に対応した補正クロススペクトル

【数５３】

$$W12(f,\theta,\phi) = W12(f)H1(f,\theta,\phi)H2^*(f,\theta,\phi) \qquad (16\text{-}3)$$

に対し、コヒーレンス関数γ^2（ｆ）と補正したパワー情報ｐｉｊ（ｆ，θ，φ）に基づいて重み付けを行う。さらに、仮想方向利得係数計算部３００７では、仮想統合パワースペクトル計算部３００６で求められた仮想的な統合信号パワーＷｚｚ（ｆ，θ，φ）に対し、コヒーレンス関数γ^2（ｆ）に基づいて重み付けを行い、先の式（２－３）により仮想方向に対応した利得係数である仮想方向利得係数α（θ，φ）を求める。

【０２５８】

音源方向検出部３００８の処理は、第１０の実施形態における音源方向推定部１０００７と同様でよい。この場合，音源方向検出部３００８が検出した音源方向に相当する利得係数α（θｏ，φｏ）を音源方向利得係数と呼ぶことにする。さらに，音源方向検出部３００８は第１０の実施形態１と同様、音源方向（θｏ，φｏ）の他に、音源方向の補正係数Ｈ１（θｏ，φｏ）を音源方向補正係数として出力する。以上により、利得係数に基づいて音源方向を推定することができる。

【０２５９】

次に、図４１を用いて本実施形態における処理の流れを説明する。

まず，初期設定として音源方向の範囲を設定する（ステップＳ５００）。次に補正係数の生成（ステップＳ５０１）、マイクロホン１０１－１～１０１－Ｍからの音声信号の入力（ステップＳ５０２）、周波数分析（ステップＳ５０３）、クロススペクトルとパワースペクトルの計算（ステップＳ５０４）及びコヒーレンス関数の計算（ステップＳ５０５）を順次行う。次に、スペクトル補正（ステップＳ５０６），パワー情報の計算（ステップＳ５０７），仮想統合パワースペクトルの計算（ステップＳ５０８）及び仮想方向利得係数の計算（ステップＳ５０９）を全ての仮想方向について繰り返し行い、最後に音源方向の検出を行う（ステップＳ５１０）。ステップＳ５０２～Ｓ５１０の処理は、ディジタル化された音声信号がステップＳ５０２においてフレーム単位で入力される毎に繰り返し行われる。

【０２６０】

（第１７の実施形態）

次に、本発明の第１７の実施形態として、第１６の実施形態で説明した利得係数に基づく音源方向推定処理により推定した音源方向を用い、目的音が移動した場合でもその方向を追尾して音声強調を行うことにより、音声強調を安定に行うことができるようにするための処理について説明する。

【０２６１】
本実施形態の音声信号処理装置は、図４２に示すように周波数分析部２０１、音源方向推定部３０００、周波数分析部２０１からの複数チャネルのスペクトル情報を音源方向補正係数に基づいて補正するスペクトル情報補正部３１００、補正したスペクトル情報を統合する信号統合部３１０１、統合スペクトル情報をコヒーレンス関数に基づいてフィルタリングするコヒーレンスフィルタ演算部３１０２、及びフィルタリングしたスペクトル情報をさらに音源利得係数に基づいて利得制御することにより雑音を抑圧する利得制御部３１０３を有する。
【０２６２】
ここで、周波数分析部２０１、音源方向推定部３０００及びスペクトル情報補正部３１００は第１６の実施形態と同様であり、またコヒーレンスフィルタ演算部３００２は第１１の実施形態と同様である。
【０２６３】
信号統合部３１０１は、音源方向推定部３０００内の図４０に示した仮想統合信号パワースペクトル計算部３００６において行っている仮想統合信号パワースペクトルの計算の際に仮定している信号の統合と同じ統合の式を用いて、補正されたスペクトル情報の統合を行う。すなわち、仮想統合信号パワースペクトル計算部３００６において２チャネルの加算平均を想定していれば、信号統合部３１０１でのスペクトル情報の統合にも加算平均を用いる。この場合、音源方向推定部３０００で得られた音源方向を（θｏ，φｏ）とし、これに対応する補正係数をＨ１（ｆ，θｏ，φｏ）、Ｈ２（ｆ，θｏ，φｏ）とする。この場合、音源方向に対応して補正した統合信号Ｚ（ｆ，θｏ，φｏ）は、次式のようになる。
【０２６４】
【数５４】

$$Z(f,\theta o,\phi o)=\frac{X1(f)H1(f,\theta o,\phi o)+X2(f)H2(f,\theta o,\phi o)}{2} \quad (17-4)$$

【０２６５】
Ｘ１（ｆ），Ｘ２（ｆ）は、周波数分析部で得られていた各チャネルのスペクトル情報である。
【０２６６】
利得制御部３１０３は、音源方向推定部３０００で推定された音源方向に対応する利得係数α（θｏ，φｏ）を用い、これに基づいて式（１６－１）に従って補正した統合信号Ｚ（ｆ，θｏ，φｏ）の振幅を制御する。制御の方法としては、単純な比例のほか、第１実施形態で述べた通りの方法を用いればよい。
【０２６７】
次に、図４３を用いて本実施形態における処理の流れを説明する。
まず、初期設定として音源方向の範囲を設定し、さらに補正係数を発生する（ステップＳ６００）。次にマイクロホン１０１－１～１０１－Ｍからの音声信号の入力（ステップＳ６０１），周波数分析（ステップＳ６０２）、音源方向の推定（ステップＳ６０３）、スペクトル情報の補正（ステップＳ６０４）、スペクトル情報の統合（ステップＳ６０５），コヒーレンスフィルタ演算（ステップＳ６０６）及び利得制御（ステップＳ６０７）をディジタル化された音声信号がステップＳ６０１においてフレーム単位で入力される毎に繰り返し行う。
【０２６８】
（第１８の実施形態）
次に、本発明の第１８の実施形態として、適応フィルタを用いて入力音声信号のチャネル間の差を適応的に補正し、目的音の方向が想定とわずかにずれる場合のほか，反射による影響も低減する音声信号処理装置について説明する。第１１の実施形態及び第１７の実施

形態で述べた音源方向推定に基づく追尾型の安定化方法は，目的音のずれには効果的であるが、反射などによるチャネル間の信号のずれには効果が小さい。反射の状況は、受音位置により異なることが多いため、チャネル間のずれを生じる原因になる。そこで、本実施形態では適応フィルタを用いた安定化方法を用いる。

【０２６９】

適応フィルタを用いた安定化方法については，既に第８の実施形態において述べている。第８の実施形態では相関係数による目的信号活性度を求める前に、相関係数により適応フィルタを制御してチャネル間の補正を行っている。この場合は、相関係数を求める際の時間遅れがあるため、この遅延よりもゆっくりと変化する外乱要因、すなわちマイクロホンのバイアス電圧変化や経年変化等による感度変化などに効果がある。これに対し、本実施形態は反射波がある場合や頻繁に目的音が動く場合など、入力音声信号のチャネル間のずれの状況の変化が比較的速い場合に効果がある。

【０２７０】

本実施形態に係る音声信号処理装置は、図示しない空間的に離れて配置された複数のマイクロホンと、該マイクロホンから入力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部に加えて、図４４に示すように周波数分析部からの複数チャネルのスペクトル情報を入力として目的信号活性度を推定する安定化目的信号活性度推定部４０００からなる。

【０２７１】

安定化目的信号活性度推定部４０００は、入力音声信号のチャネル間の修正した相互相関係数である第１の修正相互相関係数を計算する第１の修正相互相関係数計算部４００１と、第１の修正相互相関係数に基づき複数チャネルのスペクトル情報の間の差を適応的に補正して補正スペクトル情報を得る適応スペクトル補正部４００２と、補正スペクトル情報から第２の修正相互相関係数を計算する第２の修正相互相関係数計算部４００３からなる。周波数分析部、第１及び第２の修正相互相関係数計算部４００１，４００３は、既に述べたものと同一の処理を行う。

【０２７２】

適応スペクトル補正部４００２は、図４６に示すように周波数分析部で得られた各チャネルのスペクトル情報の間の伝達関数を適応フィルタ４１０３によって同定し、その差分を補正する。このとき第１の修正相互相関係数計算部４００１から出力される修正相互相関係数に基づいて適応フィルタ４１０３を制御し、目的音が到来している間のみ適応フィルタ４１０３を更新することによって雑音への適応を避け，目的音に関する伝達関数のみ推定するようにする。

【０２７３】

第１修正相互相関係数の計算は、クロススペクトルとパワースペクトルを求める際の時間平均に起因する時間遅れがあるので、第１の修正相互相関係数計算部４００１から出力される相関係数は，現時点からその時間遅れ分だけ過去の入力データに基づいて計算されたものである。従って、適応フィルタ４１０３に入力するスペクトル情報と相関係数を同期させるため、遅延回路４１０１，４１０２により相関係数計算と同じだけ遅延させたスペクトル情報を用いるようにする。

【０２７４】

時間遅れの値は、クロス・パワースペクトルの平均化に要する時間長をＴとすると、Ｔ／２である。フレーム数で見ると、平均化フレーム数をＴａとしてＴａが偶数の場合、遅れはＴａ／２フレームであるが、Ｔａが奇数の場合は（Ｔａ－１）／２で計算できる。Ｔａは奇数の方が望ましい。

【０２７５】

適応フィルタ４１０３を用いた演算は、第８の実施形態で既に述べたように例えば周波数領域のＬＭＳ適応フィルタを用いて行い、同定されたフィルタＷ（ｆ）を参照信号に用いたチャネル側のスペクトル情報に乗じて補正する。第２の修正相互相関係数計算部４００３は、適応スペクトル修正部４００２で補正されたスペクトル情報から第２の修正相互相

関係数を計算して出力する。

【０２７６】

次に、図４５を用いて本実施形態における処理の流れを説明すると、まず入力音声信号のチャネル間の修正した相互相関係数である第１の修正相互相関係数を計算し（ステップＳ７０１）、これに基づいて各チャネルのスペクトル情報の間の伝達関数の差分を補正することにより適応スペクトルの補正を行い（ステップＳ７０２）、最後に補正された適応スペクトル情報から第２の修正相互相関係数を計算して目的信号活性度として出力する（ステップＳ７０３）。

【０２７７】

本実施形態では、適応の制御とフィルタの更新を同期したデータにより行うため、修正相互相関係数計算を時間遅れを考慮して２回行っている。これにより、状況がすばやく変化する場合にも、雑音の影響を抑えてチャネル間の差を適応的に正確に補正することが可能となる。

【０２７８】

（第１９の実施形態）

第１８の実施形態では、修正相互相関係数計算に関して適応的に安定化させる場合を述べたが、修正相互相関係数の代わりに、第１２の実施形態で述べた利得係数の計算で同様な処理を行うことが可能である。

【０２７９】

本実施形態に係る音声信号処理装置は、図示しない空間的に離れて配置された複数のマイクロホンと、該マイクロホンから入力される複数チャネルの入力音声信号を周波数分析して複数チャネルのスペクトル情報を生成する周波数分析部に加えて、図４５に示すように周波数分析部からの複数チャネルのスペクトル情報を入力として目的信号活性度を推定する安定化目的信号活性度推定部５０００からなる。

【０２８０】

安定化目的信号活性度推定部５０００は、複数チャネルのスペクトル情報から目的音の活性度に相当する値である第１の修正利得係数を計算する第１の修正利得係数計算部５００１と、第１の修正利得係数に基づき複数チャネルのスペクトル情報の間の差を適応的に補正して補正スペクトル情報を得る適応スペクトル補正部５００２と、補正スペクトル情報から第２の修正利得係数を計算する第２の修正利得係数計算部５００３からなる。第１及び第２の修正利得係数計算部５００１，５００３は、第１２の実施形態で述べたものと同一の処理を行う。

【０２８１】

ところで、第１、第２、第４、第６、第１１、第１４及び第１７の各実施形態においては、相関係数または利得係数の算出結果を用いて音声強調処理を行っている。これら第１、第２、第４、第６、第１１、第１４及び第１７の各実施形態においても、図４６で説明した同様に、相関係数または利得係数の計算による時間遅延を考慮して、相関係数または利得係数と入力のスペクトル情報が同期するように、相関係数または利得係数計算時の入力のスペクトル情報を遅延させて処理することが望ましい。この場合の遅延フレーム数は、図４６で説明したと同様に、クロススペクトルとパワースペクトル推定のための時間平均化フレーム数の半分の値に選ばれる。このような遅延処理の導入は自明のことであるため、第１、第２、第４、第６、第１１、第１４及び第１７の各実施形態の説明では省略されている。

【０２８２】

なお、本発明は上記実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、上記実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせてもよい。

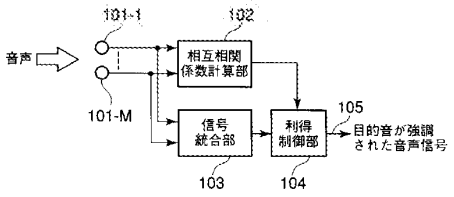
【０２８３】

【発明の効果】
以上説明したように、本発明によれば突発雑音や拡散性雑音を含む実環境雑音下で雑音を抑圧することが可能となり、雑音環境下において目的音声が到来しているか否かを高精度で検出したり、ハンズフリー通話や音声認識の前処理に好適な音声信号処理を行うことができる。

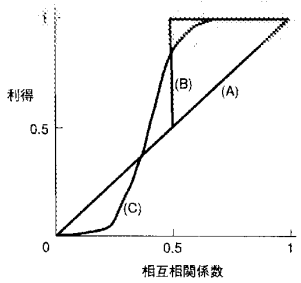
【図面の簡単な説明】
【図１】本発明の第１の実施形態に係る音声信号処理装置の構成を示すブロック図
【図２】同実施形態における統合音声信号に対する利得制御に用いる種々の関数を示す図
【図３】同実施形態における音声信号処理手順を示すフローチャート
【図４】同実施形態におけるマイクロホンの配置例を示す図
【図５】同実施形態に係る信号統合部に適応ビームフォーマを用いた音声信号処理装置の構成を示すフローチャート
【図６】本発明の第２の実施形態に係る音声信号処理装置の構成を示すフローチャート
【図７】同実施形態における音声信号処理手順を示すフローチャート
【図８】本発明の第３の実施形態に係る音声信号処理装置の構成を示すブロック図
【図９】同実施形態における音声信号処理手順を示すフローチャート
【図１０】本発明の第４の実施形態に係る音声信号処理装置の構成を示すブロック図
【図１１】同実施形態における音声信号処理手順を示すフローチャート
【図１２】同実施形態における検出処理手順を示すフローチャート
【図１３】同実施形態における検出処理の具体例を示す図
【図１４】本発明の第５の実施形態に係る音声信号処理装置の構成を示すブロック図
【図１５】同実施形態における音声信号処理手順を示すフローチャート
【図１６】本発明の第６の実施形態に係る音声信号処理装置の構成を示すブロック図
【図１７】同実施形態における音声信号処理手順を示すフローチャート
【図１８】本発明の第７の実施形態に係るマイクロホンの配置例を示す図
【図１９】同実施形態に係るマイクロホンの他の配置例を示す図
【図２０】図１９（Ｂ１）～（Ｂ４）の配置における到来方向を方位角と仰角を用いて表した図
【図２１】図１９（Ｂ１）～（Ｂ４）の配置における２つのマイクロホンの位相が一致する到来方向と２つのマイクロホンの感度が一致する到来方向の関係を示す図
【図２２】本発明の第８の実施形態に係る音声信号処理装置の構成を示すブロック図
【図２３】同実施形態におけるスペクトル補正部の構成を示すブロック図
【図２４】本発明の第９の実施形態に係る音声信号処理装置の構成を示すブロック図
【図２５】同実施形態における補正フィルタ学習指示部の構成を示すブロック図
【図２６】同実施形態におけるスペクトル補正部の構成を示すブロック図
【図２７】同実施形態におけるスペクトル補正部の処理手順を示すフローチャート
【図２８】本発明の第１０の実施形態に係る音声信号処理装置の構成を示すブロック図
【図２９】同実施形態における到来方向推定時の仮想点の設定について説明する図
【図３０】同実施形態における伝搬遅延の計算法について説明する図
【図３１】同実施形態における音声信号処理手順を示すフローチャート
【図３２】本発明の第１１の実施形態に係る音声信号処理装置の構成を示すブロック図
【図３３】同実施形態における音声信号処理手順を示すフローチャート
【図３４】本発明の第１２の実施形態に係る音声信号処理装置の構成を示すブロック図
【図３５】同実施形態における音声信号処理手順を示すフローチャート
【図３６】本発明の第１３の実施形態に係る音声信号処理装置の構成を示すブロック図
【図３７】本発明の第１４の実施形態に係る音声信号処理装置の構成を示すブロック図
【図３８】同実施形態における音声信号処理手順を示すフローチャート
【図３９】本発明の第１５の実施形態に係る音声信号処理装置の構成を示すブロック図
【図４０】本発明の第１６の実施形態に係る音声信号処理装置の構成を示すブロック図
【図４１】同実施形態における音声信号処理手順を示すフローチャート

【図４２】本発明の第１７の実施形態に係る音声信号処理装置の構成を示すブロック図
【図４３】同実施形態における音声信号処理手順を示すフローチャート
【図４４】本発明の第１８の実施形態に係る音声信号処理装置の構成を示すブロック図
【図４５】同実施形態における音声信号処理手順を示すフローチャート
【図４６】同実施形態における適応スペクトル補正部の構成を示すブロック図
【図４７】本発明の第１９の実施形態に係る音声信号処理装置の構成を示すブロック図
【符号の説明】
１０１－１～１０１－Ｍ…マイクロホン
１０２…相互相関係数計算部
１０３…信号統合部
１０４…利得制御部（調整部）
１０６…適応ビームフォーマ
２０１…周波数分析部
２０２…相互相関係数計算部
２０３…信号統合部
２０４…利得制御部（調整部）
３００…目的信号活性度計算部
３０１…クロス・パワースペクトル計算部
３０２…コヒーレンス関数計算部
３０３…パワー情報計算部
３０４…修正スペクトル計算部
３０５…重み付き相互相関係数計算部
４０１…検出処理部（判定部）
５０１…利得制御部（調整部）
６０１…コヒーレンスフィルタ演算部
７０１…無指向性マイクロホン
７０２，７１１，７１２…指向性マイクロホン
８００，９００…スペクトル補正部
８０１，９０４…適応フィルタ
８０２，９０１，９０２…補正フィルタ
９１０…補正フィルタ学習部

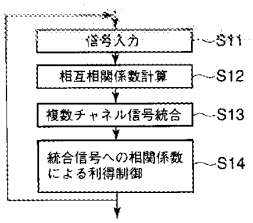
【図１】



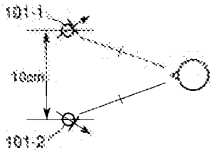
【図２】



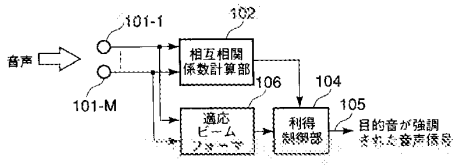
【図３】



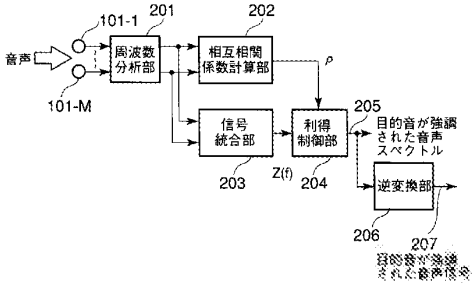
【図４】



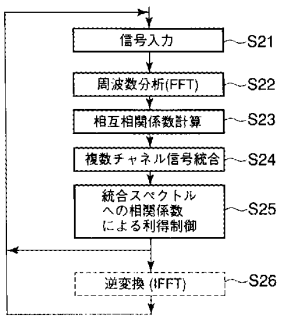
【図５】



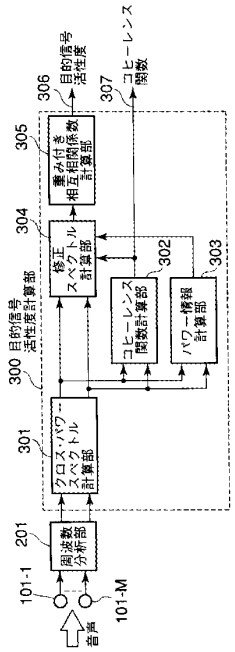
【図６】



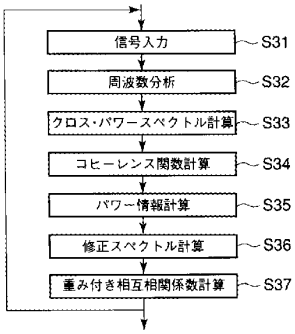
【図７】



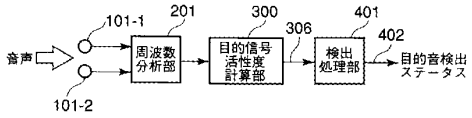
【図８】



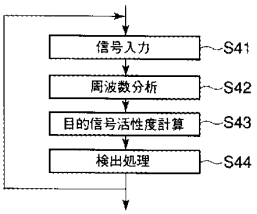
【図９】



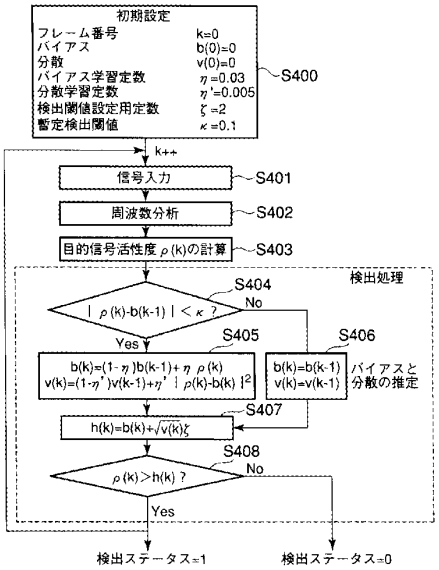
【図１０】



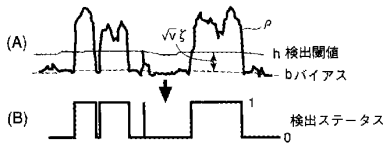
【図１１】



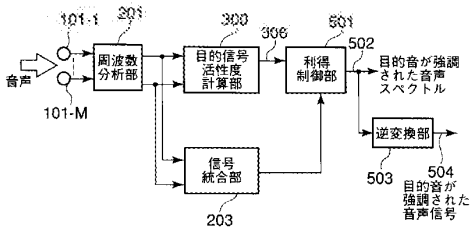
【図１２】



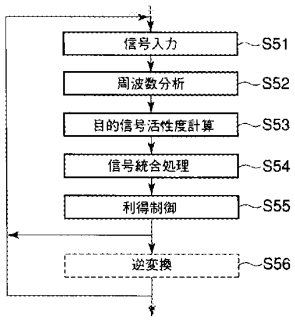
【図１３】



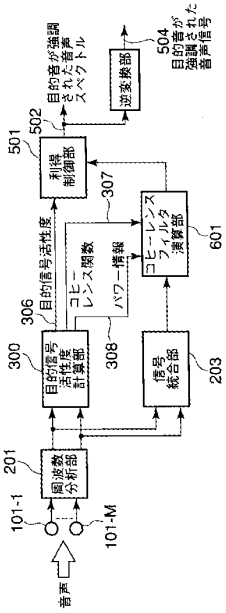
【図１４】



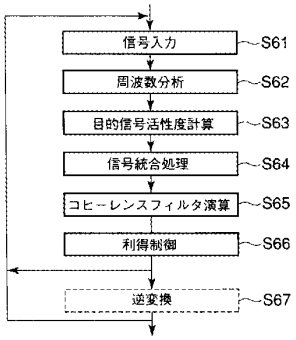
【図１５】



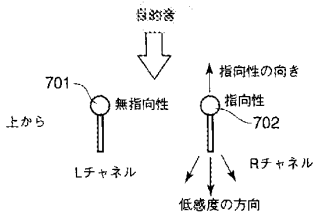
【図１６】



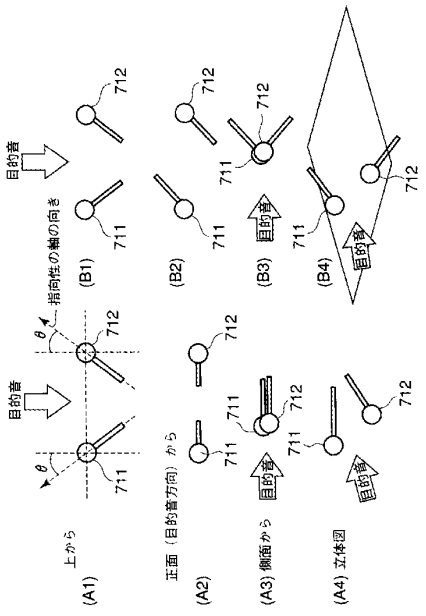
【図１７】



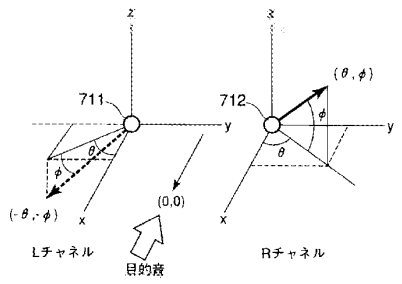
【図１８】



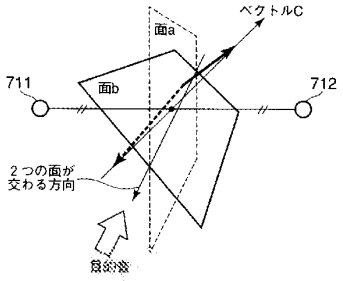
【図１９】



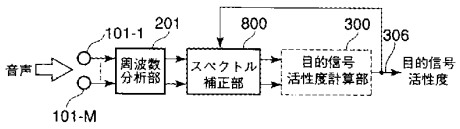
【図２０】



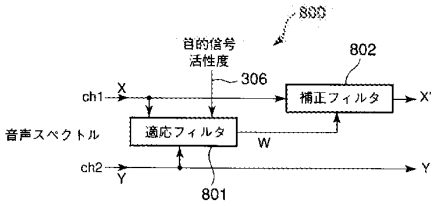
【図２１】



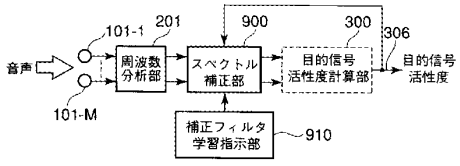
【図２２】



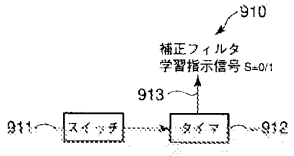
【図２３】



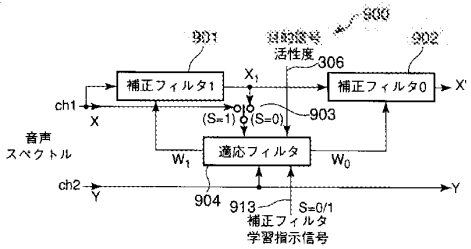
【図２４】



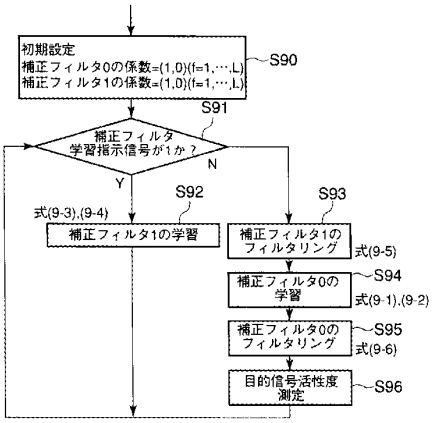
【図２５】



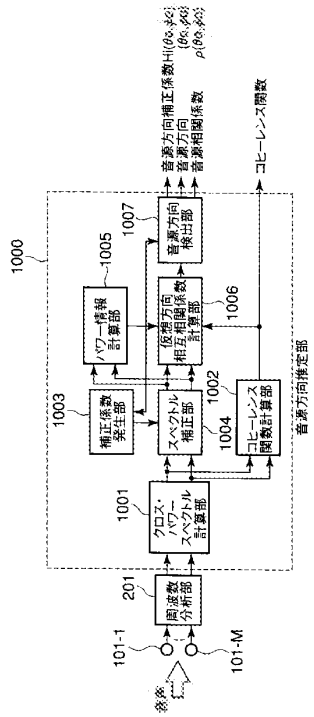
【図２６】



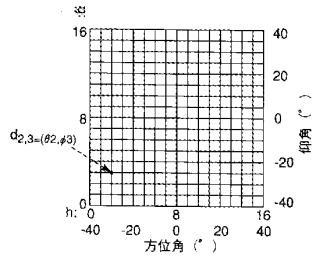
【図２７】



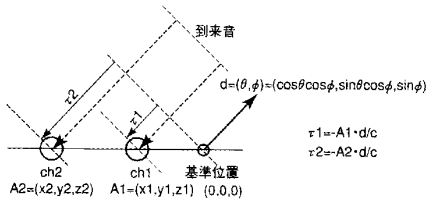
【図２８】



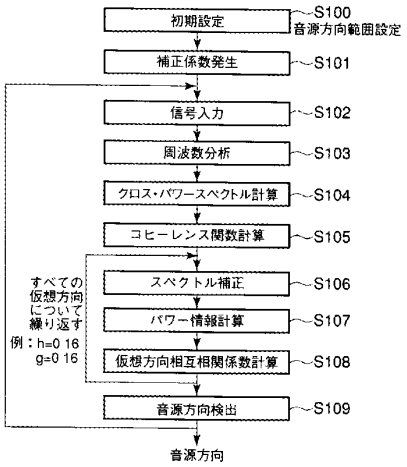
【図２９】



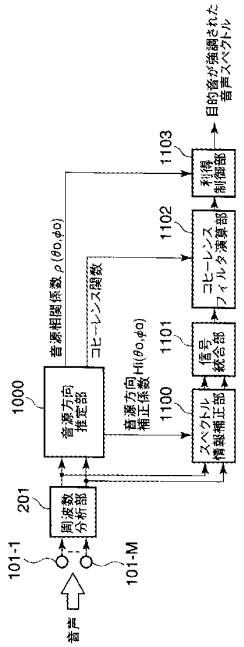
【図３０】



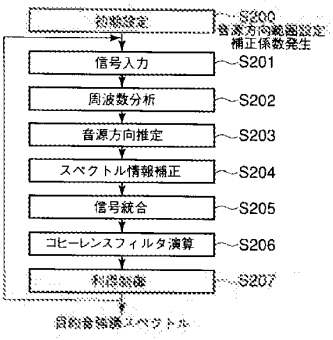
【図３１】



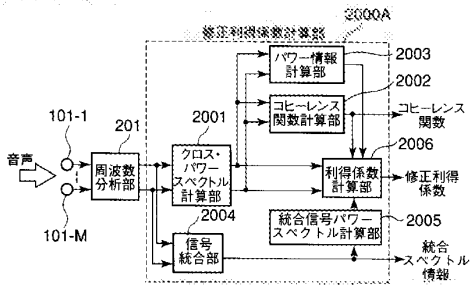
【図３２】



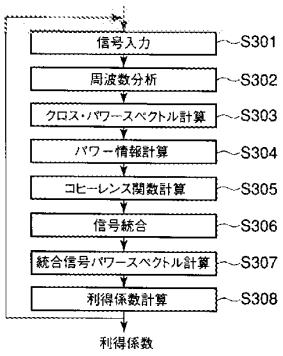
【図３３】



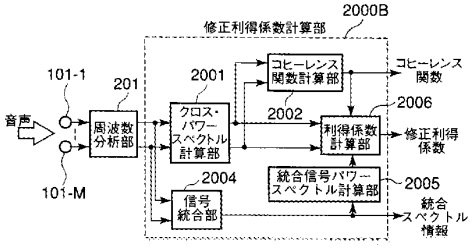
【図３４】



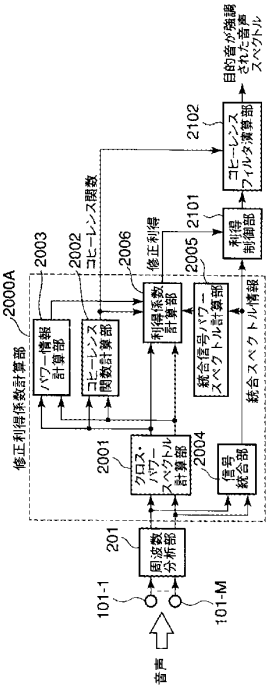
【図３５】



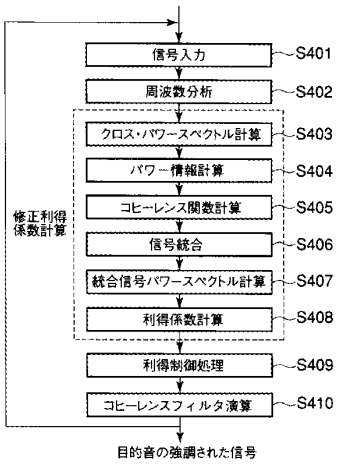
【図３６】



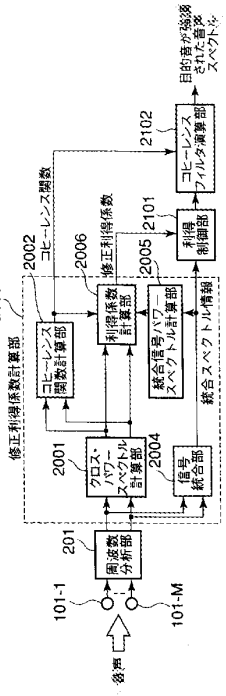
【図３７】



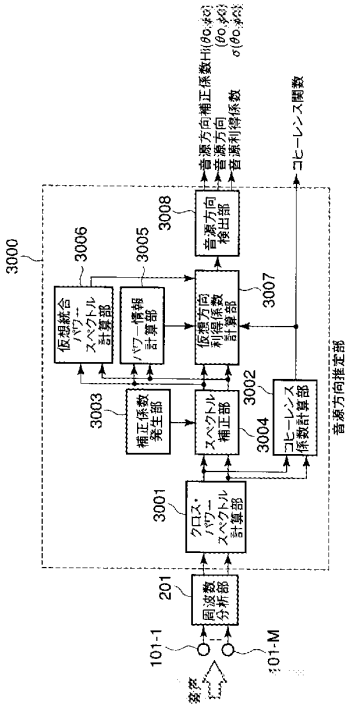
【図３８】



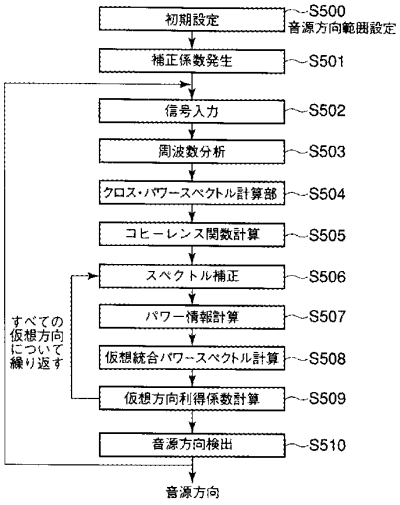
【図３９】



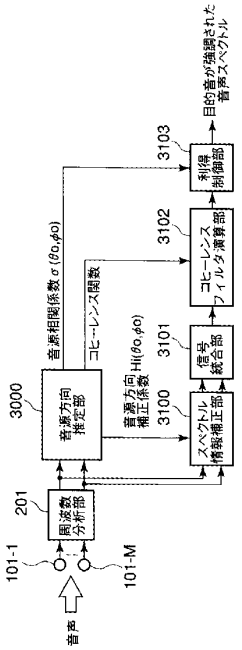
【図４０】



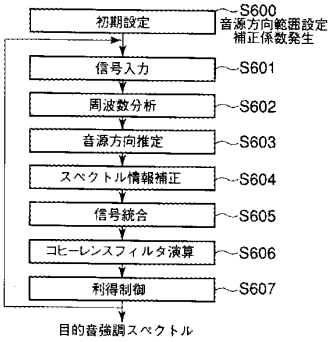
【図４１】



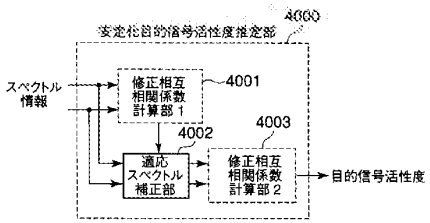
【図４２】



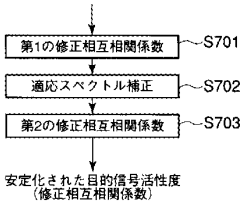
【図４３】



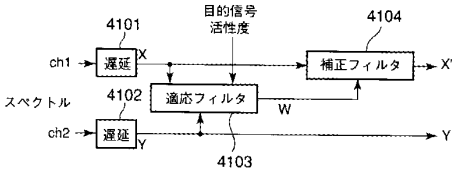
【図４４】



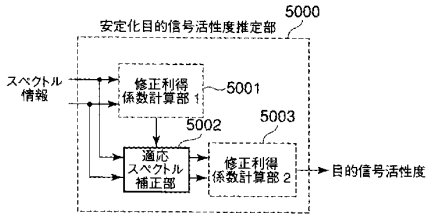
【図４５】



【図４６】



【図４７】



フロントページの続き

(51)Int.Cl.[7]	FI	テーマコード（参考）
H04R 1/40	G10L 3/00 511	
	G10L 7/02 A	

(74)代理人 100084618
弁理士 村松 貞男
(74)代理人 100092196
弁理士 橋本 良郎
(72)発明者 天田 豊
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 金澤 博史
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 永田 仁史
岩手県盛岡市青山４－１－２－４０３

Ｆターム(参考) 5D015 CC05 DD02 EE04 EE05
5D020 BB07

⑲日本国特許庁（JP）　⑪実用新案出願公開

⑫公開実用新案公報（U）　平2−97362

⑤1Int. Cl.[5]	識別記号	庁内整理番号	㊸公開　平成2年(1990)8月2日
B 65 H 35/07	L	8712−3F	
B 26 D 1/02	F	8509−3C	

審査請求　未請求　請求項の数 2　（全　頁）

㊹考案の名称　多目的テープカッター

㉑実　願　平1−6802

㉒出　願　平1(1989)1月23日

㉒考案者　細見　均　兵庫県神戸市北区北五葉2丁目11−1424

㉑出願人　細見　均　兵庫県神戸市北区北五葉2丁目11−1424

明　　細　　書

1．考案の名称　　多目的テープカッター

2．実用新案登録請求の範囲

（イ）本体（７）の側面にスライド式で取り出し自由な定規（１）を付けた。

（ロ）テープ（２）切断時において、本体（７）が動かない様固定するための重石を、脱着出来る文鎮（３）とした。

3．考案の詳細な説明

従来のテープカッターは、テープを切断するだけの物か、筆立てを付けただけで、機能的には関連性のない物が多かった。

この考案は、従来のテープカッターにまったく別の物を加え、なおかつ共に関連をもたした。

テープカッターの他に、定規（１）、文鎮（３）、クリップ等小物入れ（４）、筆立て（５）が付いており、その中の定規（１）と文鎮（３）がテープカッターと関連しています。

1323

実開 2 - 9736

Exhibit 1009
Page 1667 of 2198

定規（１）は、本体（７）側面にスライド式で取り付ける。

目盛り（８）は、通常の目盛りの他に、各種封筒等の幅を記入した目盛り（９）も書いておき、テープのカッター部（１０）を起点に定規（１）をスライドし、テープの長さを合わせて切断すれば、長すぎたり、短すぎたりしないテープカットが、可能です。



本体（７）より抜き出して、定規としてだけも使えます。

文鎮（３）は、テープカットの時に、本体（７）が動かない様、本体下部の内部にセットしておくが、文鎮として使用したい場合は、文鎮前部に取り付けたツマミ（１１）を引っ張り本体より引き抜く。

筆立て（５）は、本体側面に脱着出来るよう取り付ける。

クリップ等小物入れ（４）は、半円形で本体内部に収まる様にし、普段は、半円部が

Exhibit 1009
Page 1668 of 2198

本体内部に入っているが、使用したい時には、外側から指先で押せば、回転して収納部が表に出て来る。

小物入れは、箱型の引出し式にしてもよい。

また、小物入れ（４）の底部に磁石（１２）を付けておくことにより、金属性の小物がバラバラにならない。

４．図面の簡単な説明



第１図は、定規を取り付けた側の側面図

第２図は、上から見た図面

第３図は、第１図の反対側の側面図

第４図は、正面図

第５図は、筆立て取り付け時の拡大図

第６図は、立体図

第７図は、文鎮の拡大図

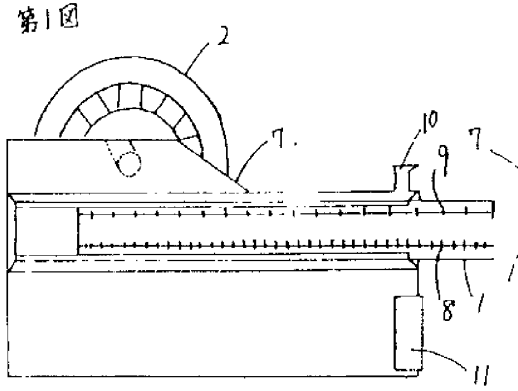
実用新案登録出願人　　細　見　　均

1325

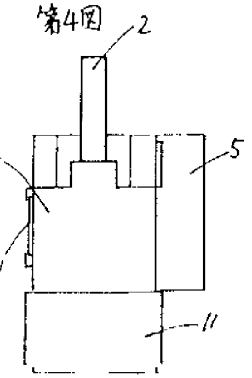
Exhibit 1009
Page 1669 of 2198

Exhibit 1009
Page 1670 of 2198

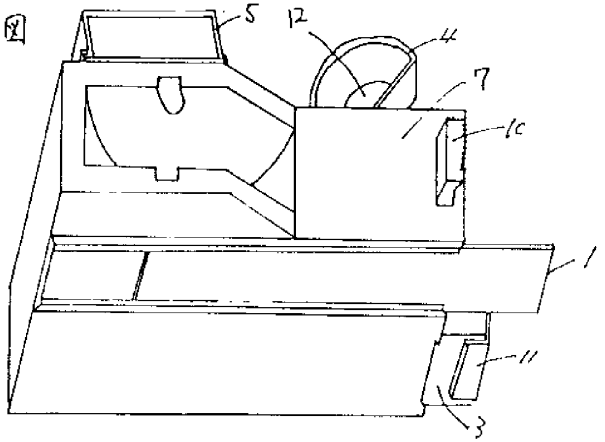
第1図



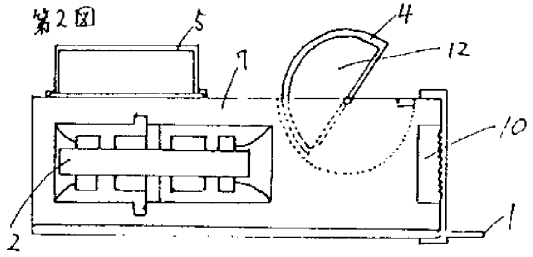
第4図



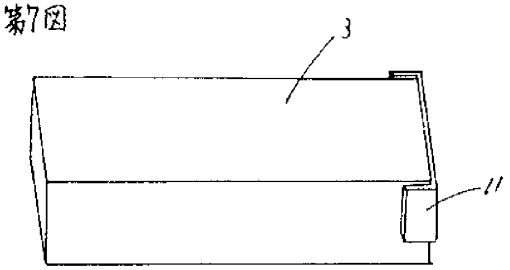
第6図



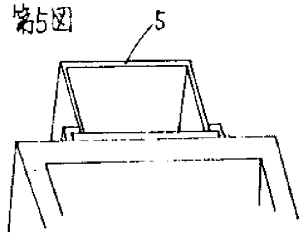
第2図



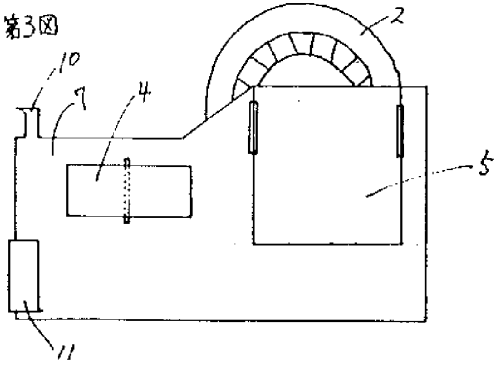
第7図



第5図



第3図



1326

実用新案登録出願人　細　見　均

(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.[7]
H04R 1/10

(11) 공개번호 특2003-0058432
(43) 공개일자 2003년07월07일

(21) 출원번호 10-2001-0088886
(22) 출원일자 2001년12월31일

(71) 출원인 삼성물산 주식회사
서울특별시 중구 태평로2가 310

케이.이.티 주식회사
경기도 안양시 만안구 안양7동 202-6 동일테크노타운 5602호

(72) 발명자 김창용
경기도 군포시 금정동 876 율곡아파트 333-1405 8/2

(74) 대리인 특허법인아람

심사청구 : 있음

(54) 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형이어폰

요약

하울링 및 에코현상을 방지할 수 있는 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형 이어폰이 개시된다. 음향 입/출력장치는 입력되는 음향신호를 출력하는 스피커, 주위의 음향을 집음하여 전기신호로 변환하는 집음마이크로폰, 사용자에 의한 발성유무를 확인하는 발성확인마이크로폰, 및 집음마이크로폰으로부터 입력되는 신호 중 스피커로부터 출력된 신호의 성분을 제거하며 발성확인마이크로폰에 의해 확인된 사용자의 발성유무에 따라 변환된 전기신호의 증폭율을 조절하는 회로부를 구비한다. 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형 이어폰은 주위 환경에서 발생된 소음이 마이크로폰을 통해 입력되는 것을 방지할 수 있을 뿐만아니라, 하울링 및 에코현상을 방지할 수 있게 된다.

대표도

도 3

색인어

마이크로폰, 이어폰, 스피커, 음향 입/출력 장치, 하울링, 에코

명세서

도면의 간단한 설명

도 1은 본 출원인에 의해 선출원된 음향 입/출력장치가 적용된 마이크로폰 일체형 이어폰을 개략적으로 나타내 보인 단면도,

도 2는 도 1의 마이크로폰 일체형 이어폰의 개략적인 블록도,

Exhibit 1009
Page 1671 of 2198

도 3은 본 발명에 따른 음향 입/출력장치가 적용된 마이크로폰 일체형 이어폰을 개략적으로 나타내 보인 단면도,

도 4는 도 3의 마이크로폰 일체형 이어폰의 측면도,

도 5는 도 3의 마이크로폰 일체형 이어폰의 개략적인 블록도, 그리고

도 6은 도 3의 마이크로폰 일체형 이어폰의 회로부를 개략적으로 나타낸 블록도이다.

* 도면의 주요부분에 대한 부호의 설명 *

1 : 귀 10 : 이어폰

11 : 본체 12 : 걸림턱

14 : 홀 15 : 지지부재

17 : 집음마이크로폰 18 : 발성확인 마이크로폰

19 : 스피커 20 : 회로부

22 : 제어신호 발생부 25 : 신호처리부

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형 이어폰에 관한 것으로서, 보다 상세하게는, 주위 환경에서 발생된 소음이 마이크로폰을 통해 입력되는 것을 방지할 수 있는 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형 이어폰에 관한 것이다.

주변 환경에서 발생된 소음의 영향을 적게 받으면서 음향을 입/출력 할 수 있는 장치가 개발되고 있다. 이러한 장치는 휴대용 오디오 플레이어, 휴대용 이동 단말기 등에 널리 채용되고 있으며, 이어폰과 마이크로폰이 분리된 형상을 이루고 있다. 일반적으로, 입/출력 장치의 이어폰은 귀에 삽입되어 이용되며, 마이크로폰은 구강부 근처에 위치되도록 형성된다. 입/출력 장치의 이어폰과 마이크로폰이 적절한 거리로 이격됨으로써, 스피커에서 출력되는 신호가 마이크로폰으로 역입력되어 신호가 왜곡되는 것이 방지된다.

그런데, 이러한 구조의 음향 입/출력 장치는 외부로부터 발생된 신호가 노출된 마이크로폰을 통해 입력될 수 있을 뿐만아니라, 이동중에는 마이크로폰과 인체의 구강부와의 이격거리가 변동될 수 있어 신호의 레벨이 변동될 수 있다. 또한, 마이크로폰과 이어폰을 상호 분리되도록 형성하여야 하기 때문에 구조가 복잡해지고, 부피가 커져 휴대상의 불편함이 있는 단점이 있다.

상기한 문제점을 개선하기 위하여 본 발명의 출원인은 음향 입/출력 장치 및 이를 적용한 마이크 일체형 이어폰에 대한 발명을 출원한 바 있다(출원번호 : 10-2001-0061260).

도 1은 본 출원인에 의해 선출원된 음향 입/출력장치가 적용된 마이크 일체형 이어폰을 개략적으로 나타내 보인 단면도이다. 도면을 참조하면, 마이크 일체형 이어폰(10)은 본체(11), 본체(11)내에 설치된 마이크로폰(17), 및 스피커(19)를 구비한다.

본체(11)는 귀에 일부를 삽입할 수 있도록 적절한 크기 및 형태로 구현된다. 또한, 본체(11)는 내부에 마이크로폰(17)과 스피커(19)를 장착할 수 있으면서 외부로부터 내부로 신호선을 삽입할 수 있도록 홈이 형성되어 있다.

마이크로폰(17)은 귀로부터 발생되는 신호를 입력받을 수 있도록 본체(11)내에 지지부재(15)에 의해 지지되어 설치되며, 신호처리부(13)가 구비된 회로기판과 접속되어 있다. 지지부재(15)는 음향신호의 입출이 용이하도록 일부가 관

Exhibit 1009
Page 1672 of 2198

통되게 형성된다. 또는, 지지부재(15)는 마이크로폰(17)의 일측과 본체가 상호 결합되도록 형성될 수도 있다.

스피커(19)는 진동판을 갖는 통상적인 것이 적용되며, 귀(1)의 고막을 향해 음향신호를 출력할 수 있도록 본체(11)내에 설치되어 있다.

참조부호 14는 외부기기로부터 입력되는 음향신호를 본체(11) 내부로 전송하고, 마이크로폰(17)으로 입력된 신호를 외부기기로 전송하기 위한 신호선들이 삽입되는 홀이다. 또한, 참조부호 12는 스폰지와 같은 소재로 된 덮개를 끼울 수 있도 록 형성된 걸림턱이다.

신호처리부(13)는 스피커(19)로부터 출력된 음향신호가 마이크로폰(17)으로 유입되지 않도록 음향신호에 대한 신호처리를 수행한다.

도 2는 도 1의 마이크로폰 일체형 이어폰의 개략적인 블록도이다. 도면을 참조하면, 신호처리부(13)는 스피커(19)로부터 출력되는 신호 및 마이크로폰(17)으로 입력되는 신호를 수신하며, 마이크로폰(17)으로 입력되는 신호 중 스피커(19)로부터 출력되는 신호의 성분이 제거되도록 스피커(19)로부터 출력되는 음향신호의 볼륨을 제어한다.

그런데, 상기한 음향 입/출력 장치와 같이 마이크로폰(17)이 귀 속에 삽입되는 경우에 주위의 소음을 차단하고 정확한 음성만을 집음할 수 있지만, 스피커(19)와 근접되어 설치되기 때문에 하울링 및 에코현상이 발생될 수 있다.

발명이 이루고자 하는 기술적 과제

본 발명은 상기와 같은 문제점을 개선함과 동시에 본 출원인에 의해 선출원된 발명을 개량하기 위하여 창안된 것으로서, 주위 환경에서 발생된 소음이 마이크로폰을 통해 입력되는 것을 방지할 수 있을 뿐만아니라, 하울링 및 에코현상을 방지할 수 있는 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형 이어폰을 제공하는데 그 목적이 있다.

발명의 구성 및 작용

상기의 목적을 달성하기 위한 본 발명에 따른 음향 입/출력 장치는, 입력되는 음향신호를 출력하는 스피커, 주위의 음향을 집음하여 전기신호로 변환하는 집 음마이크로폰, 사용자에 의한 발성유무를 확인하는 발성확인마이크로폰, 및 상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분을 제거하며 상기 발성확인마이크로폰에 의해 확인된 상기 사용자의 발성유무에 따라 변환된 상기 전기신호의 증폭율을 조절하는 회로부를 구비한다.

여기서, 상기 회로부는 상기 사용자의 발성유무에 따라 상기 전기신호의 증폭율을 조절하기 위한 제어신호를 변화시켜 출력하는 제어신호발생부, 상기 제어신호에 따라 상기 전기신호의 증폭율을 조절하는 신호처리부, 및 상기 스피커로 출력되는 상기 음향신호의 유무를 검출하는 스피커신호검출부를 구비한다. 상기 신호처리부는 검출된 상기 음향신호의 유무에 따라 상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분이 제거되도록 상기 전기신호의 증폭율을 조절한다.

한편, 본 발명에 따른 마이크로폰 일체형 이어폰은 귀에 삽입될 수 있도록 형성된 본체 내에 설치되며 입력되는 음향신호를 출력하는 스피커, 상기 본체 내에 설치되며 상기 귀로부터 발생되는 음향을 집음하여 전기신호로 변환하는 집음마이크로폰, 상기 본체의 외부에 설치되며 사용자에 의한 발성유무를 확인하는 발성확인마이크로폰, 및 상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분을 제거하며 상기 발성확인마이크로폰에 의해 확인된 상기 사용자의 발성유무에 따라 변환된 상기 전기신호의 증폭율을 조절하는 회로부를 구비한다. 여기서, 상기 스피커와 상기 집음마이크로폰은 동일축선 상에 배치된다.

이로써, 음향 입/출력 장치는 주위 환경에서 발생된 소음이 마이크로폰을 통 해 입력되는 것을 방지할 수 있을 뿐만아니라, 하울링 및 에코현상을 방지할 수 있게 된다.

이하, 첨부된 도면을 참조하여 본 발명을 보다 상세하게 설명한다.

도 3은 본 발명에 따른 음향 입/출력장치가 적용된 마이크로폰 일체형 이어폰을 개략적으로 나타내 보인 단면도이고, 도 4는 도 3의 마이크로폰 일체형 이어폰의 측면도이다. 여기서, 도 1 및 도 2에 기재된 부재와 동일한 부재에 대하여는 동일한 부재번호를 부여하였다.

도면을 참조하면, 마이크 일체형 이어폰(10)은 본체(11), 본체(11) 내에 설치된 집음마이크로폰(17), 본체(11) 외부에 설치된 발성확인 마이크로폰(18), 스피커(19), 및 본체(11) 내부에 설치된 회로부(20)를 구비한다.

본체(11)는 귀에 일부를 삽입할 수 있도록 적절한 크기 및 형태로 구현된다. 또한, 본체(11)는 내부에 집음마이크로폰(17)과 스피커(19)를 장착할 수 있으면서 외부로부터 내부로 신호선을 삽입할 수 있도록 홈이 형성되어 있다.

집음마이크로폰(17)은 귀로부터 발생되는 음향을 입력받을 수 있도록 본체(11)내에 지지부재(15)에 의해 지지되어 설치되며, 회로부(20)인 회로기판과 접속되어 있다. 집음마이크로폰(17)은 귀로부터 발생되는 음향을 집음하여 전기신호로 변환한다.

지지부재(15)는 음향신호의 입출이 용이하도록 일부가 관통되게 형성된다. 또는, 지지부재(15)는 집음마이크로폰(17)의 일측과 본체(11)가 상호 결합되도록 형성될 수도 있다.

발성확인 마이크로폰(18)은 본체(11)의 외부에 설치되며, 사용자에 의한 발성유무를 확인한다. 여기서, 발성확인 마이크로폰(18)은 귓바퀴의 연골부위에 접촉될 수 있도록 본체(11)의 하단부 외측에 설치되는 것이 바람직하다.

스피커(19)는 진동판을 갖는 통상적인 것이 적용되며, 귀(1)의 고막을 향해 음향신호를 출력할 수 있도록 본체(11)내에 설치되어 있다. 여기서, 집음마이크로폰(17)과 스피커(19)는 본체(11)의 크기를 최소화하고, 형태를 간단하게 하기 위하여 동일축선 상에 배치되는 것이 바람직하다. 또한, 스피커(19)로부터 출력된 음향이 집음마이크로폰(17)으로 직접 입력되지 않도록 음향출력방향과 음향입력방향이 서로 반대되는 것이 바람직하다.

참조부호 14는 외부기기로부터 입력되는 음향신호를 본체(11) 내부로 전송하고, 집음마이크로폰(17)으로 입력된 신호를 외부기기로 전송하기 위한 신호선들이 삽입되는 홀이다. 또한, 참조부호 12는 스폰지와 같은 소재로 된 덮개(12a)를 끼울 수 있도록 형성된 걸림턱이다.

회로부(20)는 스피커(19)로부터 출력된 음향신호가 집음마이크로폰(17)으로 유입되지 않도록 음향신호에 대한 신호처리를 수행한다. 또한, 회로부(20)는 발성확인 마이크로폰(18)에 의해 확인된 사용자의 발성유무에 따라 집음마이크로폰(17)에 의해 변환된 전기신호의 증폭율을 조절한다.

도 5는 도 3의 마이크로폰 일체형 이어폰의 개략적인 블록도이다. 도면을 참조하면, 회로부(20)는 스피커신호 검출부(21), 제어신호 발생부(22), 및 신호처리부(25)를 구비한다.

도면을 참조하여 마이크로폰 일체형 이어폰의 작용을 설명하면, 스피커신호 검출부(21)는 스피커(19)로 입력되는 음향신호의 유무를 검출한다. 음향신호가 검출되면, 스피커신호 검출부(21)는 음향신호를 제어신호 발생부(22)에 전송하여 음향신호의 존재를 통지한다.

제어신호 발생부(22)는 스피커신호 검출부(21)로부터 스피커(19)로 입력되는 음향신호의 유무에 대한 확인신호를 수신한다. 또한, 제어신호 발생부(22)는 발성확인 마이크로폰(18)으로부터 사용자의 발성유무를 확인신호를 수신한다. 제어신호 발생부(22)는 사용자의 발성유무에 따라 집음마이크로폰(17)에 의해 변환된 전기신호의 증폭율이 조절될 수 있도록 제어신호를 출력한다. 또한, 스피커(19)로 출력되는 음향신호의 유무에 따라 집음마이크로폰(17)으로 입력되는 신호 중에 스피커(19)에 의해 출력된 신호가 제거될 수 있도록 제어신호를 출력한다.

신호처리부(25)는 제어신호 발생부(22)로부터 출력되는 제어신호를 수신한다. 신호처리부(25)는 수신된 제어신호에 기초하여 집음마이크로폰(17)에 의해 출력되는 전기신호의 증폭율을 조절한다. 예컨대, 사용자에 의한 발성이 확인되지 않았지만, 집음마이크로폰(17)에 의해 음향이 집음된 경우에는 신호처리부(25)는 집음마이크로폰(17)에 의해 변환된 전기신호의 증폭율이 최저가 되도록 집음마이크로폰(17)의 볼륨을 조절한다.

사용자에 의한 발성이 확인되고 스피커(19)로 출력되는 음향신호가 확인되지 않은 경우에, 신호처리부(25)는 집음마이크로폰(17)에 의해 출력된 전기신호의 증폭율이 최대가 되도록 집음마이크로폰(17)의 볼륨을 조절한다.

사용자에 의한 발성이 확인되고 동시에 스피커(19)로 출력되는 음향신호가 확인된 경우에, 신호처리부(25)는 집음마이크로폰(17)에 의해 출력되는 전기신호의 증폭율이 중간레벨 정도가 되도록 집음마이크로폰(17)의 볼륨을 조절한다. 이 경우, 신호처리부(25)는 집음마이크로폰(17)으로부터 입력되는 신호 중 스피커(19)로부터 출력된 음향의 신호성분이 제거되도록 전기신호의 증폭율을 조절한다.

이로써, 사용자에 의해 발성된 음향은 정확하게 집음마이크로폰(17)으로 입력되고, 스피커(19)에 의해 출력된 음향은 제거되며, 동시에 집음마이크로폰(17)과 스피커(19)가 근접되어 설치됨으로 인한 하울링현상이나 에코현상은 제거되게 된다.

Exhibit 1009
Page 1674 of 2198

도 6은 도 3의 마이크로폰 일체형 이어폰의 회로부를 개략적으로 나타낸 블록도이다. 도면을 참조하면, 스피커신호 검출부(21)는 스피커(19)로 입력되는 음향신호를 증폭하며, 증폭된 음향신호를 제어신호 발생부(22)에 송신한다. 발성확인마이크로폰(18)는 발성확인신호를 증폭하는 발성확인신호 증폭부(18a)를 구비하며, 증폭된 신호를 제어신호 발생부(22)로 송신한다.

제어신호 발생부(22)는 수신된 음향신호 및/또는 발성확인신호에 따라 제어신호를 발생하는 소정의 논리회로로 구현된다. 집음마이크로폰(17)은 변환된 전기신호를 증폭하는 증폭부(17a)를 구비하며, 신호처리부(25)는 변환된 전기신호의 증폭율을 수신된 제어신호에 따라 조절한다.

발명의 효과

본 발명에 따른 음향 입/출력 장치 및 이를 적용한 마이크로폰 일체형 이어 폰에 의하면, 주위 환경에서 발생된 소음이 마이크로폰을 통해 입력되는 것을 방지할 수 있을 뿐만아니라, 하울링 및 에코현상을 방지할 수 있게 된다.

이상에서는 본 발명의 바람직한 실시예에 대해서 도시하고 설명하였으나, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형 실시가 가능한 것은 물론이고, 그와 같은 변경은 청구범위 기재의 범위 내에 있게 된다.

(57) 청구의 범위

청구항 1.
입력되는 음향신호를 출력하는 스피커;

주위의 음향을 집음하여 전기신호로 변환하는 집음마이크로폰;

사용자에 의한 발성유무를 확인하는 발성확인마이크로폰; 및

상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분을 제거하며, 상기 발성확인마이크로폰에 의해 확인된 상기 사용자의 발성유무에 따라 변환된 상기 전기신호의 증폭율을 조절하는 회로부;를 포함하는 것을 특징으로 하는 음향 입/출력 장치.

청구항 2.
제 1항에 있어서, 상기 회로부는,

상기 사용자의 발성유무에 따라 상기 전기신호의 증폭율을 조절하기 위한 제어신호를 변화시켜 출력하는 제어신호발생부; 및

상기 제어신호에 따라 상기 전기신호의 증폭율을 조절하는 신호처리부;를 포함하는 것을 특징으로 하는 음향 입/출력 장치.

청구항 3.
제 2항에 있어서, 상기 회로부는,

상기 스피커로 출력되는 상기 음향신호의 유무를 검출하는 스피커신호검출부를 더 포함하며,

상기 신호처리부는 검출된 상기 음향신호의 유무에 따라 상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분이 제거되도록 상기 전기신호의 증폭율을 조절하는 것을 특징으로 하는 음향 입/출력 장치.

청구항 4.
귀에 삽입될 수 있도록 형성된 본체 내에 설치되며, 입력되는 음향신호를 출력하는 스피커;

상기 본체 내에 설치되며, 상기 귀로부터 발생되는 음향을 집음하여 전기신호로 변환하는 집음마이크로폰;

Exhibit 1009
Page 1675 of 2198

상기 본체의 외부에 설치되며, 사용자에 의한 발성유무를 확인하는 발성확인마이크로폰; 및

상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분을 제거하며, 상기 발성확인마이크로폰에 의해 확인된 상기 사용자의 발성유무에 따라 변환된 상기 전기신호의 증폭율을 조절하는 회로부;를 포함하는 것을 특징으로 하는 마이크로폰 일체형 이어폰.

청구항 5.
제 4항에 있어서,

상기 스피커와 상기 집음마이크로폰은 동일축선 상에 배치되는 것을 특징으로 하는 마이크로폰 일체형 이어폰.

청구항 6.
제 5항에 있어서, 상기 회로부는,

상기 사용자의 발성유무에 따라 상기 전기신호의 증폭율을 조절하기 위한 제어신호를 출력하는 제어신호발생부; 및

상기 제어신호에 따라 상기 전기신호의 증폭율을 조절하는 신호처리부;를 포함하는 것을 특징으로 하는 마이크로폰 일체형 이어폰.

청구항 7.
제 6항에 있어서, 상기 회로부는,

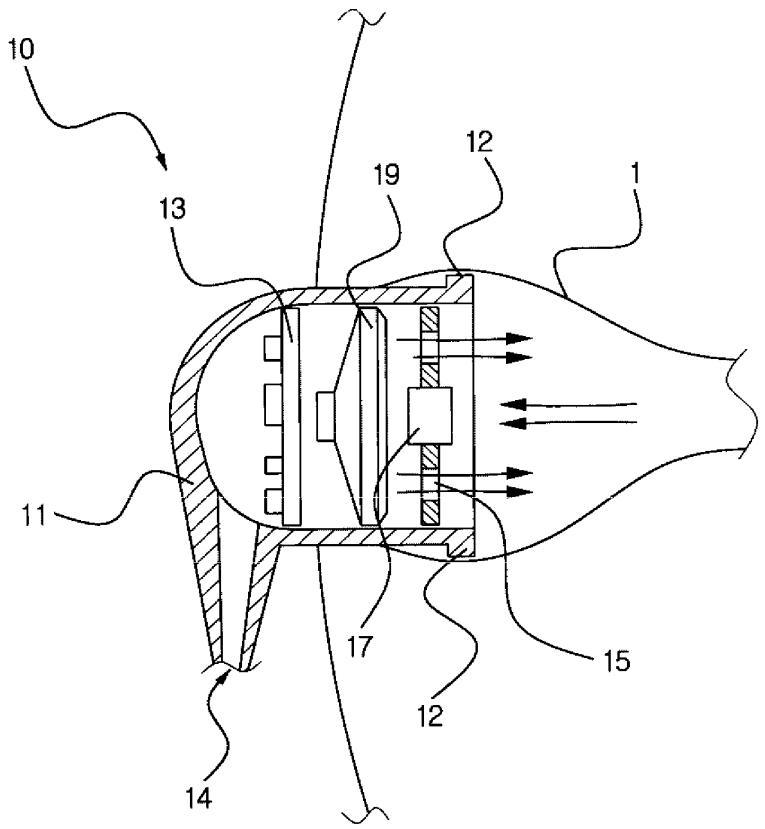
상기 스피커로 출력되는 상기 음향신호의 유무를 검출하는 스피커신호검출부를 더 포함하며,

상기 신호처리부는 검출된 상기 음향신호의 유무에 따라 상기 집음마이크로폰으로부터 입력되는 신호 중 상기 스피커로부터 출력된 신호의 성분이 제거되도록 상기 전기신호의 증폭율을 조절하는 것을 특징으로 하는 마이크로폰 일체형 이어폰.

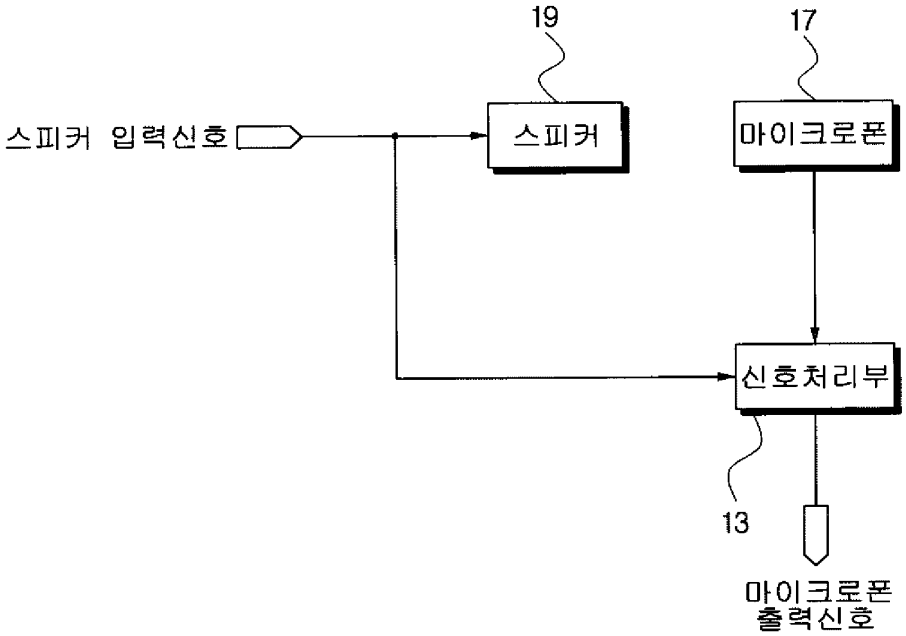
도면

Exhibit 1009
Page 1676 of 2198

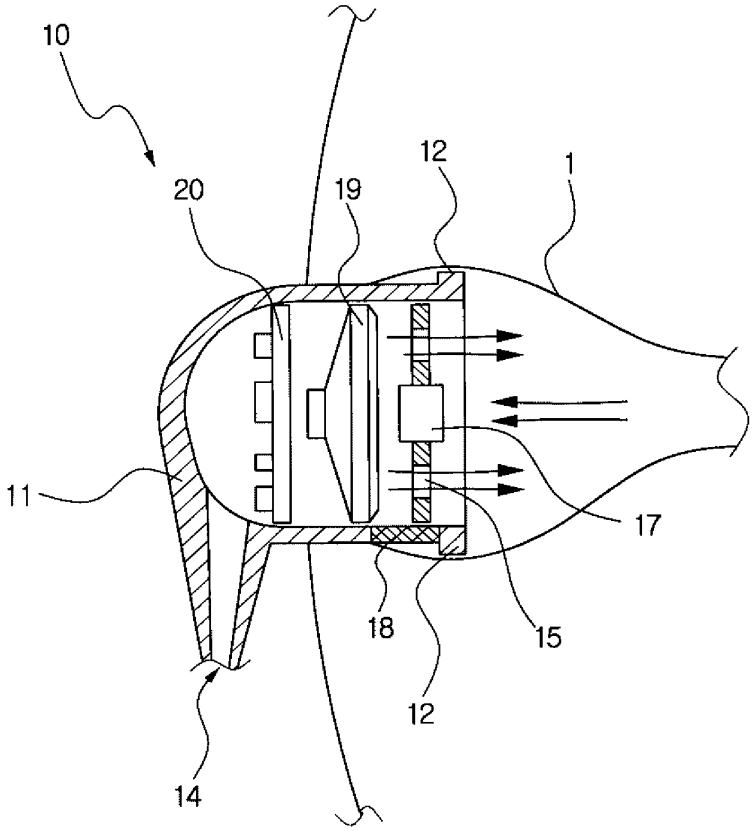
도면1



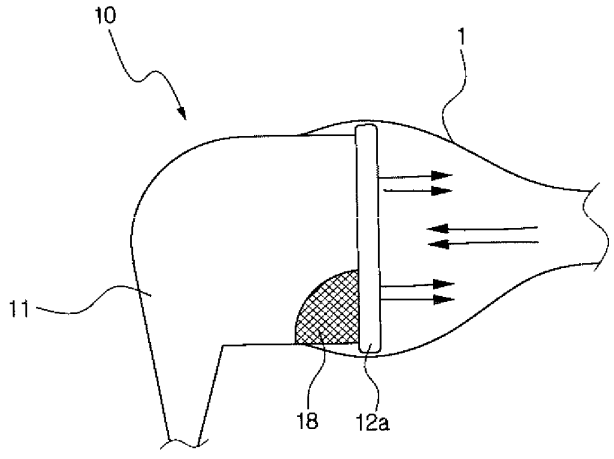
도면2



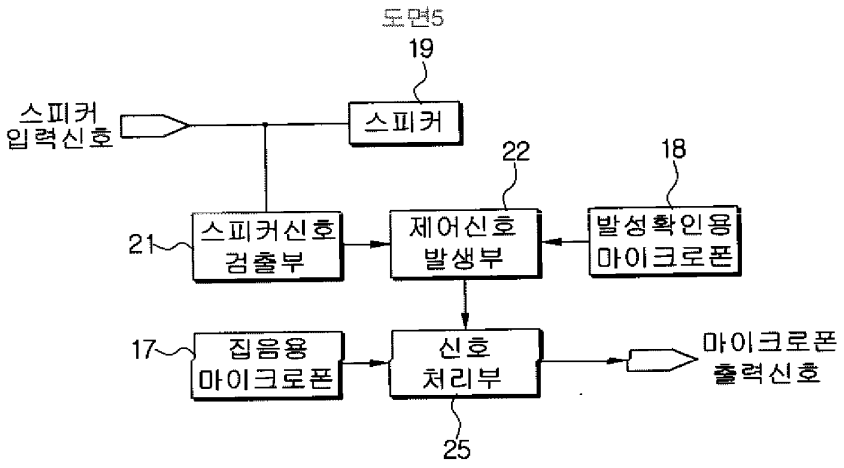
도면3



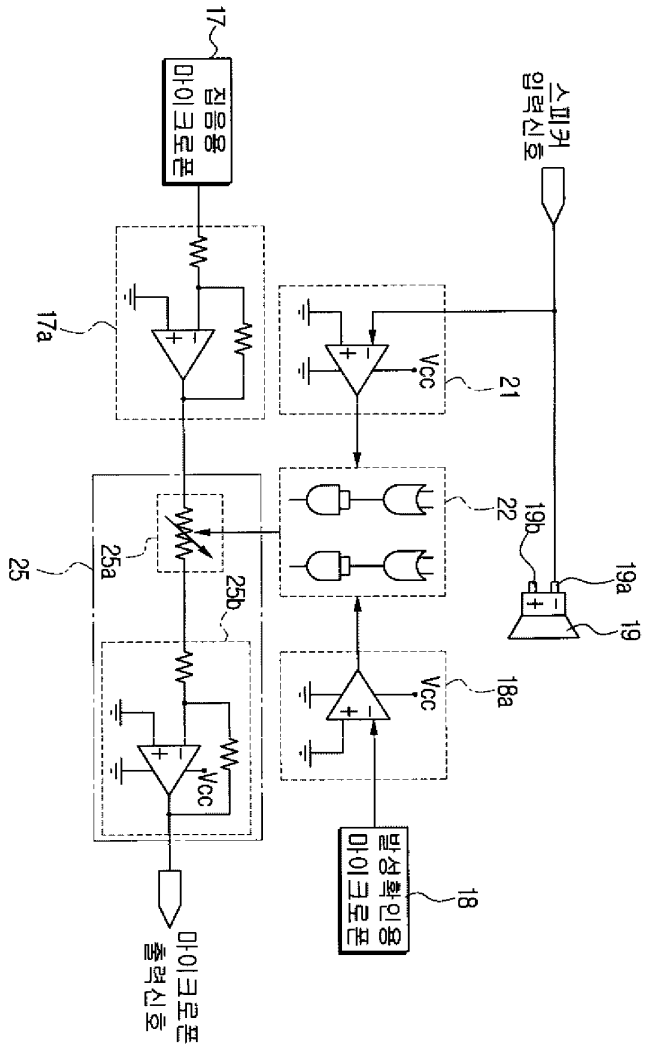
도면4



도면5


19
스피커
스피커
입력신호
22
18
21
스피커신호
검출부
제어신호
발생부
발성확인용
마이크로폰
17
집음용
마이크로폰
신호
처리부
마이크로폰
출력신호
25

도면6


17
집음용
마이크로폰
스피커
입력신호
17a
Vcc
21
22
19b
19a
19
25
25a
25b
Vcc
18a
Vcc
발성확인용
마이크로폰
18
마이크로폰
출력신호

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 특2002-0086433
(43) 공개일자 2002년11월18일

(51) 국제분류코드
H04B 1/40

(21) 출원번호 10-2002-0065280

(22) 출원일자 2002년10월24일

(71) 출원인
김효덕
서울특별시 서초구 방배1동 178-36, 대한민국

(72) 발명자
김효덕
서울특별시 서초구 방배1동 178-36, 대한민국

(74) 대리인
최용원
이상용
김상우

(77) 심사관
있음

전체 청구항 수 : 총 9 항

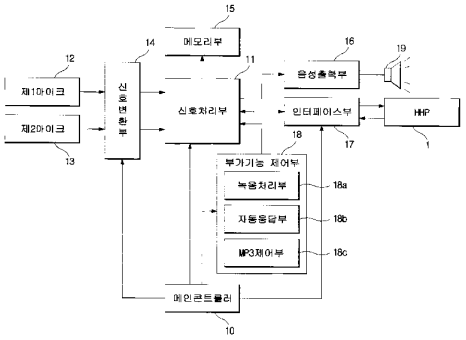
(54)소음제거 기능을 가지는 핸즈프리 장치 및 이를 위한이어마이크 구조체

(57)요 약

본 발명은 휴대전화기에 접속되는 핸즈프리장치에 관한 것으로서, 소정 길이의 코드선이 연장되어 있으며, 내부에는 음성재생을 위한 스피커가 수용되어 있는 이어폰 몸체; 음성감지를 위해 상기 코드선 혹은 이어폰 몸체의 일측에 장착되는 제1마이크; 상기 제1마이크에 근접하여 설치되어 소음을 감지하는 제2마이크; 상기 제2마이크로부터 입력되는 신호로부터 잡음신호를 분리한 후, 이 잡음신호에 상응하는 신호성분 만큼을 상기 제1마이크로부터 입력되는 신호에서 제거하여 음성신호를 추출 및 출력하는 신호처리부; 및 상기 휴대전화기 혹은 상기 신호처리부로부터 음성신호를 전달받아 처리한 후 스피커를 통해 출력하는 음성출력부;를 포함하는 핸즈프리 장치가 개시된다.

키워드 : 핸즈프리, 음향커플링, 풀듀플렉스, DSP, MP3

대 표 도 - 도 3



This Facsimile Page has been artificially created from KIPO data.

청구의 범위

청구항 1

휴대전화기에 접속되는 핸즈프리 장치로서,

소정 길이의 코드선이 연장되어 있으며, 내부에는 음성재생을 위한 스피커가 수용되어 있는 이어폰 몸체;

음성감지를 위해 상기 코드선 혹은 이어폰 몸체의 일측에 장착되는 제1마이크;

상기 제1마이크에 근접하여 설치되어 소음을 감지하는 제2마이크;

상기 제2마이크로부터 입력되는 신호로부터 잡음신호를 분리한 후, 이 잡음신호에 상응하는 신호성분 만큼을 상기 제1마이크로부터 입력되는 신호에서 제거하여 음성신호를 추출 및 출력하는 신호처리부; 및

상기 휴대전화기 혹은 상기 신호처리부로부터 음성신호를 전달받아 처리한 후 스피커를 통해 출력하는 음성출력부;를 포함하는 핸즈프리 장치.

청구항 2

휴대전화기에 접속되는 핸즈프리 장치로서,

내부에 음성재생을 위한 스피커가 수용되어 있는 이어폰 몸체;

음성감지를 위해 상기 이어폰 몸체의 일측에 장착되는 제1마이크;

상기 제1마이크에 근접하여 설치되어 소음을 감지하는 제2마이크;

상기 제2마이크로부터 입력되는 신호로부터 잡음신호를 분리한 후, 이 잡음신호에 상응하는 신호성분 만큼을 상기 제1마이크로부터 입력되는 신호에서 제거하
여 음성신호를 추출 및 출력하는 신호처리부; 및

상기 휴대전화기 혹은 상기 신호처리부로부터 음성신호를 전달받아 처리한 후 스피커를 통해 출력하는 음성출력부;를 포함하는 핸즈프리 장치.

청구항 3

제 2항에 있어서,

상기 제1마이크 혹은 제2마이크 중 어느 하나가 스피커의 전방 혹은 후방에 위치한 상태로 이어폰 몸체에 내장되는 것을 특징으로 하는 핸즈프리 장치.

청구항 4

제 2항에 있어서,

상기 제1마이크와 제2마이크가 스피커의 전방 혹은 후방에 위치한 상태로 이어폰 몸체에 내장되는 것을 특징으로 하는 핸즈프리 장치.

청구항 5

제 1항 또는 제 2항에 있어서,

음성데이터를 저장하기 위한 메모리부; 및

상기 신호처리부로부터 음성신호를 전달받아 상기 메모리부에 저장하는 녹음처리부;를 더 포함하는 핸즈프리 장치.

청구항 6

제 5항에 있어서,

상기 휴대전화기로부터 착신신호가 감지되었을 때, 상기 메모리부에 저장되어 있는 음성데이터를 자동으로 인출하여 휴대전화기 측으로 출력하는 자동응답부;
를 더 포함하는 것을 특징으로 하는 핸즈프리 장치.

청구항 7

제 1항 또는 제 2항에 있어서,

MP3 파일을 저장하기 위한 메모리부; 및

상기 메모리부로부터 MP3 파일을 인출하여 디코딩하고 이를 스피커로 출력하는 MP3 제어부;를 더 포함하는 것을 특징으로 하는 핸즈프리 장치.

청구항 8

내부에 스피커가 수용되어 있는 이어폰 몸체;

상기 이어폰 몸체 내부의 일측에 장착되어 음성을 감지하는 제1마이크; 및

소음감지를 위해 상기 스피커의 전방 혹은 후방에 위치한 상태로 이어폰 몸체에 내장되는 제2마이크;를 포함하는 이어마이크 구조체.

청구항 9

내부에 스피커가 수용되어 있는 이어폰 몸체;

상기 이어폰 몸체 바깥의 일측에 장착되어 음성을 감지하는 제1마이크; 및

소음감지를 위해 상기 스피커의 전방 혹은 후방에 위치한 상태로 이어폰 몸체에 내장되는 제2마이크;를 포함하는 이어마이크 구조체.

명세서

발명의 명칭

소음제거 기능을 가지는 핸즈프리 장치 및 이를 위한 이어마이크 구조체{Handsfree device including function of noise cancellation and ear-mic structure for the same}

도면의 간단한 설명

[0001] 본 명세서에 첨부되는 다음의 도면들은 본 발명의 바람직한 실시예를 예시하는 것이며, 후술하는 발명의 상세한 설명과 함께 본 발명의 기술사상을 더욱 이해시키는 역할을 하는 것이므로, 본 발명은 그러한 도면에 기재된 사항에만 한정되어 해석되어서는 아니된다.

[0002] 도 1은 본 발명에 따른 핸즈프리 장치의 외관의 일 예를 보여주는 사시도.

[0003] 도 2는 본 발명에 따른 핸즈프리 장치의 외관의 다른 예를 보여주는 사시도.

[0004] 도 3은 본 발명에 따른 핸즈프리 장치의 구성을 나타내는 기능 블럭도.

[0005] 도 4는 본 발명에 따른 핸즈프리 장치에 사용되는 이어마이크의 구성을 개략적으로 도시하는 측면도.

[0006] 도 5는 도 4의 변형예를 개략적으로 도시하는 측면도.

[0007] <도면의 주요 참조부호에 대한 설명>

[0008] 5...이어폰 몸체 10...메인 콘트롤러 11...신호처리부

[0009] 12...제1마이크 13...제2마이크 15...메모리부

[0010] 16...음성출력부 18...부가기능제어부 19...스피커

발명의 상세한 설명

[0011] 본 발명은 소음제거 기능을 가지는 핸즈프리 장치 및 이를 위한 이어마이크 구조체에 관한 것으로서, 보다 상세하게는 두 개의 마이크와 DSP(Digital Signal Processor)칩을 이용하여 소음을 제거함으로써 풀듀플렉스 통신과, 보이스 레코딩(Voice recording) 등의 음성/음향 관련 부가기능이 구현되도록 하는 핸즈프리 장치 및 이어마이크 구조체에 관한 것이다.

[0012] 휴대폰을 비롯하여 이동전화기능을 가지는 노트북, PDA 등 각종 휴대전화기를 편리하게 이용할 수 있도록 보조하는 통신장치인 핸즈프리는 일반적으로 이어폰과 마이크가 상호 근접하며 외부에 노출되는 구조를 가지므로 에코나, 하울링 등의 음향 커플링을 방지하기 위한 대책이 요구된다.

[0013] 이를 위해, 종래에는 이어폰으로부터 수신음이 출력되는 도중에는 마이크가 차단되도록 스위칭하는 하프 듀플렉스(Half duplex)방식이 널리 사용되었는데, 이 방식은 송신측과 수신측의 신호레벨을 서로 비교하여 신호출력이 약한 쪽을 끊어주고 강한 쪽의 신호를 전송하는 방식을 사용한다.

[0014] 따라서, 이러한 방식은 통화가 자연스럽게 이루어지지 않는 단점이 있으며, 이에 따라 최근에는 DSP칩을 이용하여 주변잡음을 제거하고, 송수신 양측의 음성데이터를 디지털 처리하여 풀듀플렉스 방식으로 동시통화를 구현하는 기술들이 소개되고 있다.

[0015] 그러나, 이러한 종래의 기술들은 단일의 마이크와 스피커를 사용하는 구조의 한계를 벗어나지 못한 상태에서 디지털 신호처리에만 의존하여 잡음제거나 듀플렉스를 구현하므로 DSP칩의 구성이 복잡해지고 제조단가가 상승할 수 밖에 없는 문제점이 있다.

[0016] 한편, 대한민국 특허공개 제1999-018766호와 같이 두 개의 마이크를 채용하여 시스템을 구성하고, 두 개의 입력신호 중 잡음이 적은 쪽을 주마이크로 선택하는 방식이 제시된 바 있으나, 이 방식의 경우 두 개의 마이크가 근접되어 있을 경우에는 잡음환경이 상호 유사하므로 큰 실효를 거둘 수 없는 한계가 있다.

[0017] 본 발명은 상기와 같은 점을 고려하여 창안된 것으로서, 소음제거 기능과 함께 풀듀플렉스 방식이 구현되도록 하는 핸즈프리 장치 및 이어마이크 구조체를 제공하는데 그 목적이 있다.

[0018] 본 발명의 다른 목적은 핸즈프리 사용자의 음성입력에 따른 신호를 디지털 처리하여 소음을 제거한 후 녹음이 이루어지도록 하는 핸즈프리 장치 및 이어마이크 구조체를 제공하는데 있다.

[0019] 본 발명의 또 다른 목적은 디지털 처리되어 녹음된 핸즈프리 사용자의 음성을 전화수신시 자동으로 재생하기 위한 핸즈프리 장치 및 이어마이크 구조체를 제공하는데 있다.

[0020] 상기와 같은 목적을 달성하기 위해 본 발명에 따른 핸즈프리 장치는, 소정 길이의 코드선이 연장되어 있으며, 내부에는 음성재생을 위한 스피커가 수용되어 있는 이어폰 몸체; 음성감지를 위해 상기 코드선 혹은 이어폰 몸체의 일측에 장착되는 제1마이크; 상기 제1마이크에 근접하여 설치되어 소음을 감지하는 제2마이크; 상기 제2마이크로부터 입력되는 신호로부터 잡음신호를 분리한 후, 이 잡음신호에 상응하는 신호성분 만큼을 상기 제1마이크로부터 입력되는 신호에서 제거하여 음성신호를 추출 및 출력하는 신호처리부; 및 상기 휴대전화기 혹은 상기 신호처리부로부터 음성신호를 전달받아 처리한 후 스피커를 통해 출력하는 음성출력부;를 포함한다.

[0021] 본 발명의 다른 실시예에 의하면, 휴대전화기에 접속되는 핸즈프리장치로서, 내부에 음성재생을 위한 스피커가 수용되어 있는 이어폰 몸체; 음성감지를 위해 상기 이어폰 몸체의 일측에 장착되는 제1마이크; 상기 제1마이크에 근접하여 설치되어 소음을 감지하는 제2마이크; 상기 제2마이크로부터 입력되는 신호로부터 잡음신호를 분리한 후, 이 잡음신호에 상응하는 신호성분 만큼을 상기 제1마이크로부터 입력되는 신호에서 제거하여 음성신호를 추출 및 출력하는 신호처리부; 및 상기 휴대전화기 혹은 상기 신호처리부로부터 음성신호를 전달받아 처리한 후 스피커를 통해 출력하는 음성출력부;를 포함하는 핸즈프리 장치가 제공된다.

[0022] 상기 제1마이크와 제2마이크 중 적어도 어느 하나는 스피커의 전방 혹은 후방에 위치한 상태로 이어폰 몸체에 내장된다.

[0023] 바람직하게 본 발명은, 음성데이터를 저장하기 위한 메모리부; 상기 신호처리부로부터 음성신호를 전달받아 상기 메모리부에 저장하는 녹음처리부;를 더 포함
한다.

[0024] 또한, 상기 휴대전화기로부터 착신신호가 감지되었을 때, 상기 메모리부에 저장되어 있는 음성데이터를 자동으로 인출하여 휴대전화기 측으로 출력하는 자동응답부;가 더 포함되며, 부가적으로 MP3 파일을 저장하기 위한 메모리부; 및 메모리부로부터 MP3 파일을 인출하여 디코딩하고 이를 스피커로 출력하는 MP3 제어부;를 더 포함하는 것이 바람직하다.

[0025] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예를 상세히 설명하기로 한다.

[0026] 먼저, 도 1은 본 발명에 따른 핸즈프리 장치의 외관을 개략적으로 보여주는 도면이다. 도면에 도시된 바와 같이 본 발명은 연결부(2)를 통해 휴대전화기(1)에 접속되는 한편, 그 내부에 각종 회로부가 수납된 본체부(3)와, 소정 길이의 코드선(4)이 연장형성되어 상기 본체부(3)에 접속되는 이어폰 몸체(5)를 포함한다. 비록, 도면에는 연결부(2)가 휴대전화기(1)의 하단 입출력포트에 접속되는 것으로 도시되어 있으나, 이에 한정되지 않고 이어폰 잭(미도시) 부분에 접속될 수도 있다.

[0027] 또한, 도면에는 본체부(3)와 휴대전화기(1)가 유선으로 접속되는 것이 도시되어 있으나, 이에 한정되지 않고 도 2에 도시된 바와 같이 공지의 무선통신모듈(6)을 부가적으로 장착하여 무선으로 접속될 수도 있음은 물론이다.

[0028] 도 3에는 본 발명에 따른 핸즈프리 장치의 기능적 구성이 도시되어 있다. 도면을 참조하면, 본 발명은 제1마이크(12) 및 제2마이크(13)를 포함하고, 상기 본체
부(3) 내에 수납되는 구성요소로서 신호처리부(11)와, 음성출력부(16)와, 부가기능제어부(18) 및 상기 요소들의 동작을 통합적으로 제어하기 위한 메인 콘트롤러(10)를 포함한다.

[0029] 제1마이크(12)와 제2마이크(13)는 각각 사용자가 입력하는 음성과 외부 소음을 감지하는 마이크로서, 에컨대 콘덴서 마이크 등 통상의 핸즈프리에 사용되는 마이크가 동일하게 채용될 수 있다. 본 발명에 있어서, 상기 제1마이크(12)와 제2마이크(13)의 구분은 설치위치와 신호처리부(11)에서의 처리동작으로 구분된다.

[0030] 바람직하게, 소음감지를 위한 상기 제2마이크(13)의 설치위치는 이어폰 몸체(5)의 내부가 되며, 음성감지를 위한 제1마이크(12)는 제2마이크(13)로부터 예컨대, 수mm 내지 수cmm의 간격을 두고 코드선(4) 혹은 이어폰 몸체(5)에 설치된다. 여기서, 제1마이크(12)와 제2마이크(13)의 위치가 상호 바뀐 구조로 본 발명이 변형될 수도 있음은 물론이다.

[0031] 도 4에는 상기 제1마이크(12)와 제2마이크(13)가 이어폰 몸체(5)에 설치되어 이어마이크 구조체를 형성하는 실시예가 개략적으로 도시되어 있다. 도면을 참조하면, 이어폰 몸체(5)의 내부에는 통상의 기술을 채용하여 이어폰용 스피커(19)가 설치되며, 이 스피커(19)의 전방에는 본 발명에 따라 소음감지를 위한 제2마이크(13)가 장착된다. 아울러, 이어폰 몸체(5)의 하단에는 음성감지를 위한 제1마이크(12)가 설치된다. 여기서, 본 발명이 이러한 구조에 반드시 한정되지 않고, 상기 제2마이크(13)가 스피커(19)의 후방에 위치하거나, 제1마이크(12)의 위치가 보다 제2마이크(13)와 근접된 위치의 이어폰 몸체(5)에 설치되도록 변형될 수도 있음은 물론
이다.

[0032] 도 5에는 상기 이어마이크 구조체의 다른 실시예가 개략적으로 도시되어 있다. 도면에 나타난 바와 같이, 제1마이크(12)와 제2마이크(13)는 스피커(19)의 전방에 위치하여 각각 음성과 외부소음을 감지하게 된다. 여기서, 변형예로서 두 개의 마이크가 모두 스피커(19)의 후방에 설치되는 것도 가능하며, 제1마이크(12)와 제2마이크(13)가 스피커(19)를 사이에 두고 대칭을 이루며 설치될 수도 있다.

[0033] 상기와 같은 구조의 제1마이크(12)와 제2마이크(13)로부터 입력되는 아날로그 신호는 신호변환부(14)를 통해 증폭 및 디지털 변환된 후 신호처리부(11)로 입력된다.

[0034] 신호처리부(11)로는 바람직하게, DSP칩이 채용되며 그 주된 기능은 입력되는 디지털 신호로부터 잡음신호를 제거(Cancellation)하는 것이다. 이를 위해, 먼저 신호처리부(11)는 제2마이크(12)로부터 입력되는 감지신호로부터 잡음대역, 즉 핸즈프리 사용자가 현재 입력하는

음성 외의 모든 음향신호를 분리한다. 이때, 잡음대역을 인식하는 방법은 스피커(19)로부터 출력되어 제2마이크(12)에 입력된 음성신호를 비롯하여, 일정레벨 이상 혹은 이하의 진폭이나 불규칙적인 파형패턴의 신호를 잡음으로 인식하는 등 통상의 핸즈프리용 DSP칩에서 사용하는 방식이 동일하게 채용될 수 있다.

[0035] 잡음대역을 분리한 후 신호처리부(11)는 분리된 잡음신호에 상응하는 신호성분 만큼을, 제1마이크(12)로부터 입력된 감지신호에서 차감한다. 이때, 잡음성분을 차감하기 위해서는, 해당 주파수 대역을 필터링하는 방식은 물론, 상기
제2마이크(13)를 통해 입력된 후 분리된 잡음신호의 위상을 180°반전시킨 후 상기 제1마이크(12)로부터 입력되는 감지신호에 인가하는 등 다양한 방법이 채용가능하다.

[0036] 이와 같이, 본 발명에서는 스피커(19)를 통해 음성이 출력되는 도중에도 제1마이크(12)를 통해 사용자의 음성입력이 이루어지게 되며, 음향커플링이나 주변잡음 신호는 제2마이크(13)를 통해 감지된 후 의사신호가 되어 상기 제1마이크(12)로부터 입력되는 신호에서 동일성분을 차감하게 되므로 사용자가 입력하는 음성성분만을 추출하는 것이 가능하다.

[0037] 상기 신호처리부(11)를 통해 분리된 음성신호는 인터페이스부(17)를 통해 아날로드 신호로 변환된 후 소정 전압으로 증폭되어 휴대전화기(1)로 전달된다.

[0038] 인터페이스부(17)는 본 발명의 핸즈프리 장치와 휴대전화기(1)를 상호 물리적으로 연결하는 부분으로서 부하의 정합(Matching)과, 변압, A/D 변환, D/A 변환 등의 기능을 수행함으로써 상기 신호처리부(11)로부터 입력되는 신호를 휴대전화기(1)로 전달하고, 이와 더불어 휴대전화기(1)로부터 입력되는 신호를 상기 신호처리부(11)로 전달한다.

[0039] 휴대전화기(1) 측에 무선통신모듈(6)이 장착될 경우, 상기 인터페이스부(17)는 이에 상응하는 표준의 무선모듈을 포함하여 구성된다. 이때의 무선표준으로는 에컨대, 블루투스(Bluetooth)가 해당될 수 있을 것이다.

[0040] 상기 인터페이스부(17)를 통해 휴대전화기(1)로부터 입력된 통화 상대방의 음성신호는 신호처리부(11)에서의 디지털 처리를 통해 잡음이 제거된 후 음성출력
부(16)로 인가된다. 이 음성출력부(16)에서는 입력된 신호를 아날로그 신호로 변환한 후 증폭하여 스피커(19)를 통해 출력한다.

[0041] 한편, 바람직하게 본 발명에는 상기 신호처리부(11)와 연결되는 부가기능 제어부(18)가 더 포함되어 보이스 레코딩, 자동응답, MP3파일 입출력 등의 기능을 수행하게 된다. 이를 위해, 부가기능 제어부(18)에는 녹음처리부(18a), 자동응답부(18b) 및 MP3 제어부(18c)가 포함된다.

[0042] 녹음처리부(18a)는 메인 콘트롤러(10)로부터 전달되는 명령신호에 따라, 상기 제1마이크(12)로부터 입력된 후 신호처리부(11)에서 잡음제거 등의 디지털 처리를 거친 음성데이터를 메모리부(15)의 할당영역에 저장하는 기능을 수행한다. 이때, 메모리부(15)로는 바람직하게 플래쉬 메모리(Flash memory)가 채용될 수 있다.

[0043] 상기 메모리부(15)에 저장된 음성데이터는 메인 콘트롤러(10)를 통한 명령신호에 따라 스피커(19)를 통해 재생이 가능하며, 휴대전화기(1)로부터 착신신호가 전달되는 순간에 자동응답부(18b)의 제어를 받아 인출되어 휴대전화기(1) 측으로 전달됨으로써 자동응답기능이 수행될 수도 있다.

[0044] MP3 제어부(18c)는 MP3 파일을 해독하기 위한 디코딩 회로를 포함하여 구성되며, 메모리부(15)에 MP3 파일을 저장하고 메인 콘트롤러(10)의 명령신호에 따라 파일을 인출하여 스

피커(19)를 통해 재생하는 부가기능을 수행한다. 이를 위해, 메모리부(15)로는 통상의 멀티미디어 카드(MMC)와 같은 메모리 카드가 바람직하게 채용될 수 있을 것이다.

[0045] 그러면, 상기와 같은 구성요소를 포함하는 본 발명에 따른 핸즈프리 장치의 동작을 설명하기로 한다.

[0046] 본 발명에 의하면 제1마이크(12)와 제2마이크(13)가 각각 별도로 구성되어 스피커(19)의 동작중에도 음성입력을 위한 제1마이크(12)를 동작시키는 것이 가능하고, 이에 따라 풀듀플렉스 방식의 통신이 가능하게 된다.

[0047] 이를 위해, 제2마이크(13)가 바람직하게 스피커(19)에 근접하여 설치되어 주변의 소음이나 스피커(19)로부터 출력되는 음성 등을 포함하는 신호를 감지한 후 신호처리부(11)로 전달한다. 신호처리부(11)에서는 상기 제2마이크(13)로부터 입력되는 신호로부터 잡음성분을 분리해 내고, 이를 의사신호로 활용하여 제1마이크(12)로부터 입력되는 신호에서 잡음성분을 제거한다.

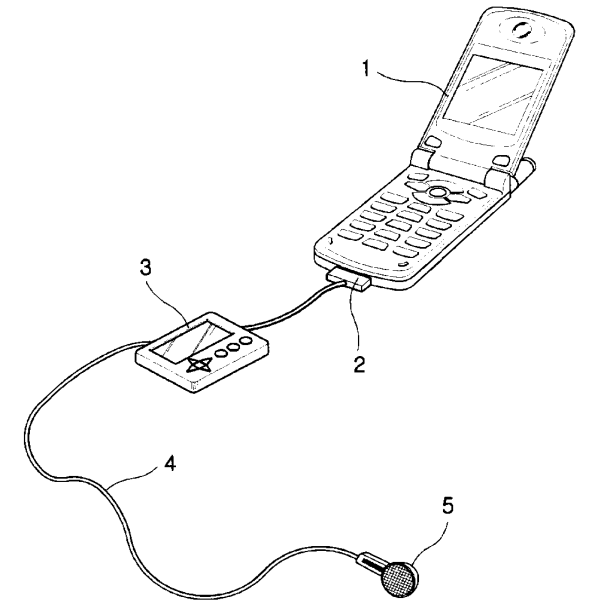
[0048] 제1마이크(12)로부터 입력되는 음성감지신호는 상기와 같이 신호처리부(11)를 통해 잡음이 제거된 후 인터페이스부(17)를 통해 휴대전화기(1) 측으로 전달된다. 이와 더불어, 휴대전화기(1)로부터 입력되는 통화상대방 측의 음성신호는 상기 신호처리부(11)를 통해 잡음제거 등의 디지털 처리를 거친 후 음성출력부(16)를 경유하여 스피커(19)로 출력된다.

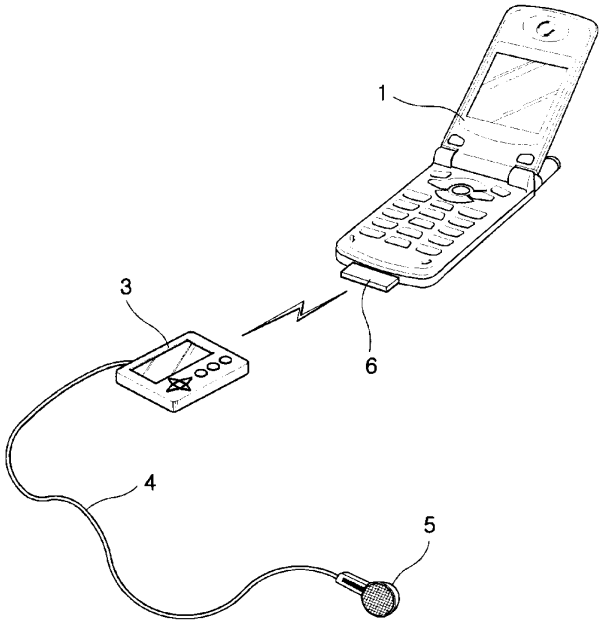
[0049] 한편, 상기와 같이 본 발명은 두 개의 마이크(12,13)와 신호처리부(11)를 통해 잡음을 제거하는 것이 가능하므로 부가적으로 녹음처리부(18a)를 통해 깨끗한 음성을 녹음하는 것이 가능하며, 착신신호 수신시 자동응답부(18b)를 통해 자동으로 음성데이터를 인출하여 재생하는 것도 가능하다.

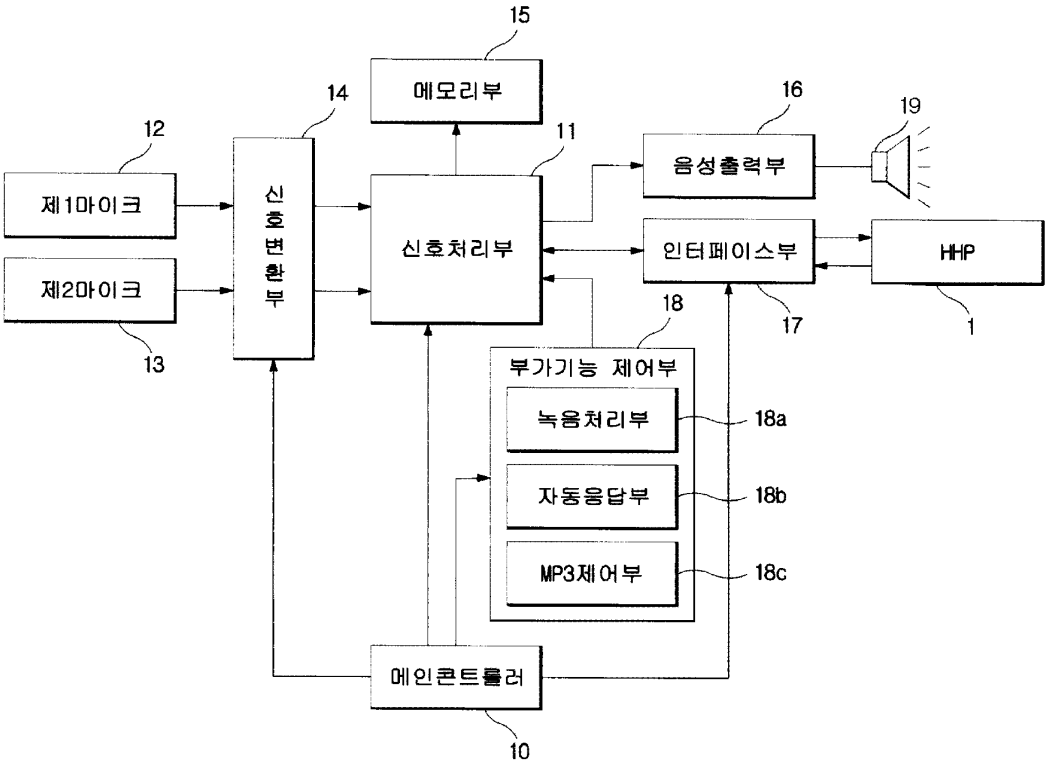
[0050] 이상, 본 발명의 바람직한 실시예를 첨부된 도면들을 참조로 설명하였다. 여기서, 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서, 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 가장 바람직한 일 실시예에 불과할 뿐이고 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.

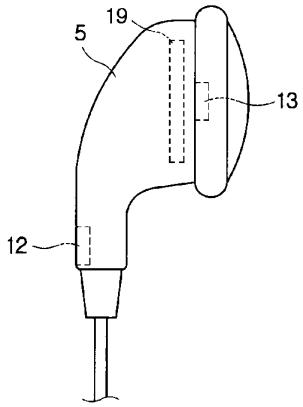
[0051] 본 발명에 의하면 두 개의 마이크와 단순한 구조의 DSP칩을 통해 저렴하게 잡음제거 및 풀듀플렉스 통신을 구현할 수 있고, 이에 따라 노래방이나 나이크클럽, 공사장 등 소음이 심한 환경하에서도 깨끗한 음질로 동시통화가 가능하게 되는 장점이 있다.

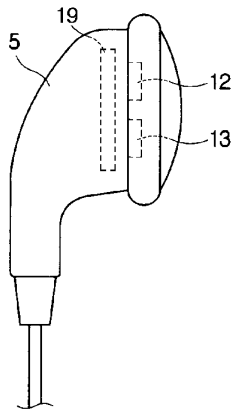
[0052] 또한, 본 발명에 따른 핸즈프리 장치에는 잡음제거 기능과 연계한 디지털 녹음처리, 자동응답 등의 기능이 구비되므로 상기와 같이 소음환경이 열악한 곳에서도 자유롭게 음성데이터를 녹음 및 재생할 수 있는 효과가 있다.


1
3
2
4
5


1
3
6
4
5


15
메모리부
14
12
11
16
19
제1마이크
신호변환부
신호처리부
음성출력부
제2마이크
인터페이스부
HHP
13
17
1
18
부가기능 제어부
녹음처리부
18a
자동응답부
18b
MP3제어부
18c
메인콘트롤러
10


19
5
13
12


19
5
12
13

(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.[7]
H04R 1/10

(11) 공개번호 특2003-0069471
(43) 공개일자 2003년08월27일

(21) 출원번호 10-2002-0009113
(22) 출원일자 2002년02월20일

(71) 출원인 삼성물산 주식회사
서울특별시 중구 태평로2가 310

케이.이.티 주식회사
경기도 안양시 만안구 안양7동 202-6 동일테크노타운 5602호

(72) 발명자 김창용
경기도 군포시 금정동 876 율곡아파트 333-1405 8/2

(74) 대리인 특허법인아람

심사청구 : 있음

(54) 음향 입출력장치 및 이를 적용한 마이크 일체형 이어폰

요약

마이크로폰 및 스피커를 하나의 몸체 내부에 내장한 음향 입출력장치 및 이를 적용한 마이크 일체형 이어폰이 개시된다. 본 발명의 마이크 일체형 이어폰은 사람의 귀에 삽입되는 이어폰 몸체 내에 내장형마이크 및 스피커가 내장되며, 그 몸체 외측으로 지향성인 외장형마이크가 더 부가된다. 그리고 내장형마이크 및 외장형마이크를 선택적으로 입력단자에 연결시키는 모드전환스위치가 구비된다. 이와 같은 이어폰은 외부 환경에 따라 사용자가 모드전환 스위치를 통해 선택적으로 내장형마이크 또는 외장형마이크를 이용하여 음성을 입력할 수 있다. 즉, 소음이 많은 지역에서는 모드전환스위치를 소음모드로 위치시켜 사용자의 귀에서 발생된 음성이 내장형마이크를 통해 상대방에게 전달되도록 하며, 소음이 적은 지역에서는 사용자의 입에서 출력된 음성이 외장형마이크를 통해 상대방에게 전달되도록 함으로서 소음에 의한 통화 간섭을 줄일 수 있다. 또한, 외장형마이크를 이용하는 경우에는 내장형마이크와 스피커의 근접설치에 따른 통화 간섭을 줄이기 위해 이용되는 회로의 전원 공급을 차단함으로서 전력소모를 감소시킬 수 있으며, 입에서 발생된 음성 입력을 통해 보다 명확하게 메시지를 전달할 수 있다.

대표도

도 3

색인어

이어폰, 내장형마이크, 외장형마이크, 스피커, 음향, 모드전환스위치,

명세서

도면의 간단한 설명

도 1은 일반적인 마이크 부가형 이어폰의 사시도,

Exhibit 1009
Page 1693 of 2198

도 2는 본 발명의 실시예에 따른 핸드폰용 이어폰의 사시도, 그리고

도 3은 도 2에 보인 핸드폰용 이어폰의 블록도이다.

도면의 주요부분에 대한 부호의 설명

10, 110: 커넥터 20, 210: 이어폰스피커

30, 220, 230: 마이크 40, 130: 제어기

42, 134: 호출응답버튼 120: 이어폰

122: 삽입부 124: 돌출부

132, 262: 모드전환조작스위치 200: 핸드폰

240: 신호검출부 250: 스위칭소자

260: 모드전환스위치 264: 모드전환스위칭부

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 음향입출력장치에 관한 것으로서, 특히, 마이크로폰(Microphone: 이하, '마이크'라 한다) 및 스피커(Speaker)를 하나의 몸체 내부에 내장한 음향 입출력장치 및 이를 적용한 마이크 일체형 이어폰에 관한 것이다.

일반적으로 이어폰(헤드폰)과 같은 장치는 휴대형 오디오 플레이어나 휴대형 라디오와 같은 메인출력장치의 보조출력수단으로서 이용되어 왔다. 그러나 최근 이동통신단말기 등과 같은 휴대형 데이터송수신장치 및 자신의 음성을 녹음한 후, 재생할 수 있도록 구현된 휴대형 언어학습기 등이 등장하면서, 이어폰(헤드폰)은 단순한 보조출력수단에 그치지 않고, 마이크를 부가하여 보조입력장치로서의 기능까지 수행할 수 있도록 개발되고 있다.

도 1은 마이크가 부가된 일반적인 이어폰의 사시도이다. 도면에서 이어폰은 핸드폰과 같은 메인장치의 입/출력단자에 결합되는 커넥터(10)와, 사용자의 귀에 삽입될 수 있도록 외관이 형성된 스피커(20)와, 커넥터(10)와 스피커(20)를 연결하는 신호선(15)과, 일반적인 사람의 구강부에 위치하도록 신호선 상에 설치된 마이크(30), 그리고 마이크(30)와 커넥터 사이의 신호선 사이에 연결되며, 핸드폰의 호출음에 응답할 수 있도록 호출응답버튼(42)을 갖는 제어기(40)를 구비한다.

위와 같은 마이크 부가형 이어폰은 핸드폰이 호출된 경우, 사용자가 제어기에 형성된 호출응답버튼(42)을 누르는 것으로, 핸드폰의 직접적인 조작없이 마이크(30)와 스피커(20)를 통해 상대방과 통화를 할 수 있게 된다.

그러나 상기와 같은 마이크 부가형 이어폰은 사용자가 몸을 움직이거나 이동중인 경우에는 마이크(30)가 사용자의 구강부로부터 이격될 수 있어 효과적으로 메시지를 전달하지 못하는 문제점이 있었다. 또한, 마이크(30)가 외부 소음에 노출되 어 있는 관계로, 소음이 많은 지역인 경우, 주변 소음까지 상대방에게 함께 전달되어 상대방에게 사용자의 메시지를 제대로 전달하지 못하는 문제점이 있었다.

위와 같은 문제점을 해결하기 위하여 본 출원인은 인체의 구강부와 귀가 서로 연통되어 있는 특징을 이용하여 마이크로의 외부소음 인입을 저감시킬 수 있으며, 사용자의 동작 중에도 효과적으로 메시지 전달을 수행할 수 있는 음향 입/출력장치 및 이를 적용한 마이크 일체형 이어폰을 제안한 바 있다(출원번호 10-2001-0061260). 또한, 상기와 같은 마이크 일체형 이어폰에서 마이크 및 스피커가 근접 설치됨에 따라 발생하는 하울링 및 에코를 저감시키도록 개선한 음향 입/출력장치 및 이를 적용한 마이크 일체형 이어폰을 제안한 바 있다(출원번호 10-2001-0088886).

Exhibit 1009
Page 1694 of 2198

제안한 종래의 음향 입출력장치 및 이를 적용한 마이크 일체형 이어폰은 신호처리부를 구비하며, 신호처리부는 송수신모드에 따라 마이크 입력레벨과 스피커 출력레벨을 조정한다. 즉, 신호처리부는 수화음이 검출되면, 마이크의 입력레벨이 낮아지도록 제어하며, 마이크 입력이 검출되면, 스피커 출력레벨이 낮아지도록 제어함으로서 마이크와 스피커의 근접설치에 따른 통화 간섭을 최소화시키며, 하울링 및 에코의 발생을 저감시키도록 하였다.

그러나 위와 같은 종래의 음향 입출력장치 및 이어폰은, 신호처리부를 이용하여 양방향통신을 수행할 수 있는 잇점은 있지만 비소음지역에서도 반드시 신호처리가 수행되므로 불필요한 전력소모를 발생시키는 문제점이 있었다. 또한, 비소음지역에서도 사람의 구강부를 통한 명확한 음성이 전달되지 못하고, 귀에서 발생된 비교적 부정확한 음성이 전달될 수 밖에 없다는 문제점이 있었다.

발명이 이루고자 하는 기술적 과제

본 발명의 목적은 상기와 같은 문제점을 해결하기 위하여 비소음지역인 경우, 전력소모를 저감시킬 수 있으면서 구강부에서 발생된 명확한 음성으로 메시지를 전달할 수 있는 음향 입출력장치 및 이를 적용한 마이크 일체형 이어폰을 제공하는 데 있다.

발명의 구성 및 작용

상기의 목적을 달성하기 위한 본 발명의 음향 입출력장치는, 전달된 전기적신호를 음향으로 변환하여 출력하는 스피커; 상기 스피커와 일체로 설치되며, 발생된 음향을 전기적신호로 변환하여 출력하는 제1 마이크로폰; 발생된 음향을 전기적신호로 변환하여 출력하는 제2 마이크로폰; 및 상기 제1 마이크로폰 및 상기 제2 마이크로폰을 선택적으로 내부입력단에 연결하는 모드전환스위치;를 포함한다.

또한, 음향 입출력장치는, 상기 스피커에 전달되는 전기적신호가 검출되면, 상기 제1 마이크로폰으로부터 발생된 전기적신호를 단속하기 위한 단속신호를 출력하는 신호검출부; 및 상기 신호검출부로부터 출력된 단속신호에 따라 상기 제1 마이크로폰으로부터 발생된 전기적신호를 스위칭하는 스위칭소자;를 더 포함한다.

그리고 상기 모드전환스위치는 상기 제2 마이크로폰을 통해 음향을 입력받도록 조작된 경우, 상기 신호검출부에 공급되는 전원을 차단하도록 구성한다.

상기의 목적을 달성하기 위한 본 발명에 따른 음향 입출력장치를 적용한 마이크 및 스피커 일체형 이어폰은, 메인장치와 전기적으로 연결하기 위한 커넥터; 사용자의 귀에 일부가 삽입될 수 있도록 외관이 형성된 몸체 내에 설치되며, 상기 커넥터를 통해 메인장치로부터 전달된 전기적신호를 음향으로 변환하여 출력하는 스피커; 상기 몸체 내부에 설치되며, 상기 사용자의 귀에서 발생된 음향을 전기적신호로 변환하여 상기 커넥터를 통해 상기 메인장치에 출력하는 제1 마이크로폰; 상기 몸체 외부에 설치되며, 발생된 음향을 전기적신호로 변환하여 상기 메인장치에 출력하는 제2 마이크로폰; 및 사용자의 조작에 의해 상기 제1 마이크로폰 및 상기 제2 마이크로폰을 선택적으로 상기 커넥터의 음성출력단에 연결하는 모드전환스위치;를 포함한다.

또한, 상기 마이크 및 스피커 일체형 이어폰은, 상기 커넥터로부터 상기 스피커에 전달되는 전기적신호가 검출되면, 상기 제1 마이크로폰으로부터 상기 커넥터에 전달되는 전기적신호를 단속하기 위한 단속신호를 출력하는 신호검출부; 및 상기 제1 마이크로폰과 상기 커넥터 사이에 연결되며, 상기 신호검출부로부터 출력된 단속신호에 따라 개폐되는 스위칭소자;를 더 포함한다. 여기서, 상기 모드전환스위치는 상기 제2 마이크로폰을 통해 음향을 입력받도록 조작된 경우, 상기 신호검출부에 공급되는 전원을 차단할 수 있도록 구현한다.

또한, 상기 제2 마이크로폰은 상기 본체 외부에 다양하게 설치할 수 있으며, 즉, 종래와 같이 신호선 상에 설치할 수도 있으며, 상기 본체로부터 사용자의 구강부측으로 돌출시킨 바 형태를 취할 수도 있다. 바람직하게는, 상기 본체의 귀 삽입부에 대향하여 상기 본체와 일체로 형성하며, 이때, 사용자의 입에서 발성되는 소리를 집음하기 위한 지향성 마이크로폰을 이용한다.

이상과 같은 본 발명의 음향 입출력장치 및 이를 적용한 마이크 및 스피커 일체형 이어폰에 의하면, 사용자는 외부환경에 따라, 즉, 주변에 소음이 많고 적음에 따라 상기 모드전환스위치를 통해 상기 제1 마이크로폰 및 상기 제2 마이크로폰을 선택적으로 사용할 수 있으며, 제2 마이크로폰 이용시는 전력소모를 감소시킬 수 있을뿐만아니라 구강부에서 발생된 음성으로 메시지를 보다 정확하게 전달할 수 있다.

이하 첨부한 도면을 참조하여 본 발명을 상세하게 설명한다.

Exhibit 1009
Page 1695 of 2198

도 2는 본 발명의 실시예에 따른 목걸이형 핸드폰용 이어폰의 사시도이다. 핸드폰용 이어폰은 커넥터(110), 마이크 일체형 이어폰(120), 그리고 제어기(130)를 포함한다.

커넥터(110)는 핸드폰(200)의 입출력단자와 연결된다.

마이크 일체형 이어폰(120)은 단일의 몸통을 가지며, 사용자의 귀에 삽입되는 삽입부(122)와 귀 외측으로 돌출되는 돌출부(124)를 갖는다. 삽입부(122)측에는 사용자의 귀를 보호하기 위하여 그리고 사용자의 귀에 밀착 또는 밀폐시키기 위하여 인체적합형 고무(또는 스폰지)가 덮어 씌어진다. 그리고 도시되지는 않았지만 이어폰 내부에는 몸체 내에 스피커와 내장형마이크가 사용자의 귀 내부를 향하여 설치되며, 돌출부측으로 스피커 및 내장형마이크와는 차폐물을 통해 분리된 외장형마이크가 내장된다. 편의상 몸통 내부에 내장된 마이크들에 대하여 삽입부(120)측으로 형성된 마이크는 내장형마이크로, 돌출부(124)측으로 형성된 마이크는 외장형마이크로 명명한다.

제어기(130)는 커넥터(110)와 마이크 일체형 이어폰(120)을 연결하는 신호선 사이에 설치되며, 사용자에 의해 소음모드 또는 일반모드로 조작할 수 있는 모드전환조작스위치(132) 및 호출응답 겸용 재다이얼링버튼(134)을 갖는다. 소음모드는 내장형마이크를 통해 입력된 사용자의 메시지를 핸드폰(200)에 출력한다. 일반모드는 외장형마이크를 통해 입력된 사용자의 메시지를 핸드폰(200)에 출력한다.

도 3은 도 2에 보인 핸드폰용 이어폰의 블록도이다. 이어폰은 스피커(210), 내장형마이크(220), 외장형마이크(230), 신호검출부(240), 스위칭소자(250), 그리고 모드전환스위치(260)를 구비한다.

스피커(210)는 사용자의 귀에 일부가 삽입될 수 있도록 외관이 형성된 몸통(120) 내부에 설치되며, 커넥터(110)를 통해 핸드폰(200)으로부터 전달된 전기적신호를 음향으로 변환하여 사용자의 귀에 출력한다.

내장형마이크(220)는 몸통 내부에 스피커와 함께 내장 설치되며, 사용자의 귀에서 발생된 음향을 전기적신호로 변환하여 커넥터(110)를 통해 핸드폰(200)에 출력한다.

외장형마이크(230)는 몸통(120)의 삽입부(122)에 대향하여 돌출부(124)측에 설치되며, 사용자의 구강부에서 발생된 음향을 전기적신호로 변환하여 커넥터(110)를 통해 핸드폰(200)에 출력한다. 이때, 외장형마이크(230)는 사용자의 구강부로부터 발생되는 음성을 집음할 수 있도록 지향성 마이크가 이용되는 것이 바람직하다.

신호검출부(240)는 스피커(210)로 출력되는 전기적신호를 검출하고, 내장형마이크(220)로부터 핸드폰(200)의 입력단에 전달되는 음성신호를 단속하기 위한 단 속신호를 출력한다.

스위칭소자(250)는 내장형마이크(220)와 마이크 출력단자 사이에 연결되며, 신호검출부(240)로부터 출력된 단속신호에 따라 개폐된다. 스위칭소자(250)는 내장형마이크(220)와 마이크 출력단자 사이를 신호검출부(240)의 차단/연결신호에 응답하여 차단/연결할 수 있는 다양한 소자가 이용될 수 있다. 예를 들어, 트랜지스터와 같은 반도체소자가 이용될 수 있으며, 릴레이와 같은 소자를 이용할 수도 있다.

모드전환스위치(260)는 사용자의 모드전환조작스위치(262)의 조작에 의해 내장형마이크 연결단자 및 외장형마이크 연결단자를 선택적으로 마이크 출력단자에 연결하는 스위칭부(264)를 구비한다. 모드전환조작스위치(260)는 슬라이드 형태로 내장형마이크(220) 및 외장형마이크(230)를 선택할 수 있으며, 푸쉬버튼 형식으로 어느 한 마이크를 선택할 수도 있다.

위와 같은 핸드폰용 이어폰은 도 2에 보인 바와 같이 제어기(130) 하단에 핸드폰고리를 형성하고, 핸드폰을 그 고리에 걸수 있도록 할 수 있으며, 사용자가 줄을 목에 걸어 핸드폰(200)을 사용자의 몸에 고정시킬 수 있다.

한편, 커넥터(110)가 핸드폰(200)의 입출력단자에 연결되고, 마이크 일체형 이어폰 몸체(120) 일부가 사용자의 귀에 삽입된 상태에서 핸드폰(200)이 호출된 경우, 사용자는 제어기(130)의 호출응답버튼(134)을 눌러 상대방과 통화를 할 수 있게 된다. 이때, 슬라이드 방식의 모드전환조작스위치(262)의 위치가 소음모드측으로 위치한 경우, 즉, 내장형마이크(220)를 연결시킨 경우, 신호검출부(240)는 핸드폰(200)으로부터 스피커(210)로 전달되는 전기적신호가 검출되면, 내장형마이크(220)를 통한 입력신호를 차단하기 위한 차단신호를 스위칭소자(250)에 출력한다. 이에 따라 스위칭소자(250)는 개방된다. 그러면 내장형마이크(220)를 통한 사용자의 음성신호는 핸드폰(200)에 전달되지 않게 된다. 반면, 신호검출부(240)는 스피커(210)로 전달되는 전기적신호가 검출되지 않는 동안은 연결신호를 스위칭소자(250)에 출력한다. 이에 따라 스위칭소자(250)는 단락상태가 되며, 내장형마이크(220)를 통해 입력된 음성신호는 핸드폰(200)에 출력된다. 이는 두 통화자가 동시에 송화를 하는 경우, 서로 메시지를 정확히 전달받을 수 없어 어느 한 측에서 송화를 중단하는 점을 감안하여 양방향 통신이 아닌 단방향 통신을 수행하도록 한 것이다. 이에 따라 별도의 신호처리로직을 이용하지 않고도 스피커(210) 및 내장형마이크(220)의 근접설치에 따른 하울링 및 에코를 방지 할 수 있게 된다.

Exhibit 1009
Page 1696 of 2198

또한, 핸드폰용 이어폰은 외부환경에 따라 내장형마이크(220) 및 외장형마이크(230)를 이용할 수 있다. 즉, 주변에 소음이 많은 지역인 경우, 사용자는 모드전환조작스위치(262)를 소음모드로 위치시켜 이어폰에 내장된 내장형마이크(220)를 통해 자신의 음성을 상대방에게 전달한다. 그리고 주변에 소음이 적은 지역인 경우에는 모드전환조작스위치(262)를 일반모드로 전환하고, 마이크 일체형 이어폰(120)의 돌출부(124)에 설치된 외장형마이크(230)를 통해 자신의 음성을 상대방에게 전달할 수 있다. 또한, 모드전환스위치(260)는 일반모드로 조작된 경우, 신호검출부(240)로 공급되는 전원을 차단한다. 이는, 사용자의 음성을 외장형마이크를 통해 전달하는 경우, 내장형마이크(220)에 의한 하울링 및 에코를 고려하지 않아도 되므로 신호검출부(240)의 동작에 따른 전력소모를 감소시키기 위한 것이다.

위와 같은 핸드폰용 마이크 및 스피커 일체형 이어폰은 사용자가 모드전환스위치 조작을 통해 소음이 많은 지역에서는 사용자의 귀에서 발생되는 소리로 메시지를 전달할 수 있으며, 주변 소음이 비교적 적은 지역에서는 사용자의 입으로 출력되는 정확한 음성으로 메시지를 전달할 수 있게 된다.

발명의 효과

이상과 같은 본 발명의 음향입출력장치 및 마이크 일체형 이어폰은 단방향 통신을 수행하므로 회로의 구현이 간단하며, 사용자가 외부환경에 따라 스위치를 조작하여 내장형마이크 및 외장형마이크를 선택할 수 있어 불필요한 전력소모를 감소시킬 수 있을 뿐만아니라 상황에 따라 상대방에게 보다 명확한 메시지를 전달할 수 있게 된다.

이상에서는 본 발명의 바람직한 실시예에 대하여 설명하였으나, 본 발명은 상술한 실시예에 한정되지 아니하며, 특허청구의 범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형실시가 가능할 것이다.

(57) 청구의 범위

청구항 1.
전달된 전기적신호를 음향으로 변환하여 출력하는 스피커;

상기 스피커와 일체로 설치되며, 발생된 음향을 전기적신호로 변환하는 제1 마이크로폰;

발생된 음향을 전기적신호로 변환하는 제2 마이크로폰; 및

사용자의 조작에 의해 상기 제1 마이크로폰 및 상기 제2 마이크로폰을 선택적으로 음성입력단에 연결하는 모드전환스위치;를 포함하는 것을 특징으로 하는 음향 입출력장치.

청구항 2.
제 1항에 있어서,

상기 스피커에 전달되는 전기적신호가 검출되면, 상기 제1 마이크로폰으로부터 발생된 전기적신호를 단속하기 위한 단속신호를 출력하는 신호검출부; 및

상기 신호검출부로부터 출력된 단속신호에 따라 상기 제1 마이크로폰으로부터 발생된 전기적신호가 스위칭될 수 있도록 개폐되는 스위칭소자;를 더 포함하는 것을 특징으로 하는 음향 입출력장치.

청구항 3.
제 2항에 있어서,

상기 모드전환스위치는 상기 제2 마이크로폰으로부터 음향을 입력받도록 조작된 경우, 상기 신호검출부에 공급되는 전원을 차단하도록 구성된 것을 특징으로 하는 음향 입출력장치.

청구항 4.
메인장치와 전기적으로 연결하기 위한 커넥터;

사용자의 귀에 일부가 삽입될 수 있도록 외관이 형성된 몸체 내부에 설치되 며, 상기 커넥터를 통해 상기 메인장치로부터 전달된 전기적신호를 음향으로 변환하여 출력하는 스피커;

Exhibit 1009
Page 1697 of 2198

상기 몸체 내부에 설치되며, 사용자의 귀에서 발생된 음향을 전기적신호로 변환하여 상기 커넥터를 통해 상기 메인장치에 출력하는 제1 마이크로폰;

발생된 음향을 전기적신호로 변환하여 상기 메인장치에 출력하는 제2 마이크로폰; 및

사용자의 조작에 의해 상기 제1 마이크로폰 및 상기 제2 마이크로폰을 선택적으로 상기 커넥터의 음성출력단에 연결하는 모드전환스위치;를 포함하는 것을 특징으로 하는 마이크 및 스피커 일체형 이어폰.

청구항 5.
제 4항에 있어서,

상기 커넥터로부터 상기 스피커에 전달되는 전기적신호가 검출되면, 상기 제1 마이크로폰으로부터 상기 커넥터에 전달되는 전기적신호를 단속하기 위한 단속신호를 출력하는 신호검출부; 및

상기 제1 마이크로폰과 상기 커넥터 사이에 연결되며, 상기 신호검출부로부터 출력된 단속신호에 따라 개폐되는 스위칭소자;를 더 포함하는 것을 특징으로 하는 마이크 및 스피커 일체형 이어폰.

청구항 6.
제 5항에 있어서,

상기 모드전환스위치는 상기 제2 마이크로폰을 통해 음향을 입력받도록 조작 된 경우, 상기 신호검출부에 공급되는 전원을 차단하는 것을 특징으로 하는 마이크 및 스피커 일체형 이어폰.

청구항 7.
제 4항에 있어서,

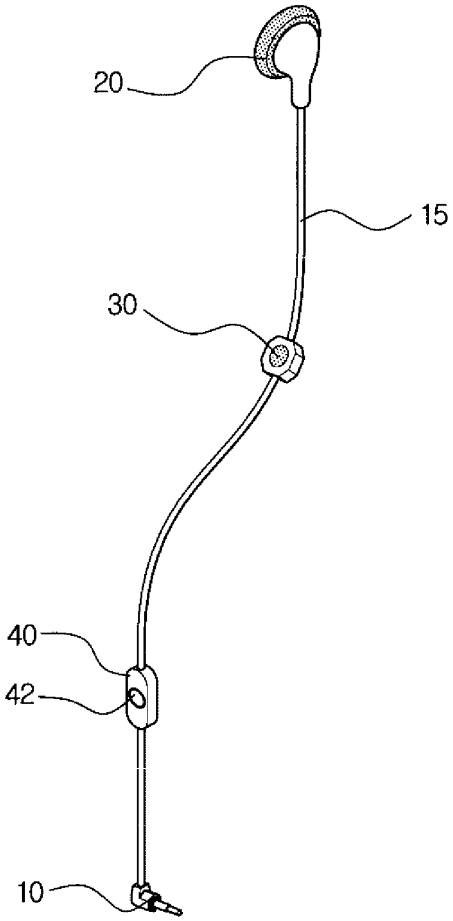
상기 제2 마이크로폰은 상기 본체의 귀 삽입부에 대향하여 상기 본체와 일체로 형성된 것을 특징으로 하는 마이크 및 스피커 일체형 이어폰.

청구항 8.
제 7항에 있어서,

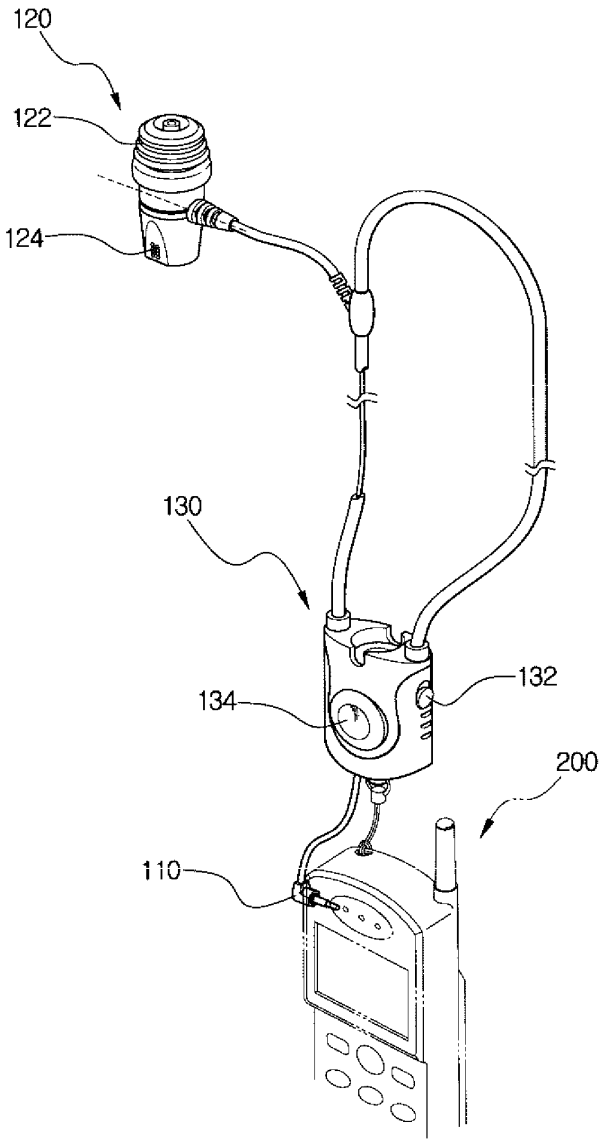
상기 제2 마이크로폰은 지향성 마이크로폰인 것을 특징으로 하는 마이크 및 스피커 일체형 이어폰.

도면

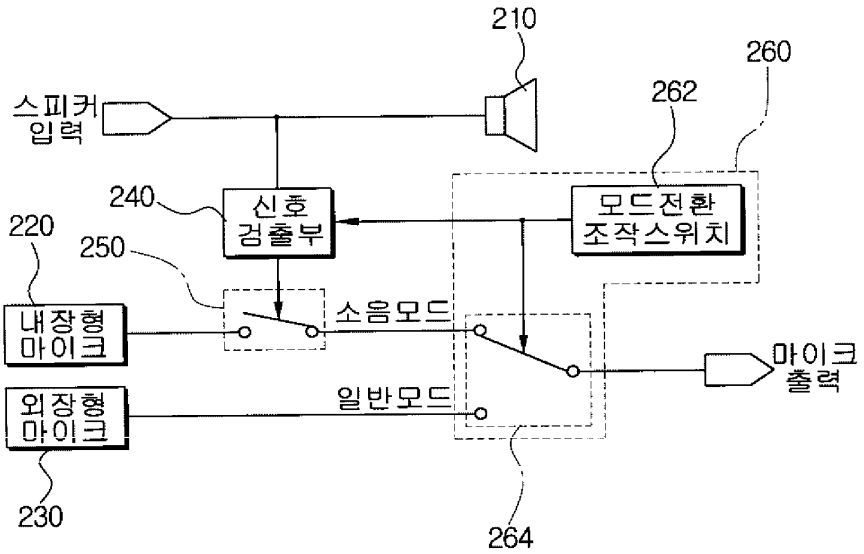
도면1



도면2



도면3


210
260
262
스피커
입력
240
신호
검출부
모드전환
조작스위치
220
250
내장형
마이크
소음모드
마이크
출력
일반모드
외장형
마이크
230
264

(19)日本国特許庁（ＪＰ）　(12)特　許　公　報（Ｂ２）　(11)特許番号
特許第3085237号
(P3085237)
(45)発行日　平成12年９月４日(2000.9.4)　(24)登録日　平成12年７月７日(2000.7.7)

(51)Int.Cl.[7]	識別記号	ＦＩ	
Ｈ０４Ｒ　25/00		Ｈ０４Ｒ　25/00	Ｈ
			Ｍ
Ｇ１０Ｌ　21/04		Ｇ１０Ｌ　3/02	Ａ

請求項の数１(全　7　頁)

(21)出願番号　特願平9－50472
(22)出願日　平成９年３月５日(1997.3.5)
(65)公開番号　特開平10－248100
(43)公開日　平成10年９月14日(1998.9.14)
審査請求日　平成10年12月14日(1998.12.14)
前置審査

(73)特許権者　000004075
ヤマハ株式会社
静岡県浜松市中沢町10番１号
(72)発明者　近藤　克文
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(72)発明者　林　好典
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(72)発明者　三木　晃
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(74)代理人　100084548
弁理士　小森　久夫
審査官　松澤　福三郎

最終頁に続く

(54)【発明の名称】　話速変換装置

(57)【特許請求の範囲】
【請求項１】　外部から入力される外部音声信号を集音する外部音声集音手段と、
装用者が発する自己発話音声信号を検出および集音する自己発話音集音手段と、
前記外部音声信号を時間的に伸長して出力バッファに書き込む手段であって、前記自己発話音声信号が検出されているとき、前記外部音声集音手段からの入力にかかわらず、無音を示す信号を前記出力バッファに書き込む話速変換手段と、
前記自己発話音声信号の検出の有無とは無関係に、先に前記出力バッファに書き込まれている過去の信号を順次読み出し、この信号および前記自己発話音声信号を合成して出力する信号出力手段と、
を備えたことを特徴とする話速変換装置。

【発明の詳細な説明】
【０００１】
【発明の属する技術分野】この発明は、補聴器などに用いられ、入力された音声信号を時間的に伸長する話速変換装置に関する。
【０００２】
【従来の技術】従来より、高齢者など聴覚機能が低下した者が装用する機能補助装置として補聴器が使用されている。ところで、老齢化による聴覚機能の低下は、最小可聴信号レベルの上昇，高音域の聴取機能の低下などの伝音系機能低下のほか音声識別臨界速度（語音を識別することができる最大の話速）の低下などの聴覚中枢系の機能低下も含まれている。
【０００３】このため、高齢者用の補聴器として、音声信号の周波数帯域の一部または全部を増幅することに加

えて、音声信号を時間的に伸長して音声信号の出力速度を入力された速度よりも低速にする話速変換処理を行う補聴器も提案されている。

【０００４】

【発明が解決しようとする課題】しかし、補聴器のマイクには装用者が自ら発した音声（自己発話音声）も入力され、話速変換処理においてはこの音声信号も時間的に伸長してしまう。人間は、自分が発した音声を耳からフィードバックしながら話しているが、このフィードバック音声が時間的に伸長されてしまうと、正常に話すことができない。そこで、装用者が話しているときには入力音声を減衰させることによって、話速変換による速さの変化で自己発話が妨害されないようにしたものが提案されている。

【０００５】しかし、入力音声を減衰させると自ら発した音声信号も減衰されてしまうため、装用者に自己発話音が正常にフィードバックされなくなり、大きな声を出したり自分の声が適当でないと不安になったりしてスムーズな会話を妨げてしまう問題点があった。

【０００６】この発明は、装用者が話しているときは、話速変換を中断してこの自己発話音をフィードバックできるようにした話速変換装置を提供することを目的とする。

【０００７】

【課題を解決するための手段】この発明は、外部から入力される外部音声信号を集音する外部音声集音手段と、装用者が発する自己発話音声信号を検出および集音する自己発話音集音手段と、前記外部音声信号を時間的に伸長して出力バッファに書き込む手段であって、前記自己発話音声信号が検出されているとき、前記外部音声集音手段からの入力にかかわらず、無音を示す信号を前記出力バッファに書き込む話速変換手段と、前記自己発話音声信号の検出の有無とは無関係に、先に前記出力バッファに書き込まれている過去の信号を順次読み出し、この信号および前記自己発話音声信号を合成して出力する信号出力手段と、を備えたことを特徴とする。

【０００８】この発明では、外部から入力される音声信号を集音して話速変換するが、この外部から入力される音声信号のなかには、この装置の装用者が発する自己発話音声信号も含まれている。したがって、装用者が話しているとき、この話速変換手段で時間的に伸長した音声を出力すると、装用者が実際に話した音声と耳からフィードバックされる音声が完全にずれてしまうため、正常に話すことができない。そこで、装用者が話しているときは、外部音声入力手段および話速変換手段を無効にしてバッファに無音を示す信号を書き込む。これで、話速変換された自己発話音により発声が妨げられることがなくなる。一方、自己発話音が全くフィードバックされないと、やはり正常な発声が困難であるため、自己発話音集音手段で集音した自己発話音を出力する。これにより、自己発話は話速変換されないでフィードバックされるため正常な発話が可能である。

【０００９】さらに、話速変換された外部音声は、時間的に伸長された出力音声信号と入力信号との時間ずれを吸収するバッファに蓄積されるが、装用者の発話中もこれを捨てずにバッファから順次読み出し、自己音声と加え併せて出力することにより、外部音声が不自然に中断されることがない。

【００１０】

【発明の実施の形態】図１はこの発明の実施形態である話速変換機能を備えた補聴器のブロック図である。この補聴器は、外部音声を収集する外部音声マイク１と装用者自身の発話音声を収集する自己発話音声マイク５とを備えている。外部音声マイク１は外部に向けて設置されるものであり、空気中を伝搬してくる音声に対して反応する。なお、この空気中を伝搬する音声には装用者が発する自己発話音声も含まれている。自己発話音声マイク５は、外耳道内に置かれた骨伝導マイクのように発声に対応して生じる振動を検出するものや発声による外耳道の形状歪みを検出するものなどが適用され、空気中を伝搬してくる音声には反応しない。空気中を伝搬してくる音声は外部音声マイク１で電気信号に変換され、前置増幅器２で一定レベルまで増幅される。増幅された信号は、アンチエリアシングフィルタ３でサンプリング周波数を超える高域成分をカットされたのち、Ａ／Ｄ変換器４によってアナログ信号からディジタル信号に変換される。このディジタル信号は信号処理プロセッサ１６に入力される。信号処理プロセッサ１６はこの信号を入力バッファ１７の外部入力バッファ１７ａに蓄積する。外部入力バッファ１７ａは３０ｍｓを１フレームとして４フレーム（１２０ｍｓ）の記憶容量を有しており、信号処理はフレームを単位として行われる。また、自己発話音声マイク５が集音した自己発話音声信号は、前置増幅器６－アンチエリアシングフィルタ７－Ａ／Ｄ変換器８を経て信号処理プロセッサ１６に入力される。信号処理プロセッサ１６はこのディジタル信号を入力バッファ１７の発話入力バッファ１７ｂに書き込む。この発話入力バッファ１７ｂも前記外部入力バッファ１７ａと同様、３０ｍｓのフレームを４フレーム備えており、１２０ｍｓ分の記憶容量を有している。また、前置増幅器６の出力は波形整形回路１１によって波形整形されたのち信号処理プロセッサ１６に入力される。信号処理プロセッサ１６は、この波形整形回路１１から入力される信号のレベルによって自己発話があるか無いかを検出する。

【００１１】信号処理プロセッサ１６は、３つの処理を並行して処理する。第１は上述した信号入力処理であり、Ａ／Ｄ変換器４から入力される外部音声の音声データを外部入力バッファ１７ａに書き込むとともに、Ａ／Ｄ変換器８から入力される自己発話音声の音声データを発話入力バッファ１７ｂに書き込む処理である。第２は

外部入力バッファ１７ａに書き込まれている音声データを時間軸伸長などして出力バッファ１８に書き込む話速変換処理である。第３は発話入力バッファ１７ｂおよび出力バッファ１８に記憶されている音声データを加算合成してＤ／Ａ変換器１５に出力する出力処理である。

【００１２】なお、Ｄ／Ａ変換器１５の後段には折り返しノイズを除去するためのローパスフィルタ１４および増幅器１３が接続されており、増幅器１３で適当なレベルに増幅された音声信号は、音声出力装置１２から音響として出力される。音声出力装置１２は、たとえば、装用者が装用するイヤホンなどである。

【００１３】話速変換処理において、信号処理プロセッサ１６は、上記外部入力バッファ１７ａおよび発話入力バッファ１７ｂに記憶されている信号を最大１２０ｍｓ遅れで処理する。通常の発話および会話ではこの程度の遅れは許容されるからである。自己発話が検出されない状態では、信号処理プロセッサ１６は、外部入力バッファ１７ａから音声データを読み出して時間軸伸長などの処理をしたのち出力バッファ１８に書き込む。一方、自己発話が検出されている場合には、外部入力バッファ１７ａから読み出した音声データを破棄し、時間軸伸長処理もすることなく０レベルを示すデータを出力バッファ１８に書き込む。これは、外部音声信号にも空気中に放出された自己発話音声が含まれているため、これを時間軸伸長して出力すると、装用者に対して時間的に遅れたフィードバックが掛かることになり自己発話を正常に行うことができなくなるためである。

【００１４】出力処理において、信号処理プロセッサ１６は、サンプリングクロックにしたがって、出力バッファ１８および発話入力バッファ１７ｂに記憶されている音声データを読み出して加算合成する。出力バッファ１８に記憶されている音声データは時間軸伸長処理などの処理が施されたものであるが、発話入力バッファ１７ｂに記憶されている信号は、入力された音声信号をディジタル化したのみのものである。したがって、両信号は時間的にずれているが自己発話音と外部音声とが多少時間的にずれて重なっても不自然ではないため時間軸伸長された信号の出力時間を長く確保するためこのような処理を行う。ただし、上述したように自己発話されている間の外部音声信号には自己発話音声信号が含まれているため、この間に出力バッファ１８に書き込まれる音声データは０レベル信号である。したがって、出力バッファ１８からこのデータを読み出している間は、外部音声信号成分は０である。

【００１５】図２を参照して、信号処理プロセッサ１６の処理について説明する。図２において、外部入力バッファ１７ａにはフレームａ～ｇの外部音声信号の音声データが書き込まれる。また、発話入力バッファ１７ｂにはフレームＡ～Ｇの自己発話音声の音声データが書き込まれる。実際には、外部入力バッファ１７ａは４フレーム設けられているが、この図においては説明を分かりやすくするため時系列に７フレーム分を示している。書込ポインタは外部入力バッファ１７ａ，発話入力バッファ１７ｂとも同じタイミングのアドレスを指示しながら移動する。また、発話入力バッファ１７ｂおよび出力バッファ１８の読出ポインタは、同じタイミングをアドレスを指示しながら移動する。一方、外部入力バッファ１７ａの読出ポインタは、話速変換処理を行うために発話入力バッファ１７ｂ，出力バッファ１８の読出ポインタよりも先を指示するように移動する。

【００１６】外部入力バッファ１７ａから読み出された信号は時間軸伸長処理などの話速変換処理を施されたのち出力バッファ１８に書き込まれる。発話入力バッファ１７ｂから読み出された信号および出力バッファから読み出された信号は加算合成されたのちＤ／Ａ変換器１５に出力される。ただし、この図では発話入力バッファ１７ｂのフレームＡ～Ｄは自己発話されておらず、信号がないため、加算合成しても出力バッファ１８の信号のみが出力される。一方、発話入力バッファ１７ｂのフレームＥ～Ｇでは装用者が発話しており、一定レベル以上の信号が入力されている。したがって、これに対応する外部入力バッファ１７ａのフレームｅ～ｇの信号は破棄され、これに対応して同じ長さ（３０ｍｓ）の０レベル信号が出力バッファ１８に書き込まれる。そして、この自己発話信号の存在するフレームＥ～Ｇが読み出されるとき、すなわち、発話入力バッファ１７ｂのフレームＥの先頭に読出ポインタが来たとき、出力バッファ１８の読出ポインタは時間軸伸長されたｄ′のフレームの読み出し途中を指示しており、この外部音声信号と自己発話信号が加算合成して出力される。このように出力バッファの信号は伸長されているため、発話入力バッファの信号とは時間ズレが生じているが、この外部音声信号には自己発話音が含まれておらず異なる種類の信号であるため問題はない。ただし、出力バッファ１８のフレームｅ′～ｇ′は０レベル信号であるためこのフレームを加算合成しても、外部音声信号の成分は０である。

【００１７】図３は前記信号処理プロセッサ１６の動作を示すフローチャートである。同図（Ａ）は信号入力処理動作を示すフローチャート、同図（Ｂ）は話速変換処理動作を示すフローチャート、同図（Ｃ）は信号出力処理動作を示すフローチャートである。これらの処理動作は並行して処理される。

【００１８】同図（Ａ）の信号入力制御動作において、まず、Ａ／Ｄ変換器４から外部音声信号のディジタルデータを入力し（ｓ１）、外部入力バッファ１７ａの書込ポインタの示すアドレスに書き込む（ｓ２）。つぎに、Ａ／Ｄ変換器８から自己発話音声信号のディジタルデータを入力して（ｓ３）、発話入力バッファ１７ｂの書込ポインタの示すアドレスに書き込む（ｓ４）。上記外部入力バッファ１７ａの書込ポインタおよび発話入力バッ

ファ１７ｂの書込ポインタは同じタイミングのアドレスを指示しながら移動する。

【００１９】同図（Ｂ）の話速変換処理動作において、外部入力バッファ１７ａの読出ポインタが示すアドレスから信号を読み出し（ｓ１０）、このタイミングに発話が検出されているかを波形整形回路１１から入力された信号に基づいて判断する（ｓ１１）。このタイミングに発話が検出されていた場合には（ｓ１２）、出力バッファ１８に０レベルを示す信号を書き込んで（ｓ１５）リターンする。一方、発話が検出されていない場合には（ｓ１２）、外部入力バッファ１７ａから読み出した外部音声信号に対して時間軸伸長処理を行い（ｓ１３）、この伸長された信号を出力バッファ１８の書込ポインタの示すアドレスに書き込む（ｓ１４）。

【００２０】同図（Ｃ）の信号出力処理において、出力バッファ１８の読出ポインタが指示するアドレスから信号を読み出し（ｓ２０）、発話入力バッファ１７ｂの読出ポインタが指示するアドレスから信号を読み出す（ｓ２１）。これらの読出ポインタは同期して移動している。これら読み出された信号を加算合成し（ｓ２２）、Ｄ／Ａ変換器１５に出力する（ｓ２３）。

【００２１】以上の動作により、自己発話されたときは同じタイミングの外部音声をマスクするとともに、この自己発話音声信号と既に時間軸伸長されて出力バッファ１８にバッファされている外部音声とを混合して出力することができる。

【００２２】なお、上記実施形態では、自己発話があるか否かを自己発話音声マイク５の前置増幅器６から分岐した波形整形回路１１の入力信号に基づいて判断するようにしているが、発話入力バッファ１７ｂに記憶されている音声データに基づいて判断するようにしてもよく。全く別の自己発話検出マイクを設けて、このマイクの入力レベルに基づいて判断するようにしてもよい。

【００２３】図４は、自己発話音声マイク５とは別の自己発話検出マイクを設けた実施形態を示すブロック図である。この図において図１に示した実施形態と同一構成の部分は同一符号を付して説明を省略する。この実施形態では、自己発話の有無を検出するために自己発話音声マイク５とは別の自己発話検出マイク９を設けている。この自己発話検出マイク９は前置増幅器１０を介して波形整形回路１１に接続されている。自己発話検出マイク９も骨伝導マイクなど装用者の発話信号のみに反応し空気中を伝搬する音声信号には反応しないものである。このマイクが集音した信号レベルを波形整形回路１１が整形して信号処理プロセッサ１６に入力する。信号処理プロセッサ１６はこの整形された信号レベルに基づいて自己発話の有無を検出する。

【００２４】図５は前記信号処理プロセッサ１６の他の動作を示す図である。上記実施形態では自己発話があるフレームでは、外部音声信号に変えて同じ長さ３０ｍｓ分の０レベル信号を書き込むようにしているが、この例では、自己発話が終了したとき、それまで書き込まれていた０レベル信号をキャンセルして時間遅れなしに時間軸伸長された外部音声信号が書き込むようにしている。これにより、時間軸伸長された外部音声信号の時間遅れを最小限にすることができる。

【００２５】

【発明の効果】以上のようにこの発明によれば、装用者が自己発話しているとき、この自己発話音声と他の音声信号とが一緒に集音されている外部音声信号を時間軸に伸長する話速変換をしないようにしたことにより、自己発話が時間遅れして装用者にフィードバックされることがなく自己発話に悪影響を及ぼさない。また、自己発話音声のみを集音した信号をフィードバックするようにしたことにより、装用者は自己発話を安定して行うことができる。さらに、このとき既に伸長された外部音声を捨てずに、自己発話音声信号と加算合成して出力するようにしたことにより、外部音声が不自然に中断されることなく、外部音声を最後まで聞かせることができる。

【図面の簡単な説明】

【図１】この発明の実施形態である話速変換機能を備えた補聴器のブロック図

【図２】同補聴器の話速変換機能を説明する図

【図３】同補聴器の信号処理プロセッサの動作を示すフローチャート

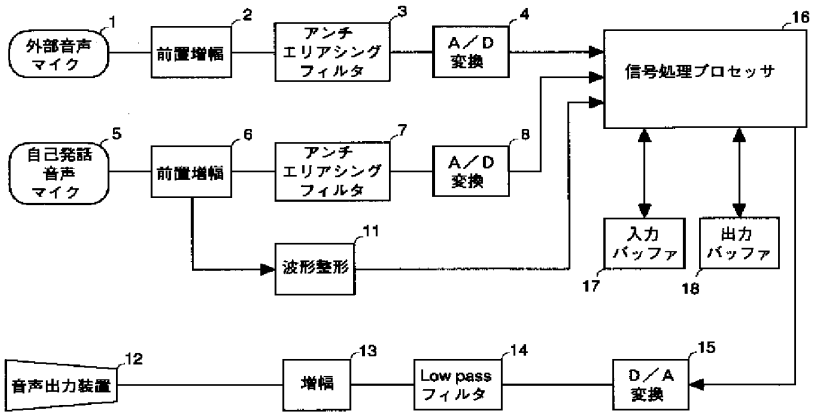
【図４】この発明の補聴器の他の実施形態を示すブロック図

【図５】この発明の話速変換機能の他の実施形態を説明する図

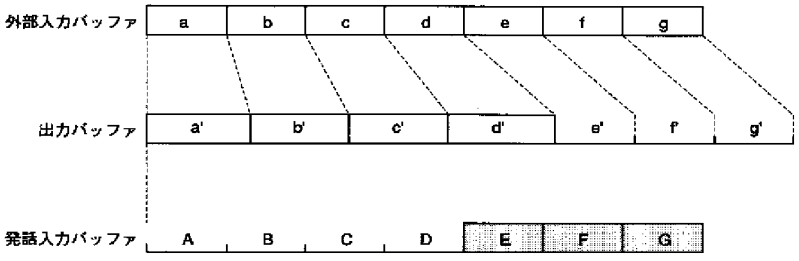
【符号の説明】

１…外部音声マイク、５…自己発話音声マイク、１６…信号処理プロセッサ、１７…入力バッファ、１７ａ…外部入力バッファ、１７ｂ…発話入力バッファ、１８…出力バッファ

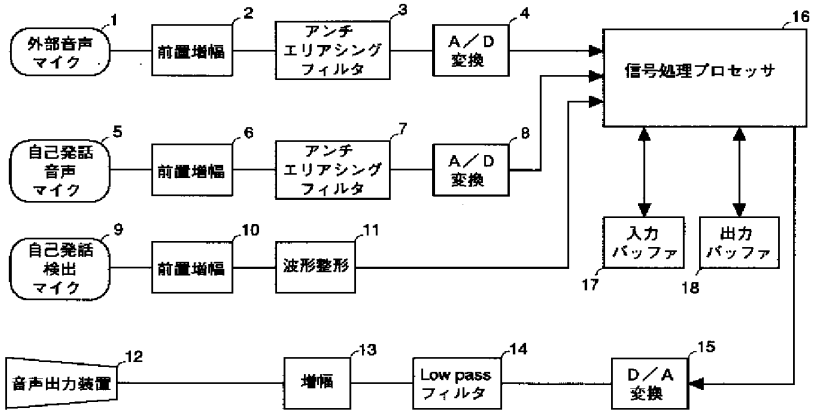
【図１】


外部音声マイク
1
前置増幅
2
アンチエリアシングフィルタ
3
A／D変換
4
信号処理プロセッサ
16
自己発話音声マイク
5
前置増幅
6
アンチエリアシングフィルタ
7
A／D変換
8
波形整形
11
入力バッファ
17
出力バッファ
18
音声出力装置
12
増幅
13
Low passフィルタ
14
D／A変換
15

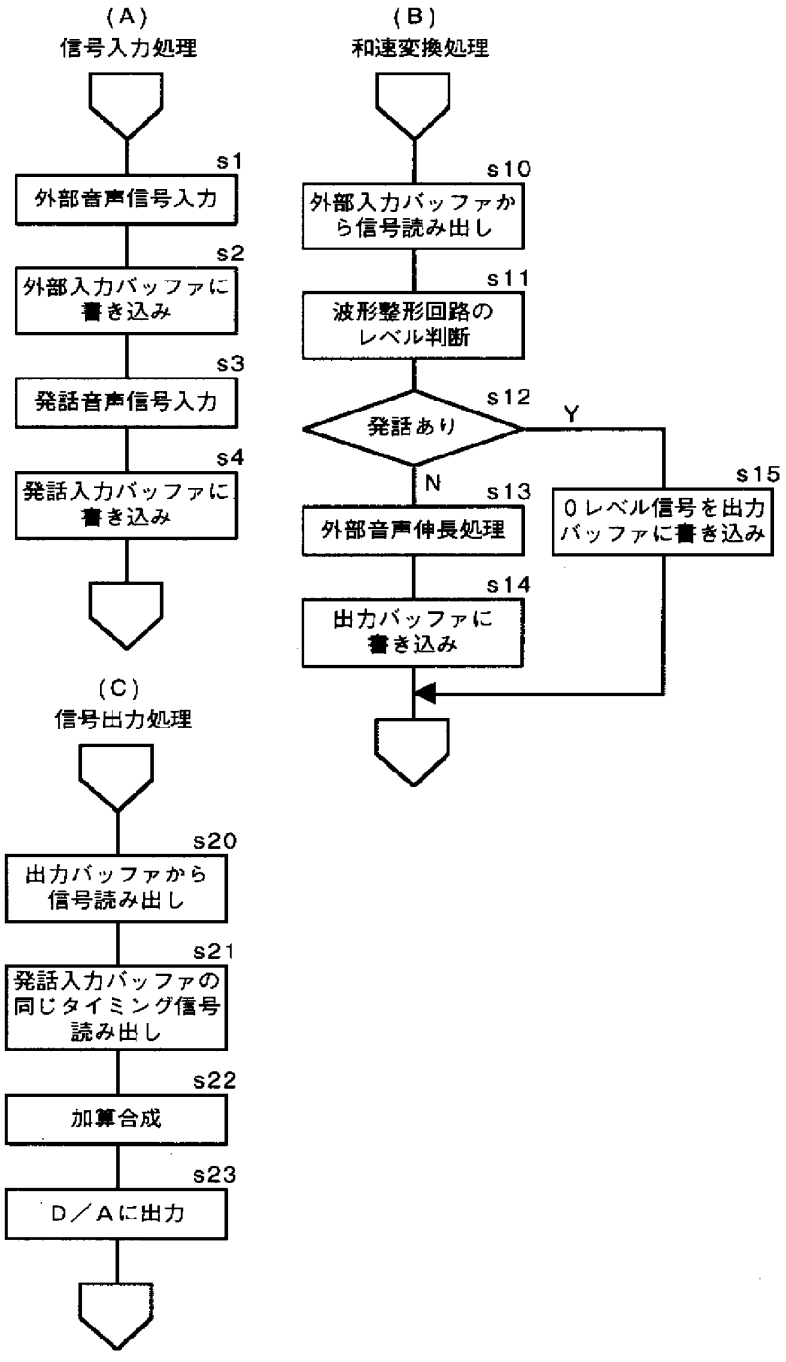
【図２】


外部入力バッファ
a
b
c
d
e
f
g
出力バッファ
a'
b'
c'
d'
e'
f'
g'
発話入力バッファ
A
B
C
D
E
F
G

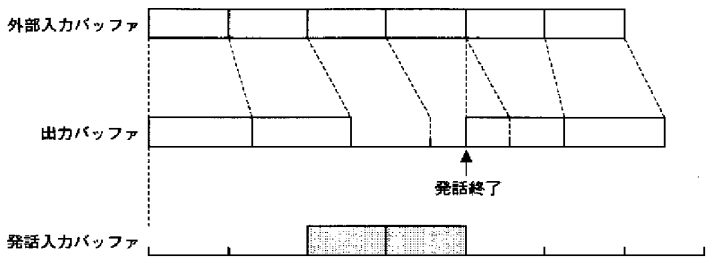
【図４】


外部音声マイク
1
前置増幅
2
アンチエリアシングフィルタ
3
A／D変換
4
信号処理プロセッサ
16
自己発話音声マイク
5
前置増幅
6
アンチエリアシングフィルタ
7
A／D変換
8
自己発話検出マイク
9
前置増幅
10
波形整形
11
入力バッファ
17
出力バッファ
18
音声出力装置
12
増幅
13
Low passフィルタ
14
D／A変換
15

【図３】


（A）
信号入力処理
s1
外部音声信号入力
s2
外部入力バッファに書き込み
s3
発話音声信号入力
s4
発話入力バッファに書き込み
（B）
和速変換処理
s10
外部入力バッファから信号読み出し
s11
波形整形回路のレベル判断
s12
発話あり
Y
N
s13
外部音声伸長処理
s14
出力バッファに書き込み
s15
０レベル信号を出力バッファに書き込み
（C）
信号出力処理
s20
出力バッファから信号読み出し
s21
発話入力バッファの同じタイミング信号読み出し
s22
加算合成
s23
D／Aに出力

【図５】



フロントページの続き

(72)発明者　谷高　幸司
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(72)発明者　大脇　浩
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

(56)参考文献　特開　平８－149599（ＪＰ，Ａ）
特開　平９－16193（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.⁷，ＤＢ名)
H04R　25/00
G10L　21/04

(19)日本国特許庁（ＪＰ）　(12)特　許　公　報（Ｂ２）　(11)特許番号
特許第3353701号
(P3353701)
(45)発行日　平成14年12月３日(2002.12.3)　(24)登録日　平成14年９月27日(2002.9.27)

(51)Int.Cl.[7]　識別記号　ＦＩ
Ｈ０４Ｒ　25/00　Ｈ０４Ｒ　25/00　Ｈ

請求項の数８(全 15 頁)

(21)出願番号　特願平10－128782
(22)出願日　平成10年５月12日(1998.5.12)
(65)公開番号　特開平11－331990
(43)公開日　平成11年11月30日(1999.11.30)
審査請求日　平成12年１月25日(2000.1.25)

(73)特許権者　000004075
ヤマハ株式会社
静岡県浜松市中沢町10番１号
(72)発明者　近藤　克文
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(72)発明者　林　好典
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(72)発明者　三木　晃
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(74)代理人　100084548
弁理士　小森　久夫

審査官　松澤　福三郎

最終頁に続く

(54)【発明の名称】　自己発話検出装置、音声入力装置および補聴器

(57)【特許請求の範囲】
【請求項１】　装用者の外耳道に挿入され、外界と外耳道との間を遮音する装着部と、
外界に向けて設けられた第１のマイクと、
外耳道内に向けて設けられた第２のマイクと、
第１のマイクの信号を所定時間遅延する遅延手段と、
遅延手段が遅延した信号を外耳道内に出力するレシーバと、
第１のマイクの信号波形と第２のマイクの信号波形との相関に基づいて装用者の自己発話を検出する信号処理部と、
を備えたことを特徴とする自己発話検出装置。
【請求項２】　前記信号処理部は、第１のマイクの信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部である請求項１に記載の自己発話検出装置。
【請求項３】　前記信号処理部は、遅延手段で遅延された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部である請求項１に記載の自己発話検出装置。
【請求項４】　テスト音信号を発生して前記遅延手段に入力する音源と、
該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする請求項３に記載の自己発話検出装

置。
【請求項５】　装用者の外耳道に挿入され、外界と外耳道との間を遮音する装着部と、
外界に向けて設けられた第１のマイクと、
外耳道内に向けて設けられた第２のマイクと、
第１のマイクが受信した信号を所定時間遅延する遅延手段と、
遅延手段が遅延した信号を外耳道内に出力するレシーバと、
遅延手段で遅延された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形を抽出し、この差分信号波形を外部出力する信号処理部と、
を備えたことを特徴とする音声入力装置。
【請求項６】　テスト音信号を発生して前記遅延手段に入力する音源と、
該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする請求項５に記載の音声入力装置。
【請求項７】　装用者の外耳道に挿入され、外界と外耳道との間を遮音する装着部と、
外界に向けて設けられた第１のマイクと、
外耳道内に向けて設けられた第２のマイクと、
第１のマイクが受信した信号を時間軸に伸長する話速変換手段と、
遅延手段が遅延した信号を外耳道内に出力するレシーバと、
話速変換手段で伸長された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出し、自己発話が検出されたとき、前記話速変換手段の動作を禁止する信号処理部と、
を備えたことを特徴とする補聴器。
【請求項８】　テスト音信号を発生して前記話速変換手段に入力する音源と、
該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする請求項７に記載の補聴器。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】この発明は、装用者が発した音声と外界から到来した音声とを識別する機能を備えた自己発話検出装置、音声入力装置および補聴器に関する。
【０００２】
【従来の技術】従来、装用者が自ら発声した音声である自己発話を外来の音声から分離識別して検出するための装置として骨伝導マイクを用いたものが実用化されている。
【０００３】骨伝導マイクは、装用者の外耳道に密着させた装着部を介して装用者の発声による頭蓋骨の振動を検出するものであり、空気振動である外来音には感度がなく、装用者の自己発話音のみをピックアップできるものである。
【０００４】
【発明が解決しようとする課題】しかし、骨伝導マイクは、自己発話音を感度よくピックアップするためには装着部と外耳道との密着性が要求される。このため、装着部の形状を各装用者の外耳道の形状に合わせる必要があり、本体が装着部に一体になったセル型のものの場合にはセル全体の交換が必要になるため、装用者毎のフィッティングが面倒であるという問題点があった。
【０００５】この発明は、より簡略な構成で自己発話を検出することができる自己発話検出装置、音声入力装置および補聴器を提供することを目的とする。
【０００６】
【課題を解決するための手段】この出願の請求項１の発明は、装用者の外耳道に挿入され外界と外耳道との間を遮音する装着部と、外界に向けて設けられた第１のマイクと、外耳道内に向けて設けられた第２のマイクと、第１のマイクの信号を所定時間遅延する遅延手段と、遅延手段が遅延した信号を外耳道内に出力するレシーバと、第１のマイクの信号波形と第２のマイクの信号波形との相関に基づいて装用者の自己発話を検出する信号処理部と、を備えたことを特徴とする。
【０００７】この出願の請求項２の発明は、請求項１の発明において、前記信号処理部は、第１のマイクの信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部であることを特徴とする。
【０００８】この出願の請求項３の発明は、請求項１の発明において、前記信号処理部は、遅延手段で遅延された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部であることを特徴とする。
【０００９】この出願の請求項４の発明は、請求項３の発明において、テスト音信号を発生して前記遅延手段に入力する音源と、該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする。

【００１０】この出願の請求項５の発明は、装用者の外耳道に挿入され外界と外耳道との間を遮音する装着部と、外界に向けて設けられた第１のマイクと、外耳道内に向けて設けられた第２のマイクと、第１のマイクが受信した信号を所定時間遅延する遅延手段と、遅延手段が遅延した信号を外耳道内に出力するレシーバと、遅延手段で遅延された第１のマイクの信号波形をレシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形を抽出し、この差分信号波形を外部出力する信号処理部と、を備えたことを特徴とする。

【００１１】この出願の請求項６の発明は、請求項５の発明において、テスト音信号を発生して前記遅延手段に入力する音源と、該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする。

【００１２】この出願の請求項７の発明は、装用者の外耳道に挿入され外界と外耳道との間を遮音する装着部と、外界に向けて設けられた第１のマイクと、外耳道内に向けて設けられた第２のマイクと、第１のマイクが受信した信号を時間軸に伸長する話速変換手段と、遅延手段が遅延した信号を外耳道内に出力するレシーバと、話速変換手段で伸長された第１のマイクの信号波形をレシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出し、自己発話が検出されたとき前記話速変換手段の動作を禁止する信号処理部と、を備えたことを特徴とする。

【００１３】この出願の請求項８の発明は、請求項７の発明において、テスト音信号を発生して前記遅延手段に入力する音源と、該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする。

【００１４】この発明では、第２のマイクは自己発話音以外にレシーバから出力された外来音声も受信する。一方、第１のマイクは外来音声を受信する。第２のマイクの信号波形と第１の信号波形の相関をとることによって外来音声の成分をキャンセルすることにより、自己発話音成分のみを分離抽出することができる。なお、第１のマイクが受信する外来音声には口から発声された自己発話音も含まれるが同様の処理でよい。

【００１５】信号処理部はＤＳＰなどで構成することができるため、複雑な信号処理をする場合でも構造的には簡略なものにすることができる。

【００１６】したがって、この発明によれば、外界と外耳道とを遮音するのみの骨伝導マイクよりも遙に簡略な装着部の構成で自己発話音のみを分離抽出することができる。

【００１７】

【発明の実施の形態】図面を参照してこの発明の実施形態について説明する。図１は、この発明が適用される自己発話検出装置の外観図である。この自己発話検出装置は補聴器や音声入力装置に用いられる。セル１の先端に外耳道４に挿入される装着部である耳栓２が設けられている。耳栓２は傘形の柔軟な樹脂で構成されており、外耳道壁３に密着して外耳道４と外界とを遮断し、相互間の音の伝搬を禁止する。

【００１８】セル１内には、外界から到来する音声を検出するためのマイク５、外耳道４に音響を出力するためのレシーバ６、外耳道４の内部の音声を検出するためのマイク７、および、レシーバ６，マイク７とを空間的に結ぶパイプ孔８が設けられている。また、マイク５，レシーバ６およびマイク７は信号処理部９に接続されている。信号処理部９はＤＳＰなどの電子回路で構成され、マイク５，マイク７から入力された音声信号を分析・処理し、処理された音声信号をレシーバ６から出力する。以下、信号処理部９を種々に構成した例とその動作について説明する。

【００１９】図２は信号処理部９を遅延装置１０、相関演算装置１１、発話判定装置１２で構成した自己発話検出装置の例を示している。図２（Ａ）において、マイク５が受信した外来音声は、遅延装置１０および相関演算装置１１に入力される。遅延装置１０は入力された音声信号を微小時間ｔだけ遅延させてレシーバ６に出力する。レシーバ６はこの音声信号を音声（空気振動）に変換して外耳道４内に放出する。マイク７は、レシーバ６が出力した音声および外耳道４から伝達された装用者の音声の両方を受信する。マイク７が受信した音声信号は相関演算装置１１に入力される。相関演算装置１１は、マイク５から入力された音声信号とマイク７から入力された音声信号の相関値を計算する。

【００２０】図２（Ｂ）、（Ｃ）を参照して外来音声が到来したときの信号波形と相関について説明する。マイク５が外来音声を受信した時刻を０ｓｅｃとする。マイク５の入力であるｐｏｉｎｔＡ波形は、同図（Ｂ）上段のような波形となる。遅延装置１０はこの音声信号をｔｓｅｃ遅延させてレシーバ６に出力する。レシーバ６に供給される波形であるｐｏｉｎｔＢ波形は、ｐｏｉｎｔＡの波形をｔｓｅｃ遅延させた同図（Ｂ）中段のような波形となる。レシーバ６はこの遅延された音声信号を外耳道４に出力する。外耳道４に出力された音声はマイク７にも受信される。マイク７の入力であるｐｏｉｎｔＣ波形は、ｐｏｉｎｔＢ波形とほぼ同様の同図（Ｂ）下段のような波形となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＣの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。同

図（Ｃ）において、０ｓｅｃの時点では外耳道４内には音声信号がなくｐｏｉｎｔＣの波形はフラットであるので相関値は０であるが、時刻ｔにｐｏｉｎｔＣに遅延された音声の波形が現れるのでこの時刻ｔの相関値が最大となる。

【００２１】なお、同図（Ｂ），（Ｃ）は入力波形のレベルをモデル化した図であるが、実際には図８のような波形を示す。同図（Ａ）がｐｏｉｎｔＡの波形であり、同図（Ｂ）がｐｏｉｎｔＣの波形である。そして同図（Ｃ）が相関値である。このようにアナログ波形の場合、相関値はパルス状にはならないが、やはりｔｓｅｃ後にピーク値をとるように推移する。

【００２２】一方、図２（Ｄ）、（Ｅ）を参照して自己発話が行われたときの信号波形と相関について説明する。装用者が発声すると、その音声信号は口から外界に放出されるとともに体内伝搬経路によって外耳道４内にも伝搬する。外界に放出された音声はマイク５に受信され、外耳道４に伝搬した音声はマイク７に受信される。双方とも伝搬距離が極めて短いため、ほぼ同時（０ｓｅｃ）に受信される。したがって、時刻０ｓｅｃにマイク５の入力であるｐｏｉｎｔＡには同図（Ｄ）上段に示すような波形が現れ、マイク７にも時刻０ｓｅｃにｐｏｉｎｔＡ波形と類似した波形が現れる。また、ｐｏｉｎｔＡ波形は遅延装置１０でｔｓｅｃ遅延されたのち（ｐｏｉｎｔＢ波形）、レシーバ６から外耳道４に出力されるので、マイク７はこの音声も受信し、マイク７は、外耳道内に発生した装用者自身の音声とレシーバ６から出力される遅延された音声の両方を受信し、受信した波形であるｐｏｉｎｔＣ波形は同図（Ｄ）下段のような合成波形となる。時刻０ｓｅｃではポイントＡとポイントＣの波形はほぼ同様であるため、相関演算装置１１は、同図（Ｅ）のように時刻０ｓｅｃに高い相関値を出力する。

【００２３】なお、同図（Ｄ），（Ｅ）は入力波形のレベルをモデル化した図であるが、実際には図９のような波形を示す。同図（Ａ）がｐｏｉｎｔＡの波形であり、同図（Ｂ）がｐｏｉｎｔＣの波形である。そして同図（Ｃ）が相関値である。このようにアナログ波形の場合、相関値はパルス状にはならないが、やはり時刻０でピーク値をとり、徐々に減衰している。

【００２４】以上説明した相関値の相違に基づき、発話判定装置１２は、０ｓｅｃに相関値のピークが出力されたとき、その相関する波形は自己発話によるものと判定し、ｔｓｅｃ後に相関値のピークが出力されたとき、その波形は外来音声によるものと判定する。

【００２５】図３は信号処理部９を遅延装置１０、相関演算装置１１、発話判定装置１２および差分処理装置１３で構成した自己発話検出装置の例を示す。図３（Ａ）において、マイク５が受信した外来音声は遅延装置１０および相関演算装置１１に入力される。遅延装置１０は入力された音声信号をｔｓｅｃだけ遅延させてレシーバ６に出力する。レシーバ６はこの音声信号を音声（空気振動）に変換して外耳道４内に放出する。遅延装置１０が遅延した信号（ｐｏｉｎｔＢ）は差分処理装置１３にも入力される。一方、マイク７は、外耳道４内に発生した音声を受信して差分処理装置１３に入力する。マイク７は、レシーバ６が出力した音声および外耳道４から伝達された装用者の音声の両方を受信する。差分処理装置１３はマイク７から入力された音声信号波形から遅延装置１０がレシーバ６に出力したものと同じ音声信号波形を減算して外来音声成分をキャンセルする。この外来音声成分がキャンセルされた音声信号波形を相関演算装置１１に入力する。相関演算装置１１は、マイク５から入力された音声信号と差分処理装置１３から入力された音声信号の相関値を計算する。

【００２６】図３（Ｂ）、（Ｃ）を参照して外来音声が到来したときの信号波形と相関について説明する。マイク５が外来音声を受信した時刻を０ｓｅｃとする。マイク５の入力であるｐｏｉｎｔＡは、同図（Ｂ）第１段のような波形となる。遅延装置１０はこの音声信号をｔｓｅｃ遅延させてレシーバ６に出力する。レシーバ６に供給されるｐｏｉｎｔＢ波形は、ｐｏｉｎｔＡ波形をｔｓｅｃ遅延させたものであるため同図（Ｂ）第２段のような波形となる。レシーバ６はこの遅延された音声信号を外耳道４に出力する。外耳道４に出力された音声はマイク７にも受信される。マイク７の入力信号波形であるｐｏｉｎｔＣ波形は、ｐｏｉｎｔＢ波形とほぼ同様の同図（Ｂ）第３段のような波形となる。差分処理装置１３はマイク７から入力された音声信号波形であるｐｏｉｎｔＣ波形と遅延装置１０から入力された音声信号波形であるｐｏｉｎｔＢ波形とを減算する。外来音声のみの場合、内部で発生した音声信号がないため、ｐｏｉｎｔＢ波形とｐｏｉｎｔＣ波形はほとんど一致し、差分処理装置１３の出力であるｐｏｉｎｔＤ波形は、同図（Ｂ）第４段に示すように殆ど０レベルのフラットな信号となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＤの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＤの波形は殆どフラットであるためｐｏｉｎｔＡの基準波形と殆ど相関性がなく、大きな相関値が出力されない。発話判定装置１２は大きな相関値が入力されないことにより、自己発話がないことを判断する。

【００２７】一方、図３（Ｄ）、（Ｅ）を参照して自己発話が行われたときの信号波形と相関について説明する。装用者が発声すると、その音声信号は口から外界に放出されるとともに体内伝搬経路によって外耳道４内にも伝搬する。外界に放出された音声はマイク５に受信され、外耳道４に伝搬した音声はマイク７に受信される。双方とも伝搬距離が極めて短いため、ほぼ同時（０ｓｅｃ）に受信される。したがって、マイク５には時刻０ｓ

ｅｃにｐｏｉｎｔＡの波形が現れ、マイク７にも時刻０ｓｅｃにｐｏｉｎｔＡと類似した波形が現れる。また、ｐｏｉｎｔＡの波形は遅延装置１０でｔｓｅｃ遅延されたのちレシーバ６から外耳道４に放出されるので、マイク７はこの音声も受信する。すなわち、マイク７は、外耳道内に発生した装用者自身の音声とレシーバ６から放出される遅延された音声の両方を受信し、その合成波形は、同図（Ｄ）第３段に示すようなｐｏｉｎｔＣの波形となる。差分処理装置１３は、マイク７から入力された音声信号波形であるｐｏｉｎｔＣ波形と遅延装置１０から入力された音声信号波形であるｐｏｉｎｔＢ波形とを減算する。ｐｏｉｎｔＣ波形は、内部伝達した波形とレシーバ６から出力された音声波形の合成信号であるため、これからｐｏｉｎｔＢ波形を減算すると、レシーバ６から出力された外来音声の波形成分がキャンセルされ、同図（Ｂ）第４段に示すようにｐｏｉｎｔＡと類似した波形となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＤの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＤの波形は殆どｐｏｉｎｔＡの波形と一致しているため、０ｓｅｃのタイミングに大きな相関値が出力される。発話判定装置１２は、大きな相関値が入力されたことにより、０ｓｅｃのタイミングに開始した自己発話があったことを判断する。このように、図３の例では、差分処理装置１３によって余分な信号を排除しているので相関値が改善され誤り判定の減少が期待できる。

【００２８】図２，図３の例は、信号経路が理想的で経路の周波数特性や位相特性の影響がほとんど無い場合について説明したが、信号経路の伝搬特性が音声信号の波形に影響を与える場合を考慮した例を図４に示す。

【００２９】図４の例では、信号処理部９を遅延装置１０、相関演算装置１１、発話判定装置１２、差分処理装置１３に加えて伝達関数補正装置１４で構成している。

【００３０】図４（Ａ）において、マイク５が受信した外来音声は遅延装置１０および相関演算装置１１に入力される。遅延装置１０は入力された音声信号をｔｓｅｃだけ遅延させてレシーバ６に出力する。レシーバ６はこの音声信号を音声（空気振動）に変換して外耳道４内に放出する。遅延装置１０が遅延したｐｏｉｎｔＢ波形は伝達関数補正装置１４にも入力される。伝達関数補正装置１４は入力された信号をフィルタ処理して差分処理装置１３に入力する。このフィルタは、この装置が外耳道４に挿入されているときの外耳道４のインパルス応答をシミュレートしたフィルタであり、このフィルタを通過した信号が外耳道４内を伝搬した信号と同じ波形になるように設計されている。すなわち、このフィルタを通過させることにより、信号波形に対して伝搬路の伝達関数が畳み込み演算される。

【００３１】一方、マイク７は、外耳道４内に発生した音声を受信して差分処理装置１３に入力する。マイク７は、レシーバ６が出力した音声および外耳道４から伝達された装用者の音声の両方を受信する。差分処理装置１３はマイク７から入力された音声信号波形から伝達関数補正装置１４から入力された音声信号波形を減算してこの成分をキャンセルする。このレシーバ６の出力分がキャンセルされた音声信号波形を相関演算装置１１に入力する。相関演算装置１１は、マイク５から入力された音声信号と差分処理装置１３から入力された音声信号の相関値を計算する。

【００３２】図４（Ｂ）を参照して外来音声が到来したときの信号波形と相関について説明する。マイク５が外来音声を受信した時刻を０ｓｅｃとする。マイク５の入力であるｐｏｉｎｔＡは、同図（Ｂ）第１段のような波形となる。遅延装置１０はこの音声信号をｔｓｅｃ遅延させてレシーバ６に出力する。レシーバ６にされる波形であるｐｏｉｎｔＢは、ｐｏｉｎｔＡの波形をｔｓｅｃ遅延させた同図（Ｂ）第２段のような波形となる。レシーバ６はこの遅延された音声信号を外耳道４に出力する。外耳道４に出力された音声はマイク７にも受信される。マイク７の入力であるｐｏｉｎｔＣは、ｐｏｉｎｔＢとタイミングはほぼ同じであるが、外耳道４を伝搬したことによって遅延成分が重畳されており同図（Ｂ）第３段のような波形となる。差分処理装置１３には、遅延装置１０から出力されるｐｏｉｎｔＢ波形を伝達関数補正装置１４で補正した波形が入力されるため、この波形をマイク７から入力された音声信号波形であるｐｏｉｎｔＣから減算する。ｐｏｉｎｔＢの波形を伝達関数補正装置１４で補正したものはほぼｐｏｉｎｔＣの波形と同じになっているため、これらの波形はほとんど一致し差分処理装置１３の出力であるｐｏｉｎｔＥ波形は、同図（Ｂ）第５段に示すように殆ど０レベルのフラットな信号となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＥの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＥの波形は殆どフラットであるためｐｏｉｎｔＡの基準波形と殆ど相関性がなく、大きな相関値が出力されない。発話判定装置１２は大きな相関値が入力されないことにより、自己発話がないことを判断する。

【００３３】一方、図４（Ｃ）を参照して自己発話が行われたときの信号波形と相関について説明する。装用者が発声すると、その音声信号は口から外界に放出されるとともに体内伝搬経路によって外耳道４内にも伝搬する。外界に放出された音声はマイク５に受信され、外耳道４に伝搬した音声はマイク７に受信される。双方とも伝搬距離が極めて短いため、ほぼ同時（０ｓｅｃ）に受信される。したがって、マイク５には時刻０ｓｅｃにｐｏｉｎｔＡの波形が現れ、マイク７にも時刻０ｓｅｃに

ｐｏｉｎｔＡ波形と類似した波形が現れる。また、ｐｏｉｎｔＡの波形は遅延装置１０でｔｓｅｃ遅延されたのちレシーバ６から外耳道４に放出されるので、マイク７はこの音声も受信する。レシーバ６から放出されマイク７が受信する信号は外耳道４を伝搬して入力される信号であるため、外耳道４による残響音成分が重畳されたものになっている。このためマイク７が受信する自己発話の直接音とレシーバ６から出力された音の合成波形は、同図（Ｃ）第３段に示すｐｏｉｎｔＣ波形となる。レシーバ６から放出され外耳道４を伝搬した音声信号は、上述したように外耳道４の伝搬特性によって残響分が重畳されているため、波形の立ち上がりと立ち下がりが緩やかになっている。伝達関数補正装置１４から出力された波形も外耳道４の伝達特性をフィルタ係数として有するフィルタを通過したものであるため、同じような残響成分が重畳されており、この信号をマイク７の入力信号から減算することにより、自己発話音のみを分離することができる（ｐｏｉｎｔＥ波形）。

【００３４】相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＤの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＤの波形は殆どｐｏｉｎｔＡの波形と一致しているため、０ｓｅｃのタイミングに大きな相関値が出力される。発話判定装置１２は、大きな相関値が入力されたことにより、０ｓｅｃのタイミングに開始した自己発話があったことを判断する。

【００３５】なお、図４（Ａ）において、差分処理装置１３の出力をそのまま取り出せば自己発話検出マイクとして機能させることができる。

【００３６】なお、図４（Ａ）の装置では、伝達関数補正装置１４を遅延装置１０と差分処理装置１３の間に挿入したが、図５のように、マイク７と差分処理装置１３の間におくこともできる。この場合、伝達関数が外耳道内に発生した自己発話音声にも畳み込まれてしまうので、差分処理後自己発話音の波形をもとに戻す伝達関数修正装置を差分処理装置１４と相関演算装置１１との間に設ける。

【００３７】なお、図３および図４の例において、相関演算装置１１は、同時に発生する２つの音声信号波形を比較するのみであるため、単純な比較演算装置でもよい。さらに、図４，図５の遅延装置１０は、ディジタル的に信号を遅延させる信号処理装置であってもよい。

【００３８】図６は、この発明を通信端末の音声入出力装置に適用した実施形態を示す図である。この装置を制御する制御装置３０は伝達関数演算モードと通常動作モードで装置を制御する。指示装置３３は装用者がこの装置のオン／オフを制御装置３０に指示するためのスイッチを含んでいる。

【００３９】図６において、図１および図４（Ａ）のマイク５－遅延装置１０－レシーバ６の系統として、マイク５（この装置も外観は図１のものと同様である：以下同じ）、増幅器３６、Ａ／Ｄ変換器３７、遅延装置３８、Ｄ／Ａ変換器３９、増幅器４０およびレシーバ６が設けられている。また、図１および図４（Ａ）のマイク７の系統として、マイク７－増幅器４１、Ａ／Ｄ変換器４２が設けられている。また、図４（Ａ）の伝達関数補正装置１４、差分処理装置１３、比較演算装置１１の機能は、制御装置３０および演算処理装置３５がソフト的に実現する。また、この装置は、外耳道４の伝達関数を割り出す伝達関数演算モードにテスト音信号を発生するための音源３４を内蔵している。この音源３４が形成したオーディオ信号は、テスト音信号としてマイク５の入力に変えてＡ／Ｄ変換器３７に入力される。

【００４０】指示装置３３によってオンが指示されると、制御部３０は最初に伝達関数モードを設定し、制御装置３０は音源３４をオンするとともに、演算処理装置３５に伝達関数演算動作を指示する。音源３４が形成するテスト音信号は、ホワイトノイズやピンクノイズでよいが、他の信号、例えばインパルスや周波数の定まった低周波信号の波形を形成するようにしてもよい。音源３４から出力されたテスト音信号はＡ／Ｄ変換器３７、遅延装置３８、Ｄ／Ａ変換器３９、増幅器４０、レシーバ６を経て外耳道４内に放出される。外耳道４内に放出されたテスト音は外耳道４に向けて設けられたマイク７で検出される。マイク７の出力は増幅器４１、Ａ／Ｄ変換器４２を経て演算処理装置３５に入力される。演算処理装置３５はＤＳＰなどで構成すればよい。一方、この演算処理装置３５には、Ａ／Ｄ変換器３７から音源３４が形成したテスト音信号をディジタル信号に変換したものも供給される。演算処理装置３５は、伝達関数演算動作として、ｐｏｉｎｔＡ信号（Ａ／Ｄ変換器３７から入力される信号）およびｐｏｉｎｔＢ信号（Ａ／Ｄ変換器４２から入力される信号）を比較し、これらの信号の相違に応じた補正係数を算出する。補正係数を求める手法は色々あるが例えば下式のようにｐｏｉｎｔＡ信号をＡ（ｋ）、ｐｏｉｎｔＢ信号をＢ（ｋ）とし、Ａ（ｋ）とＢ（ｋ）間の誤差Ｅ（ｋ）がゼロになるような関数Ｇ（ｋ）を求める。

【００４１】Ｇ（ｋ）・Ａ（ｋ）－Ｂ（ｋ）＝Ｅ（ｋ）

一度求めた関数Ｇ（ｋ）は装置を外耳道から取り外したりしなければ、比較的安定した状態にあるので、動作の最初に計算するのみでよい。演算が終了すると演算処理装置３５は待機状態になる。制御装置３０は演算処理装置３５が待機状態になったことを検出すると、音源３４を停止するとともに、演算処理装置３５に信号処理動作を指示して通常動作モードに切り換える。

【００４２】通常動作モードにおいて、マイク５から外来音声が入力されると、増幅器３６、Ａ／Ｄ変換器３７、遅延装置３８、Ｄ／Ａ変換器３９、増幅器４０を経

て、ｔｓｅｃ後にレシーバ６から外耳道４に出力される。外耳道４に出力された外来音声は外耳道４に向けて設けられたマイク７によって受信され、増幅器４１、Ａ／Ｄ変換器４２を経て演算処理装置３５に入力される。信号処理動作が設定されている演算処理装置３５は、Ａ／Ｄ変換器３７から入力されるｐｏｉｎｔＡ信号、すなわち直接の外来音声信号に上記で求めた伝達関数Ｇ（ｋ）を畳み込んだのち、ｐｏｉｎｔＢ信号、すなわち外耳道４を伝達した外来音声信号を含む信号と差演算処理を行う。この処理によってｐｏｉｎｔＢ信号から外来音声成分は消去され自己発話音声のみが抽出される。抽出された自己発話音声を通信装置の音声入力端子に接続すれば環境雑音が大きな場所でも雑音の少ない音声信号のみをピックアップして相手に送ることができる。

【００４３】なお、上記例では、参照信号としてＡ／Ｄ変換器３７の出力（ｐｏｉｎｔＡ信号）を演算処理装置３５に入力しているが、遅延装置３８の出力を入力するようにしてもよい。

【００４４】図７はこの発明を話速変換機能をもった補聴器に適用した実施形態を示す図である。この装置を制御する制御装置５０は伝達関数演算モードと通常動作モードで装置を制御する。指示装置５３は装用者がこの装置のオン／オフを制御装置５０に指示するためのスイッチを含んでいる。

【００４５】図７において、図１および図４（Ａ）のマイク５－遅延装置１０－レシーバ６の系統として、マイク５（この装置も外観は図１のものと同様である：以下同じ）、増幅器５６、Ａ／Ｄ変換器５７、話速変換・利得制御装置５８、Ｄ／Ａ変換器５９、増幅器６０およびレシーバ６が設けられている。また、図１および図４（Ａ）のマイク７の系統として、マイク７－増幅器６１、Ａ／Ｄ変換器６２が設けられている。また、図４（Ａ）の伝達関数補正装置１４、差分処理装置１３、比較演算装置１１の機能は、制御装置５０および演算処理装置５５がソフト的に実現する。また、この装置は、外耳道４の伝達関数を割り出す伝達関数演算モードにテスト音信号を発生するための音源５４を内蔵している。この音源５４が形成したオーディオ信号は、外部に向けて設けられたレシーバ９によって出力され、マイク５によって受信される。

【００４６】指示装置５３によってオンが指示されると、制御部５０は最初に伝達関数モードを設定し、制御装置５０は音源５４をオンするとともに、演算処理装置５５に伝達関数演算動作を指示する。音源５４が形成するテスト音信号は、ホワイトノイズやピンクノイズでよいが、他の信号、例えばインパルスや周波数の定まった低周波信号の波形を形成するようにしてもよい。音源５４から出力されたテスト音信号は、レシーバ９から出力され、マイク５で受信・電気信号に変換されたのち、Ａ／Ｄ変換器５７、話速変換・利得制御装置５８、Ｄ／Ａ変換器５９、増幅器６０、レシーバ６を経て外耳道４内に放出される。外耳道４内に放出されたテスト音は外耳道４に向けて設けられたマイク７で検出される。マイク７の出力は増幅器６１、Ａ／Ｄ変換器６２を経て演算処理装置５５に入力される。演算処理装置５５はＤＳＰなどで構成すればよい。一方、この演算処理装置５５には、話速変換・利得制御装置５８から音源５４が形成し、レシーバ９から出力されたテスト音信号を話速変換したものも供給される。演算処理装置３５は、伝達関数演算動作として、ｐｏｉｎｔＣ信号（話速変換・利得制御装置５８から入力される信号）およびｐｏｉｎｔＢ信号（Ａ／Ｄ変換器６２から入力される信号）を比較し、これらの信号の相違に応じた補正係数を算出する。補正係数を求める手法は色々あるが例えば下式のようにｐｏｉｎｔＡ信号をＡ（ｋ）、ｐｏｉｎｔＢ信号をＢ（ｋ）とし、Ａ（ｋ）とＢ（ｋ）間の誤差Ｅ（ｋ）がゼロになるような関数Ｇ（ｋ）を求める。

【００４７】Ｇ（ｋ）・Ａ（ｋ）－Ｂ（ｋ）＝Ｅ（ｋ）

一度求めた関数Ｇ（ｋ）は装置を外耳道から取り外したりしなければ、比較的安定した状態にあるので、動作の最初に計算するのみでよい。演算が終了すると演算処理装置５５は待機状態になる。制御装置５０は演算処理装置５５が待機状態になったことを検出すると、音源５４を停止するとともに、演算処理装置５５に信号処理動作を指示して通常動作モードに切り換える。

【００４８】通常動作モードにおいて、マイク５から外来音声が入力されると、増幅器５６、Ａ／Ｄ変換器５７、話速変換・利得制御装置５８、Ｄ／Ａ変換器５９、増幅器６０を経て、ｔｓｅｃ後にレシーバ６から外耳道４に出力される。外耳道４に出力された外来音声は外耳道４に向けて設けられたマイク７によって受信され、増幅器６１、Ａ／Ｄ変換器６２を経て演算処理装置５５に入力される。信号処理動作が設定されている演算処理装置５５は、話速変換・利得制御装置５８から入力されるｐｏｉｎｔＣ信号、すなわち直接の外来音声信号に上記で求めた伝達関数Ｇ（ｋ）を畳み込んだのち、ｐｏｉｎｔＢ信号、すなわち外耳道４を伝達した外来音声信号を含む信号と差演算処理を行う。この処理によってｐｏｉｎｔＢ信号から外来音声成分は消去され自己発話音声成分のみが抽出される。マイク５が外来音声のみを受信しているときは、この自己発話音声成分は０であるが、自己発話しているときには、そのレベルの成分が抽出される。制御装置５０は演算処理装置５５が自己発話音声成分を抽出したとき、話速変換・利得制御装置５８に対して話速変換を禁止し、利得を小さめに制御する。これにより、自己発話に対する話速変換が禁止され、自己発話が滑らかに行われるようになる。

【００４９】なお、上記例では、参照信号としてＡ／Ｄ変換器５７の出力（ｐｏｉｎｔＡ信号）を演算処理装置５５にも入力するようにすればよい。

【００５０】これにより、話速変換は外来音声に作用させたいが、自分の発声音声には作用させたくない。自分が発話するとき、人は自分の声を聞きながら発声を制御しているため、発声音声が遅れて聞こえるとうまく話せないという問題が起こる。そのため、自己発話に対しては、発話区間を検出して話速変換処理を適応しないようにする必要がある。また外来音声に比べ、自己発話音声は大きく聞こえるので適当な大きさになるよう利得制御が必要となる。

【００５１】このように、簡略な構成で自己発話と外来音とを分離することができるため、話速変換補聴器にこれを用いた場合に、装用者の外耳道の形状にかかわらず同じセルでこれを使用することができる。

【００５２】

【発明の効果】この発明によれば、外耳道に向けて設けられた第２のマイクの入力信号から外来音声信号をキャンセルして自己発話音のみを抽出できるようにしたことにより、構造が複雑で装用者毎のセッティングが微妙な骨伝導マイクを用いなくても環境ノイズなどの外来音声を除去した自己発話音声のみを抽出処理することができる。

【００５３】また、これを補聴器にもちいれば自己発話音声の区間のみ話速変換を禁止し、発話が容易な補聴器を構成することができる。

【図面の簡単な説明】

【図１】この発明の実施形態である自己発話検出装置の外観図

【図２】請求項１の発明の実施形態である自己発話検出装置のブロック図および信号波形を示す図

【図３】請求項２の発明の実施形態である自己発話検出装置のブロック図および信号波形を示す図

【図４】請求項３の発明の実施形態である自己発話検出装置のブロック図および信号波形を示す図

【図５】道自己発話検出装置の他の実施形態を示す図

【図６】請求項４の発明の実施形態である音声入力装置のブロック図号波形を示す図

【図７】この発明の自己発話検出装置を補聴器に応用した例を示す図

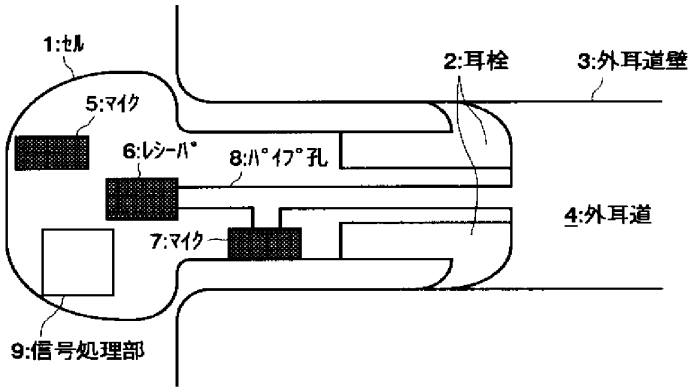
【図８】自己発話検出装置の各部における信号波形の例を示す図

【図９】自己発話検出装置の各部における信号波形の例を示す図

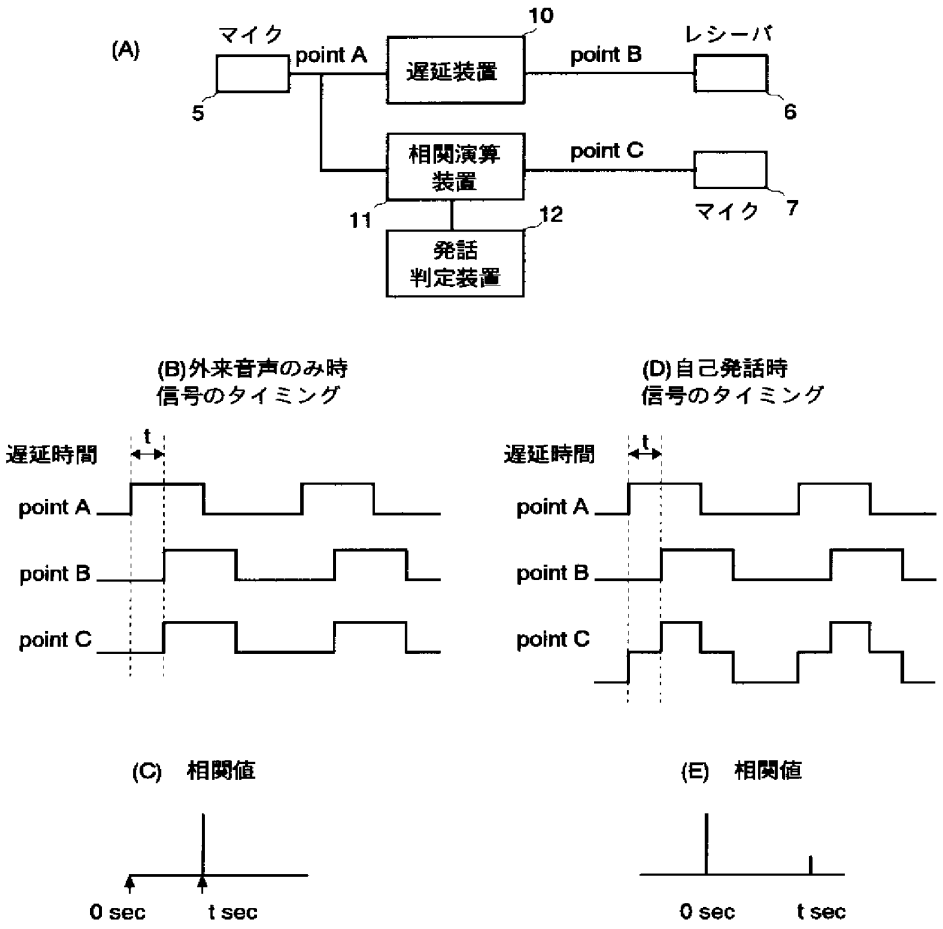
【符号の説明】

１…セル、２…耳栓、３…外耳道壁、４…外耳道、５…（外界に向けられた）マイク、６…レシーバ、７…（外耳道に向けられた）マイク、８…パイプ孔、９…信号処理部、１０…遅延装置、１１…相関演算装置、１２…発話判定装置、１３…差分処理装置、１４…伝達関数補正装置、

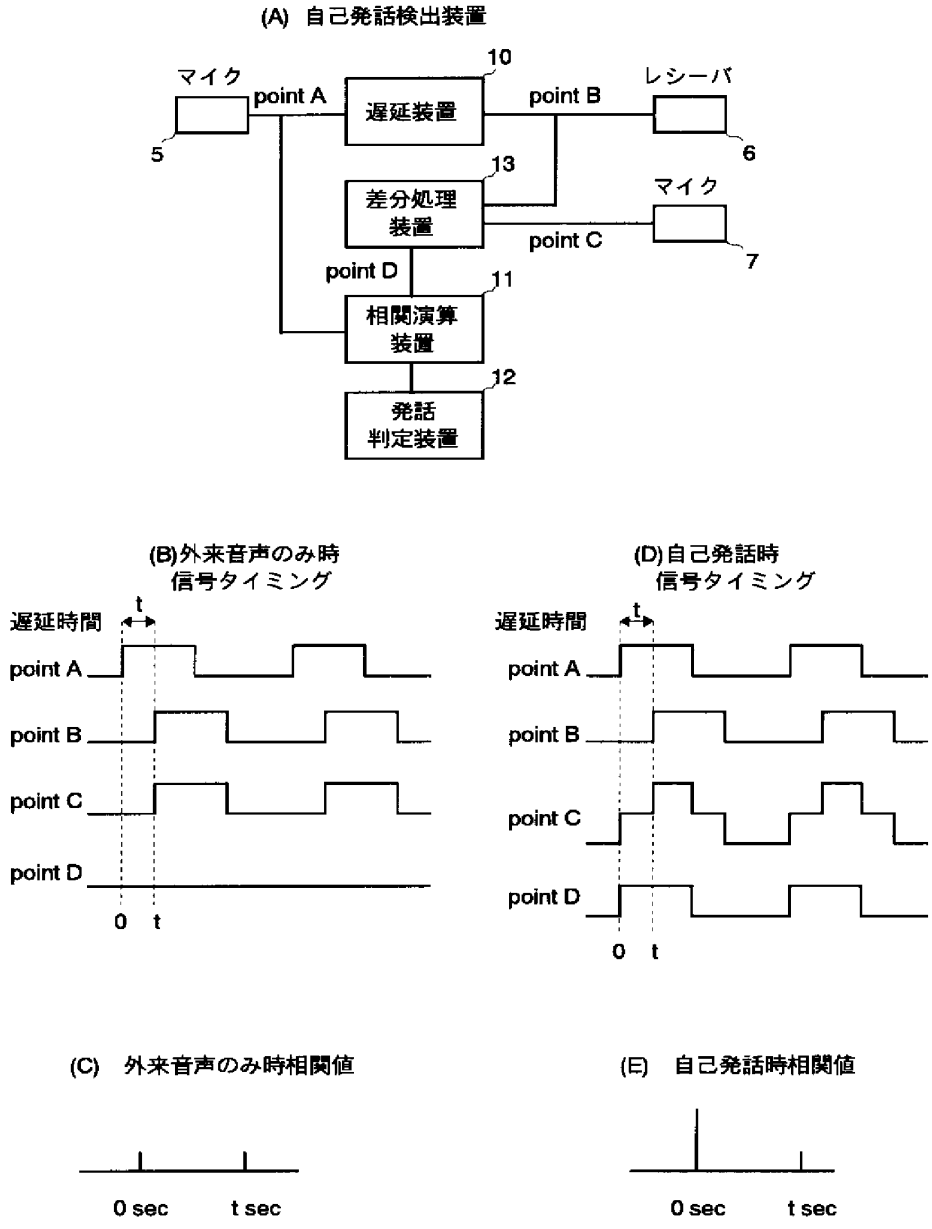
【図１】



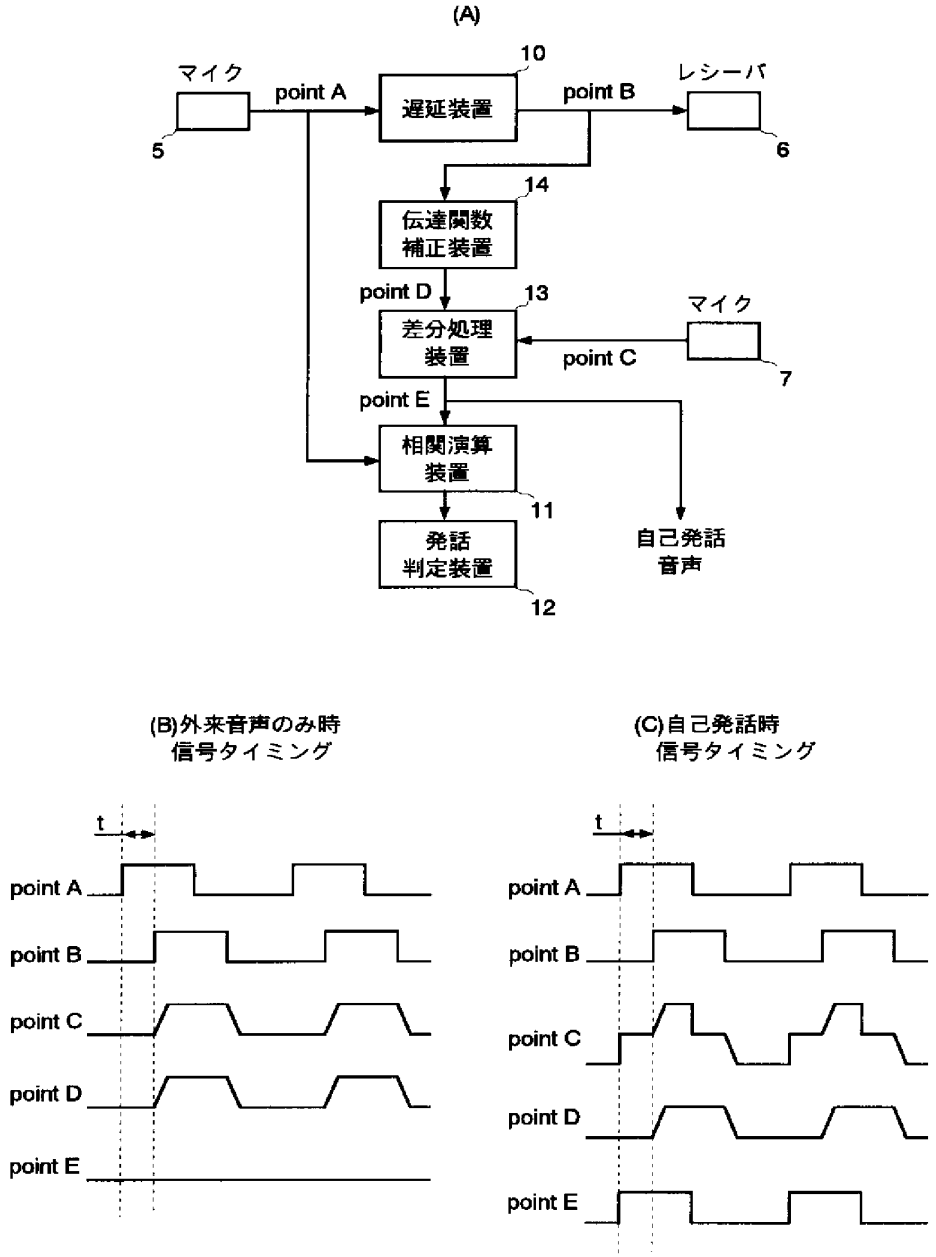
【図２】


(A)
マイク
point A
10
遅延装置
point B
レシーバ
5
6
相関演算
装置
point C
マイク
7
11
12
発話
判定装置
(B)外来音声のみ時
信号のタイミング
遅延時間
t
point A
point B
point C
(D)自己発話時
信号のタイミング
遅延時間
t
point A
point B
point C
(C)　相関値
0 sec
t sec
(E)　相関値
0 sec
t sec

【図３】


(A) 自己発話検出装置
マイク
point A
遅延装置
10
point B
レシーバ
5
6
13
差分処理装置
マイク
point C
7
point D
11
相関演算装置
12
発話判定装置
(B)外来音声のみ時信号タイミング
遅延時間
t
point A
point B
point C
point D
0
t
(D)自己発話時信号タイミング
遅延時間
t
point A
point B
point C
point D
0
t
(C) 外来音声のみ時相関値
0 sec
t sec
(E) 自己発話時相関値
0 sec
t sec

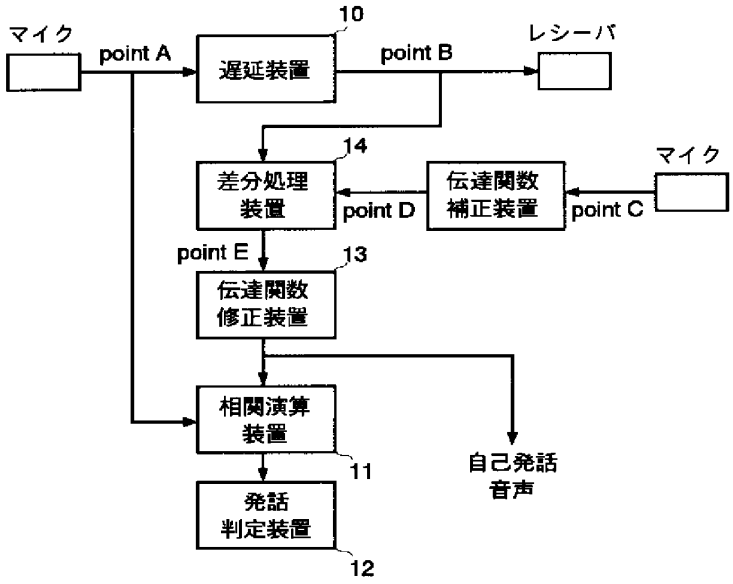
【図４】


(A)
マイク
point A
10
遅延装置
point B
レシーバ
5
6
14
伝達関数
補正装置
point D
13
差分処理
装置
マイク
point C
7
point E
相関演算
装置
11
発話
判定装置
12
自己発話
音声
(B)外来音声のみ時
信号タイミング
t
point A
point B
point C
point D
point E
(C)自己発話時
信号タイミング
t
point A
point B
point C
point D
point E

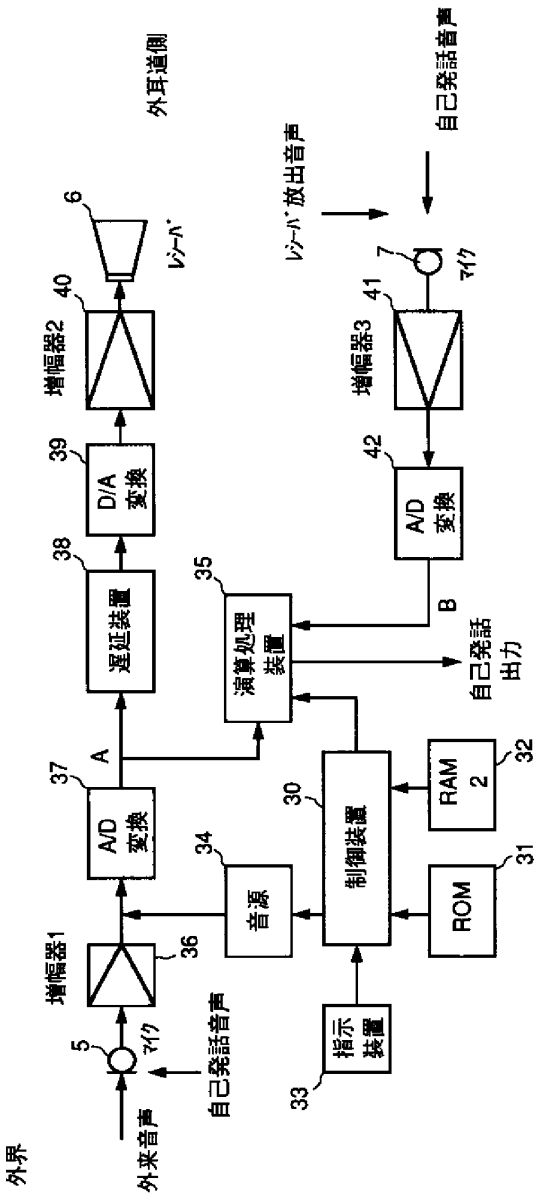
（１２）　　自己発話検出装置、音声入力装置および補聴器

【図５】

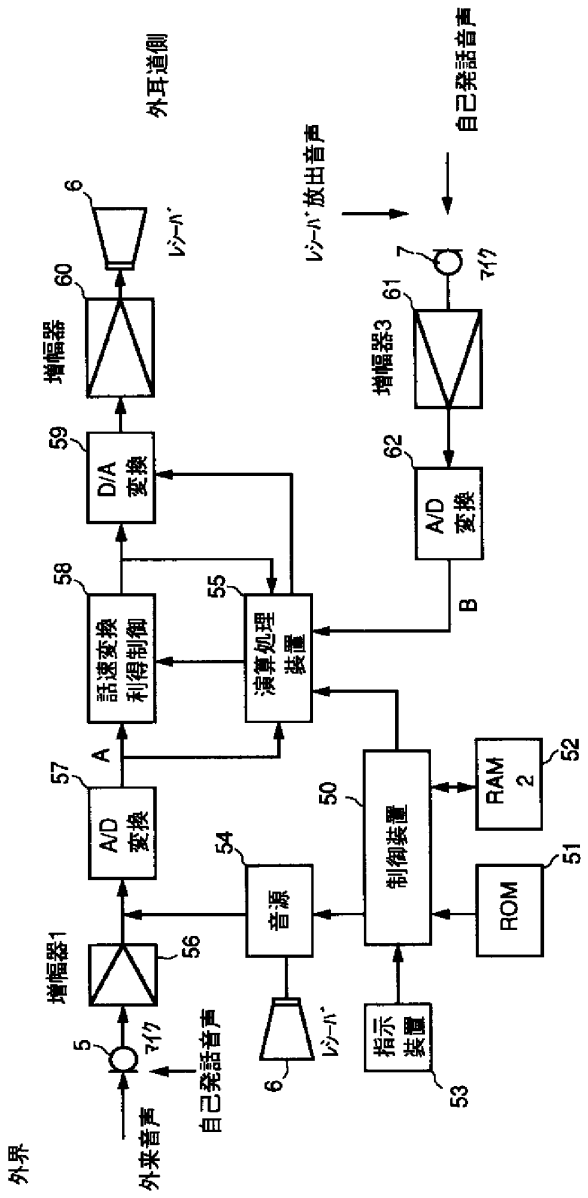
】



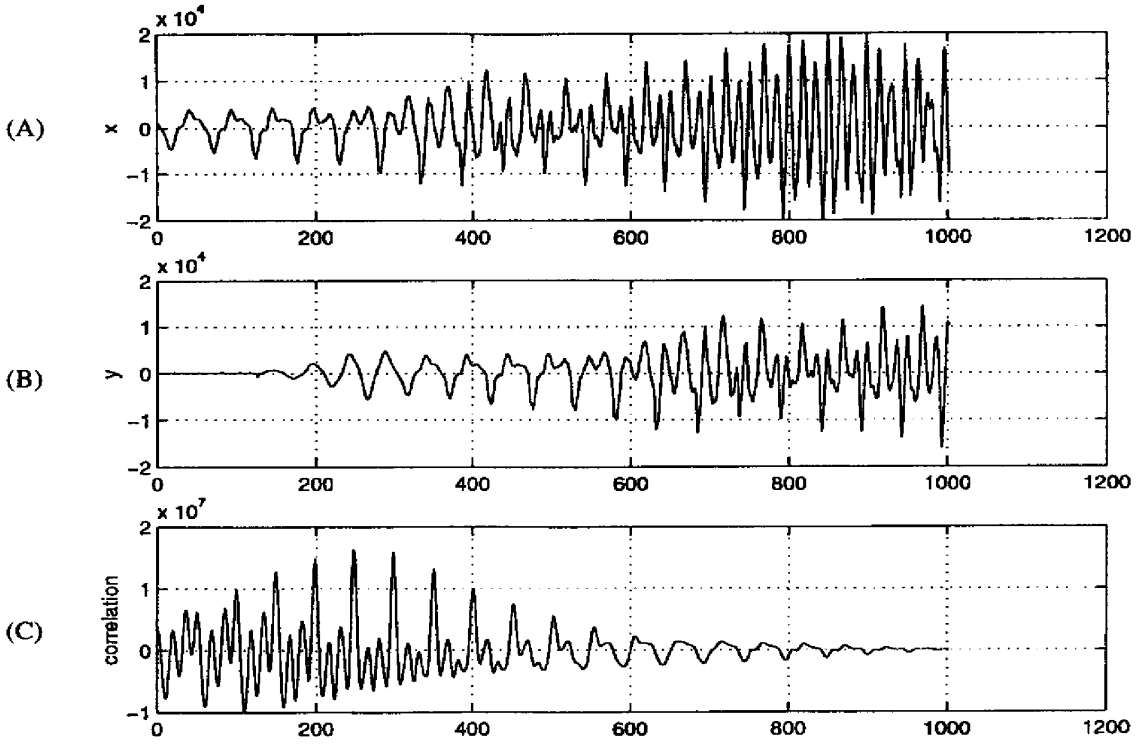
【図６】


外界
外来音声
5
マイク
自己発話音声
増幅器1
36
A/D変換
37
A
遅延装置
38
D/A変換
39
増幅器2
40
6
レシーバ
外耳道側
レシーバ放出音声
自己発話音声
7
マイク
増幅器3
41
A/D変換
42
B
演算処理装置
35
自己発話出力
音源
34
制御装置
30
ROM
31
RAM 2
32
指示装置
33

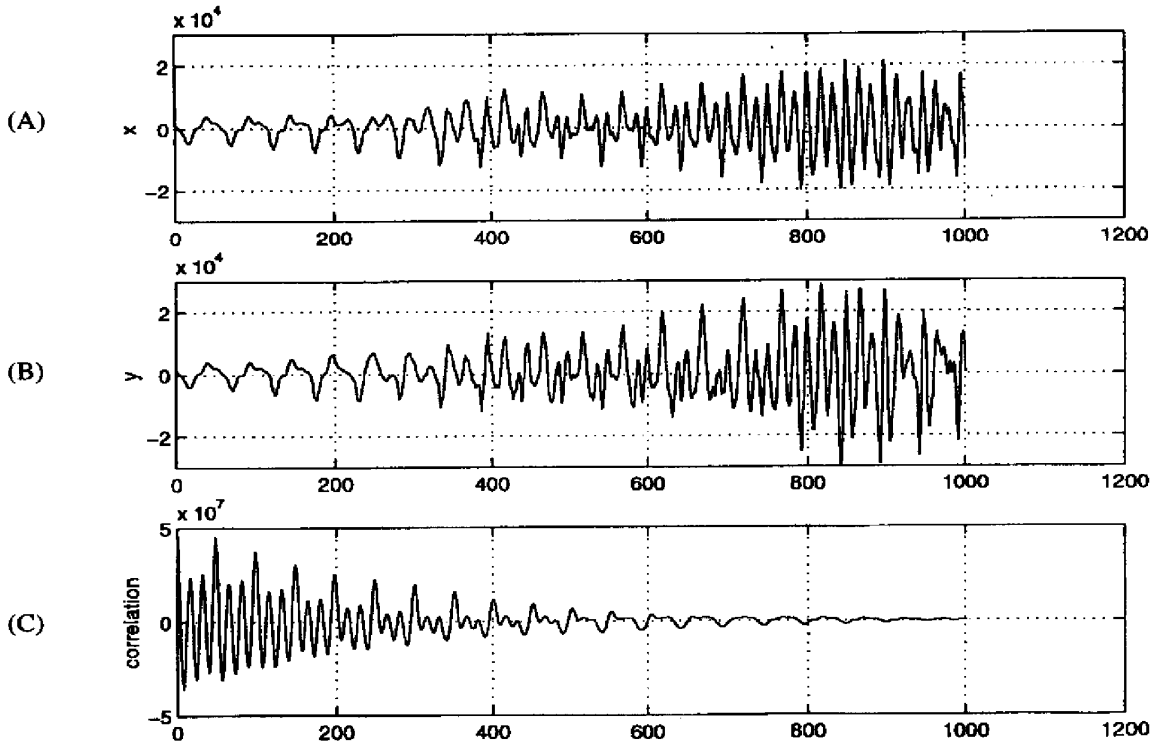
【図７】


外界
外来音声
5
マイク
自己発話音声
増幅器1
56
A/D変換
57
A
話速変換利得制御
58
D/A変換
59
増幅器
60
6
レシーバ
外耳道側
レシーバ放出音声
自己発話音声
7
マイク
増幅器3
61
A/D変換
62
B
演算処理装置
55
音源
54
制御装置
50
ROM
51
RAM 2
52
指示装置
53

【図８】


(A)
x 10^4
x
2
1
0
−1
−2
0
200
400
600
800
1000
1200
(B)
x 10^4
y
2
1
0
−1
−2
0
200
400
600
800
1000
1200
(C)
x 10^7
correlation
2
1
0
−1
0
200
400
600
800
1000
1200

【図９】



フロントページの続き

(72)発明者　谷高　幸司
静岡県浜松市中沢町10番１号　ヤマハ株式会社内
(72)発明者　大脇　浩
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

(56)参考文献　特開　平５－244696（ＪＰ，Ａ）
特開　平１－93298（ＪＰ，Ａ）
特表　平８－502396（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.⁷，ＤＢ名)
H04R　25/00
H04R　3/00

(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11－331990

(43)公開日　平成11年(1999)11月30日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｈ０４Ｒ　25/00		Ｈ０４Ｒ　25/00	Ｈ

審査請求　未請求　請求項の数８　ＯＬ　（全 15 頁）

(21)出願番号　特願平10－128782

(22)出願日　平成10年(1998)５月12日

(71)出願人　000004075
ヤマハ株式会社
静岡県浜松市中沢町10番１号

(72)発明者　近藤　克文
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

(72)発明者　林　好典
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

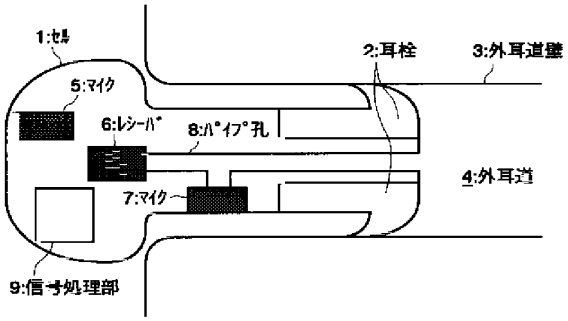
(72)発明者　三木　晃
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

(74)代理人　弁理士　小森　久夫

最終頁に続く

(54)【発明の名称】　自己発話検出装置、音声入力装置および補聴器

(57)【要約】
【課題】簡略な構成で自己発話を検出する。
【解決手段】　外界と外耳道とを遮断し、外界で受信した信号を一定時間遅延して外耳道のレシーバ６から出力する。外耳道にもマイク７が設けられており、レシーバ６が出力した信号と、装用者が自己発話し内部伝搬した音声信号が受信される。マイク７が受信した信号のうちマイク５が受信した信号成分を差し引くことにより、外来の音声信号成分がキャンセルされ、自己発話音声成分のみを検出・抽出することができる。



【特許請求の範囲】
【請求項１】　装用者の外耳道に挿入され、外界と外耳道との間を遮音する装着部と、
外界に向けて設けられた第１のマイクと、
外耳道内に向けて設けられた第２のマイクと、
第１のマイクの信号を所定時間遅延する遅延手段と、
遅延手段が遅延した信号を外耳道内に出力するレシーバと、
第１のマイクの信号波形と第２のマイクの信号波形との相関に基づいて装用者の自己発話を検出する信号処理部と、
を備えたことを特徴とする自己発話検出装置。
【請求項２】　前記信号処理部は、第１のマイクの信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部である請求項１に記載の自己発話検出装置。
【請求項３】　前記信号処理部は、遅延手段で遅延された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部である請求項１に記載の自己発話検出装置。
【請求項４】　テスト音信号を発生して前記遅延手段に入力する音源と、
該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする請求項３に記載の自己発話検出装置。
【請求項５】　装用者の外耳道に挿入され、外界と外耳道との間を遮音する装着部と、
外界に向けて設けられた第１のマイクと、
外耳道内に向けて設けられた第２のマイクと、
第１のマイクが受信した信号を所定時間遅延する遅延手段と、
遅延手段が遅延した信号を外耳道内に出力するレシーバと、
遅延手段で遅延された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形を抽出し、この差分信号波形を外部出力する信号処理部と、
を備えたことを特徴とする音声入力装置。
【請求項６】　テスト音信号を発生して前記遅延手段に入力する音源と、
該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする請求項５に記載の音声入力装置。
【請求項７】　装用者の外耳道に挿入され、外界と外耳道との間を遮音する装着部と、
外界に向けて設けられた第１のマイクと、
外耳道内に向けて設けられた第２のマイクと、
第１のマイクが受信した信号を時間軸に伸長する話速変換手段と、
遅延手段が遅延した信号を外耳道内に出力するレシーバと、
話速変換手段で伸長された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出し、自己発話が検出されたとき、前記話速変換手段の動作を禁止する信号処理部と、
を備えたことを特徴とする補聴器。
【請求項８】　テスト音信号を発生して前記話速変換手段に入力する音源と、
該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする請求項７に記載の補聴器。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】この発明は、装用者が発した音声と外界から到来した音声とを識別する機能を備えた自己発話検出装置、音声入力装置および補聴器に関する。
【０００２】
【従来の技術】従来、装用者が自ら発声した音声である自己発話を外来の音声から分離識別して検出するための装置として骨伝導マイクを用いたものが実用化されている。
【０００３】骨伝導マイクは、装用者の外耳道に密着させた装着部を介して装用者の発声による頭蓋骨の振動を検出するものであり、空気振動である外来音には感度がなく、装用者の自己発話音のみをピックアップできるものである。
【０００４】
【発明が解決しようとする課題】しかし、骨伝導マイクは、自己発話音を感度よくピックアップするためには装着部と外耳道との密着性が要求される。このため、装着部の形状を各装用者の外耳道の形状に合わせる必要があり、本体が装着部に一体になったセル型のものの場合にはセル全体の交換が必要になるため、装用者毎のフィッティングが面倒であるという問題点があった。
【０００５】この発明は、より簡略な構成で自己発話を検出することができる自己発話検出装置、音声入力装置および補聴器を提供することを目的とする。

【０００６】
【課題を解決するための手段】この出願の請求項１の発明は、装用者の外耳道に挿入され外界と外耳道との間を遮音する装着部と、外界に向けて設けられた第１のマイクと、外耳道内に向けて設けられた第２のマイクと、第１のマイクの信号を所定時間遅延する遅延手段と、遅延手段が遅延した信号を外耳道内に出力するレシーバと、第１のマイクの信号波形と第２のマイクの信号波形との相関に基づいて装用者の自己発話を検出する信号処理部と、を備えたことを特徴とする。
【０００７】この出願の請求項２の発明は、請求項１の発明において、前記信号処理部は、第１のマイクの信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部であることを特徴とする。
【０００８】この出願の請求項３の発明は、請求項１の発明において、前記信号処理部は、遅延手段で遅延された第１のマイクの信号波形を、レシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分を計算し、この差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出する処理部であることを特徴とする。
【０００９】この出願の請求項４の発明は、請求項３の発明において、テスト音信号を発生して前記遅延手段に入力する音源と、該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする。
【００１０】この出願の請求項５の発明は、装用者の外耳道に挿入され外界と外耳道との間を遮音する装着部と、外界に向けて設けられた第１のマイクと、外耳道内に向けて設けられた第２のマイクと、第１のマイクが受信した信号を所定時間遅延する遅延手段と、遅延手段が遅延した信号を外耳道内に出力するレシーバと、遅延手段で遅延された第１のマイクの信号波形をレシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形を抽出し、この差分信号波形を外部出力する信号処理部と、を備えたことを特徴とする。
【００１１】この出願の請求項６の発明は、請求項５の発明において、テスト音信号を発生して前記遅延手段に入力する音源と、該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする。
【００１２】この出願の請求項７の発明は、装用者の外耳道に挿入され外界と外耳道との間を遮音する装着部と、外界に向けて設けられた第１のマイクと、外耳道内に向けて設けられた第２のマイクと、第１のマイクが受信した信号を時間軸に伸長する話速変換手段と、遅延手段が遅延した信号を外耳道内に出力するレシーバと、話速変換手段で伸長された第１のマイクの信号波形をレシーバから第２のマイクに至る外耳道を含む伝搬路の伝達関数で畳み込み演算し、この畳込信号波形と第２のマイクの信号波形の差分信号波形と第１のマイクの信号波形との相関に基づいて装用者の自己発話を検出し、自己発話が検出されたとき前記話速変換手段の動作を禁止する信号処理部と、を備えたことを特徴とする。
【００１３】この出願の請求項８の発明は、請求項７の発明において、テスト音信号を発生して前記遅延手段に入力する音源と、該テスト音信号の直接波形およびレシーバから出力され第２のマイクによって受信されたテスト音信号の伝搬波形に基づいて前記伝搬路の伝達関数を割り出す手段を備えたことを特徴とする。
【００１４】この発明では、第２のマイクは自己発話音以外にレシーバから出力された外来音声も受信する。一方、第１のマイクは外来音声を受信する。第２のマイクの信号波形と第１の信号波形の相関をとることによって外来音声の成分をキャンセルすることにより、自己発話音成分のみを分離抽出することができる。なお、第１のマイクが受信する外来音声には口から発声された自己発話音も含まれるが同様の処理でよい。
【００１５】信号処理部はＤＳＰなどで構成することができるため、複雑な信号処理をする場合でも構造的には簡略なものにすることができる。
【００１６】したがって、この発明によれば、外界と外耳道とを遮音するのみの骨伝導マイクよりも遙に簡略な装着部の構成で自己発話音のみを分離抽出することができる。
【００１７】
【発明の実施の形態】図面を参照してこの発明の実施形態について説明する。図１は、この発明が適用される自己発話検出装置の外観図である。この自己発話検出装置は補聴器や音声入力装置に用いられる。セル１の先端に外耳道４に挿入される装着部である耳栓２が設けられている。耳栓２は傘形の柔軟な樹脂で構成されており、外耳道壁３に密着して外耳道４と外界とを遮断し、相互間の音の伝搬を禁止する。
【００１８】セル１内には、外界から到来する音声を検出するためのマイク５、外耳道４に音響を出力するためのレシーバ６、外耳道４の内部の音声を検出するためのマイク７、および、レシーバ６，マイク７とを空間的に結ぶパイプ孔８が設けられている。また、マイク５，レシーバ６およびマイク７は信号処理部９に接続されている。信号処理部９はＤＳＰなどの電子回路で構成され、マイク５，マイク７から入力された音声信号を分析・処理し、処理された音声信号をレシーバ６から出力する。以下、信号処理部９を種々に構成した例とその動作につ

いて説明する。

【００１９】図２は信号処理部９を遅延装置１０、相関演算装置１１、発話判定装置１２で構成した自己発話検出装置の例を示している。図２（Ａ）において、マイク５が受信した外来音声は、遅延装置１０および相関演算装置１１に入力される。遅延装置１０は入力された音声信号を微小時間ｔだけ遅延させてレシーバ６に出力する。レシーバ６はこの音声信号を音声（空気振動）に変換して外耳道４内に放出する。マイク７は、レシーバ６が出力した音声および外耳道４から伝達された装用者の音声の両方を受信する。マイク７が受信した音声信号は相関演算装置１１に入力される。相関演算装置１１は、マイク５から入力された音声信号とマイク７から入力された音声信号の相関値を計算する。

【００２０】図２（Ｂ）、（Ｃ）を参照して外来音声が到来したときの信号波形と相関について説明する。マイク５が外来音声を受信した時刻を０ｓｅｃとする。マイク５の入力であるｐｏｉｎｔＡ波形は、同図（Ｂ）上段のような波形となる。遅延装置１０はこの音声信号をｔｓｅｃ遅延させてレシーバ６に出力する。レシーバ６に供給される波形であるｐｏｉｎｔＢ波形は、ｐｏｉｎｔＡの波形をｔｓｅｃ遅延させた同図（Ｂ）中段のような波形となる。レシーバ６はこの遅延された音声信号を外耳道４に出力する。外耳道４に出力された音声はマイク７にも受信される。マイク７の入力であるｐｏｉｎｔＣ波形は、ｐｏｉｎｔＢ波形とほぼ同様の同図（Ｂ）下段のような波形となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＣの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。同図（Ｃ）において、０ｓｅｃの時点では外耳道４内には音声信号がなくｐｏｉｎｔＣの波形はフラットであるので相関値は０であるが、時刻ｔにｐｏｉｎｔＣに遅延された音声の波形が現れるのでこの時刻ｔの相関値が最大となる。

【００２１】なお、同図（Ｂ），（Ｃ）は入力波形のレベルをモデル化した図であるが、実際には図８のような波形を示す。同図（Ａ）がｐｏｉｎｔＡの波形であり、同図（Ｂ）がｐｏｉｎｔＣの波形である。そして同図（Ｃ）が相関値である。このようにアナログ波形の場合、相関値はパルス状にはならないが、やはりｔｓｅｃ後にピーク値をとるように推移する。

【００２２】一方、図２（Ｄ）、（Ｅ）を参照して自己発話が行われたときの信号波形と相関について説明する。装用者が発声すると、その音声信号は口から外界に放出されるとともに体内伝搬経路によって外耳道４内にも伝搬する。外界に放出された音声はマイク５に受信され、外耳道４に伝搬した音声はマイク７に受信される。双方とも伝搬距離が極めて短いため、ほぼ同時（０ｓｅｃ）に受信される。したがって、時刻０ｓｅｃにマイク５の入力であるｐｏｉｎｔＡには同図（Ｄ）上段に示すような波形が現れ、マイク７にも時刻０ｓｅｃにｐｏｉｎｔＡ波形と類似した波形が現れる。また、ｐｏｉｎｔＡ波形は遅延装置１０でｔｓｅｃ遅延されたのち（ｐｏｉｎｔＢ波形）、レシーバ６から外耳道４に出力されるので、マイク７はこの音声も受信し、マイク７は、外耳道内に発生した装用者自身の音声とレシーバ６から出力される遅延された音声の両方を受信し、受信した波形であるｐｏｉｎｔＣ波形は同図（Ｄ）下段のような合成波形となる。時刻０ｓｅｃではポイントＡとポイントＣの波形はほぼ同様であるため、相関演算装置１１は、同図（Ｅ）のように時刻０ｓｅｃに高い相関値を出力する。

【００２３】なお、同図（Ｄ），（Ｅ）は入力波形のレベルをモデル化した図であるが、実際には図９のような波形を示す。同図（Ａ）がｐｏｉｎｔＡの波形であり、同図（Ｂ）がｐｏｉｎｔＣの波形である。そして同図（Ｃ）が相関値である。このようにアナログ波形の場合、相関値はパルス状にはならないが、やはり時刻０でピーク値をとり、徐々に減衰している。

【００２４】以上説明した相関値の相違に基づき、発話判定装置１２は、０ｓｅｃに相関値のピークが出力されたとき、その相関する波形は自己発話によるものと判定し、ｔｓｅｃ後に相関値のピークが出力されたとき、その波形は外来音声によるものと判定する。

【００２５】図３は信号処理部９を遅延装置１０、相関演算装置１１、発話判定装置１２および差分処理装置１３で構成した自己発話検出装置の例を示す。図３（Ａ）において、マイク５が受信した外来音声は遅延装置１０および相関演算装置１１に入力される。遅延装置１０は入力された音声信号をｔｓｅｃだけ遅延させてレシーバ６に出力する。レシーバ６はこの音声信号を音声（空気振動）に変換して外耳道４内に放出する。遅延装置１０が遅延した信号（ｐｏｉｎｔＢ）は差分処理装置１３にも入力される。一方、マイク７は、外耳道４内に発生した音声を受信して差分処理装置１３に入力する。マイク７は、レシーバ６が出力した音声および外耳道４から伝達された装用者の音声の両方を受信する。差分処理装置１３はマイク７から入力された音声信号波形から遅延装置１０がレシーバ６に出力したものと同じ音声信号波形を減算して外来音声成分をキャンセルする。この外来音声成分がキャンセルされた音声信号波形を相関演算装置１１に入力する。相関演算装置１１は、マイク５から入力された音声信号と差分処理装置１３から入力された音声信号の相関値を計算する。

【００２６】図３（Ｂ）、（Ｃ）を参照して外来音声が到来したときの信号波形と相関について説明する。マイク５が外来音声を受信した時刻を０ｓｅｃとする。マイク５の入力であるｐｏｉｎｔＡは、同図（Ｂ）第１段のような波形となる。遅延装置１０はこの音声信号をｔｓｅｃ遅延させてレシーバ６に出力する。レシーバ６に供

給されるｐｏｉｎｔＢ波形は、ｐｏｉｎｔＡ波形をｔｓｅｃ遅延させたものであるため同図（Ｂ）第２段のような波形となる。レシーバ６はこの遅延された音声信号を外耳道４に出力する。外耳道４に出力された音声はマイク７にも受信される。マイク７の入力信号波形であるｐｏｉｎｔＣ波形は、ｐｏｉｎｔＢ波形とほぼ同様の同図（Ｂ）第３段のような波形となる。差分処理装置１３はマイク７から入力された音声信号波形であるｐｏｉｎｔＣ波形と遅延装置１０から入力された音声信号波形であるｐｏｉｎｔＢ波形とを減算する。外来音声のみの場合、内部で発生した音声信号がないため、ｐｏｉｎｔＢ波形とｐｏｉｎｔＣ波形はほとんど一致し、差分処理装置１３の出力であるｐｏｉｎｔＤ波形は、同図（Ｂ）第４段に示すように殆どＯレベルのフラットな信号となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＤの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＤの波形は殆どフラットであるためｐｏｉｎｔＡの基準波形と殆ど相関性がなく、大きな相関値が出力されない。発話判定装置１２は大きな相関値が入力されないことにより、自己発話がないことを判断する。

【００２７】一方、図３（Ｄ）、（Ｅ）を参照して自己発話が行われたときの信号波形と相関について説明する。装用者が発声すると、その音声信号は口から外界に放出されるとともに体内伝搬経路によって外耳道４内にも伝搬する。外界に放出された音声はマイク５に受信され、外耳道４に伝搬した音声はマイク７に受信される。双方とも伝搬距離が極めて短いため、ほぼ同時（０ｓｅｃ）に受信される。したがって、マイク５には時刻０ｓｅｃにｐｏｉｎｔＡの波形が現れ、マイク７にも時刻０ｓｅｃにｐｏｉｎｔＡと類似した波形が現れる。また、ｐｏｉｎｔＡの波形は遅延装置１０でｔｓｅｃ遅延されたのちレシーバ６から外耳道４に放出されるので、マイク７はこの音声も受信する。すなわち、マイク７は、外耳道内に発生した装用者自身の音声とレシーバ６から放出される遅延された音声の両方を受信し、その合成波形は、同図（Ｄ）第３段に示すようなｐｏｉｎｔＣの波形となる。差分処理装置１３は、マイク７から入力された音声信号波形であるｐｏｉｎｔＣ波形と遅延装置１０から入力された音声信号波形であるｐｏｉｎｔＢ波形とを減算する。ｐｏｉｎｔＣ波形は、内部伝達した波形とレシーバ６から出力された音声波形の合成信号であるため、これからｐｏｉｎｔＢ波形を減算すると、レシーバ６から出力された外来音声の波形成分がキャンセルされ、同図（Ｂ）第４段に示すようにｐｏｉｎｔＡと類似した波形となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＤの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＤの波形は殆どｐｏｉｎｔＡの波形と一致しているため、０ｓｅｃのタイミングに大きな相関値が出力される。発話判定装置１２は、大きな相関値が入力されたことにより、０ｓｅｃのタイミングに開始した自己発話があったことを判断する。このように、図３の例では、差分処理装置１３によって余分な信号を排除しているので相関値が改善され誤り判定の減少が期待できる。

【００２８】図２，図３の例は、信号経路が理想的で経路の周波数特性や位相特性の影響がほとんど無い場合について説明したが、信号経路の伝搬特性が音声信号の波形に影響を与える場合を考慮した例を図４に示す。

【００２９】図４の例では、信号処理部９を遅延装置１０、相関演算装置１１、発話判定装置１２、差分処理装置１３に加えて伝達関数補正装置１４で構成している。

【００３０】図４（Ａ）において、マイク５が受信した外来音声は遅延装置１０および相関演算装置１１に入力される。遅延装置１０は入力された音声信号をｔｓｅｃだけ遅延させてレシーバ６に出力する。レシーバ６はこの音声信号を音声（空気振動）に変換して外耳道４内に放出する。遅延装置１０が遅延したｐｏｉｎｔＢ波形は伝達関数補正装置１４にも入力される。伝達関数補正装置１４は入力された信号をフィルタ処理して差分処理装置１３に入力する。このフィルタは、この装置が外耳道４に挿入されているときの外耳道４のインパルス応答をシミュレートしたフィルタであり、このフィルタを通過した信号が外耳道４内を伝搬した信号と同じ波形になるように設計されている。すなわち、このフィルタを通過させることにより、信号波形に対して伝搬路の伝達関数が畳み込み演算される。

【００３１】一方、マイク７は、外耳道４内に発生した音声を受信して差分処理装置１３に入力する。マイク７は、レシーバ６が出力した音声および外耳道４から伝達された装用者の音声の両方を受信する。差分処理装置１３はマイク７から入力された音声信号波形から伝達関数補正装置１４から入力された音声信号波形を減算してこの成分をキャンセルする。このレシーバ６の出力分がキャンセルされた音声信号波形を相関演算装置１１に入力する。相関演算装置１１は、マイク５から入力された音声信号と差分処理装置１３から入力された音声信号の相関値を計算する。

【００３２】図４（Ｂ）を参照して外来音声が到来したときの信号波形と相関について説明する。マイク５が外来音声を受信した時刻を０ｓｅｃとする。マイク５の入力であるｐｏｉｎｔＡは、同図（Ｂ）第１段のような波形となる。遅延装置１０はこの音声信号をｔｓｅｃ遅延させてレシーバ６に出力する。レシーバ６にされる波形であるｐｏｉｎｔＢは、ｐｏｉｎｔＡの波形をｔｓｅｃ遅延させた同図（Ｂ）第２段のような波形となる。レシーバ６はこの遅延された音声信号を外耳道４に出力す

る。外耳道４に出力された音声はマイク７にも受信される。マイク７の入力であるｐｏｉｎｔＣは、ｐｏｉｎｔＢとタイミングはほぼ同じであるが、外耳道４を伝搬したことによって遅延成分が重畳されており同図（Ｂ）第３段のような波形となる。差分処理装置１３には、遅延装置１０から出力されるｐｏｉｎｔＢ波形を伝達関数補正装置１４で補正した波形が入力されるため、この波形をマイク７から入力された音声信号波形であるｐｏｉｎｔＣから減算する。ｐｏｉｎｔＢの波形を伝達関数補正装置１４で補正したものはほぼｐｏｉｎｔＣの波形と同じになっているため、これらの波形はほとんど一致し差分処理装置１３の出力であるｐｏｉｎｔＥ波形は、同図（Ｂ）第５段に示すように殆どＯレベルのフラットな信号となる。相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＥの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＥの波形は殆どフラットであるためｐｏｉｎｔＡの基準波形と殆ど相関性がなく、大きな相関値が出力されない。発話判定装置１２は大きな相関値が入力されないことにより、自己発話がないことを判断する。

【００３３】一方、図４（Ｃ）を参照して自己発話が行われたときの信号波形と相関について説明する。装用者が発声すると、その音声信号は口から外界に放出されるとともに体内伝搬経路によって外耳道４内にも伝搬する。外界に放出された音声はマイク５に受信され、外耳道４に伝搬した音声はマイク７に受信される。双方とも伝搬距離が極めて短いため、ほぼ同時（０ｓｅｃ）に受信される。したがって、マイク５には時刻０ｓｅｃにｐｏｉｎｔＡの波形が現れ、マイク７にも時刻０ｓｅｃにｐｏｉｎｔＡ波形と類似した波形が現れる。また、ｐｏｉｎｔＡの波形は遅延装置１０でｔｓｅｃ遅延されたのちレシーバ６から外耳道４に放出されるので、マイク７はこの音声も受信する。レシーバ６から放出されマイク７が受信する信号は外耳道４を伝搬して入力される信号であるため、外耳道４による残響音成分が重畳されたものになっている。このためマイク７が受信する自己発話の直接音とレシーバ６から出力された音の合成波形は、同図（Ｃ）第３段に示すｐｏｉｎｔＣ波形となる。レシーバ６から放出され外耳道４を伝搬した音声信号は、上述したように外耳道４の伝搬特性によって残響分が重畳されているため、波形の立ち上がりと立ち下がりが緩やかになっている。伝達関数補正装置１４から出力された波形も外耳道４の伝達特性をフィルタ係数として有するフィルタを通過したものであるため、同じような残響成分が重畳されており、この信号をマイク７の入力信号から減算することにより、自己発話音のみを分離することができる（ｐｏｉｎｔＥ波形）。

【００３４】相関演算装置１１は、ｐｏｉｎｔＡの波形を基準波形とし、ｐｏｉｎｔＤの波形が基準波形にどの程度類似しているかを時系列に検査する装置であり、その相関の程度に比例した相関値を出力する。この例では、ｐｏｉｎｔＤの波形は殆どｐｏｉｎｔＡの波形と一致しているため、０ｓｅｃのタイミングに大きな相関値が出力される。発話判定装置１２は、大きな相関値が入力されたことにより、０ｓｅｃのタイミングに開始した自己発話があったことを判断する。

【００３５】なお、図４（Ａ）において、差分処理装置１３の出力をそのまま取り出せば自己発話検出マイクとして機能させることができる。

【００３６】なお、図４（Ａ）の装置では、伝達関数補正装置１４を遅延装置１０と差分処理装置１３の間に挿入したが、図５のように、マイク７と差分処理装置１３の間におくこともできる。この場合、伝達関数が外耳道内に発生した自己発話音声にも畳み込まれてしまうので、差分処理後自己発話音の波形をもとに戻す伝達関数修正装置を差分処理装置１４と相関演算装置１１との間に設ける。

【００３７】なお、図３および図４の例において、相関演算装置１１は、同時に発生する２つの音声信号波形を比較するのみであるため、単純な比較演算装置でもよい。さらに、図４，図５の遅延装置１０は、ディジタル的に信号を遅延させる信号処理装置であってもよい。

【００３８】図６は、この発明を通信端末の音声入出力装置に適用した実施形態を示す図である。この装置を制御する制御装置３０は伝達関数演算モードと通常動作モードで装置を制御する。指示装置３３は装用者がこの装置のオン／オフを制御装置３０に指示するためのスイッチを含んでいる。

【００３９】図６において、図１および図４（Ａ）のマイク５－遅延装置１０－レシーバ６の系統として、マイク５（この装置も外観は図１のものと同様である：以下同じ）、増幅器３６、Ａ／Ｄ変換器３７、遅延装置３８、Ｄ／Ａ変換器３９、増幅器４０およびレシーバ６が設けられている。また、図１および図４（Ａ）のマイク７の系統として、マイク７－増幅器４１、Ａ／Ｄ変換器４２が設けられている。また、図４（Ａ）の伝達関数補正装置１４、差分処理装置１３、比較演算装置１１の機能は、制御装置３０および演算処理装置３５がソフト的に実現する。また、この装置は、外耳道４の伝達関数を割り出す伝達関数演算モードにテスト音信号を発生するための音源３４を内蔵している。この音源３４が形成したオーディオ信号は、テスト音信号としてマイク５の入力に変えてＡ／Ｄ変換器３７に入力される。

【００４０】指示装置３３によってオンが指示されると、制御部３０は最初に伝達関数モードを設定し、制御装置３０は音源３４をオンするとともに、演算処理装置３５に伝達関数演算動作を指示する。音源３４が形成するテスト音信号は、ホワイトノイズやピンクノイズでよいが、他の信号、例えばインパルスや周波数の定まった

低周波信号の波形を形成するようにしてもよい。音源３４から出力されたテスト音信号はＡ／Ｄ変換器３７、遅延装置３８、Ｄ／Ａ変換器３９、増幅器４０、レシーバ６を経て外耳道４内に放出される。外耳道４内に放出されたテスト音は外耳道４に向けて設けられたマイク７で検出される。マイク７の出力は増幅器４１、Ａ／Ｄ変換器４２を経て演算処理装置３５に入力される。演算処理装置３５はＤＳＰなどで構成すればよい。一方、この演算処理装置３５には、Ａ／Ｄ変換器３７から音源３４が形成したテスト音信号をディジタル信号に変換したものも供給される。演算処理装置３５は、伝達関数演算動作として、ｐｏｉｎｔＡ信号（Ａ／Ｄ変換器３７から入力される信号）およびｐｏｉｎｔＢ信号（Ａ／Ｄ変換器４２から入力される信号）を比較し、これらの信号の相違に応じた補正係数を算出する。補正係数を求める手法は色々あるが例えば下式のようにｐｏｉｎｔＡ信号をＡ（ｋ）、ｐｏｉｎｔＢ信号をＢ（ｋ）とし、Ａ（ｋ）とＢ（ｋ）間の誤差Ｅ（ｋ）がゼロになるような関数Ｇ（ｋ）を求める。

【００４１】Ｇ（ｋ）・Ａ（ｋ）－Ｂ（ｋ）＝Ｅ（ｋ）

一度求めた関数Ｇ（ｋ）は装置を外耳道から取り外したりしなければ、比較的安定した状態にあるので、動作の最初に計算するのみでよい。演算が終了すると演算処理装置３５は待機状態になる。制御装置３０は演算処理装置３５が待機状態になったことを検出すると、音源３４を停止するとともに、演算処理装置３５に信号処理動作を指示して通常動作モードに切り換える。

【００４２】通常動作モードにおいて、マイク５から外来音声が入力されると、増幅器３６、Ａ／Ｄ変換器３７、遅延装置３８、Ｄ／Ａ変換器３９、増幅器４０を経て、ｔｓｅｃ後にレシーバ６から外耳道４に出力される。外耳道４に出力された外来音声は外耳道４に向けて設けられたマイク７によって受信され、増幅器４１、Ａ／Ｄ変換器４２を経て演算処理装置３５に入力される。信号処理動作が設定されている演算処理装置３５は、Ａ／Ｄ変換器３７から入力されるｐｏｉｎｔＡ信号、すなわち直接の外来音声信号に上記で求めた伝達関数Ｇ（ｋ）を畳み込んだのち、ｐｏｉｎｔＢ信号、すなわち外耳道４を伝達した外来音声信号を含む信号と差演算処理を行う。この処理によってｐｏｉｎｔＢ信号から外来音声成分は消去され自己発話音声のみが抽出される。抽出された自己発話音声を通信装置の音声入力端子に接続すれば環境雑音が大きな場所でも雑音の少ない音声信号のみをピックアップして相手に送ることができる。

【００４３】なお、上記例では、参照信号としてＡ／Ｄ変換器３７の出力（ｐｏｉｎｔＡ信号）を演算処理装置３５に入力しているが、遅延装置３８の出力を入力するようにしてもよい。

【００４４】図７はこの発明を話速変換機能をもった補聴器に適用した実施形態を示す図である。この装置を制御する制御装置５０は伝達関数演算モードと通常動作モードで装置を制御する。指示装置５３は装用者がこの装置のオン／オフを制御装置５０に指示するためのスイッチを含んでいる。

【００４５】図７において、図１および図４（Ａ）のマイク５－遅延装置１０－レシーバ６の系統として、マイク５（この装置も外観は図１のものと同様である：以下同じ）、増幅器５６、Ａ／Ｄ変換器５７、話速変換・利得制御装置５８、Ｄ／Ａ変換器５９、増幅器６０およびレシーバ６が設けられている。また、図１および図４（Ａ）のマイク７の系統として、マイク７－増幅器６１、Ａ／Ｄ変換器６２が設けられている。また、図４（Ａ）の伝達関数補正装置１４、差分処理装置１３、比較演算装置１１の機能は、制御装置５０および演算処理装置５５がソフト的に実現する。また、この装置は、外耳道４の伝達関数を割り出す伝達関数演算モードにテスト音信号を発生するための音源５４を内蔵している。この音源５４が形成したオーディオ信号は、外部に向けて設けられたレシーバ９によって出力され、マイク５によって受信される。

【００４６】指示装置５３によってオンが指示されると、制御部５０は最初に伝達関数モードを設定し、制御装置５０は音源５４をオンするとともに、演算処理装置５５に伝達関数演算動作を指示する。音源５４が形成するテスト音信号は、ホワイトノイズやピンクノイズでよいが、他の信号、例えばインパルスや周波数の定まった低周波信号の波形を形成するようにしてもよい。音源５４から出力されたテスト音信号は、レシーバ９から出力され、マイク５で受信・電気信号に変換されたのち、Ａ／Ｄ変換器５７、話速変換・利得制御装置５８、Ｄ／Ａ変換器５９、増幅器６０、レシーバ６を経て外耳道４内に放出される。外耳道４内に放出されたテスト音は外耳道４に向けて設けられたマイク７で検出される。マイク７の出力は増幅器６１、Ａ／Ｄ変換器６２を経て演算処理装置５５に入力される。演算処理装置５５はＤＳＰなどで構成すればよい。一方、この演算処理装置５５には、話速変換・利得制御装置５８から音源５４が形成し、レシーバ９から出力されたテスト音信号を話速変換したものも供給される。演算処理装置３５は、伝達関数演算動作として、ｐｏｉｎｔＣ信号（話速変換・利得制御装置５８から入力される信号）およびｐｏｉｎｔＢ信号（Ａ／Ｄ変換器６２から入力される信号）を比較し、これらの信号の相違に応じた補正係数を算出する。補正係数を求める手法は色々あるが例えば下式のようにｐｏｉｎｔＡ信号をＡ（ｋ）、ｐｏｉｎｔＢ信号をＢ（ｋ）とし、Ａ（ｋ）とＢ（ｋ）間の誤差Ｅ（ｋ）がゼロになるような関数Ｇ（ｋ）を求める。

【００４７】Ｇ（ｋ）・Ａ（ｋ）－Ｂ（ｋ）＝Ｅ（ｋ）

一度求めた関数Ｇ（ｋ）は装置を外耳道から取り外したりしなければ、比較的安定した状態にあるので、動作の

最初に計算するのみでよい。演算が終了すると演算処理装置５５は待機状態になる。制御装置５０は演算処理装置５５が待機状態になったことを検出すると、音源５４を停止するとともに、演算処理装置５５に信号処理動作を指示して通常動作モードに切り換える。

【００４８】通常動作モードにおいて、マイク５から外来音声が入力されると、増幅器５６、Ａ／Ｄ変換器５７、話速変換・利得制御装置５８、Ｄ／Ａ変換器５９、増幅器６０を経て、ｔｓｅｃ後にレシーバ６から外耳道４に出力される。外耳道４に出力された外来音声は外耳道４に向けて設けられたマイク７によって受信され、増幅器６１、Ａ／Ｄ変換器６２を経て演算処理装置５５に入力される。信号処理動作が設定されている演算処理装置５５は、話速変換・利得制御装置５８から入力されるｐｏｉｎｔＣ信号、すなわち直接の外来音声信号に上記で求めた伝達関数Ｇ（ｋ）を畳み込んだのち、ｐｏｉｎｔＢ信号、すなわち外耳道４を伝達した外来音声信号を含む信号と差演算処理を行う。この処理によってｐｏｉｎｔＢ信号から外来音声成分は消去され自己発話音声成分のみが抽出される。マイク５が外来音声のみを受信しているときは、この自己発話音声成分は０であるが、自己発話しているときには、そのレベルの成分が抽出される。制御装置５０は演算処理装置５５が自己発話音声成分を抽出したとき、話速変換・利得制御装置５８に対して話速変換を禁止し、利得を小さめに制御する。これにより、自己発話に対する話速変換が禁止され、自己発話が滑らかに行われるようになる。

【００４９】なお、上記例では、参照信号としてＡ／Ｄ変換器５７の出力（ｐｏｉｎｔＡ信号）を演算処理装置５５にも入力するようにすればよい。

【００５０】これにより、話速変換は外来音声に作用させたいが、自分の発声音声には作用させたくない。自分が発話するとき、人は自分の声を聞きながら発声を制御しているため、発声音声が遅れて聞こえるとうまく話せないという問題が起こる。そのため、自己発話に対しては、発話区間を検出して話速変換処理を適応しないようにする必要がある。また外来音声に比べ、自己発話音声は大きく聞こえるので適当な大きさになるよう利得制御が必要となる。

【００５１】このように、簡略な構成で自己発話と外来音とを分離することができるため、話速変換補聴器にこれを用いた場合に、装用者の外耳道の形状にかかわらず同じセルでこれを使用することができる。

【００５２】

【発明の効果】この発明によれば、外耳道に向けて設けられた第２のマイクの入力信号から外来音声信号をキャンセルして自己発話音のみを抽出できるようにしたことにより、構造が複雑で装用者毎のセッティングが微妙な骨伝導マイクを用いなくても環境ノイズなどの外来音声を除去した自己発話音声のみを抽出処理することができる。

【００５３】また、これを補聴器にもちいれば自己発話音声の区間のみ話速変換を禁止し、発話が容易な補聴器を構成することができる。

【図面の簡単な説明】

【図１】この発明の実施形態である自己発話検出装置の外観図

【図２】請求項１の発明の実施形態である自己発話検出装置のブロック図および信号波形を示す図

【図３】請求項２の発明の実施形態である自己発話検出装置のブロック図および信号波形を示す図

【図４】請求項３の発明の実施形態である自己発話検出装置のブロック図および信号波形を示す図

【図５】道自己発話検出装置の他の実施形態を示す図

【図６】請求項４の発明の実施形態である音声入力装置のブロック図号波形を示す図

【図７】この発明の自己発話検出装置を補聴器に応用した例を示す図

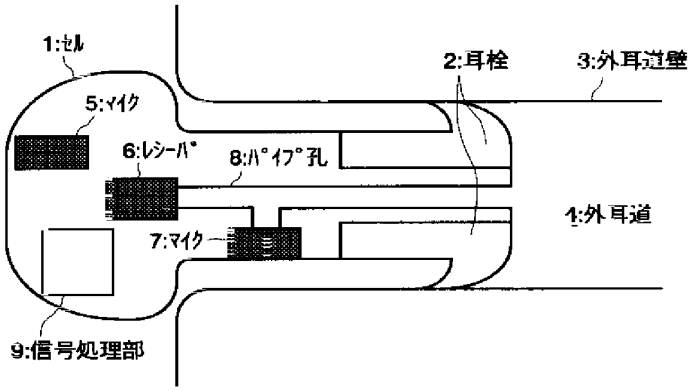
【図８】自己発話検出装置の各部における信号波形の例を示す図

【図９】自己発話検出装置の各部における信号波形の例を示す図

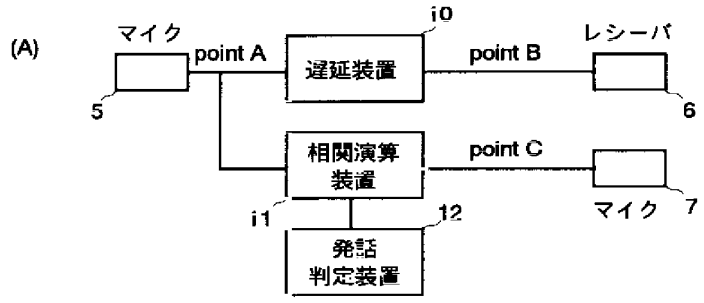
【符号の説明】

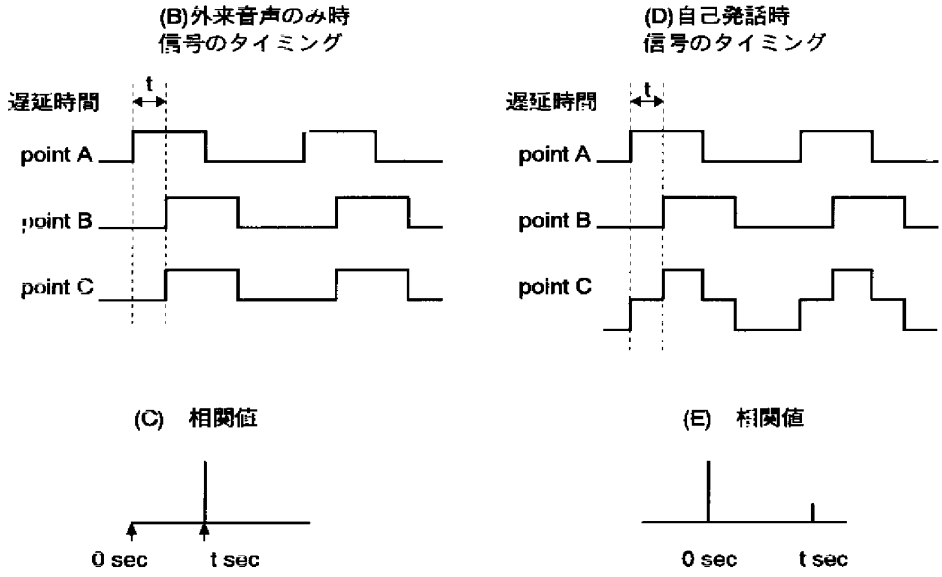
１…セル、２…耳栓、３…外耳道壁、４…外耳道、５…（外界に向けられた）マイク、６…レシーバ、７…（外耳道に向けられた）マイク、８…パイプ孔、９…信号処理部、１０…遅延装置、１１…相関演算装置、１２…発話判定装置、１３…差分処理装置、１４…伝達関数補正装置、

【図１】

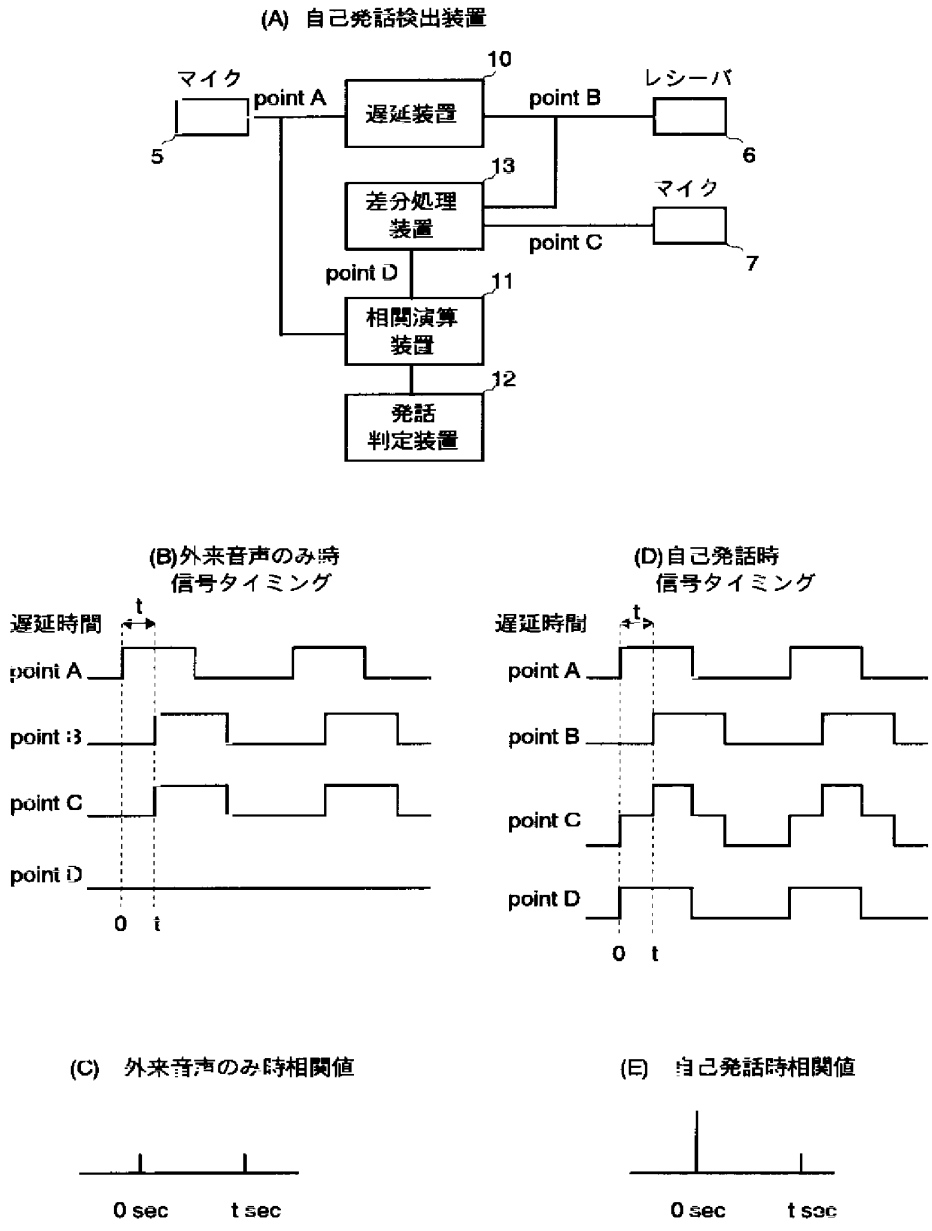

1:セル
2:耳栓
3:外耳道壁
5:マイク
6:レシーバ
8:パイプ孔
4:外耳道
7:マイク
9:信号処理部

【図２】

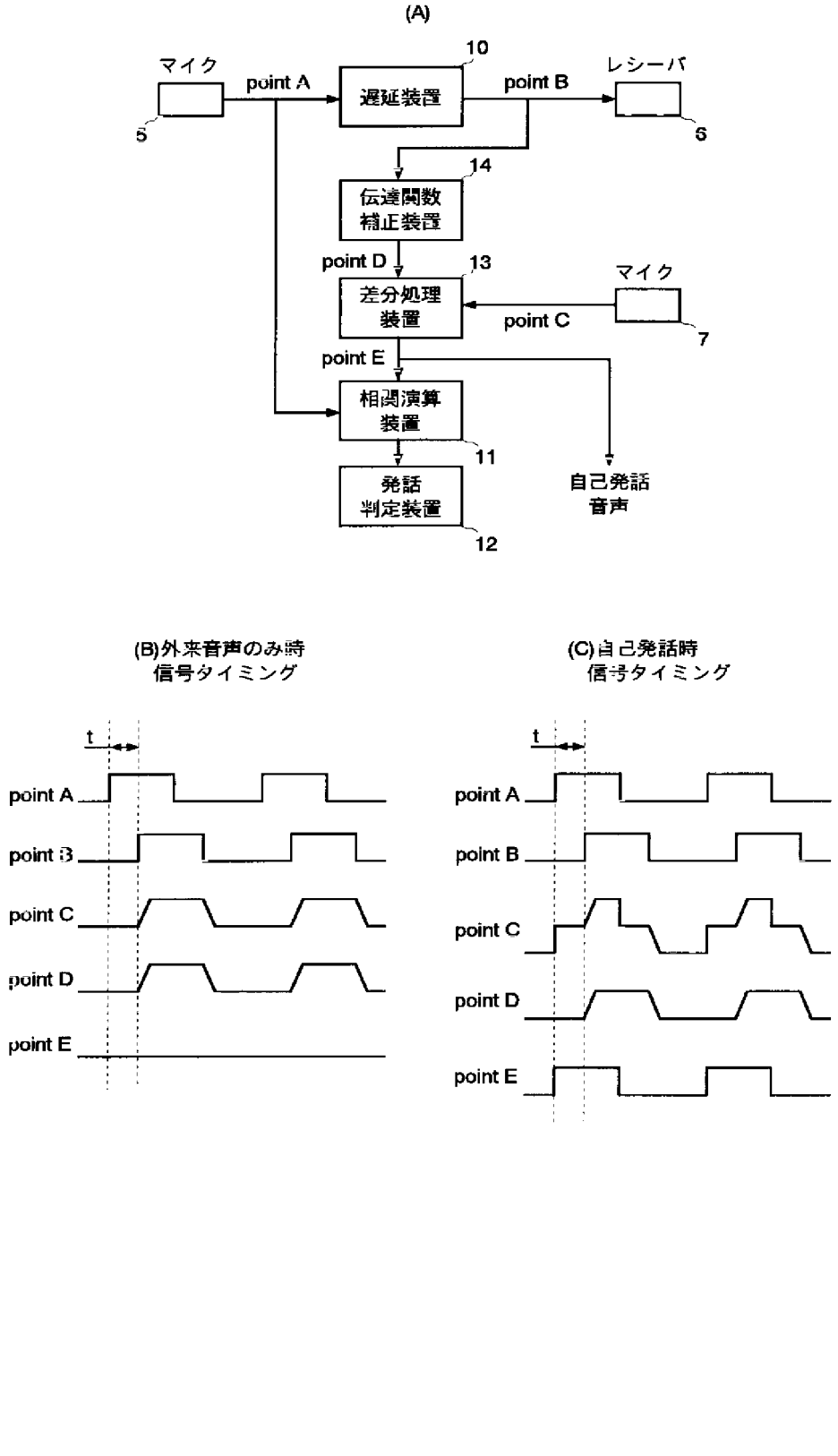

(A)
マイク
point A
10
遅延装置
point B
レシーバ
5
6
相関演算
装置
point C
マイク
7
11
12
発話
判定装置


(B)外来音声のみ時
信号のタイミング
遅延時間
t
point A
point B
point C
(D)自己発話時
信号のタイミング
遅延時間
t
point A
point B
point C
(C)　相関値
0 sec
t sec
(E)　相関値
0 sec
t sec

【図３】

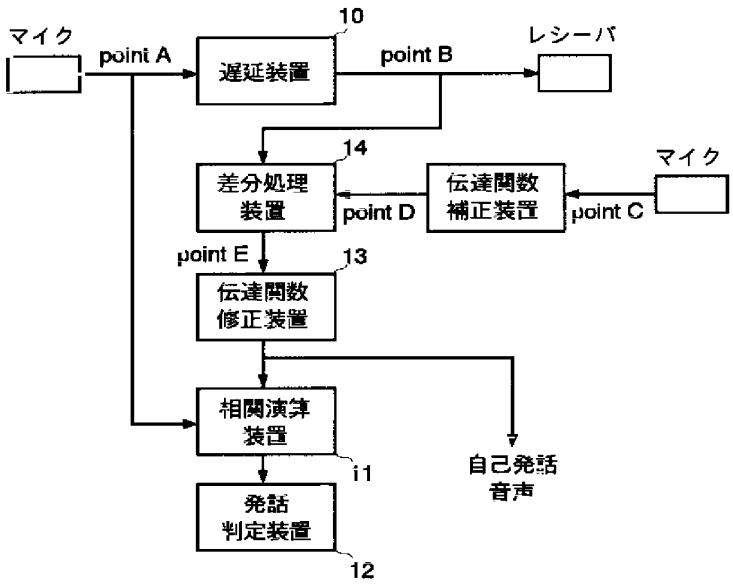

(A) 自己発話検出装置
マイク
point A
遅延装置
10
point B
レシーバ
5
6
差分処理装置
13
マイク
point C
7
point D
11
相関演算装置
12
発話判定装置
(B)外来音声のみ時信号タイミング
遅延時間
t
point A
point B
point C
point D
0
t
(D)自己発話時信号タイミング
遅延時間
t
point A
point B
point C
point D
0
t
(C) 外来音声のみ時相関値
0 sec
t sec
(E) 自己発話時相関値
0 sec
t sec

【図４】

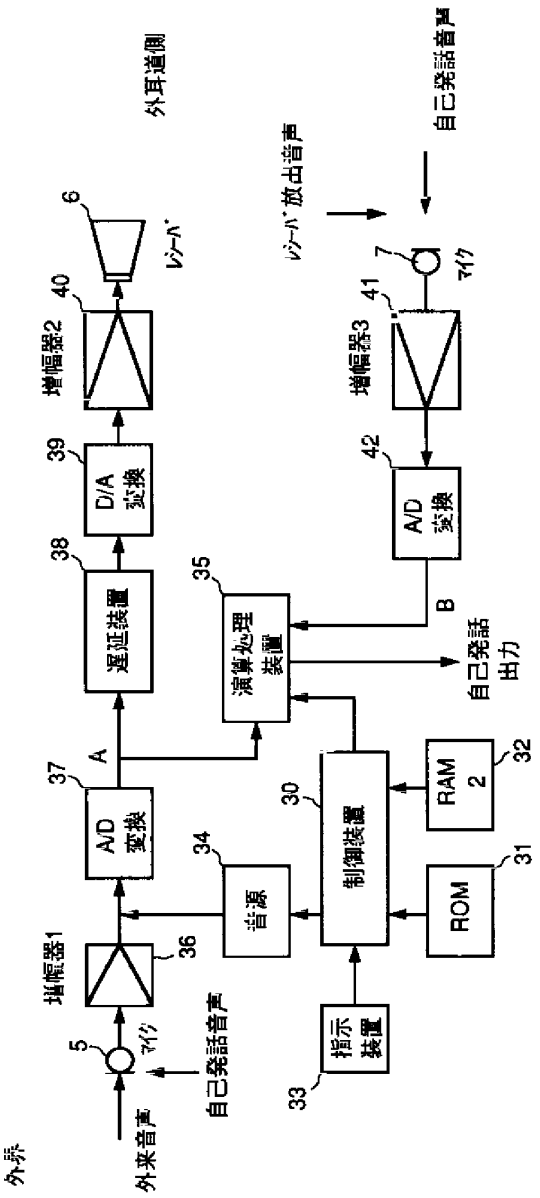

(A)
マイク
point A
10
遅延装置
point B
レシーバ
5
6
14
伝達関数
補正装置
point D
13
差分処理
装置
マイク
point C
7
point E
相関演算
装置
11
発話
判定装置
12
自己発話
音声
(B)外来音声のみ時
信号タイミング
t
point A
point B
point C
point D
point E
(C)自己発話時
信号タイミング
t
point A
point B
point C
point D
point E

【図５】

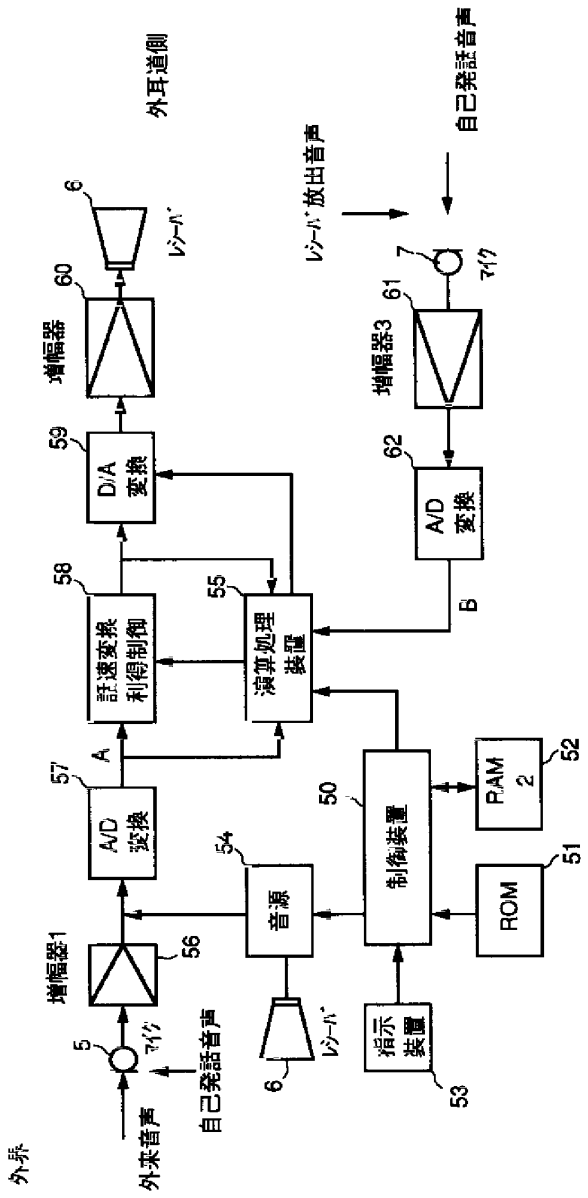
】


マイク
point A
10
遅延装置
point B
レシーバ
14
差分処理
装置
伝達関数
補正装置
マイク
point D
point C
point E
13
伝達関数
修正装置
相関演算
装置
11
自己発話
音声
発話
判定装置
12

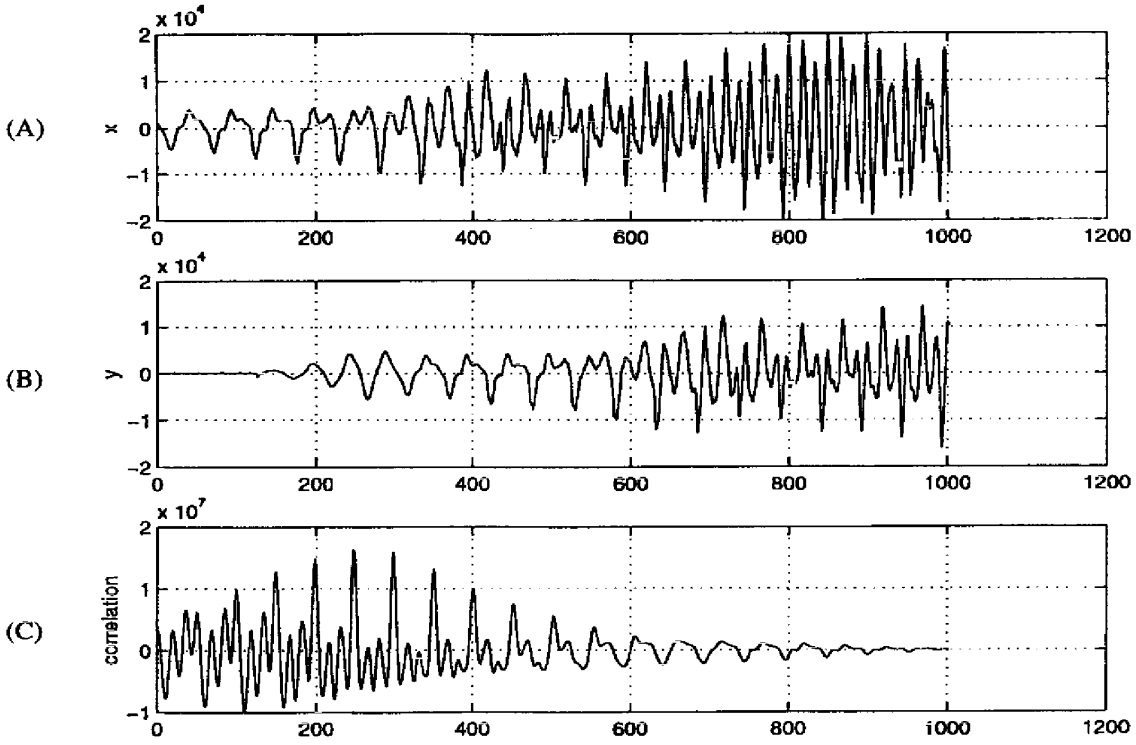
【図６】


外界
外来音声
5
マイク
増幅器1
36
A/D変換
37
A
遅延装置
38
D/A変換
39
増幅器2
40
6
レシーバ
外耳道側
レシーバ放出音声
自己発話音声
7
マイク
増幅器3
41
A/D変換
42
B
演算処理装置
35
自己発話出力
音源
34
制御装置
30
指示装置
33
ROM
31
RAM 2
32
自己発話音声

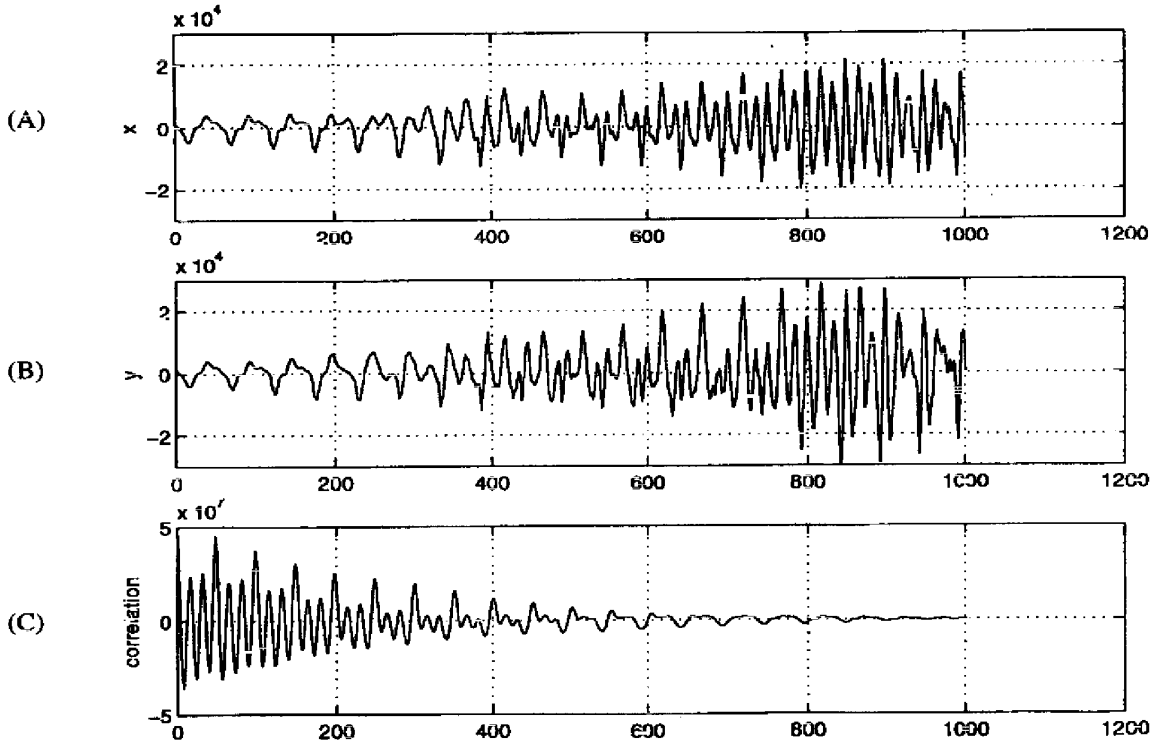
【図７】


外界
外来音声
5
マイク
増幅器1
56
A/D変換
57
A
証速変換利得制御
58
D/A変換
59
増幅器
60
6
レシーバ
外耳道側
レシーバ放出音声
自己発話音声
7
マイク
増幅器3
61
A/D変換
62
B
演算処理装置
55
音源
54
制御装置
50
指示装置
53
ROM
51
RAM 2
52
自己発話音声

【図８】


(A)
x 10^4
x
2
1
0
−1
−2
0
200
400
600
800
1000
1200
(B)
x 10^4
y
2
1
0
−1
−2
0
200
400
600
800
1000
1200
(C)
x 10^7
correlation
2
1
0
−1
0
200
400
600
800
1000
1200

【図９】



フロントページの続き

(72)発明者　谷高　幸司
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

(72)発明者　大脇　浩
静岡県浜松市中沢町10番１号　ヤマハ株式会社内

(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3556987号
(P3556987)
(45)発行日　平成16年8月25日(2004.8.25)　(24)登録日　平成16年5月21日(2004.5.21)

(51)Int.Cl.[7]　F I
H04R 1/10　H04R 1/10　101A
H04R 5/033　H04R 5/033　Z

請求項の数 4　(全 9 頁)

(21)出願番号　特願平7-18867
(22)出願日　平成7年2月7日(1995.2.7)
(65)公開番号　特開平8-214392
(43)公開日　平成8年8月20日(1996.8.20)
審査請求日　平成13年1月29日(2001.1.29)

(73)特許権者　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　100087848
弁理士　小笠原　吉義
(74)代理人　100074848
弁理士　森田　寛
(74)代理人　100087147
弁理士　長谷川　文廣
(72)発明者　増井　誠生
神奈川県川崎市中原区上小田中１０１５番地　富士通株式会社内

審査官　大野　弘

最終頁に続く

(54)【発明の名称】環境音伝送型ヘッドセット装置

(57)【特許請求の範囲】
【請求項1】
会話用マイクを備えたヘッドセット装置において，
オーディオ信号を入力するオーディオ信号入力端子と，
外部環境音を収録する環境音収録用マイクと，
音響提示用イヤースピーカと，
前記オーディオ信号入力端子から入力したオーディオ信号に所定の周波数帯域の信号が含まれるかどうかをチェックするオーディオ帯域チェッカと，
前記オーディオ信号入力端子から入力したオーディオ信号と前記環境音収録用マイクから入力した外部環境音とを，前記オーディオ帯域チェッカのチェック結果に基づき音量比率を変化させて混合し，前記音響提示用イヤースピーカへ出力する混合ボリュームとを備える
ことを特徴とする環境音伝送型ヘッドセット装置。
【請求項2】
会話用マイクを備えたヘッドセット装置において，
外部環境音を収録する環境音収録用マイクと，
オーディオ信号を出力するオーディオ信号出力端子と，
前記会話用マイクから入力した音声と前記外部環境音とを混合し前記オーディオ信号出力端子へ出力する混合ボリュームとを備える
ことを特徴とする環境音伝送型ヘッドセット装置。

【請求項3】
可動式のブーム付き会話用マイクを備えたヘッドセット装置において，
外部環境音を収録する環境音収録用マイクと，
オーディオ信号を出力するオーディオ信号出力端子と，
前記ブーム付き会話用マイクのブーム位置を検出するブーム位置検出器と，
前記ブーム付き会話用マイクから入力した会話音と前記環境音収録用マイクから入力した外部環境音とを，前記ブーム位置検出器により検出したブーム位置に基づき音量比率を変化させて混合し，前記オーディオ信号出力端子へ出力する混合ボリュームとを備える
ことを特徴とする環境音伝送型ヘッドセット装置。
【請求項4】
可動式のブーム付き会話用マイクを備えたヘッドセット装置において，
オーディオ信号を入力するオーディオ信号入力端子と，
オーディオ信号を出力するオーディオ信号出力端子と，
外部環境音を収録する環境音収録用マイクと，
音響提示用イヤースピーカと，
前記オーディオ信号入力端子から入力したオーディオ信号に所定の周波数帯域の信号が含まれるかどうかをチェックするオーディオ帯域チェッカと，
前記オーディオ信号入力端子から入力したオーディオ信号と前記環境音収録用マイクから入力した外部環境音とを，前記オーディオ帯域チェッカのチェック結果に基づき音量比率を変化させて混合し，前記音響提示用イヤースピーカへ出力する第１の混合ボリュームと，
前記ブーム付き会話用マイクのブーム位置を検出するブーム位置検出器と，
前記ブーム付き会話用マイクから入力した会話音と前記環境音収録用マイクから入力した外部環境音とを，前記ブーム位置検出器により検出したブーム位置に基づき音量比率を変化させて混合し，前記オーディオ信号出力端子へ出力する第２の混合ボリュームとを備える
ことを特徴とする環境音伝送型ヘッドセット装置。

【発明の詳細な説明】
【０００１】
【産業上の利用分野】
本発明は，外部環境音を収録して伝送することができる音声対話用のヘッドセット装置に関するものである。
【０００２】
特に，遠隔地間で利用することにより，会話と同時に現場の環境を音響として伝達することができ，臨場感のある双方向音声伝達が可能となる。自分が体験している音響空間の雰囲気を自己の音声による解説等とともに遠隔地にいる相手に対して伝えるというような臨場感通信の分野において，特に有効である。
【０００３】
【従来の技術】
従来のヘッドセットマイクは主に会話や歌唱等の音声収録用であり，耳掛け型マイク，ヘッドホン一体型マイク等がある。また，可動式のブームに設けられた音声収録用マイクを上下することで，音声収録用マイクのスイッチをＯＮ／ＯＦＦする機能を備えるものもある。このように，従来のヘッドセットマイクは，会話等のための音声収録用の装置として利用されている。
【０００４】
【発明が解決しようとする課題】
従来の音声対話用のマイクが付いたヘッドホンでは，通常，利用者の音声だけをモノラルで伝えることができるだけであり，ヘッドホン利用者のいる場所の雰囲気，状況等を相手の対話者へ伝えることができなかった。
【０００５】

本発明は上記問題点の解決を図り，会話音声と同時に，利用者が存在する場所の雰囲気をステレオ音響として伝えることができるようにし，臨場感のある遠隔地対話を可能にする手段を提供することを目的とする。ここで，会話音声と外部環境音を自然に混合して伝えること，および，聴取する側で，複数箇所の外部環境音が混乱して聴こえることのないように，状況に応じて自分のいる場所の環境音が遮断されるような配慮が望ましい。

【０００６】

【課題を解決するための手段】

図１は，本発明の原理説明図である。図１において，１はオーディオ信号入力端子，２は環境音収録用マイク，３はブーム付き会話用マイク，４，５は混合ボリューム，６は音響提示用イヤースピーカ（ヘッドホン），７はオーディオ信号出力端子，８はオーディオ帯域チェッカ，９はブーム位置検出器，１０は可動式ブームを表す。

【０００７】

本発明は，上記の課題を解決するため，特にヘッドセット装置に外部環境音を収録するための環境音収録用マイク２を備え，この環境音収録用マイク２から入力した外部環境音と相手話者の音声または自分の音声とを音量比率を変えて混合する混合ボリューム４，５を備えることを主要な特徴とする。

【０００８】

オーディオ信号入力端子１は，遠隔地の対話者の会話音声と外部環境音との混合音のオーディオ信号を入力する入力端子である。環境音収録用マイク２は，利用者の存在する周囲の外部環境音を収録するマイクであり，例えば利用者本人が聴こえる環境音と同様な音を収録するために，音響提示用イヤースピーカ６の近辺に設けられる。ブーム付き会話用マイク３は，可動式ブーム１０を備え，利用者の会話音を収録するマイクである。

【０００９】

混合ボリューム４は，オーディオ信号入力端子１からのオーディオ信号と環境音収録用マイク２からの外部環境音とを混合する回路である。混合ボリューム４は，オーディオ帯域チェッカ８の制御により，混合するオーディオ入力信号と外部環境音との音量比率を変化させる。

【００１０】

混合ボリューム５は，環境音収録用マイク２からの外部環境音とブーム付き会話用マイク３からの会話音とを混合する回路である。混合ボリューム５は，ブーム位置検出器９の検出結果により，混合する外部環境音と会話音との音量比率を変化させる。

【００１１】

音響提示用イヤースピーカ６は，混合ボリューム４で混合されたオーディオ信号（相手の話声と相手側の外部環境音）と利用者のいる外部環境音との混合音を出力する手段である。オーディオ信号出力端子７は，環境音収録用マイク２から入力した利用者のいる場所の外部環境音と利用者自身の話声とを混合ボリューム５で混合した音声信号を，相手側へ送るためにオーディオ出力する出力端子である。

【００１２】

オーディオ帯域チェッカ８は，オーディオ信号入力端子１から入力される音声の周波数帯域を監視することにより，入力したオーディオ信号に外部環境音が含まれるか相手の話声だけであるかを判定して，オーディオ信号入力端子１からのオーディオ信号と環境音収録用マイク２からの外部環境音との音量比率を変化させる混合ボリューム４の制御信号を送出する。

【００１３】

ブーム位置検出器９は，ブーム付き会話用マイク３に備えられた可動式ブーム１０のブーム位置を検出し，環境音収録用マイク２からの外部環境音とブーム付き会話用マイク３からの会話音との音量比率を変化させ，または，ブーム付き会話用マイク３および環境音収録用マイク２のＯＮ／ＯＦＦの制御信号を送出する。

【００１４】

ブーム位置検出器９は，例えば図１（Ｂ）に示すように，その可動式ブーム１０の上下動

に応じた角度を検出することより，ブーム付き会話用マイク３が利用者の口の近くにある場合には，利用者の会話音の混合比率が大きくなり，それより下方にある場合には，環境音収録用マイク２からの環境音の混合比率がブームの角度に応じて増加し，最上位のＯＦＦ位置にある場合には環境音収録用マイク２およびブーム付き会話用マイク３のスイッチがＯＦＦになるように混合ボリューム５およびマイクスイッチ（図示省略）を制御する。

【００１５】

【作用】

本発明では，ブーム付き会話用マイク３の可動式ブーム１０の上下動によって，オーディオ信号出力端子７への音声出力における，外部環境音と会話音との比率を変更できるようにしている。さらに，利用者自身の音響提示用イヤースピーカ６への出力について外部環境音と相手側からのオーディオ入力信号との混合比率を変更できるようにしている。

【００１６】

例えば，ユーザＡとユーザＢの間でユーザＡが発話する場合，ユーザＢが受けるユーザＡからのオーディオ信号は，外部環境音に対して会話音の音量比率が強くなっている。ユーザＢは，通常の場合，ユーザＢ自身の周辺の外部環境音が聴こえるが，ユーザＡからのオーディオ信号の入力に，ユーザＡの外部環境音が強く含まれる場合には，ユーザＢの外部環境音の比率が弱まり，ユーザＡの外部環境音が聴こえるように制御される。これにより，ユーザＡとユーザＢとの会話時には，互いの環境音が同時に聴こえて混乱するといった状況が防止される。

【００１７】

【実施例】

以下，本発明の一実施例について説明する。

図２は，本発明の実施例の外観構成を示す図である。

【００１８】

図２において，２０はオーディオ帯域チェッカ／混合ボリューム部，２１はヘアーバンド，２２はステレオ入出力用４線ケーブル，２３はオーディオＩ／Ｏコネクタ，２４はステレオ入力端子，２５はステレオ出力端子を表す。

【００１９】

本装置は，クローズドエアー型ヘッドホン６Ｌ，６Ｒの左右イヤーパッド内に単一指向性マイクによって構成される環境音収録用マイク２Ｌ，２Ｒを備える。またはヘッドバンド２１の左右の耳付近に環境音収録用マイク２Ｌ’，２Ｒ’を設けてもよい。これらの左右の単一指向性マイクにより，利用者のいる周囲の環境音をステレオ音響で収録することができる。

【００２０】

会話用マイク３は，利用者の会話を収録するマイクである。この会話用マイク３は，可動式ブーム１０によって最上位のＯＦＦ位置と口元のセット位置と顎下の付近の最下位置の間を自由に動かすことができる。

【００２１】

利用者は，可動式ブーム１０を口元のセット位置と最下位置との間で上下に移動させることにより，会話用マイク３で収録する会話音と環境音収録用マイク２Ｌ，２Ｒで収録する利用者の外部の環境音との混合比率を変化させる。口元のセット位置では，会話音の比率が最大であり，最下位置では環境音の比率が最大となる。また，使用しないときには，可動式ブーム１０を押し上げて会話用マイク３をＯＦＦ位置にする。これにより，会話用マイク３および環境音収録用マイク２Ｌ，２Ｒのスイッチが共にＯＦＦとなり，会話音も外部環境音も収録されない。

【００２２】

会話用マイク３で収録された会話音と環境音収録用マイク２Ｌ，２Ｒで収録された外部環境音とは，可動式ブーム１０の位置により設定された混合比率により左右別々に混合され，ステレオ入出力用４線ケーブル２２を介して，オーディオＩ／Ｏコネクタ２３のステレオ出力端子２５からオーディオ出力される。

【００２３】
オーディオ帯域チェッカ／混合ボリューム部２０は，図１に示すオーディオ帯域チェッカ８および混合ボリューム４，５の回路が配置された部分であり，ここでは重量バランスおよび内部配線の関係上，ヘッドバンド２１の中央に位置するように設けられている。
【００２４】
相手話者からの音声および環境音は，ステレオ入力端子２４からオーディオ入力される。オーディオ入力信号内に環境音が強く含まれる場合には，オーディオ帯域チェッカ／混合ボリューム部２０によりステレオ入力端子２４からのオーディオ入力信号の比率を高くして，またはオーディオ入力信号だけをヘッドホン６Ｌ，６Ｒへ出力し，自分の外部環境音の混合比率を減少させる。オーディオ入力信号内に相手環境音が含まれない場合には，ステレオ入力端子２４からのオーディオ入力信号に，環境音収録用マイク２Ｌ，２Ｒで収録した環境音をある程度混合してヘッドホン６Ｌ，６Ｒへ出力する。これにより，利用者と対話者の会話時に，双方の環境音が同時に聴こえて混乱する状況を防止できる。
【００２５】
図３は，本発明の実施例の回路・配線図である。
図３に示すように，ステレオ入力端子からのオーディオ入力（右）１Ｒと環境音収録用マイク２Ｒからの入力は，混合ボリューム４Ｒを介してヘッドホン（右）６Ｒへ通じている。また，ステレオ入力端子からのオーディオ入力（右）１Ｒは，オーディオ帯域チェッカ８に接続され，オーディオ帯域チェッカ８は，オーディオ入力の周波数が音声帯域のみか否かを判定する。
【００２６】
同様に，ステレオ入力端子からのオーディオ入力（左）１Ｌと環境音収録用マイク２Ｌからの入力は，混合ボリューム４Ｌを介してヘッドホン（左）６Ｌへ通じている。オーディオ入力（左）１Ｌも，同じようにオーディオ帯域チェッカ８に接続され，周波数が判定される。
【００２７】
オーディオ帯域チェッカ８は，混合ボリューム４Ｒ，４Ｌに接続されて，それぞれに対しボリューム制御信号を送出する。
また，環境音収録用マイク２Ｒ，２Ｌからの入力は，それぞれ混合ボリューム５Ｒ，５Ｌを介してオーディオ出力（右／左）７Ｒ，７Ｌへ通じている。会話用マイク３（モノラル）からの入力は，混合ボリューム５Ｒ，５Ｌを介して，オーディオ出力（右／左）７Ｒ，７Ｌへ通じている。
【００２８】
可動式ブーム１０のブーム位置検出器９は，混合ボリューム５Ｒ，５Ｌと接続され，会話用マイクブームの角度に応じたボリューム制御信号を送出する。
図４は，実施例におけるヘッドホン出力に関する利用状態説明図であり，特に図４（Ａ）はオーディオ入力の周波数が音声帯域のみの場合，図４（Ｂ）はオーディオ入力の周波数に音声帯域以外が含まれる場合の例を示している。
【００２９】
ステレオ入力端子２４からのオーディオ入力１Ｒ，１Ｌが，オーディオ帯域チェッカ８により，相手話者の会話音声のみであると判断された場合，図４（Ａ）に示すように，オーディオ入力（右）１Ｒの相手の話声と環境音収録用マイク２Ｒで収録した自己の環境音（右）とを混合ボリューム４Ｒで混合して，ヘッドホン（右）６Ｒへ送出する。同様に，オーディオ入力（左）１Ｌの相手話声と環境音収録用マイク２Ｌで収録した自己の環境音（左）とを混合して，ヘッドホン（左）６Ｌへ送出する。
【００３０】
一方，ステレオ入力端子２４からのオーディオ入力１Ｒ，１Ｌに，オーディオ帯域チェッカ８により，相手の会話音以外の音が含まれると判断された場合，図４（Ｂ）に示すように，オーディオ入力（右）１Ｒの相手の話声と相手の環境音（右）を混合した音をそのままヘッドホン（右）６Ｒへ送出し，環境音収録用マイク２Ｒからの自己の環境音は遮断す

る。同様に，オーディオ入力（左）１Ｌの相手の話声と相手の環境音（左）を混合した音をそのままヘッドホン（左）６Ｌへ送出し，環境音収録用マイク２Ｌからの自己の環境音は遮断する。

【００３１】

図５は，実施例におけるオーディオ出力に関する利用状態説明図であり，図５（Ａ）は会話用マイクブーム（可動式ブーム１０）を口元にセットした状態，図５（Ｂ）は会話用マイクブームを押し下げた状態，図５（Ｃ）は会話用マイクブームを押し上げた状態の場合の例を示している。

【００３２】

会話用マイク３の可動式ブーム１０がセット位置の場合には，図５（Ａ）に示すように，自己の話声と自己の環境音の録音の音量比率については，自己の話声を最大とする音量比率で混合する。したがって，環境音収録用マイク２Ｒからの自己の環境音（右）を少なく，会話用マイク３の自己の話声を多く混合した信号を，オーディオ出力（右）７Ｒとしてステレオ出力端子２５から出力する。同様に，環境音収録用マイク２Ｌからの自己の環境音（左）を少なく，会話用マイク３の自己の話声を多く混合した信号を，オーディオ出力（左）７Ｌとしてステレオ出力端子２５から出力する。

【００３３】

また，会話用マイク３の可動式ブーム１０が押し下げられた状態の場合には，図５（Ｂ）に示すように，自己の話声と自己の環境の録音の音量比率については，可動式ブーム１０の位置が下がるにつれて，自己の環境音が大きな比率で混合されるように混合ボリューム５Ｒ，５Ｌを制御する。したがって，環境音収録用マイク２Ｒからの自己の環境音（右）を多く，会話用マイク３の自己の話声を少なく混合した信号を，オーディオ出力（右）７Ｒとしてステレオ出力端子２５から出力する。同様に，環境音収録用マイク２Ｌからの自己の環境音（左）を多く，会話用マイク３の自己の話声を少なく混合した信号を，オーディオ出力（左）７Ｌとしてステレオ出力端子２５から出力する。

【００３４】

また，会話用マイク３の可動式ブーム１０が最上位のＯＦＦ位置にあると，図５（Ｃ）に示すように，会話用マイク３および環境音収録用マイク２Ｒ，２Ｌのスイッチが共にＯＦＦとなり，会話音も環境音も入力されず，オーディオ出力は行われない。

【００３５】

以上のように，クローズドエアー型のヘッドホン６Ｌ，６Ｒの左右イヤーパッド内，またはヘッドバンド２１に単一指向性マイクを内蔵して，ステレオで外部環境音を収録可能とすると共に，ブーム付き会話用マイク３によって音声会話収録も可能としたステレオヘッドセット装置を実現することができる。遠隔地での音声通信において，音声と共に利用者の周囲の環境音を混合して伝送することにより，会話の臨場感を高めることができる。この装置の利用者は通常は自分の場所で取得した環境音が聴こえるが，対話者から環境音混合音声が聴こえると，自分の側の環境音の比重が下がり，相手側の音響空間を遠隔地で聴取することが可能になる。

【００３６】

【発明の効果】

以上説明したように，本発明では，会話音声と同時に，利用者が存在する場所の雰囲気をステレオ音響として伝えることが可能になる。また，状況に応じて自分のいる場所の環境音を遮断することにより，聴取する場合に複数箇所の外部環境音が混乱して聴こえることのないようにして，臨場感のある遠隔地対話が可能となる。

【図面の簡単な説明】

【図１】本発明の原理説明図である。

【図２】本発明の実施例の外観構成を示す図である。

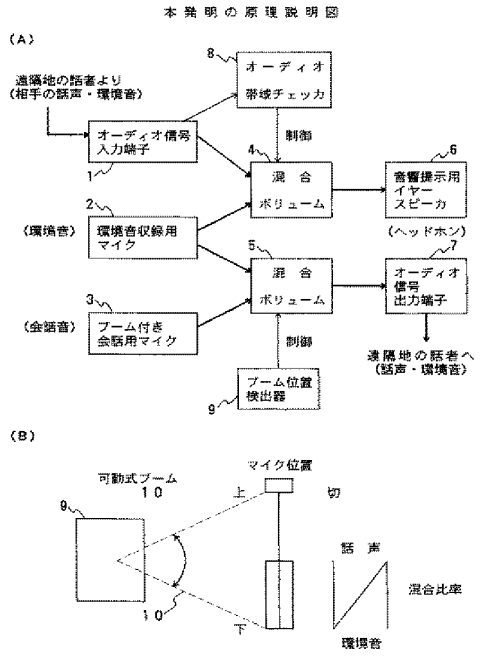
【図３】本発明の実施例の回路・配線図である。

【図４】実施例におけるヘッドホン出力に関する利用状態説明図である。

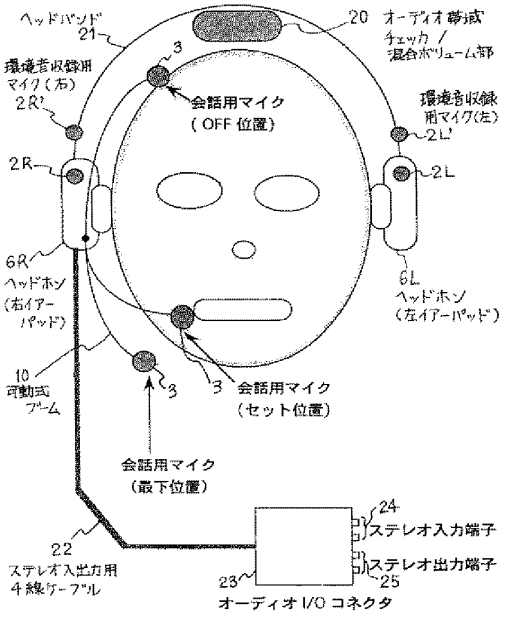
【図５】実施例におけるオーディオ出力に関する利用状態説明図である。

【符号の説明】
１　オーディオ信号入力端子
２　環境音収録用マイク
３　ブーム付き会話用マイク
４　混合ボリューム
５　混合ボリューム
６　音響提示用イヤースピーカ（ヘッドホン）
７　オーディオ信号出力端子
８　オーディオ帯域チェッカ
９　ブーム位置検出器
１０　可動式ブーム

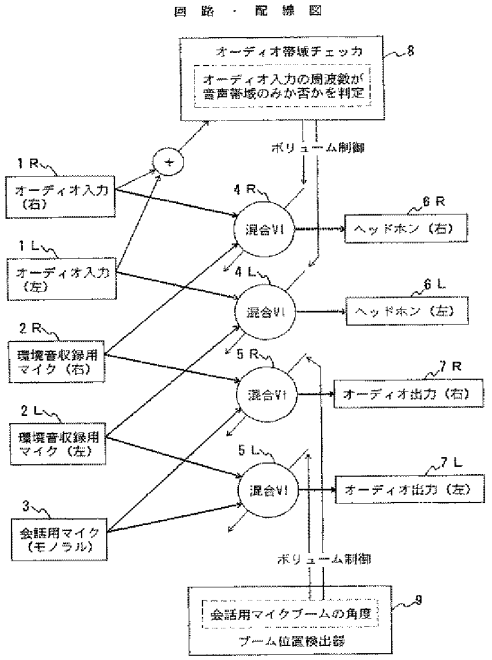
【図1】



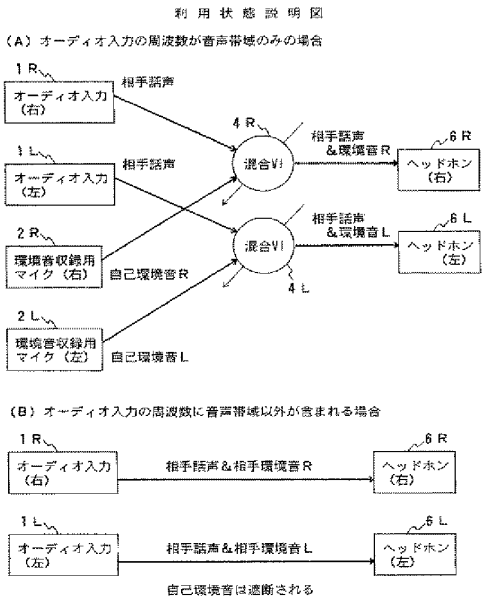
【図2】



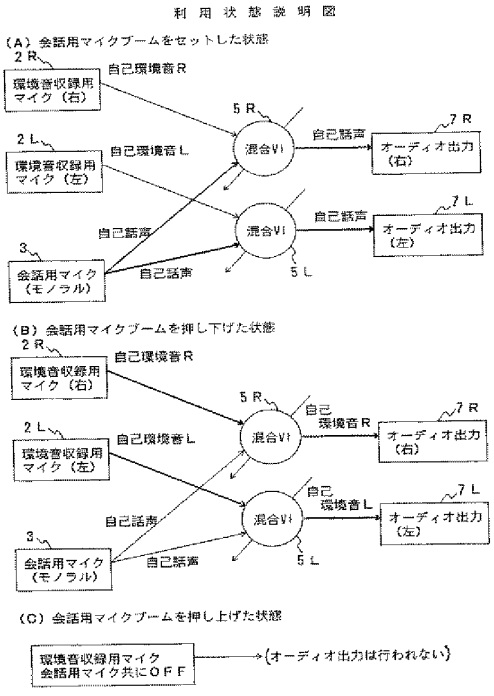
【図3】


回路・配線図
オーディオ帯域チェッカ
オーディオ入力の周波数が音声帯域のみか否かを判定
8
ボリューム制御
1 R
オーディオ入力（右）
1 L
オーディオ入力（左）
2 R
環境音収録用マイク（右）
2 L
環境音収録用マイク（左）
3
会話用マイク（モノラル）
4 R
混合VI
4 L
混合VI
5 R
混合VI
5 L
混合VI
6 R
ヘッドホン（右）
6 L
ヘッドホン（左）
7 R
オーディオ出力（右）
7 L
オーディオ出力（左）
ボリューム制御
会話用マイクブームの角度
ブーム位置検出器
9

【図4】


利用状態説明図
（A）オーディオ入力の周波数が音声帯域のみの場合
1 R
オーディオ入力（右）
相手話声
1 L
オーディオ入力（左）
相手話声
2 R
環境音収録用マイク（右）
自己環境音R
2 L
環境音収録用マイク（左）
自己環境音L
4 R
混合VI
4 L
混合VI
相手話声&環境音R
相手話声&環境音L
6 R
ヘッドホン（右）
6 L
ヘッドホン（左）
（B）オーディオ入力の周波数に音声帯域以外が含まれる場合
1 R
オーディオ入力（右）
相手話声&相手環境音R
6 R
ヘッドホン（右）
1 L
オーディオ入力（左）
相手話声&相手環境音L
6 L
ヘッドホン（左）
自己環境音は遮断される

【図5】


利用状態説明図
（A）会話用マイクブームをセットした状態
2 R
環境音収録用マイク（右）
自己環境音R
2 L
環境音収録用マイク（左）
自己環境音L
3
会話用マイク（モノラル）
自己話声
自己話声
5 R
混合VI
5 L
混合VI
自己話声
自己話声
7 R
オーディオ出力（右）
7 L
オーディオ出力（左）
（B）会話用マイクブームを押し下げた状態
2 R
環境音収録用マイク（右）
自己環境音R
2 L
環境音収録用マイク（左）
自己環境音L
3
会話用マイク（モノラル）
自己話声
自己話声
5 R
混合VI
5 L
混合VI
自己環境音R
自己環境音L
7 R
オーディオ出力（右）
7 L
オーディオ出力（左）
（C）会話用マイクブームを押し上げた状態
環境音収録用マイク
会話用マイク共にOFF
（オーディオ出力は行われない）

(56)参考文献　特開平０４－１１３９４７（ＪＰ，Ａ）
実開昭５９－００１２２０（ＪＰ，Ｕ）

(58)調査した分野(Int.Cl.[7]，ＤＢ名)
H04R　1/10　101
H04R　5/033

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 April 2006 (13.04.2006)

PCT

(10) International Publication Number
WO 2006/037156 A1

(51) International Patent Classification[7]: H04R 29/00, A61B 5/00

(21) International Application Number:
PCT/AU2005/001506

(22) International Filing Date:
30 September 2005 (30.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2004905703 1 October 2004 (01.10.2004) AU

(71) Applicant *(for all designated States except US)*: **HEAR WORKS PTY LTD** [AU/AU]; 384 Albert Street, East Melbourne, VIC 3002 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MEJIA, Jorge, Patricio** [AU/AU]; 126 Greville Street, Chatswood, NSW 2067 (AU). **DILLON, Harvey, Albert** [AU/AU]; 17 Vernon Street, Turramurra, NSW 2074 (AU). **FISHER, Michael, John, Amiel** [AU/AU]; 36 Grasmere Road, Cremorne, New South Wales 2090 (AU).

(74) Agent: **SHELSTON IP**; 60 Margaret Street, Sydney NSW 2000 (AU).

(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, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

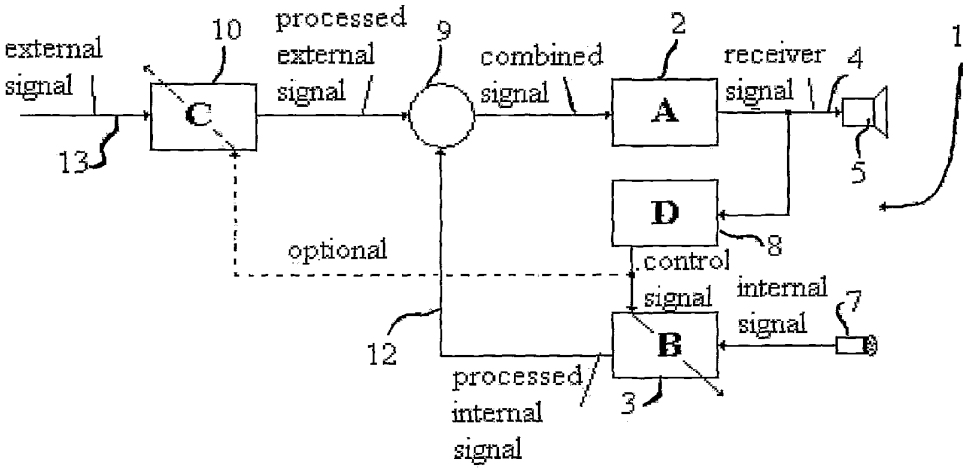
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, 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).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

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: ACOUSTICALLY TRANSPARENT OCCLUSION REDUCTION SYSTEM AND METHOD



(57) Abstract: A system and method that reduces the perceptual effect resulting from ear occlusion is disclosed. It comprises of an electro-acoustic feedback network that produces phase cancelling sounds in the ear, where a receiver and a microphone are located. A novel control mechanism is disclosed that controls the response of the feedback network to minimise distortion in the ear while maintaining a desired frequency response for external signals. Devices producing the external signal include hearing aids, personal sound devices, in ear monitors, communications headsets and hearing protectors. The integration of the present invention with these devices improves the user's perception of their own voice.

WO 2006/037156 A1

Acoustically Transparent Occlusion Reduction System and Method

Field of Invention

The present invention relates to an improved occlusion reduction system for applications such as hearing aids, personal sound devices, in ear monitors, communications headsets and hearing protection devices.

Description of Background Art

An electro-acoustic negative feedback scheme was originally presented by H.F Olson, in 1961 in U.S Patent No 2,983,790. A more comprehensive implementation was later proposed by Bose et al, in 1982 in U.S Patent No 4,494,074 under the title "Feedback Control". Using headphones, Bose proposed the idea of an electro-acoustic feedback in the proximity of the ear canal. The concept was later used by Langberg et al, in 1988 in U.S. Patent No 4,985,925, describing a system functioning as a bilateral transducer drive with a shunt feedback correction network. Later on in 1991, Langberg et al, in U.S Patent No. 5,267,321 entitled "Active Sound Absorption" describes an electro-acoustical feedback system, with the receiver acting as both a diaphragm actuator and motion sensor. In 1996, U.S Patent No. 5,774,565 to Benning et al describes an electro-acoustic feedback subsystem with oscillation prevention circuit in the forward path of the loop. In 2002, U.S Patent Application No. 2003/0012391 A1 to Armstrong et al, entitled "Digital Hearing Aids System" discloses a hearing aid including an occlusion processing subsystem.

The occlusion effect is commonly described as a hollow or echoing like sound of a person's own voice. The occlusion effect results from acoustically sealing or partially sealing the ear, or to a greater effect sealing or partially sealing the ear canal from the external acoustic environment. As a result, the occlusion effect creates discomfort and/or an unnatural sound sensation. This problem is commonly reported to clinicians in the

hearing aid industry as it affects a large number of hearing aid wearers (those with mild low-frequency hearing loss). Until now, there were at least two common schemes to decrease the occlusion effect in hearing aids, either using a vent or by increasing the insertion depth of the earmold into the ear canal. To restore naturalness of a hearing aid wearer's voice, vents of up to 3.5 mm in diameter may be employed. These vents need to be sufficiently large so that the residual sound pressure in the canal due to the occlusion effect is not significant. On the other hand, a sufficiently large vent limits the hearing aid amplification due to oscillations created by positive feedback occurring around the loop defined by an external microphone, amplifier, receiver, and path through the vent back to the external microphone.

The soft tissue in the ear canal is excited by vibration energy propagated by the skull and jaw due to the wearer's voice and this results in an increased sound pressure within the occluded of partially occluded ear relative to an open ear. Another solution is to insert the earmold further into the ear canal to fill the cartilaginous portion of the canal and hence reduce the occlusion effect at its source. However there are a number of practical problems relating to the deep insertion of an earmold, for instance physical discomfort.

While some of the aforementioned patents include methods to alter the response of the electro-acoustic feedback system, none of these patents incorporate methods to compensate for the effect that response alteration within the feedback system has on external signals such as from an external microphone or hearing aid processor. In addition, the aforementioned patents do not provide a mechanism to prevent the receiver from overloading by sensing the level of signal at the optimum point, being the input to the receiver.

Summary of Invention

It is an object of the present invention to provide an improved occlusion reduction system for applications such as hearing aids, personal sound devices, in ear monitors, communications headsets and hearing protection devices.

In accordance with a first aspect of the present invention, there is provided an electro-acoustic system comprising of: an electro-acoustic circuit including a negative-feedback loop comprising of: a microphone for generating an internal signal from an acoustical signal located within or closely coupled to an occluded or partially occluded ear; a second electronic circuit, (B) for modifying the internal signal to produce a processed internal signal; a combiner for combining the processed internal signal with a processed external signal to produce a combined signal; a first electronic circuit (A), for modifying the combined signal to produce an receiver signal; and a receiver for generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear; a third electronic circuit (C) for modifying an external signal such as from an external microphone or hearing aid processor to produce an external processed signal; a fourth electronic circuit (D), for estimating the level of the receiver signal optimally from the receiver signal and producing a control signal; wherein the processing performed by the second electronic circuit (B) preferably is controlled by the control signal.

The negative-feedback loop preferably can include a filter that alters the open loop response so that the real component of the response is large and negative at frequencies where the occlusion effect can be typically the greatest. Consequently reducing all signals introduced into the closed loop within this frequency range.

The third electronic circuit (C) preferably can include filtering that provides compensation for the closed loop response of the negative-feedback loop.

Optionally the filtering provided by the third electronic circuit (C) adapts to compensate for changes in the closed response of the negative-feedback loop.

Optionally the control signal can control the response of the first electronic circuit (A) while controlling the response of the third electronic circuit (C).

The microphone for generating the internal signal optionally can be coupled to the occluded or partially occluded ear by a tube. The receiver optionally can be coupled to the occluded or partially occluded ear by a tube. The microphone and receiver optionally can be combined in a single unit or jointly coupled to the occluded or partially occluded ear by a common tube. At least one of the electronic circuits can be implemented digitally. At least one of the digital electronic circuits can performed signal processing at a sampling rate at least four times that of the bandwidth of the signal.

In accordance with a further aspect of the present invention, there is provided an electro-acoustic system including a negative-feedback loop comprising of: a microphone for generating an internal signal from an acoustical signal located within or closely coupled to an occluded or partially occluded ear; a second electronic circuit (B) for modifying the internal signal to produce a processed internal signal; a combiner for combining the processed internal signal with a processed external signal to produce a combined signal; a first electronic circuit (A) for modifying the combined signal to produce a receiver signal; and a receiver for generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear; a third electronic circuit (C) for modifying an external signal such as from external microphone to produce a processed external signal; a fourth electronic circuit (D) for estimating the level of the receiver signal and producing a control signal from this estimate; wherein the control signal controls the processing performed by the second

electronic circuit (B) and controls the processing performed by the third electronic circuit (C).

In accordance with a further aspect of the present invention, there is provided an electro-acoustic system including a negative-feedback loop comprising of: a microphone for generating an internal signal from an acoustical signal located within or closely coupled to an occluded or partially occluded ear; a second electronic circuit (B) for modifying the internal signal to produce a processed internal signal; a combiner for combining the processed internal signal with a processed external signal to produce a combined signal; a first electronic circuit (A) for modifying the combined signal to produce a receiver signal; and a receiver for generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear; a third electronic circuit (C) for modifying an external signal to produce a processed external signal; a fourth electronic circuit (D) for estimating the level of the receiver signal and producing a control signal from this estimate; wherein the control signal controls the processing performed by first electronic circuit (A) and controls the processing performed by the third electronic circuit (C).

In accordance with a further aspect of the present invention, there is provided a method of providing a negative feedback loop for an electro-acoustic system, the method including the steps of: (a) generating an internal signal representing an acoustical signal located within or closely coupled to an occluded or partially occluded ear; (b) modifying the internal signal to produce a processed internal signal; (c) combining the processed internal signal with a processed external signal to produce a combined signal; (d) modifying the combined signal to produce a receiver signal; (e) generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear; (f) modifying an external signal to produce a processed external

signal; (g) estimating the level of the receiver signal either directly from the receiver signal or from the combined signal and producing a control signal from this estimate; wherein the degree of modification in step (b) is controlled by the control signal. The control signal can be utilised to control the amount of modification occurring in step (f).

In any of the aforementioned aspects of the present invention the fourth electronic circuit (D) optionally produces a control signal using estimates of the signal levels from the internal signal or the processed internal signal and the receiver signal or the combined signal.

Brief Description of the Drawings

Preferred embodiments of the present invention will now be described with reference to the accompanying drawings in which:

Fig. 1 is a block diagram of the occlusion reduction scheme of the preferred embodiment;

Fig. 2 is a schematic diagram of an analog electronic circuit for the electro-acoustic negative feedback loop;

Fig. 3 is a schematic diagram of an analog adaptive pre-compensation electronic circuit;

Fig. 4 is a schematic diagram of an analog level estimator and control signal generation electronic circuit;

Fig. 5 is a schematic diagram of a digital implementation of the occlusion reduction scheme;

Fig. 6 is a schematic diagram of an adaptive IIR filter;

Fig. 7 is a schematic diagram of a level estimator and control signal generator ;

Fig. 8 is a graph of the simulated closed loop response in a Zwislocki coupler;

Fig. 9 is a graph of the measured response of the effective gain reduction inside a real ear;

Fig. 10 is a graph of the measured responses from an external signal to a Zwislocki coupler microphone; and

Fig. 11 is an illustration of the occlusion reduction scheme implemented as an in-the-canal hearing aid device.

Detail Description of the Preferred and Other Embodiments

The preferred embodiment operates to reduce the level of signals generated within an electro-acoustic negative feedback loop, such as signals produced by vibration in the ear canal walls due to bone conduction of a user's voice. The reduction occurs in the low to mid audible frequencies, typically ranging from 80 Hz up to 1 kHz, where the occlusion effect is more predominant and perceptually apparent.

A negative feedback scheme is provided which combines a processed externally generated signal such as from an external microphone or a sound system with a processed internal signal such as from a microphone located within or closely coupled to the occluded or partially occluded ear. The combined signal after optional further processing is applied to a receiver located within or closely coupled to the occluded or partially occluded ear. The level of the signal to be applied to the receiver is optimally estimated either from the signal applied to the receiver or from the combined processed external and processed internal signals. Optionally this signal level can be estimated from signals at other points within the scheme. High signal levels applied to the receiver may produce a distorted output from the receiver. This distortion is reduced by applying active gain reduction in the feedback path in response to estimated high signal levels being present. Optionally low signal levels can be detected so that noise inherent in the negative feedback components such as the internal microphone can be minimised by

applying active gain reduction in the feedback path. Thus, the high and the low signal level thresholds for gain reduction in the loop can depend on the dynamic operational range of the discrete components within the system. In addition, an adaptive equalisation filter is applied to the external signal to compensate for variations of the transfer response of the closed loop.

The preferred embodiment includes of a microphone to sense the sound pressure in the ear. The preferred embodiment also includes a novel design of estimating the level of the signal to be applied to the receiver and reducing the gain in the feedback path of the loop when this level is high. This mechanism effectively improves the robustness of the closed loop system by limiting excessive driving levels being applied to the receiver. A filter within the feedback path of the loop yields the necessary phase and gain around the loop to generate a phase cancelling sound in the ear without creating acoustic feedback. This negative feedback response also causes a sound pressure reduction for external signals thus affecting the response from the external processed signal to the receiver signal. As a result, an adaptive pre-compensation filter is provided. The adaptive pre-compensation filter performs adaptive equalisation to maintain a constant frequency response between the external signal and the receiver signal in response to changes in loop response.

In addition, the negative feedback response also causes a reduction in external sounds transmitted through a vent or leakage, thus minimising both effects.

Referring initially to **Fig. 1**, a schematic diagram 1 illustrates the occlusion reduction scheme of the preferred embodiment. This electronic circuit can be encapsulated in an earmold 100 as discussed hereinafter with reference to **Fig. 11**. The earmold optionally contains tubing for coupling the receiver to the ear. The earmold optionally contains tubing for coupling the internal microphone to the ear. The

microphone and receiver optionally can be combined into a single unit or jointly coupled to the occluded or partially occluded ear by a common tube. Optionally these tubes can be protected from wax blockage using wax guards. The lengths of these tubes are preferably as short as possible to minimise delays around the feedback loop, but can be any length. The earmold can optionally contain an open vent to depressurise the ear thus reducing the sensation of stuffiness in the ear. Optionally, the vent can be fitted with an acoustic damper for compensating for the vent resonance that may affect the closed loop response.

The internal signal from an internal microphone 7 is proportional to the ear canal sound pressure. This internal signal is filtered in a feedback loop shown as first electronic circuit **(A)** 2 and second electronic circuit (**B) 3** to produce the receiver signal 4 output to the receiver 5. The aim is to produce cancellation around the loop, limited to a given low to mid frequency band.

An analog implementation of the occlusion reduction scheme is depicted in **Figs. 2-4** and a digital implementation is shown in **Figs. 5-7**.

In addition, **Fig. 1** shows fourth electronic circuit (D) 8 that estimates the level of the receiver signal to produce a control signal. Optionally the input to the fourth electronic circuit (D) 8 can be obtained from the output of a combiner 9 with appropriate compensation for the effects of the first electronic circuit (A) 2. The response of the negative feedback loop is controlled by the control signal. The estimated level of the receive signal 4 can be compared to a reference level which is not shown in this figure. The control signal reduces the gain in the loop as the level of the receive signal increases above the reference level. The reference level is set to a level to minimise distortion occurring within the loop.

The combined signal within the loop results from a combination 9 of processed internal signal 12 and processed external signal. The processed external signal results from a filtered external signal 13. The pre-compensation filter 10 depicted as the third electronic circuit **(C)** 10 in **Fig. 1**, equalises the magnitude of the transfer function from the external signal 13 to the receiver signal 4 so that it is approximately constant across frequency, assuming a fixed closed loop response. Optionally, the filter in the third electronic circuit (C) 10 is adaptively controlled by fourth electronic circuit (D) 8 so that the magnitude of the transfer function from the external signal 13 to the receiver signal 4 is approximately constant across frequency regardless of changes in the closed loop response.

Optionally the fourth electronic circuit (D) uses estimates of signal levels from the receiver signal or combined signal and from the internal signal or processed internal signal to produce a control signal to control the loop response. This control signal can control the loop response by directly applying gain reduction to the first electronic circuit (A) or to the second electronic circuit (B). The control signal produce from this arrangement can be used in combination with the fixed or adaptive third electronic circuit (C).

Analog Implementation

Fig. 2, depicts an analog implementation 20 of the negative feedback loop of **Fig. 1**. Firstly, the internal signal is applied to a high pass filter 21 and lead-lag filter 22 in order to reduce effects from jaw movements and very low frequency instability in the loop. The equations for the transfer function in relation to the corresponding discrete components are shown below:

$$H_{HP}(j\omega) = \frac{j\omega \cdot R_1 C_2}{1 + j\omega \cdot R_1 \times C_1}$$

if $R_4 = R_5, C_2 = C_3$ then the lead lag transfer function equation is shown below

$$H_{Lead-Lag}(j\omega) = \frac{(R_3 + R_6) + j\omega \bullet C_2 \bullet (2R_3R_6 + R_6R_4 + R_3R_4)}{R_6 + j\omega \bullet C_2 \bullet (R_3R_6 + R_6R_4)}$$

The next filtering stage reduces the dominant transducer resonance within the loop and provides greater open loop gain at frequencies at which the occlusion effect is greatest. This is achieved by using a bi-quadratic filter 24, and placing a complex pair of zeros at the dominant transducer resonance frequency followed by a pair of real poles to provide low frequency emphasis in the open loop response.

In the arrangement of Fig. 2, if we allow $R_8 = R_{10}, C_4 = C_6$, $R_9 = R_{11}, C_5 = C_7$ and $R_{12} = R_{13}, C_8 = C_9$,then the biquadratic transfer function equation can be computed as follows

$$H_{bi-quad}(j\omega) = \frac{(j\omega)^2 C_4R_9C_5 + j\omega\left(C_4 + \frac{C_5R_9}{R_8} - C_5\right) + \frac{1}{R_8}}{(j\omega)^2 C_5R_9C_8 + j\omega\left(\frac{C_5R_9}{R_{12}} + C_8\right) + \frac{1}{R_{12}}}$$

To determine the pole zero placement, the biquadratic transfer function equation can be directly related to a well known second order transfer function equation.

$$K(j\omega) = \frac{(j\omega)^2 + 2\xi_z\omega_z j\omega + \omega_z^2}{(j\omega)^2 + 2\zeta_p\omega_p j\omega + \omega_p^2}$$

Where the ω_p and ω_z are the location of the pole and zero frequencies. Similarly ξ_z and ζ_p are the damping factors for the poles and zeros respectively. Thus, this relationship can be used to position the poles and zeros at the desired frequencies with the desired damping.

With appropriate phase compensation, amplification is added to the loop. The gain amplification is chosen in accordance to a well-known gain and phase margin criteria (e.g. Linear Control Systems Analysis and Designs, John J.D'Azzo, Constantine H.

Houpis, 2nd Edition, McGraw-Hill, 1981). A loop gain of less than or equal to –3 dB is chosen at frequencies likely to produce positive feedback. The filtering arrangement of **Fig. 1** shown in **Fig 2**, produces a gain greater than unity, for an open loop response between 80 Hz and 1 kHz and less than unity at other frequencies, where positive feedback may occur. The frequency band ranging from 80 Hz up to 1 kHz is an appropriated choice as the occlusion effect is subjectively more apparent at these frequencies, as described in literature (e.g. Hearing Aids, Harvey Dillon, Boomerang-Press, 2001).

The bi-quadratic filter 24 is followed up by a voltage controlled variable resistance 25, referred to as **VVR**. This control produces up to – 20 dB of gain around the loop, by controlling the voltage at the gate of the JFET. Thus, the gain across the VVR network is found by observing that the JFET is series with R_7 forms a voltage divider, with its gain given by:

$$Gain_{VVR} = \frac{R_{Q1}}{R_7 + R_{Q1}}$$

Where R_{Q1}, is the variable resistance across drain to source junction.

The VVR is followed by a bufferstage 26, that is subsequently followed by a variable resistor, R_{14}. 27. This latter resistor is used to fined tune the gain around the loop manually. The variable resistor 27 is followed by an amplifier and a combiner 9, combining a processed external signal 28 with the processed internal signal 12. Finally, this buffer functions as a voltage controlled voltage source to the receiver, optionally a **class D** amplifier may be used.

Referring now to **Fig. 3**, there is illustrated the precompensation circuit 10 of **Fig. 1** in more detail. The circuit may be used to pre-compensate an external signal 13 to produce a processed external signal 28. The external signal is pre-compensated with a

fixed high pass filter 32, a variable notch filter 33, and an adjustable gain control 34. Note that the notch and the gain are also controlled with voltage variable resistances, using JFET transistors.

If $C_{13} = C_{14}$,then the transfer function equation for the notch filter, positioned between U6 and U7 is shown below

$$H_{Notch}(j\omega) = \frac{1 + j\omega(2R_x C_{13}) + \left(R_x R_{19} C_{13}^{\ 2}\right)\cdot(j\omega)^2}{1 + j\omega(R_{19} + 2R_x)C_{13} + \left(R_x R_{19} C_{13}^{\ 2}\right)\cdot(j\omega)^2}$$

Where $R_x = R_{20} + R_{Q3} + R_{Q4}$ and $R_{Q3} + R_{Q4}$ is the combined resistance across drain to source junction of the two JFET transistors. Also note that the amplification at U7 is controlled by the source to drain resistance at Q5, and the low pass filter between U5 and U6 is essentially switched on and off by Q2.

The voltage variable resistances shown in **Fig. 2** and **3** are driven by the circuit 40 shown in **Fig. 4**. This circuit functions as a signal level estimator and threshold detector and produces a control voltage 41. In this circuit the receiver signal 4 is compared to a reference voltage 43, shown as v-Ref in the figure, at the negative input of U8 (44). If this voltage does not exceed the reference level, the current source, made up by BJ3 and BJ4 transistor pair pulls down the voltage across the RC tank towards the negative supply voltage, with a time constant equal to the release time. Note that the voltage across the RC tank feeds directly into the gate of the JFET control transistor in **Fig. 2** from U9 buffer (1), resulting in a VVR signal 25. Similarly from U10 and U11 buffer (2) and buffer (3), results is an inverted VVR signal 35, that is directly applied to the gates of all the JFET transistors shown in **Fig. 3**. As the voltage exceeds the threshold reference level, the current source, shown as BJ6 and BJ5 transistor pair, charges the RC tank, towards the base voltage at BJ6, with a time constant set by the attack time. The

attack and release time constants can be approximated by finding the appropriate ratio between resistors R_{31} and R_{30}.

The voltage across the capacitor C_{15} ranges from a voltage at the based of BJ7 , say 0.7V or fully charged to the voltage at the collector of BJ3, say -2.8V or fully discharged, then the change in voltage $\Delta V \approx 3.5V$, from which the current needed to produce the discharging rate can be evaluated as follows:

$$\text{Discharging } I = C_{15} \times \frac{\Delta V}{\text{Re}lease\_Time} = 7uA$$

thus effectively $I_{charging} \approx 11 \times I_{discharging}$

Then $R_{31} = \frac{0.7}{77u} \approx 8.2k$ $\qquad R_{30} = 11 \times R_{31} \approx 91k$

Digital Implementation

A digital implementation of the arrangement of **Fig. 1** typically requires the signal within the loop to be oversampled to reduce the delay introduced into the loop by sampling. The signal processing performed by electronic circuits A, B, C and D and the combiner 9 may be performed digitally. The digital processing optionally is performed by a digital signal processor which is instructed to perform the signal processing. Various techniques are available to those skilled in the art to perform these processes. Examples of these are performing the filtering operations using FIR or IIR filters or modifying the signals in the frequency domain using the short-time Fourier transform techniques. One implementation of the scheme described for **Fig 1** is shown in **Figs. 5-7**.

In **Fig. 5,** after A/D conversion 50, the internal signal is filtered with an IIR filter 51 in order to provide the required gain and phase compensation, with the same criteria applied in the description of **Fig. 2**.

In this digital implementation the level of the receiver signal 4 is directly estimated and the control signal *r(n)* is produced by the level estimator and control signal generator

8 which is shown in greater detailed in **Fig. 7.**. Referring now to **Fig. 7,** the receive signal level is estimated with a digital 1st order switchable time-constant envelope detector. The envelope signal is compared 71 to a threshold value and the maximum of either the ratio of its exceedance of this threshold or unity is produced. This maximum is inverted to produce the control signal. As the level of the envelope signal exceeds the threshold, the control signal shown as *r(n)* decreases at a rate determined by the attack time constant λ_{attack}. As the envelope signal decreases in level, the control signal increases at a rate determined by the release time constant $\lambda_{release}$. Alternative signal level compression strategies employing different envelope detector designs well known to those skilled in the art may be used.

The internal signal, after being filtered by the IIR filter 51, is scaled by the control signal *r(n)* to produce the processed internal signal. The processed external signal is added to the processed internal signal by the combiner 9 to produce the combined signal. The combined signal is scaled 2 to produce the receiver signal 4 which after D/A conversion is applied to the receiver 5.

As in the analog description, the negative feedback loop requires pre-compensation so that the gain response measured from the external signal to the receiver signal is not altered by the closed loop response. In **Fig. 5,** the pre-compensation filter 10 is realised as a cascaded IIR filter, with a structural realisation shown in **Fig. 6** . The first filter stage shown in **Fig. 6** is a high pass filter, comprising of a zero at 0 Hz and a pole at f_{HP}, the coefficient being a function of the sampling rate, Fs. The second stage of the filter shown in **Fig. 6** is adaptive, the locations of poles f_p and zeros f_z of the digital filter are varied in accordance with the output of a mapping function *p(n)* 61. Alternatively the damping factors for the poles and zeros can be controlled by two independent mapping functions derived from the control signal *r(n)*. **Fig. 5** shows *p(n)*

as a single output from a polynomial mapping function with input $r(n)$. The polynomial's coefficients are found using an autoregressive analysis of the required compensating response as a function of the control signal $r(n)$.

Any combination of analog and digital electronic circuits is possible in addition to the all analogue and the all digital implementations described with the appropriate conversions between analogue and digital formats. In particular, electronic circuits B, C and D may be implemented digitally with the combiner and electronic circuit A implemented in analog circuitry. Furthermore, the processed internal and processed external signals may be combined while both in a 1-bit format and applied directly to the receiver without electronic circuit A performing any function or with it simply providing the 1-bit drive current for the receiver.

Referring now to **Fig. 8** the simulated closed loop response in a Zwislocki coupler is shown. The scheme yields approximately 16 dB of occlusion reduction at 300 Hz, and 18 dB at 100 Hz. The occlusion reduction response is a measure of the level difference between the sound level in the coupler with the occlusion reduction system active and the sound level in the coupler without any occlusion reduction. Referring to **Fig. 9** occlusion reduction in a real ear is shown. This was calculated by taking the linear average of the energy in 1/12 octave bands recorded using an in-the-ear-canal microphone while a subject talked for 2 minutes. This signal was recorded while the feedback loop gain is maximum, and then again while the loop was open.

Referring now to **Fig. 10**, three responses measured in a Zwislocki coupler are shown. These are the transfer functions from an external signal to the coupler microphone for the compensated 93 and un-compensated responses 92. Also shown is the natural receiver response 94 employing a voltage controlled voltage source to the receiver. Note that the compensated response is not adversely affected by the occlusion

reduction scheme compared to that of the natural receiver response 94. Also similar responses are derived for minimally open vented cavities with the appropriate changes to the mappings, filtering and gain equalisation in electronic circuits A, B, and C.

Referring now to **Fig. 11**, there is shown one form of actual implementation of the present invention as an in-the-canal hearing aid device 100 placed in a user's ear 101 that substantially eliminates the discomfort associated with the occlusion effect. The device 100 includes a sealed or optionally vented 102 earmold that occludes or partially occludes the ear canal. The earmold encapsulates a hearing aid electronic circuit comprising of: a hearing aid processor 103 directly connected to an external microphone 104 and powered by small battery 105. The hearing aid processor includes occlusion reduction electronic circuits A, B, C, D and the combiner. A separate receiver 5 and an internal microphone 7 are also interconnected. These transducer elements are acoustically coupled to the inside of the ear canal with short tubes 108, 109; and protected from ear wax with wax-guarding devices 110, 111.

The forgoing describes preferred embodiments of the present invention. Modifications, obvious to those skilled in the art, can be made thereto without departing from the scope of the invention.

CLAIMS

1. An electro-acoustic system including:

an electro-acoustic circuit including a negative-feedback loop comprising of:

a microphone for generating an internal signal from an acoustical signal located within or closely coupled to an occluded or partially occluded ear;

a second electronic circuit (B) for modifying the internal signal to produce a processed internal signal;

a combiner for combining the processed internal signal with a processed external signal to produce a combined signal;

a first electronic circuit (A) for modifying the combined signal to produce a receiver signal;

and

a receiver for generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear;

a third electronic circuit (C) for modifying an external signal to produce a processed external signal;

a fourth electronic circuit (D) for estimating the level of the receiver signal either directly from the receiver signal or from the combined signal and producing a control signal from this estimate;

wherein the processing performed by the second electronic circuit (B) is controlled by the control signal.

2. A system as claimed in claim 1 wherein the negative-feedback loop includes filtering that alters the open loop response so that the real component of the response is large and negative at frequencies where the occlusion effect is most apparent.

3. A system as claimed in claim 2 wherein the third electronic circuit (C) includes filtering that provides compensation for the closed response of the negative-feedback loop.

4. A system as claimed in claim 3 wherein the filtering provided by the third electronic circuit (C) is adaptive and controlled by the control signal.

5. A system as claimed in any previous claim wherein the fourth electronic circuit (D) estimates the level of the internal signal or the level of the processed internal signal and uses this estimate with the estimate of the receiver signal to produce a control signal.

6. A system as claimed in any previous claim wherein the gain of the second electronic circuit (B) is controlled by the control signal.

7. A system as claimed in any previous claim wherein the ear is occluded or partially occluded by an earmold.

8. A system as claimed in claim 7 wherein the earmold contains a vent.

9. A system as claimed in claim 8 wherein the vent contains an acoustic damper

10. A system as claimed in any previous claim wherein the microphone for generating the internal signal is coupled to the occluded or partially occluded ear by a tube.

11. A system as claimed in any previous claim wherein the receiver is coupled to the occluded or partially occluded ear by a tube.

12. A system as claimed in any previous claim wherein the microphone and the receiver are jointly coupled to the occluded or partially occluded ear by a common tube.

13. A system as claimed in any previous claim wherein at least one of the electronic circuits is digital.

14. A system as claimed in any previous claim wherein at least one of the digital electronic circuits uses a sampling rate at least four times that of the bandwidth of the signal.

15. A system as claimed in claims 13 or claim 14 wherein the first electronic circuit (A) includes a digital to analogue converter.

16. A system as claimed in claim 13, claim 14 or claim 15 wherein the second electronic circuit (B) includes an analogue to digital converter.

17. A system as claimed in claim 16 wherein the second electronic circuit (B) includes a digital to analogue converter.

18. A system as claimed in claim 14 wherein at least one of the input signals to the combiner is a binary signal.

19. An electro-acoustic system including a negative-feedback loop comprising of:

a microphone for generating an internal signal from an acoustical signal located within or closely coupled to an occluded or partially occluded ear;

a second electronic circuit (B) for modifying the internal signal to produce a processed internal signal;

a combiner for combining the processed internal signal with a processed external signal to produce a combined signal;

a first electronic circuit (A) for modifying the combined signal to produce a receiver signal;

and

a receiver for generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear;

a third electronic circuit (C) for modifying an external signal such as from external microphone to produce a processed external signal;

a fourth electronic circuit (D) for estimating the level of the receiver signal and producing a control signal from this estimate;

wherein the control signal controls the processing performed by the second electronic circuit (B) and controls the processing performed by the third electronic circuit (C).

20. An electro-acoustic system including a negative-feedback loop comprising of:

a microphone for generating an internal signal from an acoustical signal located within or closely coupled to an occluded or partially occluded ear;

a second electronic circuit (B) for modifying the internal signal to produce a processed internal signal;

a combiner for combining the processed internal signal with a processed external signal to produce a combined signal;

a first electronic circuit (A) for modifying the combined signal to produce a receiver signal;

and

a receiver for generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear;

a third electronic circuit (C) for modifying an external signal to produce a processed external signal;

a fourth electronic circuit (D) for estimating the level of the receiver signal and producing a control signal from this estimate;

wherein the control signal controls the processing performed by first electronic circuit (A) and controls the processing performed by the third electronic circuit (C).

21. A method of providing a negative feedback loop for an electro-acoustic system, the method including the steps of:

(a) generating an internal signal representing an acoustical signal located within or closely coupled to an occluded or partially occluded ear;

(b) modifying the internal signal to produce a processed internal signal;

(c) combining the processed internal signal with a processed external signal to produce a combined signal;

(d) modifying the combined signal to produce a receiver signal;

(e) generating an acoustical signal from the receiver signal at a location within or closely coupled to the occluded or partially occluded ear;

(f) modifying an external signal to produce a processed external signal;

(g) estimating the level of the receiver signal either directly from the receiver signal or from the combined signal and producing a control signal from this estimate;

wherein the degree of modification in step (b) is controlled by the control signal.

22. A method as claimed in claim 21 wherein the control signal is utilised to control the amount of modification occurring in step (f).

23. A method substantially as herein described with reference to any one of the embodiments of the invention illustrated in the accompanying drawings and/or examples.

24. A system substantially as herein described with reference to any one of the embodiments of the invention illustrated in the accompanying drawings and/or examples.

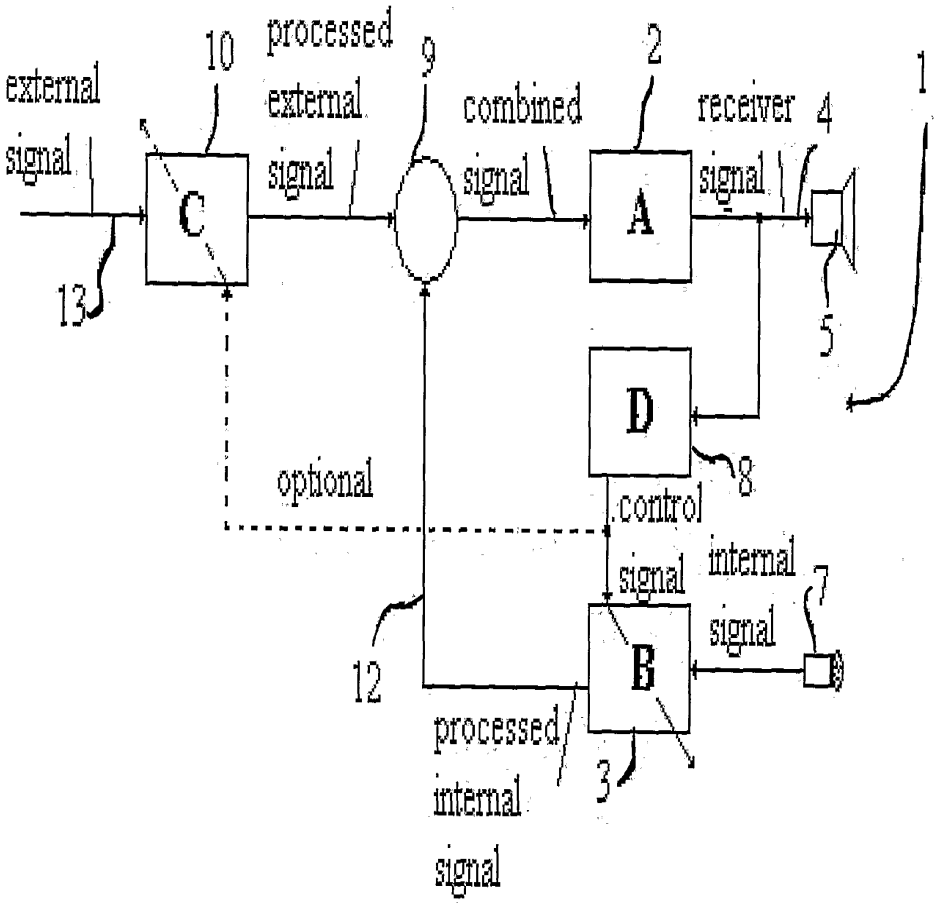

external
signal
10
processed
external
signal
9
combined
signal
2
receiver
signal
4
1
C
A
13
5
D
8
optional
control
signal
internal
signal
7
12
B
processed
internal
signal
3

Fig. 1

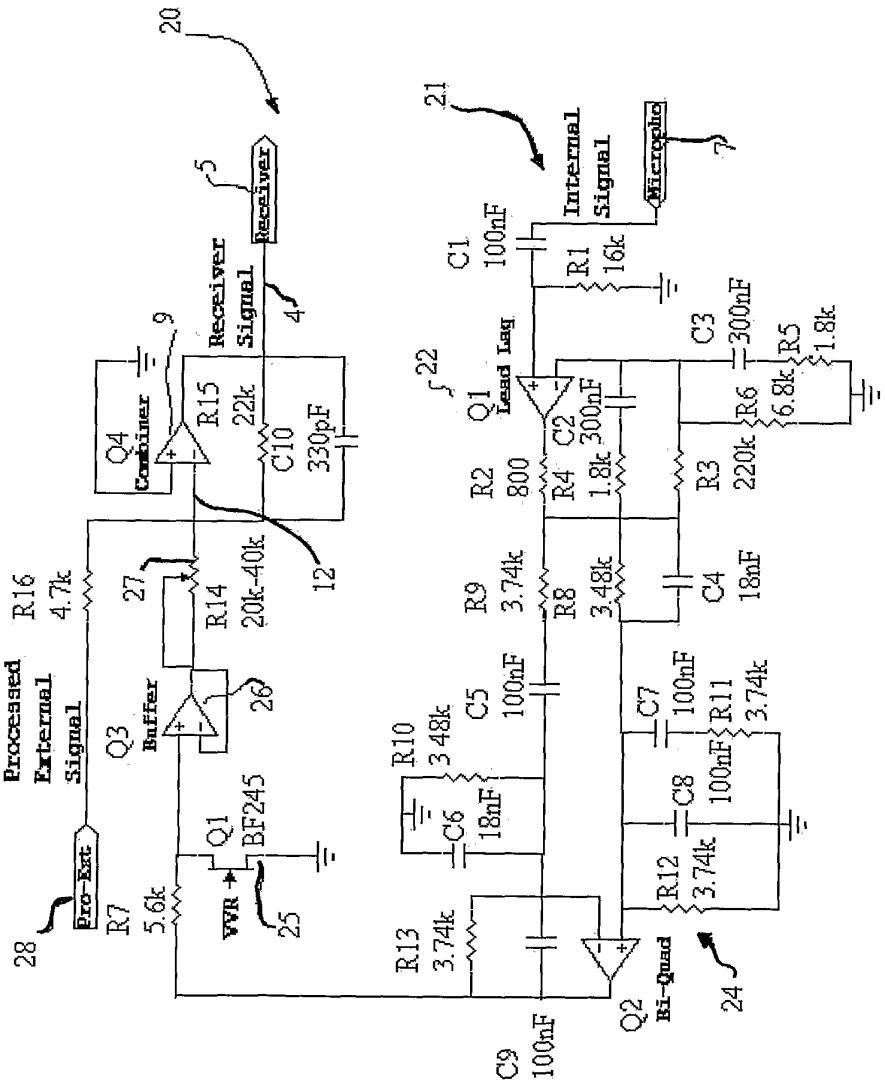

20
21
22
Receiver
Receiver Signal
Internal Signal
Microphone
Processed External Signal
Pre-Ext
Buffer
Lead Lag
Bi-Quad
Container
VVR
BF245
R1 16k
R2 800
R3 220k
R4 1.8k
R5 1.8k
R6 6.8k
R7 5.6k
R8 3.48k
R9 3.74k
R10 3.48k
R11 3.74k
R12 3.74k
R13 3.74k
R14 20k-40k
R15 22k
R16 4.7k
C1 100nF
C2 300nF
C3 300nF
C4 18nF
C5 100nF
C6 18nF
C7 100nF
C8 100nF
C9 100nF
C10 330pF
Q1
Q2
Q3
Q4

Fig. 2

13
External
External Signal
C11
100nF
32
Q5
buffer (1)
R18
22k
33
Q6
Buffer (2)
R19
47k
C13
6.8nF
C14
6.8nF
34
U7
Gain
Processed External Signal
28
Pro-Ext
R17
16k
C12
18nF
Q2
BF245
R20
560
Q3
BF245
Q4
BF245
R22
1.8k
R21
680
Q5
BF245
10
Inv (VVR)
35

Fig. 3

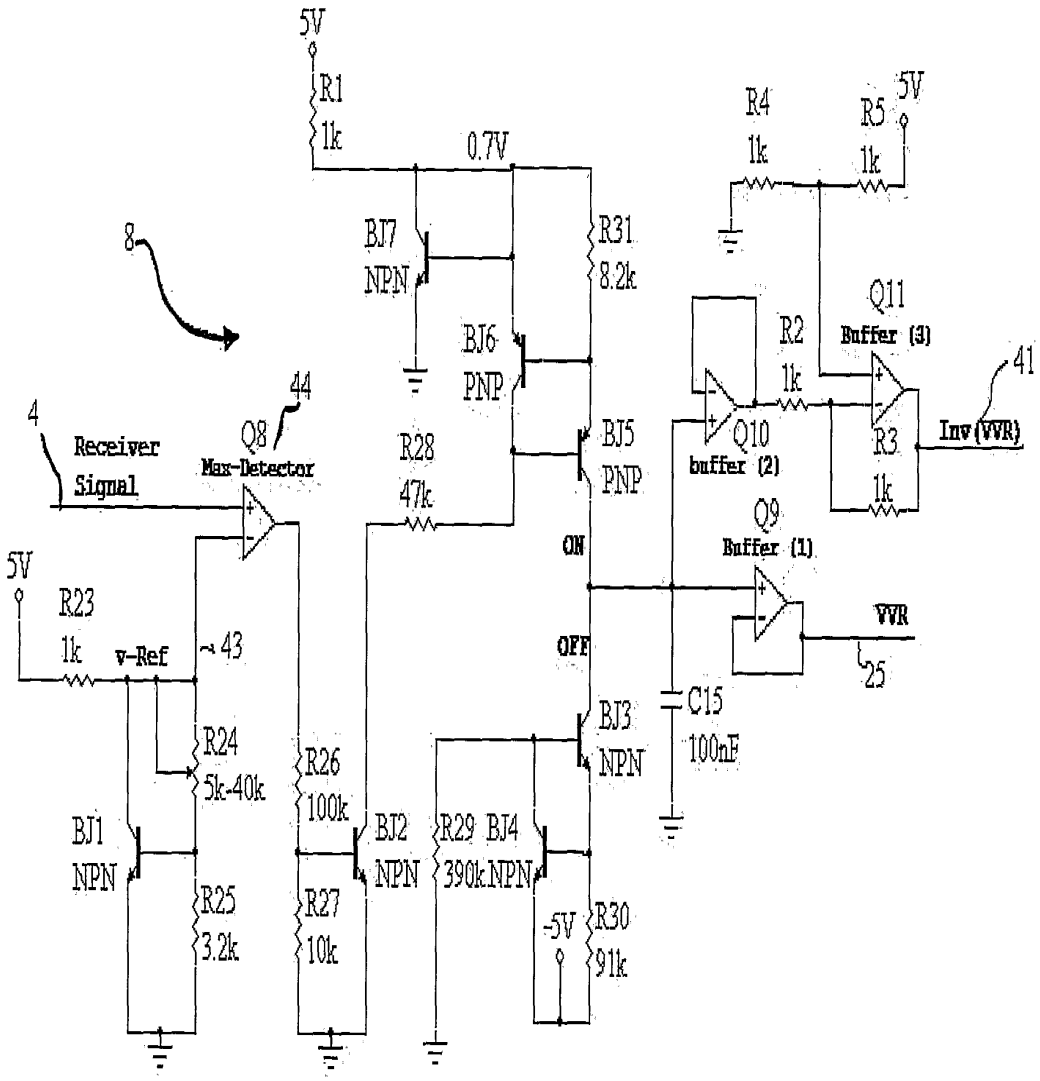

5V
R1
1k
0.7V
R4
1k
R5
1k
5V
8
BJ7
NPN
R31
8.2k
Q11
Buffer (3)
BJ6
PNP
R2
1k
41
4
44
Receiver
Signal
Q8
Max-Detector
R28
47k
BJ5
PNP
Q10
buffer (2)
R3
1k
Inv (VVR)
ON
Q9
Buffer (1)
5V
R23
1k
v-Ref
43
VVR
OFF
25
C15
100nF
BJ3
NPN
R24
5k-40k
R26
100k
BJ1
NPN
BJ2
NPN
R29
390k
BJ4
NPN
R25
3.2k
R27
10k
-5V
R30
91k

Fig. 4

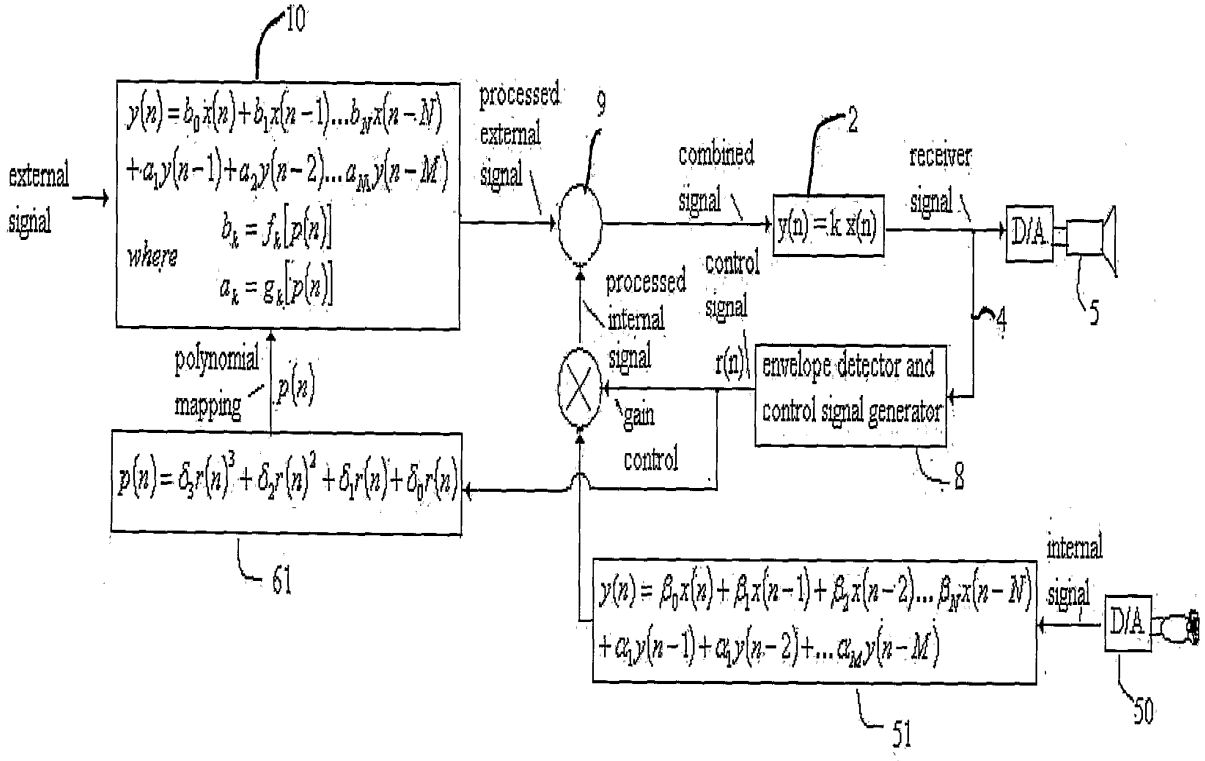

10
external signal
$y(n)=b_0x(n)+b_1x(n-1)...b_Nx(n-N)$
$+a_1y(n-1)+a_2y(n-2)...a_My(n-M)$
where
$b_k=f_k[p(n)]$
$a_k=g_k[p(n)]$
processed external signal
9
combined signal
2
y(n)=k x(n)
receiver signal
D/A
5
4
processed internal signal
control signal
r(n)
envelope detector and control signal generator
8
gain control
polynomial mapping
p(n)
$p(n)=\delta_3r(n)^3+\delta_2r(n)^2+\delta_1r(n)+\delta_0r(n)$
61
internal signal
$y(n)=\beta_0x(n)+\beta_1x(n-1)+\beta_2x(n-2)...\beta_Nx(n-N)$
$+\alpha_1y(n-1)+\alpha_1y(n-2)+...\alpha_My(n-M)$
D/A
50
51

Fig. 5

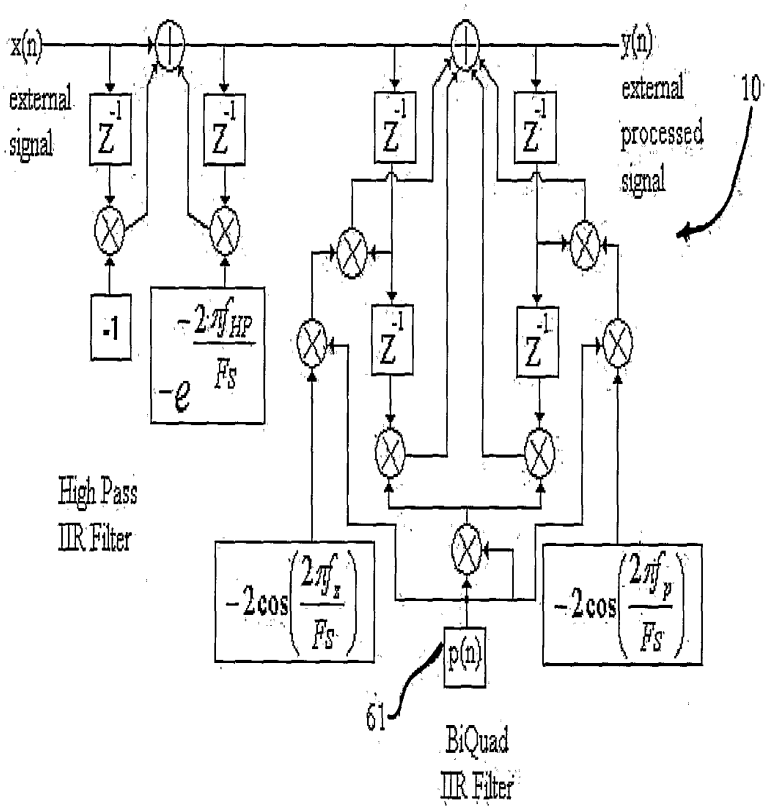

x(n)
external
signal
y(n)
external
processed
signal
10
-1
$-2\pi f_{HP}/Fs$
$-e$
High Pass
IIR Filter
$-2\cos(2\pi f_z/Fs)$
$-2\cos(2\pi f_p/Fs)$
p(n)
61
BiQuad
IIR Filter

Fig. 6

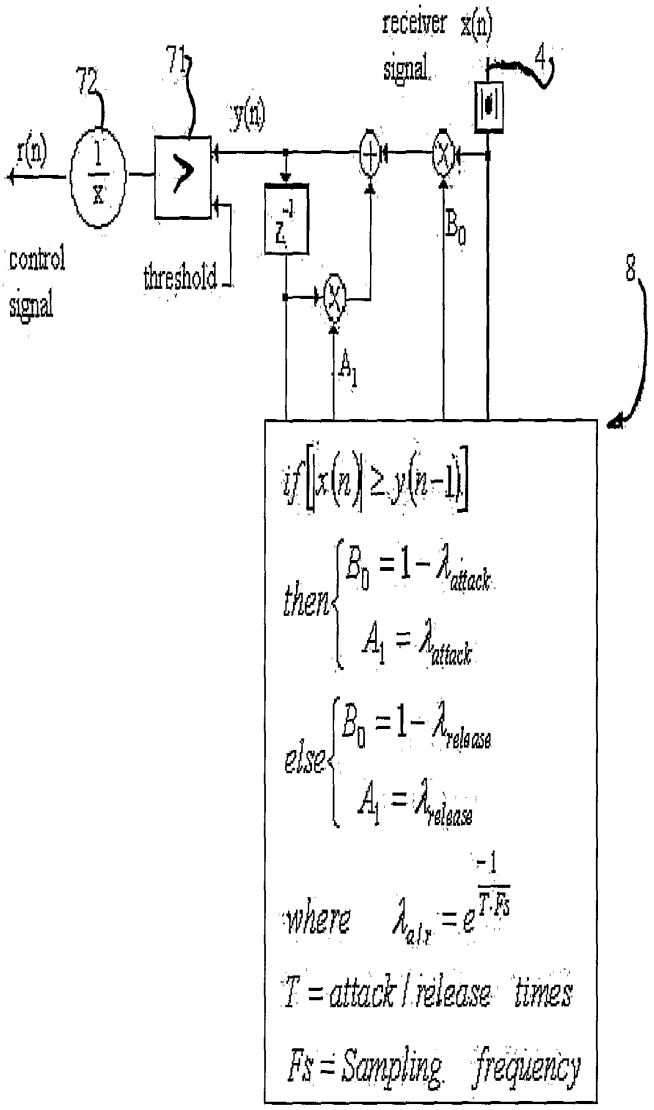

receiver signal
x(n)
4
72
71
y(n)
r(n)
1/x
control signal
threshold
z^{-1}
B_0
A_1
8
$if\left[x(n) \geq y(n-1)\right]$
$then\begin{cases} B_0 = 1-\lambda_{attack} \\ A_1 = \lambda_{attack} \end{cases}$
$else\begin{cases} B_0 = 1-\lambda_{release} \\ A_1 = \lambda_{release} \end{cases}$
$where \quad \lambda_{a/r} = e^{\frac{-1}{T \cdot Fs}}$
$T = attack / release \ \ times$
$Fs = Sampling \ \ frequency$

Fig. 7