fon (440) dafür eingerichtet ist, in dem ersten Gehörgang (303) verwendet zu werden, und dieses den verdeckten Ton in dem ersten Gehörgang (303) empfängt, wobei der wenigstens eine Abschnitt der ersten Ohrform zur Anordnung in dem ersten Gehörgang (303) eingerichtet ist, und ferner mit einer zweiten ohrseitigen Vorrichtung (210B), die beinhaltet:

einen Drahtlosempfänger, um ein erstes Drahtlossignal zu empfangen, das einen Fernton darstellt;
einen zweiten Prozessor (260), der mit dem Drahtlosempfänger verbunden ist, um das erste Drahtlossignal in den Fernton umzuwandeln;
einen Lautsprecher (470), der mit dem zweiten Prozessor verbunden ist, wobei der Lautsprecher zur Verwendung in dem zweiten Gehörgang (303) eingerichtet ist, um den Fernton an den zweiten Gehörgang (303) zu übertragen; und
einer zweiten Ohrform, die wenigstens den Lautsprecher einschließt, wobei wenigstens ein Teil der zweiten Ohrform zur Anordnung in dem zweiten Gehörgang (303) eingerichtet ist.

17. System nach Anspruch 16, bei dem die zweite ohrseitige Vorrichtung (210B) ferner ein zweites Mikrofon (420) aufweist, das mit dem zweiten Prozessor (460) verbunden ist, wobei das zweite Mikrofon (420) dafür eingerichtet ist, einen Umgebungston von außerhalb des zweiten Gehörgangs (303) zu empfangen, und bei dem:

der zweite Prozessor (460) den Umgebungston weiterverarbeitet; und
der Lautsprecher (470) dafür eingerichtet ist, den Fernton und den verarbeiteten Umgebungston an den zweiten Gehörgang (303) zu senden.

18. System nach einem der Ansprüche 16 oder 17, bei dem die erste ohrseitige Vorrichtung (110, 210A) die erste sprachbetriebene Austauschschaltung (VOX) (430) mit dem ersten Prozessor (460) verbunden aufweist, und wobei die zweite ohrseitige Vorrichtung (210B) ferner eine zweite VOX-Schaltung (430) aufweist, die mit dem zweiten Prozessor verbunden ist.

19. System nach einem der Ansprüche 16 bis 18, bei dem der erste Prozessor und der zweite Prozessor jeweils ein Spracherkennungsmodul aufweisen.

20. System nach einem der Anspruch 16 bis 19, bei dem die ersten und zweiten ohrseitigen Vorrichtungen jeweils die BTE-vorrichtungen darstellen.

21. System nach einem der Ansprüche 16 bis 19, bei dem die ersten und zweiten ohrseitigen Vorrichtungen jeweils die ITE-Vorrichtungen darstellen.

22. System nach Anspruch 21, bei dem die ersten und zweiten ohrseitigen Vorrichtungen jeweils eine In-Gehörgangs-Vorrichtung (ITC) (310B) darstellen.

23. System nach Anspruch 22, bei dem die ersten und zweiten ohrseitigen Vorrichtungen jeweils Vollständig-In-Gehörgangs-Vorrichtungen (CIC) (310C) darstellen.

24. System nach einem der Ansprüche 16 bis 23, das ferner eine entfernte Vorrichtung (190, 290) aufweist, die kommunikativ mit den ersten und zweiten ohrseitigen Vorrichtungen verbunden ist.

25. System nach Anspruch 24, bei dem die entfernte Vorrichtung (190, 290) wenigstens einen Computer, einen Personal Digital Assistant (PDA), ein Mobiltelefon, ein Walkie Talkie und/oder einen Sprachübersetzer aufweist.

26. Verfahren, das beinhaltet:

Bestimmen eines verdeckten Tons unter Verwendung eines ersten Mikrofons, das zur Verwendung in einem Gehörgang (303) eingerichtet ist;
Bestimmen eines eingehenden Funksignals, das einen Fernton von einer entfernten Vorrichtung (190, 290) darstellt;
Starten einer sprachbetriebenen Austauscheinrichtung (VOX), die in einer ohrseitigen Kommunikationsvorrichtung eingeschlossen ist, wenn der verdeckte Ton und/oder das eingehende Funksignal bestimmt wird, wenn die VOX nicht bereits an ist, wobei die ohrseitige Kommunikationsvorrichtung eine In-Ohr-Vorrichtung (ITE) oder eine Hinter-Ohr-Vorrichtung (BTE) darstellt;
Einblenden des verdeckten Tons und Ausblenden des Ferntons, wenn der verdeckte Ton bestimmt wird, während das eingehende Funksignal nicht bestimmt wird; und
Einblenden des Ferntons und Ausblenden des verdeckten Tons, wenn das eingehende Funksignal bestimmt wird, während der verdeckte Ton nicht bestimmt wird.

27. Verfahren nach Anspruch 26, bei dem das Bestimmen des verdeckten Tons, das Vergleichen einer Amplitude des verdeckten Tons mit einem Schwellwert für einen verdeckten Ton beinhaltet, wobei der verdeckte Ton bestimmt wird, wenn die Amplitude des verdeckten Tones den Schwellwert für den verdeckten Ton überschreitet.

Exhibit 1004
Page 0370 of 1001

28. Verfahren nach einem der Ansprüche 26 oder 27, bei dem das Bestimmen des eingehenden Funksignals das Vergleichen einer Amplitude des Ferntons mit einem Ferntonschwellwert beinhaltet, wobei das eingehende Funksignal bestimmt wird, wenn die Amplitude des Ferntons den Ferntonschwellwert überschreitet.

29. Verfahren nach einem der Ansprüche 26 bis 28, das ferner das Bestimmen eines Umgebungstons unter Verwendung eines zweiten Mikrofons beinhaltet, und wobei das Starten der VOX das Starten der VOX beinhaltet, wenn wenigstens der verdeckte Ton, das eingehende Funksignal und/oder der Umgebungston bestimmt wird, während die VOX noch nicht bereits an ist.

30. Verfahren nach Anspruch 29, das ferner das Ausblenden des verdeckten Tons, das Einblenden des Ferntons und das Ausblenden des Umgebungstons beinhaltet, wenn das eingehende Funksignal bestimmt wird, während der verdeckte Ton nicht bestimmt wird.

31. Verfahren nach Anspruch 29, das ferner das Ausblenden des verdeckten Tons, das Einblenden des Ferntons und das Einblenden des Umgebungstons beinhaltet, wenn das eingehende Funksignal bestimmt wird, während der verdeckte Ton nicht bestimmt wird.

32. Verfahren nach Anspruch 29, das ferner das Einblenden des verdeckten Tones, das Ausblenden des Ferntons und das Ausblenden des Umgebungstons beinhaltet, wenn der verdeckte Ton bestimmt wird, während das eingehende Funksignal nicht bestimmt wird.

Revendications

1. Un système adapté en vue d'une utilisation dans au moins une parmi une première oreille (101A) ayant un premier conduit auditif (303) et une seconde oreille (101B) ayant un second conduit auditif (303), comprenant:

   un premier dispositif d'oreille (110, 210A) configuré pour être un parmi un dispositif placé à l'intérieur de l'oreille (ITE) et un dispositif contour d'oreille (BTE), le premier dispositif d'oreille comprenant :

   un premier microphone (440) adapté en vue d'une utilisation dans un parmi le premier conduit auditif et le second conduit auditif, et adapté pour recevoir un son occlus (305) à l'intérieur de celui parmi le premier conduit auditif (303) et le second conduit auditif (303) ; et
   un récepteur sans fil, connecté à un premier processeur (460), pour recevoir un premier signal sans fil représentant un son distant ;

   dans lequel le premier processeur (460) est connecté au premier microphone (440) et est adapté pour traiter le son occlus (305) et le son distant; le premier dispositif d'oreille comprenant en outre :

   un haut-parleur adapté en vue d'une utilisation dans celui parmi le premier conduit auditif (303) et le second conduit auditif (303) et adapté pour transmettre le son distant traité à celui parmi le premier conduit auditif (303) et le second conduit auditif (303) ; un émetteur sans fil, connecté au premier processeur (460) et adapté pour recevoir le son occlus traité du premier processeur (460) et émettre un second signal sans fil représentant le son occlus ;
   un premier embout auriculaire logeant au moins le premier microphone (440) et le haut-parleur (470), au moins une portion du premier embout auriculaire étant configurée en vue d'être placée dans celui parmi le premier conduit auditif (303) et le second conduit auditif (303) ; et
   un premier circuit d'échange à commande vocale (VOX) (430) connecté au premier processeur (460) ;
   le premier circuit VOX (430) comprenant :

   un module de contrôle de son occlus (441) adapté pour laisser passer le son occlus et pour bloquer le son distant lorsque le son occlus est détecté alors que le son distant n'est pas détecté ; et
   un module de contrôle de son distant (451) adapté pour laisser passer le son distant et pour bloquer le son occlus lorsque le son distant entrant est détecté alors que le son occlus n'est pas détecté.

2. Système selon la revendication 1, dans lequel le premier embout auriculaire loge le premier microphone (440), le premier processeur et l'émetteur sans fil, et dans lequel le premier embout auriculaire est configuré pour être utilisé comme dispositif ITE (310 A).

3. Système selon la revendications 1, dans lequel le premier dispositif d'oreille (110, 210A) comprend un module ITE comprenant le premier microphone (440) et un module BTE fixé au module ITE, le module BTE comprenant le premier processeur et l'émetteur sans fil.

4. Système selon la revendication 1, dans lequel le premier microphone (440) et le haut-parleur (470) sont

Exhibit 1004
Page 0371 of 1001

mis en oeuvre sous la forme d'un dispositif commun recevant le son occlus et émettant le son distant traité.

5. Système selon la revendication 4, comprenant en outre un second microphone (420) connecté au premier processeur (460), le second microphone (420) étant adapté pour recevoir un son ambiant de l'extérieur de celui parmi le premier conduit auditif (303) et le second conduit auditif (303), et dans lequel :

le premier processeur (460) traite en outre le son ambiant ; et
le haut-parleur (470) est configuré pour transmettre les sons distant et ambiant traités à celui parmi le premier conduit auditif (303) et le second conduit auditif (303).

6. Système selon la revendication 5, dans lequel le premier circuit VOX (430) comprend un module de contrôle de son ambiant (421), couplé au premier microphone (420), pour occulter ou atténuer sensiblement le son ambiant.

7. Système selon l'une quelconque des revendications précédentes, dans lequel le premier processeur (460) comprend un module de reconnaissance vocale.

8. Système selon l'une quelconque des revendications 4 à 7, dans lequel le premier embout auriculaire loge le premier microphone (440), le premier processeur (460), l'émetteur sans fil, le récepteur sans fil, et le haut-parleur (470), et dans lequel le premier embout auriculaire est configuré pour être utilisé comme dispositif ITE.

9. Système selon la revendication 8, dans lequel le premier embout auriculaire est configuré pour être utilisé comme dispositif intra-conduit (ITC) (310B).

10. Système selon la revendication 9, dans lequel le premier embout auriculaire est configuré pour être utilisé comme dispositif complètement intra-conduit (CIC) (310C).

11. Système selon l'une quelconque des revendications 4 à 7, dans lequel le premier dispositif d'oreille (110, 210A) comprend un module ITE comprenant le premier microphone (440) et le haut-parleur (470) et un module BTE fixé au module ITE et comprenant le second microphone (420), le premier processeur (460), l'émetteur sans fil et le récepteur sans fil.

12. Système selon la revendication 11, dans lequel le module BTE comprend le premier circuit d'échange à commande vocale (VOX) (430).

13. Système selon la revendication 11, dans lequel le module BTE comprend une batterie rechargeable (480).

14. Système selon l'une quelconque des revendications 7 à 13, comprenant en outre un dispositif distant (190, 290) couplé en communication à l'émetteur sans fil et au récepteur sans fil via une liaison de télémesure (195, 295) assurant une communication bidirectionnelle simultanée entre le premier dispositif d'oreille (110, 210A) et le dispositif distant (190, 290).

15. Système selon la revendication 14, dans lequel le dispositif distant (190, 290) comprend au moins un parmi un ordinateur, un assistant numérique personnel (PDA), un téléphone cellulaire, un talkie-walkie, ou un traducteur de langue.

16. Système selon la revendication 1, dans lequel le premier microphone (440) est adapté en vue d'une utilisation dans le premier conduit auditif (303) et reçoit le son occlus à l'intérieur du premier conduit auditif (303), ans lequel au moins la portion du premier embout auriculaire est configurée en vue d'être placée dans le premier conduit auditif (303), et comprenant en outre un second dispositif d'oreille (210B) comprenant :

un récepteur sans fil pour recevoir un premier signal sans fil représentant un son distant ;
un second processeur (260), connecté au récepteur sans fil, pour convertir le premier signal sans fil dans le son distant ;
un haut-parleur (470) connecté au second processeur, le haut-parleur étant adapté en vue d'une utilisation dans le second conduit auditif (303) pour transmettre le son distant au second conduit auditif (303) ; et
un second embout auriculaire logeant au moins le haut-parleur, au moins une portion du second embout auriculaire étant configurée en vue d'être placée dans le second conduit auditif (303).

17. Système selon la revendication 16, dans lequel le second dispositif d'oreille (210B) comprend en outre un second microphone (420) connecté au second processeur (460), le second microphone (420) étant adapté pour recevoir un son ambiant de l'extérieur du second conduit auditif (303), et dans lequel :

le second processeur (460) traite en outre le son ambiant ; et le haut-parleur (470) est configuré pour transmettre le son distant et le son ambiant traité au second conduit auditif (303).

18. Système selon l'une quelconque des revendications 16 et 17, dans lequel le premier dispositif d'oreille

Exhibit 1004
Page 0372 of 1001

(110, 210) comprend le premier circuit d'échange à commande vocale (VOX) (430) connecté au premier processeur (460), et le second dispositif d'oreille (210B) comprend en outre un second circuit VOX (430) connecté au second processeur.

19. Système selon l'une quelconque des revendications 16 à 18, dans lequel le premier processeur et le second processeur comprennent chacun un module de reconnaissance vocale.

20. Système selon l'une quelconque des revendications 16 à 19, dans lequel les premier et second dispositifs d'oreille sont respectivement le dispositif BTE.

21. Système selon l'une quelconque des revendications 16 à 19, dans lequel les premier et second dispositifs d'oreille sont respectivement le dispositif ITE.

22. Système selon la revendication dans lequel les premier et second dispositifs d'oreille sont respectivement un dispositif intra-conduit (ITC) (310B).

23. Système selon la revendication 22, dans lequel les premier et second dispositifs d'oreille sont respectivement un dispositif complètement intra-conduit (CIC) (310C).

24. Système selon l'une quelconque des revendications 16 à 23, comprenant en outre un dispositif distant (190, 290) couplé en communication aux premier et second dispositifs d'oreille.

25. Système selon la revendication 24, dans lequel le dispositif distant (190, 290) comprend au moins un parmi un ordinateur, un assistant numérique personnel (PDA), un téléphone cellulaire, un talkie walkie, ou un traducteur de langue.

26. Un procédé comprenant les étapes consistant à :

détecter un son occlus au moyen d'un premier microphone adapté en vue d'une utilisation dans un conduit auditif (303) ;
détecter un signal radio entrant représentant un son distant provenant d'un dispositif distant (190, 290) ;
lancer un échange à commande vocale (VOX) logé à l'intérieur d'un dispositif de communication d'oreille lorsqu'au moins un parmi le son occlus et le signal radio entrant est détecté si le VOX n'est pas déjà en marche, le dispositif de communication d'oreille étant un parmi un dispositif placé à l'intérieur de l'oreille (ITE) et un dispositif contour d'oreille (BTE) ;
laisser passer le son occlus, et bloquer le son distant lorsque le son distant occlus est détecté alors que le signal radio entrant n'est pas détecté ; et
laisser passer le son distant, et bloquer le son occlus lorsque le signal radio entrant est détecté alors que le son occlus n'est pas détecté.

27. Procédé selon la revendication 26, dans lequel la détection du son occlus comprend la comparaison d'une amplitude du son occlus avec un seuil de son occlus, dans lequel le son occlus est détecté lorsque l'amplitude du son occlus dépasse le seuil du son occlus.

28. Procédé selon l'une quelconque des revendications 26 et 27, dans lequel la détection du signal radio entrant comprend la comparaison d'une amplitude du son distant avec un seuil de son distant, dans lequel le signal radio entrant est détecté lorsque l'amplitude du son distant dépasse le seuil du son distant.

29. Procédé selon l'une quelconque des revendications 26 à 28, comprenant en outre la détection d'un son ambiant au moyen d'un second microphone, et dans lequel le lancement du VOX comprend le lancement du VOX lorsqu'au moins un parmi le son occlus, le signal radio entrant et le son ambiant est détecté si le VOX n'est pas déjà en marche.

30. Procédé selon la revendications 29, comprenant en outre l'étape consistant à bloquer le son occlus, laisser passer le son distant et bloquer le son ambiant lorsque le signal radio entrant est détecté alors que le son occlus n'est pas détecté.

31. Procédé selon la revendication 29, comprenant en outre l'étape consistant à bloquer le son occlus, laisser passer le son distant et laisser passer le son ambiant lorsque le signal radio entrant est détecté alors que le son occlus n'est pas détecté.

32. Procédé selon la revendication 29, comprenant en outre l'étape consistant à laisser passer le son occlus, bloquer le son distant, et bloquer le son ambiant lorsque le son occlus est détecté alors que le signal radio entrant n'est pas détecté.

Exhibit 1004
Page 0373 of 1001

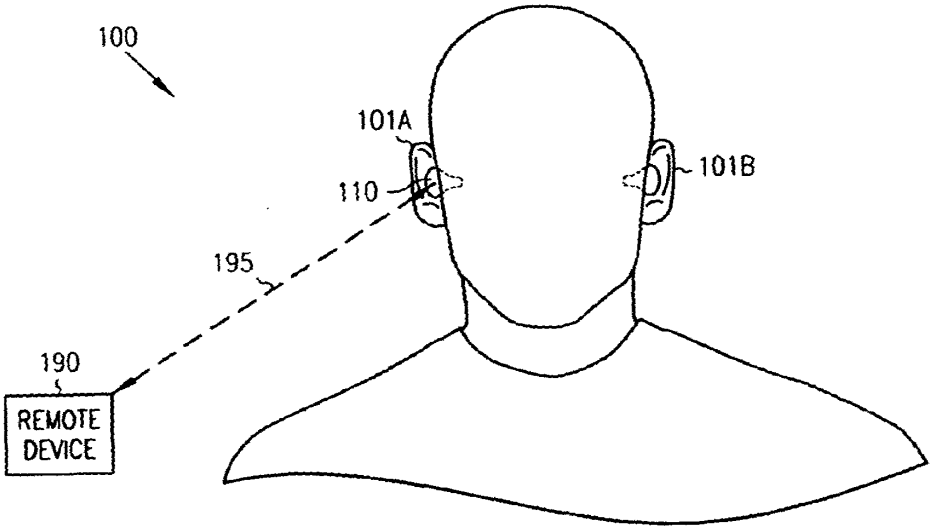

100
101A
101B
110
195
190
REMOTE DEVICE

FIG. 1

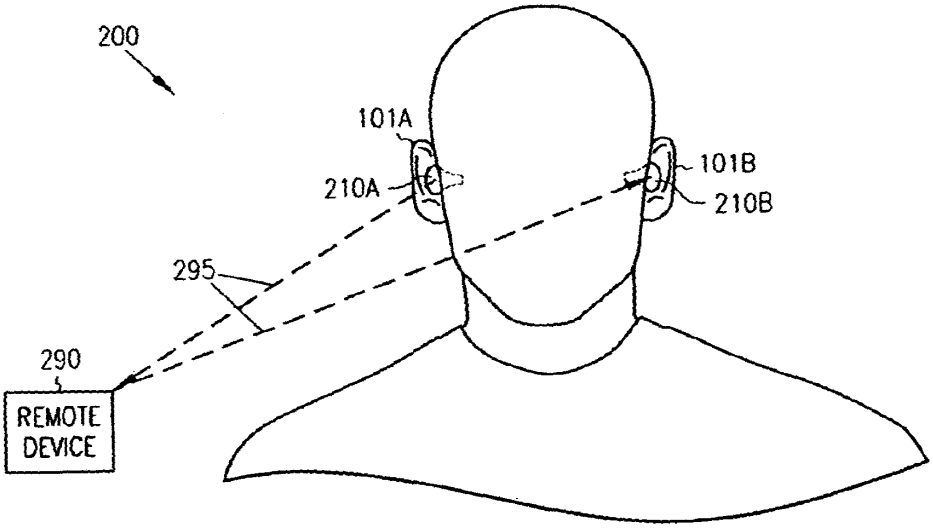

200
101A
101B
210A
210B
295
290
REMOTE DEVICE

FIG. 2

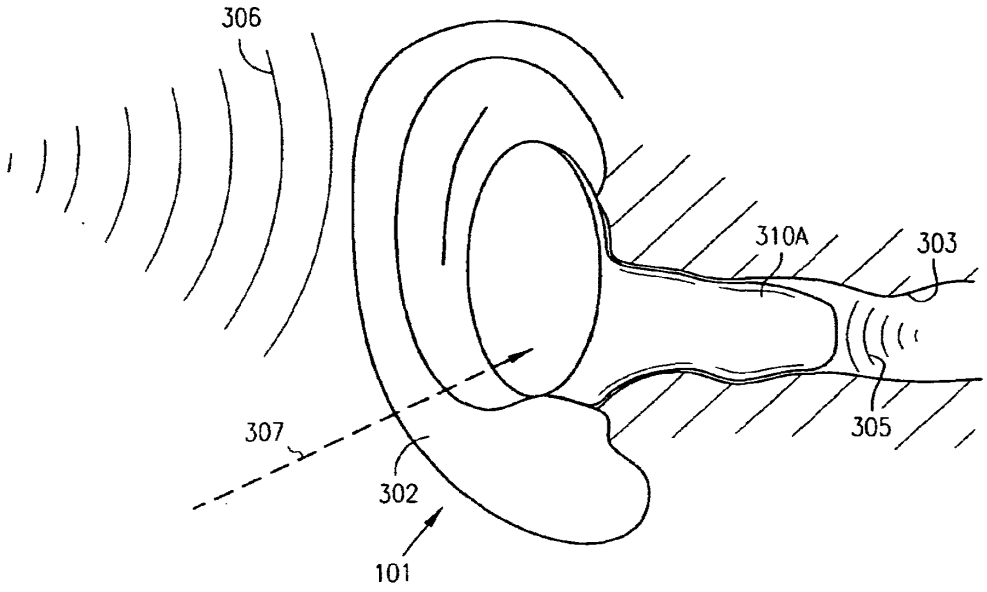


FIG. 3A

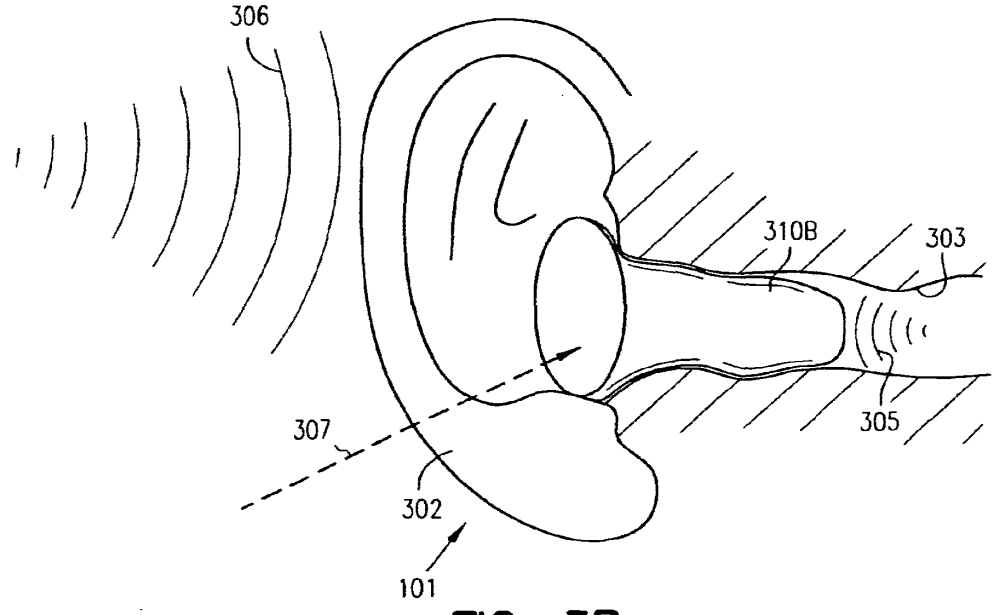


FIG. 3B

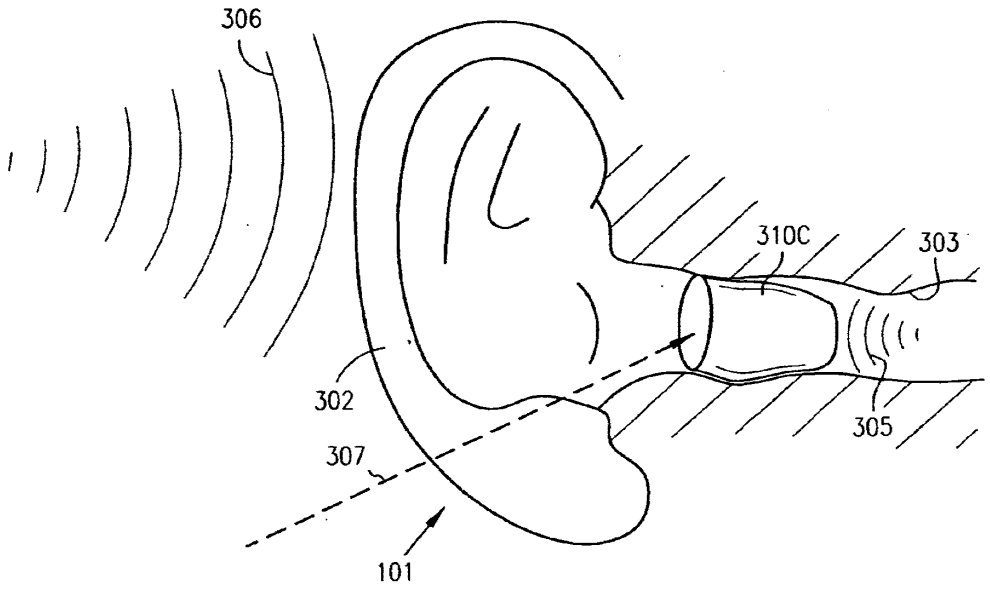


FIG. 3C

Exhibit 1004
Page 0376 of 1001

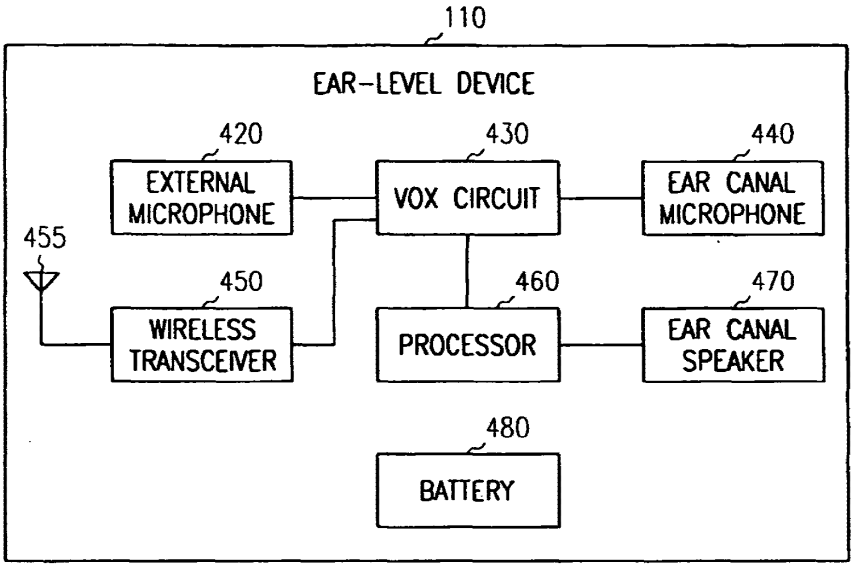


FIG. 4A

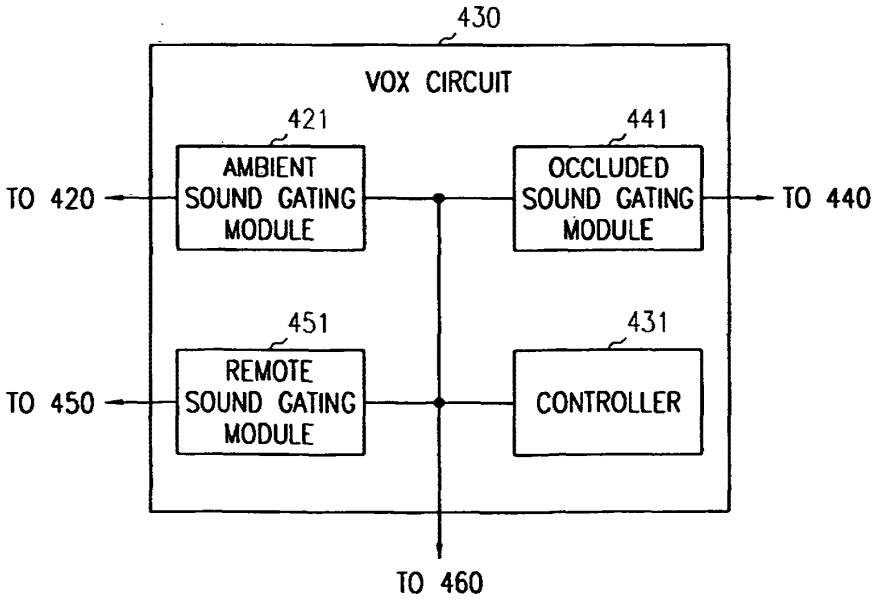


FIG. 4B

Exhibit 1004
Page 0377 of 1001

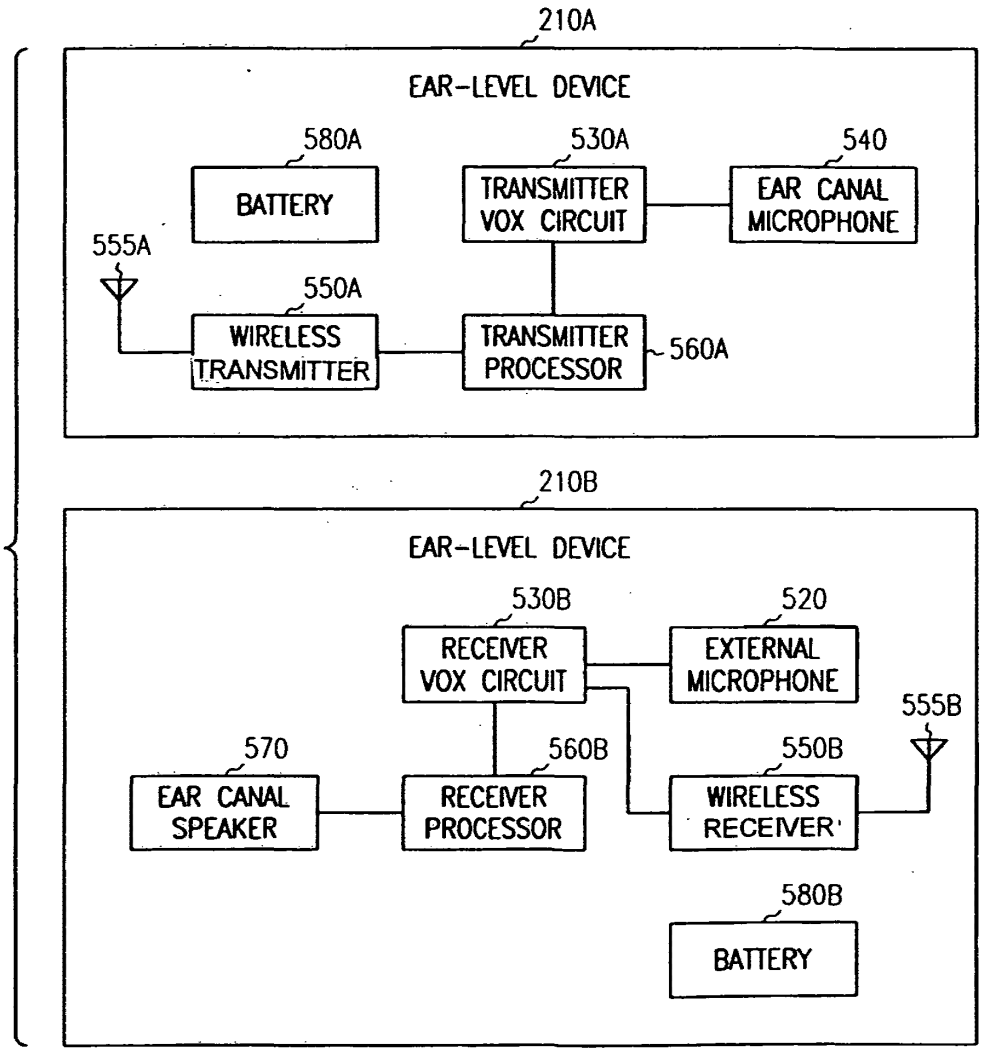

210A
EAR-LEVEL DEVICE
580A
BATTERY
530A
TRANSMITTER VOX CIRCUIT
540
EAR CANAL MICROPHONE
555A
550A
WIRELESS TRANSMITTER
TRANSMITTER PROCESSOR
560A
210B
EAR-LEVEL DEVICE
530B
RECEIVER VOX CIRCUIT
520
EXTERNAL MICROPHONE
555B
570
EAR CANAL SPEAKER
560B
RECEIVER PROCESSOR
550B
WIRELESS RECEIVER
580B
BATTERY

FIG. 5

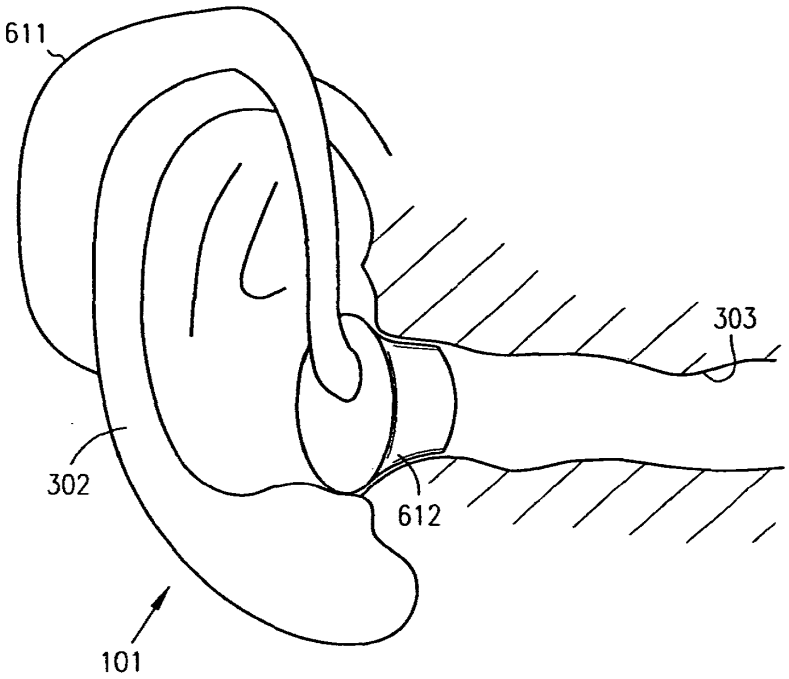

611
303
302
612
101

FIG. 6

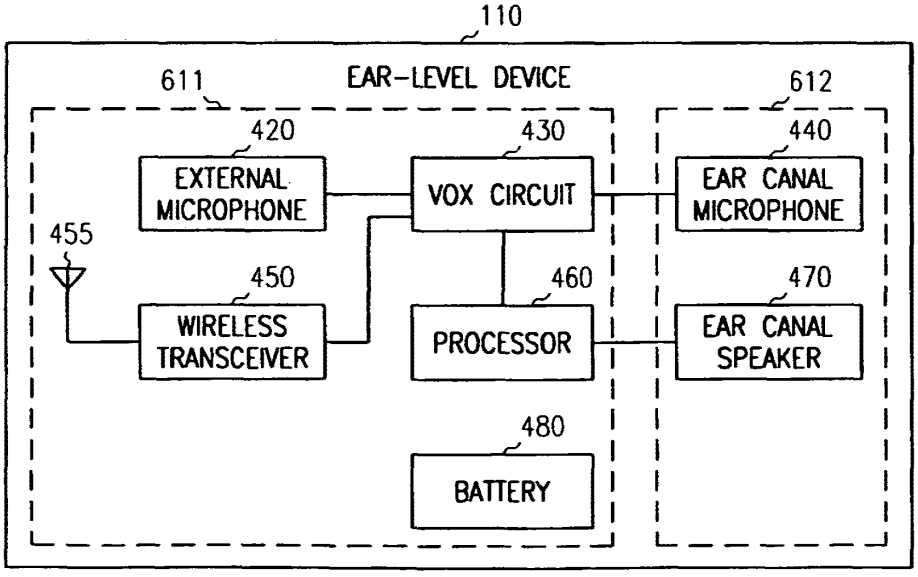

110
611
EAR-LEVEL DEVICE
612
420
EXTERNAL MICROPHONE
430
VOX CIRCUIT
440
EAR CANAL MICROPHONE
455
450
WIRELESS TRANSCEIVER
460
PROCESSOR
470
EAR CANAL SPEAKER
480
BATTERY

FIG. 7

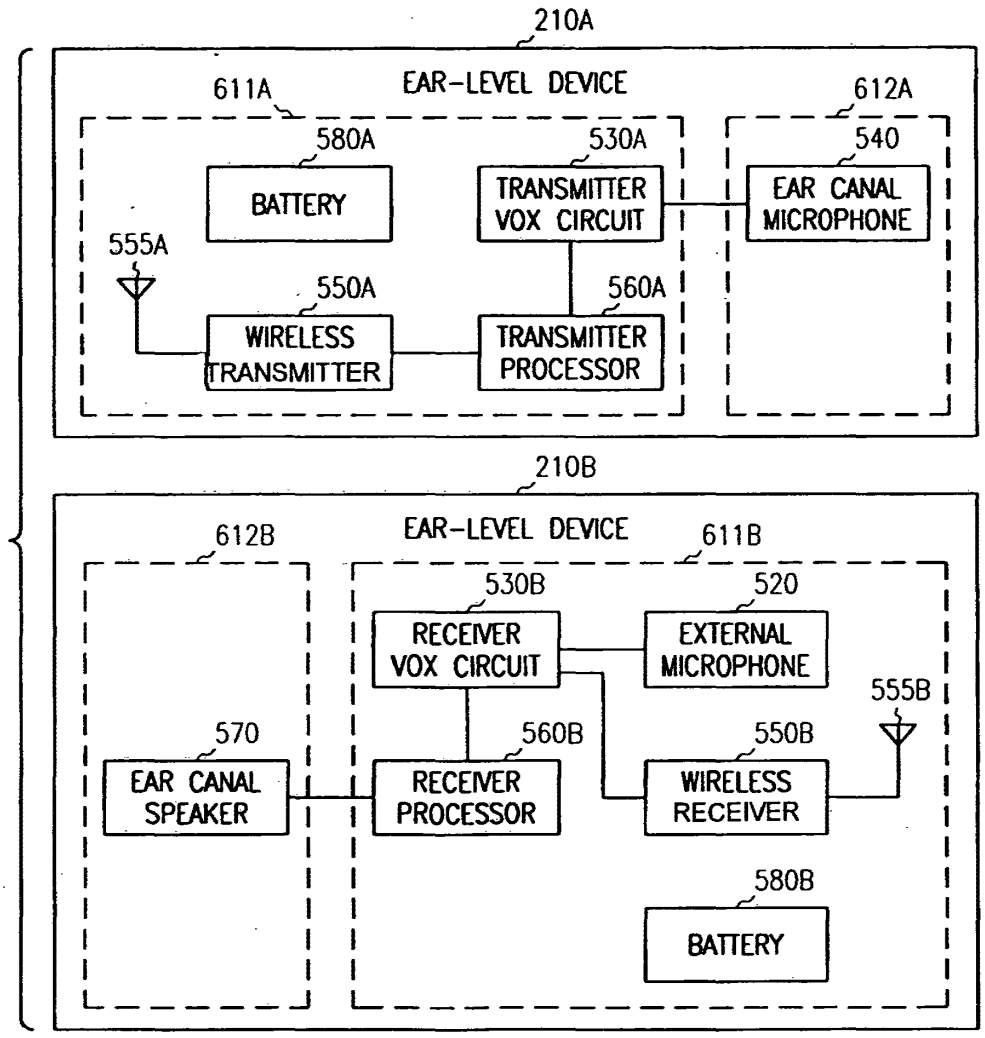

210A
EAR-LEVEL DEVICE
611A
612A
580A
530A
540
BATTERY
TRANSMITTER VOX CIRCUIT
EAR CANAL MICROPHONE
555A
550A
560A
WIRELESS TRANSMITTER
TRANSMITTER PROCESSOR
210B
EAR-LEVEL DEVICE
612B
611B
530B
520
RECEIVER VOX CIRCUIT
EXTERNAL MICROPHONE
555B
570
560B
550B
EAR CANAL SPEAKER
RECEIVER PROCESSOR
WIRELESS RECEIVER
580B
BATTERY

FIG. 8

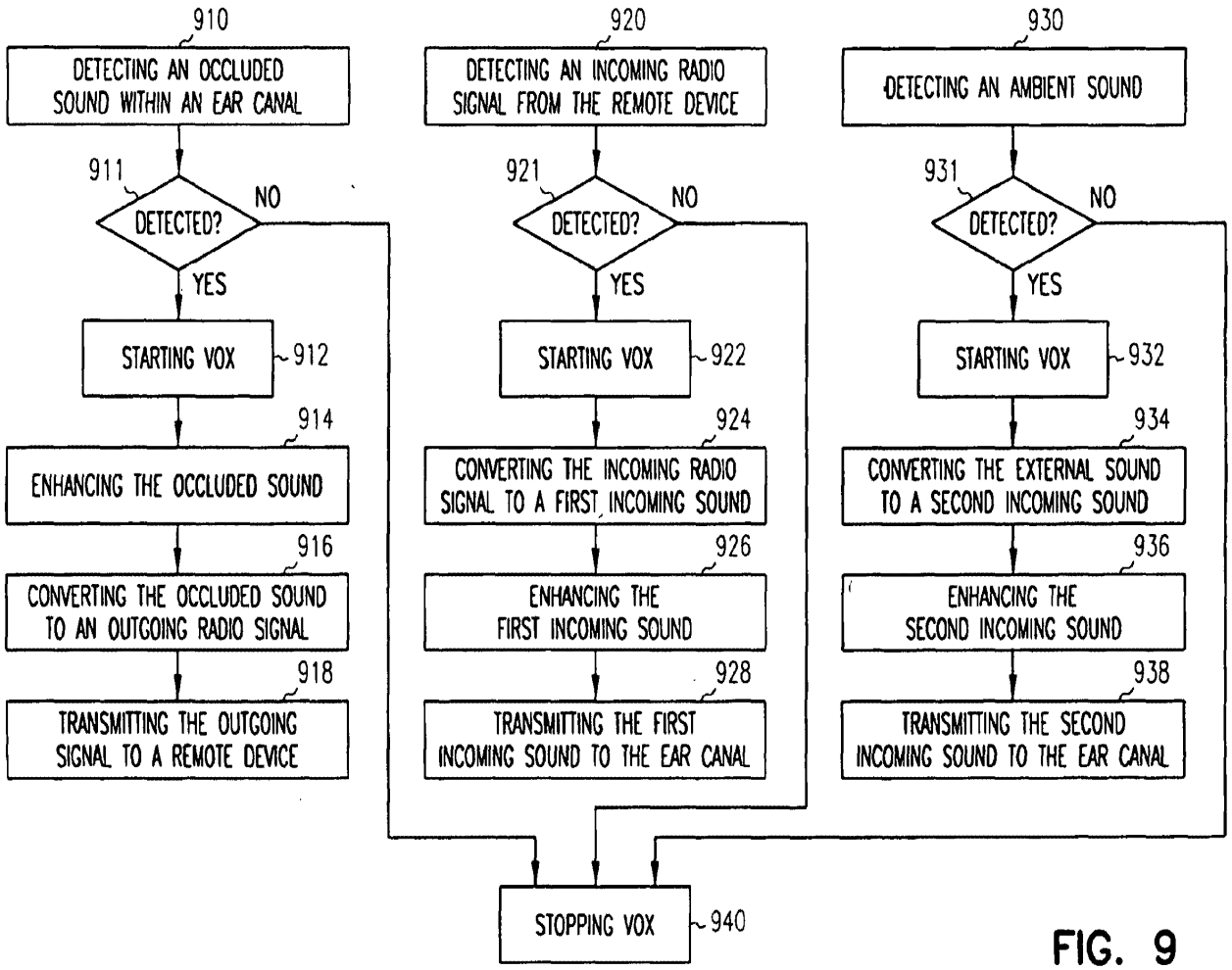

910
DETECTING AN OCCLUDED SOUND WITHIN AN EAR CANAL
911
DETECTED?
NO
YES
STARTING VOX
912
914
ENHANCING THE OCCLUDED SOUND
916
CONVERTING THE OCCLUDED SOUND TO AN OUTGOING RADIO SIGNAL
918
TRANSMITTING THE OUTGOING SIGNAL TO A REMOTE DEVICE
920
DETECTING AN INCOMING RADIO SIGNAL FROM THE REMOTE DEVICE
921
DETECTED?
NO
YES
STARTING VOX
922
924
CONVERTING THE INCOMING RADIO SIGNAL TO A FIRST INCOMING SOUND
926
ENHANCING THE FIRST INCOMING SOUND
928
TRANSMITTING THE FIRST INCOMING SOUND TO THE EAR CANAL
930
DETECTING AN AMBIENT SOUND
931
DETECTED?
NO
YES
STARTING VOX
932
934
CONVERTING THE EXTERNAL SOUND TO A SECOND INCOMING SOUND
936
ENHANCING THE SECOND INCOMING SOUND
938
TRANSMITTING THE SECOND INCOMING SOUND TO THE EAR CANAL
STOPPING VOX
940

FIG. 9

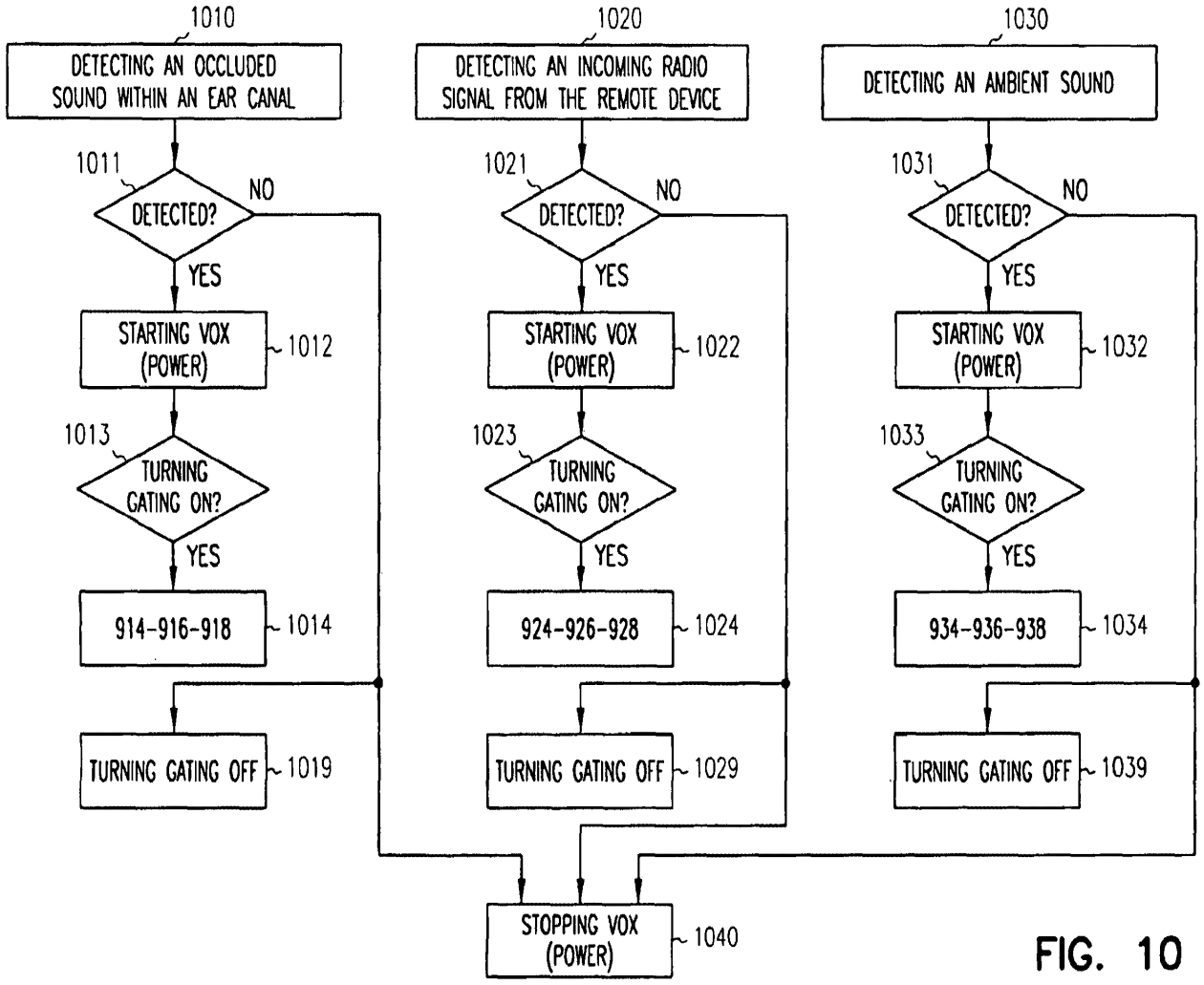

1010
DETECTING AN OCCLUDED SOUND WITHIN AN EAR CANAL
1011
DETECTED?
NO
YES
STARTING VOX (POWER)
1012
1013
TURNING GATING ON?
YES
914-916-918
1014
TURNING GATING OFF
1019
1020
DETECTING AN INCOMING RADIO SIGNAL FROM THE REMOTE DEVICE
1021
DETECTED?
NO
YES
STARTING VOX (POWER)
1022
1023
TURNING GATING ON?
YES
924-926-928
1024
TURNING GATING OFF
1029
1030
DETECTING AN AMBIENT SOUND
1031
DETECTED?
NO
YES
STARTING VOX (POWER)
1032
1033
TURNING GATING ON?
YES
934-936-938
1034
TURNING GATING OFF
1039
STOPPING VOX (POWER)
1040

FIG. 10

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 9845937 A **[0008]**
- WO 03073790 A **[0010]**

Exhibit 1004
Page 0383 of 1001

(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 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.

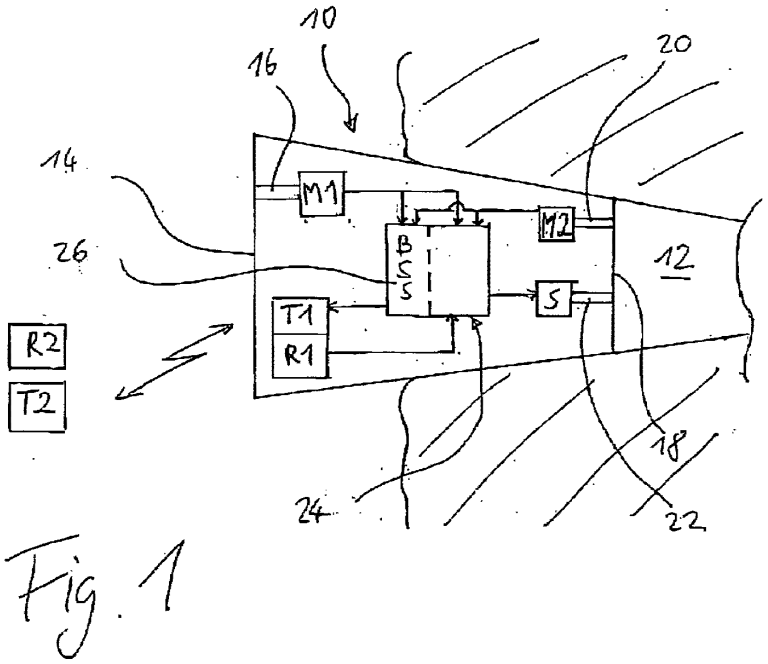


Fig. 1

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 1004
Page 0385 of 1001

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 1004
Page 0386 of 1001

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 1004
Page 0387 of 1001

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 1004
Page 0388 of 1001

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 1004
Page 0389 of 1001

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 1004
Page 0390 of 1001

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 1004
Page 0391 of 1001

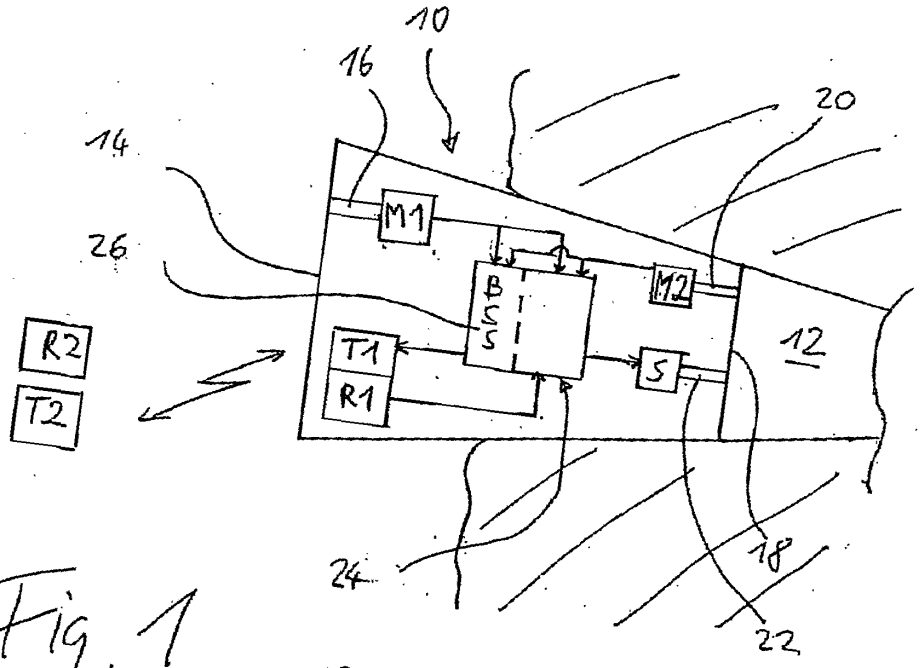


Fig. 1

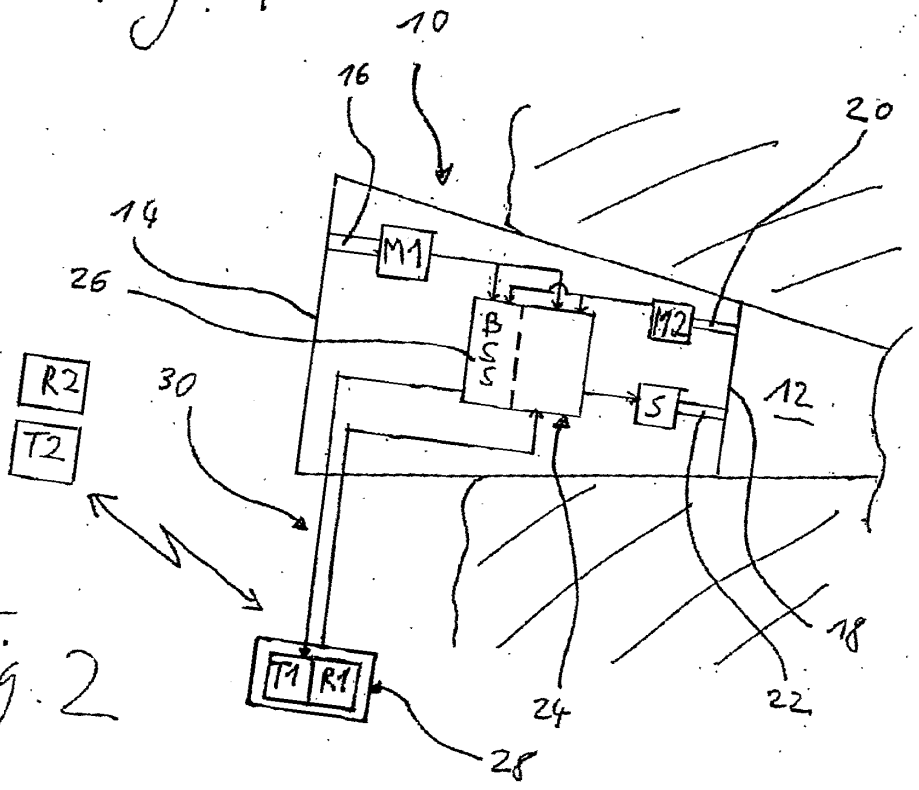


Fig. 2

Exhibit 1004
Page 0392 of 1001

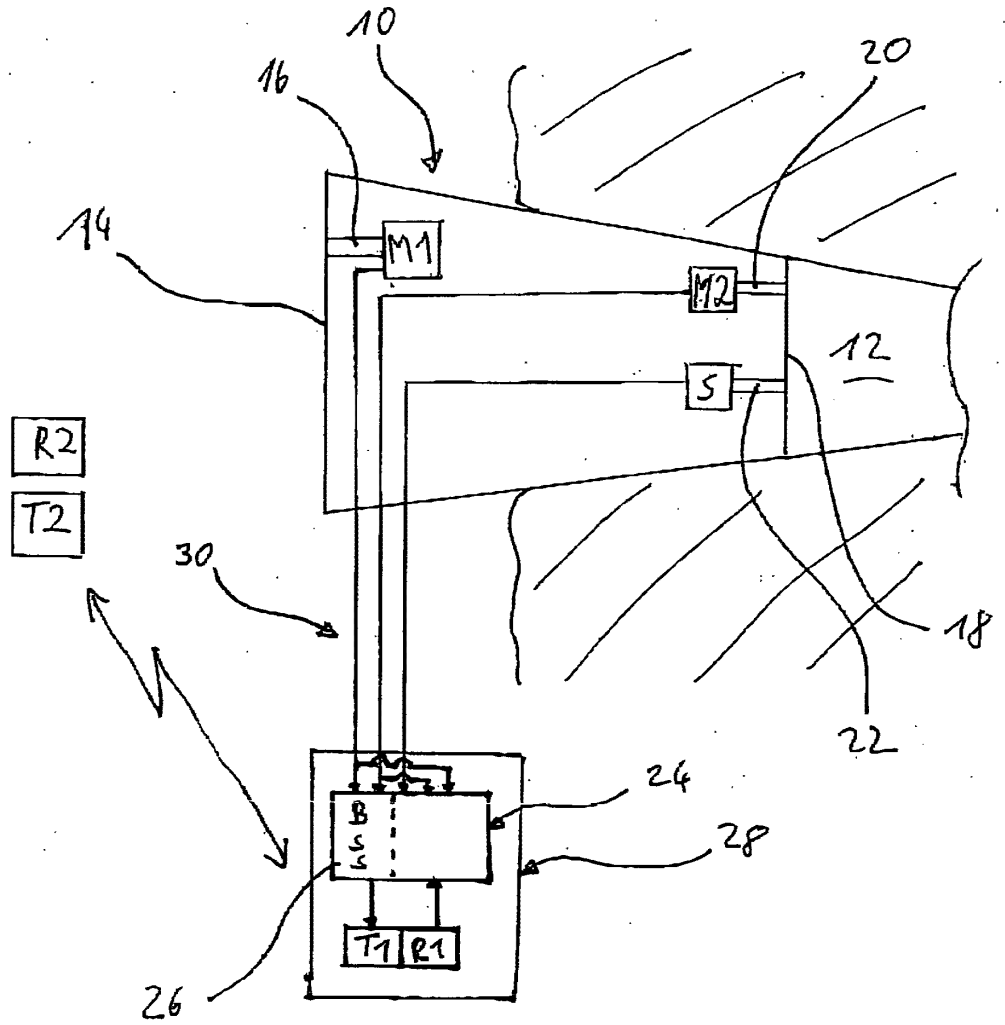


Fig. 3

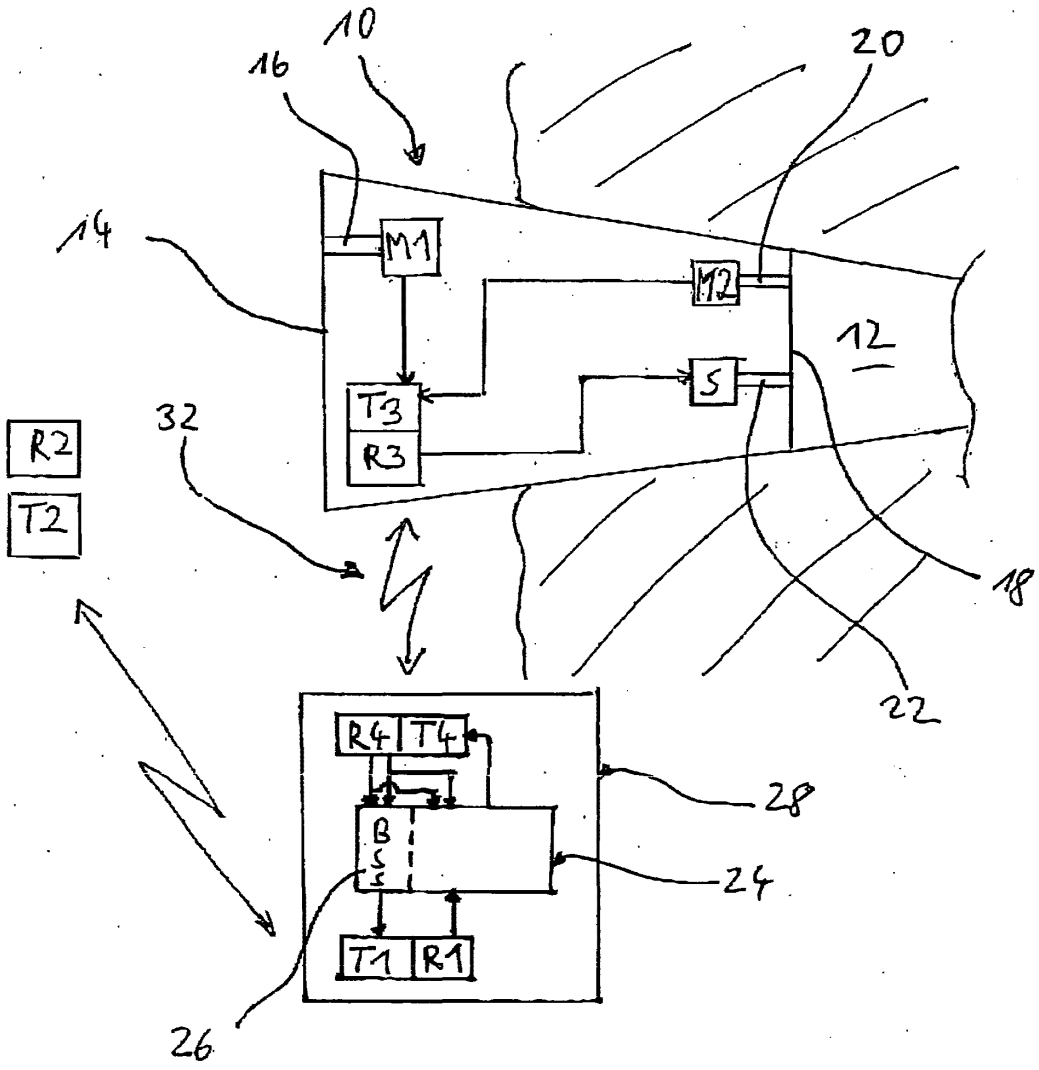


Fig. 4

Exhibit 1004
Page 0394 of 1001

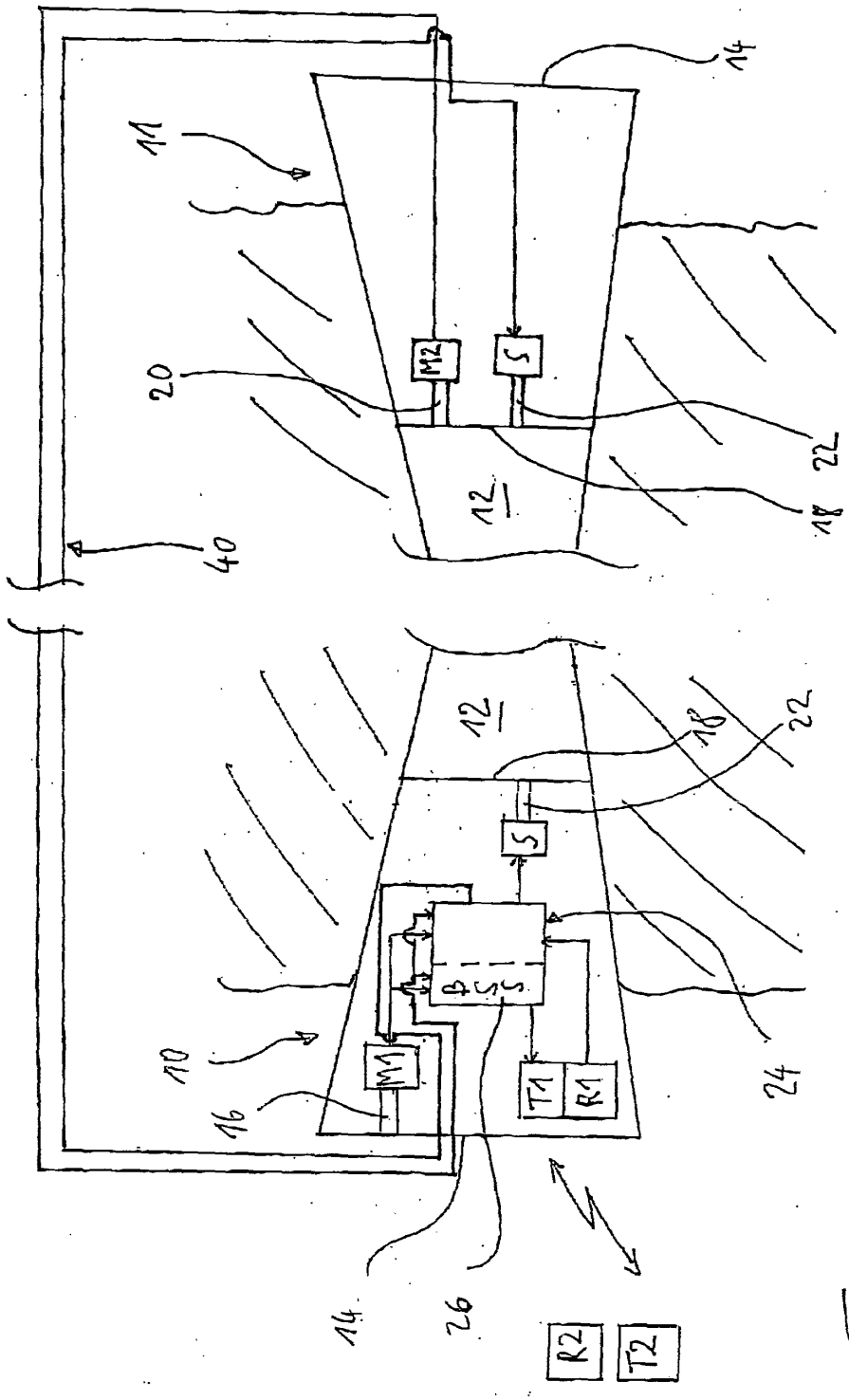


Fig. 5

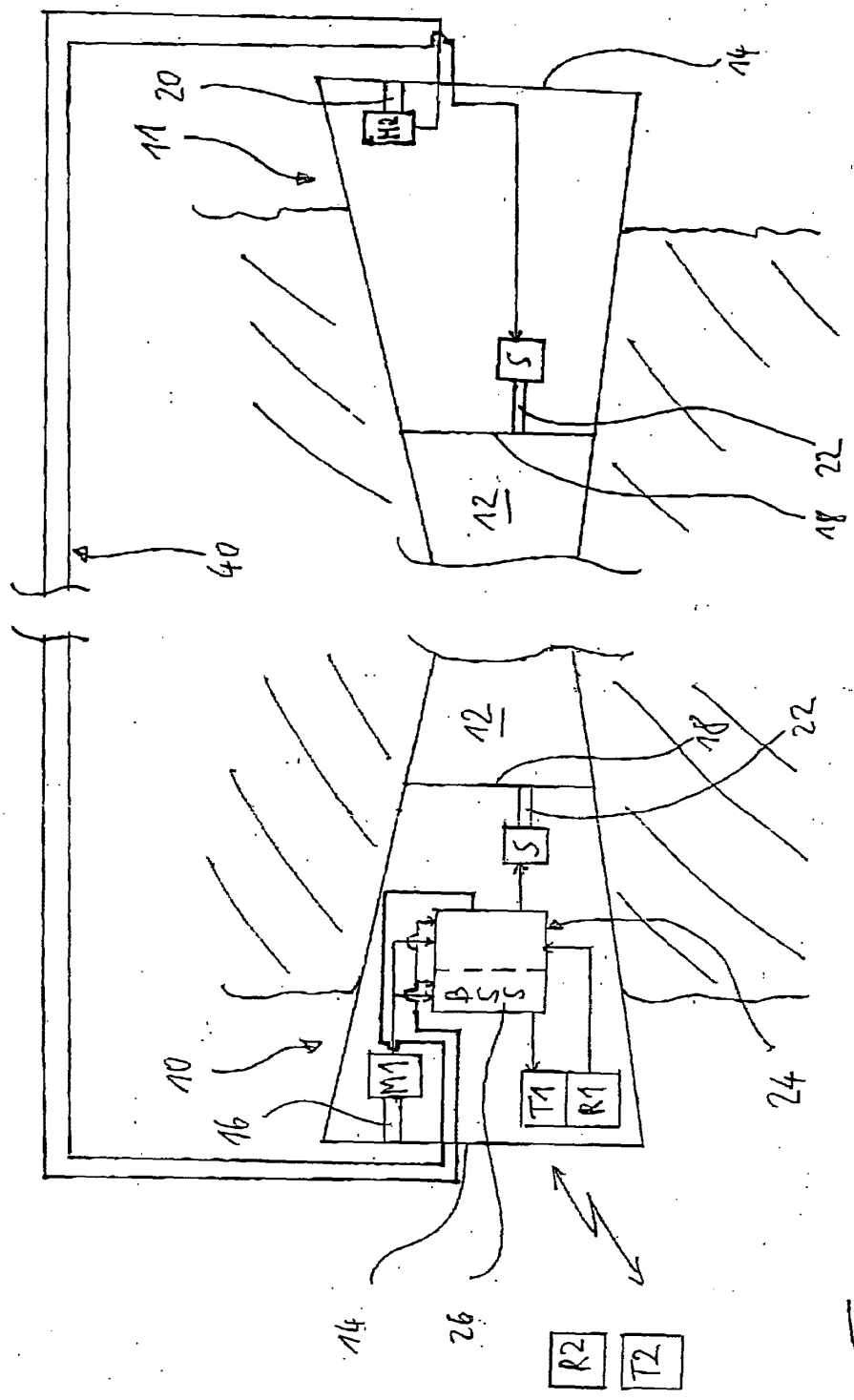


Fig. 6

Exhibit 1004
Page 0396 of 1001

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	
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	

TECHNICAL FIELDS SEARCHED (IPC)
G10L
H04R

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 1004
Page 0397 of 1001

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)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8－77468

(43)公開日　平成８年(1996) ３月22日

(51)Int.Cl.[6]	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０８Ｂ　13/196		9419－2E		
Ｈ０４Ｎ　5/915				
			Ｈ０４Ｎ　5/ 91	Ｋ

審査請求　未請求　請求項の数１　ＯＬ　（全 5 頁）

(21)出願番号　特願平6－214581

(22)出願日　平成６年(1994) ９月８日

(71)出願人　591114559
小野電機株式会社
福岡県田川市大字川宮110番地

(72)発明者　小野　敬三郎
福岡県田川市大字川宮110番地　小野電機株式会社内

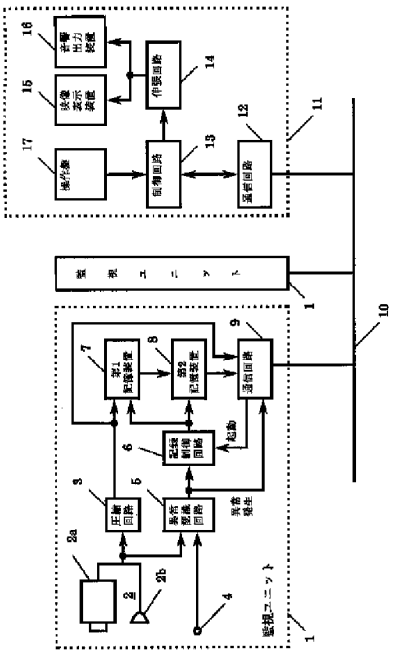
(74)代理人　弁理士　小堀　益

(54)【発明の名称】　監視装置

(57)【要約】

【目的】　重要な事象が発生した前後の情報をリアルタイムで連続的に記録することができると共に、その情報を長期間にわたって確実に保存できるようにすること。

【構成】　監視対象を撮像するビデオカメラ２ａと、監視対象の音を拾うマイクロホン２ｂと、ビデオカメラ２ａからの映像信号とマイクロホン２ｂからの音響信号とを圧縮する圧縮回路３と、この圧縮回路３の出力を所定時間分エンドレスで記録する第１記憶装置７と、この第１記憶装置７からの再生信号を記録し長期間にわたって保存可能な第２記憶装置８と、ビデオカメラ２ａからの映像信号とマイクロホン２ｂからの音響信号の内容を解析して異常事象の発生を検出する異常認識回路５と、この異常認識回路５により異常事象の発生が検出されたときに第１記憶装置７を一定時間再生状態とするとともに、第１記憶装置７の再生期間中に第２記憶装置８を記録状態とする記録制御回路６とを備えている。



【特許請求の範囲】
【請求項１】　監視対象の映像情報を検出する映像情報検出装置と、
監視対象の音響情報を検出する音響情報検出装置と、
前記映像情報検出装置からの映像信号と前記音響情報検出装置からの音響信号とを圧縮する圧縮回路と、
該圧縮回路の出力を所定時間分エンドレスで記録する第１記憶装置と、
該第１記憶装置からの再生信号を記録し長期間にわたって保存可能な第２記憶装置と、
前記映像情報検出装置からの映像信号と前記音響情報検出装置からの音響信号の少なくとも一方の信号の内容を解析して異常事象の発生を検出する異常認識回路と、
該異常認識回路により異常事象の発生が検出されたときに前記第１記憶装置を一定時間再生状態とするとともに、前記第１記憶装置の再生期間中に前記第２記憶装置を記録状態とする記録制御回路とからなることを特徴とする監視装置。
【発明の詳細な説明】
【０００１】
【産業上の利用分野】本発明は、商店、銀行等において使用する監視装置に関する。
【０００２】
【従来の技術】商店、銀行等において、防犯のため、或いは、犯罪が発生したときの記録のために、防犯カメラ等の監視装置を設けている。
【０００３】従来の監視装置においては、防犯カメラからの映像信号及び音声信号を、タイム・ラスプ・レコーダと呼ばれる間欠動作型式の磁気記録装置に記録している。このタイム・ラスプ・レコーダは、防犯カメラからの信号を数秒～数十秒の一定時間間隔で１フレームずつ記録するもので、数十時間の連続記録が可能である。
【０００４】
【発明が解決しようとする課題】しかしながら、このタイム・ラスプ・レコーダは、間欠的にしか信号を記録しないので、犯罪発生の瞬間や、被疑者の人相着衣等の重要な情報が記録されていないことがある。また、記録が間欠的であるため被疑者の挙動の詳細を知ることが出来なかった。
【０００５】また、防犯カメラからの信号を連続的に磁気記録装置に記録することも考えられるが、この場合は、使用する磁気テープの長さの制限から、連続記録可能な時間はせいぜい数時間であり、一定時間おきテープを交換しなければならないので実際的ではない。
【０００６】また、人間が操作して記録を開始することも考えられるが、防犯カメラの特性から考えて、記録が必要になったような状態で装置の操作を行なうのは危険性がある。また、操作が煩雑であり操作を忘れやすいため、これも実際的ではない。
【０００７】そこで、本発明は、重要な事象が発生した前後の情報をリアルタイムで連続的に記録することができると共に、その情報を長期間にわたって確実に保存できるようにすることを目的とする。
【０００８】
【課題を解決するための手段】本発明の監視装置は、前記目的を達成するため、監視対象の映像情報を検出する映像情報検出装置と、監視対象の音響情報を検出する音響情報検出装置と、前記映像情報検出装置からの映像信号と前記音響情報検出装置からの音響信号とを圧縮する圧縮回路と、該圧縮回路の出力を所定時間分エンドレスで記録する第１記憶装置と、該第１記憶装置からの再生信号を記録し長期間にわたって保存可能な第２記憶装置と、前記映像情報検出装置からの映像信号と前記音響情報検出装置からの音響信号の少なくとも一方の信号の内容を解析して異常事象の発生を検出する異常認識回路と、該異常認識回路により異常事象の発生が検出されたときに前記第１記憶装置を一定時間再生状態とするとともに、前記第１記憶装置の再生期間中に前記第２記憶装置を記録状態とする記録制御回路と　からなることを特徴とする。
【０００９】
【作用】本発明の監視装置においては、第１記憶装置は、常時記録動作を行っており、定常時には、監視対象近傍の映像情報及び音響情報が一定時間分エンドレスに記録される。異常認識回路により異常事象の発生が検出されると、第１記憶装置における再生動作が開始されるとともに、第２記憶装置の記録動作が開始され、所定時間後元の状態に戻る。したがって、第２記憶装置においては、異常事象の発生前後の監視対象近傍の映像情報及び音響情報が保持される。
【００１０】
【実施例】以下、図面を参照しながら実施例に基づいて本発明の特徴を具体的に説明する。
【００１１】図１は、本発明の監視装置の実施例を示すブロック図である。
【００１２】図において、１は、監視対象となる建物の中の互いに異なった場所に複数個配置された監視ユニットである。各監視ユニット１は、ＣＣＤカメラ等のビデオカメラ２ａとマイクロホン２ｂからなる監視カメラ装置２、ビデオカメラ２ａからの映像信号とマイクロホン２ｂからの音声信号を圧縮する圧縮回路３、人体の存在或いは動きを検出する人体センサ４、ビデオカメラ２ａからの映像信号とマイクロホン２ｂからの音声信号と人体センサ４からの出力が供給され、これらの信号の内容を解析して異常な事態発生したかどうかを検出する異常認識回路５、この異常認識回路５の出力に基づき後述する記憶装置の記録再生動作を制御する記録制御回路６、前記圧縮回路３からの圧縮信号を一定時間分エンドレスで記憶する第１記憶装置７、この第１記憶装置７から再生された情報を記録して長期間にわたって保存可能な第

２記憶装置８、この第２記憶装置８から再生された情報を電話回線等の通信回線１０を介して警備会社の中に設置されている集中管理システム１１に送信すると共に、集中管理システム１１からの各種の指示を受信するための通信回路９等を備えている。

【００１３】前記人体センサ４としては、たとえば、赤外線センサ、ドップラレーダ等が使用される。また、前記第１記憶装置７としては、たとえば、ハードディスク装置や、半導体メモリからなるソリッド・ステート・ファイルが使用され、また、第２記憶装置８としては、たとえば、ビデオ・テープ・レコーダ等の磁気記録再生装置が使用される。

【００１４】集中管理システム１１は、通信回線１０を介して各監視ユニット１と通信するための通信回路１２と、集中管理システム１１全体の制御を行なう制御回路１３と、圧縮された映像信号及び音響信号を伸張して元の信号に復元する伸張回路１４と、復元された映像信号に基づきビデオカメラ２ａで撮像された映像を表示する映像表示装置１５と、復元された音響信号に基づきマイクロホン２ｂで拾われた音響を出力する音響出力装置１６と、各監視ユニット１に対して各種の指示を与えるためのキーボードや異常発生を表示するためのランプやＣＲＴディスプレイを含む操作盤１７とを備えている。前記制御回路１３は、各監視ユニット１からの情報を受信して、伸張回路１４を介して映像出力装置１５及び音響出力装置１６に供給したり、操作盤１７から入力された各種の指示を各監視ユニット１に対して送り出す処理を行なう。

【００１５】次に、上述した監視装置の動作について説明する。ここでは、閉店後の店内を店内の各所に配置された複数の監視ユニット１で監視する場合を例に挙げて説明する。ビデオカメラ２ａからの映像信号とマイクロホン２ｂからの音声信号は圧縮回路３に記録され、符号化等によりデータ量の圧縮が行なわれる。圧縮方式としては種々の方式が採用可能であるが、たとえば、ＭＰＥＧ（ｍｏｖｉｎｇ　ｐｉｃｔｕｒｅ　ｃｏｄｉｎｇ　ｅｘｐｅｒｔｓ　ｇｒｏｕｐ）方式と呼ばれるフレーム内符号化とフレーム間符号化を併用した標準化された符号化方式を採用することができる。

【００１６】圧縮回路３からの圧縮データは、第１記憶装置７に常時供給されている。第１記憶装置７は、記録制御回路６からのモード切り換え信号により記録モードと再生モードの切り換えが可能となっている。通常状態では、第１記憶装置７は記録モードとなっており、第１記憶装置７は圧縮回路３からの一定時間分の圧縮データをエンドレスで記録する。また、第２記憶装置８は、通常状態では停止状態となっており、第１記憶装置７の再生時に記録状態となる。

【００１７】たとえば、第１記憶装置７が半導体メモリからなるソリッド・ステート・ファイルである場合、メモリの先頭アドレスから順番に圧縮データを書き込んでいき、メモリ領域が一杯になったら、最初に書き込んだアドレスから重ねて書き込んでいく。ソリッド・ステート・ファイルは、機械的な可動部品が全くないので、半永久的に使用が可能である。ハードディスク装置の場合も同様に、回転するハードディスク上の各セクタ及びトラックに順次連続的に記録している。ハードディスク装置においては、ハードディスクは常時回転した状態にあり、ハードディスクの回転により発生する空気流により、書き込み・読み取りヘッドはハードディスクの表面から浮いた状態にあるので、ハードディスクの表面が摩耗することはない。したがって、数カ月～数年にわたって極めて長期間の使用が可能となる。

【００１８】これにより第１記憶装置７には、現在の時点から、第１記憶装置７の記憶容量に対応した一定時間だけ遡った時点までの情報が記録されることになる。通常状態においては、上述したように第１記憶装置７の記録内容が、最新の一定期間の情報で常時更新される。すなわち、第１記憶装置７はエンドレスの記憶装置として機能する。

【００１９】次に、店内に不審者が侵入し人体センサ４の近傍を通過した場合には、人体センサ４の出力が異常認識回路５に供給される。異常認識回路５は、人体センサ４から出力があったときには、異常が発生したと判断する。

【００２０】また、窓ガラスが割られたり、ドアがこじ開けられた場合には、その時に発生する音がマイクロホン２ｂにより拾われ、異常認識回路５に供給される。異常認識回路５においては、入力された可聴周波数領域の信号の周波数成分及び振幅成分のパターンが、予め登録されている窓ガラスが割られたときに発生する音のパターンや、ドアがこじ開けられた場合に発生する音のパターンと一致するかどうか判別し、一致したときには、異常が発生したと判断する。

【００２１】また、ビデオカメラ２ａからの映像信号は異常認識回路５に供給され、撮像範囲内で映像の変化があったかどうか、すなわち、撮像範囲内に動体が存在するか否かが判別される。この動体判別は、たとえば、一定時間離れた二つの時点における映像信号の内容の差異の有無を調べることにより検出することができる。すなわち、二つの映像信号の内容に差異があるときには、異常が発生したと判断する。また、映像の内容に予め登録された所定パターン以外のパターンが表れたときに異常が発生したと判断することもできる。

【００２２】上述したように、ビデオカメラ２ａ、マイクロホン２ｂ、人体センサ４の出力を異常認識回路５に供給することにより、各種の状況に対応可能となる。

【００２３】異常認識回路５により異常が発生したと判断されると、異常認識回路５から異常検出信号が記録制御回路６に供給され、第１記憶装置７が記録モードから

再生モードに切り換えられると共に、第２記憶装置８の記録動作が開始され、第１記憶装置７に記録されていた最新の一定期間の情報が第２記憶装置８に記録される。第１記憶装置７に記録されていた情報が第２記憶装置８に記録されると、第１記憶装置７は再度記録モードとされると共に、第２記憶装置８の記録動作が停止される。

【００２４】上述の動作により、異常事象発生毎に第１記憶装置７の一定時間の再生動作と、この再生動作に同期した第２記憶装置８の一定時間の記録動作が繰り返される。したがって、第２記憶装置８には、異常事象発生毎に、異常事象発生直前の一定の映像情報及び音響情報が順次記録されることになる。

【００２５】また、異常認識回路５により異常が発生したと判断されると、異常が発生した事を示す異常発生データが、どの監視ユニットからのデータであるかを示す識別データと共に、通信回路９及び通信回線１０を介して警備会社の中に設置されている集中管理システム１１に対して送信される。この異常発生データは、通信回路１２を介して制御回路１３に供給され、操作盤１７にどこでどのような種類の異常が発生したのが表示される。また、第１記憶装置７に記録されている情報が再生されて第２記憶装置８に記録されると同時に、第１記憶装置７からの再生信号は通信回線１０を介して集中管理システム１１の制御回路１３に供給され、更に、伸張回路１４で元の映像信号及び音響信号に復元され、映像出力装置１５及び音響出力装置１６に供給され、異常事象発生直前の映像及び音が再生される。なお、第１記憶装置７に記録されている情報が再生された後には、ビデオカメラ２ａ及びマイクロホン２ｂからのリアルタイムの情報を自動的に、或いは、操作盤１７からの指示に基づき送信するようにしてもよい。

【００２６】これにより、警備会社は異常の詳細を知ることができ、警備員派遣、所轄の警察署への通報等の適切な処置をとることができる。

【００２７】また、操作盤１７から、異常を検出した監視ユニット以外の監視ユニット１の記録制御回路６に対して起動指示を与えることにより、異常を検出した監視ユニットの近傍の他の監視ユニットにおいても、第２記憶装置８への記録が開始されると共に、リアルタイムの情報が通信回線１０介して送信され、集中管理システム１１側では、侵入者の侵入経路や逃亡経路を容易に把握することができる。

【００２８】また、操作盤１７から監視ユニット１の記録制御回路６に対して第２記憶装置８を再生する指示を与えることにより、集中管理システム１１側で第２記憶装置８に記録されている過去分の異常事象の状態も知ることができる。なお、監視ユニット１の第２記憶装置８に直接テレビモニタを接続して、第２記憶装置８の記録内容を再生することもできる。

【００２９】なお、上述の実施例においては、防犯装置を例に挙げて説明したが、本発明はこれに限定されるものではなく、たとえば、化学プラントや製造工場における異常事象の監視にも適用することができる。

【００３０】

【発明の効果】以上に述べたように、本発明においては、監視対象の状況を第１記憶装置に常時エンドレスに記録しておき、異常事象発生時には第１記憶装置で再生動作を行わせて、この再生信号を長期間の保存が可能な第２記憶装置に記録するようにしている。これにより、エンドレス記録を行なう第１記憶装置の記録可能時間に制限されることなく、異常事象発生時の状況を確実に保持することができる。

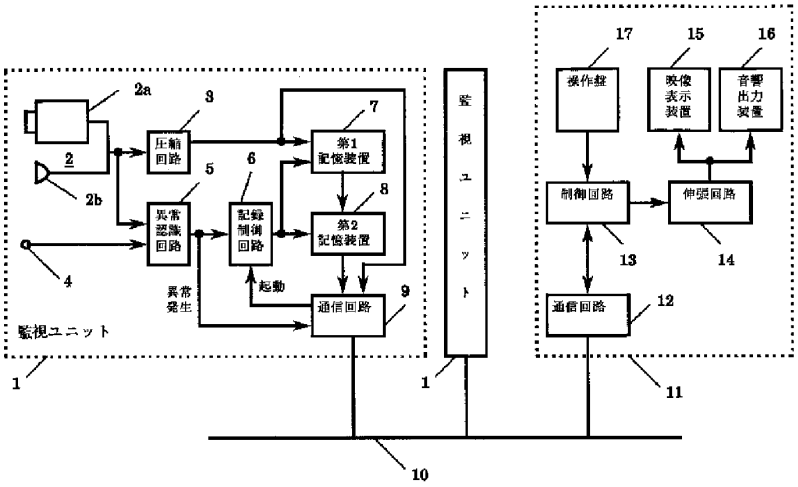
【図面の簡単な説明】

【図１】　本発明の監視装置の実施例を示すブロック図である。

【符号の説明】

１…監視ユニット、２…監視カメラ装置、２ａ…ビデオカメラ、２ｂ…マイクロホン、３…圧縮回路、４…人体センサ、５…異常認識回路、６…記録制御回路、７…第１記憶装置、８…第２記憶装置、９…通信回路、１０…通信回線、１１…集中管理システム、１２…通信回路、１３…制御回路、１４…伸張回路、１５…映像表示装置、１６…音響表示装置、１７…操作盤

【図１】


2a
3
7
圧縮回路
第1記憶装置
2
5
6
8
2b
異常認識回路
記録制御回路
第2記憶装置
4
異常発生
起動
9
通信回路
監視ユニット
1
監視ユニット
1
17
15
16
操作盤
映像表示装置
音響出力装置
制御回路
伸張回路
13
14
通信回路
12
11
10

Espacenet

Bibliographic data: JPH0877468 (A) — 1996-03-22

MONITOR DEVICE

Inventor(s): ONO KEIZABURO ± (ONO KEIZABURO)

Applicant(s): ONO DENKI KK ± (ONO DENKI KK)

Classification: **- international:** ***G08B13/196; H04N5/915;*** (IPC1-7): G08B13/196; H04N5/915
- cooperative:

Application number: JP19940214581 19940908

Priority number(s): JP19940214581 19940908

Abstract of JPH0877468 (A)

PURPOSE: To record information before and after an important event occurs successively in real time and securely store the information for a long period. CONSTITUTION: This device is equipped with a video camera 2a which picks up an image of an object to be monitored, a microphone 2b which picks up the sound of the object to be monitored, a compressing circuit 3 which compresses the video signal from the video camera 2a and the sound signal from the microphone 2b, a 1st storage device 7 wherein the output of the compressing circuit 3 is recorded endlessly for a specific time, a 2nd storage device 8 wherein the reproduced signal from the 1st storage device 7 is stored for a long period, an abnormality recognizing circuit 5 which analyzes the contents of the video signal from the video camera 2a and the sound signal from the microphone 2b and detects the occurrence of an abnormal event, and a recording control circuit 6 which places the 1st storage device 7 in a reproduction state for a certain time when the abnormality recognizing circuit 5 detects the abnormal event and places the 2nd storage device 8 in a recording state during the reproduction period of the 1st storage device 7.

JPH0877468A MONITOR DEVICE

Data originating from sources other than the EPO may not be accurate, complete, or up to date.

The wording below is an initial machine translation of the original publication. To generate a version using the latest translation technology, go to the original language text and use Patent Translate.

[0001] BACKGROUND OF THE INVENTION 1. Field of the Invention The present invention relates to a monitoring device used in stores, banks and the like.

[0002] 2. Description of the Related Art In a store, a bank or the like, a monitoring device such as a security camera is provided for security or for recording when a crime has occurred.

[0003] In a conventional monitoring apparatus, a video signal and an audio signal from a security camera are recorded in an intermittent operation type magnetic recording apparatus called a time-rasp recorder. The time-rasp recorder records a signal from a security camera one frame at a time at intervals of several seconds to several tens of seconds, and is capable of continuous recording for several tens of hours.

[0004] However, since the time-rasp recorder only records signals intermittently, important information such as the moment of a crime or the clothing of a suspect is not known. It may not be recorded. In addition, since the record was intermittent, it was not possible to know details of the behavior of the suspect.

[0005] It is also conceivable to continuously record signals from a security camera on a magnetic recording device. In this case, however, the continuous recording time is several hours at most due to the limitation of the length of a magnetic tape used. This is not practical because the tape must be changed at regular intervals.

[0006] It is also conceivable that a recording operation is started by a human being. However, there is a danger in operating the apparatus in a state where recording is necessary, in view of the characteristics of the security camera. This is not practical because the operation is complicated and the operation is easy to forget.

[0007] It is therefore an object of the present invention to be able to continuously record information before and after an important event has occurred in real time, and to reliably store the information for a long period of time.

[0008] In order to achieve the above object, a monitoring device according to the present invention has a video information detecting device for detecting video information to be monitored and an audio information for detecting acoustic information to be monitored. A detection device, a compression circuit that compresses a video signal from the video information detection device and an audio signal from the audio information detection device, and a first storage device that stores an output of the compression circuit endlessly for a predetermined time, A second storage device capable of recording a reproduction signal from the first storage device and storing the reproduced signal for a long period of time; and a content of at least one of a video signal from the video information detection device and an audio signal from the audio information detection device. An abnormality recognition circuit that detects the occurrence of an abnormal event by analyzing the first storage device when the occurrence of the abnormal event is detected by the abnormality recognition circuit for a certain period of time, It said second storage device in the serial reproduction period of the first storage device, characterized by comprising a recording control circuit for a recording state.

[0009] In the monitoring apparatus according to the present invention, the first storage device performs a recording operation at all times, and in a steady state, video information and audio information near the monitoring target are recorded endlessly for a fixed time. When the occurrence of an abnormal event is detected by the abnormality recognizing circuit, the reproducing operation in the first storage device is started, and the recording operation in the second storage device is started, and returns to the original state after a predetermined time. Therefore, in the second storage device, video information and audio information near the monitoring target before and after the occurrence of the abnormal event are held.

[0010] DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS Hereinafter, the features of the present invention will be specifically described based on embodiments with reference to the drawings.

[0011] FIG. 1 is a block diagram showing an embodiment of a monitoring apparatus according to the present invention.

[0012] Referring to FIG. 1, reference numeral 1 denotes a plurality of monitoring units arranged at different places in a building to be monitored. Each monitoring unit 1 includes a monitoring camera device 2 including a video camera 2a such as a CCD camera and a microphone 2b, a compression circuit 3 for compressing a video signal from the video camera 2a and an audio signal from the microphone 2b, and the presence or movement of a human body. The human body sensor 4 to be detected, the video signal from the video camera 2a, the audio signal from the microphone 2b, and the output from the human body sensor 4 are supplied, and the contents of these signals are analyzed to detect whether an abnormal situation has occurred. An abnormality recognizing circuit 5, a recording control circuit 6 for controlling a recording / reproducing operation of a storage device to be described later based on an output of the abnormality recognizing circuit 5, and a first storage device for storing a compression signal from the compression circuit 3 endlessly for a fixed time. 7, a second storage device 8 capable of recording information reproduced from the first storage device 7 and storing the information for a long period of time; The information reproduced from the device 8 is transmitted to a centralized management system 11 installed in a security company via a communication line 10 such as a telephone line, and various instructions from the centralized management system 11 are received. The communication circuit 9 is provided.

[0013] As the human body sensor 4, for example, an infrared sensor, a Doppler radar, or the like is used. As the first storage device 7, for example, a hard disk device or a solid state file made of a semiconductor memory is used. As the second storage device 8, for example, a magnetic tape device such as a video tape recorder is used. A recording / reproducing device is used.

[0014] The centralized management system 11 includes a communication circuit 12 for communicating with each monitoring unit 1 via the communication line 10, a control circuit 13 for controlling the entire centralized management system 11, a compressed video signal and audio signal. An expansion circuit 14 for expanding the signal to restore the original signal, an image display device 15 for displaying an image captured by the video camera 2a based on the restored image signal, and a microphone 2b based on the restored audio signal. A sound output device 16 for outputting the picked-up sound, and a control panel 17 including a keyboard for giving various instructions to each monitoring unit 1, a lamp for displaying occurrence of abnormality, and a CRT display are provided. . The control circuit 13 receives information from each monitoring unit 1, supplies the information to the video output device 15 and the audio output device 16 via the expansion circuit 14, and receives various instructions input from the operation panel 17. The process of sending out to the monitoring unit 1 is performed.

[0015] Next, the operation of the above-described monitoring device will be described. Here, a case where the inside of the store after the store is closed is monitored by a plurality of monitoring units 1 arranged at various places in the store will be described as an example. The video signal from the video camera 2a and the audio signal from the microphone 2b are recorded in the compression circuit 3, and the data amount is compressed by encoding or the like. Various methods can be adopted as the compression method. For example, it is possible to adopt a standardized coding method that uses both intra-frame coding and inter-frame coding called MPEG (moving picture coding experts group). it can.

[0016] The compressed data from the compression circuit 3 is always supplied to the first storage device 7. The first storage device 7 can switch between a recording mode and a reproduction mode by a mode switching signal from the recording control circuit 6. In the normal state, the first storage device 7 is in the recording mode, and the first storage device 7 records the compressed data from the compression circuit 3 for a certain period of time endlessly. Further, the second storage device 8 is in a stopped state in a normal state, and is in a recording state when the first storage device 7 is reproduced.

[0017] For example, when the first storage device 7 is a solid state file composed of a semiconductor memory, compressed data is written in order from the head address of the memory, and when the memory area is full, the compressed data is written first. Writing is repeated starting from the address. Solid state files can be used semi-permanently because there are no mechanical moving parts. Similarly, in the case of a hard disk device, the data is sequentially and continuously recorded on each sector and track on the rotating hard disk. In a hard disk drive, the hard disk is constantly rotating, and the write / read head is floating above the surface of the hard disk due to the airflow generated by the rotation of the hard disk, so that the surface of the hard disk does not wear. Therefore, it can be used for a very long time over several months to several years.

[0018] As a result, the information from the current time to the time that has been traced back by a certain time corresponding to the storage capacity of the first storage device 7 is recorded in the first storage device 7. In the

normal state, as described above, the recorded contents of the first storage device 7 are constantly updated with the latest information of a certain period. That is, the first storage device 7 functions as an endless storage device.

[0019] Next, when a suspicious person enters the store and passes near the human body sensor 4, the output of the human body sensor 4 is supplied to the abnormality recognition circuit 5. When there is an output from the human body sensor 4, the abnormality recognition circuit 5 determines that an abnormality has occurred.

[0020] When the window glass is broken or the door is opened, the sound generated at that time is picked up by the microphone 2b and supplied to the abnormality recognition circuit 5. In the abnormality recognition circuit 5, the pattern of the frequency component and the amplitude component of the input signal in the audible frequency range is changed to a pattern of a sound generated when a window glass registered in advance is broken or a door is opened. It is determined whether or not the pattern matches the sound pattern generated in the case. If the pattern matches, it is determined that an abnormality has occurred.

[0021] The video signal from the video camera 2a is supplied to the abnormality recognizing circuit 5, and it is determined whether or not the video has changed within the imaging range, that is, whether or not a moving object exists within the imaging range. This moving object discrimination can be detected, for example, by examining the presence or absence of a difference in the content of the video signal at two points separated by a certain time. That is, when there is a difference between the contents of the two video signals, it is determined that an abnormality has occurred. It is also possible to determine that an abnormality has occurred when a pattern other than the predetermined pattern registered in advance appears in the content of the video.

[0022] As described above, by supplying the outputs of the video camera 2a, the microphone 2b, and the human body sensor 4 to the abnormality recognition circuit 5, it becomes possible to cope with various situations.

[0023] When the abnormality recognition circuit 5 determines that an abnormality has occurred, an abnormality detection signal is supplied from the abnormality recognition circuit 5 to the recording control circuit 6, and the first storage device 7 is switched from the recording mode to the reproduction mode, and Then, the recording operation of the second storage device 8 is started, and the information of the latest certain period recorded in the first storage device 7 is recorded in the second storage device 8. When the information recorded in the first storage device 7 is recorded in the second storage device 8, the first storage device 7 is again set to the recording mode, and the recording operation of the second storage device 8 is stopped.

[0024] By the above-described operation, every time an abnormal event occurs, the reproduction operation of the first storage device 7 for a certain time and the recording operation of the second storage device 8 for a certain time synchronized with this reproduction operation are repeated. Therefore, every time an abnormal event occurs, certain video information and audio information immediately before the occurrence of the abnormal event are sequentially recorded in the second storage device 8.

[0025] When it is determined by the abnormality recognition circuit 5 that an abnormality has occurred, the communication circuit 9 provides the abnormality occurrence data indicating that the abnormality has occurred together with the identification data indicating which monitor unit the data originated from. And transmitted to the centralized management system 11 installed in the security company via the communication line 10. The abnormality occurrence data is supplied to the control circuit 13 via the communication circuit 12, and the operation panel 17 displays where and what kind of abnormality has occurred. At the same time as the information recorded in the first storage device 7 is reproduced and recorded in the second storage device 8, the reproduced signal from the first storage device 7 is transmitted to the central management system 11 via the communication line 10. The video signal and the audio signal are restored to the original video signal and the audio signal by the decompression circuit 14 and supplied to the video output device 15 and the audio output device 16 to reproduce the video and the sound immediately before the occurrence of the abnormal event. After the information recorded in the first storage device 7 is reproduced, real-time information from the video camera 2a and the microphone 2b is transmitted automatically or based on an instruction from the operation panel 17. You may.

[0026] As a result, the security company can know the details of the abnormality, and can take appropriate measures such as dispatching a security guard or reporting to a police station in charge.

[0027] Also, by giving an activation instruction from the operation panel 17 to the recording control circuit 6 of the monitoring unit 1 other than the monitoring unit that has detected the abnormality, the other monitoring units

in the vicinity of the monitoring unit that has detected the abnormality can receive a start instruction. Also, the recording in the second storage device 8 is started, real-time information is transmitted via the communication line 10, and the centralized management system 11 can easily grasp the intruding route and the escape route of the intruder. it can.

[0028] Further, by giving an instruction to reproduce the second storage device 8 from the operation panel 17 to the recording control circuit 6 of the monitoring unit 1, the data is recorded in the second storage device 8 on the centralized management system 11 side. The status of abnormal events in the past can also be known. Note that a television monitor can be directly connected to the second storage device 8 of the monitoring unit 1 to reproduce the recorded contents of the second storage device 8.

[0029] In the above-described embodiment, the security device has been described as an example. However, the present invention is not limited to this. For example, the invention can be applied to monitoring of abnormal events in a chemical plant or a manufacturing plant. can do.

[0030] As described above, according to the present invention, the status of a monitoring target is always endlessly recorded in the first storage device, and when an abnormal event occurs, the reproduction operation is performed in the first storage device. Then, the reproduced signal is recorded in the second storage device capable of storing the data for a long time. Thus, the situation at the time of occurrence of the abnormal event can be reliably maintained without being limited by the recordable time of the first storage device that performs the endless recording.

[0031] BRIEF DESCRIPTION OF THE FIGURES

[0032] FIG. 1 is a block diagram showing an embodiment of a monitoring device according to the present invention.

[0033] Explanation of reference numerals

[0034] DESCRIPTION OF SYMBOLS 1 ... Surveillance unit, 2 ... Surveillance camera device, 2a ... Video camera, 2b ... Microphone, 3 ... Compression circuit, 4 ... Human body sensor, 5 ... Abnormality recognition circuit, 6 ... Recording control circuit, 7 ... First storage device, 8 .. 2nd storage device, 9 communication circuit, 10 communication line, 11 centralized management system, 12 communication circuit, 13 control circuit, 14 expansion circuit, 15 video display device, 16 audio display device, 17 …Operation board

JPH0877468A MONITOR DEVICE

Data originating from sources other than the EPO may not be accurate, complete, or up to date.

The wording below is an initial machine translation of the original publication. To generate a version using the latest translation technology, go to the original language text and use Patent Translate.

1. A video information detection device that detects video information to be monitored, an audio information detection device that detects audio information to be monitored, and a video signal from the video information detection device and an audio signal from the audio information detection device are compressed. A first storage device for recording an output of the compression circuit endlessly for a predetermined time, a second storage device for recording a reproduction signal from the first storage device and capable of storing the signal for a long period of time, An abnormality recognition circuit that analyzes the content of at least one of the video signal from the information detection device and the audio signal from the audio information detection device to detect the occurrence of an abnormal event; A recording control circuit that, when detected, sets the first storage device to a reproduction state for a certain period of time and sets the second storage device to a recording state during a reproduction period of the first storage device. Monitoring apparatus according to claim and.

(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10－162283

(43)公開日　平成10年(1998)６月19日

(51)Int.Cl.[6]	識別記号	ＦＩ	
Ｇ０８Ｇ　1/00		Ｃ０８Ｇ　1/00	Ｊ
Ｇ０１Ｓ　15/88		Ｃ０１Ｓ　15/88	
Ｇ０８Ｂ　21/00		Ｇ０８Ｂ　21/00	Ｕ
25/04		25/04	Ｃ
25/10		25/10	Ａ

審査請求　未請求　請求項の数８　ＯＬ　（全 ７ 頁）

(21)出願番号　特願平８－318197

(22)出願日　平成８年(1996)11月28日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　末吉　賢浩
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内

(72)発明者　深見　明久
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内

(72)発明者　片倉　景義
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内

(74)代理人　弁理士　小川　勝男

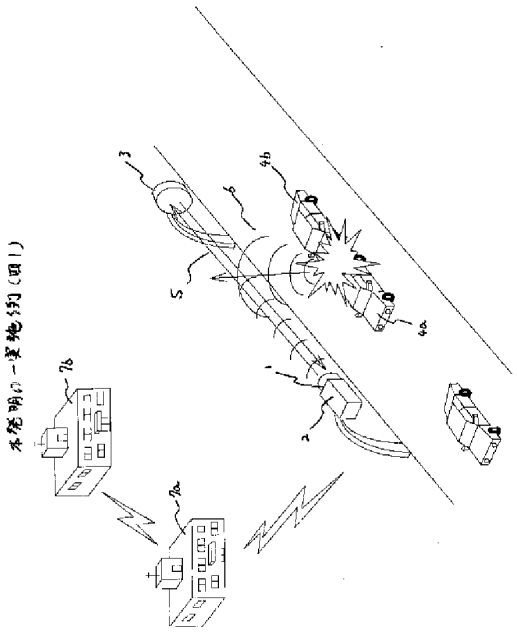
最終頁に続く

(54)【発明の名称】　道路状況監視装置

(57)【要約】

【課題】本発明は、道路上での車輌事故等の発生を天候、時間、視界の善し悪しに関係なく、また事故当事者及び目撃者の通報を待つことなく瞬時に警察及び消防署等へ通報すると同時に、後続車輌への警告を発することにより二次災害の発生を防止することを目的としている。

【解決手段】道路上方で車輌の通行を妨げない位置に設置し、空気中に音波を送波又は受波するパラメトリック送受波器と、道路上方に車輌の通行を妨げない位置でパラメトリック送受波器と対向して設置し、パラメトリック送受波器から放射された音波を反射する反射装置と、パラメトリック送受波器に送波信号を送信し、当該反射装置からの反射音波を受信処理して記録又は通報する制御装置とからなり、パラメトリック送受波器から周波数の僅かに異なる音波を送波し、反射装置により反射された音波が道路を走行中の車輌による異常音の影響で変調を受けているかどうかを制御装置で検出する構成とすることにより解決される。



【特許請求の範囲】
【請求項１】道路を走行中の車輌による異常音を検出し、車輌事故等の道路上の異常を監視する装置において、道路上方等で車輌の通行を妨げない位置に設置し、空気中に音波を送波又は受波する音波送受波器と、道路上方等に車輌の通行を妨げない位置で当該音波送受波器と対向して設置し、当該音波送受波器から放射された音波を反射する反射装置と、当該音波送受波器に送波信号を送信し、当該反射装置からの反射音波を受信処理して記録し、交通管制センタ等に通報する制御装置とから成り、当該音波送受波器から送波した音波が当該反射装置で反射されて帰ってくる間に車輌事故等により発生した異常音波と干渉して変調を受けているかどうかを当該制御装置で受信処理し、異常を検出したときは交通管制センタ等に通報することを特徴とした道路状況監視装置。
【請求項２】道路を走行中の車輌による異常音を検出し、車輌事故等の道路上の異常を監視する装置において、道路上方等で車輌の通行を妨げない位置に設置し、空気中に音波を送波又は受波する音波送受波器と、道路上方等に車輌の通行を妨げない位置で当該音波送受波器と対向して設置し、当該音波送受波器から放射された音波を反射する反射装置と、当該音波送受波器に送波信号を送信し、当該反射装置からの反射音波を受信処理して記録し、交通管制センタ等に通報する制御装置と、道路状況を車輌に知らせる情報表示装置とから成り、当該音波送受波器から送波した音波が当該反射装置で反射されて帰ってくる間に車輌事故等により発生した異常音波と干渉して変調を受けているかどうかを当該制御装置で受信処理し、異常を検出したときは交通管制センタ等に通報すると同時に、当該情報表示装置に後続の車輌に注意を促すよう表示させることを特徴とした道路状況監視装置。
【請求項３】道路を走行中の車輌による異常音を検出し、車輌事故等の道路上の異常を監視する装置において、道路上方で車輌の通行を妨げない位置に設置し、空気中に音波を送波する音波送波器と、当該音波送波器に送波信号を送信する送信制御装置と、当該音波送波器と対向して設置し、当該音波送波器から放射された音波を受波する音波受波器と、当該音波受波器の受波信号を処理して記録し、交通管制センタ等に通報する受信制御装置と、当該制御装置で検出した道路状況を表示する情報表示装置とから成り、当該音波送波器から送波した音波を当該音波受波器で受波する間に、車輌事故等により発生した異常音波と干渉して変調を受けているかどうかを当該制御装置で受信処理し、異常を検出したときは交通管制センタ等に通報すると同時に、当該情報表示装置に後続の車輌に注意を促すよう表示させることを特徴とした道路状況監視装置。
【請求項４】請求項１又は請求項２において前記音波送受波器は、例えば圧電セラミック等の電気音響変換素子を用い、周波数の僅かに異なる二種類の音波を送受波し、源周波数よりも周波数が低く且つビーム幅の狭い音波を発生させるパラメトリックアレイを用いることを特徴とした道路状況監視装置。
【請求項５】請求項３において、前記音波送波器及び音波受波器は、少なくともどちらか一方が請求項４記載のパラメトリックアレイを用いることを特徴とした道路状況監視装置。
【請求項６】請求項１、請求項２、請求項３のいずれかにおいて、前記制御装置及び受信制御装置は、請求項４及び請求項５記載のパラメトリックアレイにより受波した音響信号を周波数分析して車輌の衝突音、爆発音及び車輪の軋み音等の車輌事故等の道路通行上異常な音波に特徴的な周波数スペクトルパターンを検出することによりにより事故等の異常を検出、記録及び交通管制センタ等へ通報、情報表示装置に後続車輌に注意を促す表示をさせることを特徴とした道路状況監視装置。
【請求項７】請求項６において、制御装置は受波音響信号を一つ以上のバンドパスフィルタにより車輌事故等に発生する道路通行上異常な音波の周波数帯域に制限した後、予め登録しておいた異常音の信号波形と比較することにより事故等の異常を検出、記録及び交通管制センタ等へ通報、情報表示装置に後続車輌に注意を促す表示をさせることを特徴とした道路状況監視装置。
【請求項８】請求項７において、制御装置は受波音響信号の周波数パターンを記録・蓄積し、当該データを利用することにより、より精度良く異常の検出を可能とすることを特徴とした道路状況監視装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、道路走行における車輌による事故等の異常を検知し、迅速的、且つ自動的に交通管制センタ等へ通報し、また後続車輌への情報表示を行い二次災害発生を防止する為の道路状況監視装置に関するものである。
【０００２】
【従来の技術】従来、車輌の走行中に事故等が発生した場合、事故当事者又は目撃者が、例えば公衆電話機及び非常用電話機等の固定型電話機、車載用電話機及び携帯用電話機等の移動型電話機で警察署や消防署等へ通報するのが一般的である。しかしながら、固定型電話機で通報する場合、事故当事者や目撃者が設置されている場所へ移動するまで多少の時間を要すること、付近に設置されていない場合があること、目撃者が無く事故当事者が通報困難な場合などで事故発生と同時に通報できるとは限らない。移動型電話機においても同様のことが言える。この為事故の発見・通報が遅くなってしまい、負傷者救護の遅れ、事故発生を知らない他の車輌による二次災害発生の恐れがある。
【０００３】これを解決する為に、車輌に衝突事故が発

生したか否かを検知する衝突センサ、車輌の現在位置を検出する車輌位置センサ、車輌の運転者が自分の負傷状態を入力するスイッチ等で構成される入力装置及び通信機を含む車載事故通報装置と、交通管制センタと、警察署及び消防署を含む緊急連絡施設を含むシステムが検討されている。

【０００４】また、車輌に車載事故自動通報装置を携帯させ、事故発生を通報するシステムが考案されている。この例として、特開平８－３０８８５号公報が挙げられる。

【０００５】この方法によれば加速度検出手段、判別手段及び送信手段を１ユニットで構成できるので車輌への搭載が簡単に行える。また、高速道路の入口で貸し出せば全車輌に搭載することができ、万一事故が発生した場合でも警察等への通報、後続車輌への通報が事故発生とほぼ同時に可能となる。

【０００６】

【発明が解決しようとする課題】しかしながら、上記のように車輌にセンサを取り付ける場合、製造段階での取り付けが必要であり、車輌製造業者への負担が増加し、既存車輌への取り付けでは利用者の費用面での負担が増加する。また車輌の運転者が自分の負傷状態を入力するにしても、程度にもよるが、入力不可能である場合が多く、また事故により装置が故障してしまうことがある等の問題がある。

【０００７】車載事故自動通報装置を携帯させる場合は、事故により装置が故障し通報できない場合があること、利用者の負担が増加すること、貸し出す場合には、数量に制限ができてしまうこと、整備・再利用方法等の運用面では解決すべき問題が多々ある。

【０００８】本発明は、この様な種々の問題を解決し、道路上での車輌事故等の発生を天候、時間、視界の善し悪しに関係なく、また事故当事者及び目撃者の通報を待つことなく瞬時に警察及び消防署等へ通報すると同時に、後続車輌への警告を発することにより二次災害の発生を防止することを目的としている。

【０００９】

【課題を解決する為の手段】上記課題は、以下のような装置構成とすることによって解決される。

【００１０】（１）道路上方等で車輌の通行を妨げない位置に設置し、空気中に音波を送波又は受波する音波送受波器と、道路上方等に車輌の通行を妨げない位置で音波送受波器と対向して設置し、音波送受波器から放射された音波を反射する反射装置と、音波送受波器に送波信号を送信し、反射装置からの反射音波を受信処理して記録し、交通管制センタ等に通報する制御装置とで構成することにより解決される。

【００１１】（２）道路上方等で車輌の通行を妨げない位置に設置し、空気中に音波を送波又は受波する音波送受波器と、道路上方等に車輌の通行を妨げない位置で音波送受波器と対向して設置し、音波送受波器から放射された音波を反射する反射装置と、音波送受波器に送波信号を送信し、反射装置からの反射音波を受信処理して記録し、交通管制センタ等に迅速的且つ自動的に通報する制御装置と、道路状況を車輌に知らせる情報表示装置とで構成することにより解決される。

【００１２】（３）道路上方で車輌の通行を妨げない位置に設置し、空気中に音波を送波する音波送波器と、音波送波器に送波信号を送信する送信制御装置と、音波送波器と対向して設置し、音波送波器から放射された音波を受波する音波受波器と、音波受波器の受波信号を処理して記録し、交通管制センタ等に迅速的且つ自動的に通報する受信制御装置と、制御装置で検出した道路状況を表示する情報表示装置とで構成することにより解決される。

【００１３】（４）（１）及び（２）において音波送受波器は、例えば圧電セラミック等の電気音響変換素子を用い、周波数の僅かに異なる二種類の音波を送受波し、源周波数よりも周波数が低く且つビーム幅の狭い音波を発生させるパラメトリックアレイを用いる音波送受波器とすることで解決される。

【００１４】（５）（３）において音波送波器及び音波受波器は、少なくともどちらか一方にパラメトリックアレイを用いることで解決される。

【００１５】（６）（１）、（２）及び（３）において制御装置及び受信制御装置は、（４）及び（５）のパラメトリックアレイにより受波した音響信号を周波数分析して車輌の衝突音、爆発音及び車輪の軋み音等の車輌事故等の道路通行上異常な音波に特徴的な周波数スペクトルパターンを検出することによりにより事故等の異常を検出、記録及び交通管制センタ等へ迅速的且つ自動的に通報、情報表示装置に後続車輌に注意を促す表示をさせることで解決される。

【００１６】（７）（６）において制御装置は受波音響信号を一つ以上のバンドパスフィルタにより車輌事故等に発生する道路通行上異常な音波の周波数帯域に制限した後、予め登録しておいた異常音の信号波形と比較することにより事故等の異常を検出、記録及び交通管制センタ等へ迅速的か辻堂的に通報し、情報表示装置に後続車輌に注意を促す表示をさせることで解決される。

【００１７】（８）（７）において制御装置は受波音響信号の周波数パターンを記録・蓄積し、当該データを利用することにより、より精度良く異常の検出を可能とすることで解決される。

【００１８】

【発明の実施の形態】本発明の実施例を図を用いて以下説明する。

【００１９】図１は本発明における一実施例である。１はパラメトリック送受波器、２は送波信号の生成と受波信号の処理を行う制御装置、３はパラメトリック送受波

器１から送波された音波を反射する反射装置、４ａ及び４ｂは道路を走行する車両、５はパラメトリック送受波器１の送波音波伝搬経路、６は車輌４ａと４ｂが衝突したとき発生する音である。７ａと７ｂは、交通管制センタ、警察署等の緊急連絡施設である。

【００２０】ここで、パラメトリック送波器・受波器についてその原理を簡単に説明しておく。一般に二つの異なった周波数を持つ有限振幅音波を同時に同方向に放射した場合、各々の高調波成分だけでなく、非線形な相互作用によって差音、和音等たくさんの周波数成分が生まれ、例えば、二つの周波数をｆ１、ｆ２とすると、ｆ１，２ｆ１，・・・、ｆ２，２ｆ２，・・・、成分は勿論のこと、ｆ１－ｆ２，ｆ１＋ｆ２，２ｆ１－ｆ２，・・・、の多くの成分が生ずること、これらの成分のうちｆ１－ｆ２の差音の音響特性が極めて優れていることが知られている。この差音を利用する送受波器をパラメトリックアレイと言う。この音源の特徴は、（１）指向性が鋭くしかも副極が存在しないか、たとえあっても非常に小さい、（２）指向性は送波器の寸法に関係しないから送波器を小型軽量化できる、（３）差音は広帯域な周波数特性を持つ。（４）指向性は広い周波数帯域にわたってほぼ一定である、（５）源周波数よりも差音は非常に低いことから、送波器付近での音圧が小さくても差音は遠方まで伝搬する、等である。原理図を図２（ａ）に示す。８は電力増幅器、９は送波器、１０は送波ビームパターンである。

【００２１】パラメトリック受波器について説明すると、図２（ｂ）に示したように、送波器から放射される音波をｆｐ、周囲からの外来信号をｆｓとする。１１は送波器、１２は受波器、１３はフィルタ、１４は受波ビームパターンである。ｆｐの音波の振幅を大きく有限振幅の領域にしたとき、周波数ｆｐ－ｆｓの差音や、ｆｐ＋ｆｓの和音が現れる。差音や和音は信号波の情報を含むことから、信号波を受けなくても送波器に対向して設置した差音、和音を受け適切な処理を施せば信号波に対し指向性の優れた受波器ができる。

【００２２】一般に音波（振動）は、音波送波器の物理的寸法が同一の場合、周波数が高いほど指向性は狭くなるが伝搬距離が短く、周波数が低いと伝搬距離は長くなるが指向性が広くなってしまう。低い周波数で指向性を狭くするには物理的寸法が大きなものとなってしまう。

【００２３】本装置の場合、道路上方あるいは近傍への装置設置を考慮すると、物理寸法は小さくなければならないが、音波送受波器と反射装置の距離を可能な限り広くするためには音波伝搬距離を長くしなければならないと言う矛盾が生ずることになる。この問題を解決するためにはパラメトリック送受波器のように、送受波器の物理寸法は小さくても、差音により低い周波数を発生でき、指向性を狭くできる送受波器が必要になる。

【００２４】さて、図１の動作であるが、道路上方で車輌通行を妨害しない位置に設置した制御装置２で僅かに異なる周波数ｆ１と周波数ｆ２を生成し、同じく道路上方に設置したパラメトリック送受波器１により電気音響変換し空中に送波する。送波された音波は、伝搬経路５をたどり、また同じく道路上方でパラメトリック送受波器１と対向して設置した反射装置３で反射され伝搬経路５を逆にたどってパラメトリック送受波器１に到達する。パラメトリック送受波器１より送波された周波数ｆ１と周波数ｆ２の音波は、和周波数ｆ１＋ｆ２と差周波数ｆ１－ｆ２を形成するが、和周波数ｆ１＋ｆ２は周波数ｆ１と周波数ｆ２より大きい吸収を受けるのでほぼ無視できる。この差周波数ｆ１－ｆ２が伝搬経路５を往復する間に、車輌４ａに車輌４ｂが衝突したとする。このとき発生した異常音ｆ３による非線形音波６が空中を伝搬し、伝搬経路５を伝搬している差周波数ｆ１－ｆ２の音波に干渉して変調せしめ、これが反射音波となってパラメトリック送受波器１にて受波される。変調された差周波数音波は、（ｆ１－ｆ２）－ｆ３音波、（ｆ１－ｆ２）＋ｆ３音波としてパラメトリック送受波器１において受波され、制御装置２において受信処理され周波数スペクトルパターン分析、異常音データと比較され異常発生を検出し、制御装置２内の記録器に記録されると同時に交通管制センタ又は警察署や消防署等の緊急連絡施設へ迅速的且つ自動的に通報される。通報される情報には、差周波数音波から計算した事故発生地点の絶対位置情報が含まれるので、警察及び消防等への的確な情報提供が可能となる。

【００２５】道路上の異常を検知する方法として、道路側方にマイクロホンを並べておく方法も考えられるが、この場合、非常に多数のマイクロホンを設置する必要があり費用がかかるという問題がある。この点本発明の装置では、マイクロホンを多数並べる必要もなく、また道路上方へ設置することができるので、天候の影響、例えば積雪による埋没、等の影響を受けることが無く、また音波を利用するので視界の善し悪し及び時間に関係なく迅速的且つ自動的な通報が可能となる点で非常に優れている。

【００２６】図３は、カーブ等の見通しの悪い道路において、情報表示装置１５を用いた実施例である。４ｃは後続車輌である。

【００２７】上記図１の実施例のごとく、道路上方で車輌通行を妨害しない位置に設置した制御装置２で僅かに異なる周波数ｆ１と周波数ｆ２を生成し、同じく道路上方に設置したパラメトリック送受波器１により電気音響変換し空中に送波する。送波された音波は、伝搬経路５をたどり、また同じく道路上方でパラメトリック送受波器１と対向して設置した反射装置３で反射され伝搬経路５を逆にたどってパラメトリック送受波器１に到達する。パラメトリック送受波器１より送波された周波数ｆ１と周波数ｆ２の音波は、和周波数ｆ１＋ｆ２と差周波

数ｆ１－ｆ２を形成するが、和周波数ｆ１＋ｆ２は周波数ｆ１と周波数ｆ２より大きい吸収を受けるのでほぼ無視できる。この差周波数ｆ１－ｆ２が伝搬経路５を往復する間に、車輌４ａに車輌４ｂが衝突したとする。このとき発生した異常音ｆ３による非線形音波６が空中を伝搬し、伝搬経路５を伝搬している差周波数ｆ１－ｆ２の音波に干渉して変調せしめ、これが反射音波となってパラメトリック送受波器１にて受波される。変調された差周波数音波は、（ｆ１－ｆ２）－ｆ３音波、（ｆ１－ｆ２）＋ｆ３音波としてパラメトリック送受波器１において受波され、制御装置２において受信処理され周波数スペクトルパターン分析、異常音データと比較され異常発生を検出し、制御装置２内の記録器に記録されると同時に交通管制センタ又は警察署や消防署等の緊急連絡施設へ迅速的且つ自動的に通報される。通報される情報には、差周波数音波から計算した事故発生地点の絶対位置情報が含まれるので、警察及び消防等への的確な情報提供が可能となる。また本実施例の場合、情報表示装置１５に前方において事故等の異常があることを表示させ、後続の車輌４ｃ以後に対し、注意を促し４ａ、４ｂに追突する等の二次災害を防ぐことが可能となる。この情報表示装置は、既存の表示装置と兵横死手も構わないし、赤い警告灯が点灯するようにしてもよい。この場合情報表示装置の新規製造・設置の費用削減が可能である。

【００２８】図４は、パラメトリック送受波器１を送波器と受波器に分けた例である。１６は送信制御装置、１７はパラメトリック送波器、１８はパラメトリック受波器、１９は受信制御装置、２０は伝搬経路である。

【００２９】上記図３の実施例のごとく、道路上方で車輌通行を妨害しない位置に設置した送信制御装置１６で僅かに異なる周波数ｆ１と周波数ｆ２を生成し、同じく道路上方に設置したパラメトリック送波器１７により電気音響変換し空中に送波する。送波された音波は、伝搬経路２０をたどり、また同じく道路上方でパラメトリック送波器１７と対向して設置したパラメトリック受波器１８で受波される。パラメトリック受波器１８で受波された差周波数は受信制御装置１９で受信処理される。パラメトリック送波器１７より送波された周波数ｆ１と周波数ｆ２の音波は、和周波数ｆ１＋ｆ２と差周波数ｆ１－ｆ２を形成するが、和周波数ｆ１＋ｆ２は周波数ｆ１と周波数ｆ２より大きい吸収を受けるのでほぼ無視できる。この差周波数ｆ１－ｆ２が伝搬経路２０を伝搬する間に、車輌４ａに車輌４ｂが衝突したとする。このとき発生した異常音ｆ３による非線形音波６が空中を伝搬し、伝搬経路２０を伝搬している差周波数ｆ１－ｆ２の音波に干渉して変調せしめ、これが変調差周波数となってパラメトリック受波器１８にて受波される。変調された差周波数音波は、（ｆ１－ｆ２）－ｆ３音波、（ｆ１－ｆ２）＋ｆ３音波としてパラメトリック受波器１８において受波され、受信制御装置１９において受信処理され周波数スペクトルパターン分析、異常音データと比較され異常発生を検出し、受信制御装置１９内の記録器に記録されると同時に交通管制センタ又は警察署や消防署等の緊急連絡施設へ迅速的且つ自動的に通報される。通報される情報には、差周波数音波から計算した事故発生地点の絶対位置情報が含まれるので、警察及び消防等への的確な情報提供が可能となる。また本実施例の場合図３と同様に、情報表示装置１５に前方において事故等の異常があることを表示させ、後続の車輌４ｃ以後に対し、注意を促し４ａ、４ｂに追突する等の二次災害を防ぐことが可能となる。また、本実施例においては、送波系と受波系を分離し更に前記の実施例よりも小型軽量化が可能となっている。情報表示装置１５に対する情報の表示指示・制御は、交通管制センタ等からでもよいが、受信制御装置１９から指示・制御をするようにすることも可能である。

【００３０】本発明における装置では各制御装置、送波器・受波器はいずれも非常に小型軽量化できるので、道路側方に充分な設置空間が確保できない場合でも、道路上方、走行中の車輌の上空に設置できるので何等問題が無い。

【００３１】上記のいずれの実施例においても、送波周波数を任意に変更可能とすることで、環境等の状況の変化に対応した最適な設定とすることができ、かつ周波数帯域幅を自由に変化させることができるので、検出対象の異常音を限定することや広い範囲を一度に検出するなど検出条件設定に自由度を持たせることが可能である。

【００３２】図５に示すように、道路上で発生する音波を周波数分析すると、通常走行音、衝突音及び車輪の軋み音は各々異なる特徴的な周波数スペクトルパターンを有している。これに着目し道路上の種々の音波の中から異常音を検出するように、パラメトリック受波器で受波した音波をバンドパスフィルタに通し、予め登録しておいた衝突音等の周波数スペクトルパターンと比較して検出するよう受波制御装置を構成している。

【００３３】また、受波音波から得た特徴的な周波数スペクトルパターンは、記録器にて記録データ蓄積を行うようにすることで次に比較検出するときの検出精度向上、異常判定基準の多様化にも対応可能な装置構成とすることができる。

【００３４】

【発明の効果】以上述べたように、本発明によれば、道路上での車輌事故等の発生を天候、時間、視界の善し悪しに関係なく、また事故当事者及び目撃者の通報を待つことなく瞬時に警察及び消防署等へ通報すると同時に、後続車輌への警告を発することにより二次災害の発生を防止する効果がある。

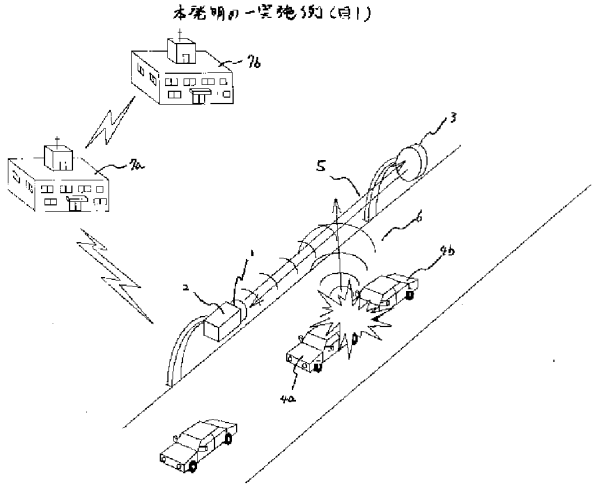
【図面の簡単な説明】

【図１】本発明における一実施例。

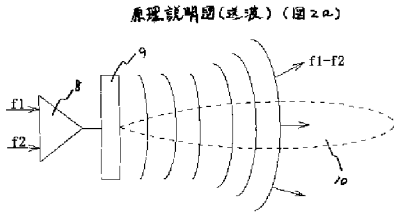
【図２】パラメトリック送波器。受波器の原理説明図。

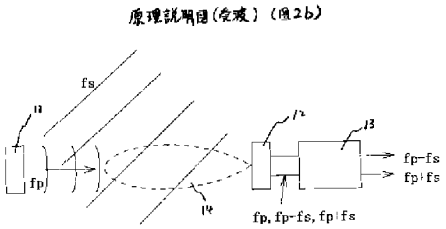
【図３】本発明における一実施例。
【図４】本発明における一実施例。
【図５】周波数スペクトルパターンの例。
【符号の説明】
１・・・パラメトリック送受波器
２・・・制御装置
３・・・反射装置
４ａ、４ｂ、４ｃ・・・車輌
５・・・送波音波往復伝搬経路
６・・・車輌衝突時の発生音
７ａ、７ｂ・・・緊急連絡施設
８・・・電力増幅器
９・・・送波器
１０・・・送波ビームパターン
１１・・・送波器
１２・・・受波器
１３・・・フィルタ
１４・・・受波ビームパターン
１５・・・情報表示装置
１６・・・送信制御装置
１７・・・パラメトリック送波器
１８・・・パラメトリック受波器
１９・・・受信制御装置
２０・・・送波音波伝搬経路
２１・・・通常走行音の周波数スペクトルパターン
２２・・・衝突音の周波数スペクトルパターン
２３・・・車輪軋み音の周波数スペクトルパターン

【図１】

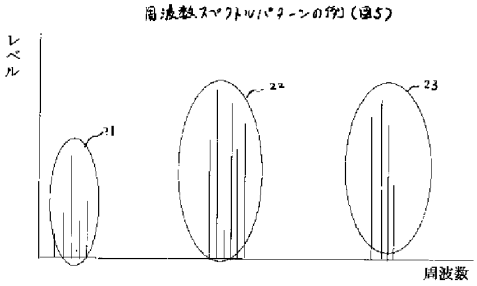


【図２】

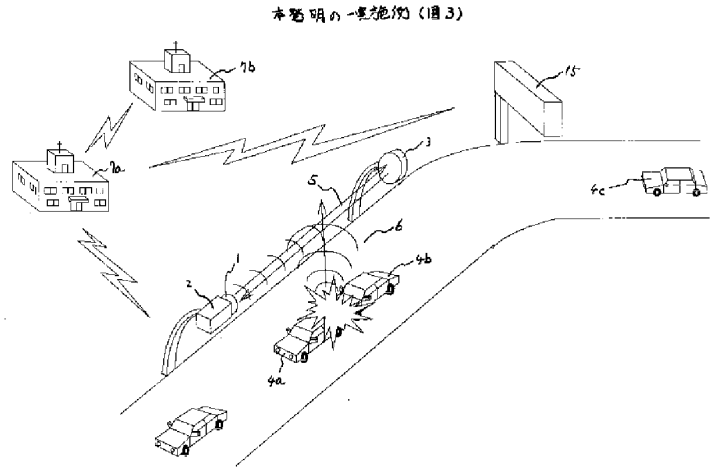




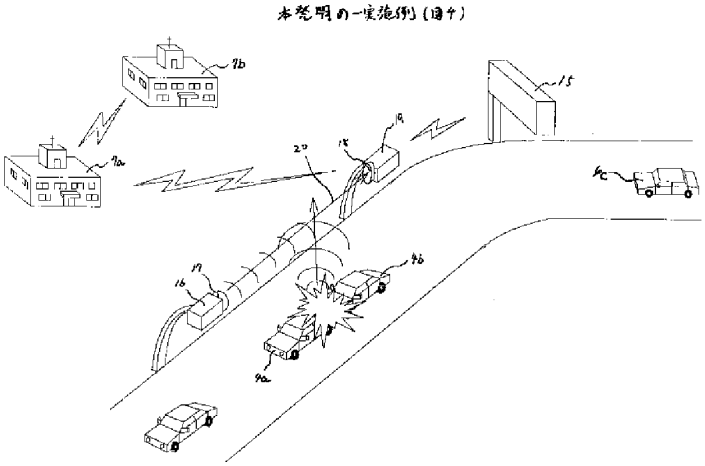
【図５】



【図３】



【図４】



フロントページの続き

(72)発明者　長山　暁司
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内
(72)発明者　神田　義信
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内
(72)発明者　船崎　好助
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内
(72)発明者　増田　俊也
神奈川県横浜市戸塚区戸塚町216番地株式会社日立製作所情報通信事業部内

Espacenet

Bibliographic data: JPH10162283 (A) — 1998-06-19

ROAD CONDITION MONITORING DEVICE

Inventor(s): SUEYOSHI MASAHIRO; FUKAMI AKIHISA; KATAKURA KAGEYOSHI; NAGAYAMA AKIJI; KANDA YOSHINOBU; FUNAZAKI YOSHISUKE; MASUDA TOSHIYA ± (SUEYOSHI MASAHIRO, ; FUKAMI AKIHISA, ; KATAKURA KAGEYOSHI, ; NAGAYAMA AKIJI, ; KANDA YOSHINOBU, ; FUNAZAKI YOSHISUKE, ; MASUDA TOSHIYA)

Applicant(s): HITACHI LTD ± (HITACHI LTD)

Classification: **- international:** ***B60W30/00; G01S15/88; G08B21/00; G08B25/04; G08B25/10; G08G1/00;*** (IPC1-7): G01S15/88; G08B21/00; G08B25/04; G08B25/10; G08G1/00

- cooperative:

Application number: JP19960318197 19961128

Priority number(s): JP19960318197 19961128

Abstract of JPH10162283 (A)

PROBLEM TO BE SOLVED: To instantaneously report the occurrence of car accident to a police and a firehouse and to prevent the occurrence of secondary disaster by simultaneously issuing an alarm to a following vehicle by providing a reflector for reflecting sonic waves and a controller for receiving the reflected sonic waves, recording these sonic waves and reporting them to a traffic control center or the like. SOLUTION: Frequencies f1 and f2 different a little are generated by a controller 2 installed above a road, and sent into air after performing electric/ acoustic transduction through a parametric transmitter/receiver 1. On the assumption that a vehicle 4b collides with a vehicle 4a while this differential frequency f1-f2 is reflected by a reflector 3 and reciprocated through a propagation path 5, a non-linear acoustic wave 6 caused by an abnormal sound f3 generated at such a time is propagated in air, the sonic wave with differential frequency f1-f2 propagated through the propagation path 5 is interferred and modulated, and this sonic wave is turned to a reflected sonic wave and received by the parametric transmitter/receiver 1. Then, the generation of abnormality is detected by a controller 2, recorded and speedily and automatically reported to emergency contact facilities 7a and 7b at the same time.

JPH10162283A ROAD CONDITION MONITORING DEVICE

Data originating from sources other than the EPO may not be accurate, complete, or up to date.

The wording below is an initial machine translation of the original publication. To generate a version using the latest translation technology, go to the original language text and use Patent Translate.

[0001] The present invention relates to is, detects an abnormality such as an accident caused by a vehicle in the road traveling, rapid manner, and automatically Problem to the traffic control center or the like, also two performs a display of information to a subsequent vehicle order those related to road condition monitoring apparatus for preventing accidents.

[0002] Conventionally, when the accident or the like during the running of the vehicle has occurred, accident parties or witnesses, for example, public phone and fixed phone of emergency phones, and the like, in-vehicle phone and mobile, such as a portable phone it is general to report to the police station and fire station and the like on the phone. However, if the report at a fixed phone, it takes some time before moving to the location where the accident parties or witnesses are installed, that it may not be installed near, accident parties without witnesses not always possible to report the accident and at the same time, such as when the report difficult. Similarly, in the mobile phone it can be said. Because of this slows down the discovery and reporting of the accident, there is a risk of secondary disaster caused by other vehicles do not know delay of injured relief, the accident occurred.

[0003] To solve this, composed of a switch or the like collision sensor collision in the vehicle for detecting whether or not occurs, the vehicle position sensor for detecting the current position of the vehicle, the driver of the vehicle to enter their wounded a vehicle accident notification apparatus comprising an input device and a communication device is a traffic control center, the system including emergency facilities, including police and fire departments have been studied.

[0004] In addition, to carry a vehicle accident automatic reporting device in the vehicle, it has been devised system to report the accident. Examples of this include Hei 8-30885 JP.

[0005] Acceleration detection means According to this method, mounted to the vehicle can be carried out easily because it constitutes a discrimination means and transmitting means in one unit. Also, Kashidase at the inlet of the highway can be mounted to all vehicles, emergency report to the police or the like even when an accident occurs, reporting to a subsequent vehicle is substantially simultaneously possible with the accident.

[0006] The present invention is, however, if you are installing a sensor in the vehicle as described above, it is necessary to mounting at the manufacturing stage, and increase the burden on the vehicle manufacturer, in the attachment to an existing vehicle users burden on the cost side of the increases. Also even if the driver of the vehicle to enter their wounded, depending on the degree, there is a problem that often is not possible input and device by accident which may cause malfunctioning.

[0007] Case of mobile-vehicle accident automatic reporting device is that it may device by accident can not report a failure, the burden on the user is increased, when lending is that she can limit the quantity, re-establish and operational aspects such as how to use is there are many problems to be solved.

[0008] The present invention, such to solve the various problems, police the occurrence of a vehicle accident on the road weather, time, regardless of the field of view of the quality, also instantly without waiting for the report of the accident parties and witnesses and at the same time report to the fire station or the like, and aims to prevent the occurrence of secondary disasters by issuing a warning to following vehicles.

[0009] The above-mentioned problems, according to means to solve the problem] is solved by the device configuration, such as the following.

[0010] (1) was placed in a position that does not interfere with vehicle traffic in the road upper or the like, a sonic transducer for transmitting or reception waves in the air, sonic wave transceiver at a position that does not interfere with the passage of vehicles on the road upward like vessels facing to be installed, a reflector for reflecting the sound waves emitted from the ultrasonic transducer transmits a transmit signal into sound waves

transducer and recorded by the reception processing the reflected sound wave from the reflecting device, It is solved by configuring by the controller to notify the traffic control center or the like.

[0011] (2) it was placed in a position that does not interfere with vehicle traffic in the road upper or the like, a sonic transducer for transmitting or reception waves in the air, sonic wave transceiver at a position that does not interfere with the passage of vehicles on the road upward like vessels facing to be installed, a reflector for reflecting the sound waves emitted from the ultrasonic transducer transmits a transmit signal into sound waves transducer and recorded by the reception processing the reflected sound wave from the reflecting device, a control device for Problem rapid manner and automatically to a traffic control center or the like, are solved by constructing in an information display device to inform the road condition to the vehicle.

[0012] (3) was placed in a position that does not interfere with vehicle traffic in the road upper, a sonic wave transmitter for transmitting a sound wave into the air, and a transmission control unit for transmitting a transmit signal to the sonic wave transmitter, ultrasonic wave transmission vessels facing to be installed, and ultrasonic wave receiver for reception of sound waves emitted from the ultrasonic wave transmitter, and recorded by processing the received signals of the sonic receivers, rapid manner and the traffic control center or the like a reception control device that automatically notify, is solved by configuring in an information display device for displaying the road conditions detected by the control device.

[0013] (4) (1) and ultrasonic transducer (2), for example using an electro-acoustic transducer of a piezoelectric ceramic or the like, a slightly different two kinds of waves of frequency transmission and reception waves, the source lower than the frequency and it is solved by a sound wave transducer using a parametric array to generate a narrow sound wave of the beam width.

[0014] (5) sonic wave transmitter and ultrasonic wave receiver (3) it is solved by using a parametric array in at least one.

[0015] (6) (1), (2) and the control apparatus and the reception control unit (3) is, (4) and (5) vehicle collision noise acoustic signal by performing frequency analysis with reception by parametric array of explosion detecting an abnormality such as an accident by by detecting the sound and the wheels of the creaking sound such as a vehicle accident such as road traffic on the unusual waves characteristic frequency spectrum pattern of, rapid manner and automatically to the recording and the traffic control center or the like report, be resolved by causing the attention displayed in the following vehicle in the information display device.

[0016] (7) After the controller is limited to the frequency band of the road traffic on abnormal waves generated in a vehicle accident or the like by one or more band-pass filters the reception acoustic signal (6), previously registered abnormal detecting an abnormality such as an accident by comparing the signal waveform of sound, it is resolved by the recording and transportation quick manner or Tsujido to Problem to the control center or the like, be the attention displayed in the subsequent vehicle on the information display device .

[0017] (8) (7) the controller will record and store the frequency pattern of the received wave an acoustic signal in, by utilizing the data, it is solved by allowing a more accurate abnormality detection.

[0018] It is described below with reference to FIG the embodiment of the embodiment of the present invention.

[0019] FIG. 1 is an embodiment of the present invention. 1 parametric transducer, 2 control unit for the processing of generation and received signal of the transmit signal, 3 reflector for reflecting the sound waves transmitted from a parametric transducer 1, 4a and 4b the road traveling vehicle, the transmitting sound wave propagation path of parametric transducer 1 5, 6 is the sound generated when the vehicle 4a and 4b has collided. 7a and 7b, the traffic control center, an emergency contact facilities such as police stations.

[0020] Here, a brief description of the principle for parametric wave transmitter-receivers. In general, when the radiation at the same time in the same direction the finite amplitude waves with different frequencies of the two, not only the harmonic components of each, born difference tone, chord, etc. a lot of frequency components by non-linear interaction, for example, of two If the frequency is f1, f2, f1,2f1, ⋯, f2,2f2, ⋯, component is of course that, f1-f2, f1 + f2,2f1-f2, ⋯, many of the components of the resulting it, acoustic characteristics of the difference tone of f1-f2 of these components are known to be extremely excellent. A transducer to take advantage of this difference tone say parametric array. The feature of this sound source is, (1) or sharp directivity And Vice pole does not exist, even if very small, even, (2) small size and light weight of the wave transmitter since the directivity is not related to the size of the wave transmitter possible reduction, (3) the

difference tone will have a wide-band frequency characteristics. (4) directivity is approximately constant over a wide frequency band, (5) from the difference tone is very low than the source frequency, the difference tone even with a small sound pressure in the vicinity of the wave transmitter is propagated to distant , and the like. The principle diagram shown in FIG. 2 (a). 8 power amplifier, 9 wave transmitter, is 10, which is a transmit beam pattern.

[0021] Referring to the parametric wave receiver, as shown in FIG. 2 (b), the sound waves radiated from the wave transmitter fp, and extraneous signals from the surrounding fs. 11 wave transmitter, 12 receivers, 13 filters, 14 is a receiving beam pattern. When in the area of large finite amplitude of the amplitude of the sound waves of fp, difference tone and the frequency fp-fs, chord of fp + fs appears. Difference tone and chords from it, including the information of the signal wave, excellent difference tone even without receiving a signal wave was set to face the wave transmitter, for the signal wave if Hodokose an appropriate process received a chord of directivity the wave receiver can be.

[0022] Are generally acoustic wave (vibration), if the physical dimensions of the sonic wave transmitter is the same, the directivity higher frequency is shorter propagation distance narrows, the propagation distance frequency is low is but wider directivity long and will. To narrow the directivity in the low frequency becomes physical dimension is a big thing.

[0023] In the case of the present system, in view of the device installation to the road above or near, the physical dimensions are not should be small, to lengthen the sound wave propagation distance in order to widen as much as possible the distance of the reflecting device and ultrasonic transducer so that the inconsistency to say there must be caused. As in parametric transducer in order to solve this problem, physical dimensions of the transducer be small, can generate a lower frequency by the difference tone, it is necessary to transducer capable of narrow directivity.

[0024] Now, the operation of FIG 1, to produce a slightly different frequencies f1 and f2 by a control device 2 installed in a position that does not interfere with vehicle traffic on the road above, likewise the parametric transducer 1 installed in a road upward electro-acoustic transducer and to transmit in the air. Transmitting sound waves will follow the propagation path 5, also similarly reach the propagation path 5 is reflected by the parametric transducer 1 and the opposing reflecting device 3 was placed on the road upward traced back to the parametric transducer 1 to. Almost ignored because waves of frequencies f1 and f2 that has been transmitting from parametric transducer 1, but to form a sum frequency f1 + f2 and the difference frequency f1-f2, sum frequency f1 + f2 is subjected to greater absorption frequencies f1 and f2 it can. During this difference frequency f1-f2 is a round trip propagation paths 5, the vehicle 4b has collided with the vehicle 4a. At this time, non-linear wave 6 is propagated in the air due to the abnormal sound f3 generation, interfere with the sound waves propagation difference frequency that route is 5 propagates through the f1-f2 is allowed to modulation, parametric transducer 1 this becomes a reflected sound wave It is reception at. Modulated difference frequency sound waves, (f1-f2) -f3 waves, (f1-f2) + f3 is reception in parametric transducer 1 as a sound wave, receiving the processed frequency spectral pattern analysis in the control unit 2, abnormal sound data detecting a comparison is abnormal occurrence and, is notified rapid manner and automatically to the emergency contact facilities such as traffic control center or police station and fire station at the same time when it is recorded on the recording device in the control device 2. The information is notified, because it contains the absolute position information of the accident point was calculated from the difference frequency wave, it is possible to accurately provide information to the police and fire the like.

[0025] As a method for detecting an abnormality on the road, a method may be considered to be side-by-side microphone on the road side, in this case, there is a problem in that a very large number of the microphone must be installed expensive. In apparatus of this point the present invention, there is no need to arrange a large number of microphones, and because it can be installed into road upward, the influence of the weather, for example, buried by snow, without being affected etc., also utilize sonic It has a very good point that it is possible to quickly specific and automatic report regardless of the quality and time of sight since.

[0026] 3, in the road with bad visibility, such as curves, an example in which using the information display device 15. 4c is a subsequent vehicle.

[0027] As indicated above Embodiment FIG 1, to produce a slightly different frequencies f1 and f2 by a control device 2 installed in a position that does not interfere with vehicle traffic on the road above, also electricity through parametric transducer 1 installed in a road upward to transmit in the air and acoustic conversion. Transmitting sound waves will follow the propagation path 5, also similarly reach the propagation path 5 is

reflected by the parametric transducer 1 and the opposing reflecting device 3 was placed on the road upward traced back to the parametric transducer 1 to. Almost ignored because waves of frequencies f1 and f2 that has been transmitting from parametric transducer 1, but to form a sum frequency f1 + f2 and the difference frequency f1-f2, sum frequency f1 + f2 is subjected to greater absorption frequencies f1 and f2 it can. During this difference frequency f1-f2 is a round trip propagation paths 5, the vehicle 4b has collided with the vehicle 4a. At this time, non-linear wave 6 is propagated in the air due to the abnormal sound f3 generation, interfere with the sound waves propagation difference frequency that route is 5 propagates through the f1-f2 is allowed to modulation, parametric transducer 1 this becomes a reflected sound wave It is reception at. Modulated difference frequency sound waves, (f1-f2) -f3 waves, (f1-f2) + f3 is reception in parametric transducer 1 as a sound wave, receiving the processed frequency spectral pattern analysis in the control unit 2, abnormal sound data detecting a comparison is abnormal occurrence and, is notified rapid manner and automatically to the emergency contact facilities such as traffic control center or police station and fire station at the same time when it is recorded on the recording device in the control device 2. The information is notified, because it contains the absolute position information of the accident point was calculated from the difference frequency wave, it is possible to accurately provide information to the police and fire the like. In the case of this example, to display that the information display device 15 is abnormal such as an accident at the front, with respect to the subsequent vehicle 4c thereafter, 4a attention, to prevent secondary disasters such as rear-end to 4b it is possible. The information display device, to may be an existing display device and Heiyoko death hand, may be a red warning light is illuminated. Cost savings of the new production and installation of this case information display device are possible.

[0028] Figure 4 is an example of dividing the parametric transducer 1 to the wave transmitter and wave receiver. 16 transmission control device, 17 parametric wave transmitter, 18 parametric receivers, 19 receiving control device, 20 is a propagation path.

[0029] As examples of FIG 3, to produce a slightly different frequencies f1 and f2 in the transmission control apparatus 16 which is installed at a position that does not interfere with vehicle traffic on the road above, likewise the parametric wave transmitter 17 installed in the road upward electro-acoustic transducer and to transmit in the air. Transmitting sound waves follows a propagation path 20, and is received at the parametric wave receiver 18 also installed to face the parametric wave transmitter 17 on the road above. Difference frequency that is received at the parametric receivers 18 are receiving process by the reception control unit 19. Almost ignored because waves of frequencies f1 and f2, which is transmitting from the parametric wave transmitter 17 is to form the sum frequency f1 + f2 and the difference frequency f1-f2, sum frequency f1 + f2 is subjected to greater absorption frequencies f1 and f2 it can. During this difference frequency f1-f2 is propagating in the propagation path 20, the vehicle 4b has collided with the vehicle 4a. At this time, non-linear wave 6 is propagated in the air due to the abnormal sound f3 generation, interfere with the sound waves propagation difference frequency of the route 20 propagating f1-f2 is allowed to modulation, parametric receivers this becomes a modulation difference frequency 18 is reception at. Modulated difference frequency sound waves, (f1-f2) -f3 waves, (f1-f2) + f3 waves are received wave in parametric wave receiver 18 as, receiving the processed frequency spectral pattern analysis in the receiving control unit 19, abnormal sound detecting a comparison is abnormal occurrence and data, it is notified rapid manner and automatically to the emergency contact facilities such as traffic control center or police station and fire station at the same time when it is recorded in a recording device of the receiving control device 19. The information is notified, because it contains the absolute position information of the accident point was calculated from the difference frequency wave, it is possible to accurately provide information to the police and fire the like. Also in the same manner as in the case Figure 3 of the present example, to display that the information display device 15 is abnormal such as an accident at the front, with respect to the subsequent vehicle 4c thereafter, 4a attention, such as the rear-end collision to 4b two it is possible to prevent the following accident. In the present embodiment, it becomes possible to reduce the size and weight than the further above embodiment separates the transmitting system and the reception system. Display instruction and control information to the information display device 15, may be any from the traffic control center or the like, can also be adapted to the instructions and control of the reception control unit 19.

[0030] Each control device in the apparatus of the present invention, since none of the wave transmitter-receivers can be very compact and lightweight, even if a sufficient installation space on the road side can not be secured, the road above, the sky of the vehicle while driving there is no what problems such as so can be installed in.

[0031] In either embodiment described above, by making it possible to arbitrarily change the transmit frequency can be the optimum setting corresponding to the changing conditions in the environment, and changing the

frequency bandwidth free since it is, it is possible to provide a degree of freedom in such detection condition setting detects that and wide range to limit the abnormal sound to be detected at a time.

[0032] As shown in FIG. 5, when the waves generated on the road to frequency analysis, usually running sound, squeak the collision sound and the wheel has a respective different characteristic frequency spectrum pattern. This to detect abnormal sounds from various sound wave on the road by focusing, through a sound wave received at the parametric wave receiver to the bandpass filter, the frequency spectrum pattern such as collision sound that has been previously registered constitute the reception control device to detect compared to the.

[0033] Further, the characteristic frequency spectrum pattern obtained from the reception wave, the detection accuracy at the time of next compare detected by to perform the recording data storage in a recording device, to the diversification of the abnormality judgment reference corresponding can allow device configurations.

[0034] As described above, according to the present invention, according to the present invention, the occurrence of the vehicle accident on the road weather, time, irrespective of the field of view of quality, and without waiting for the notification of an accident parties and witnesses instant and at the same time to report to the police and fire department, etc., there is an effect to prevent the occurrence of secondary disasters by issuing a warning to the following vehicle.

[0035] Brief description of the drawings

[0036] One embodiment in FIG present invention.

[0037] Figure 2 parametric wave transmitter. Explanatory view of the principle of receivers.

[0038] One embodiment shown in FIG. 3 the present invention.

[0039] One embodiment shown in FIG. 4 the present invention.

[0040] The example of FIG. 5 frequency spectrum pattern.

[0041] Description of the code

[0042] 1 ... parametric transducer 2 ... control device 3 ... reflecting device 4a, 4b, 4c ··· vehicle 5 ... transmit sound waves round-trip propagation path 6 ... vehicle collision at the time of the generated sound 7a , 7b ··· emergency contact facility 8 ... power amplifier 9 ... wave transmitter 10 ... transmission beam pattern 11 ... wave transmitter 12 ... wave-receiver 13 ... filter 14, ... reception beam pattern 15 ... information display device 16 ... transmission control device 17 ... parametric wave transmitter 18 ... parametric receivers 19 ... reception control device 20 ... transmitting sound waves frequency spectrum pattern of the propagation path 21 ... usually frequency spectrum pattern of the frequency spectrum pattern 22 ... collision sound of running sound 23 ... the wheels squeak

JPH10162283A ROAD CONDITION MONITORING DEVICE

Data originating from sources other than the EPO may not be accurate, complete, or up to date.

The wording below is an initial machine translation of the original publication. To generate a version using the latest translation technology, go to the original language text and use Patent Translate.

1. Detects an abnormal sound due to vehicles traveling on a road, in an apparatus for monitoring abnormality on the road, such as a vehicle accident, was placed in a position that does not interfere with vehicle traffic in the road upward like, transmitting or sound waves in air a sonic transducer which received wave, at a position that does not interfere with the passage of vehicles on the road upward like disposed opposite to the said ultrasonic transducer, a reflector for reflecting the sound waves emitted from the ultrasonic transducer, It sends a transmit signal to the ultrasonic transducer, and recorded receives and processes the reflected sound wave from the reflecting device, consists of a control unit to report to the traffic control center or the like, was transmitted from the ultrasonic transducer traffic control center when whether interfering with the abnormal sound waves generated by the vehicle accidents are receiving the modulated reception processing in the control unit, detects an abnormality while the sound wave comes back after being reflected in the reflecting device road conditions monitoring device is characterized in that report to equal.

2. Detects an abnormal sound due to vehicles traveling on a road, in an apparatus for monitoring abnormality on the road, such as a vehicle accident, was placed in a position that does not interfere with vehicle traffic in the road upward like, transmitting or sound waves in air a sonic transducer which received wave, at a position that does not interfere with the passage of vehicles on the road upward like disposed opposite to the said ultrasonic transducer, a reflector for reflecting the sound waves emitted from the ultrasonic transducer, It sends a transmit signal to the ultrasonic transducer, from recorded receives and processes the reflected sound wave from the reflecting device, and a control device to report to the traffic control center or the like, and the information display device to inform the road conditions in the vehicle becomes, the reception processing whether sound waves transmitted from the ultrasonic transducer is undergoing modulation interferes with abnormal waves generated by the vehicle accident while coming back after being reflected in the reflecting device in the control device and, when detecting abnormality when Problem to traffic control center or the like at the same time, the information display device to the road condition monitoring apparatus characterized thereby alert display on subsequent vehicles.

3. Waves road to detect the abnormal sound by the vehicle while driving, in an apparatus for monitoring the abnormality on the road, such as a vehicle accident, which was placed in a position that does not interfere with the vehicle of the traffic on the road above, to transmit sound waves in the air a wave transmitter, a transmission control unit for transmitting a transmit signal to the sonic wave transmitter, and disposed to face the said ultrasonic wave transmitter, sound waves received for reception of sound waves emitted from the ultrasonic wave transmitter a duplexer, and recorded by processing the reception signal of the ultrasonic wave receiver, consists of a receiving control device to notify the traffic control center or the like, an information display device for displaying the road conditions detected by the control device, the sound waves transmitted from the ultrasonic wave transmitter during the reception in the ultrasonic wave receiver, whether interfering with the abnormal sound waves generated by the vehicle accidents are receiving the modulated reception processing in the control unit, At the same time when detecting an abnormality is reported to the traffic control center or the like, road condition monitoring apparatus characterized thereby alert display subsequent vehicle to the information display device.

4. The acoustic wave transducer according to claim 1 or claim 2, for example using an electro-acoustic transducer of a piezoelectric ceramic or the like, a slightly different two kinds of waves of frequency transmission and reception waves, the source lower than the frequency and beam road condition monitoring apparatus characterized by using a parametric array for generating a narrow wave width.

5. Device according to claim 3, wherein the ultrasonic wave transmitter and ultrasonic wave receiver is traffic condition monitoring apparatus at least one is characterized by using a parametric array according to claim 4.

6. Claim 1, Claim 2, in any one of claims 3, wherein the control device and the reception control device, the vehicle collision by the frequency analyzing acoustic signals received wave by claim 4 and claim 5 parametric array according sound, detecting an abnormality such as an accident by by detecting the characteristic frequency spectrum pattern to road traffic on the abnormal waves such as a vehicle accident of creaking sound, etc. of explosion sound and the wheels, report to the recording and the traffic control center or the like, road conditions

monitoring device is characterized in that the information display device to the attention displayed in the subsequent vehicle.

7. In claim 6, the control device after limiting the frequency band of the road traffic on abnormal waves generated in a vehicle accident or the like by one or more band-pass filter a receive acoustic signals, the abnormal sound is registered in advance detecting an abnormality of an accident such as by comparing the signal waveform, recording and traffic control Problem to center or the like, the information display device to the road condition monitoring apparatus characterized in that to the attention displayed on subsequent vehicles.

8. Method according to claim 7, the control unit will record and store the frequency pattern of the received wave an acoustic signal, by utilizing the data, road condition monitoring apparatus which is characterized in that to enable more accurate abnormality detection.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 December 2004 (29.12.2004)

PCT

(10) International Publication Number
WO 2004/114722 A1

(51) International Patent Classification[7]: H04R 25/00

(21) International Application Number:
PCT/DK2004/000442

(22) International Filing Date: 23 June 2004 (23.06.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2003 00944 24 June 2003 (24.06.2003) DK

(71) Applicant *(for all designated States except US)*: **GN RESOUND A/S** [DK/DK]; Maarkaervej 2A, DK-2630 Taastrup (DK).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **PEDERSEN, Brian, Dam** [DK/DK]; Kildebrinken 13, DK-4100 Ringsted (DK).

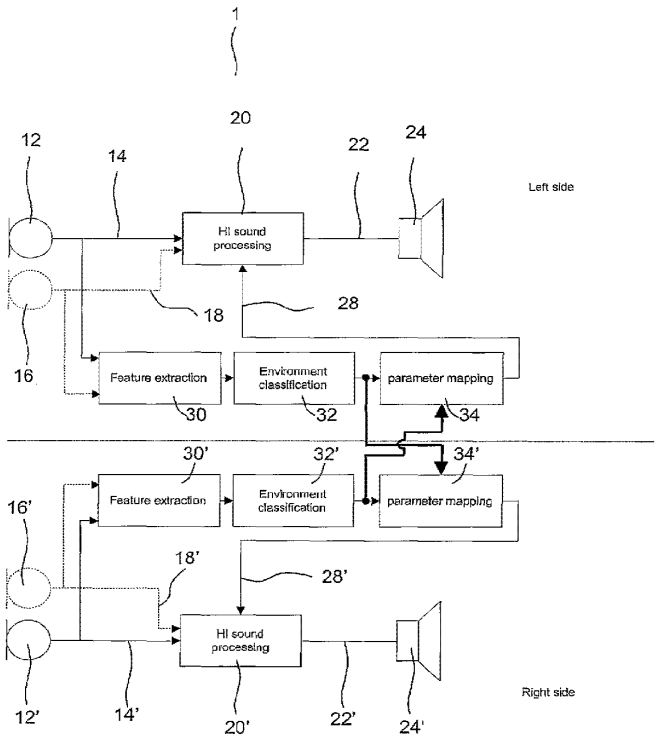
(74) Agent: **JAKOBSEN, Gert, Høy**; Albihns A/S, H.C. Andersens Boulevard 49, DK-1553 Copenhagen (DK).

(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, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, 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, IT, LU, MC, NL, PL, PT, RO, SE, SI,

[Continued on next page]

(54) Title: A BINAURAL HEARING AID SYSTEM WITH COORDINATED SOUND PROCESSING



(57) Abstract: The present invention relates to a binaural hearing aid system comprising a first hearing aid and a second hearing aid, each of which comprises a microphone and an A/D converter for provision of a digital input signal in response to sound signals received at the respective microphone in a sound environment, a processor that is adapted to process the digital input signals in accordance with a predetermined signal processing algorithm to generate a processed output signal, and a D/A converter and an output transducer for conversion of the respective processed sound signal to an acoustic output signal, and a binaural sound environment detector for binaural determination of the sound environment surrounding a user of the binaural hearing aid system based on at least one signal from the first hearing aid and at least one signal from the second hearing aid for provision of outputs for each of the first and second hearing aids for selection of the signal processing algorithm of each of the respective hearing aid processors so that the hearing aids of the binaural hearing aid system perform coordinated sound processing.

WO 2004/114722 A1

SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

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.

A BINAURAL HEARING AID SYSTEM WITH COORDINATED SOUND PROCESSING

FIELD OF THE INVENTION

The present invention relates to a binaural hearing aid system with a first hearing aid and a second hearing aid, each of which comprises a microphone, an A/D converter for provision of a digital input signal in response to sound signals received at the respective microphone in a sound environment, a processor that is adapted to process the digital input signals in accordance with a predetermined signal processing algorithm to generate a processed output signal, and a D/A converter and an output transducer for conversion of the respective processed sound signal to an acoustic output signal.

BACKGROUND OF THE INVENTION

Today's conventional hearing aids typically comprise a Digital Signal Processor (DSP) for processing of sound received by the hearing aid for compensation of the user's hearing loss. As is well known in the art, the processing of the DSP is controlled by a signal processing algorithm having various parameters for adjustment of the actual signal processing performed. The gains in each of the frequency channels of a multi-channel hearing aid are examples of such parameters.

The flexibility of the DSP is often utilized to provide a plurality of different algorithms and/or a plurality of sets of parameters of a specific algorithm. For example, various algorithms may be provided for noise suppression, i.e. attenuation of undesired signals and amplification of desired signals. Desired signals are usually speech or music, and undesired signals can be background speech, restaurant clatter, music (when speech is the desired signal), traffic noise, etc.

The different algorithms or parameter sets are typically included to provide comfortable and intelligible reproduced sound quality in different sound environments, such as speech, babble speech, restaurant clatter, music, traffic noise, etc. Audio signals obtained from different sound environments may possess very different characteristics, e.g. average and maximum sound pressure levels (SPLs) and/or frequency content. Therefore, in a hearing aid with a DSP, each type of sound environment may be associated with a particular program wherein a particular setting of algorithm parameters of a signal processing algorithm provides processed sound of optimum signal quality in a specific sound environment. A set of such parameters may typically include parameters related to broadband gain, corner frequencies or slopes of frequency-selective filter algorithms and

parameters controlling e.g. knee-points and compression ratios of Automatic Gain Control (AGC) algorithms.

Consequently, today's DSP based hearing instruments are usually provided with a number of different programs, each program tailored to a particular sound environment category and/or particular user preferences. Signal processing characteristics of each of these programs is typically determined during an initial fitting session in a dispenser's office and programmed into the instrument by activating corresponding algorithms and algorithm parameters in a non-volatile memory area of the hearing aid and/or transmitting corresponding algorithms and algorithm parameters to the non-volatile memory area.

Some known hearing aids are capable of automatically classifying the user's sound environment into one of a number of relevant or typical everyday sound environment categories, such as speech, babble speech, restaurant clatter, music, traffic noise, etc.

Obtained classification results may be utilised in the hearing aid to automatically select signal processing characteristics of the hearing aid, e.g. to automatically switch to the most suitable algorithm for the environment in question. Such a hearing aid will be able to maintain optimum sound quality and/or speech intelligibility for the individual hearing aid user in various sound environments.

US 5,687,241 discloses a multi-channel DSP based hearing instrument that utilises continuous determination or calculation of one or several percentile values of input signal amplitude distributions to discriminate between speech and noise input signals. Gain values in each of a number of frequency channels are adjusted in response to detected levels of speech and noise.

However, it is often desirable to provide a more subtle characterization of a sound environment than only discriminating between speech and noise. As an example, it may be desirable to switch between an omni-directional and a directional microphone preset program in dependence of, not just the level of background noise, but also on further signal characteristics of this background noise. In situations where the user of the hearing aid communicates with another individual in the presence of the background noise, it would be beneficial to be able to identify and classify the type of background noise. Omni-directional operation could be selected in the event that the noise being traffic noise to allow the user to clearly hear approaching traffic independent of its direction of arrival. If, on the other hand, the background noise was classified as being babble-noise, the directional listening program could be selected to allow the user to hear a target speech signal with improved signal-to-noise ratio (SNR) during a conversation.

Applying Hidden Markov Models for analysis and classification of the microphone signal may obtain a detailed characterisation of e.g. a microphone signal. Hidden Markov Models are capable of modelling stochastic and non-stationary signals in terms of both short and long time temporal variations. Hidden Markov Models have been applied in speech recognition as a tool for modelling statistical properties of speech signals. The article "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", published in Proceedings of the IEEE, VOL 77, No. 2, February 1989 contains a comprehensive description of the application of Hidden Markov Models to problems in speech recognition.

WO 01/76321 discloses a hearing aid that provides automatic identification or classification of a sound environment by applying one or several predetermined Hidden Markov Models to process acoustic signals obtained from the listening environment. The hearing aid may utilise determined classification results to control parameter values of a signal processing algorithm or to control switching between different algorithms so as to optimally adapt the signal processing of the hearing aid to a given sound environment.

The different available signal processing algorithms may change the signal characteristics significantly. In binaural hearing aid systems, it is therefore important that the determination of sound environment does not differ for the two hearing aids. However, since sound characteristics may differ significantly at the two ears of a user, it will often occur that sound environment determination at the two ears of a user differs, and this leads to undesired different signal processing of sounds for each of the ears of the user.

SUMMARY OF THE INVENTION

Thus, there is a need for a binaural hearing aid system wherein sound environment determination does not differ for the two hearing aids so that signal processing in the two hearing aids may be coordinated and the user be provided with desired processed sound in both ears at the same time.

According to the present invention, this and other objects are solved by provision of a binaural hearing aid system of the above-mentioned type wherein the hearing aids are connected either by wire or by a wireless link to at least one binaural sound environment detector for binaural determination of the sound environment surrounding a user of the binaural hearing aid system based on at least one signal from the first hearing aid and at least one signal from the second hearing aid whereby the sound environment is determined and classified based on binaural signals. The one or more binaural sound environment detectors provide outputs for each of the first and second hearing aids for selection of the

signal processing algorithm of each of the hearing aid processors so that the hearing aids of the binaural hearing aid system perform coordinated sound processing.

In this way both hearing aids may process sound in response to a common determination of sound environment. Sound environment determination may be performed by one common environment detector, for example situated in one of the hearing aids or in a remote control, or, by a plurality of environment detectors, such as an environment detector in each of the first and second hearing aids.

In the event that the user has substantially the same hearing loss in both ears and the sound environment is omni-directional, i.e. the sound environment does not change with direction, coordination of sound processing in the hearing aids leads to execution of identical signal processing algorithms in the respective signal processors of the hearing aids. In the event that the hearing aid user suffers from a binaural hearing loss, the signal processing algorithms may desirably differ for compensation of the different binaural hearing losses.

It is an important advantage of the present invention that binaural sound environment detection is more accurate than monaural detection since signals from both ears are taken into account.

It is a further advantage of the present invention that signal processing in the hearing aids of the binaural hearing aid system is coordinated since the sound environment detection is the same for both hearing aids.

BRIEF DESCRIPTION OF THE DRAWINGS

For a better understanding of the present invention reference will now be made, by way of example, to the accompanying drawings, in which:

Fig. 1 illustrates schematically a prior art monaural hearing aid with sound environment classification,

Fig. 2 illustrates schematically a first embodiment of the present invention,

Fig. 3 illustrates schematically a second embodiment of the present invention,

Fig. 4 illustrates schematically a third embodiment of the present invention, and

Fig. 5 illustrates schematically a fourth embodiment of the present invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Fig. 1 illustrates schematically a prior art monaural hearing aid 10 with sound environment classification.

The monaural hearing aid 10 comprises a first microphone 12 and a first A/D converter (not shown) for provision of a digital input signal 14 in response to sound signals received at the microphone 12 in a sound environment, and a second microphone 16 and a second A/D converter (not shown) for provision of a digital input signal 18 in response to sound signals received at the microphone 16, a processor 20 that is adapted to process the digital input signals 14, 18 in accordance with a predetermined signal processing algorithm to generate a processed output signal 22, and a D/A converter (not shown) and an output transducer 24 for conversion of the respective processed sound signal 22 to an acoustic output signal.

The hearing aid 10 further comprises a sound environment detector 26 for determination of the sound environment surrounding a user of the hearing aid 10. The determination is based on the output signals of the microphones 12, 16. Based on the determination, the sound environment detector 26 provides outputs 28 to the hearing aid processor 20 for selection of the signal processing algorithm appropriate in the determined sound environment. Thus, the hearing aid processor 20 is automatically switched to the most suitable algorithm for the determined environment whereby optimum sound quality and/or speech intelligibility is maintained in various sound environments.

The signal processing algorithms of the processor 20 may perform various forms of noise reduction and dynamic range compression as well as a range of other signal processing tasks.

The sound environment detector 26 comprises a feature extractor 30 for determination of characteristic parameters of the received sound signals. The feature extractor 30 maps the unprocessed sound inputs 14, 18 sound features, i.e. the characteristic parameters. These features can be signal power, spectral data and other well-known features.

The sound environment detector 26 further comprises an environment classifier 32 for categorizing the sound environment based on the determined characteristic parameters. The environment classifier categorizes the sounds into a number of environmental classes, such as speech, babble speech, restaurant clatter, music, traffic noise, etc. The classification process may consist of a simple nearest neighbour search, a neural network, a Hidden Markov Model system or another system capable of pattern recognition. The output of the environmental classification can be a "hard" classification containing one

single environmental class or a set of probabilities indicating the probabilities of the sound belonging to the respective classes. Other outputs may also be applicable.

The sound environment detector 26 further comprises a parameter map 34 for the provision of outputs 28 for selection of the signal processing algorithms.

The parameter map 34 maps the output of the environment classification 32 to a set of parameters for the hearing aid sound processor 20. Examples of such parameters are amount of noise reduction, amount of gain and amount of HF gain. Other parameters may be included.

Figs. 2-5 illustrate various preferred embodiments of the present invention. The illustrated binaural hearing aid system 1 comprises a first hearing aid 10 and a second hearing aid 10', each of which comprises a first microphone 12, 12 ' and an A/D converter (not shown) and a second microphone 16, 16' and A/D converter (not shown) for provision of a digital input signals 14, 14', 18, 18' in response to sound signals received at the respective microphones 12, 12', 16, 16' in a sound environment, a processor 20, 20' that is adapted to process the digital input signals 14, 18, 14', 18' in accordance with a predetermined signal processing algorithm to generate a processed output signal 22, 22', and a D/A converter (not shown) and an output transducer 24, 24' for conversion of the respective processed sound signals 22, 22' to an acoustic output signal

In the embodiments illustrated in Figs. 2-4, each of the hearing aids 10, 10' of the binaural hearing aid system 1 further comprises a binaural sound environment detector 26, 26' for determination of the sound environment surrounding a user of the binaural hearing aid system 1. The determination is based on the output signals of the microphones 12, 12', 16, 16'. Based on the determination, the binaural sound environment detector 26, 26' provides outputs 28, 28' to the hearing aid processors 20, 20' for selection of the signal processing algorithm appropriate in the determined sound environment. Thus, the binaural sound environment detectors 26, 26' determines the sound environment based on signals from both hearing aids, i.e. binaurally, whereby hearing aid processors 20, 20' is automatically switched in co-ordination to the most suitable algorithm for the determined environment whereby optimum sound quality and/or speech intelligibility is maintained in various sound environments by the binaural hearing aid system 1.

The binaural sound environment detectors 26, 26' illustrated in Figs. 2-4 are both similar to the binaural sound environment detector shown in Fig. 1 apart from the fact that the monaural environment detector only receives inputs from one hearing aid while each of the binaural sound environment detectors 26, 26' receives inputs from both hearing aids. Thus,

according to the present invention, signals are transmitted between the hearing aids 10, 10' so that the algorithms executed by the signal processors 20, 20' are selected in co-ordination, e.g. in case of an omni-directional sound environment, i.e. the sound environment does not change with direction, the algorithms are selected to be identical apart from possible differences in hearing loss compensation of the two ears.

In the embodiment of Fig. 2, the unprocessed signals 14, 14', 18, 18' from the microphones 12, 12', 16, 16' of one hearing aid 10, 10' are transmitted to the other hearing aid and inputted to the respective feature extractor 30, 30'. Thus, feature extraction in each of the hearing aids is based on the identical four input signals so that identical sound environment characteristic parameters will be determined binaurally in both hearing aids 10, 10'.

The signals may be transmitted in analogue form or in digital form, and the communication channel may be wired or wireless.

In the embodiment shown in Fig. 3, the output 36, 36' of the feature extractor 30, 30' of one hearing aid 10, 10' is transmitted to the respective other hearing aid 10', 10. The environment classifier 32, 32' then operates on two sets of features 36, 36' to determine the environment. Since both environment classifiers 32, 32' receive the same data, they will produce the same output.

In the embodiment shown in Fig. 4, the output 38, 38' of the environment classifier 32, 32' of one hearing aid 10, 10' is transmitted to the respective other hearing aid 10, 10'. The parameter map 34, 34' then operates on two inputs 38, 38' to produce the parameters for the processor algorithms, but since both parameter mapping units 34, 34' receive identical inputs, identical parameter values will be produced.

This embodiment has a number of advantages: Usually classification systems take both past and present data into account – they have memory. This makes them sensitive to missing data, since a classification requires a complete data set. Therefore it is required that the data link is safe, in the sense that data is guaranteed to be transmitted. The parameter mapping can be implemented without memory so that only present data is taken into account when generating parameters. This makes the system robust to packet loss and latency, since the parameter mapping may simply re-use old data in the event that data is missing. This will of course delay the correct action, but to the user the systems will appear to be synchronized.

The transmission data rate is low, since only a set of probabilities or logic values for the environment classes has to be transmitted.

Rather high latency can be accepted. By applying time constants to the variables that will change according to the output of the parameter mapping it is possible to smooth out any differences that is caused by latency. As described earlier it is important that signal processing in the two hearing instruments is coordinated. However if transition periods of a few seconds are allowed the system can operate with only 3-4 transmissions per second. Hereby, power consumption is kept low.

A binaural hearing aid system 1 with a remote control 40 is shown in Fig. 5. The environment detector 26 is positioned in the remote control 40. The required signals are transmitted to and from both hearing aids.

CLAIMS

1. A binaural hearing aid system comprising

a first hearing aid and a second hearing aid, each of which comprises

a microphone and an A/D converter for provision of a digital input signal in response to sound signals received at the respective microphone in a sound environment,

a processor that is adapted to process the digital input signals in accordance with a predetermined signal processing algorithm to generate a processed output signal, and

a D/A converter and an output transducer for conversion of the respective processed sound signal to an acoustic output signal, and

a binaural sound environment detector for binaural determination of the sound environment surrounding a user of the binaural hearing aid system based on at least one signal from the first hearing aid and at least one signal from the second hearing aid for provision of outputs for each of the first and second hearing aids for selection of the signal processing algorithm of each of the respective hearing aid processors so that the hearing aids of the binaural hearing aid system perform coordinated sound processing.

2. A system according to claim 1, wherein the binaural sound environment detector having inputs for each of the digital input signals is provided in a remote control of the system.

3. A system according to claim 1, wherein the binaural sound environment detector is provided in one of the hearing aids and provides an output for the other hearing aid.

4. A system according to claim 1, wherein each of the hearing aids comprises a binaural sound environment detector.

5. A system according to any of the preceding claims, wherein the binaural sound environment detector comprises

a feature extractor for determination of characteristic parameters of the received sound signals,

an environment classifier for categorizing the sound environment based on the determined characteristic parameters, and

a parameter map for the provision of outputs for selection of the signal processing algorithms.

6. A system according to claim 4 and 5, wherein each of the feature extractors has inputs for each of the digital input signals.

7. A system according to claim 4 and 5, wherein each of the environment classifiers has inputs connected to each of the feature extractors.

8. A system according to claim 4 and 5, wherein each of the parameter maps has inputs for connected to each of the environment classifiers.

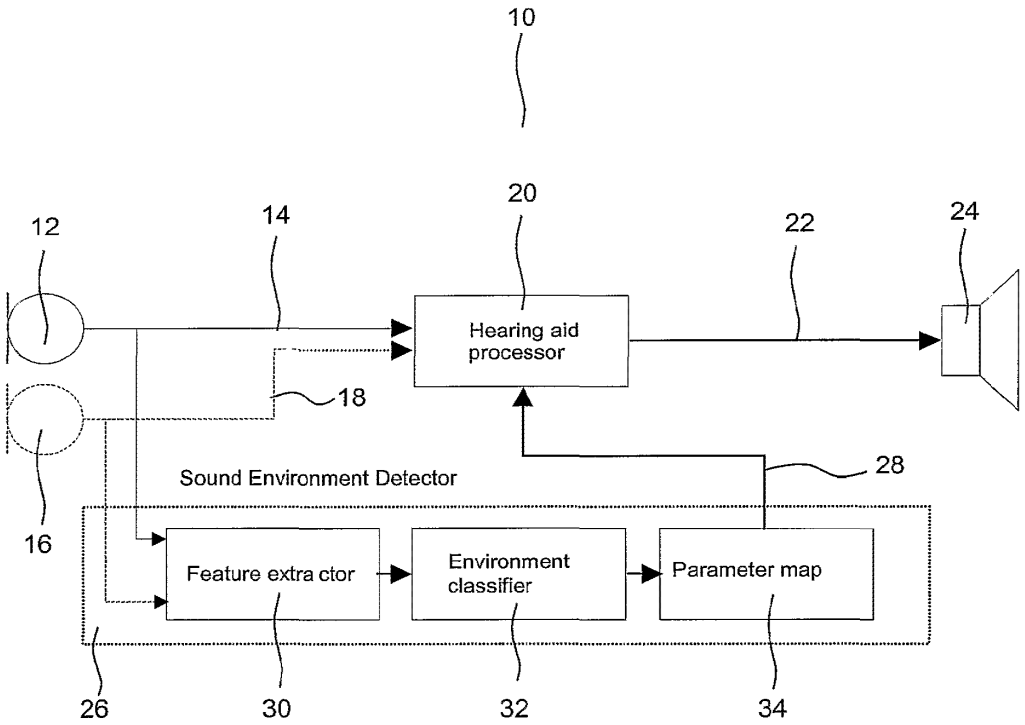

10
12
14
20
22
24
Hearing aid processor
18
16
Sound Environment Detector
28
Feature extra ctor
Environment classifier
Parameter map
26
30
32
34

Fig. 1

2/5

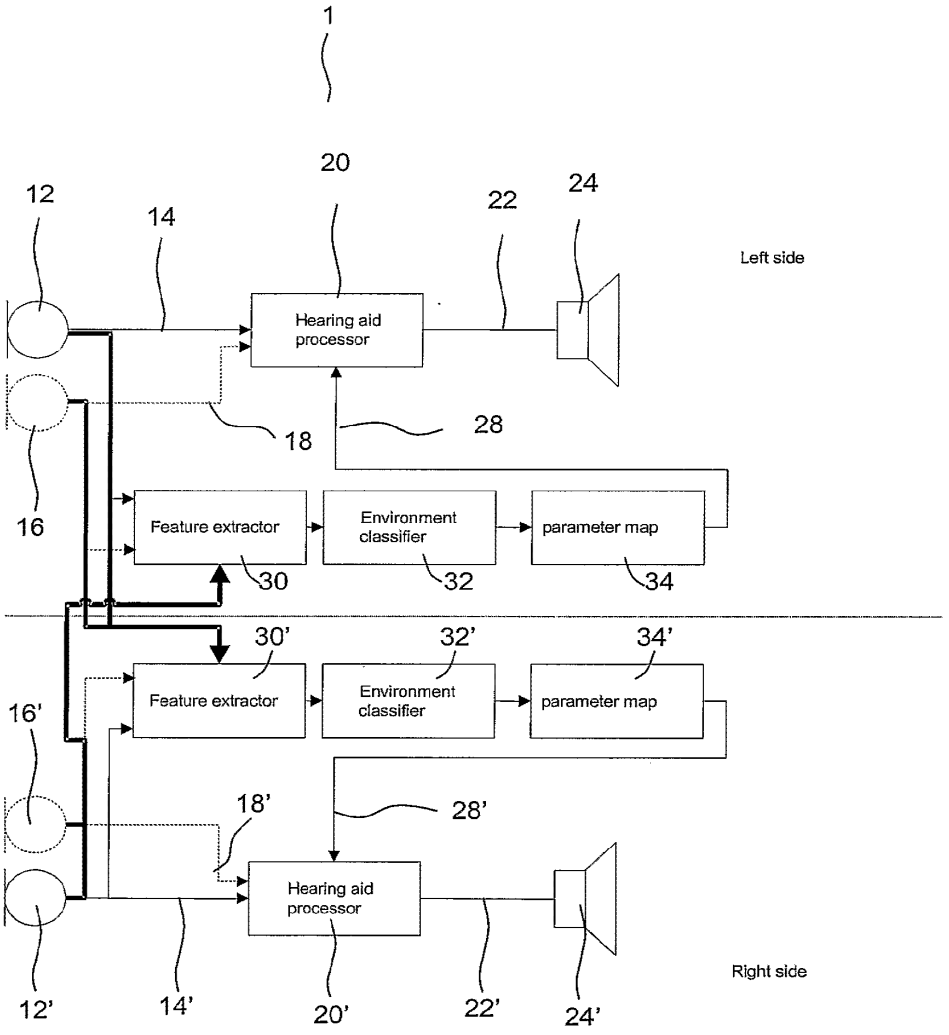

1
20
12
14
22
24
Left side
Hearing aid processor
18
28
16
Feature extractor
Environment classifier
parameter map
30
32
34
30'
32'
34'
Feature extractor
Environment classifier
parameter map
16'
18'
28'
Hearing aid processor
12'
14'
20'
22'
24'
Right side

Fig. 2

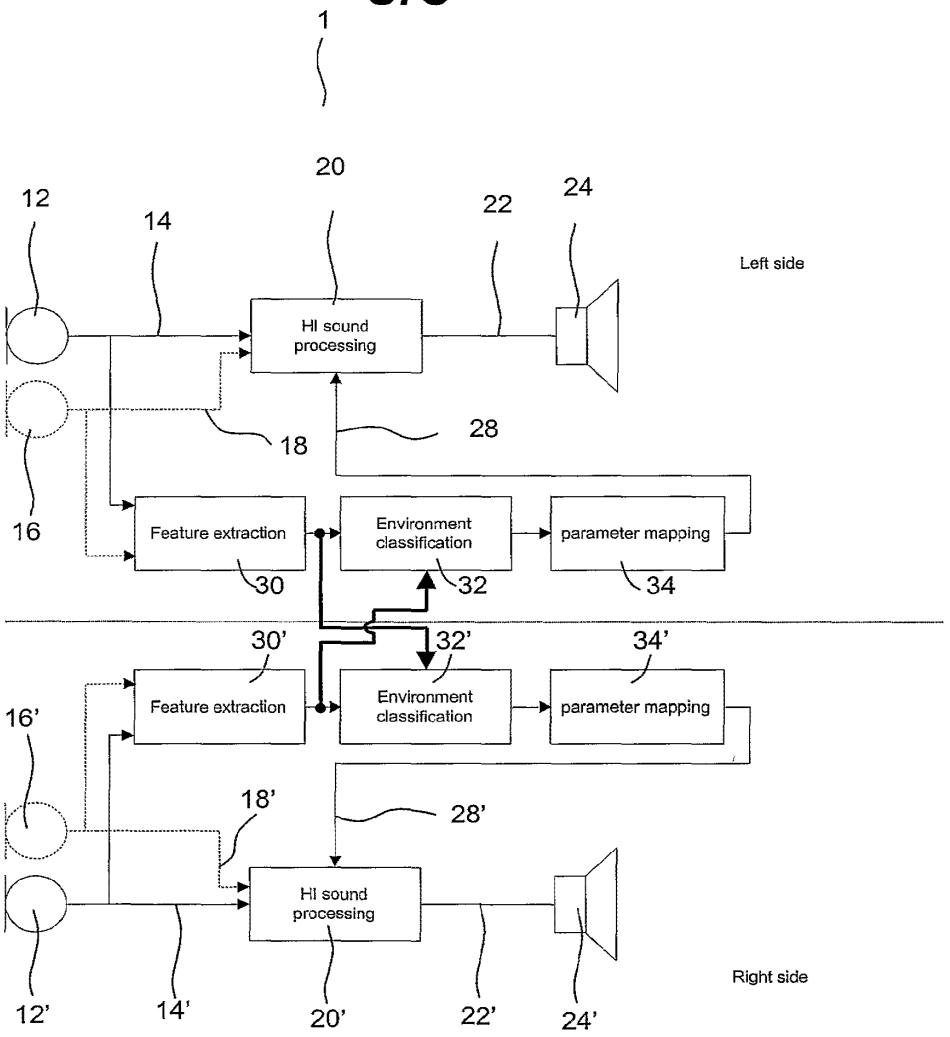

1
20
12
14
22
24
Left side
HI sound processing
18
28
16
Feature extraction
Environment classification
parameter mapping
30
32
34
30'
32'
34'
16'
Feature extraction
Environment classification
parameter mapping
18'
28'
HI sound processing
Right side
12'
14'
20'
22'
24'

Fig. 3

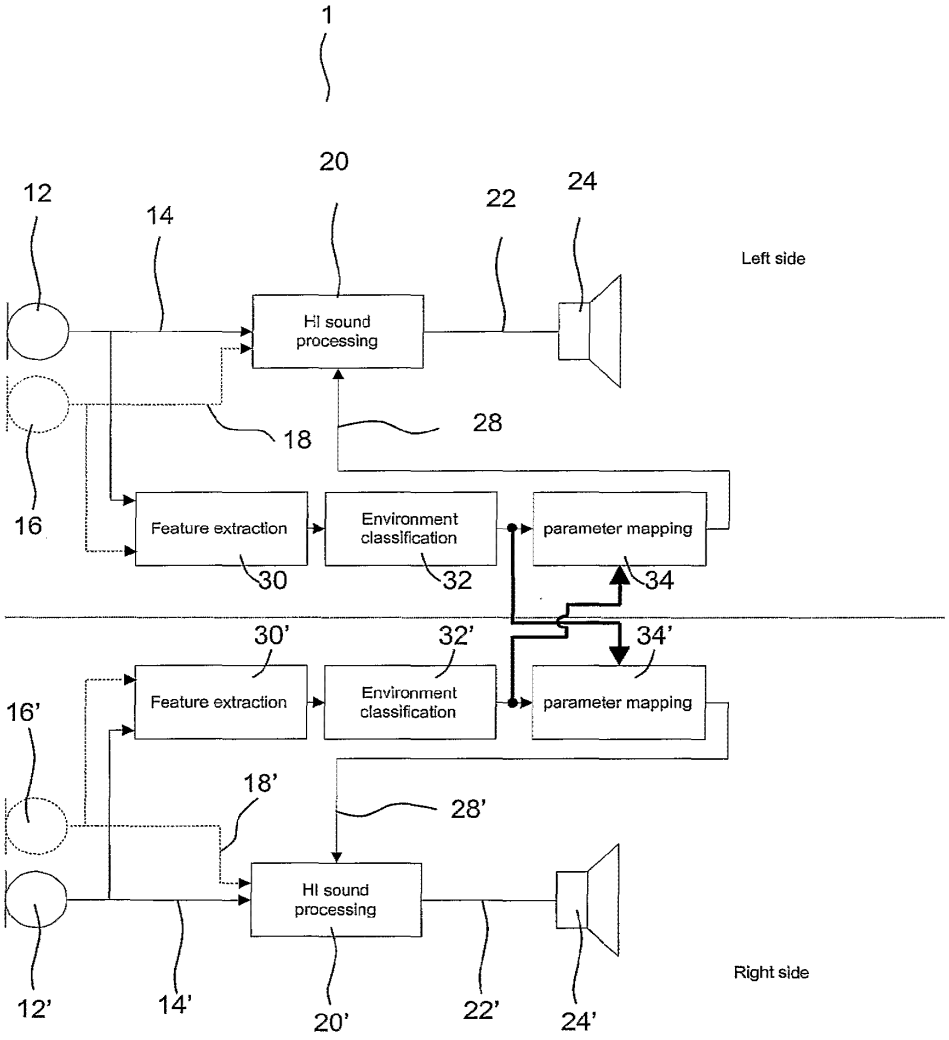

1
20
12
14
22
24
Left side
HI sound processing
18
28
16
Feature extraction
Environment classification
parameter mapping
30
32
34
30'
32'
34'
Feature extraction
Environment classification
parameter mapping
16'
18'
28'
HI sound processing
12'
14'
20'
22'
24'
Right side

Fig. 4

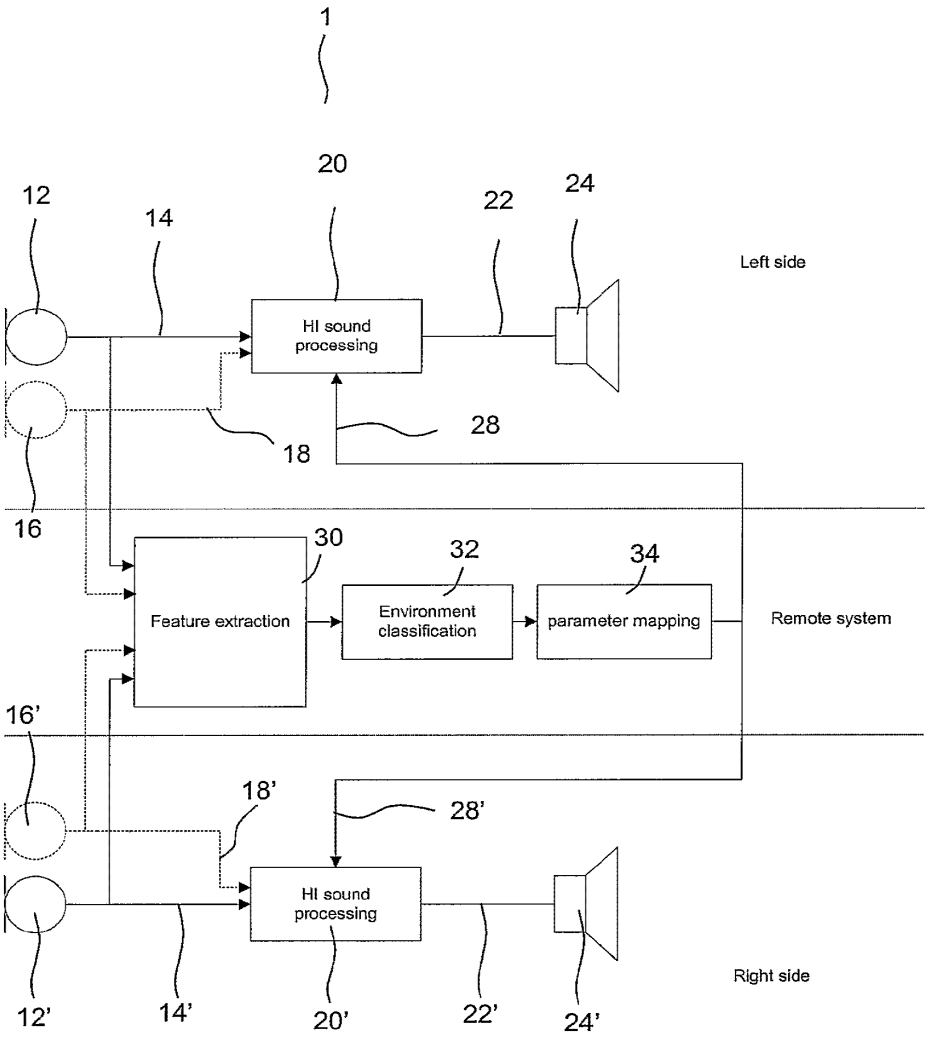

1
20
12
14
22
24
Left side
HI sound processing
28
18
16
30
32
34
Feature extraction
Environment classification
parameter mapping
Remote system
16'
18'
28'
HI sound processing
Right side
12'
14'
20'
22'
24'

Fig. 5

INTERNATIONAL SEARCH REPORT

International Application No
PCT/DK2004/000442

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 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 7 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)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 02/28143 A (SIEMENS AUDIOLOGISCHE TECHNIK ; NIEDERDRAENK TORSTEN (DE)) 4 April 2002 (2002-04-04) the whole document	1-8
X	US 5 757 932 A (BISGAARD NIKOLAI ET AL) 26 May 1998 (1998-05-26) the whole document	1-8
A	US 6 549 633 B1 (WESTERMANN SOEREN ERIK) 15 April 2003 (2003-04-15) the whole document	1-8
A	US 2002/037087 A1 (ALLEGRO SYLVIA ET AL) 28 March 2002 (2002-03-28) the whole document	5
A	WO 02/32208 A (PHONAK AG ; ALLEGRO SILVIA (CH)) 25 April 2002 (2002-04-25)	

[ ] 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
9 November 2004	18/11/2004

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	Fülöp, I

1

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/DK2004/000442

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0228143 A	04-04-2002	DE 10048354 A1	08-05-2002
		CN 1470149 T	21-01-2004
		WO 0228143 A2	04-04-2002
		EP 1360870 A2	12-11-2003
		US 2004013280 A1	22-01-2004
US 5757932 A	26-05-1998	US 5479522 A	26-12-1995
		AU 7248796 A	30-04-1997
		DE 69631781 D1	08-04-2004
		DK 855130 T3	12-07-2004
		EP 0855130 A1	29-07-1998
		NO 981559 A	08-06-1998
		WO 9714268 A1	17-04-1997
US 6549633 B1	15-04-2003	WO 9943185 A1	26-08-1999
		AU 733433 B2	17-05-2001
		AU 5982598 A	06-09-1999
		CA 2311405 A1	26-08-1999
		EP 1057367 A1	06-12-2000
		JP 2002504794 T	12-02-2002
US 2002037087 A1	28-03-2002	WO 0120965 A2	29-03-2001
		US 2002090098 A1	11-07-2002
		AU 2139901 A	24-04-2001
		CA 2400104 A1	29-03-2001
		EP 1348315 A2	01-10-2003
		JP 2004500592 T	08-01-2004
WO 0232208 A	25-04-2002	WO 0232208 A2	25-04-2002
		AU 2472202 A	29-04-2002
		CA 2439427 A1	25-04-2002
		EP 1470735 A2	27-10-2004

(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.

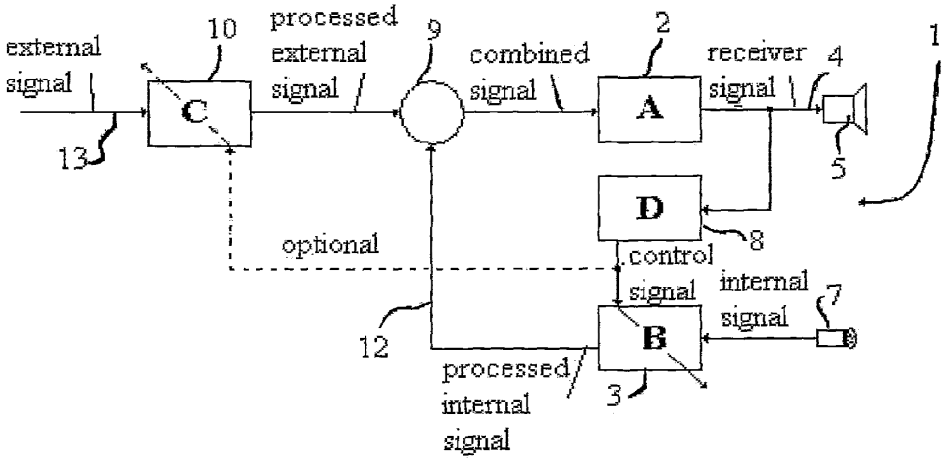
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, 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 $R_{31} = \frac{0.7}{77u} \approx 8.2k$ $R_{30} = 11 \times R_{31} \approx 91k$

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.

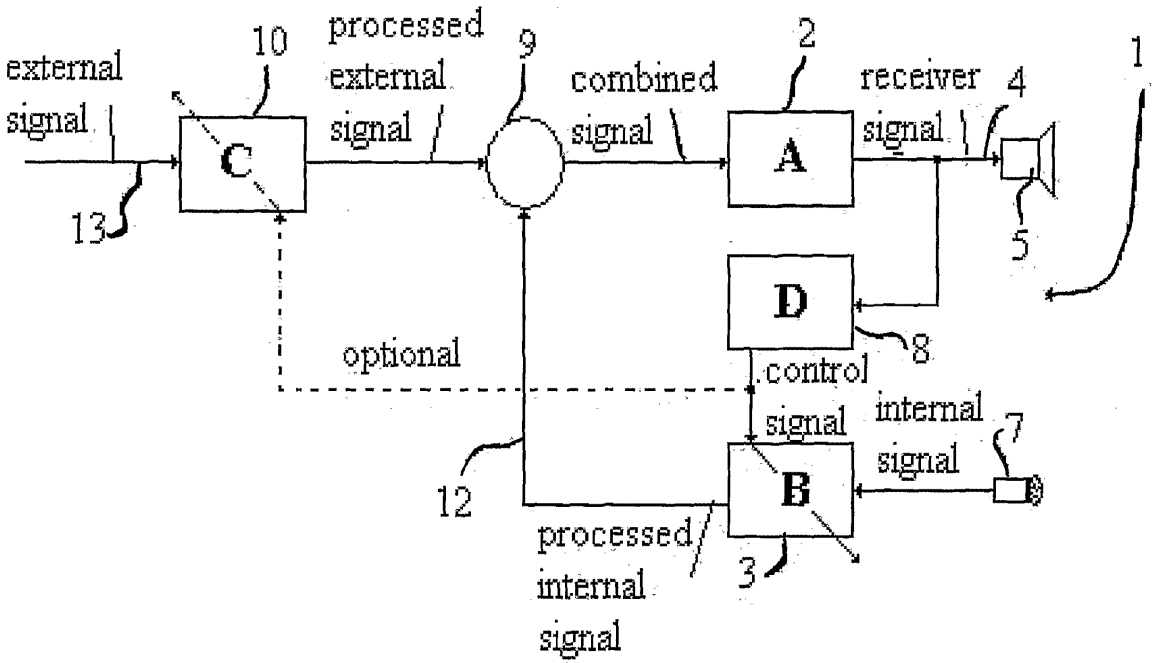

external
signal
10
C
13
processed
external
signal
9
combined
signal
2
A
receiver
signal
4
1
5
D
8
optional
control
signal
internal
signal
7
B
12
processed
internal
signal
3

Fig. 1

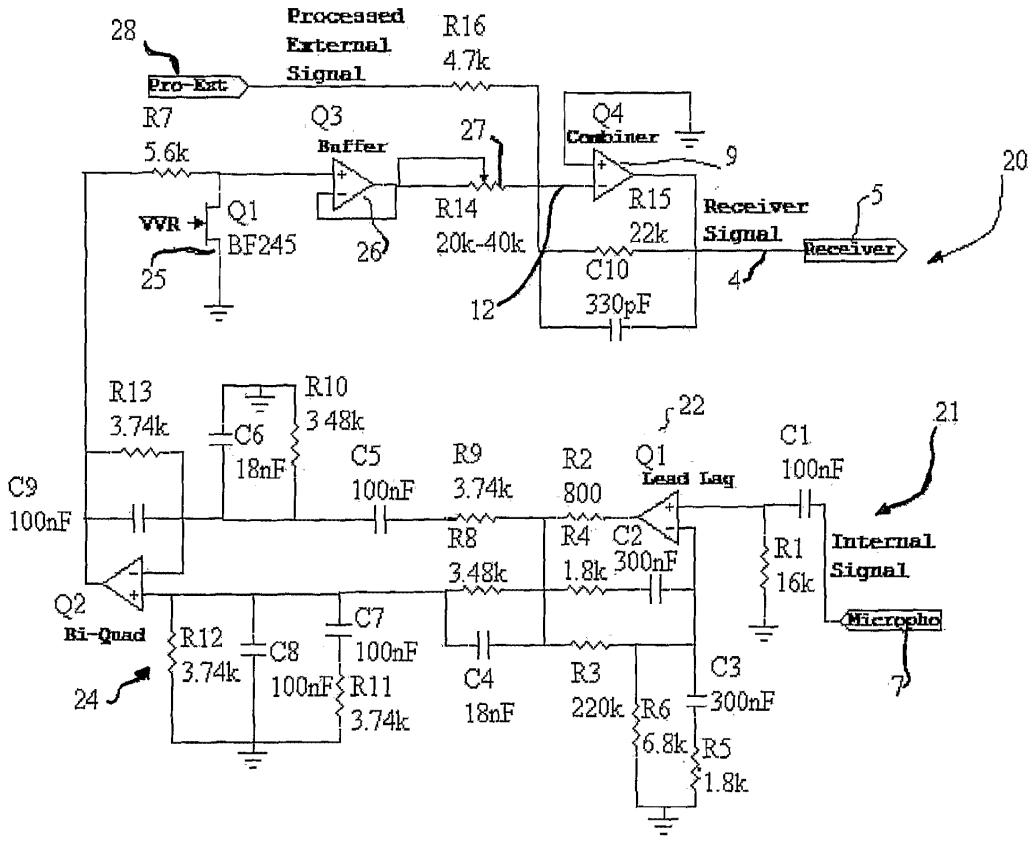

28
Processed
External
Signal
R16
4.7k
Pro-Ext
R7
5.6k
Q3
Buffer
27
Q4
Combiner
9
20
VVR
Q1
BF245
25
26
R14
20k-40k
R15
22k
Receiver
Signal
5
Receiver
C10
330pF
12
4
R13
3.74k
R10
3.48k
C6
18nF
22
C1
100nF
21
C9
100nF
C5
100nF
R9
3.74k
R2
800
Q1
Lead Lag
R8
3.48k
R4
1.8k
C2
300nF
R1
16k
Internal
Signal
Q2
Bi-Quad
R12
3.74k
C8
100nF
C7
100nF
R11
3.74k
C4
18nF
R3
220k
R6
6.8k
C3
300nF
R5
1.8k
24
7

Fig. 2

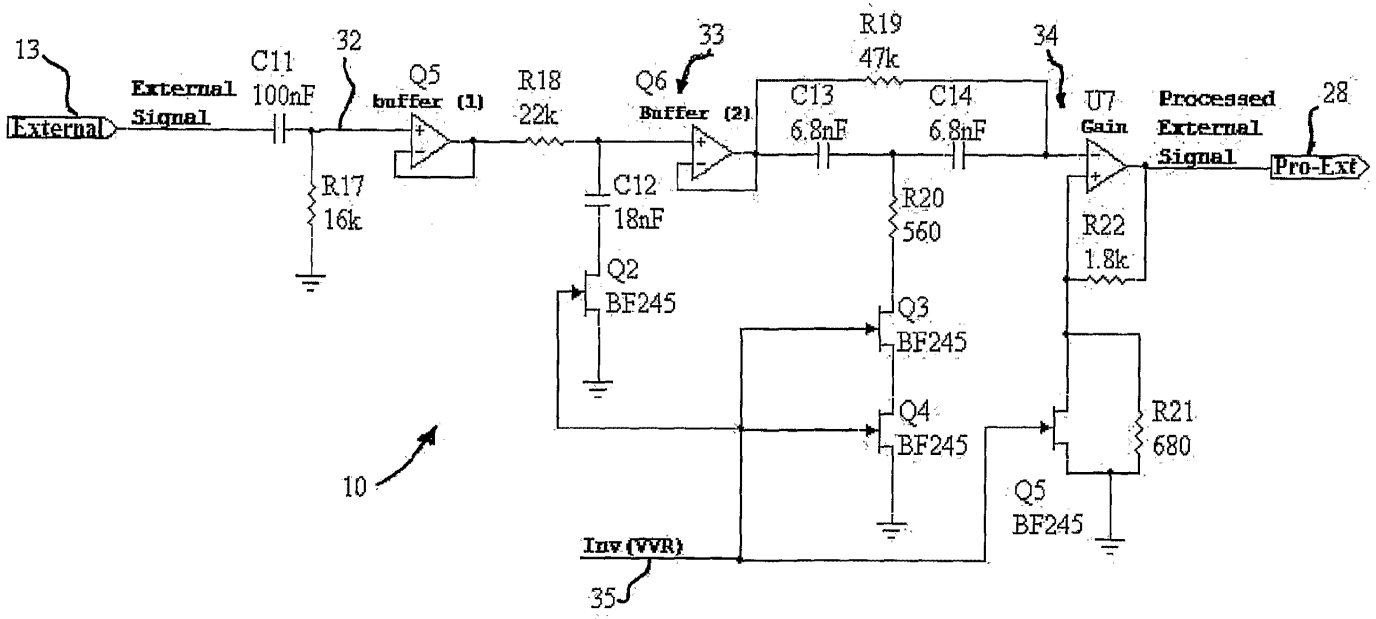

13
External
External Signal
C11
100nF
32
Q5
buffer (1)
R18
22k
33
Q6
Buffer (2)
R19
47k
34
C13
6.8nF
C14
6.8nF
U7
Gain
Processed External Signal
28
Pro-Ext
R17
16k
C12
18nF
R20
560
R22
1.8k
Q2
BF245
Q3
BF245
Q4
BF245
R21
680
Q5
BF245
10
Inv (VVR)
35

Fig. 3

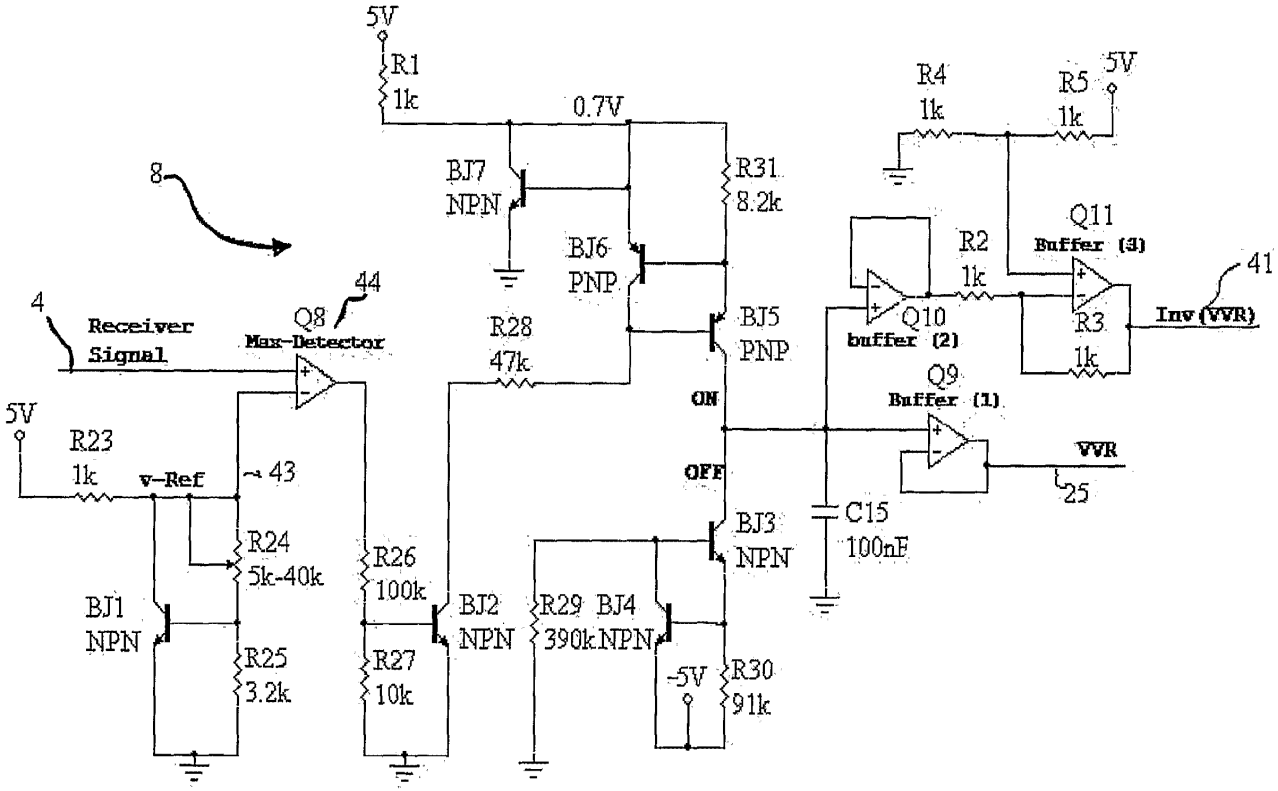

5V
R1
1k
0.7V
8
BJ7
NPN
R31
8.2k
BJ6
PNP
BJ5
PNP
R4
1k
R5
1k
5V
Q11
Buffer (3)
R2
1k
41
Q10
buffer (2)
R3
1k
Inv (VVR)
4
Receiver
Signal
Q8
Max-Detector
44
R28
47k
ON
Q9
Buffer (1)
5V
R23
1k
v-Ref
43
OFF
VVR
25
BJ3
NPN
C15
100nF
R24
5k-40k
R26
100k
BJ1
NPN
BJ2
NPN
R29
390k
BJ4
NPN
R25
3.2k
R27
10k
-5V
R30
91k

Fig. 4

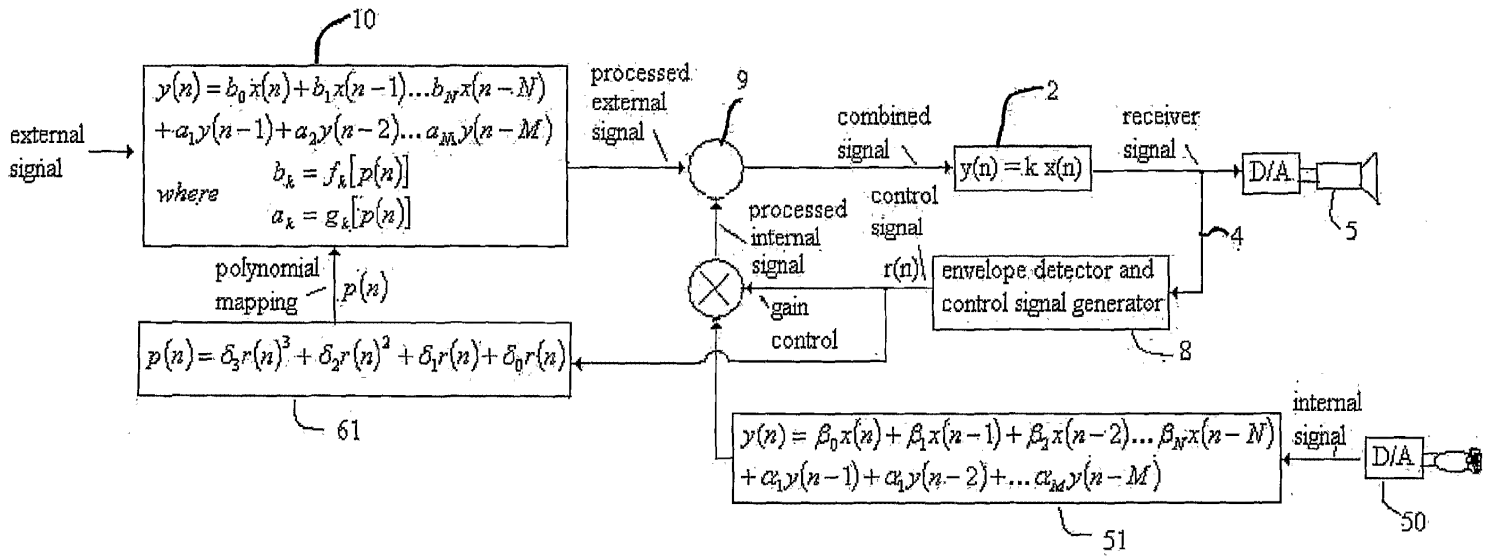

10
$y(n) = b_0 x(n) + b_1 x(n-1) \ldots b_N x(n-N)$
$+ a_1 y(n-1) + a_2 y(n-2) \ldots a_M y(n-M)$
where $b_k = f_k[p(n)]$
$a_k = g_k[p(n)]$
external signal
processed external signal
9
combined signal
2
y(n) = k x(n)
receiver signal
D/A
5
4
processed internal signal
control signal
r(n)
envelope detector and control signal generator
8
gain control
polynomial mapping
p(n)
$p(n) = \delta_3 r(n)^3 + \delta_2 r(n)^2 + \delta_1 r(n) + \delta_0 r(n)$
61
$y(n) = \beta_0 x(n) + \beta_1 x(n-1) + \beta_2 x(n-2) \ldots \beta_N x(n-N)$
$+ \alpha_1 y(n-1) + \alpha_1 y(n-2) + \ldots \alpha_M y(n-M)$
51
internal signal
D/A
50

Fig. 5

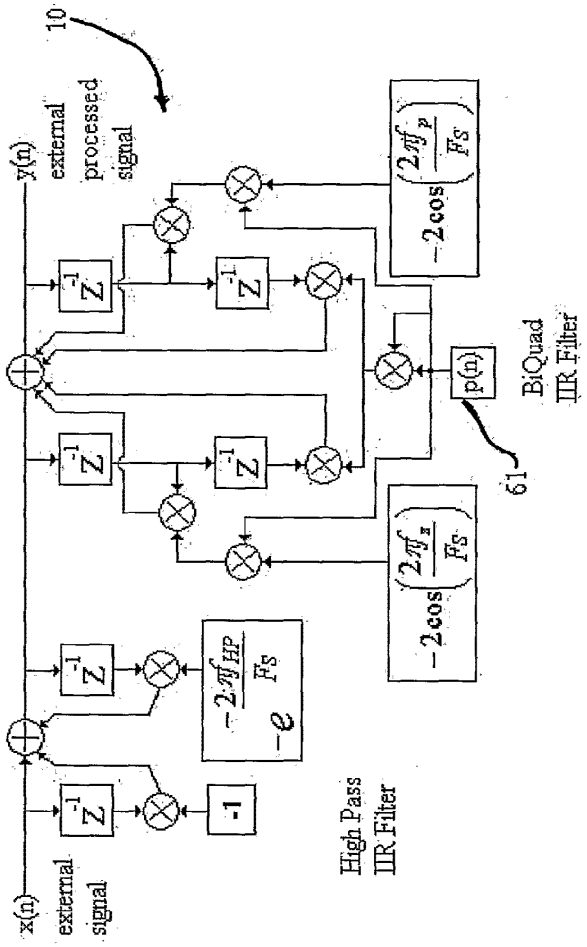


Fig. 6

Exhibit 1004
Page 0472 of 1001

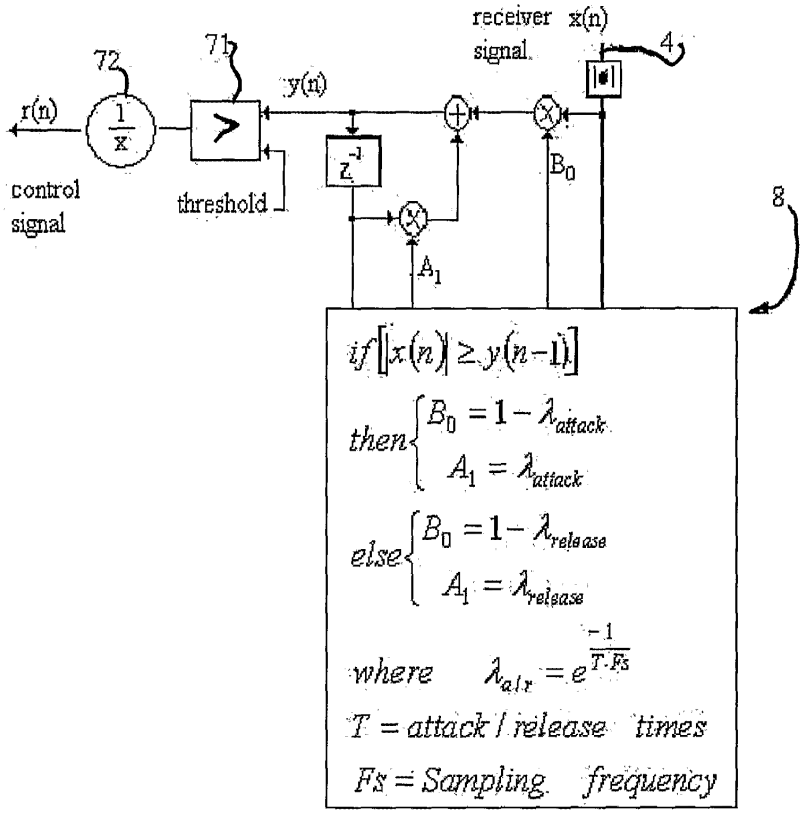

receiver x(n)
signal
4
72
71
y(n)
r(n)
1/x
>
control
signal
threshold
z^-1
B0
8
A1
if [|x(n)| ≥ y(n-1)]
then { B0 = 1 - λattack
A1 = λattack
else { B0 = 1 - λrelease
A1 = λrelease
where λa/r = e^(-1/(T·Fs))
T = attack / release times
Fs = Sampling frequency

Fig. 7

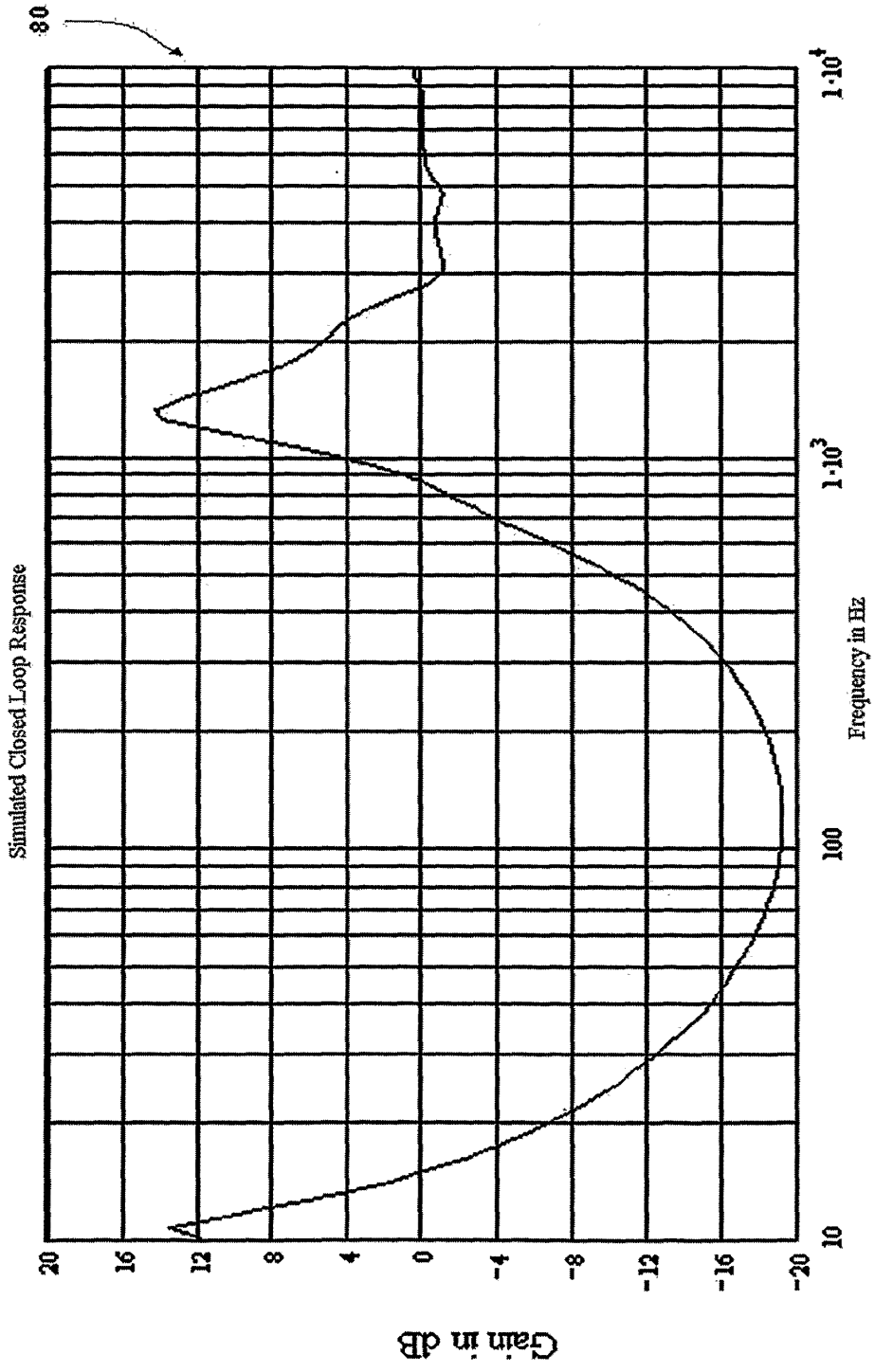

80
Simulated Closed Loop Response
Gain in dB
20
16
12
8
4
0
-4
-8
-12
-16
-20
10
100
1·10^3
1·10^4
Frequency in Hz

Fig. 8

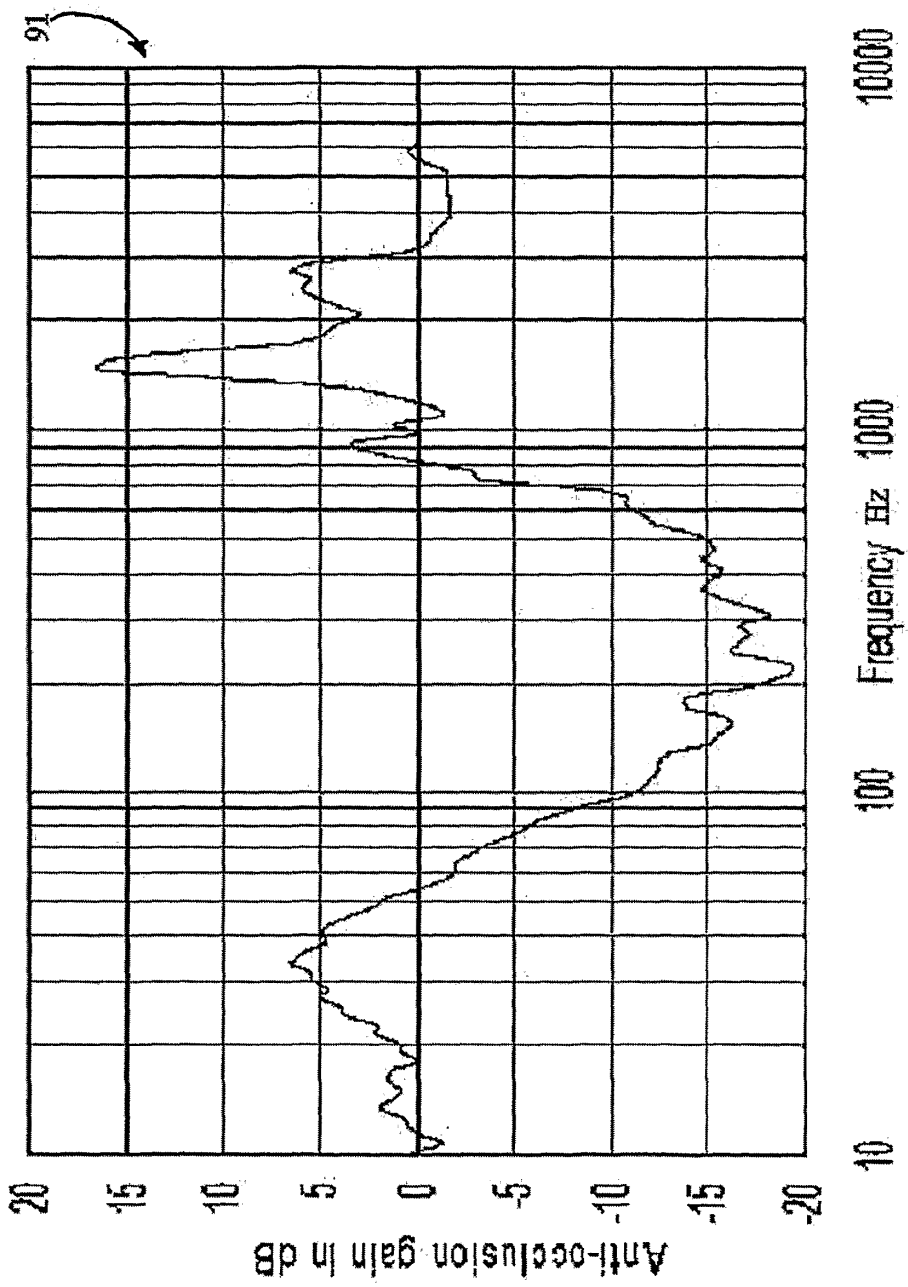

91
Anti-occlusion gain in dB
20
15
10
5
0
-5
-10
-15
-20
10
100
1000
10000
Frequency Hz

Fig. 9

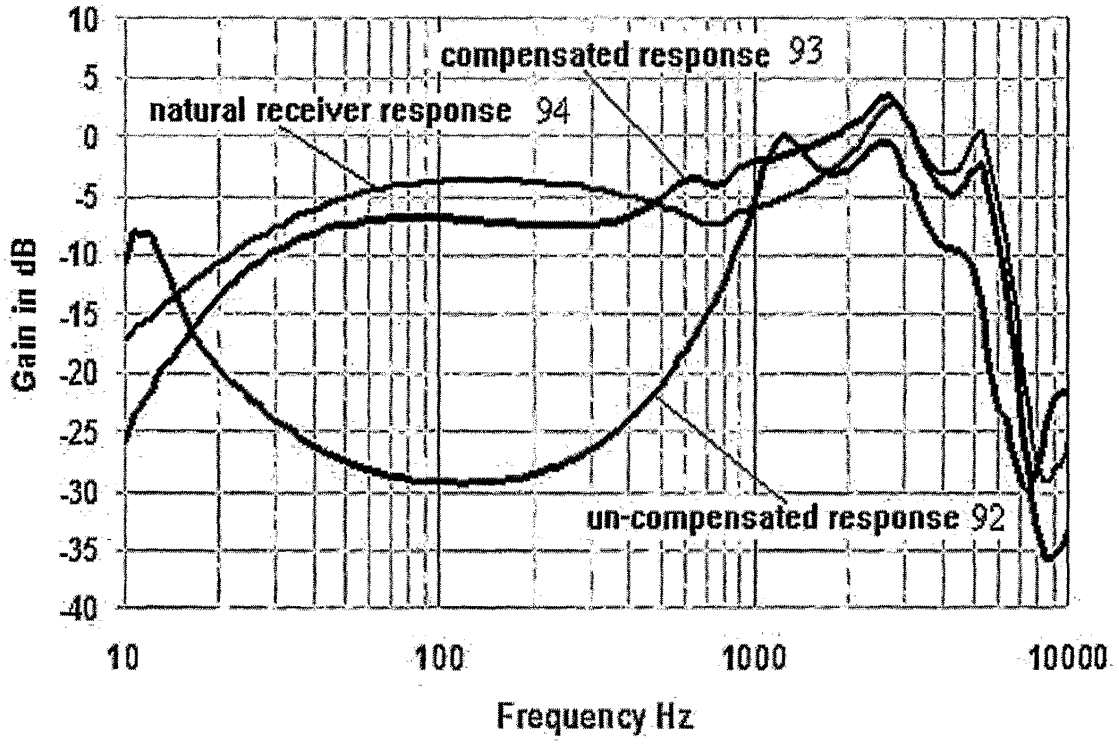

compensated response 93
natural receiver response 94
un-compensated response 92
Gain in dB
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40
10
100
1000
10000
Frequency Hz

Fig. 10

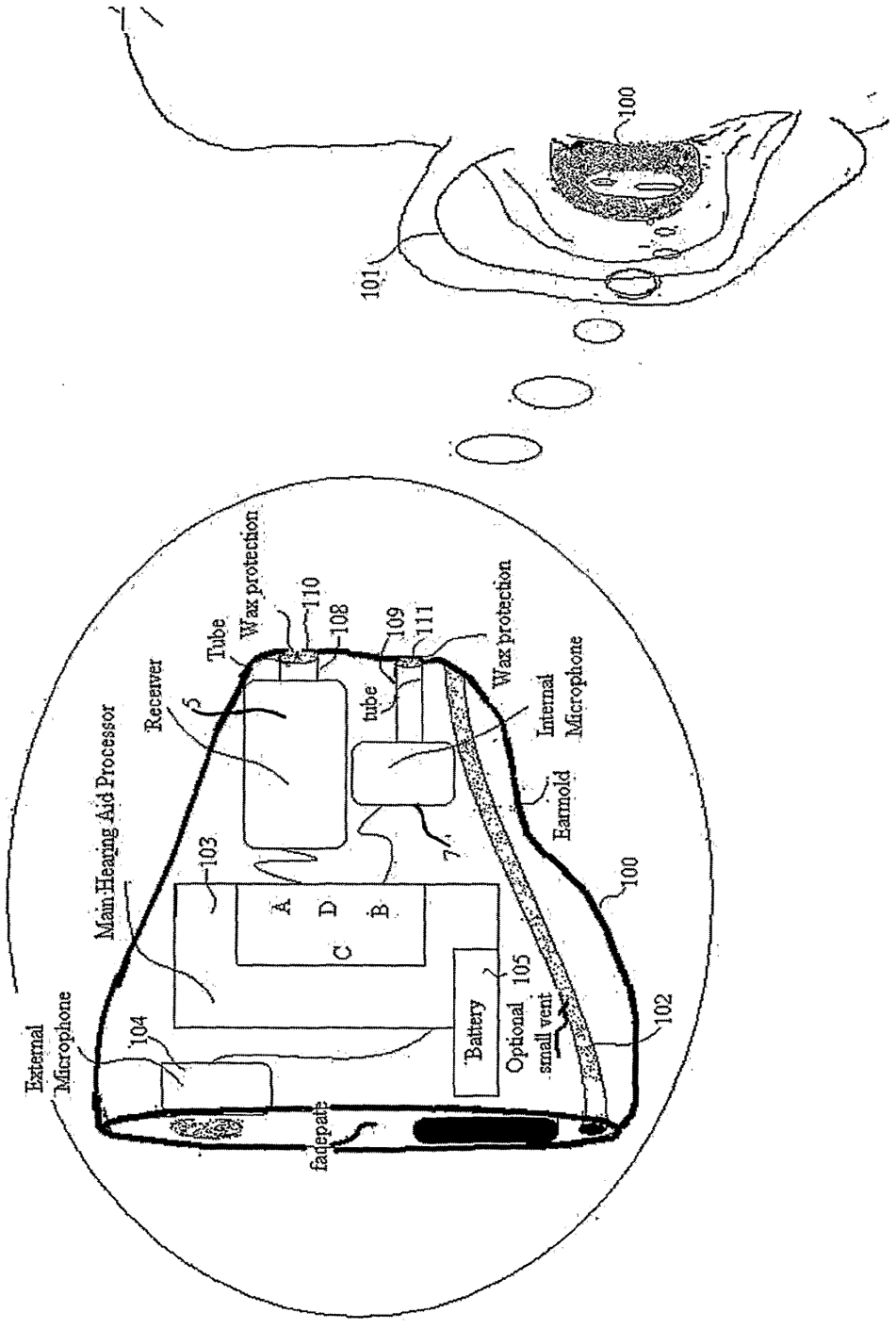


Fig. 11

Exhibit 1004
Page 0477 of 1001

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2005/001506

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]: H04R 29/00; A61B 5/00

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)

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)
WPAT: hearing aid, occlusion, gain, frequency range, sound level, signal strength and similar terms.
USPTO: hearing aid, occlusion, reduce

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6687377 B2 (VOIX et al.) 3 February 2004 Col. 3 line 33-Col. 6 line 22	1-24
A	WO 2002/085063 A2 (WIDEX A/S et al.) 24 October 2002 Page 4 line 14-Page 10 line 32	1-24
A	US 6766031 B1 (WESTERMANN) 20 July 2004 Col. 2 line 18-Col. 5 line 28	1-24
A	CA 2381516 A1 (GENNUM CORPORATION, CA) 12 October 2002 Page 1 line 20-Page 18 line 2	1-24

[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 application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 November 2005	9 NOV 2005
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized officer **JUZER KHANBHAI** Telephone No : (02) 6283 2176

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2005/001506

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 5195139 A (GAUTHIER) 16 March 1993 Col. 1 line 39-Col. 3 line 22	1-24

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2005/001506

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
US	6687377	AU	20394/02	CA	2477010	CN	1483298
		EP	1356255	US	2002076057	WO	0250499
		ZA	200305523				
WO	2002/085063	CA	2442538	EP	1250027	US	2002150269
US	6766031	AU	69189/98	CA	2286037	DK	42297
		EP	0980641	WO	9847318		
CA	2381516	N/A					
US	5195139	AU	20280/92	EP	0584259	WO	9221218

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2006 年5 月26 日 (26.05.2006)

PCT

(10) 国際公開番号
WO 2006/054698 A1

(51) 国際特許分類:
H04S 5/02 (2006.01) *H04S 7/00* (2006.01)
H04S 1/00 (2006.01)

(21) 国際出願番号: PCT/JP2005/021247

(22) 国際出願日: 2005 年11 月18 日 (18.11.2005)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2004-335667
2004 年11 月19 日 (19.11.2004) JP
特願 2004-337165
2004 年11 月22 日 (22.11.2004) JP
特願 2004-369406
2004 年12 月21 日 (21.12.2004) JP
特願2005-149276 2005 年5 月23 日 (23.05.2005) JP
特願2005-178919 2005 年6 月20 日 (20.06.2005) JP

(71) 出願人 (米国を除く全ての指定国について): 日本ビクター株式会社 (VICTOR COMPANY OF JAPAN, LIMITED) [JP/JP]; 〒2218528 神奈川県横浜市神奈川区守屋町３丁目１２番地 Kanagawa (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 村田 寿子 (MURATA, Hisako). 鈴木 琢磨 (SUZUKI, Takuma). 安良 定浩 (YASURA, Sadahiro). 羽石 実希 (HANEISHI, Miki). 長谷川 勝巳 (HASEGAWA, Katsumi). 相崎 正子 (AIZAKI, Naoko). 藤浪 喜久 (FUJINAMI, Yoshihisa). 武田 正美 (TAKEDA, Masami). 菅原 隆幸 (SUGAHARA, Takayuki).

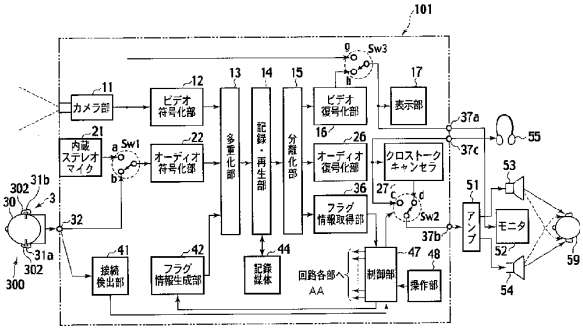
(74) 代理人: 三好 秀和, 外 (MIYOSHI, Hidekazu et al.); 〒1050001 東京都港区虎ノ門１丁目２番８号 虎ノ門琴平タワー Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,

[続葉有]

(54) Title: VIDEO/AUDIO RECORDING APPARATUS AND METHOD, AND VIDEO/AUDIO REPRODUCING APPARATUS AND METHOD

(54) 発明の名称: 映像音声記録装置及び方法、並びに、映像音声再生装置及び方法



11 CAMERA PART
21 INTERNALLY-MOUNTED STEREO MICROPHONES
41 CONNECTION DETECTING PART
12 VIDEO ENCODING PART
22 AUDIO ENCODING PART
42 FLAG INFORMATION GENERATING PART
13 MULTIPLEXING PART
14 RECORDING/REPRODUCING PART
44 RECORDING MEDIUM
15 SEPARATING PART
16 VIDEO DECODING PART
26 AUDIO DECODING PART
36 FLAG INFORMATION ACQUIRING PART
AA TO VARIOUS CIRCUIT COMPONENTS
17 DISPLAY PART
27 CROSSTALK CANCELLER
47 CONTROL PART
48 OPERATING PART
51 AMPLIFIER
52 MONITOR

(57) Abstract: A video/audio recording/reproducing apparatus (101) has an internally-mounted stereo microphones (21a,21b) and an external microphone connection terminal (32). The external microphone connection terminal (32) is connected to binaural microphones (3) to be attached to the ears of a recording user (300). In a case of collecting ambient sounds by use of the binaural microphones (3), the user switches the audio signals to be recorded in a recording medium from the audio signals obtained from the internally-mounted stereo microphones (21a,21b) to the binaural audio signals obtained from the binaural microphones (3). The recording user (300) attaches the binaural microphones (3(31a,31b)) to his ears to collect sounds around him including sounds generated from an imaged subject. A camera part (11) images the subject. The binaural audio signals, video signals obtained by imaging the subject, and binaural flag signals are recorded into the recording medium.

[続葉有]

WO 2006/054698 A1

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) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (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).

添付公開書類:
— 国際調査報告書

２文字コード及び他の略語については、定期発行される各*PCT*ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(57) 要約: 映像音声記録再生装置（１０１）は内蔵ステレオマイク（２１ａ，２１ｂ）と外部マイク接続端子（３２）を備える。外部マイク接続端子（３２）には撮影者（３００）の耳に装着するバイノーラルマイク（３）が接続されている。バイノーラルマイク（３）で周囲音を集音する場合は記録媒体に記録する音声信号を内蔵ステレオマイク（２１ａ，２１ｂ）からの音声信号からバイノーラルマイク（３）からのバイノーラル音声信号へと切り替える。撮影者（３００）はバイノーラルマイク（３（３１ａ．３１ｂ））を耳に装着して被写体から発せられた音を含む撮影者（３００）の周囲音を集音する。カメラ部（１１）は被写体を撮像する。記録媒体にバイノーラル音声信号と撮像した映像信号とバイノーラルフラグ信号とを記録する。

明 細 書

映像音声記録装置及び方法、並びに、映像音声再生装置及び方法

技術分野

[0001] 本発明は、被写体を撮像した映像信号と被写体から発せられた音を含む撮影者の周囲音を集音した音声信号とを記録する映像音声記録装置及び方法、並びに、記録媒体に記録した映像信号と音声信号とを再生する映像音声再生装置及び方法に係り、特に撮像した映像と共に臨場感のある音声を再生することができる映像音声記録装置及び方法、並びに、映像音声再生装置及び方法に関する。

背景技術

[0002] 被写体を撮像した映像信号と被写体から発せられた音を含む撮影者の周囲音を集音した音声信号とを記録する映像音声記録再生装置(いわゆるビデオカメラ)が普及している。この種の映像音声記録再生装置は音声をステレオ録音するためのステレオマイクを備えている。近年、映像音声記録再生装置は小型化が一段と進んでおり、小型化された映像音声記録再生装置に搭載したステレオマイクでは臨場感のある音声を記録しにくいという問題がある。そこで、より臨場感のある音声を記録することができる映像音声記録再生装置が望まれている。

[0003] 国際公開第96／10884号パンフレットには、映像音声記録再生装置本体の両側に耳構造体を配置することにより、被写体を撮像した映像信号と共に撮影者の周囲の音場をバイノーラル集音した音声を記録することができる映像音声記録再生装置が開示されている。

[0004] しかしながら、上記文献に開示されている、バイノーラルマイクを装置本体に取り付けた映像音声記録再生装置では、装置本体の横幅、即ち左右のマイクの間隔が人頭と同程度でなければ臨場感のある音声を記録することができない。近年製品化されている映像音声記録再生装置本体は、高密度記録技術、ディジタル信号記録技術及び映像圧縮処理技術の進歩に伴って小型化されているので、映像音声記録再生装置本体にバイノーラル集音を目的とするマイクを搭載しても期待するほどの効果が得られない。また、装置の形状が人頭とは大きく異なるので、上記文献に開示され

ているような効果を得ることは困難と考えられる。

発明の開示

[0005] 本発明はこのような問題点に鑑みなされたものであり、装置本体の大きさや形状に関係なく、撮像した映像と共に臨場感のある音声を再生することができる映像音声記録装置及び方法、並びに、映像音声再生装置及び方法を提供することを目的とする。

[0006] また、被写体をズームアップして撮影する場合に、被写体のズームアップに連動してさらに臨場感のある音声を再生することができる映像音声記録装置及び方法、並びに、映像音声再生装置及び方法を提供することを目的とする。

[0007] さらに、再生信号を視聴する視聴者とバイノーラル集音した撮影者とが異なる場合のように、任意の撮影者が撮影して任意の視聴者が視聴する場合でも、違和感のほとんどない臨場感のある音声を受聴することができる映像音声再生装置及び方法を提供することを目的とする。

[0008] 本発明は、上述した従来の技術の課題を解決するため、被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者の周囲音を集音した音声信号とを記録する映像音声記録装置において、前記被写体を撮像するカメラ部と、前記周囲音を集音するマイクとして、前記撮影者の耳に装着するバイノーラルマイクを使用するかバイノーラルマイク以外のマイクを使用するかを切り替える切り替え部と、前記カメラ部より出力された映像信号を処理するビデオ処理部と、前記周囲音を集音するマイクより出力された音声信号を処理するオーディオ処理部と、前記切り替え部によって前記周囲音を集音するマイクとして前記バイノーラルマイクに切り替えた場合に、前記周囲音を集音する際の音声モードをバイノーラルモードとして、バイノーラルモードであることを示すバイノーラルフラグ信号を生成するフラグ情報生成部と、前記ビデオ処理部によって処理された映像信号と前記オーディオ処理部によって処理された音声信号と前記バイノーラルフラグ信号とを記録媒体に記録する記録部とを備えて構成したことを特徴とする映像音声記録装置を提供する。

[0009] 本発明によれば、装置本体の大きさや形状に関係なく、撮像した映像と共に臨場感のある音声を再生することができる。また、被写体をズームアップして撮影する場合

、被写体のズームアップに連動してさらに臨場感のある音声を再生することができる。さらに、再生信号を視聴する視聴者とバイノーラル集音した撮影者とが異なる場合のように、任意の撮影者が撮影して任意の視聴者が視聴する場合でも、違和感のほとんどない臨場感のある音声を受聴することができる

この構成において、前記映像音声記録装置に内蔵されている内蔵マイクと、外部マイク接続端子と、前記外部マイク接続端子に接続して使用する外部マイクとして、前記バイノーラルマイクかバイノーラルマイク以外のマイクかを設定する設定部と、前記外部マイク接続端子に外部マイクが接続されているか否かを検出する接続検出部と、前記オーディオ処理部に供給する音声信号として、前記内蔵マイクより出力された音声信号と前記外部マイクより出力された音声信号とを切り替えるスイッチと、前記設定部によって前記外部マイクがバイノーラルマイクに設定され、前記接続検出部によって前記外部マイク接続端子に外部マイクが接続されていることが検出されたとき、音声モードを前記バイノーラルモードとする制御部とを備え、前記制御部は、音声モードが前記バイノーラルモードのとき、前記スイッチが前記外部マイクより出力された音声信号を前記オーディオ処理部に供給するよう前記スイッチを切り替え制御し、前記フラグ情報生成部が前記バイノーラルフラグ信号を生成するよう前記フラグ情報生成部を制御することが好ましい。

[0010] また、前記カメラ部より出力された映像信号を表示する表示部を備え、音声モードが前記バイノーラルモードのとき、前記表示部に前記バイノーラルモードであることを示すバイノーラルマークを表示させる表示制御部を備えることが好ましい。

[0011] さらに、前記カメラ部は前記被写体を拡大して撮像するズーム機能を有し、前記バイノーラルマイクによって集音された音声信号を前記カメラ部における拡大率に応じて増幅する音声ズーム処理部を備えることが好ましい。

[0012] さらにまた、前記カメラ部は前記被写体を拡大して撮像するズーム機能を有し、前記バイノーラルマイクが集音する音声信号の音源を仮想的に受聴者に接近させて仮想音源を形成するための頭部伝達関数を前記仮想音源と前記受聴者との複数の距離に対応させて記憶する伝達関数メモリと、この伝達関数メモリに記憶された複数の頭部伝達関数のいずれかを前記カメラ部における拡大率に応じて選択的に取得す

る関数取得器と、前記バイノーラルマイクによって集音された音声信号に対して前記関数取得器にて取得された頭部伝達関数を畳み込む畳み込み演算器とを有する音声ズーム処理部を備えることが好ましい。

[0013] 本発明は、上述した従来の技術の課題を解決するため、被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者の周囲音を集音した音声信号とを記録する映像音声記録方法において、前記被写体を撮像する撮像ステップと、前記周囲音を集音するマイクとして、前記撮影者の耳に装着するバイノーラルマイクを使用するかバイノーラルマイク以外のマイクを使用するかを切り替える切り替えステップと、前記被写体を撮像した映像信号を処理するビデオ処理ステップと、前記周囲音を集音するマイクより出力された音声信号を処理するオーディオ処理ステップと、前記切り替えステップにて前記周囲音を集音するマイクとして前記バイノーラルマイクに切り替えた場合に、前記周囲音を集音する際の音声モードをバイノーラルモードとして、バイノーラルモードであることを示すバイノーラルフラグ信号を生成するフラグ情報生成ステップと、前記ビデオ処理ステップにて処理された映像信号と前記オーディオ処理ステップにて処理された音声信号と前記バイノーラルフラグ信号とを記録媒体に記録する記録ステップとを有することを特徴とする映像音声記録方法を提供する。

[0014] 本発明は、上述した従来の技術の課題を解決するため、被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者の周囲音を集音した音声信号とが記録された記録媒体を再生する映像音声再生装置において、前記記録媒体に記録された記録信号を再生する再生部と、前記再生部によって再生された前記記録信号から前記映像信号と前記音声信号とを分離する分離化部と、前記分離化部によって分離された前記映像信号を処理するビデオ処理部と、前記分離化部によって分離された前記音声信号を処理するオーディオ処理部と、前記記録媒体に、前記周囲音を集音するマイクとして前記撮影者の耳に装着するバイノーラルマイクを使用したことを示すバイノーラルフラグ信号が記録されているとき、前記バイノーラルフラグ信号を取得するフラグ情報取得部と、前記フラグ情報取得部が前記バイノーラルフラグ信号を取得したとき、前記オーディオ処理部によって処理された前記音声信号をスピーカによって発音した場合に発生するクロストーク信号を打ち消すよう前記音声信号を処理

するクロストークキャンセラとを備え、前記クロストークキャンセラは前記音声信号に対して予め求めたフィルタ特性を畳み込むフィルタを有し、前記フィルタ特性は、測定信号を円筒型構造体に装着した一対のマイクで集音した音声信号を用いて計測した頭部伝達関数に基づいたフィルタ特性であることを特徴とする映像音声再生装置を提供する。

[0015] また、被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者の周囲音を集音した音声信号とが記録された記録媒体を再生する映像音声再生方法において、前記記録媒体に記録された記録信号を再生する再生ステップと、前記再生ステップにて再生された前記記録信号から前記映像信号と前記音声信号とを分離する分離化ステップと、前記分離化ステップにて分離された前記映像信号を処理するビデオ処理ステップと、前記分離化ステップによって分離された前記音声信号を処理するオーディオ処理ステップと、前記記録媒体に、前記周囲音を集音するマイクとして前記撮影者の耳に装着するバイノーラルマイクを使用したことを示すバイノーラルフラグ信号が記録されているとき、前記バイノーラルフラグ信号を取得するフラグ情報取得ステップと、前記フラグ情報取得ステップにて前記バイノーラルフラグ信号を取得したとき、前記オーディオ処理ステップにて処理された前記音声信号をスピーカによって発音した場合に発生するクロストーク信号を打ち消すよう前記音声信号を処理するクロストークキャンセルステップとを有し、前記クロストークキャンセルステップは前記音声信号に対して予め求めたフィルタ特性を畳み込むステップであり、前記フィルタ特性は、測定信号を円筒型構造体に装着した一対のマイクで集音した音声信号を用いて計測した頭部伝達関数に基づいたフィルタ特性であることを特徴とする映像音声再生方法を提供する。

図面の簡単な説明

[0016] [図1]図1は、本発明の第1の実施形態に係る映像音声記録再生装置を示す外観斜視図である。

[図2]図2は、本発明の第1の実施形態に係る映像音声記録再生装置を用いて被写体を撮影している状態を示す図である。

[図3]図3は、本発明の第1の実施形態に係る映像音声記録再生装置の内部構成例

を示すブロック図である。

[図4]図4は、本発明の各実施形態に係る映像音声記録再生装置において音声モードの初期設定を行う表示画面を示す図である。

[図5]図5は、本発明の各実施形態に係る映像音声記録再生装置におけるバイノーラルマイクの表示例を示す図である。

[図6]図6は、本発明の各実施形態に係る映像音声記録再生装置で用いるバイノーラルマイクの変形例を示す図である。

[図7]図7は、本発明の各実施形態に係る映像音声記録再生装置で用いるバイノーラルマイクの変形例を示す図である。

[図8]図8は、本発明の各実施形態に係る映像音声記録再生装置で用いるバイノーラルマイクの変形例を示す図である。

[図9]図9は、本発明の各実施形態に係る映像音声記録再生装置におけるバイノーラルフラグ信号の記述フォーマットの一例を示す図である。

[図10]図10は、本発明の各実施形態に係る映像音声記録再生装置におけるバイノーラルフラグ信号の記述フォーマットの他の例を示す図である。

[図11]図11は、本発明の各実施形態に係る映像音声記録再生装置におけるバイノーラルフラグ信号の記述フォーマットのさらに他の例を示す図である。

[図12]図12は、本発明の第1の実施形態に係る映像音声記録再生装置における記録動作を説明するためのフローチャートである。

[図13]図13は、本発明の第1の実施形態に係る映像音声記録再生装置における再生動作を説明するためのフローチャートである。

[図14]図14は、本発明の各実施形態に係る映像音声記録再生装置で用いるクロストークキャンセラの構成例を示すブロック図である。

[図15]図15は、本発明の各実施形態に係る映像音声記録再生装置で用いるクロストークキャンセラで用いる頭部伝達関数特性を求めるための頭部伝達関数測定装置を示す図である。

[図16]図16は、図15に示す頭部伝達関数測定装置で用いるマイクユニットを取り付けた円筒形構造体と比較のためのダミーヘッドマイクを示す図である。

[図17]図17は、図15に示す頭部伝達関数測定装置で測定したインパルス応答波形を示す波形図である。

[図18]図18は、図15に示す頭部伝達関数測定装置で測定した周波数特性を示す波形図である。

[図19]図19は、ダミーヘッドマイクを用いて測定したインパルス応答波形を示す波形図である。

[図20]図20は、ダミーヘッドマイクを用いて測定した周波数特性を示す波形図である。

[図21]図21は、マイクユニットを取り付けた円筒形構造体を用いて測定した頭部伝達関数に基づくフィルタ特性によるクロストークキャンセル特性を説明するための図である。

[図22]図22は、ダミーヘッドマイクを用いて測定した頭部伝達関数に基づくフィルタ特性によるクロストークキャンセル特性を説明するための図である。

[図23]図23は、マイクユニットを取り付けた円筒形構造体を用いて測定した頭部伝達関数に基づくフィルタ特性によるクロストークキャンセル特性を説明するための図である。

[図24]図24は、ダミーヘッドマイクを用いて測定した頭部伝達関数に基づくフィルタ特性によるクロストークキャンセル特性を説明するための図である。

[図25]図25は、本発明の各実施形態に係る映像音声記録再生装置で用いるクロストークキャンセラの他の構成例を示すブロック図である。

[図26]図26は、本発明の各実施形態に係る映像音声記録再生装置で用いるクロストークキャンセラのさらに他の構成例を示すブロック図である。

[図27]図27は、本発明の各実施形態に係る映像音声記録再生装置におけるヘッドフォンでの再生動作を示すフローチャートである。

[図28]図28は、本発明の第2の実施形態に係る映像音声記録再生装置の内部構成例を示すブロック図である。

[図29]図29は、本発明の第2の実施形態に係る映像音声記録再生装置における音声ズーム処理部の構成例を示すブロック図である。

[図30]図30は、本発明の第2の実施形態に係る映像音声記録再生装置において音声ズーム処理を行う場合の動作を説明するためのフローチャートである。

[図31]図31は、本発明の第2の実施形態に係る映像音声記録再生装置における音声ズーム処理部の他の構成例を示すブロック図である。

[図32]図32は、図31の音声ズーム処理部の他の構成例で用いる頭部伝達関数を求めるための頭部伝達関数測定装置を示す図である。

[図33]図33は、図32の頭部伝達関数測定装置で用いるダミーヘッドマイクを示す断面図である。

[図34]図34は、図32の頭部伝達関数測定装置で測定して得られた頭部伝達関数の特性を示す図である。

[図35]図35は、図32の頭部伝達関数測定装置で測定して得られた頭部伝達関数の特性を示す図である。

[図36]図36は、図32の頭部伝達関数測定装置で測定して得られた頭部伝達関数の特性を示す図である。

[図37]図37は、図32の頭部伝達関数測定装置で測定して得られた頭部伝達関数の特性を示す図である。

[図38]図38は、図32の頭部伝達関数測定装置で測定して得られた頭部伝達関数の特性を示す図である。

[図39]図39は、図32の頭部伝達関数測定装置で測定して得られた頭部伝達関数の特性を示す図である。

[図40]図40は、本発明の第2の実施形態に係る映像音声記録再生装置において図31に示す音声ズーム処理部の他の構成例を用いて音声ズーム処理を行う場合の動作を説明するためのフローチャートである。

[図41]図41は、本発明の第3の実施形態に係る映像音声記録再生装置の内部構成例を示すブロック図である。

[図42]図42は、本発明の第3の実施形態に係る映像音声記録再生装置における音声ズーム処理部の構成例を示すブロック図である。

[図43]図43は、本発明の第3の実施形態に係る映像音声記録再生装置における音

声ズーム処理部の他の構成例を示すブロック図である。

[図44]図44は、本発明の第4の実施形態に係る映像音声記録再生装置の内部構成例を示すブロック図である。

[図45]図45は、本発明の第4の実施形態に係る映像音声記録再生装置における音声ズーム処理部の構成例を示すブロック図である。

[図46]図46は、本発明の第4の実施形態に係る映像音声記録再生装置におけるマニュアル音声ズーム処理を説明するためのフローチャートである。

[図47]図47は、本発明の第5の実施形態に係る映像音声記録再生装置の内部構成例を示すブロック図である。

[図48]図48は、本発明の第5の実施形態に係る映像音声記録再生装置における音声ズーム処理部の構成例を示すブロック図である。

[図49]図49は、本発明の第6の実施形態に係る映像音声記録再生装置を示す外観斜視図である。

[図50]図50は、本発明の第6の実施形態に係る映像音声記録再生装置の内部構成例を示すブロック図である。

[図51]図51は、本発明の第6の実施形態に係る映像音声記録再生装置におけるコード収納部の構造例を示す平面図である。

[図52]図52は、本発明の第7の実施形態に係る映像音声記録再生装置を示す外観斜視図である。

[図53]図53は、本発明の第7の実施形態に係る映像音声記録再生装置の内部構成例を示すブロック図である。

[図54]図54は、本発明の第7の実施形態に係る映像音声記録再生装置における無線型バイノーラルマイクと無線送受信部の具体的構成例を示すブロック図である。

[図55]図55は、本発明の第7の実施形態に係る映像音声記録再生装置において無線型バイノーラルマイクが通信可能圏を越えた場合の警告を説明するための図である。

[図56]図56は、本発明の第7の実施形態に係る映像音声記録再生装置において無線型バイノーラルマイクが通信可能圏を越えた場合に表示部に表示させる警告マー

クの例を示す図である。

[図57]図57は、本発明の第7の実施形態に係る映像音声記録再生装置における動作を説明するためのフローチャートである。

発明を実施するための最良の形態

[0017] 以下、本発明の映像音声記録装置及び方法、並びに、映像音声再生装置及び方法の実施の形態について図面を参照しながら説明する。

[0018] ≪第1の実施形態≫

図1は、第1の実施形態である映像音声記録再生装置101の外観構成例を示す斜視図である。

[0019] 図1に示す映像音声記録再生装置101は、カメラ部11、表示部17、内蔵ステレオマイク21a, 21b、外部マイク接続端子32を備えて構成されている。外部マイク接続端子32には無指向性の左右のマイク31a及び31bを有するイヤホン型のバイノーラルマイク3が着脱自在に接続される。ここでは、バイノーラルマイク3を外部マイク接続端子32に接続した状態を図示している。マイク31a, 31bは内部に振動板を備える。映像音声記録再生装置101は後に詳述するように、内蔵ステレオマイク21a, 21bを用いた撮影(音声収録)とバイノーラルマイク3を用いた撮影(音声収録)とを選択的に行うことができるようになされている。なお、撮影とは、単に被写体の映像を撮像するだけでなく、被写体の映像を撮像することと被写体から発せられた音を含む撮影者の周囲音を集音することの双方を意味する場合がある。

[0020] 図2は、撮影者300が映像音声記録再生装置101を用いて図示しない被写体を撮影している状態を示している。バイノーラルマイク3を用いて集音しながら被写体を撮影する場合には、図2に示すように、撮影者300は左右の耳302に左右のマイク31a, 31bを装着する。これにより、撮影者300の頭部30とマイク31a, 31bとの位置関係により与えられるバイノーラル音響特性で被写体から発せられた音を含む撮影者300の周囲音が集音される。撮影者300は、表示部17に表示される被写体のモニタ映像を目視しながら、被写体をカメラ部11で撮影すると共に周囲音をバイノーラルマイク3で集音する。後に詳述するように、カメラ部11で撮像した映像信号とバイノーラルマイク3で集音した音声信号がここでは図示していない記録媒体に記録される。また

、後に詳述するように、記録媒体に記録された映像信号は、視聴者があたかも撮影者300と同じ撮影環境で視聴しているような臨場感を有する音声と共に再生される。

[0021] 図3は、映像音声記録再生装置101の具体的な内部構成例を示すブロック図である。

[0022] 映像音声記録再生装置101は、カメラ部11、ビデオ符号化部12、多重化部13、記録・再生部14、分離化部15、ビデオ復号化部16、表示部17、内蔵ステレオマイク21(21は21a, 21bの総称)、オーディ符号化部22、オーディオ復号化部26、クロストークキャンセラ27、外部マイク接続端子32、フラグ情報取得部36、映像出力端子37a、音声出力端子37b、接続検出部41、フラグ情報生成部42、記録媒体44、制御部47、操作部48、スイッチSw1, Sw2, Sw3を備えて構成されている。なお、記録媒体44としては、ディスク状記録媒体やテープカセットのような着脱自在の記録媒体でもよいし、ハードディスクのように映像音声記録再生装置101内に予め装着されている記録媒体であってもよい。

[0023] 映像出力端子37aにはテレビジョン受像機等のモニタ52が接続される。音声出力端子37bにはアンプ51を介してスピーカ53, 54が接続される。スピーカ53, 54から発音された音声は視聴者59により受聴される。なお、図3では便宜上、撮影者300と視聴者59とを同時に図示しているが、通常は撮影者300による撮影と視聴者59による再生映像・音声の視聴は別個に行われることは言うまでもない。

[0024] ＜記録動作＞

映像音声記録再生装置101における記録動作について説明する。

[0025] まず、撮影者300が操作部48により音声モードの初期設定画像(ウインドウ)を表示させる操作を行うと、制御部47は一例として図4に示すような初期設定画像170を表示部17に表示させる。外部マイク接続端子32には、外部マイクとして、図1, 図2で説明したようなバイノーラルマイク3と通常の外部マイクとのいずれも接続することができる。操作部48による所定の操作により、撮影者300はバイノーラルマイク3を用いて集音する場合には、図4に示すように“バイノーラル”を選択し、通常の外部マイクを用いて集音する場合には“ノーマル”を選択する。制御部47は、外部マイク接続端子32に接続して使用する外部マイクとして、バイノーラルマイク3かバイノーラルマイク

以外のマイクかを設定する設定部となっている。外部マイク入力として“バイノーラル”が選択され、かつ、接続検出部41が外部マイク接続端子32に外部マイクのプラグが挿入されていることを検出した場合に、制御部47は、映像音声記録再生装置101がバイノーラルマイク3を用いた撮影に対応した記録動作を行うよう回路各部を制御する。バイノーラルマイク3を用いて周囲音を集音して、集音した音声信号を記録する音声モードをバイノーラルモードと称することとする。内蔵ステレオマイク21または通常の外部マイクを用いて周囲音を集音して、集音した音声信号を記録する音声モードをノーマルモードとする。

[0026] なお、バイノーラルマイク3のプラグを通常の外部マイクとは異なるバイノーラルマイク3専用の形状とし、外部マイク接続端子32をバイノーラルマイク3専用の接続端子とすれば、前述の音声モードの初期設定を省略することができる。

[0027] 具体的には、図3において、接続検出部41は外部マイク接続端子32に外部マイクが接続されていることを検出すると、検出信号を制御部47に供給する。制御部47は外部マイクが“バイノーラル”に設定された状態で外部マイクが接続されていることを示す検出信号が入力されたら、スイッチSw1を内蔵ステレオマイク21からの音声信号を入力する端子aから、バイノーラルマイク3からの音声信号を入力する端子bへと切り替える。これにより、バイノーラルマイク3からの音声信号がオーディオ符号化部22へと供給される。スイッチSw1は周囲音を集音するマイクとして、撮影者300の耳302に装着するバイノーラルマイクを使用するかバイノーラルマイク以外のマイクを使用するかを切り替える切り替え部となっている。

[0028] また、制御部47は、フラグ情報生成部42がバイノーラルモードであることを示すフラグ情報（バイノーラルフラグ信号）を生成して出力するよう制御する。バイノーラルフラグ信号は多重化部13に入力される。

[0029] このようにバイノーラルモードの設定がされた場合、制御部47は表示部17にバイノーラルモードとなっていることを示すマークを表示させることが好ましい。図5にマークの例を示す。図5（Λ）に示すマーク171はバイノーラルマイク3を使用する撮影者300を模式的に示すマークであり、図5（B）に示すマーク172はバイノーラル音声のスピーカ再生を模式的に示すマークである。バイノーラルモードを示すマークとしては図

5(A), (B)いずれのマークを用いてもよく、勿論これ以外のマークを用いてもよい。これらのマークは、前述の初期設定終了後や外部マイク接続端子32にバイノーラルマイク3が接続された場合にカメラ部11で撮影された映像に重畳して表示部17に表示させればよい。このようなマーク表示により撮影者300はバイノーラルマイク3を使用する際に、音声モードがバイノーラルモードに設定されているか否かを確認することができる。制御部47は、音声モードがバイノーラルモードのとき、表示部17にバイノーラルモードであることを示すバイノーラルマーク(マーク171,172)を表示させる表示制御部となっている。

[0030] バイノーラルマイク3の左右のマイク31a, 31bを左右の耳302に装着した撮影者300は被写体をカメラ部11で撮像する。カメラ部11から出力された映像信号はビデオ符号化部(ビデオ処理部)12及びスイッチSw3の端子gに入力される。映像音声記録再生装置101で撮影(記録)を行っている場合には、スイッチSw3は端子gに切り替えられていて、カメラ部11からの映像信号が表示部17に供給されて被写体の画像が表示される。これと同時に、撮影者300の頭部30とマイク31a, 31bとの位置関係により、マイク31a, 31からは被写体を正中面方向とするバイノーラル集音した音声信号が得られる。この音声信号はスイッチSw1を介してオーディオ符号化部(オーディオ処理部)22に入力される。

[0031] ここでは記録媒体44がDVカセットの場合について説明すると、ビデオ符号化部12は入力された映像信号をA/D変換すると共にDV圧縮方式により符号化して符号化映像信号を生成する。オーディオ符号化部22は入力された音声信号をA/D変換すると共に非圧縮の音声信号のデータ位置をシャフリングにより並び替えて符号化音声信号を生成する。

[0032] 多重化部13は、入力された符号化映像信号と符号化音声信号とバイノーラルフラグ信号とを民生用ディジタルVCR仕様で定められる信号フォーマットに従って時分割多重して多重化信号を生成する。多重化部13より出力された多重化信号は記録・再生部14に入力される。記録・再生部14は入力された多重化信号を民生用ディジタルVCR仕様で定められた記録フォーマットに従って記録媒体44に記録する。バイノーラルフラグ信号の具体的な記録方法は後に詳述する。

[0033] ここで、バイノーラルマイク3の変形例について説明する。

[0034] ＜バイノーラルマイク3の変形例＞

図6(A)はマイク31a, 31bの第1の変形例であるマイク31cを示しており、図6(B)はマイク31a, 31bの第2の変形例であるマイク31dを示している。図6(A)に示すマイク31cは、撮影者300の耳302に挿入するマイク保持部312と、マイク保持部312の上部に連結して設けられており、振動板等のマイクユニットを収納するマイク収納部311とを備えている。バイノーラルマイク3をマイク収納部311とマイク保持部312とに分離した構造のマイク31cとすることにより、撮影者300はバイノーラルマイク3を装着しても外部から到来する音を良好に聞くことができる。図6(B)に示すマイク31dは、マイク保持部312の下部に連結してマイク収納部311を設けたものである。マイク31dでもマイク31cと同様の効果がある。

[0035] 図7はマイク保持部312の具体的構成例を示す斜視図である。ここではマイク保持部312の上部にマイク収納部311を設けた図6(A)のマイク31cについて示している。図7(A)に示すマイク保持部312は、保持部本体312aに、耳302の内部に向かうに従って径が小さくなるテーパ状の通音孔313aを形成したものである。図7(B)に示すマイク保持部312は、保持部本体312aに円柱状の音孔313bを形成したものである。図7(A)の保持部本体312aは撮影者300の耳302に挿入が容易であり、図7(B)の保持部本体312bは装着時の外音の減衰が小さいという特長を有する。

[0036] 図8は、図6(A)のマイク31cのマイク保持部312の外形を異ならせた例を示している。図8において、(A)は大の外形を有するマイク保持部312、(B)は中の外形を有するマイク保持部312、(C)は小の外形を有するマイク保持部312である。このようにマイク保持部312として複数の大きさのものを用意すれば撮影者300は自分の耳302の大きさに合ったものを選択することができる。なお、図8(A)～(C)はマイク収納部311(マイクユニット)の形状及び大きさは同じであり、マイクユニットの感度及び応答周波数特性も同一である。

[0037] ＜バイノーラルフラグ信号の記録＞

撮影者300がバイノーラルマイク3を装着して集音したバイノーラル音声と内蔵ステレオマイク21で集音したステレオ音声とを識別するため、バイノーラル音声集音時に

はバイノーラル音声と共にバイノーラルフラグ信号を記録媒体44に記録する。バイノーラルフラグ信号はフラグ情報生成部42で生成される。

[0038] まず、記録媒体44がDVカセットの場合のバイノーラルフラグ信号の具体的な記録方法について説明する。

[0039] 図9はDVカセットにオーディオデータを記録する場合のデータフォーマットである。0～89バイトのオーディオデータを記録するに際し、第0，1バイトは同期コードを、第2～4バイトはID(識別)コードを、第5～9バイトはオーディオ用補助データ(AUX)を、第10～81はオーディオデータを、そして第82～89は誤りデータ検出及び訂正用の内符号パリティデータを記録する。フラグ情報生成部42は一例としてバイノーラルモードであればバイノーラルフラグ信号として1を出力し、バイノーラルモードでなければ(ノーマルモードであれば)0を出力する。多重化部13は図9に示すデータフォーマットの信号を生成する。

[0040] 次に、記録媒体44が記録用のディスクの場合のバイノーラルフラグ信号の具体的な記録方法について説明する。記録用のディスクとしては、DVD－RAM,DVD－RW,DVD－Rのような赤色レーザ光を用いて記録再生するディスクと、Blu-ray DiscやHD－DVDのような青色レーザ光を用いて記録再生するディスクがある。ここではこれらの記録用のディスク全般に適用することができるDVDビデオ規格を用いてバイノーラルフラグ信号を重畳する場合について説明する。

[0041] DVDビデオ規格を用いてバイノーラルフラグ信号を重畳する第1の方法としては、DVDビデオ規格におけるDVD-Video zoneにバイノーラルフラグ信号を重畳する方法が考えられる。

[0042] 図10に示すように、DVD規格のボリューム空間(Volume space)は、Volume and File structureとDVD-Video zoneとDVD others zoneとからなる。DVD-Video zoneはVMG(Video Manager)とVTS(Video Title Set)＃1～DVTS＃nを含む。nは2以上の所定の整数である。それぞれのVTSは制御データ(Control Data)とVOBS(Video Object Set)とを含む。VOBSは複数のVOB(Video Object)を含む。VOBSは複数のCELLを含む。CELLは複数のVOBU(Video Object Unit)を含む。VOBUはナビゲーションパック(NV_PACK)，オーディオパック(A_PACK)，ビデオパック(V_PAC

K)を含む。本実施形態では、VOBU内にバイノーラルフラグ信号を含むデータパック(D_PACK)を設ける。

[0043] データパック(D_PACK)は、パックヘッダ, パケットヘッダ, サブストリームID, オーディオフレーム情報, オーディオデータ情報, バイノーラルフラグ信号を含む。バイノーラルフラグを信号は複数のオーディオフレームレイヤからなる。

[0044] このようにDVDビデオ規格に準拠した形式でバイノーラルフラグ信号を含む情報をデータパック(D_PACK)としてパック化し、MPEG多重化する。これにより、DVDビデオ規格との互換性を保ちながら、オーディオ信号のどのオーディオフレーム部分がバイノーラル音声信号であり、どのオーディオフレーム部分が通常のステレオ音声であるのかが分かる。従って、クロストークキャンセル処理をどのオーディオフレーム部分で施せばよいのかを簡単に判別することができる。

[0045] DVDビデオ規格を用いてバイノーラルフラグ信号を重畳する第2の方法としては、DVDビデオ規格におけるDVD others zoneにバイノーラルフラグ信号を重畳する方法が考えられる。DVD others zoneは映像や音声の本体のデータに関連する補助データを記録する領域であり、ユーザデータ記録領域でもある。

[0046] 図10と図11とを比較すれば分かるように、本実施形態ではユーザデータ記録領域であるDVD others zoneのデータ構造をDVD-Video zoneのデータ構造と類似した構造としている。図11に示すように、DVD others zoneもVMG, VTS, VOBS, VOB, CELL, VOBUなる各情報を含んでいる。図11に示すDVD others zoneにおけるそれらの各情報を図10のそれと区別するため、頭にDを付した名称としている。

[0047] 図11に示すように、DVD others zoneはDVMGとDVTS＃1～DVTS＃nを含む。それぞれのDVTSはDVTSI(Video Title Set Information)とDVOBSとを含む。DVOBSは複数のDVOBを含む。DVOBSは複数のDCELLを含む。DCELLは複数のDVOBUを含む。DVOBUは複数のオーディオフレームレイヤを含む。オーディオフレームレイヤはオーディオ信号の符号化パラメータなどのオーディオフレームデータを記録する領域であり、このオーディオフレームレイヤの一部の領域をバイノーラルフラグ信号の記録領域とする。

[0048] このようにDVDビデオ規格に準拠した形式でDVD others zoneにバイノーラルフラ

グ信号を記述することにより、DVD-Video zoneに含まれるオーディオ信号(バイノーラル音声信号または通常のステレオ音声)とバイノーラルフラグ信号とが互いに関連付けられる。DVDビデオ規格との互換性を保ちながら、オーディオ信号のどのオーディオフレーム部分がバイノーラル音声信号であり、どのオーディオフレーム部分が通常のステレオ音声であるのかが分かる。従って、クロストークキャンセル処理をどのオーディオフレーム部分で施せばよいのかを簡単に判別することができる。

[0049] 図10, 図11の場合、例えばスタートボタンを操作して撮影を開始してストップボタンを操作して撮影を終了する間に集音されたオーディオ信号を1または複数のオーディオフレームレイヤとし、それぞれのオーディオフレームレイヤに対して音声モード情報を付加する。そして、音声モード情報内にバイノーラルフラグ信号を記述し、バイノーラル情報パケットとして管理する。バイノーラルフラグ信号をバイノーラル情報パケットとして管理することにより、オーディオフレーム毎に音声モード情報を容易に取得することができる。これにより、記録媒体44を再生する際に、バイノーラル音声信号と通常のステレオ音声信号とが混在して記録されている場合であっても、音声モードに応じて後で詳述するクロストークキャンセラ27の切り替えを適切に行うことができる。音声モード情報の記録は少なくとも撮影を開始する操作毎に行うことが必要であり、所定時間毎に記録することはより好ましい。

[0050] また、記録媒体44として例えば半導体メモリを用いる場合であっても、上記と同様にバイノーラルフラグ信号の記録領域を定義し、記録される音声信号の音声モードを指定することにより、バイノーラルで記録された音声信号を識別し、クロストークキャンセラ27の動作・不動作を適切に切り替えながら音声信号を再生させることができる。

[0051] さらに、MPEGなどの符号化方式の多重化レイヤのユーザデータ内にバイノーラルフラグ信号を記述することも考えられる。一例として、映像及び音声の通信機能を有する携帯電話を用い、送信側の携帯電話からバイノーラルマイク3で集音した音声信号と共に撮影した映像信号を受信側の携帯電話に伝送する場合を想定する。このような場合にも、送信側の携帯電話から受信側の携帯電話へとバイノーラルフラグ信号を伝送することができる。音声信号に対してバイノーラルフラグ信号を付して伝送することにより、バイノーラル音声の場合に臨場感のある音声を再生することが可能となる

。この場合、携帯電話で伝送する映像及び音声パケットデータの所定の箇所にバイノーラルフラグ信号を記述してパケットデータとして伝送する。例えば、MPEG－4で規定される伝送方式を用いる場合には、エレメンタリストリームのユーザデータ記録領域で、図9に示したようなバイノーラルフラグ信号を伝送することができる。MPEG－4標準で規定されるトランスポートストリームを用いる場合には、プライベートデータ領域（private_data_type）にバイノーラルフラグ信号を記述することができる。

[0052] また、映像データ及び音声データを添付ファイルの形式でファイル化されたデータとして伝送する場合は、そのファイルヘッダの部分にバイノーラルフラグ信号を記述するようにしてもよい。

[0053] 映像音声記録再生装置101における記録動作について図12に示すフローチャートを用いてさらに説明する。

[0054] まず、制御部47は、ステップS151にて、図4で説明した初期設定が、外部マイク接続端子32に接続する外部マイクをバイノーラルマイク3に設定されているか否かを判定する。ステップS151で初期設定がバイノーラルに設定されていると判定されれば（YES）、ステップS152に移る。バイノーラルに設定されていると判定されなければ（NO）、制御部47はスイッチSw1を端子aへと切り替え、ステップS154にて、映像音声記録再生装置101は内蔵ステレオマイク21からの音声信号を取得する。制御部47は、ステップS152にて、接続検出部41が外部マイク接続端子32に外部マイクのプラグが挿入されていることを検出しているか否かを判定する。ステップS152で外部マイク接続端子32に外部マイクが接続されている判定されれば（YES）、制御部47は、スイッチSw1を端子bへと切り替え、ステップS153にて、映像音声記録再生装置101はバイノーラルマイク3からの音声信号を取得する。ステップS152で外部マイク接続端子32に外部マイクが接続されていると判定されなければ（NO）、制御部47は、スイッチSw1を端子aへと切り替え、ステップS154にて、映像音声記録再生装置101は内蔵ステレオマイク21からの音声信号を取得する。

[0055] 次に、ステップS155にて、カメラ部11からの映像信号をビデオ符号化部12内の図示しないメモリに一時記憶し、バイノーラルマイク3または内蔵ステレオマイク21からの音声信号をオーディオ符号化部22内の図示しないメモリに一時記憶する。ステッ

プS156にて、ビデオ符号化部12は映像信号を符号化し、オーディオ符号化部22は音声信号を符号化する。テップS157にて、符号化映像信号をビデオ符号化部12内の図示しないバッファに一時記憶し、符号化音声信号をオーディオ符号化部22内の図示しないバッファに一時記憶する。さらに、ステップS158にて、フラグ情報生成部42はバイノーラルモードのとき制御部47の指示に基づいてバイノーラルフラグ信号を生成する。

[0056] そして、多重化部13は、ステップS159にて、符号化映像信号と符号化音声信号とバイノーラルフラグ信号とを多重化し、ステップS160にて、パケット化ストリーム信号を生成する。記録・再生部14は、ステップS161にて、媒体記録44にストリーム信号を記録する。ステップS162にて、ビデオ符号化部12とオーディオ符号化部22は符号化すべき映像信号と音声信号があるか判定し、符号化すべき映像信号と音声信号がまだあると判定されれば(YES)ステップS152に戻り、以上の動作を繰り返す。ステップS162で符号化すべき映像信号と音声信号がまだあると判定されなければ(NO)、処理を終了する。

[0057] ＜再生動作＞

図3に戻り、映像音声記録再生装置101における再生動作について説明する。図3において、操作部48の図示しない再生ボタンが操作されると、記録・再生部14は制御部47の制御により、記録媒体44に記録された記録信号である多重化信号を再生する。記録・再生部14で再生された多重化信号は分離化部15に供給され、分離化部15は入力された多重化信号を符号化映像信号と符号化音声信号とバイノーラルフラグ信号とに分離する。

[0058] 符号化映像信号はビデオ復号化部(ビデオ処理部)16に供給され、符号化音声信号はオーディオ復号化部(オーディオ処理部)26に供給され、バイノーラルフラグ信号はフラグ情報取得部36に供給される。ビデオ復号化部16は入力された符号化映像信号を復号化して映像信号を得る。再生ボタンが操作に伴って、制御部47の制御によりスイッチSw3が端子hに切り替えられる。ビデオ復号化部16より出力された映像信号は表示部17で表示されると共に、映像出力端子37aを介してモニタ52に供給されて表示される。オーディオ復号化部26は入力された符号化音声信号を復号

化して音声信号を得る。音声信号はクロストークキャンセラ27とスイッチSw2の端子cへと供給される。

[0059] クロストークキャンセラ27は、バイノーラル集音された音声信号をスピーカ53, 54で再生する際に発生する、左側のスピーカ54から発音された信号が受聴者59の右側の耳で受聴される第1のクロストーク成分と、右側のスピーカ53から発音された信号が受聴者59の左側の耳で受聴される第2のクロストーク成分とを打ち消すための信号を生成して音声信号に付加することによりクロストーク処理信号を生成する。フラグ情報取得部36は分離化部15で得られたバイノーラルフラグ信号を保持する。制御部47はフラグ情報取得部36がバイノーラルフラグ信号を保持しているか否かによってスイッチSw2を切り替え制御する。即ち、フラグ情報取得部36がバイノーラルフラグ信号を保持していれば、スイッチSw2が端子dに接続してクロストークキャンセラ27からのクロストーク処理信号を音声出力端子37bに対して出力するよう制御し、バイノーラルフラグ信号を保持していなければ、スイッチSw2が端子cに接続してオーディオ復号化部26からのクロストーク処理していない音声信号を音声出力端子37bに対して出力するよう制御する。

[0060] 音声出力端子37bより出力され、アンプ51で増幅された音声信号は左右のスピーカ53, 54から発音される。音声出力端子37bより出力される音声信号がクロストークキャンセラ27から出力されたクロストーク処理信号の場合には、視聴者59は、モニタ52に表示された映像を見ながら、撮影者300が撮影を行った際の撮影者300の周囲音を臨場感のある音声として受聴することができる。このとき、クロストークキャンセラ27は、後に詳述する頭部伝達関数を用いてクロストーク成分を打ち消しているので、撮影者300と視聴者59とが異なる場合でも、任意の撮影者300が撮影して任意の視聴者59が視聴する場合でも、違和感のほとんどない臨場感のある音声を受聴することができる。

[0061] 映像音声記録再生装置101における再生動作について図13に示すフローチャートを用いてさらに説明する。

[0062] 図13において、ステップS181にて、記録・再生部14は記録媒体44を再生し、多重化信号に基づいたストリーム信号を得る。ステップS182にて、記録・再生部14は

ストリーム信号を復号してパケット信号を生成する。ステップS183にて、分離化部15はパケット信号をビデオ信号、オーディオ信号、バイノーラルフラグ信号に分離する。ステップS184にて、ビデオ復号化部16はビデオ信号を復号し、オーディオ復号化部26はオーディオ信号を復号する。ステップS185にて、ビデオ復号化部16とオーディオ復号化部26は復号したビデオ信号とオーディオ信号を図示しないバッファに一時記憶する。ステップS186にて、フラグ情報取得部36はバイノーラルフラグ信号を取得する。

[0063] そして、制御部47は、ステップS187にて、フラグ情報取得部36が取得したバイノーラルフラグ信号に基づいて、再生したオーディオ信号が通常のステレオ音声信号であるかバイノーラル音声信号であるかを判定する。ステップS187でバイノーラル音声信号であると判定した場合には(YES)、ステップS188に移る。ステップS187でバイノーラル音声信号でないと判定した場合には(NO)、ステップS191に移り、制御部47は、スイッチSw2を端子cに切り替えて、ビデオ信号とオーディオ信号を同期再生するよう回路各部を制御する。

[0064] バイノーラル音声信号であれば、ステップS188にて、制御部47は、スイッチSw2を端子dに切り替えることによりクロストークキャンセラ27によるクロストークキャンセル処理を有効な状態とする。ステップS189にて、制御部47は、ビデオ信号とクロストークキャンセラ27によってクロストークキャンセル処理したオーディオ信号を同期再生するよう回路各部を制御する。ステップS190にて、再生すべきビデオ信号とオーディオ信号がまだあると判定されれば(YES)、ステップS182に戻り、以上の動作を繰り返す。ステップS190で再生すべきビデオ信号とオーディオ信号がまだあると判定されなければ(NO)、処理を終了する。

[0065] ＜クロストークキャンセル＞

ここで、図14を用いてクロストークキャンセラ27の具体的構成と動作について説明する。図14に示すように、クロストークキャンセラ27は、フィルタ272a～272d、加算器274a, 274b、フィルタ275a, 275bを備えて構成されている。

[0066] 図14において、バイノーラル音声信号のうち左チャンネルの信号$P_L(t)$はフィルタ272a, 272bに、右チャンネルの信号$P_R(t)$はフィルタ272c, 272dに入力される。フィ

ルタ272a～272dは順に、後述する頭部伝達関数$h_{rs}(t)$, $h_{lo}(t)$, $h_{ro}(t)$, $h_{ls}(t)$に基づいて作成したフィルタ特性(フィルタ係数)を記憶している。フィルタ272a, 272dは頭部伝達関数$h_{rs}(t)$, $h_{ls}(t)$と同等のフィルタ特性を有しており、フィルタ272b, 272cは頭部伝達関数$h_{lo}(t)$, $h_{ro}(t)$を反転したのと同等のフィルタ特性を有する。便宜上、フィルタ272a～272dのフィルタ特性を$h_{rs}(t)$, $-h_{lo}(t)$, $-h_{ro}(t)$, $h_{ls}(t)$と表記する。フィルタ272a～272dは、入力される信号$P_L(t)$, $P_R(t)$にそれぞれのフィルタ特性を与えて出力する。

[0067] 加算器274aはフィルタ272a, 272cから出力される信号を加算し、フィルタ275aは加算された信号にd(t)なるフィルタ特性を与える。加算器274bはフィルタ272b, 272dから出力される信号を加算し、フィルタ275bは加算された信号にd(t)なるフィルタ特性を与える。

[0068] ここで、フィルタ275a, 275bに記憶されているフィルタ特性d(t)は次の(1)式である。

[0069] $$d(t)=\{h_{ls}(t)\times h_{rs}(t)-h_{lo}(t)\times h_{ro}(t)\}^{-1} \quad \cdots(1)$$

フィルタ275a, 275bから出力される信号は、スピーカ53, 54から発音される際にクロストークがキャンセルされるよう、クロストークキャンセル処理が施されたクロストーク処理信号である。フィルタ275a, 275bから出力されたロストーク処理信号は、アンプ51の左チャンネルアンプ51aと右チャンネルアンプ51bでそれぞれ増幅され、スピーカ53, 54から発音される。

[0070] スピーカ53から発音された信号(音声)は視聴者59の左耳で受聴されると共に、発音された信号の一部は破線にて示す第1のクロストーク信号として視聴者59の右耳で受聴される。クロストークキャンセラ27は視聴者59の右耳で受聴される第1のクロストーク信号を打ち消すための第1のクロストーク打消し信号を生成してスピーカ54から発音させる。第1のクロストーク打消し信号により第1のクロストーク信号は打ち消される(減衰される)。同様に、スピーカ54から発音された信号(音声)は視聴者59の右耳で受聴されると共に、発音された信号の一部は破線にて示す第2のクロストーク信号として視聴者59の左耳で受聴される。クロストークキャンセラ27は視聴者59の左耳で受聴される第2のクロストーク信号を打ち消すための第2のクロストーク打消し信号

を生成してスピーカ53から発音させる。第2のクロストーク打消し信号により第2のクロストーク信号は打ち消される(減衰される)。結果として、視聴者59は左耳でクロストークがキャンセルされた音声信号$P_l(t)$を受聴し、右耳でクロストークがキャンセルされた音声信号$P_r(t)$を受聴する。

[0071] ＜頭部伝達関数の測定＞

図15を用いて、フィルタ272a～272d, 275a, 275bに格納する頭部伝達関数特性を求めるための頭部伝達関数測定装置6について説明する。図15に示すように、頭部伝達関数測定装置6は、パーソナルコンピュータ61、増幅器62、スピーカ63、マイクユニット65a, 65b、円筒形構造体65e、増幅器66a, 66bを備えて構成されている。

[0072] 頭部伝達関数の測定方法について説明する。

[0073] まず、パーソナルコンピュータ61により例えばインパルス音よりなる測定信号を生成する。測定信号は増幅器62で増幅される。左側のスピーカ63から発音された測定信号は左右のマイクユニット65a, 65bで受音される。受音して得られた左右の信号は増幅器66a, 66bで増幅されてパーソナルコンピュータ61に入力される。この入力信号はスピーカ63から発音された音声を円筒形構造体65eに取り付けられた左右のマイクユニット65a, 65bで受音した信号の頭部伝達関数$h_{ls}(t)$, $h_{lo}(t)$である。頭部伝達関数$h_{ls}(t)$は左側のスピーカ63から発音された信号が左側のマイクユニット65aで受聴される特性であり、頭部伝達関数$h_{lo}(t)$は左側のスピーカ63から発音された信号が右側のマイクユニット65bで受聴されるクロストーク成分の特性である。

[0074] 同様に、右側のスピーカ64から発音された測定信号は左右のマイクユニット65a, 65bで受音される。受音して得られた左右の信号は増幅器66a, 66bで増幅されてパーソナルコンピュータ61に入力される。パーソナルコンピュータ61は生成した測定信号と受音した信号とを比較することにより、スピーカ64から発音された音声を円筒形構造体65eに取り付けられた左右のマイクユニット65a, 65bで受音した信号の頭部伝達関数$h_{rs}(t)$, $h_{ro}(t)$を求める。頭部伝達関数$h_{rs}(t)$は右側のスピーカ64から発音された信号が右側のマイクユニット65bで受聴される特性であり、頭部伝達関数$h_{ro}(t)$は右側のスピーカ64から発音された信号が左側のマイクユニット65aで受聴さ

れるクロストーク成分の特性である。

[0075] 図16を用いて円筒形構造体65eについて説明する。図16において、(A)は円筒形構造体65eの上面図、(B)は円筒形構造体65eの斜視図であり、(C)は比較のため示すいわゆるダミーヘッドマイクの断面図である。

[0076] 図16(A),(B)に示すように、円筒形構造体65eの表面にはマイクユニット65a,65bが180°離間した状態で装着されている。図示のように、マイクユニット65a,65bは耳介及び外耳道を有していない。マイクユニット65a,65bの図示しない振動板は円筒形構造体65eの表面とほぼ一致する位置に配置されている。一方、図16(C)に示すダミーヘッドマイク69は、人工頭691の両側部に耳介形部材692a,692bと耳道693a,693bとを備え、耳道693a,693bの奥部にマイクユニット694a,694bを備えたものである。マイクユニット694a,694bは人の鼓膜の位置に相当する位置に設けられており、人が聴取する音に近似する音声信号を集音するものである。

[0077] 図16(A),(B)に示す円筒形構造体65eに取り付けられたマイクユニット65a,65bの受音特性は、人それぞれで大きさや形状が異なる耳介や外耳道により与えられる特性差の影響を受けることがなく、頭部伝達関数を計測することができる特性である。スピーカ63,64から発音された音波は円筒形構造体65eにより遮蔽されるものの、円筒形構造体65eに添って回折してマイクユニット65a,65bに到達する。マイクユニット65a,65bはスピーカ63,64から発音されて直接到達する音波と円筒形構造体65eに添って回折して到達する音波とよりなる特性を計測することになる。円筒形構造体65eを用いることにより、平均的な頭部遮蔽特性を有する頭部伝達関数を得ることができる。従って、頭部の大きさや形状が異なり頭部遮蔽特性が異なる種々の視聴者がバイノーラル音声信号を受聴しても、違和感のほとんどない臨場感のある音声を受聴することができる。

[0078] 図17(A)～(D)は、前述の音響信号伝達特性測定装置6が発生したインパルス音に、音響信号伝達特性測定装置6によって測定した円筒形構造体65eの頭部伝達関数$h_{ls}(t)$, $h_{lo}(t)$, $h_{rs}(t)$, $h_{ro}(t)$が畳み込まれたインパルス応答波形を示している。図17(D)は前述の(1)式に示すフィルタ特性d(t)を示している。図17(A)～(E)において、縦軸は所定の出力電圧で正規化した信号電圧の振幅を示し、横軸は測

定された信号を48kHzで標本化した場合のサンプルの個数で表した時間を示している。

[0079] 図18(A)～(E)は、図17(A)～(E)に示す信号をフーリエ解析して得られる周波数特性を示している。図18(A)～(E)において、100Hz, 1kHz, 10kHzの周波数位置を破線の縦線で示しており、縦軸の応答特性における2つの破線の横線間の利得差は10dBである。

[0080] 図14のフィルタ272a～272dには、以上のようにして得られた頭部伝達関数$h_{rs}(t)$, $h_{lo}(t)$, $h_{ro}(t)$, $h_{ls}(t)$に基づいたフィルタ特性が与えられる。前述のように、フィルタ272a, 272dには頭部伝達関数$h_{rs}(t)$, $h_{ls}(t)$に相当するフィルタ特性が与えられ、フィルタ272b, 272cには頭部伝達関数$h_{lo}(t)$, $h_{ro}(t)$を極性反転した$-h_{lo}(t)$, $-h_{ro}(t)$に相当するフィルタ特性が与えられる。

[0081] 比較のため、円筒形構造体65eに取り付けたマイクユニット65a, 65bの代わりに図16(C)に示すダミーヘッドマイク69を用いて測定した特性を図19, 図20に示す。図19は図17と同様に測定して得られた特性である。図17と図19とを比較すれば分かるように、円筒形構造体65eに取り付けたマイクユニット65a, 65bを用いて測定したインパルス応答波形は、ダミーヘッドマイク69を用いて測定したインパルス応答波形よりも、測定用入力信号であるインパルス音に近い波形である。

[0082] 図20はダミーヘッドマイク69を用いて測定した周波数応答特性である。図18と図20とを比較すれば分かるように、円筒形構造体65eに取り付けたマイクユニット65a, 65bを用いて得た特性の方が周波数特性の乱れが小さく平坦に近い。図20(A)～(E)に示すそれぞれの特性は、1．5～7kHzの周波数において応答特性が増強されたり減衰されたりしている。図18(A)～(E)に示す特性の方が応答特性の増強及び減衰の程度が小さい。これは円筒形構造体65e取り付けたマイクユニット65a, 65bを用いた場合には耳介や外耳道により生じる特性の乱れがないためである。ダミーヘッドマイク69では、スピーカ63, 64より発音された音波の一部が耳介で反射して、反射した音波と直接到来する音波とが同相で合成されて増強されたり、逆相で合成されて減衰したりする。また、外耳道の共振や反共振などの影響により特定の周波数で増強されたり減衰したりする。円筒形構造体65eに取り付けたマイクユニット65a, 65b

を用いることによりダミーヘッドマイク69を用いた場合のような悪影響を抑えることができる。

[0083] クロストークキャンセラ27のフィルタ272a～272d及びフィルタ275a，275bに、円筒形構造体65eに取り付けたマイクユニット65a，65bを用いて測定した頭部伝達関数に基づいたフィルタ特性（第1の条件とする）と、ダミーヘッドマイク69を用いて測定した頭部伝達関数に基づいたフィルタ特性とを与え（第2の条件とする）、複数の視聴者による両者の比較視聴テストを行った。視聴者の耳道に細い小型マイクを挿入し、その小型マイクで受音される音が視聴者により受聴される音であるとして受音特性を測定した。

[0084] 図21は、第1の条件で、ある視聴者により測定した特性を示している。図21において、（A）は左側の入力信号$P_L(t)$をインパルス信号、右側の入力信号$P_R(t)$を無信号としてスピーカ53，54から発音させ、視聴者の左耳に装着した小型マイクで受音されたインパルス応答信号の波形である。（B）は（A）と同じ条件で発音させ、視聴者の右耳に装着した小型マイクで受音されたクロストーク成分の波形である。図21（A）では大きなレベルのインパルス応答波形が得られ、図21（B）では小さなレベルの波形が得られている。図21（C）は応答波形の周波数分析結果であり、Caは（A）の応答波形を周波数分析した応答特性であり、Cbは（B）の応答波形を周波数分析した応答特性である。100Hz～2kHzで20dBを越えるクロストークキャンセル効果が得られている。

[0085] さらに、図21において、（D）は左側の入力信号$P_L(t)$を無信号、右側の入力信号$P_R(t)$をインパルス信号としてスピーカ53，54から発音させ、視聴者の左耳に装着した小型マイクで受音されたクロストーク成分の波形である。（E）は（D）と同じ条件で発音させ、視聴者の右耳に装着した小型マイクで受音されたインパルス応答の波形である。図21（D）では小さなレベルの波形が得られ、図21（E）では大きなレベルのインパルス応答波形が得られている。図21（F）は応答波形の周波数分析結果であり、Fdは（D）の応答波形を周波数分析した応答特性であり、Feは（E）の応答波形を周波数分析した応答特性である。100Hz～2kHzで16dB程度のクロストークキャンセル効果が得られている。

[0086] 図22は、第2の条件で、図21と同じ視聴者により測定した特性を示している。測定条件は図21と同一である。図22(C)のクロストークキャンセル効果は約14dBであり、図22(F)のクロストークキャンセル効果は約11dBである。第2の条件では第1の条件と比較して効果が劣っていることが分かる。

[0087] 図23は、図21，図22とは異なる視聴者とし、第1の条件にて図21と同一の測定条件で測定した特性を示している。図23(C)のクロストークキャンセル効果は約22dBであり、図23(F)のクロストークキャンセル効果は約18dBである。異なる視聴者であっても良好な効果が得られている。

[0088] 図24は、図23と同じ視聴者で、第2の条件にて図22と同一の測定条件で測定した特性を示している。図24(C)のクロストークキャンセル効果は約14dBであり、図24(F)のクロストークキャンセル効果は約10dBである。異なる視聴者であっても、第2の条件では第1の条件と比較して効果が劣っていることが分かる。さらに異なる視聴者により同様の測定を行い、第1の条件と第2の条件それぞれで上記と同様の効果が得られることを確認した。

[0089] 以上の測定結果より、クロストークキャンセラ27のフィルタ272a～272d及びフィルタ275a，275bに与えるフィルタ特性は、ダミーヘッドマイク69を用いて測定した頭部伝達関数に基づいたフィルタ特性よりも円筒形構造体65eに取り付けたマイクユニット65a，65bを用いて測定した頭部伝達関数に基づいたフィルタ特性の方が、左側のスピーカから発音されて右側の耳で受聴されるクロストーク成分と右側のスピーカから発音されて左側の耳で受聴されるクロストーク成分を打ち消す効果が大きいことが確認された。

[0090] 円筒形構造体65eに取り付けたマイクユニット65a，65bを用いて測定した頭部伝達関数に基づいたフィルタ特性では、高音域での周波数特性の乱れが少な。即ち、円筒形構造体65eを用いることにより、特定の周波数部分で周波数特性が大きく落ち込んだり上昇したりすることを少なくすることができ、音質の劣化を極力抑えることができる。これにより視聴者はほとんど違和感を覚えることなく臨場感のある音声を受聴することが可能となる。

[0091] クロストークキャンセラ27のフィルタ272a～272d及びフィルタ275a，275bに与え

るフィルタ特性を円筒形構造体65eに取り付けたマイクユニット65a，65bを用いて測定した頭部伝達関数に基づいたフィルタ特性とした場合には、クロストークキャンセルを、バイノーラル音声信号を受聴する際の構造体である視聴者59の外耳道入り口近傍で行うことになる。従って、耳介や外耳道の音響特性が異なる複数の視聴者59に対してクロストーク成分の打消しが効果的になされることになる。

[0092] なお、円筒形構造体65eは完全な円筒形に限定されるものではなく、円筒形を若干変形させた形状であってもよい。耳介や外耳道などが与えるような応答特性の変化を与える凹凸部分を有さない形状とするのが好ましく、円筒形構造体65eにマイクユニット65a，65bを取り付けた状態で応答特性の乱れが極力少ないことが好ましい。

[0093] クロストークキャンセラ27は図14に示す構成に限定されることはなく、低域周波数帯域で生じる逆相感をさらに減少させることができる帯域分割型のクロストークキャンセラとしてもよい。帯域分割型のクロストークキャンセラは、全周波数帯域の信号として入力されるバイノーラル音声信号を低域周波数帯域の信号と中高域周波数帯域の信号とに分割し、中高域周波数帯域のバイノーラル音声信号に対してのみクロストークキャンセル処理を行うものである。

[0094] 図25に帯域分割型のクロストークキャンセラ27aを示し、その構成と動作について説明する。図14に示したクロストークキャンセラ27と同一の機能を有する構成部分には同一符号を付し、その説明を省略する。

[0095] 図25に示すように、クロストークキャンセラ27aは、図14のクロストークキャンセラ27と比較して、ローパスフィルタ(LPF)271a，271d、ハイパスフィルタ(HPF)271b，271c、遅延器273a，273b、ゲインコントロールアンプ(GC)276a～276d、加算器277a，277bが追加されている点で相違している。

[0096] クロストークキャンセラ27aに入力されたバイノーラル音声信号のうち、左チャンネルの信号$P_L(t)$はLPF271aとHPF271bに入力され、右チャンネルの信号$P_R(t)$はLPF271dとHPF271cに入力されて、低域周波数帯と中高域周波数帯とに分割される。LPF271a，271dとHPF271b，271cのカットオフ周波数は100～200Hz程度に設定される。

[0097] HPF271b, 271cより出力された中高域周波数帯の信号は、クロストークキャンセラ27と同様に構成されるフィルタ272a～272dと加算器274a, 274bとフィルタ275a, 275bとよりなる回路部分でクロストークキャンセル処理が施される。クロストークキャンセル処理が施された中高域周波数帯の信号はゲインコントロールアンプ276b, 276cに入力され、ゲインが調整される。

[0098] LPF271a, 271dより出力された低域周波数帯の信号は遅延器273a, 273bに入力され、中高域周波数帯の信号に対するクロストークキャンセル処理に要する時間とほぼ同じ時間だけ遅延される。遅延器273a, 273bより出力された低域周波数帯の信号はゲインコントロールアンプ276a, 276dに入力され、中高域周波数帯の信号とのレベル差がなくなるようゲインが調整される。

[0099] そして、加算器277a, 277bは、ゲインコントロールアンプ276a～276dから出力された低域周波数帯の信号と中高域周波数帯の信号とを加算する。加算器277a, 277bより出力された信号は、中高域周波数帯の信号に対してのみクロストークキャンセル処理が施されたクロストーク処理信号である。加算器277a, 277bから出力されたロストーク処理信号は、アンプ51の左チャンネルアンプ51aと右チャンネルアンプ51bでそれぞれ増幅され、スピーカ53, 54から発音される。

[0100] この図25の構成では、低域周波数帯の信号に対してはクロストークキャンセル処理が施されないため、低域周波数帯で逆相感のない信号として再生される。

[0101] 図14に示すクロストークキャンセラ27では、前述の図21～図24にて説明したように、クロストークキャンセラ27によるクロストークキャンセル効果は100Hz以下では十分に得られていない。100Hz以下の低域周波数帯は音源の定位に与える影響が少ない帯域である。クロストークをキャンセルしない信号は逆相の信号として違和感を有して受聴されてしまう。

[0102] 図25に示すクロストークキャンセラ27aでは、100～200Hz以下の低域周波数帯でクロストークキャンセルを行わないようにし、低域周波数帯で逆相信号が生じないようにしたクロストークキャンセラを実現している。

[0103] 図26に、図25とはフィルタの構成方法が異なる周波数帯域分割型のクロストークキャンセラ27bを示し、その構成と動作について説明する。図25に示したクロストークキ

ャンセラ27aと同一の機能を有する構成部分には同一符号を付し、その説明を省略する。

[0104] 図26に示すクロストークキャンセラ27bは、図25のクロストークキャンセラ27aと比較して、フィルタ272a～272d及びフィルタ275a, 275bの代わりにフィルタ278a, 278b及びフィルタ279a, 279bを備える点で相違しており、結線方法も異なっている。クロストークキャンセラ27aはフィルタ特性をフィードフォワード型のFIR(finite impulse response:有限インパルス応答)形で構成しているのに対し、クロストークキャンセラ27bはフィルタ特性をフィードバック型のFIR形で構成している。

[0105] 図26において、HPF271b, 271cより出力された中高域周波数帯の信号は、FIR形のフィルタ278a, 278b, 279a, 279b及び加算器274a, 274bによってクロストークキャンセル処理が施される。頭部伝達関数測定装置6で求められたフィルタ特性は、フィルタ278a, 278b, 279a, 279b内の図示しない記憶領域に格納され、フィルタ278a, 278b, 279a, 279bは入力される信号にそれぞれのフィルタ特性を畳み込んで出力する。クロストークキャンセラ27bによれば、低域周波数帯で逆相信号が生じないようにして違和感を軽減するという点でクロストークキャンセラ27aと同様の作用効果を奏する。図26に示すクロストークキャンセラ27bの方が図25に示すクロストークキャンセラ27aよりも使用するフィルタの次数を少なくすることができ、構成を簡略化できる。なお、FIR形のフィルタの代わりにIIR形(infinite impulse response:無限インパルス応答)のフィルタを用いてもよい。

[0106] 図3に示す構成においては、クロストークキャンセラ27(または27a, 27b)を制御部47とは別構成としているが、制御部47がDSP(Digital signal processor)を備えたマイクロプロセッサの場合には、クロストークキャンセラ27, 27a, 27bの機能を制御部47に実行させてもよい。また、クロストークキャンセラ27, 27a, 27bはハードウェアに限定されず、ソフトウェアによって実現することができる。

[0107] ＜ヘッドフォン再生＞

ところで、図3に示す映像音声記録再生装置101においては、オーディオ復号化部26より出力される音声信号をヘッドフォンでも受聴できるようにしている。バイノーラル音声信号をヘッドフォンで受聴する場合には、上述したクロストーク成分は生じな

い。クロストークキャンセラ27にて処理したクロストーク処理信号をヘッドフォンで受聴するとバイノーラル音声信号の逆相成分が左右の耳で受聴されてしまう。逆相成分は自然界にない音響信号成分であり、受聴上好ましくないため、バイノーラル音声信号をヘッドフォンでも受聴する場合にはクロストークキャンセル処理を行わない。

[0108] そこで、図3に示すように、オーディオ復号化部26より出力された音声信号を、クロストークキャンセラ27を介することなく音声出力端子37cに出力するようにしている。音声出力端子37cから出力された音声信号はヘッドフォン55へ供給される。視聴者59は前述のようにクロストークキャンセラ27より出力されたクロストーク処理信号をスピーカ53, 54によって受聴することもできるし、クロストークキャンセラ27による処理を施していない音声信号をヘッドフォン55によって受聴することもできる。

[0109] 図27を用いてバイノーラル音声信号のヘッドフォン55での再生手順について説明する。図13に示すフローチャートと同一の処理には同一の符号を付し、その説明を省略することとする。

[0110] 図27において、ステップS181～S186までは図13で説明した通りであり、ステップS192にて、ヘッドフォン55での再生か否かを判定する。ヘッドフォン55の接続端子である音声出力端子37cにプラグが挿入されているかを図示していない接続検出部によって検出すればよい。ステップS192でヘッドフォン再生であると判定されれば(YES)、ステップS193にて、ビデオ信号に同期してクロストークキャンセル処理を行わないバイノーラル音声信号をヘッドフォン55にて再生して、ステップS190に移る。なお、再生した音声信号がバイノーラル音声信号ではない通常のステレオ音声の場合も同様にオーディオ復号化部26より出力された音声信号をヘッドフォン55に供給すればよい。

[0111] ステップS192でヘッドフォン再生であると判定されなければ(NO)、図13と同様、ステップS187～S190の処理を行う。ステップS189での処理は、ステップS193でのヘッドフォン55による再生処理とは異なり、バイノーラル音声信号をスピーカ53, 54によって再生する処理である。

[0112] ≪第2の実施形態≫

撮影者300は左右の耳302にバイノーラルマイク3を装着して集音しながら被写体

の撮影及び記録媒体44への記録を行うことにより、視聴者59は撮影者300が撮影を行った際の全方向の周囲音を受聴することができる。ところが、通常の映像音声記録再生装置(ビデオカメラ)により撮影される映像はカメラ前方約60度の範囲の映像である。ズーム撮影を行う場合にはさらに狭い視野角で撮影される。そこで、ズーム撮影を行う場合には、ズームアップされた被写体の近傍から発音される音声を増強して集音できることが望ましい。第2の実施形態は、被写体のズームアップに併せて、被写体近傍の音声を増強して集音して記録するように構成したものである。

[0113] 図28に第2の実施形態である音声のズーム処理部を備えた映像音声記録再生装置102を示し、その構成と動作について説明する。図3に示した第1の実施形態の映像音声記録再生装置101と同じ機能を有する構成部分には同一符号を付し、その説明を省略する。映像音声記録再生装置102は映像音声記録再生装置101と比較して音声ズーム処理部33を備える点が相違している。なお、図28では、ヘッドフォン55とヘッドフォン55の接続端子である音声出力端子37cの図示を省略している。

[0114] 図28において、外部マイク接続端子32を介して入力されるバイノーラルマイク3からの音声信号は音声ズーム処理部33に入力される。カメラ部11は複数のレンズ(図示せず)を備え、一または複数のレンズを移動させてレンズ間距離を可変することによって被写体をズームアップ/ダウンするズーム機能を有している。操作部48によってズームアップ操作がなされると、制御部47からズームアップ制御信号がカメラ部11に入力されて被写体がズームアップされて撮像されると共に、ズームアップ制御信号が音声ズーム処理部33に入力されて入力された音声信号に対して音声のズームアップ処理が施される。

[0115] 音声ズーム処理部33は、ズームアップ制御信号に基づいて、バイノーラル音声信号のうち被写体近傍から発せられる音声信号を多く含む撮影者300の正中面で集音される音声信号を増強してズームアップ音声信号を生成する。ズームアップ音声信号はスイッチSw1を介してオーディオ符号化部22に入力される。ズームアップ撮像された被写体の映像信号はビデオ符号化部12で符号化され、ズームアップ音声信号はオーディオ符号化部22で符号化され、第1の実施形態と同様に記録媒体44に記録される。

[0116] 図29は音声ズーム処理部33の具体的な構成例である。図29に示すように、音声ズーム処理部33は、ズーム倍率検出器331，係数算出器332，加算器335、可変増幅器337、加算器338a，338bを備えて構成されている。

[0117] 図29において、ズーム倍率検出器331は制御部47から供給されるズームアップ制御信号を基にズーム倍率を検出する。係数算出器332は検出されたズーム倍率に応じて被写体近傍から発せられる音声の増強度を表す係数αを算出する。加算器335は外部マイク接続端子32より入力された左右チャンネルのバイノーラル音声信号を加算する。可変増幅器337は加算器335の出力信号を係数算出器332からの係数αに従って増幅する。加算器338a，338bは、可変増幅器337の出力信号にそれぞれ左，右チャンネルのバイノーラル音声信号を加算する。バイノーラルマイク3のマイク31a，31bを撮影者300の左右の耳302に装着した状態でマイク31a，31b内の振動板はほぼ平行となる。撮影者300の左右方向からの音声は互いに逆相の成分を含むことがあるので、加算器335での左右チャンネルの加算の結果、左右方向からの音声は部分的に相殺されて減衰する。従って、音声ズーム処理部33からは、撮影者300の正中面で集音される音声が増強されたズームアップ音声信号が出力されることになる。

[0118] 図30を用いて、音声ズーム処理部33での動作を含む映像音声記録再生装置102の動作についてさらに説明する。図30において、ステップS201にて、加算器335はバイノーラルマイク3からの左右チャンネルのバイノーラル音声信号を加算して加算信号Sを得る。ステップS202にて、ズーム倍率検出器331は操作部48の操作に応じて制御部47で得られるズーム倍率を検出する。ズーム倍率をカメラ部11のレンズを駆動するモータへの印加電圧と駆動時間との関係から求めるようにしてもよい。ステップS203にて、係数算出器332はズーム倍率に応じて増強度を表す係数αを算出する。ステップS204にて、可変増幅器337は加算器335の出力信号に係数αを乗じてαSを求める。ステップS205にて、加算器338a，338bは、左，右チャンネルのバイノーラル音声信号と可変増幅器337の出力信号（αS）とを加算する。そして、ステップS206にて、ズームアップ音声信号を記録媒体44に記録する。ステップS207にて、制御部47は記録が終了したか判定し、終了していなければ（NO）、ステップ

S201から繰り返す。ステップS207で記録が終了していれば(YES)、音声ズーム処理部33での処理を終了する。

[0119] 図31は、音声ズーム処理部33の他の構成例を示す音声ズーム処理部33aを示している。正中面の音声信号に対して音源が受聴者に接近したような効果を与えるような頭部伝達関数を畳み込むと、カメラ部11での被写体のズームアップに連動して音源があたかも受聴者に接近したように感じられ、さらに臨場感を有する音声信号を受聴することができる。図31に示す音声ズーム処理部33aは、正中面の音声信号に音源を接近させるような効果を与える頭部伝達関数を畳み込むように構成したものである。

[0120] 図31に示す音声ズーム処理部33aは、図29の音声ズーム処理部33と比較して、関数取得器333, 伝達関数メモリ334, 畳み込み演算器336が追加されている点で相違している。

[0121] 図31において、伝達関数メモリ334には音源を仮想的に近接して定位させた仮想音源を形成するための頭部伝達関数を記憶しておく。その頭部伝達関数は仮想音源から発せられる音の受聴特性を得るための関数であり、受聴特性は仮想音源と受聴者との距離によって決まる。

[0122] 関数取得器333は、係数算出器332で算出される係数αに従って想定される音源の定位位置に対応する頭部伝達関数を伝達関数メモリ334から取得する。なお、図29における係数αと図31及び他の図における係数αとは必ずしも同じものではないが、便宜上同一の符号を付す。畳み込み演算器336は、加算器335で生成されたバイノーラル音声の加算信号に関数取得器333で取得された頭部伝達関数を畳み込む。可変増幅器337は、頭部伝達関数が畳み込まれた加算信号を係数算出器332で得られた係数αに従って増幅する。加算器338a, 338bは、左, 右チャンネルのバイノーラル音声信号と可変増幅器337の出力信号とを加算する。ここでは、可変増幅器337を備える構成としたが、音源を仮想的に近接して定位さるのみでも十分に効果的であり、可変増幅器337を省略することも可能である。また、関数取得器333が頭部伝達関数を選択するための係数αと、可変増幅器337における増幅の程度を決める係数αとを別の係数としてもよい。

[0123] 図32を用いて、仮想音源を形成するための頭部伝達関数の測定方法について説明する。

[0124] 図32に示す頭部伝達関数測定装置6aは、パーソナルコンピュータ61、増幅器62、スピーカ63、増幅器66a, 66b、ダミーヘッドマイク68を備えて構成されている。ダミーヘッドマイク68は、人工頭681にマイクユニット684a, 684bを装着している。頭部伝達関数測定装置6aは、図15に示した頭部伝達関数測定装置6と比較して、円筒形構造体65eに装着したマイクユニット65a, 65bの代わりにダミーヘッドマイク68を用いている点、及び、スピーカを左右のスピーカ63, 64の一方のみとして(ここではスピーカ63を使用)ダミーヘッドマイク68の正中面に配置している点で相違している。

[0125] 図33にダミーヘッドマイク68の断面図を示す。ダミーヘッドマイク68は、耳介形部材682a, 682bを備える人工頭681の耳道683a, 683bの入口近傍にマイクユニット684a, 684bを配置したものである。図16(C)に示したダミーヘッドマイク69は、耳道693a, 693bの奥部の人の鼓膜の位置に相当する位置にマイクユニット694a, 694bを配置したものであり、ダミーヘッドマイク68では耳道683a, 683bの入口近傍にマイクユニット684a, 684bを配置している点でダミーヘッドマイク69と相違している。なお、通常、ダミーヘッドマイクとは、図16(C)に示すように耳道693a, 693bの奥部の人の鼓膜の位置に相当する位置にマイクユニット694a, 694bを配置するものをいうが、耳介形部材682a, 682bを備える人工頭681の耳道683a, 683bの入口近傍にマイクユニット684a, 684bを配置した図33のものも便宜上ダミーヘッドマイクと称している。

[0126] ダミーヘッドマイク68を用いることにより、スピーカ63からの音声を耳道683a, 683bによる影響が除かれたバイノーラル音声として集音することができる。

[0127] 図32に戻り、パーソナルコンピュータ61により例えばインパルス音よりなる測定信号を生成する。測定信号は増幅器62で増幅される。スピーカ63から発音された測定信号はダミーヘッドマイク68の左右のマイクユニット684a, 684bで受音される。受音して得られた左右の信号は増幅器66a, 66bで増幅されてパーソナルコンピュータ61に入力される。パーソナルコンピュータ61は生成した測定信号と受音した信号とを

比較することにより、ダミーヘッドマイク68の頭部伝達関数$h_l(t)$, $h_r(t)$を求める。頭部伝達関数$h_l(t)$は左側のマイクユニット684aでの受音信号により求めた頭部伝達関数、頭部伝達関数$h_r(t)$は右側のマイクユニット684bでの受音信号により求めた頭部伝達関数である。スピーカ63とダミーヘッドマイク68との距離Dを、例えば0．5m，1m，2mなどのように設定し、それぞれの距離Dにおける頭部伝達関数を順次求める。

[0128] 図34～図39に、図32に示した頭部伝達関数測定装置6aで測定して得られた頭部伝達関数の特性を示す。

[0129] 図34(A)に示すインパルス応答波形は、スピーカ63とダミーヘッドマイク68との距離Dが50cmの場合に左側のマイクユニット684aで受音された波形である。縦軸は正規化された振幅（電圧）である。横軸は時間であり、標本化周波数が48kHzである信号の標本点の数で示してある。図34（B）は図34（A）に示すインパルス応答波形をパーソナルコンピュータ61でフーリエ解析して得られた周波数応答特性である。横軸が周波数（Hz）であり、縦軸は応答特性である。

[0130] 図35（A）は距離Dが50cmの場合に右側のマイクユニット684bで受音されたインパルス応答波形である。図35（B）は図35（A）に示すインパルス応答波形をフーリエ解析して得られた周波数応答特性である。測定条件は図34と同じである。

[0131] 以下、同様にして、図36（A）は距離Dが1mの場合に左側のマイクユニット684aで受音されたインパルス応答波形、図36（B）はその周波数応答特性である。

[0132] 図37（A）は距離Dが1mの場合に右側のマイクユニット684bで受音されたインパルス応答波形、図37（B）はその周波数応答特性である。

[0133] 図38（A）は距離Dが2mの場合に左側のマイクユニット684aで受音されたインパルス応答波形、図38（B）はその周波数応答特性である。

[0134] 図39（A）は距離Dが2mの場合に右側のマイクユニット684bで受音されたインパルス応答波形、図39（B）はその周波数応答特性である。

[0135] それらの特性を比較すると、図34～図39の（A）に示すインパルス応答波形は距離Dの値を0．5m，1m，2mと増加させると波形の振幅は減少する。図34～図39の（B）に示す周波数応答特性は、距離Dが0．5mの場合は破線の楕円で囲んだ1kHz～

4kHzの周波数部分で約400Hzの間隔の規則的な凹凸特性(ピークディップ)が生じており、距離Dが1mの場合ではやや不規則な凹凸特性となり、距離Dが2mの場合では周波数間隔が異なる複数の凹凸特性が組み合わされた特性となっている。また、距離Dが同一の場合は左右でほぼ同一の特性が得られている。

[0136] パーソナルコンピュータ61は、生成したインパルス音の測定信号と増幅器66a, 66bから入力されるインパルス応答信号の波形とを比較して、距離Dごとの頭部伝達特性を求める。距離Dごとに求めた頭部伝達特性が音源を仮想的に距離Dの位置に定位させて、その仮想的な音源からの音声信号を受聴者に受聴させるための特性である。なお、本実施形態では距離Dを0.5m, 1m, 2mとしたが、さらに多くの距離においてそれぞれの特性を求めてもよく、距離Dの間隔を0.5mよりも狭くしてもよい。

[0137] 以上のようにして得られた頭部伝達特性は図31の伝達関数メモリ334に記憶される。ズーム倍率検出器331で検出されたズーム倍率に対して、記憶された伝達関数のいずれを用いるかは、一例として、カメラ部11における図示しない自動焦点測定機能により計測される被写体までの距離をズーム倍率で除して得られる係数αを用いる。例えば被写体までの距離が10mでズーム倍率が5の場合、係数αは2となる。被写体までの距離が10mでズーム倍率が10の場合、係数αは1となり、ズーム倍率が20の場合には係数αは0.5となる。

[0138] 図40を用いて、音声ズーム処理部33aでの動作を含む映像音声記録再生装置102の動作についてさらに説明する。図40において、ステップS211にて、加算器335はバイノーラルマイク3からの左右チャンネルのバイノーラル音声信号を加算して加算信号Sを得る。ステップS212にて、ズーム倍率検出器331は操作部48の操作に応じて制御部47で得られるズーム倍率を検出する。ステップS213にて、係数算出器332はズーム倍率に応じて伝達関数メモリ334に記憶された複数の伝達関数の内のいずれを選択するかを決めたり、可変増幅器337での増強度を表す係数αを算出したりする。なお、係数αとしては被写体までの距離をズーム倍率で除して得られる数をそのまま用いてもよいし、被写体までの距離をズーム倍率で除して得られる数に基づいて生成した値であってもよい。

[0139] ステップS214にて、関数取得器333は係数αに応じて伝達関数メモリ334より伝

達関数を取得し、畳み込み演算器336は加算器335から出力された加算信号に伝達関数を畳み込む。ステップS215にて、可変増幅器337は畳み込み演算器336の出力信号に係数αを乗じて増強する。ステップS216にて、加算器338a，338bは、左，右チャンネルのバイノーラル音声信号と可変増幅器337の出力信号とを加算する。そして、ステップS216にて、ズームアップ音声信号を記録媒体44に記録する。ステップS218にて、制御部47は記録が終了したか判定し、終了していなければ(NO)、ステップS211から繰り返す。ステップS218で記録が終了していれば(YES)、音声ズーム処理部33aでの処理を終了する。

[0140] ≪第3の実施形態≫

第2の実施形態は音声のズームアップ処理を記録側で行うものであるが、第3の実施形態は音声のズームアップ処理を再生側で行うものである。図41に示す第3の実施形態の映像音声記録再生装置103において、図3に示した第1の実施形態の映像音声記録再生装置101と同じ機能を有する構成部分には同一符号を付し、その説明を省略する。映像音声記録再生装置103は映像音声記録再生装置101と比較して、音声ズーム処理部33bを分離化部15の後段に設けており、ズーム倍率検出器331を多重化部13の前段に設けている点が相違している。なお、図41では、ヘッドフォン55とヘッドフォン55の接続端子である音声出力端子37cの図示を省略している。

[0141] 映像音声記録再生装置103の動作について説明する。操作部48が操作され、制御部47で生成されるカメラ部11のレンズ駆動用信号はカメラ部11とズーム倍率検出器331に供給される。ズーム倍率検出器331は、レンズ駆動用信号のズーム方向，ズーム速度，レンズ駆動時間を解析してズーム倍率を検出する。検出されたズーム倍率を示すズーム倍率情報は多重化部13に供給される。多重化部13は符号化映像信号と符号化音声信号とバイノーラルフラグ信号に併せてズーム倍率情報も多重化する。記録・再生部14はズーム倍率情報も含む多重化信号を記録媒体44に記録する。

[0142] 記録・再生部14は記録媒体44に記録された多重化信号を再生し、分離化部15は符号化映像信号と符号化音声信号とバイノーラルフラグ信号に併せてズーム倍率情

報も分離する。ズーム倍率情報は音声ズーム処理部33bに入力される。

[0143] 図42は音声ズーム処理部33bの具体的な構成例である。図42に示すように、音声ズーム処理部33bは図29の音声ズーム処理部33と比較して、ズーム倍率検出器331が省かれている点と加算器338a，338bに入力される信号がクロストークキャンセラ27の出力信号である点で相違している。

[0144] 図42において、係数算出器332は分離化部15で分離されて入力されたズーム倍率情報を用いて可変増幅器337で入力信号を増幅する際に用いる係数αを算出する。加算器335はオーディオ復号化部26から入力されるバイノーラル音声信号を加算する。可変増幅器337は加算器335からの出力信号を係数算出器332から入力された係数αに応じて増幅する。加算器338a，338bはクロストークキャンセラ27からの出力信号と可変増幅器337からの増幅信号とを加算する。

[0145] ところで、カメラ部11におけるズーム処理を例えばDSPを用いてソフトウェアによって演算処理して実行させる場合がある。音声ズーム処理を再生時に行う場合にはズーム処理に必要なDSPの信号処理時間を確保する必要がない。そのため、記録時にDSPによって撮影した映像信号の最適化処理や映像信号の符号化、または記録制御などの信号処理を十分に行うことができる。音声ズーム処理を再生時に行うことにより、DSPでの演算工数がズーム処理のための演算処理に割り当てられて、記録を行うための演算時間が不足するという不具合が発生するおそれを回避することができる。

[0146] 図43は、図42の音声ズーム処理部33bに対して、図31と同様、正中面の音声信号に近接効果を与える頭部伝達関数を畳み込むように構成した音声ズーム処理部33cを示している。音声ズーム処理部33cは音声ズーム処理部33bと比較して、関数取得器333、伝達関数メモリ334、畳み込み演算器336が追加されている点で相違している。関数取得器333、伝達関数メモリ334、畳み込み演算器336の動作は図31と同様であり、説明を省略する。

[0147] ≪第4の実施形態≫

図44に示す第4の実施形態の映像音声記録再生装置104は、音声のズームアップ処理を記録媒体44の再生時に外部操作により手動で行うよう構成したものである。

即ち、ズーム倍率情報が記録媒体44に記録されていない場合に、視聴者59がモニタ52で再生した映像信号を見ながら音声のズームアップ処理を行うものである。視聴者59による手動で行う音声のズームアップ処理をマニュアル音声ズーム処理と称することとする。

[0148] 図44に示す映像音声記録再生装置104は、映像音声記録再生装置103と比較して、ズーム倍率検出器331が省略され、音声ズーム処理部33bの代わりに音声ズーム処理部33dを備えている点で相違している。

[0149] 視聴者59が操作部48を操作してマニュアル音声ズームを行うよう指示すると、制御部47はズームアップ制御信号を音声ズーム処理部33dに入力される。音声ズーム処理部33dはズームアップ制御信号に従ってオーディオ復号化部26で復号されたバイノーラル音声信号に対してズームアップ処理を施す。

[0150] 図45は音声ズーム処理部33dの具体的な構成例である。図45に示すように、音声ズーム処理部33dは、図42のズーム処理部33bと比較して、制御部47からのズームアップ制御信号が入力されるズーム倍率検出器331aが設けられ、係数算出器332には分離化部15からのズーム倍率情報ではなくズーム倍率検出器331aで生成したズーム倍率情報を入力している点で相違している。その他の部分はズーム処理部33bと同様の動作であり、説明を省略する。

[0151] 図46を用いて第4の実施形態におけるマニュアル音声ズーム処理についてさらに説明する。図46において、ステップS221にて、加算器335は再生して得られたバイノーラル音声信号の左右の信号を加算して加算信号Sを得る。ステップS222にて、制御部47は操作部48によって音声ズーム倍率が可変されたか否かを判定する。ステップS222で音声ズーム倍率が可変された場合は(YES)、ステップS223に移り、可変されなかった場合は(NO)、ステップS226に移る。

[0152] 音声ズーム倍率が可変された場合にはステップS223にて、ズーム倍率検出器331aはズームアップ制御信号に基づいてズーム倍率を算出する。ステップS224にて、係数算出器332はズーム倍率検出器331aから入力されたズーム倍率に基づいて係数αを算出する。係数αは、音源を視聴者の前方に定位させるための頭部伝達関数の特性を含んでいてもよい。ステップS225にて、係数αの値を新たに算出された

値に更新する。

[0153] そして、ステップS226にて、可変増幅器337は加算信号Sに係数αを乗算してαSを生成する。なお、ステップS223～S225を経ていない場合は、係数αは音声ズーム倍率が可変する前の値である。ステップS227にて、加算器338a，338bはクロストークキャンセラ27によってクロストークキャンセル処理が施されたバイノーラル音声信号に信号αSを加算する。ステップS228にて、スイッチSw2及び音声出力端子37bを介してステップS227で得られた音声信号を出力する。ステップS229にて、制御部47は再生が終了したか否かを判定し、終了していなければ(NO)、ステップS221に戻り、終了していれば(YES)、処理を終了する。

[0154] ≪第5の実施形態≫

図47に示す第5の実施形態の映像音声記録再生装置105は、ズームアップ音声信号をヘッドフォン55で受聴するのに好適な構成を示している。図47に示す映像音声記録再生装置105は、図41の映像音声記録再生装置103と比較して、音声ズーム処理部33bの代わりに音声ズーム処理部33eを備えており、音声ズーム処理部33eから出力された音声信号を音声出力端子37cを介してヘッドフォン55へと供給している点で相違している。ヘッドフォン55には、クロストークキャンセラ27によってクロストークキャンセル処理を施しておらず、音声ズーム処理部33eによってズームアップ処理を施したバイノーラル音声信号を供給する。

[0155] 図48は音声ズーム処理部33eの具体的な構成例である。音声ズーム処理部33eは図42の音声ズーム処理部33bと比較して、加算器338c，338dが追加されている点で相違している。加算器338c，338dは、オーディオ復号化部26で復号して得られたバイノーラル音性信号に可変増幅器337から出力されたズームアップ音声信号を加算する。加算器338c，338dによって加算されたヘッドフォン受聴用音声信号は音声出力端子37cを介してヘッドフォン55へと供給される。

[0156] 以上説明した第2～第5の実施形態によるズームアップ音声信号の再生効果は以下の通りである。

[0157] カメラ部11の視野角がワイドに設定され、ズーム倍率が小さい場合には加算器335で加算された信号は可変増幅器337で増強されないため、視聴者59は、モニタ52

に表示された映像信号と撮影者300を囲む360度の音声信号をスピーカ53，54で再生した臨場感を有する音声信号とを視聴することになる。視野角をワイドに設定した場合でも視野角は60度程度である。視聴者59は、映像の視野角と音声信号を集音した際の角度範囲との違いにより、モニタ52に表示される被写体から発する音が不足するいわゆる中抜け音声であると認識する場合がある。これに対して、前述の如く音声ズームアップ処理が施された音声信号は、撮影者300に対して正中面からの信号成分を増強してバイノーラル音声信号に加算するため、中抜けが補償された音声信号となる。よって、視聴者59が中抜け音声であると認識することを防止し、視聴者59は第1実施例と比較してさらに違和感の少ない臨場感のある音声を受聴することが可能となる。

[0158] ≪第6の実施形態≫

図49，図50に示す第6の実施形態の映像音声記録再生装置106は、第1～第5の実施形態のように内蔵ステレオマイク21とバイノーラルマイク3とを別々に設けたものではなく、通常のステレオ集音用のマイクとバイノーラルマイクとを共用させた構成を示している。図49は第6の実施形態も映像音声記録再生装置106の外観構成例を示す平面図であり、図50は映像音声記録再生装置106の具体的な内部構成例を示すブロック図である。図49，図50において、図1，図3と同一の機能を有する構成部分には同一符号を付し、その説明を省略する。

[0159] 図49に示すように、映像音声記録再生装置106は上面にマイク31e，31fを載置するためのマイク載置部35a，35bとマイク31e，31fに接続されたマイクコード310e，310fを収納するコード収納部34を備えている。

[0160] 図49において、マイク31e，31fを用いて通常のステレオ音声にて集音する場合は、撮影者300はマイク31e，31fをマイク載置部35a，35bに載置して集音する。バイノーラル音声の集音を行う場合は、撮影者300はコード収納部34からマイクコード310e，310fを引き出してマイク31e，31fを耳302に装着して集音する。映像音声記録再生装置106はマイク載置部35a，35bに載置されるマイク31e，31fを検出する図示しない突起状検出部を有する。突起状検出部に連動して動作するスイッチ（図50のスイッチSw4に相当）のオン／オフにより、映像音声記録再生装置106はマイク3

1e, 31fがマイク載置部35a, 35bに載置されているか否かを検出する。マイク載置部35a, 35bに載置されているか否かの検出はこれに限定されることはなく、例えばマイク31e, 31fに内蔵されている永久磁石から発生する磁界をホール素子や磁気抵抗素子を用いて検出するようにしてもよい。

[0161] 具体的には、図50において、スイッチSw4は、マイク31e, 31fがマイク載置部35a, 35bに載置されていなければ端子eに接続してオフ状態となり、マイク31e, 31fが載置されていれば端子fに接続してオン状態となる。載置検出部41aは、スイッチSw4が端子eに接続しているか端子fに接続しているかによってマイク31e, 31fがマイク載置部35a, 35bに載置されているか否かを検出する。載置検出部41aによる検出信号は制御部47へ供給される。

[0162] 載置検出部41aが、マイク31e, 31fが載置されていることを検出した場合には、マイク31e, 31fは通常のステレオ集音を行い、制御部47は、映像音声記録再生装置106がノーマルモードでの撮影に対応した記録動作を行うよう回路各部を制御する。この場合、マイク31e, 31fの役割は図3における内蔵ステレオマイク21と等価であり、フラグ情報生成部42はバイノーラルモードであることを示すバイノーラルフラグ信号を生成しない。一方、載置検出部41aが、マイク31e, 31fが載置されていないことを検出した場合には、撮影者300がマイク31e, 31fを耳302に装着してバイノーラル集音するバイノーラルモードであるとみなし、制御部47は、映像音声記録再生装置106がバイノーラルモードでの撮影に対応した記録動作を行うよう回路各部を制御する。この場合、フラグ情報生成部42は制御部47による制御に基づいてバイノーラルフラグ信号を生成する。

[0163] 第6の実施形態では、マイク31e, 31f、マイク載置部35a, 35b、載置検出部41a、制御部47が全体として、周囲音を集音するマイクとして、撮影者の耳に装着するバイノーラルマイクを使用するかバイノーラルマイク以外のマイクを使用するかを切り替える切り替え部として動作している。

[0164] ここで、図51を用いてコード収納部34の構造の一例について説明する。図51(A)はコード収納部34を上面側から見た内部構造であり、マイクコード310e, 310fは回転軸343を有するリール341に巻き付けられている。図51(B)はコード収納部34を

下面側から見た内部構造であり、リール341の内部には渦巻状のスプリング342が備えられている。マイク31e，31fがマイク載置部35a，35bに載置されているか否かを検出する代わりに、またはこれに加えて、リール341の回転角を検出することによりバイノーラルモードであるか否かを検出してもよい。

[0165] ≪第7の実施形態≫

図52に示す第7の実施形態の映像音声記録再生装置107は、集音して得られた音声信号を装置本体に無線伝送する無線型バイノーラルマイクを用いた構成を示している。図52において、図1と同一の機能を有する構成部分には同一符号を付し、その説明を省略する。

[0166] 図52において、映像音声記録再生装置107は、図1の外部マイク接続端子32の代わりに無線送受信部39を備え、バイノーラルマイク3の代わりに無線型バイノーラルマイク38を用いて音声収録を行う。撮影者300は、装置本体と無線結合される無線型バイノーラルマイク38を頭部に装着し左右のマイク38a，38bを耳302に挿入して集音することにより、マイクコードに邪魔されることなく撮影を行うことができる。また、撮影者300と、図示しない集音者の2人による撮影も可能となる。

[0167] 図53を用いて映像音声記録再生装置107の内部構成について説明する。図53において、図3と同一の機能を有する構成部分には同一符号を付し、その説明を省略する。図53に示す映像音声記録再生装置107は、図3の映像音声記録再生装置101と比較して、外部マイク接続端子32及び接続検出部41の代わりに無線送受信部39を備えている点で相違している。

[0168] 制御部47は、無線型バイノーラルマイク38が装置本体から所定の距離にあると判断し、無線送受信部39が無線型バイノーラルマイク38からバイノーラル音声信号を受信する場合に、スイッチSw1を端子bに接続させて無線型バイノーラルマイク38からバイノーラル音声信号をオーディオ符号化部22に供給させる。この場合、制御部47は、フラグ情報生成部42がバイノーラルフラグ信号を生成するよう制御する。また、制御部47は、無線型バイノーラルマイク38が装置本体から所定の距離を越えると判断した場合には、スイッチSw1を端子aに接続させ、内蔵ステレオマイク21からのステレオ音声信号をオーディオ符号化部22に供給させる。この場合、フラグ情報生成

部42はバイノーラルフラグ信号を生成しない。

[0169] 図54に無線型バイノーラルマイク38及び無線送受信部39の内部構成例を示し、それらの動作について説明する。

[0170] 図54に示すように、無線型バイノーラルマイク38のマイク38aは、マイクユニット381、マイクアンプ382、送受信ユニット383、アンテナ384、警告信号発信部385を備えて構成されている。図示を省略しているが、マイク38bは、警告信号発信部385を備えていない点以外はマイク38aと同様の構成である。無線送受信部39は、送受信ユニット391、マイク判定部392、距離測定部393、通信可能圏内判定部394、警告信号発信部395、アンテナ396を備えて構成されている。

[0171] マイク38a(38b)のマイクユニット381はバイノーラル音声信号を生成する。マイクアンプ382はマイクユニット381からのバイノーラル音声信号を増幅する。送受信ユニット383はマイクアンプ382からの増幅されたバイノーラル音声信号を所定の変調方式により変調してアンテナ384を介して送信する。警告信号発信部385は後述する無線送受信部39の警告信号発信部395で生成され、送受信ユニット391及び送受信ユニット383を介して伝送された警告信号を基に警告信号を発生する。

[0172] 無線送受信部39のアンテナ396は左右のマイク38a, 38bから送信される変調信号を受信する。送受信ユニット391は受信した変調信号を復調してバイノーラル音声信号を生成すると共に、変調信号の受信電力の強度を測定する。距離測定部393は測定された受信電力の強度を基に無線送受信部39から無線型バイノーラルマイク38までの距離を推定する。通信可能圏内判定部394は推定された距離が通信可能な所定の距離以内であるか否かを判定する。通信可能圏内判定部394の判定結果は制御部47に供給される。制御部47は、推定された距離が通信可能な所定の距離以内であればスイッチSw1を端子bに接続させ、フラグ情報生成部42がバイノーラルフラグ信号を生成するよう制御する。制御部47は、推定された距離が通信可能な所定の距離を超える場合にはスイッチSw1を端子aに接続させる。

[0173] 警告信号発信部395は、通信可能圏内判定部394で推定された距離が通信可能な所定の距離を超えると判定した場合に警告信号を生成する。警告信号は制御部47に供給される。制御部47は警告マークを生成して表示部17に供給して、表示部17

に所定の警告マークを表示させる。マイク判定部392は警告信号発信部395で警告信号を生成していない場合は正常にバイノーラル音声信号が得られていると判定し、送受信ユニット391にて復調して得られたバイノーラル音声信号を、スイッチSw1を介してオーディオ符号化部22に供給する。

[0174] 図55に、無線型バイノーラルマイク38及び映像音声記録再生装置107でなされる警告表示の例を示す。マイク38aには発光ダイオード(LED)386を先端部に装着した棒状部材が設けられている。LED386には警告信号発信部385で生成された警告信号が供給され、警告信号に基づいてLED386が点滅(または点灯)する。LED386が点滅に加えて、またはLED386が点滅の代わりに警告音を発生させてもよい。この場合には、警告音がマイクユニット381で受聴されないよう(受聴されにくいよう)、警告音のレベルを小さくしたり、警告音の周波数を例えば数十Hz以下の低周波数としたりすることが好ましい。

[0175] 警告信号発信部395は、無線型バイノーラルマイク38が破線の円で示す通信可能圏を越えたと判定した場合に警告信号を生成する。図55に示すように、無線型バイノーラルマイク38が通信可能圏を越えていれば、表示部17に所定の警告マークが表示される。

[0176] 図56は表示部17に表示される警告マークの例を示している。図56(A)に示す警告マーク171aは、図5(A)に示したバイノーラルマーク171上に×印を重ねて表示したものである。図56(B)に示す警告マーク172aは、図5(B)に示すマーク172をグレー表示したものである。警告マークとしては図56(A),(B)いずれのマークを用いてもよく、勿論これ以外のマークを用いてもよい。警告表示を行ったにもかかわらず無線送受信部39での受信電力が受信限界以下となると想定される場合には、制御部47は無線型バイノーラルマイク38から内蔵ステレオマイク21へと切り替える。

[0177] 図57を用いて映像音声記録再生装置107の動作についてさらに説明する。図57において、ステップS251にて、制御部47はバイノーラルモードであるか否かを判定する。ステップS251でバイノーラルモードであると判定されなければ(NO)、ステップS253にて、記録・再生部14は内蔵ステレオマイク21で集音して通常のステレオ音声信号を記録媒体44に記録する。ステップS251でバイノーラルモードであると判定

されれば(YES)、ステップS252にて無線送受信部39は無線型バイノーラルマイク38からの送信信号を受信する。ステップS254にて、距離測定部393は受信電力の強度(受信強度)に基づいて無線送受信部39から無線型バイノーラルマイク38までの距離を検出する。ステップS255にて、通信可能圏内判定部394は検出した距離が所定の距離以内であるか否かを判定する。

[0178] ステップS255で所定の距離以内であると判定されなければ(NO)、ステップS257にて、制御部47は、警告表示時間tが0(非提示)であるか否か判定し、警告表示時間tが0であればステップS300にて警告信号を発生するよう警告信号発信部395を制御する。警告信号の発生後、ステップS254に戻り同じ処理を繰り返す。ステップS257にて、警告表示時間tが0でなければ(NO)、ステップS258にて、制御部47は、警告表示時間tが予め定めた最大時間tmaxを越えたか否か判定し、最大時間tmaxを越えていなければ(NO)、ステップS300に移り、同様にテップS254に戻る。最大時間tmaxを越えていれば(YES)、ステップS259にて、制御部47は、無線型バイノーラルマイク38から内蔵ステレオマイク21へと切り替えるようスイッチSw1を制御し、警告信号の発生を停止するよう警告信号発信部395を制御する。そして、ステップS253に移る。

[0179] ステップS255で所定の距離以内であると判定されれば(YES)、ステップS256にて、制御部47は、警告信号発信部395が警告信号を発生中であれば、警告信号の発生を停止するよう警告信号発信部395を制御する。そして、ステップS301にて、記録・再生部14は無線型バイノーラルマイク38で集音してバイノーラル音声信号を記録媒体44に記録する。ステップS302にて、制御部47は、記録終了の操作がなされたか否か判定し、記録終了の操作がなされていなければ(NO)、ステップS251に戻り、記録終了の操作がなされていれば(YES)、処理を終了する。

産業上の利用可能性

[0180] 本発明は、民生用のビデオカメラのみならず、業務用のビデオカメラにおいても、撮像した映像と共に臨場感のある音声を再生することが必要な映像音声記録再生装置に適用できる。また、動画撮影機能を有するデジタルカメラや携帯電話にも適用できる。さらには、本発明は映像信号と音声信号との双方を記録再生する映像音声記録

再生装置に用いて好適であるが、音声信号のみを記録再生する音声記録再生装置に用いても十分に効果的である。

請求の範囲

[1] 被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者(300)の周囲音を集音した音声信号とを記録する映像音声記録装置(101, 102, 103, 104, 105, 107)において、

前記被写体を撮像するカメラ部(11)と、

前記周囲音を集音するマイクとして、前記撮影者(300)の耳に装着するバイノーラルマイク(3)を使用するかバイノーラルマイク(3)以外のマイクを使用するかを切り替える切り替え部(Sw1)と、

前記カメラ部(11)より出力された映像信号を処理するビデオ処理部(12)と、

前記周囲音を集音するマイクより出力された音声信号を処理するオーディオ処理部(22)と、

前記切り替え部(Sw1)によって前記周囲音を集音するマイクとして前記バイノーラルマイク(3)に切り替えた場合に、前記周囲音を集音する際の音声モードをバイノーラルモードとして、バイノーラルモードであることを示すバイノーラルフラグ信号を生成するフラグ情報生成部(42)と、

前記ビデオ処理部(12)によって処理された映像信号と前記オーディオ処理部(22)によって処理された音声信号と前記バイノーラルフラグ信号とを記録媒体に記録する記録部(14)と、

を備えて構成したことを特徴とする映像音声記録装置(101, 102, 103, 104, 105, 107)。

[2] 前記映像音声記録装置(101, 102, 103, 104, 105, 107)に内蔵されている内蔵マイク(21)と、

外部マイク接続端子(32)と、

前記外部マイク接続端子(32)に接続して使用する外部マイクとして、前記バイノーラルマイク(3)かバイノーラルマイク以外のマイクかを設定する設定部(48, 47)と、

前記外部マイク接続端子(32)に外部マイクが接続されているか否かを検出する接続検出部(41)と、

前記オーディオ処理部(22)に供給する音声信号として、前記内蔵マイク(21)より

出力された音声信号と前記外部マイクより出力された音声信号とを切り替えるスイッチ(Sw1)と、

前記設定部(47, 48)によって前記外部マイクがバイノーラルマイク(3)に設定され、前記接続検出部(41)によって前記外部マイク接続端子(32)に外部マイクが接続されていることが検出されたとき、音声モードを前記バイノーラルモードとする制御部(47)とを備え、

前記制御部(47)は、音声モードが前記バイノーラルモードのとき、前記スイッチ(Sw1)が前記外部マイクより出力された音声信号を前記オーディオ処理部(22)に供給するよう前記スイッチ(Sw1)を切り替え制御し、前記フラグ情報生成部(42)が前記バイノーラルフラグ信号を生成するよう前記フラグ情報生成部(42)を制御することを特徴とする請求の範囲第1項に記載の映像音声記録装置(101, 102, 103, 104, 105, 107)。

[3] 前記カメラ部(11)より出力された映像信号を表示する表示部(17)を備え、

音声モードが前記バイノーラルモードのとき、前記表示部(17)に前記バイノーラルモードであることを示すバイノーラルマークを表示させる表示制御部(47)を備えたことを特徴とする請求の範囲第1項に記載の映像音声記録装置(101, 102, 103, 104, 105, 107)。

[4] 前記カメラ部(11)は前記被写体を拡大して撮像するズーム機能を有し、

前記バイノーラルマイク(3)によって集音された音声信号を前記カメラ部(11)における拡大率に応じて増幅する音声ズーム処理部(33)を備えたことを特徴とする請求の範囲第1項に記載の映像音声記録装置(102)。

[5] 前記カメラ部(101)は前記被写体を拡大して撮像するズーム機能を有し、

前記バイノーラルマイク(3)が集音する音声信号の音源を仮想的に受聴者に接近させて仮想音源を形成するための頭部伝達関数を前記仮想音源と前記受聴者との複数の距離に対応させて記憶する伝達関数メモリ(334)と、この伝達関数メモリ(334)に記憶された複数の頭部伝達関数のいずれかを前記カメラ部(11)における拡大率に応じて選択的に取得する関数取得器(333)と、前記バイノーラルマイク(3)によって集音された音声信号に対して前記関数取得器(333)にて取得された頭部伝達

関数を畳み込む畳み込み演算器(336)とを有する音声ズーム処理部(33a, 33c)を備えたことを特徴とする請求の範囲第1項に記載の映像音声記録装置(102, 103)。

[6] 被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者(300)の周囲音を集音した音声信号とを記録する映像音声記録方法において、

前記被写体を撮像する撮像ステップ(S155)と、

前記周囲音を集音するマイクとして、前記撮影者(300)の耳に装着するバイノーラルマイク(3)を使用するかバイノーラルマイク以外のマイクを使用するかを切り替える切り替えステップ(S151)と、

前記被写体を撮像した映像信号を処理するビデオ処理ステップ(S156)と、

前記周囲音を集音するマイクより出力された音声信号を処理するオーディオ処理ステップ(S156)と、

前記切り替えステップ(S151)にて前記周囲音を集音するマイクとして前記バイノーラルマイク(3)に切り替えた場合に、前記周囲音を集音する際の音声モードをバイノーラルモードとして、バイノーラルモードであることを示すバイノーラルフラグ信号を生成するフラグ情報生成ステップ(S158)と、

前記ビデオ処理ステップ(S156)にて処理された映像信号と前記オーディオ処理ステップ(S156)にて処理された音声信号と前記バイノーラルフラグ信号とを記録媒体(44)に記録する記録ステップ(S161)と、

を有することを特徴とする映像音声記録方法。

[7] 被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者(300)の周囲音を集音した音声信号とが記録された記録媒体(44)を再生する映像音声再生装置(101, 102, 103, 104, 105, 106, 107)において、

前記記録媒体に記録された記録信号を再生する再生部(14)と、

前記再生部(14)によって再生された前記記録信号から前記映像信号と前記音声信号とを分離する分離化部(15)と、

前記分離化部(15)によって分離された前記映像信号を処理するビデオ処理部(16)と、

前記分離化部(15)によって分離された前記音声信号を処理するオーディオ処理部(26)と、

前記記録媒体(44)に、前記周囲音を集音するマイクとして前記撮影者(300)の耳に装着するバイノーラルマイクを使用したことを示すバイノーラルフラグ信号が記録されているとき、前記バイノーラルフラグ信号を取得するフラグ情報取得部(36)と、

前記フラグ情報取得部(36)が前記バイノーラルフラグ信号を取得したとき、前記オーディオ処理部(26)によって処理された前記音声信号をスピーカ(53, 54)によって発音した場合に発生するクロストーク信号を打ち消すよう前記音声信号を処理するクロストークキャンセラ(27)とを備え、

前記クロストークキャンセラ(27)は前記音声信号に対して予め求めたフィルタ特性を畳み込むフィルタ(272a－272d)を有し、前記フィルタ特性は、測定信号を円筒型構造体に装着した一対のマイクで集音した音声信号を用いて計測した頭部伝達関数に基づいたフィルタ特性であることを特徴とする映像音声再生装置(101, 102, 103, 104, 105, 106, 107)。

[8] 被写体を撮像した映像信号と、前記被写体から発せられた音を含む撮影者(300)の周囲音を集音した音声信号とが記録された記録媒体(44)を再生する映像音声再生方法において、

前記記録媒体(44)に記録された記録信号を再生する再生ステップ(S181)と、

前記再生ステップにて再生された前記記録信号から前記映像信号と前記音声信号とを分離する分離化ステップ(S183)と、

前記分離化ステップ(S183)にて分離された前記映像信号を処理するビデオ処理ステップ(S184)と、

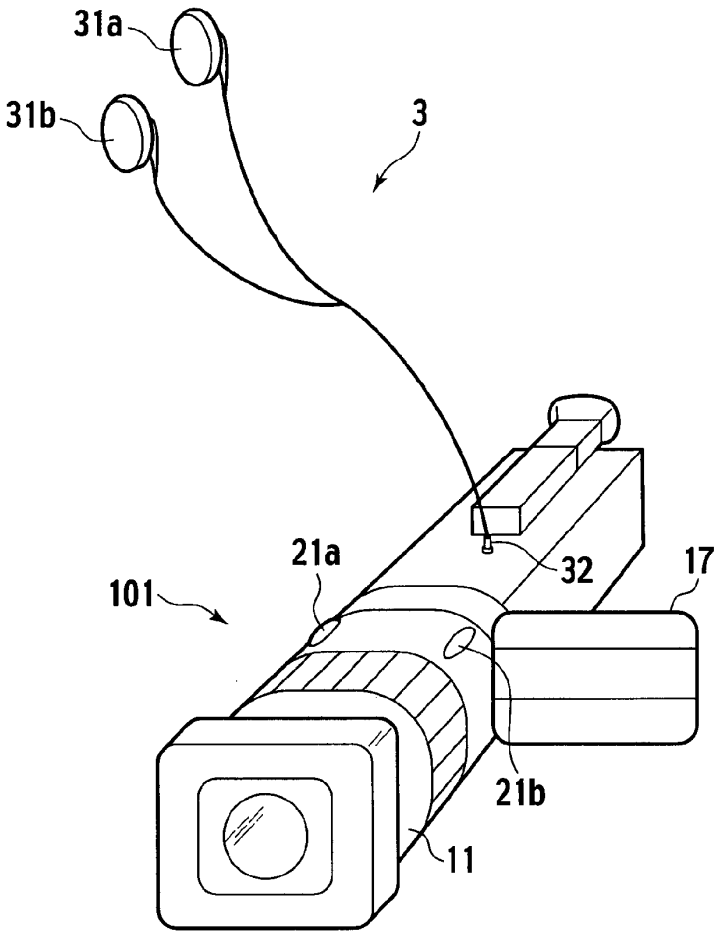
前記分離化ステップ(S183)によって分離された前記音声信号を処理するオーディオ処理ステップ(S184)と、

前記記録媒体(44)に、前記周囲音を集音するマイクとして前記撮影者(300)の耳に装着するバイノーラルマイク(3)を使用したことを示すバイノーラルフラグ信号が記録されているとき、前記バイノーラルフラグ信号を取得するフラグ情報取得ステップ(S186)と、

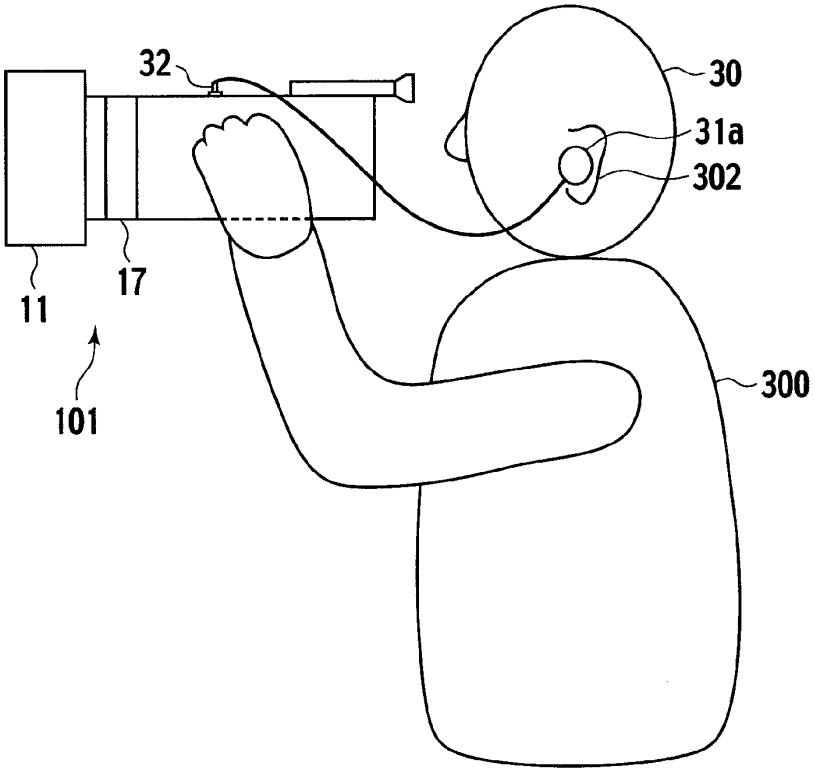
前記フラグ情報取得ステップ（S186）にて前記バイノーラルフラグ信号を取得したとき、前記オーディオ処理ステップ（S184）にて処理された前記音声信号をスピーカ（53，54）によって発音した場合に発生するクロストーク信号を打ち消すよう前記音声信号を処理するクロストークキャンセルステップ（S188）とを有し、

前記クロストークキャンセルステップ（S188）は前記音声信号に対して予め求めたフィルタ特性を畳み込むステップであり、前記フィルタ特性は、測定信号を円筒型構造体に装着した一対のマイクで集音した音声信号を用いて計測した頭部伝達関数に基づいたフィルタ特性であることを特徴とする映像音声再生方法。

[図1]



[図2]



[図3]

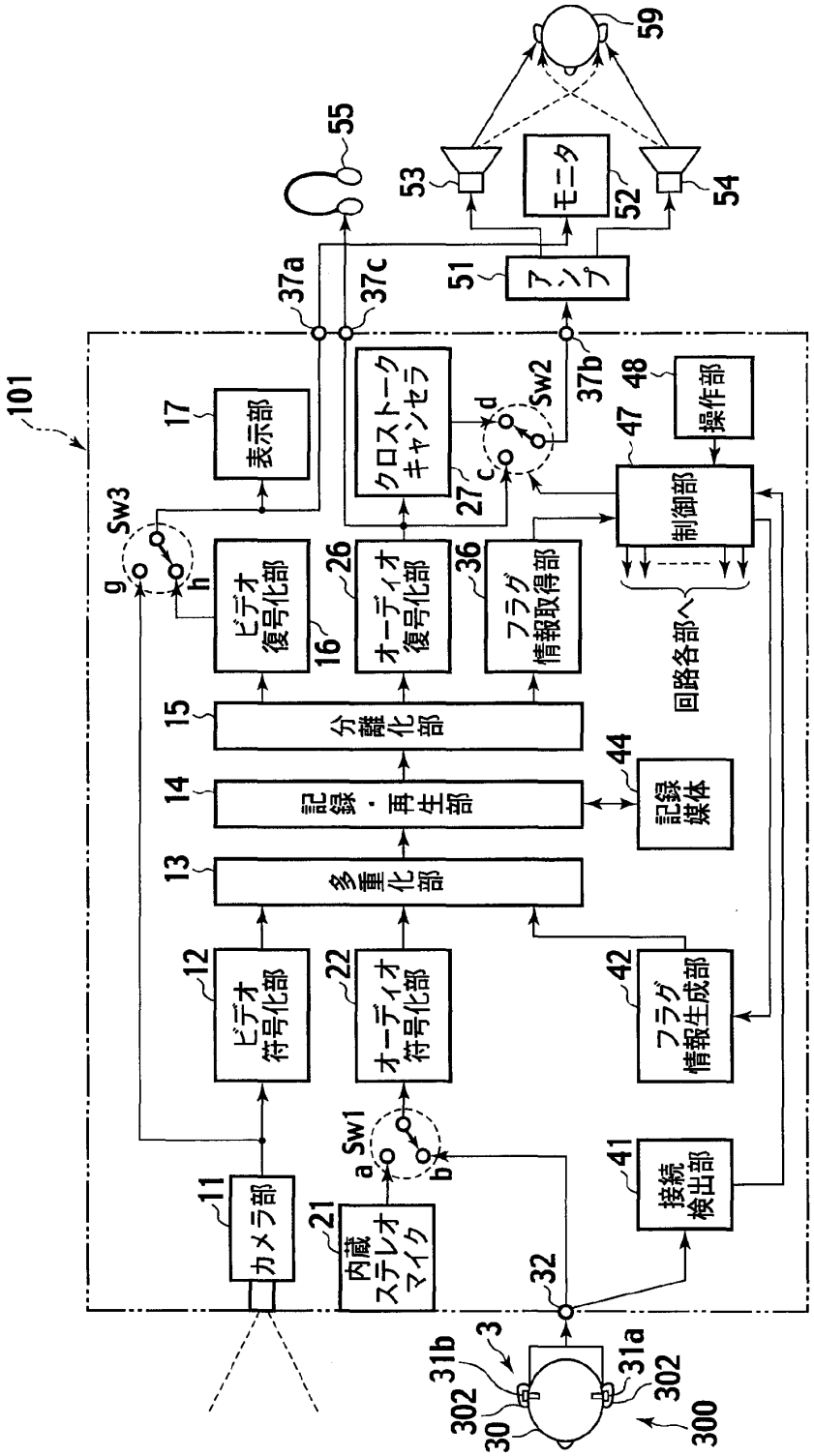
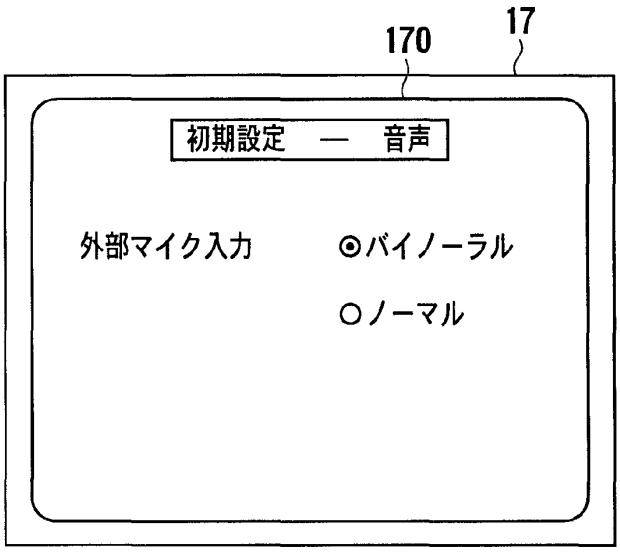
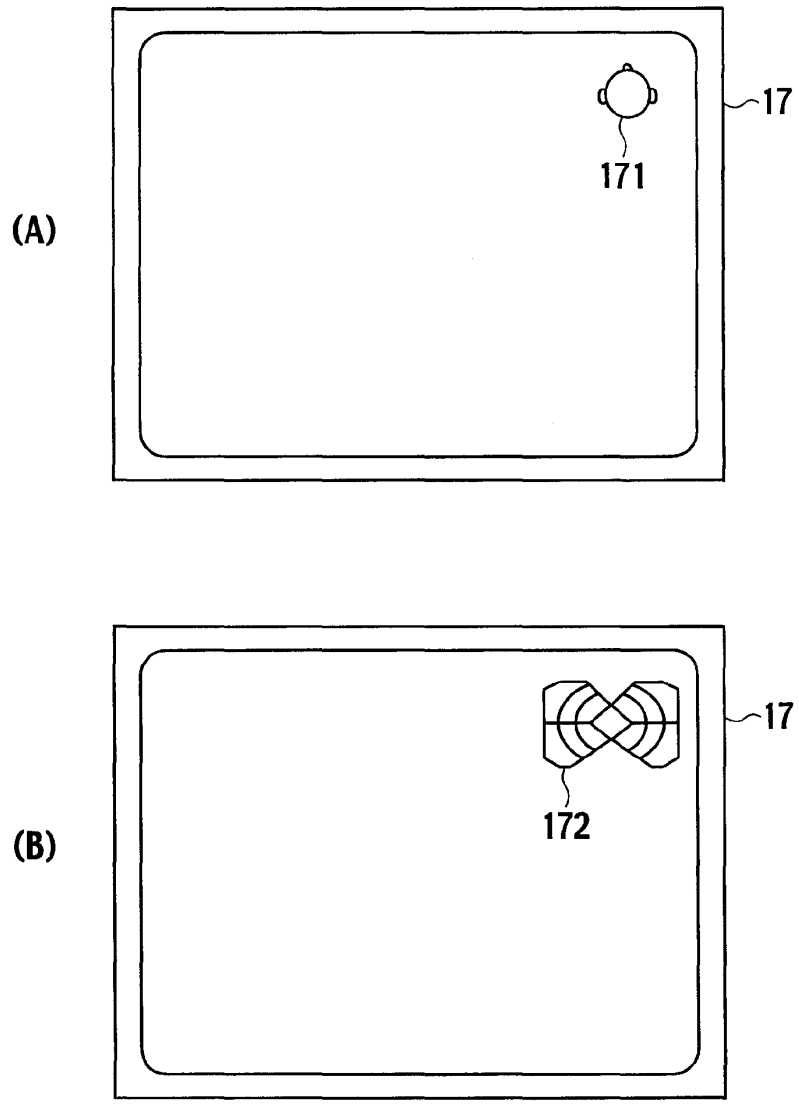


Exhibit 1004
Page 0538 of 1001

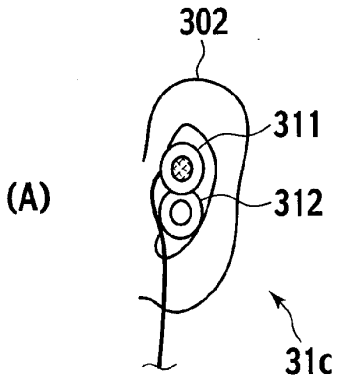
[図4]

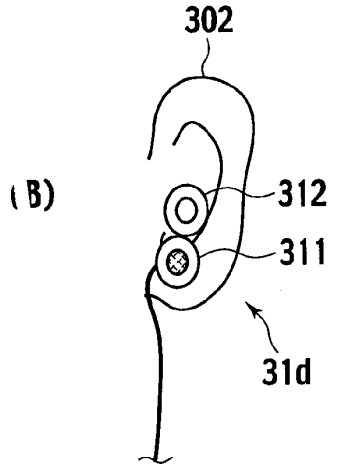


[図5]

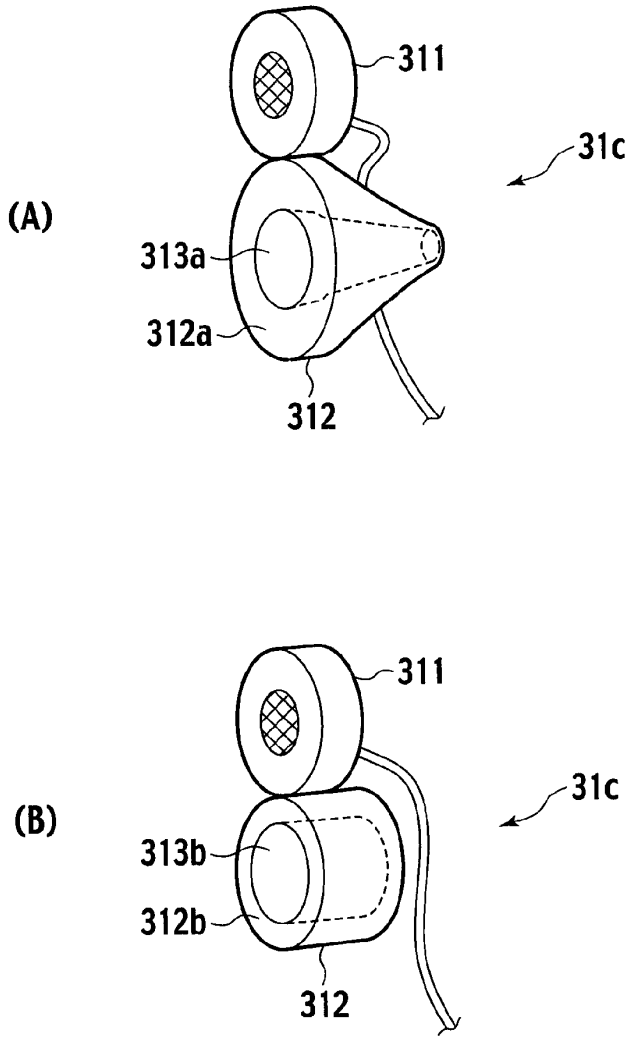


[図6]

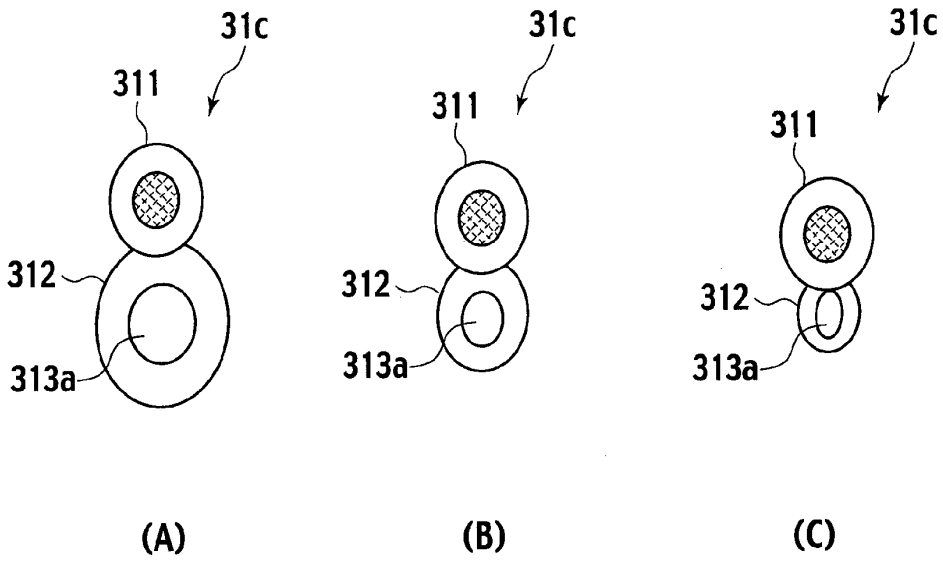




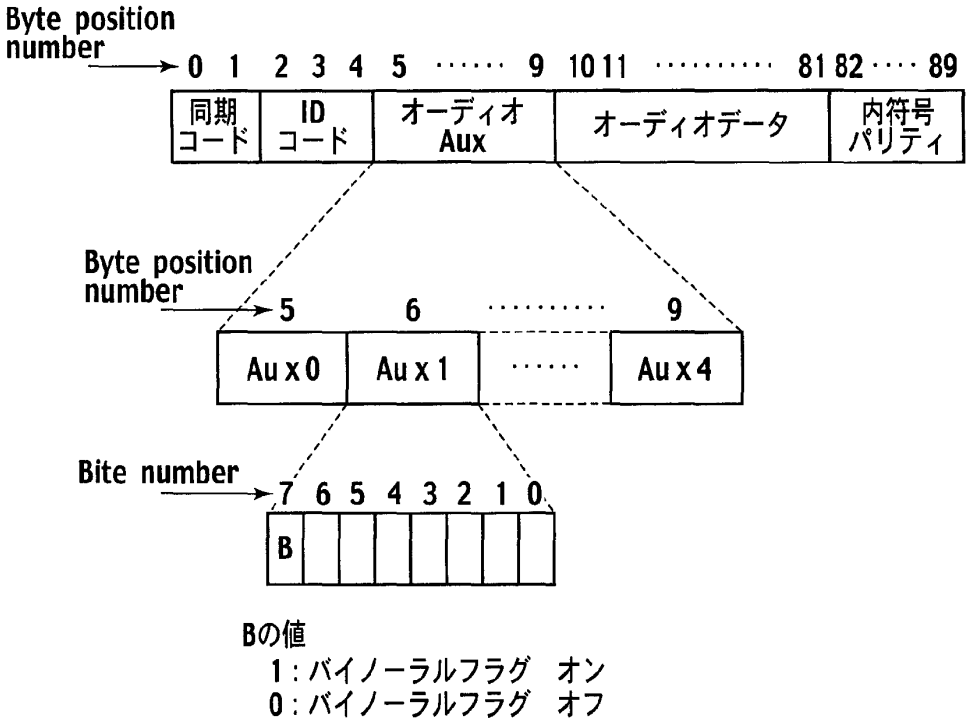
[図7]



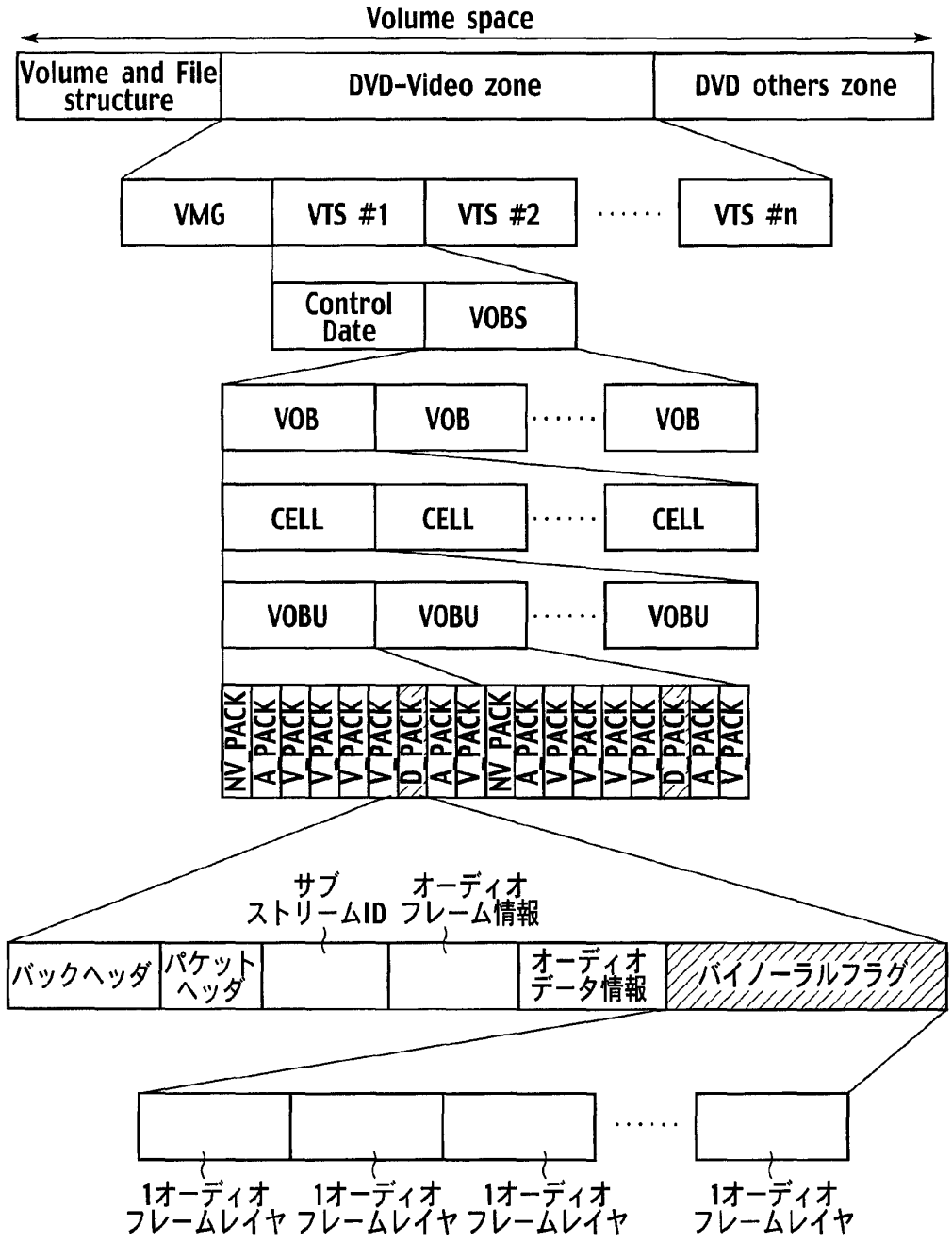
[図8]



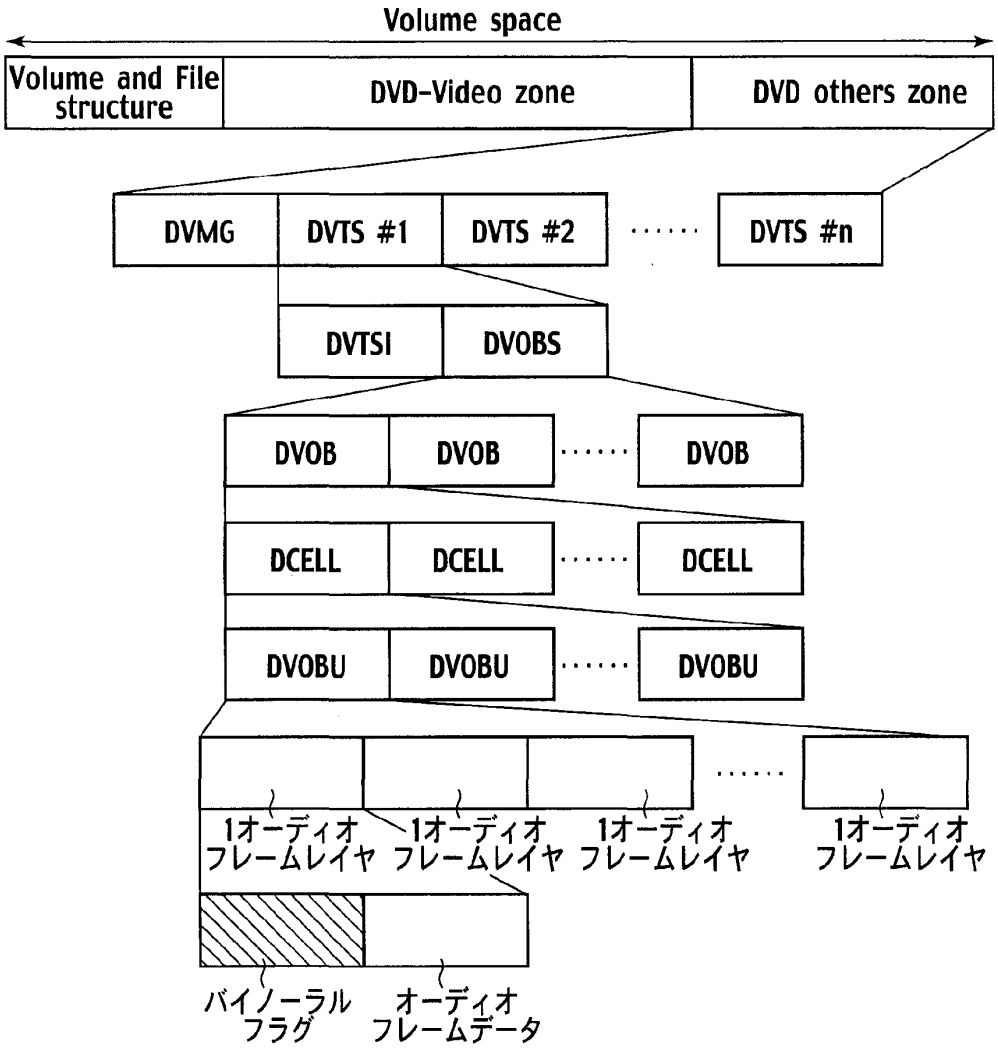
[図9]


Byte position number
0 1 2 3 4 5 ····· 9 10 11 ····· 81 82 ···· 89
同期コード
IDコード
オーディオAux
オーディオデータ
内符号パリティ
Byte position number
5 6 ····· 9
Au x 0
Au x 1
Au x 4
Bite number
7 6 5 4 3 2 1 0
B
Bの値
1：バイノーラルフラグ オン
0：バイノーラルフラグ オフ

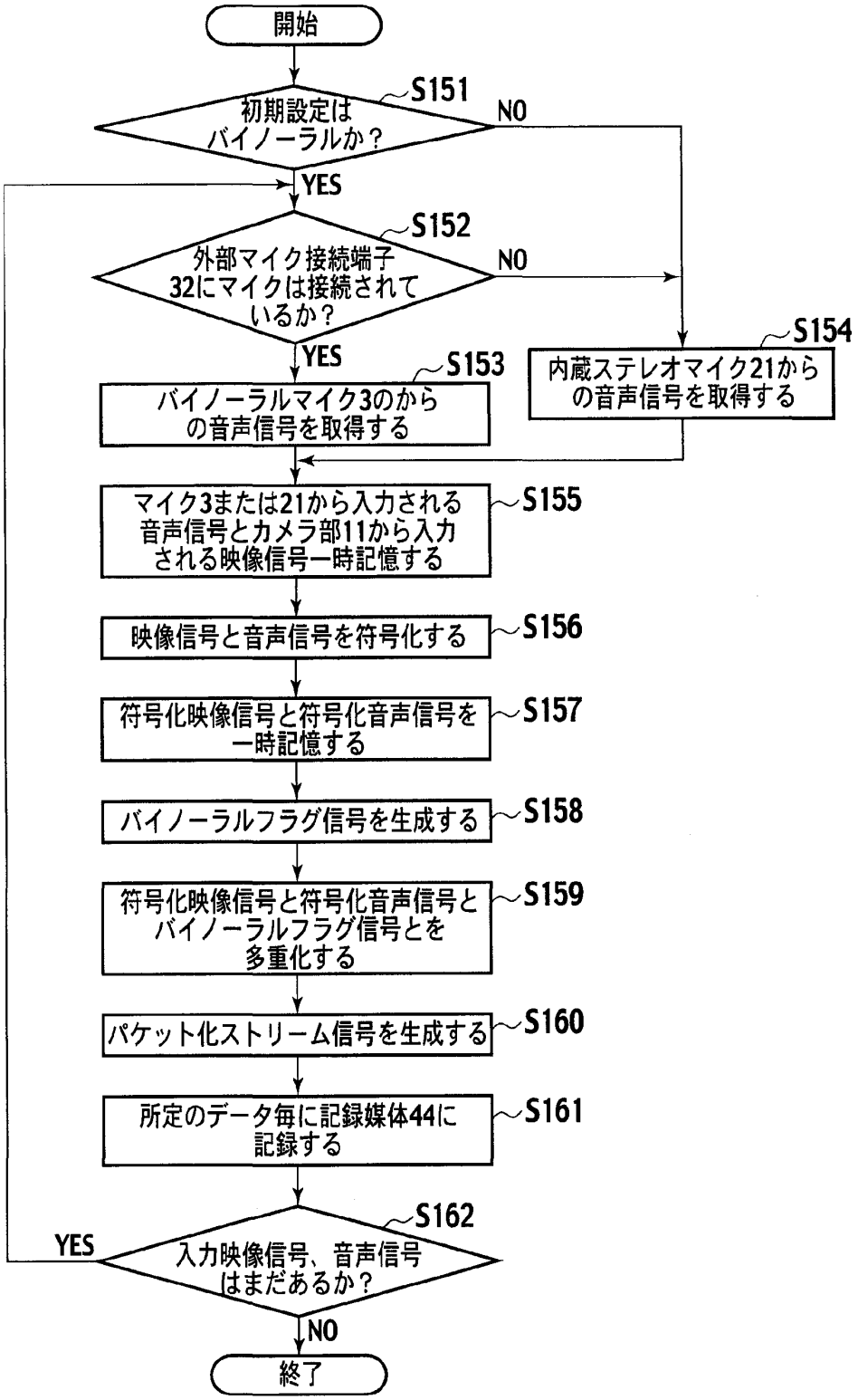
[図10]


Volume space
Volume and File structure
DVD-Video zone
DVD others zone
VMG
VTS #1
VTS #2
VTS #n
Control Date
VOBS
VOB
CELL
VOBU
NV PACK
A PACK
V PACK
D PACK
サブストリームID
オーディオフレーム情報
バックヘッダ
パケットヘッダ
オーディオデータ情報
バイノーラルフラグ
1オーディオフレームレイヤ

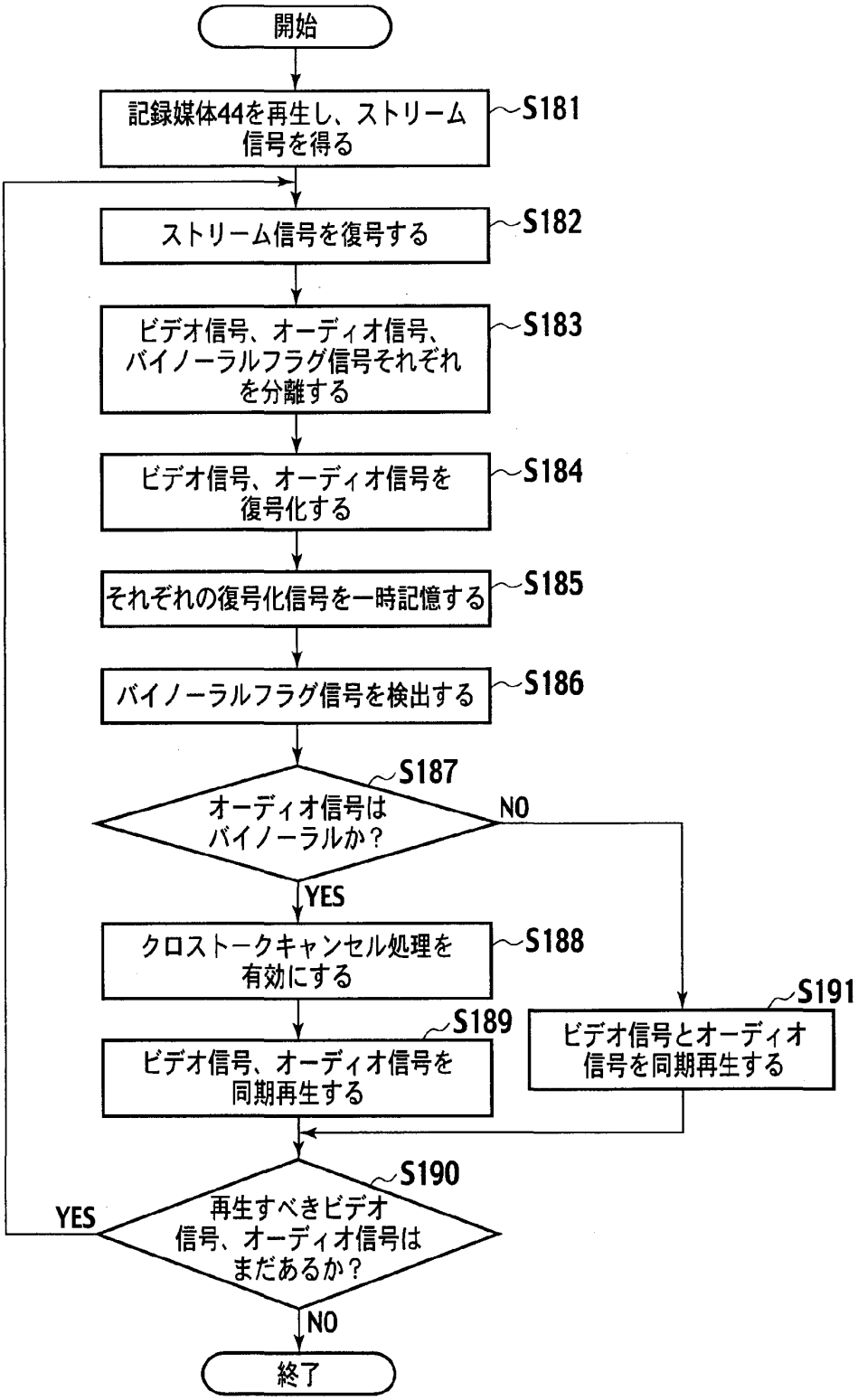
[図11]


Volume space
Volume and File structure
DVD-Video zone
DVD others zone
DVMG
DVTS #1
DVTS #2
DVTS #n
DVTSI
DVOBS
DVOB
DVOB
DVOB
DCELL
DCELL
DCELL
DVOBU
DVOBU
DVOBU
1オーディオフレームレイヤ
1オーディオフレームレイヤ
1オーディオフレームレイヤ
1オーディオフレームレイヤ
バイノーラルフラグ
オーディオフレームデータ

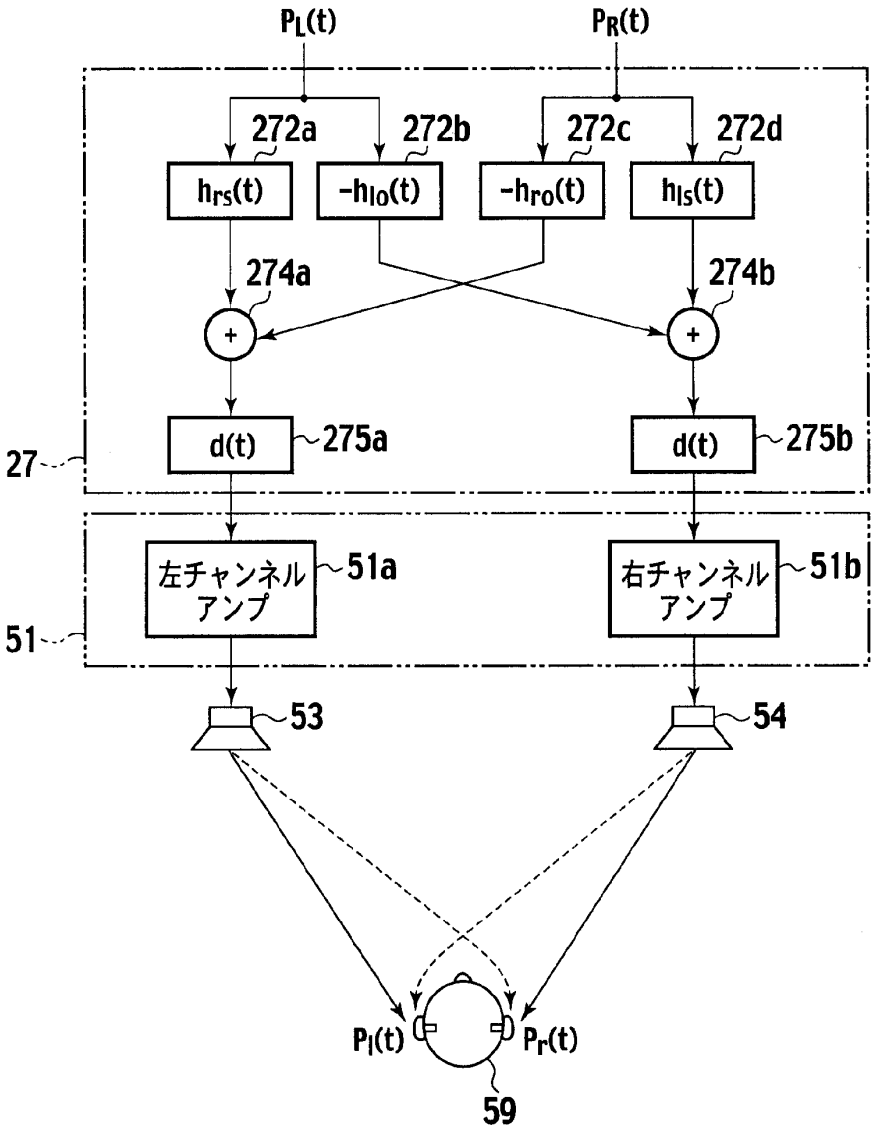
[図12]


開始
S151
初期設定は
バイノーラルか？
NO
YES
S152
外部マイク接続端子
32にマイクは接続されて
いるか？
NO
YES
S153
バイノーラルマイク3のから
の音声信号を取得する
S154
内蔵ステレオマイク21から
の音声信号を取得する
マイク3または21から入力される
音声信号とカメラ部11から入力
される映像信号一時記憶する
S155
映像信号と音声信号を符号化する
S156
符号化映像信号と符号化音声信号を
一時記憶する
S157
バイノーラルフラグ信号を生成する
S158
符号化映像信号と符号化音声信号と
バイノーラルフラグ信号とを
多重化する
S159
パケット化ストリーム信号を生成する
S160
所定のデータ毎に記録媒体44に
記録する
S161
S162
YES
入力映像信号、音声信号
はまだあるか？
NO
終了

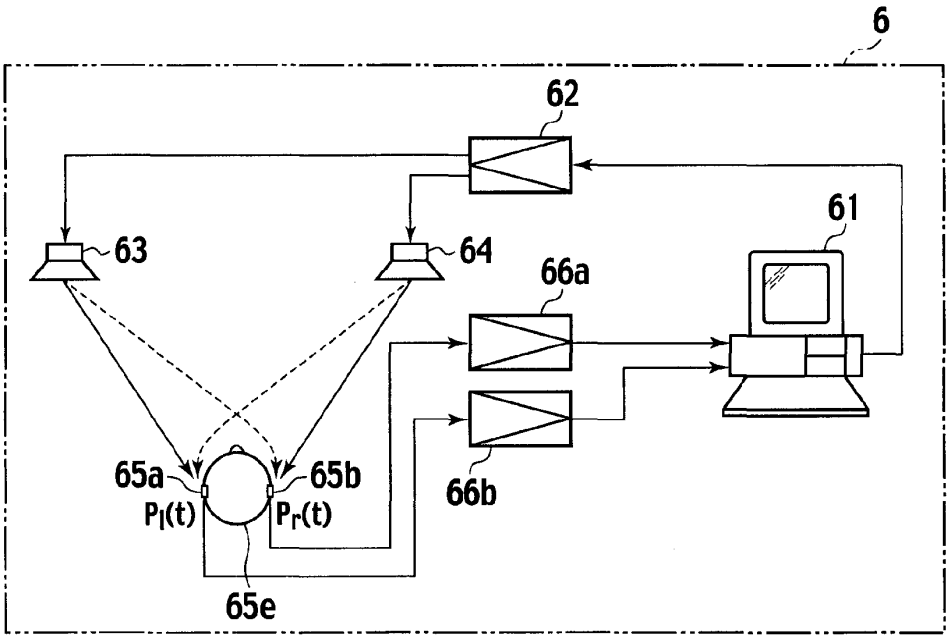
[図13]


開始
記録媒体44を再生し、ストリーム信号を得る
S181
ストリーム信号を復号する
S182
ビデオ信号、オーディオ信号、バイノーラルフラグ信号それぞれを分離する
S183
ビデオ信号、オーディオ信号を復号化する
S184
それぞれの復号化信号を一時記憶する
S185
バイノーラルフラグ信号を検出する
S186
S187
オーディオ信号はバイノーラルか？
NO
YES
クロストークキャンセル処理を有効にする
S188
S189
ビデオ信号、オーディオ信号を同期再生する
S191
ビデオ信号とオーディオ信号を同期再生する
S190
再生すべきビデオ信号、オーディオ信号はまだあるか？
YES
NO
終了

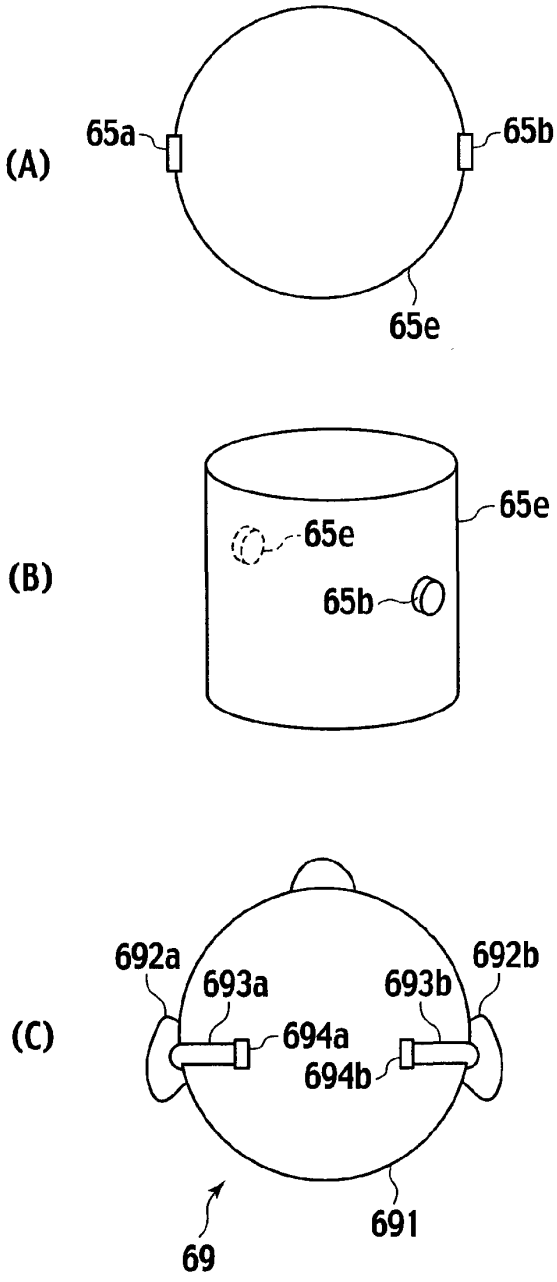
[図14]


PL(t)
PR(t)
272a
272b
272c
272d
hrs(t)
-hlo(t)
-hro(t)
hls(t)
274a
274b
+
+
d(t)
275a
d(t)
275b
27
左チャンネル
アンプ
51a
右チャンネル
アンプ
51b
51
53
54
Pl(t)
Pr(t)
59

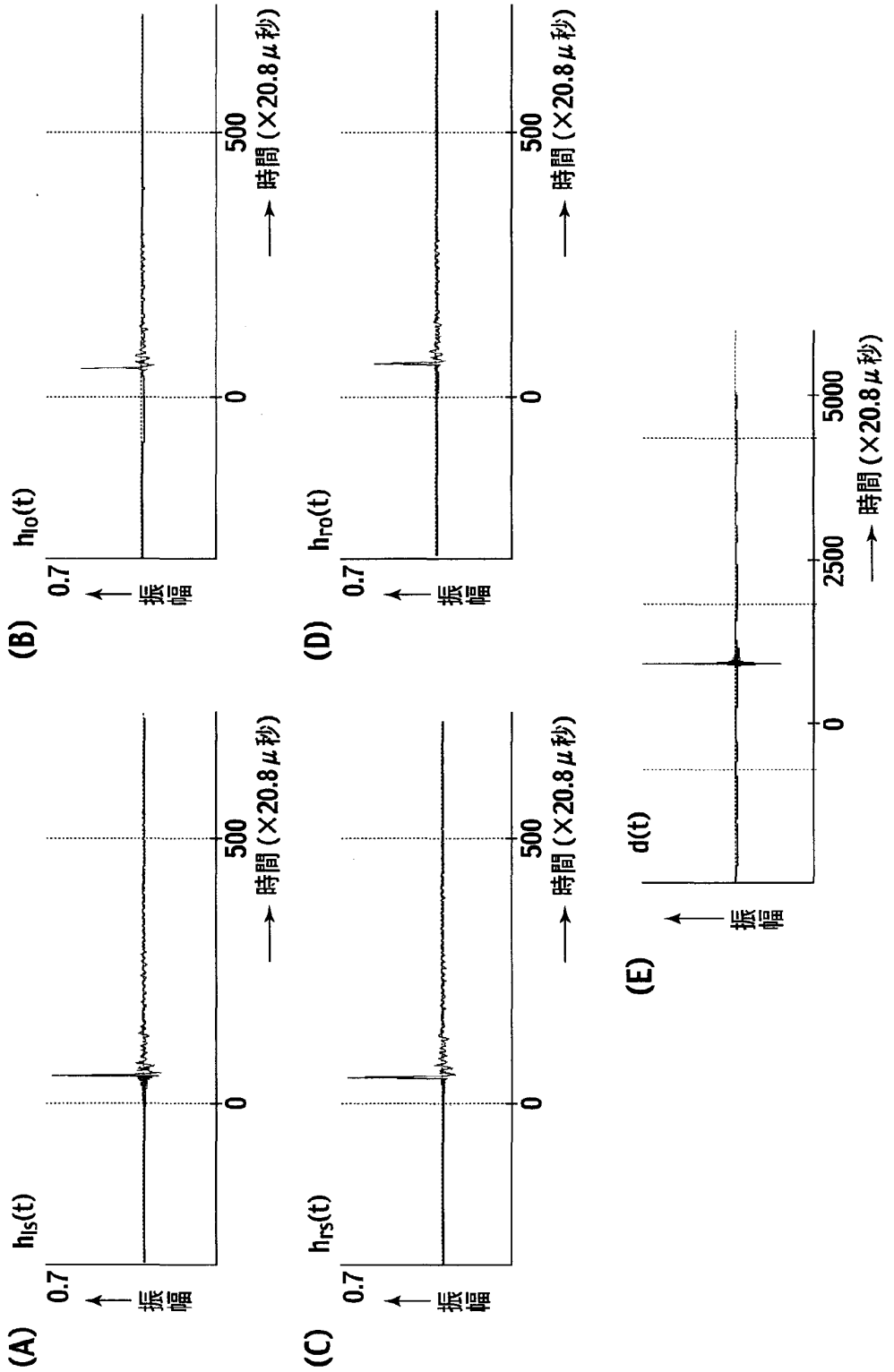
[図15]



[図16]


(A)
65a
65b
65e
(B)
65e
65e
65b
(C)
692a
693a
694a
693b
692b
694b
691
69

[図17]


(A)
$h_{IS}(t)$
0.7
振幅
0
500
時間 (×20.8μ秒)
(B)
$h_{IO}(t)$
0.7
振幅
0
500
時間 (×20.8μ秒)
(C)
$h_{rS}(t)$
0.7
振幅
0
500
時間 (×20.8μ秒)
(D)
$h_{rO}(t)$
0.7
振幅
0
500
時間 (×20.8μ秒)
(E)
d(t)
振幅
0
2500
5000
時間 (×20.8μ秒)

[図18]

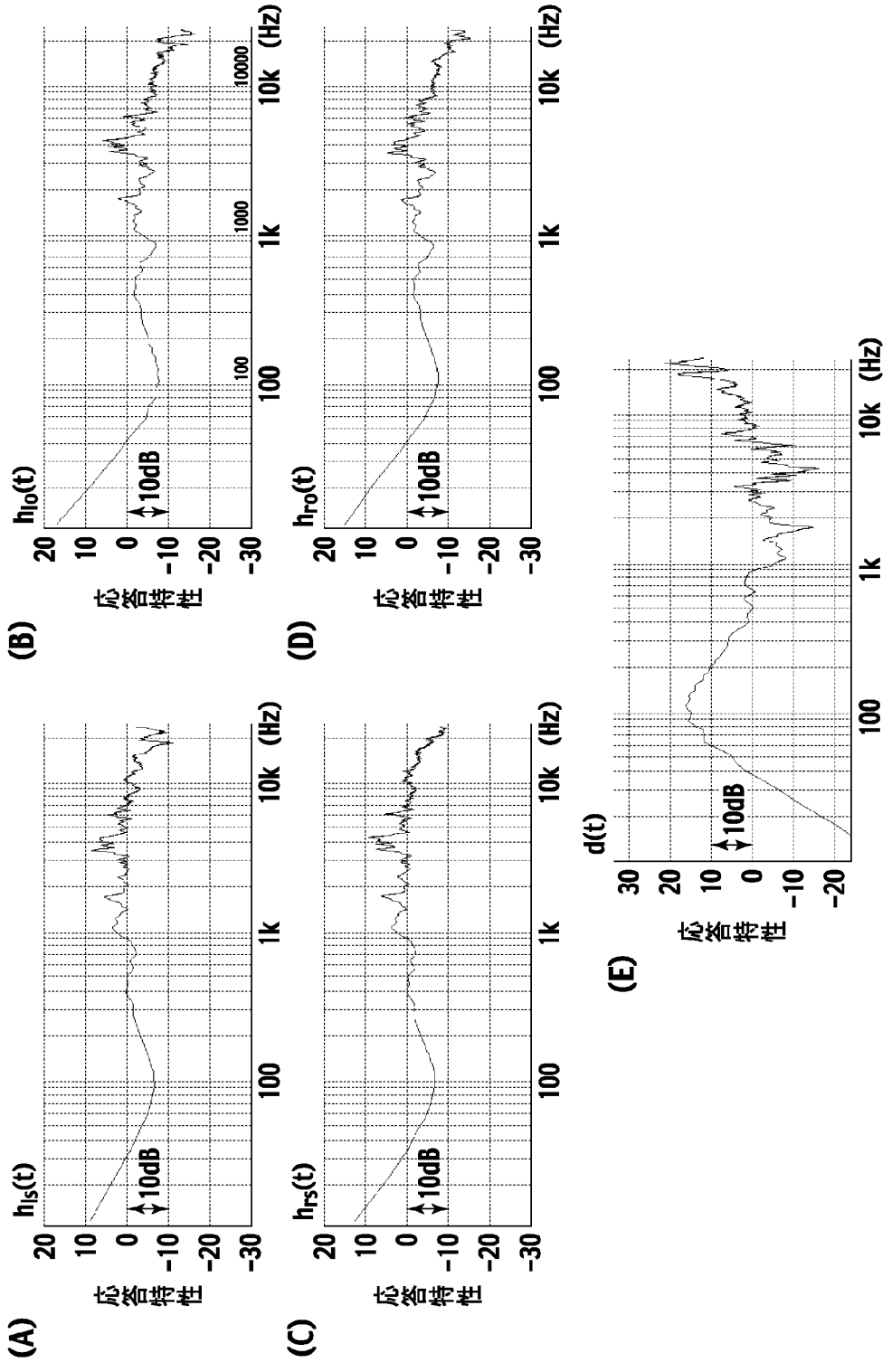
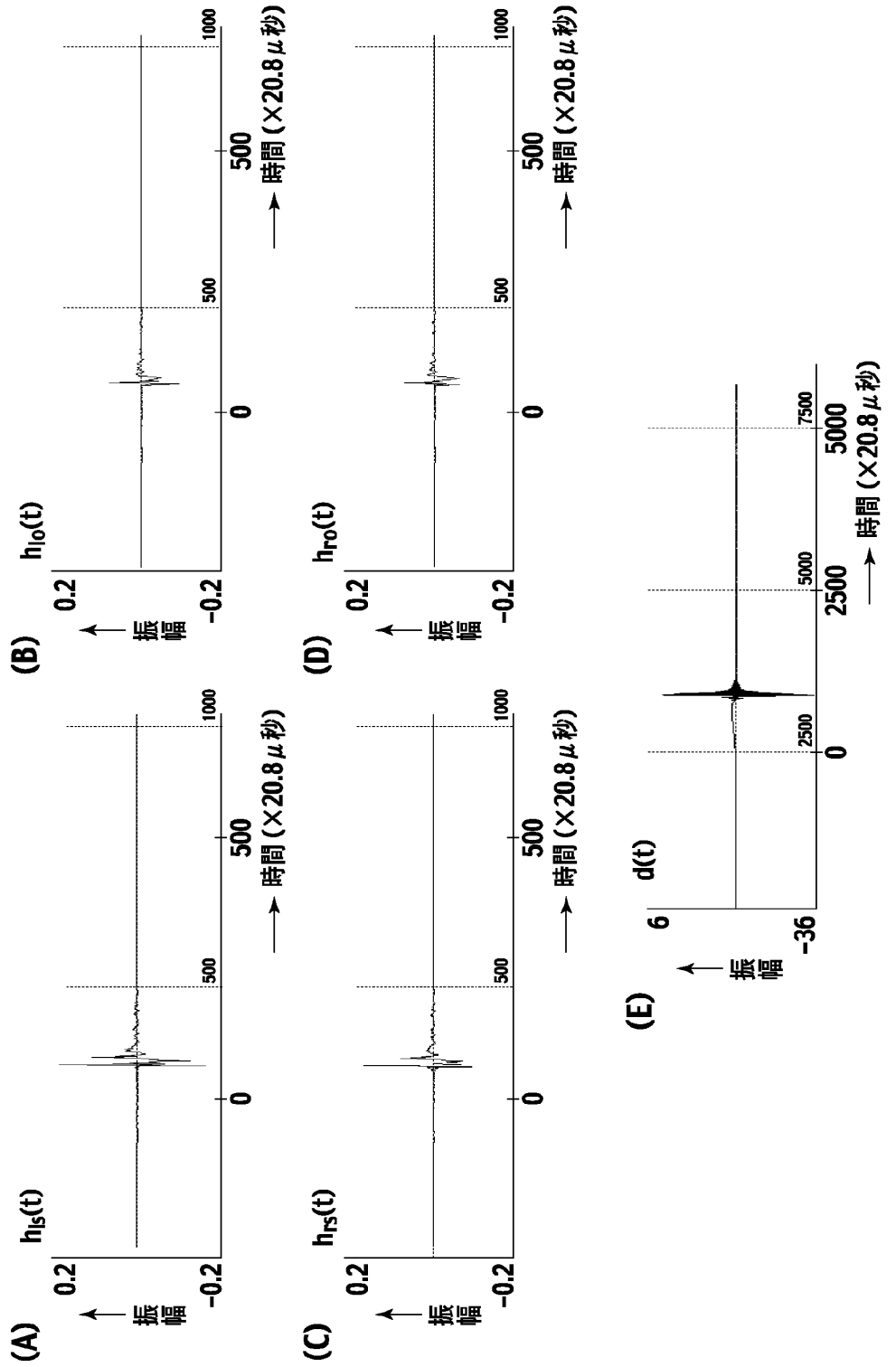


Exhibit 1004
Page 0553 of 1001

[図19]


(A)
$h_{ls}(t)$
0.2
-0.2
振幅
0
500
1000
時間 (×20.8μ秒)
(B)
$h_{lo}(t)$
0.2
-0.2
振幅
0
500
1000
時間 (×20.8μ秒)
(C)
$h_{rs}(t)$
0.2
-0.2
振幅
0
500
1000
時間 (×20.8μ秒)
(D)
$h_{ro}(t)$
0.2
-0.2
振幅
0
500
1000
時間 (×20.8μ秒)
(E)
d(t)
6
-36
振幅
0
2500
5000
7500
時間 (×20.8μ秒)

[図20]

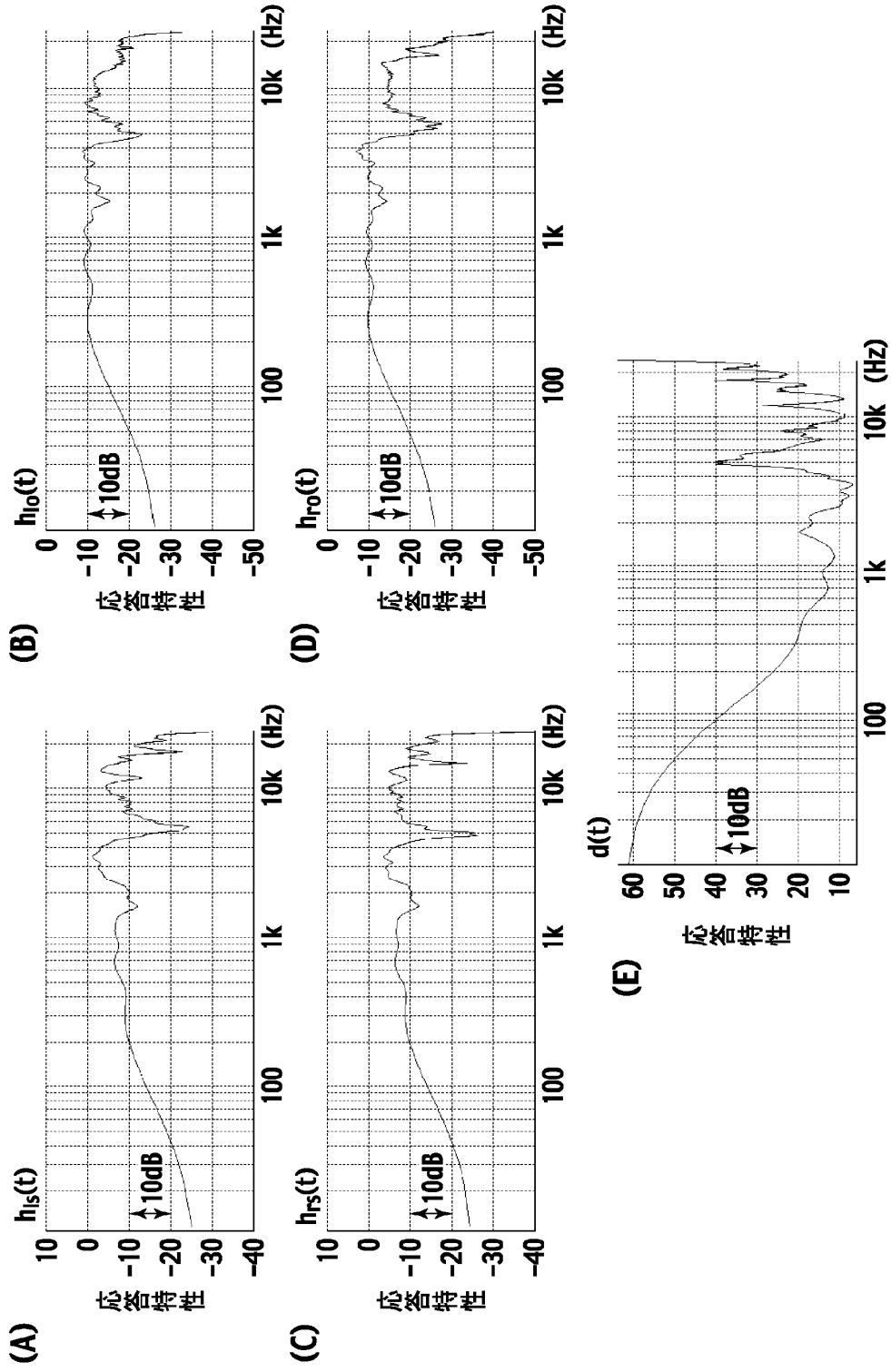
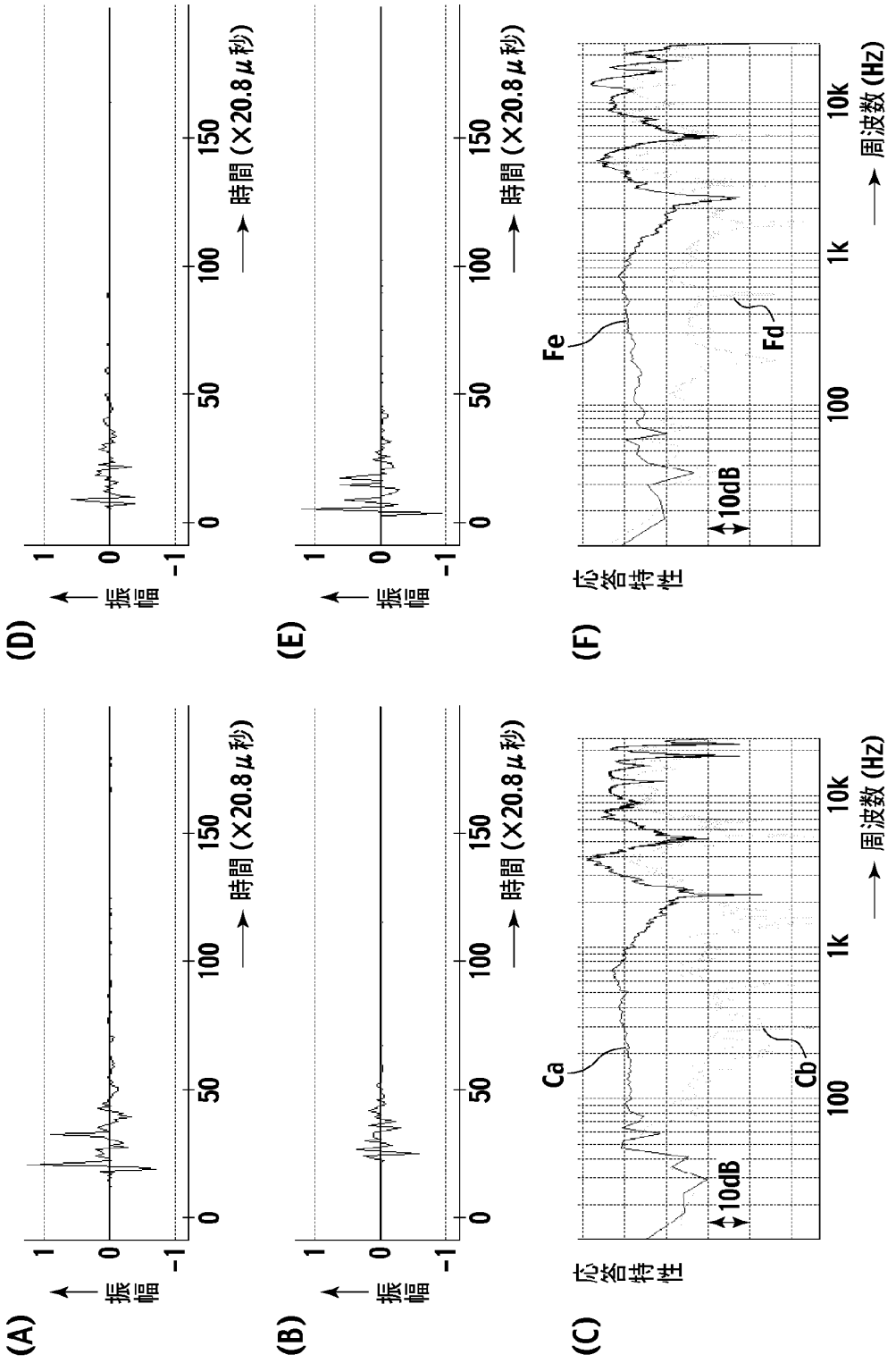


Exhibit 1004
Page 0555 of 1001

[図21]


(A)
(B)
(C)
(D)
(E)
(F)
振幅
時間 (×20.8μ秒)
周波数 (Hz)
応答特性
Ca
Cb
Fd
Fe
10dB
1
0
-1
50
100
150
1k
10k

[図22]

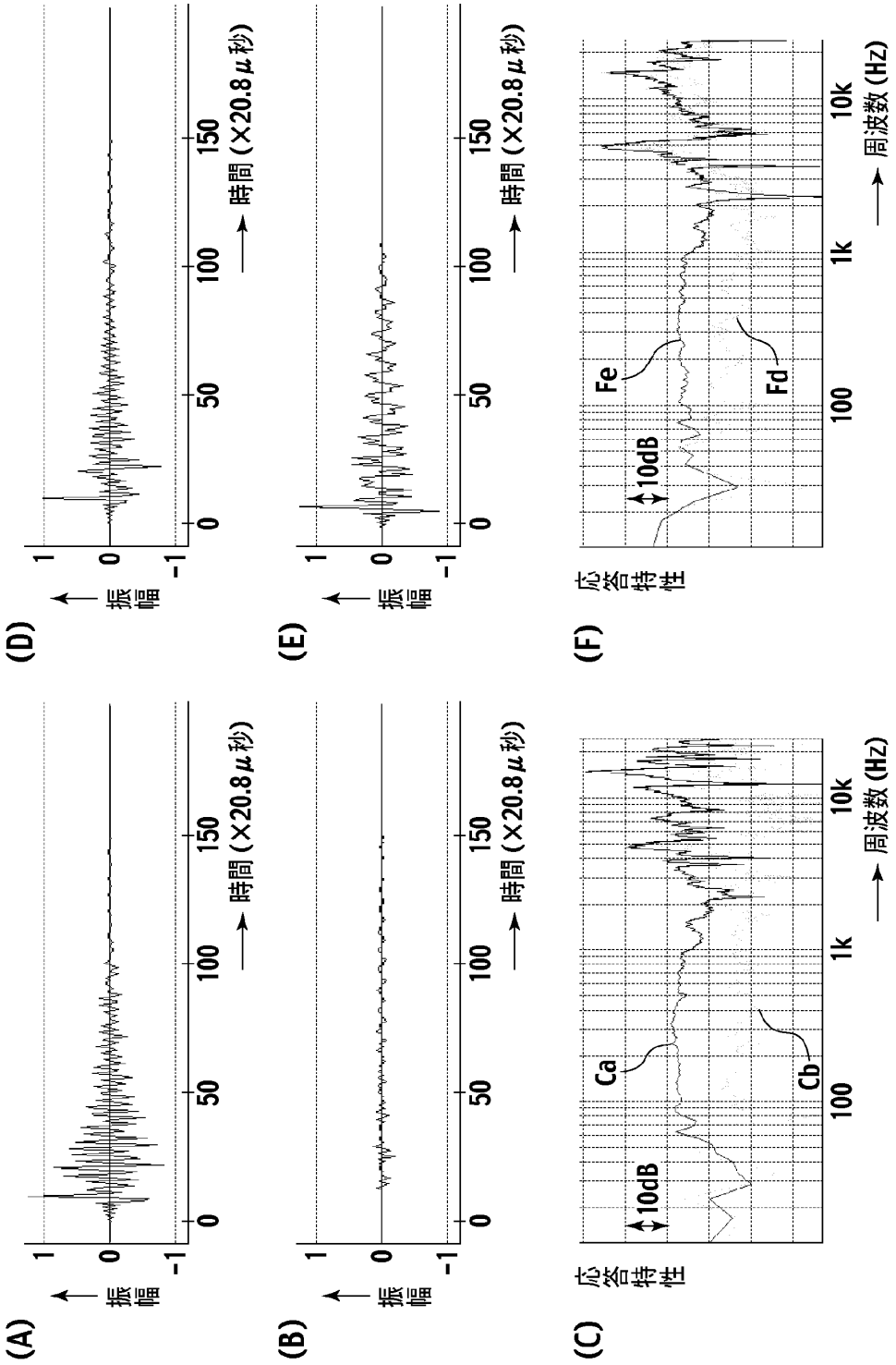
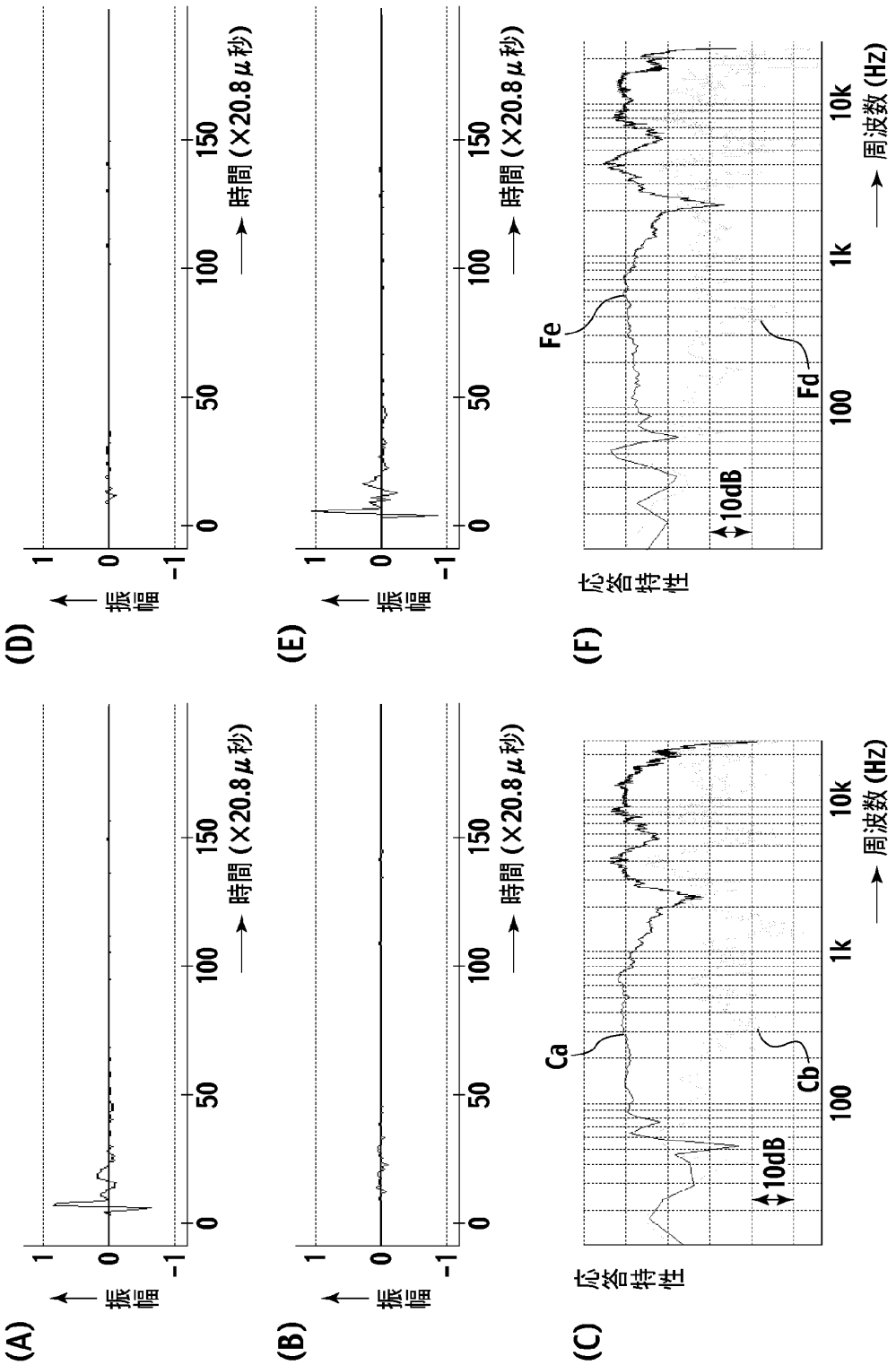


Exhibit 1004
Page 0557 of 1001

[図23]


(A)
(B)
(C)
(D)
(E)
(F)
振幅
時間 (×20.8μ秒)
周波数 (Hz)
応答特性
Ca
Cb
Fd
Fe
10dB
1
0
-1
50
100
150
1k
10k

[図24]

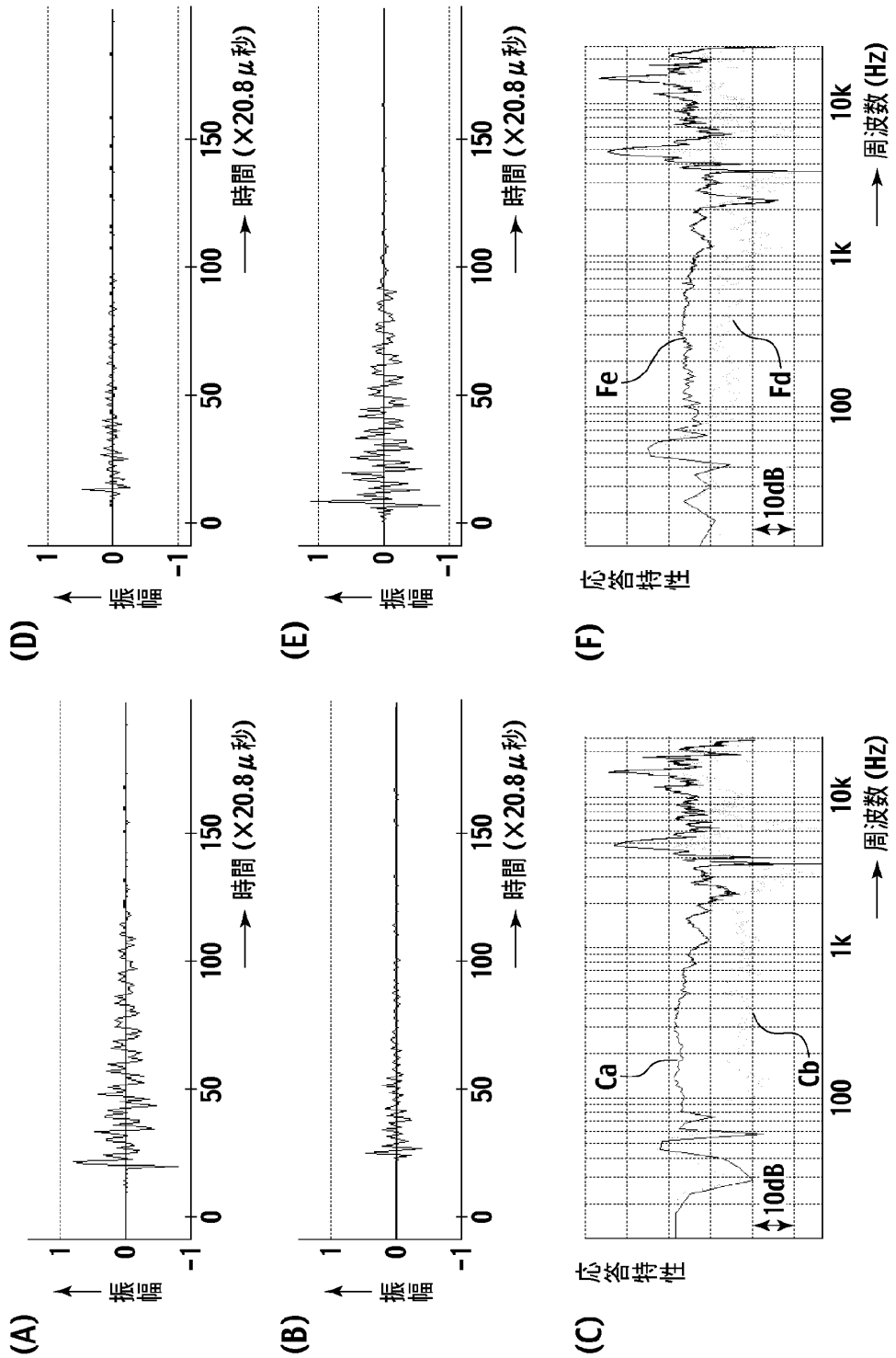
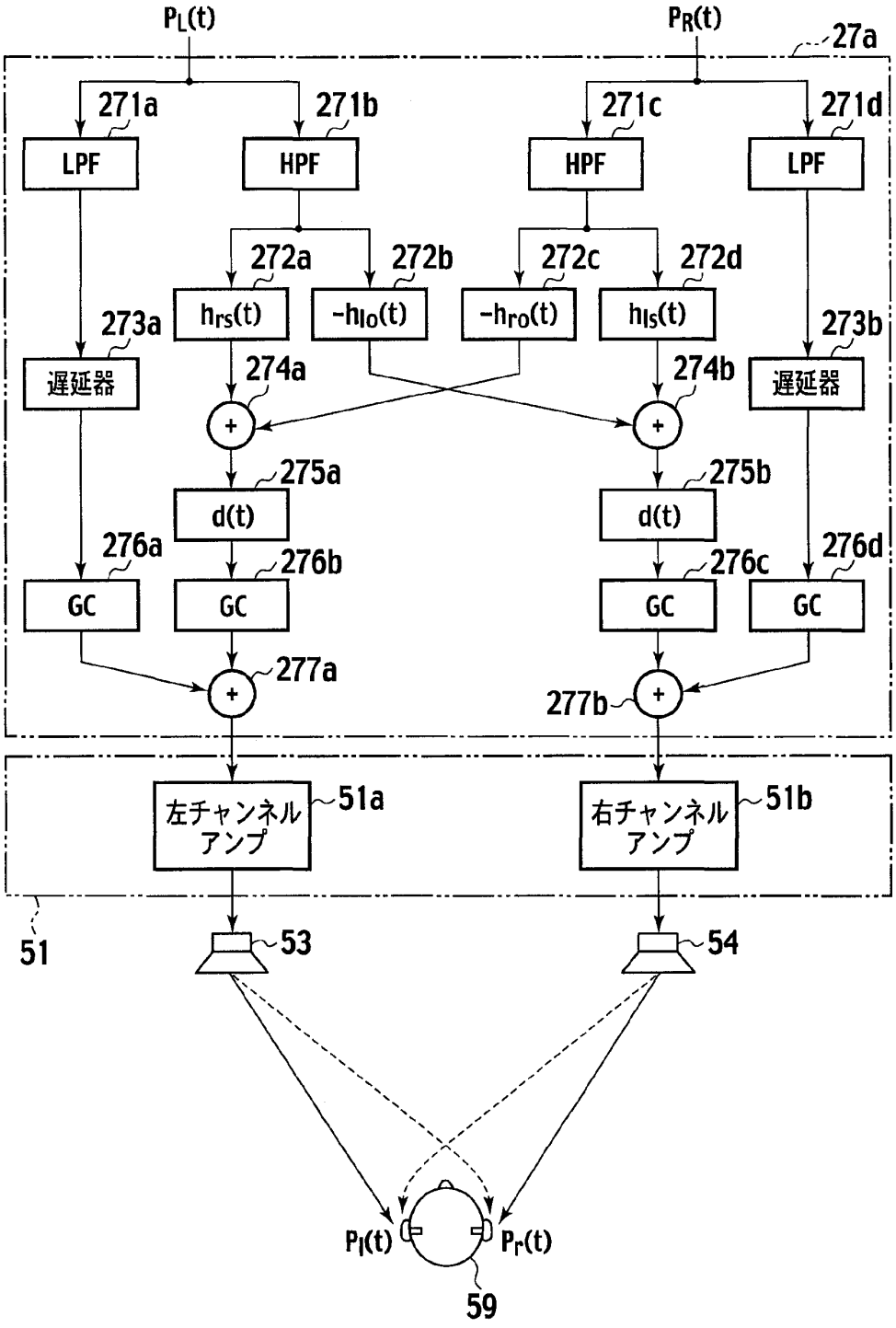
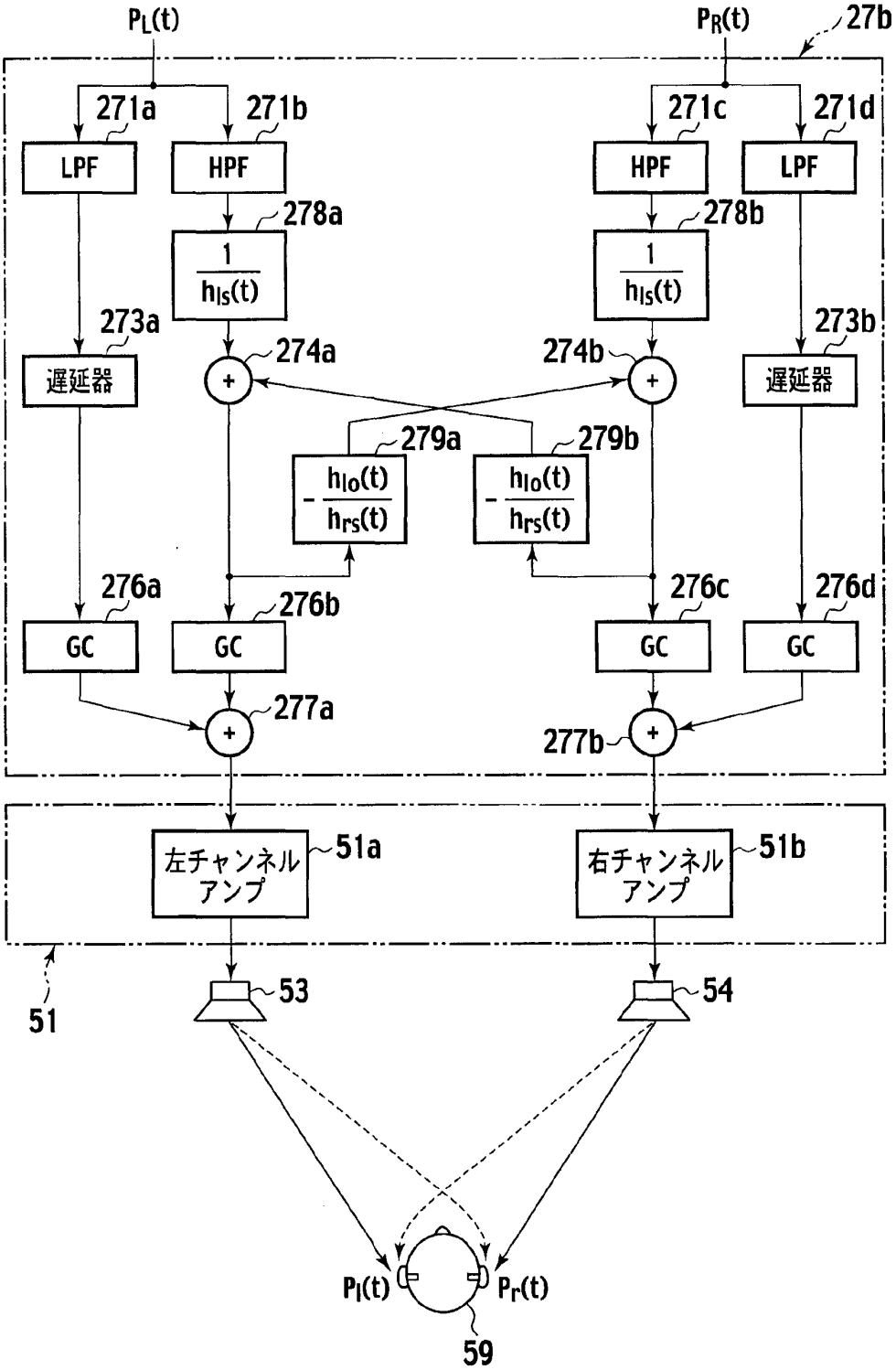


Exhibit 1004
Page 0559 of 1001

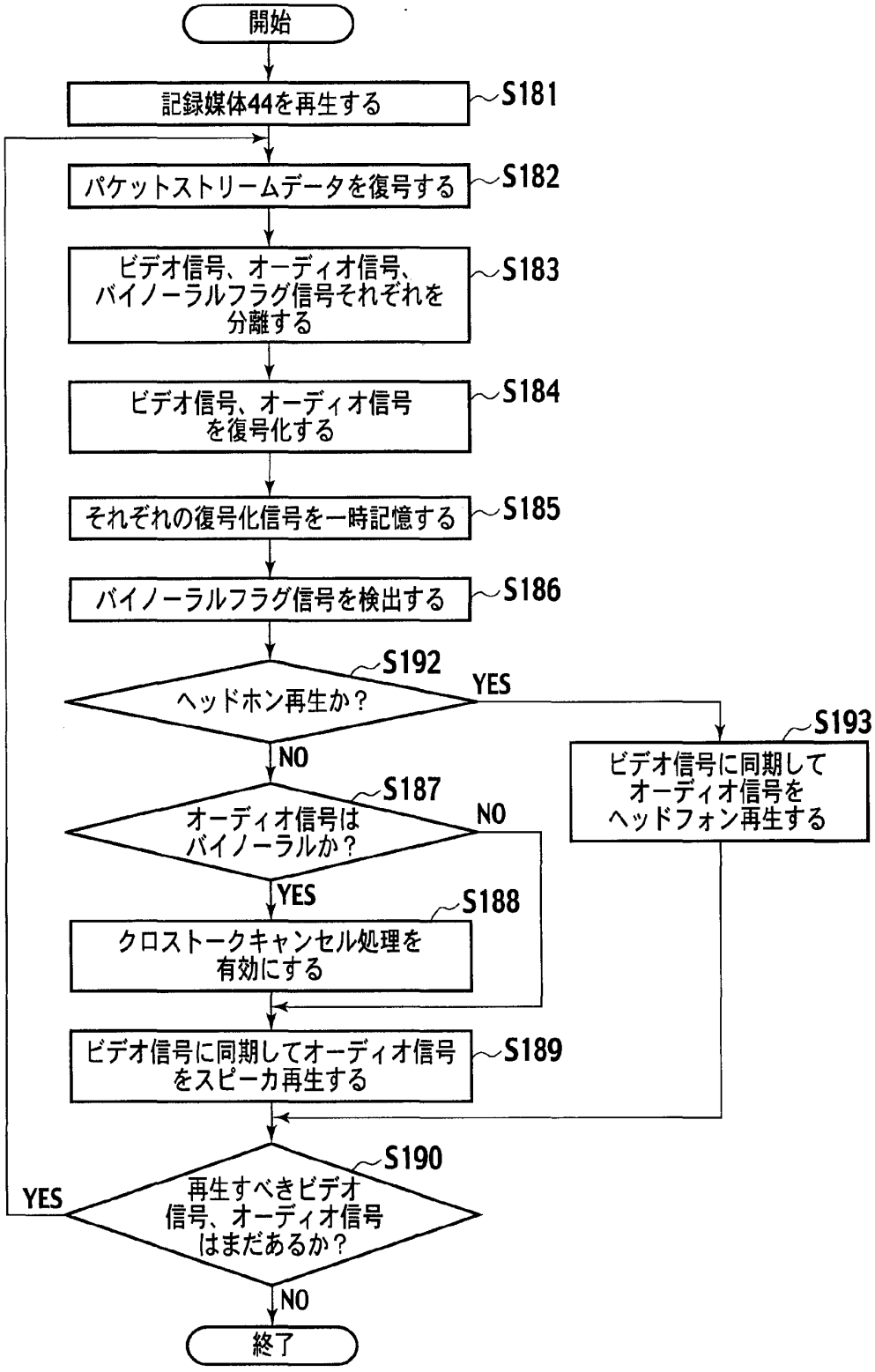
[図25]


PL(t)
PR(t)
27a
271a
271b
271c
271d
LPF
HPF
HPF
LPF
272a
272b
272c
272d
hrs(t)
-hlo(t)
-hro(t)
hls(t)
273a
273b
遅延器
遅延器
274a
274b
+
+
275a
275b
d(t)
d(t)
276a
276b
276c
276d
GC
GC
GC
GC
277a
277b
+
+
左チャンネル
アンプ
51a
右チャンネル
アンプ
51b
51
53
54
Pl(t)
Pr(t)
59

[図26]


$P_L(t)$
$P_R(t)$
27b
271a
271b
271c
271d
LPF
HPF
HPF
LPF
278a
278b
$\frac{1}{h_{ls}(t)}$
$\frac{1}{h_{ls}(t)}$
273a
273b
遅延器
遅延器
274a
274b
279a
279b
$-\frac{h_{lo}(t)}{h_{rs}(t)}$
$-\frac{h_{lo}(t)}{h_{rs}(t)}$
276a
276b
276c
276d
GC
GC
GC
GC
277a
277b
左チャンネル
アンプ
51a
右チャンネル
アンプ
51b
51
53
54
$P_l(t)$
$P_r(t)$
59

[図27]


開始
記録媒体44を再生する
S181
パケットストリームデータを復号する
S182
ビデオ信号、オーディオ信号、
バイノーラルフラグ信号それぞれを
分離する
S183
ビデオ信号、オーディオ信号
を復号化する
S184
それぞれの復号化信号を一時記憶する
S185
バイノーラルフラグ信号を検出する
S186
S192
ヘッドホン再生か？
YES
NO
S193
ビデオ信号に同期して
オーディオ信号を
ヘッドフォン再生する
S187
オーディオ信号は
バイノーラルか？
NO
YES
S188
クロストークキャンセル処理を
有効にする
ビデオ信号に同期してオーディオ信号
をスピーカ再生する
S189
S190
再生すべきビデオ
信号、オーディオ信号
はまだあるか？
YES
NO
終了

[図28]

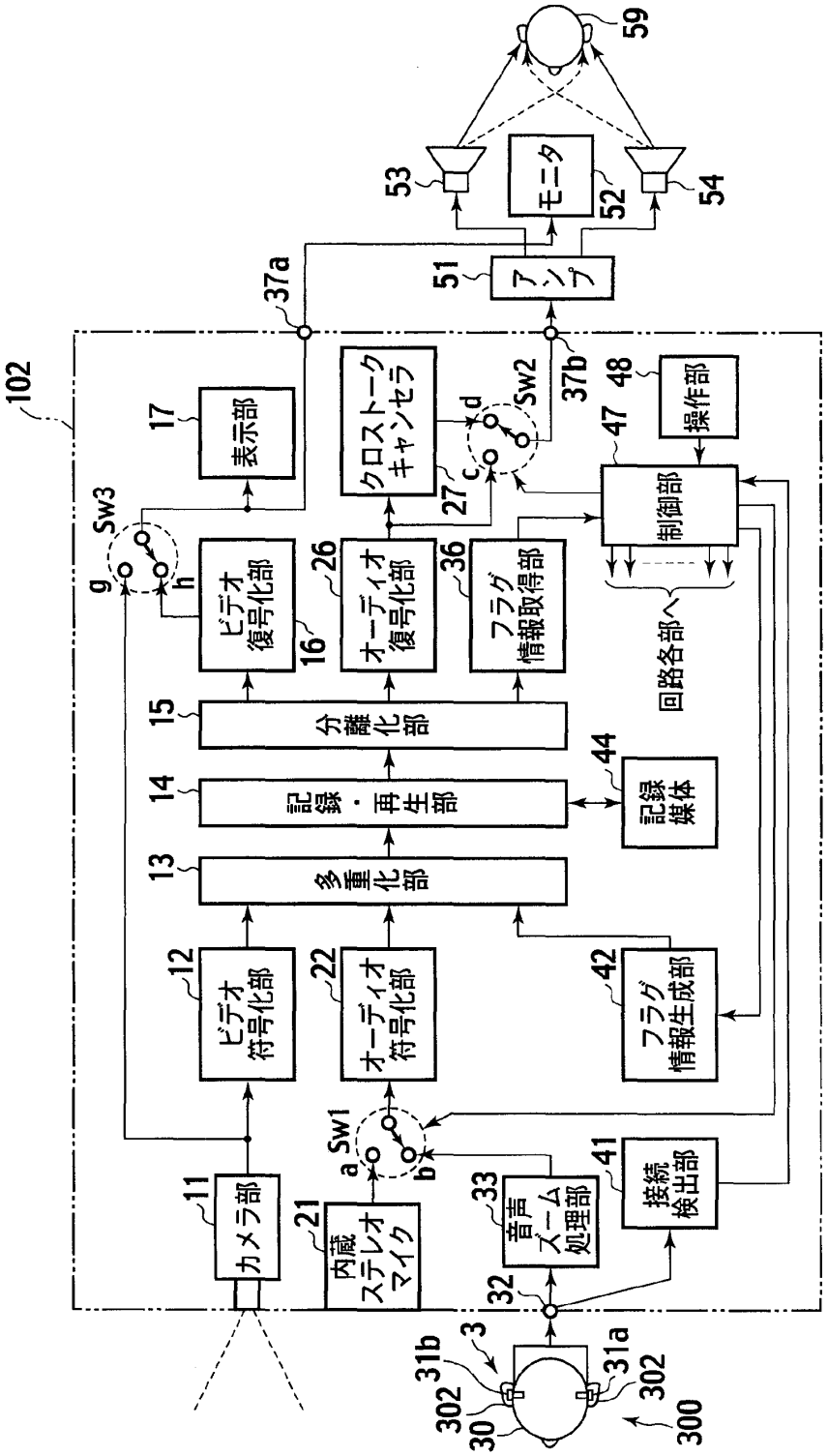
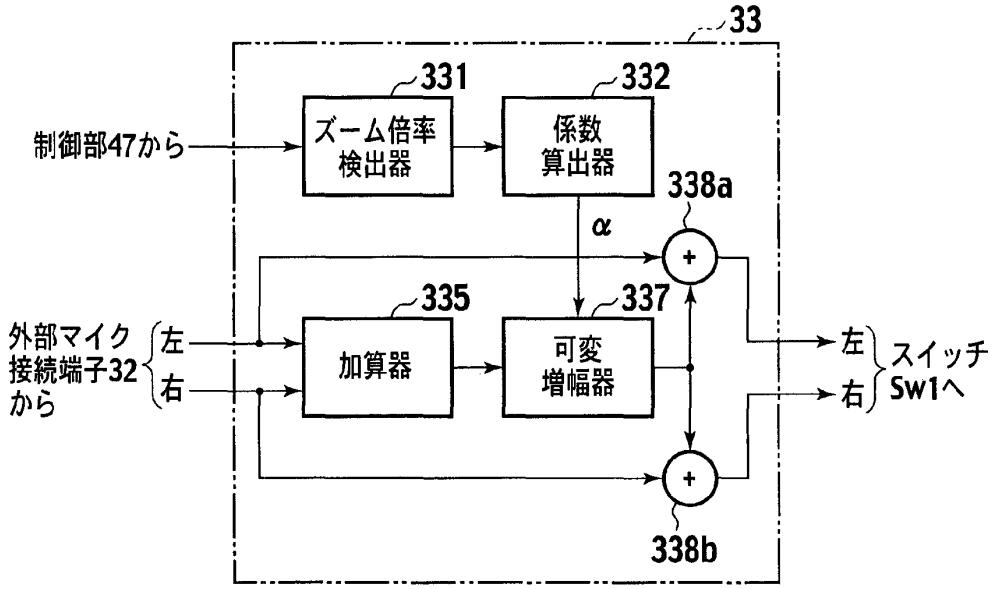
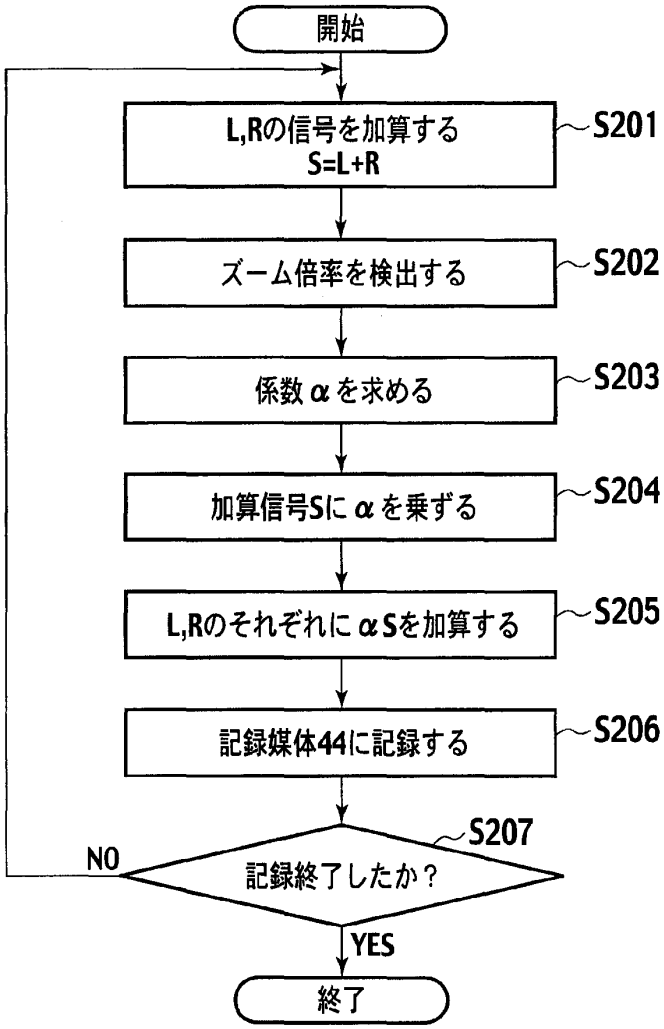


Exhibit 1004
Page 0563 of 1001

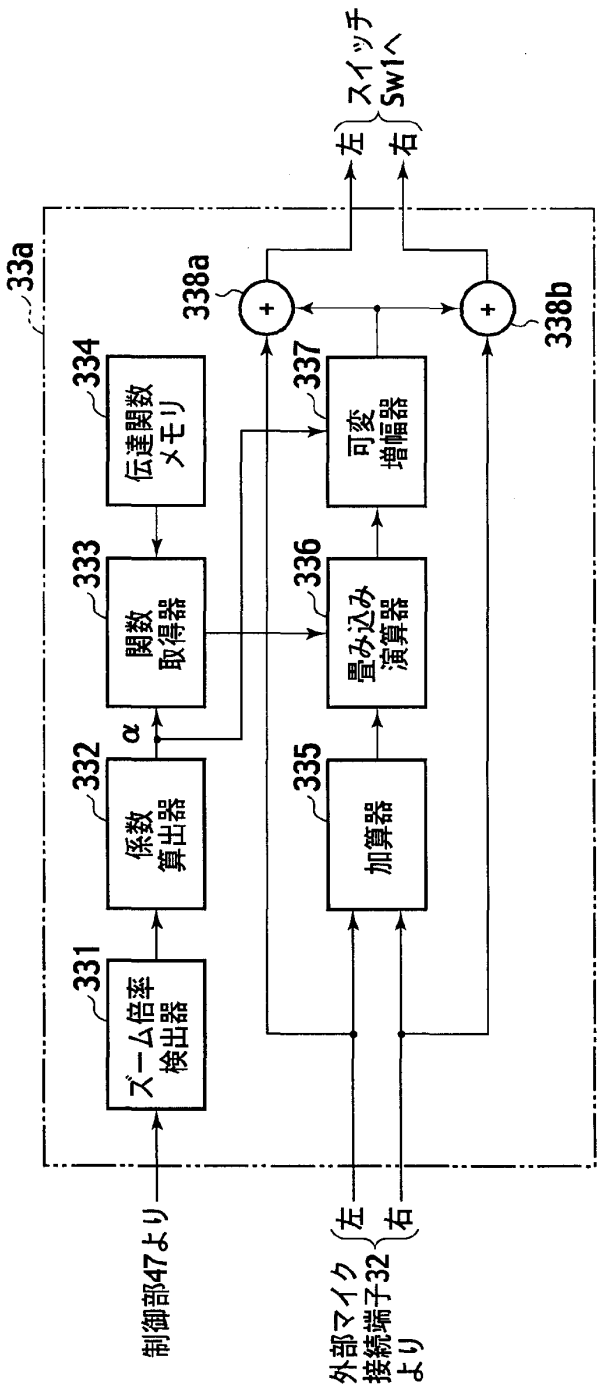
[図29]


33
331
332
制御部47から
ズーム倍率検出器
係数算出器
338a
α
335
337
外部マイク接続端子32から
左
右
加算器
可変増幅器
左
右
スイッチSw1へ
338b

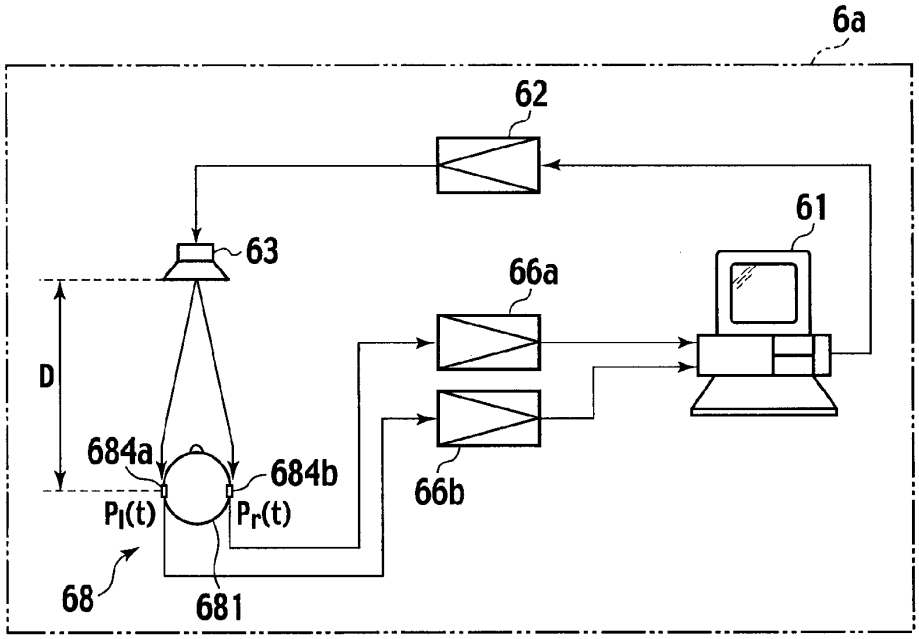
[図30]


開始
L,Rの信号を加算する
S=L+R
S201
ズーム倍率を検出する
S202
係数αを求める
S203
加算信号Sにαを乗ずる
S204
L,Rのそれぞれにα Sを加算する
S205
記録媒体44に記録する
S206
S207
NO
記録終了したか？
YES
終了

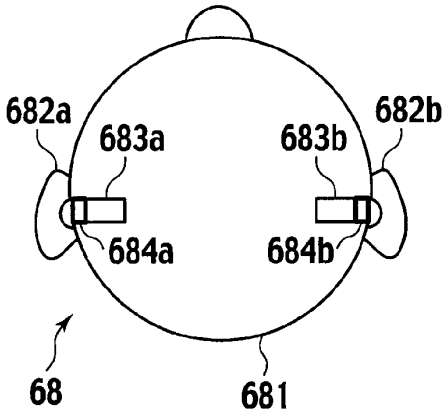
[図31]


33a
338a
338b
334
伝達関数メモリ
337
可変増幅器
333
関数取得器
336
畳み込み演算器
α
332
係数算出器
335
加算器
331
ズーム倍率検出器
左
右
スイッチSw1へ
制御部47より
外部マイク接続端子32より
左
右

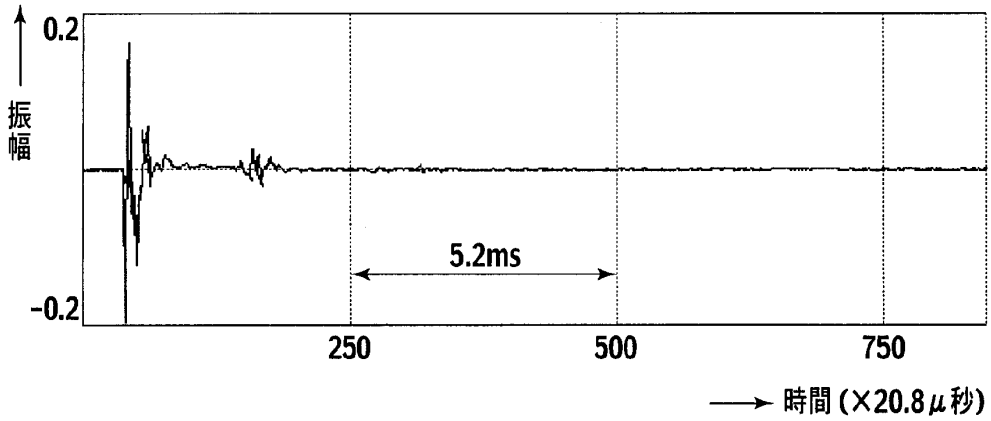
[図32]



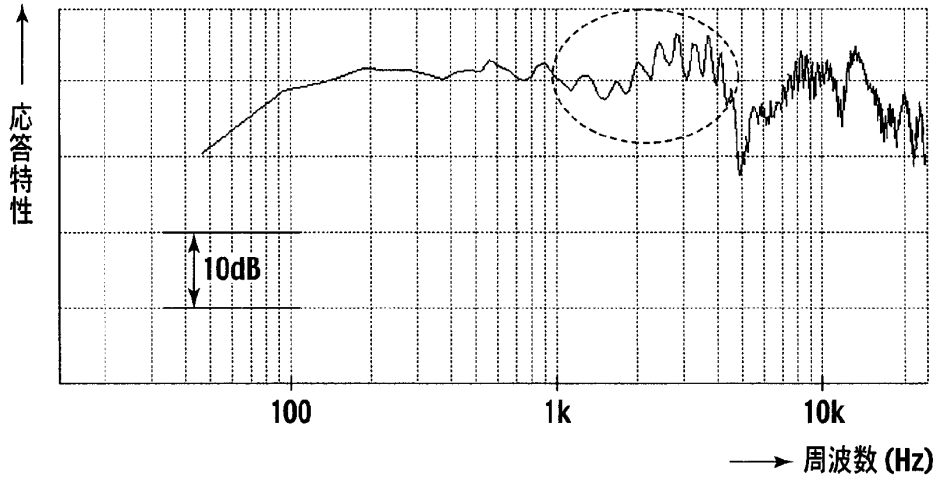
[図33]



[図34]



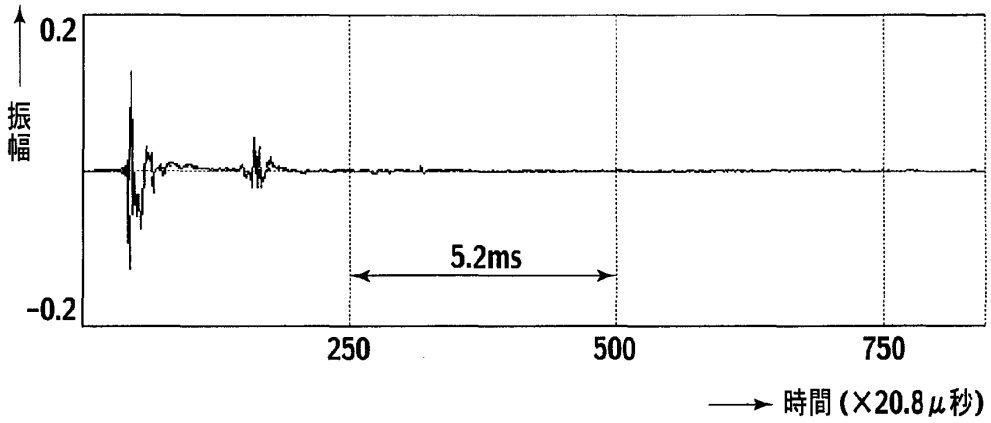
(A) インパルス応答波形



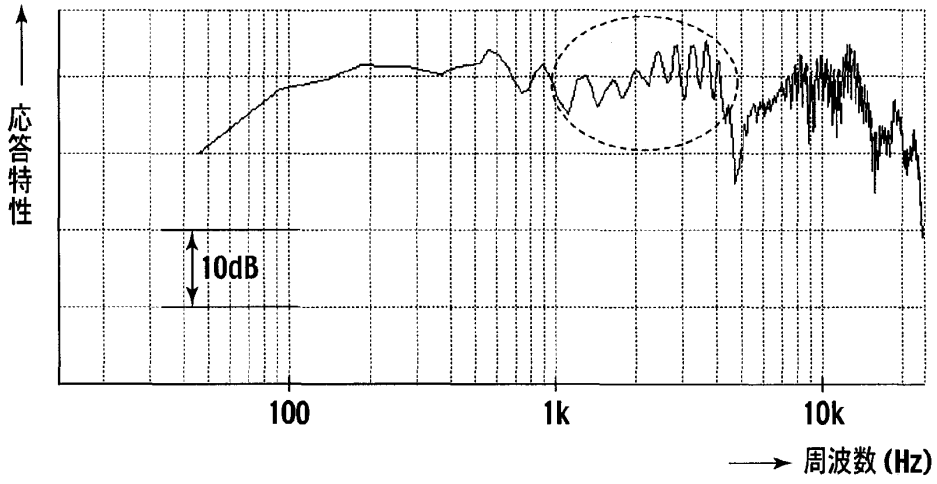
(B) 周波数応答特性

頭部伝達関数hl (t) D=0.5m

[図35]



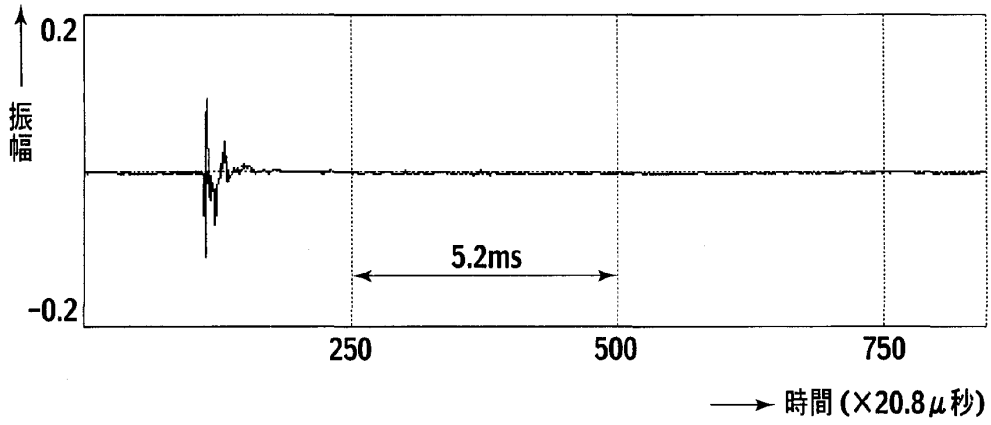
(A) インパルス応答波形



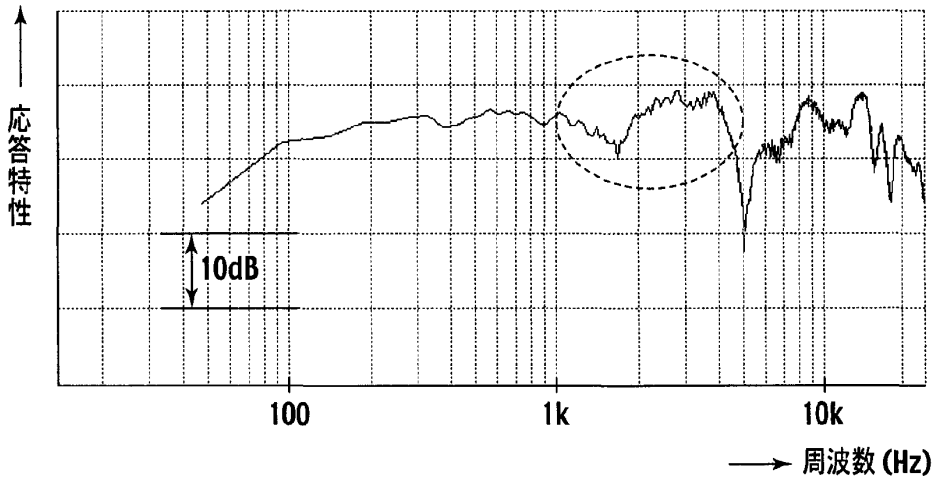
(B) 周波数応答特性

頭部伝達関数hr (t) D=0.5m

[図36]



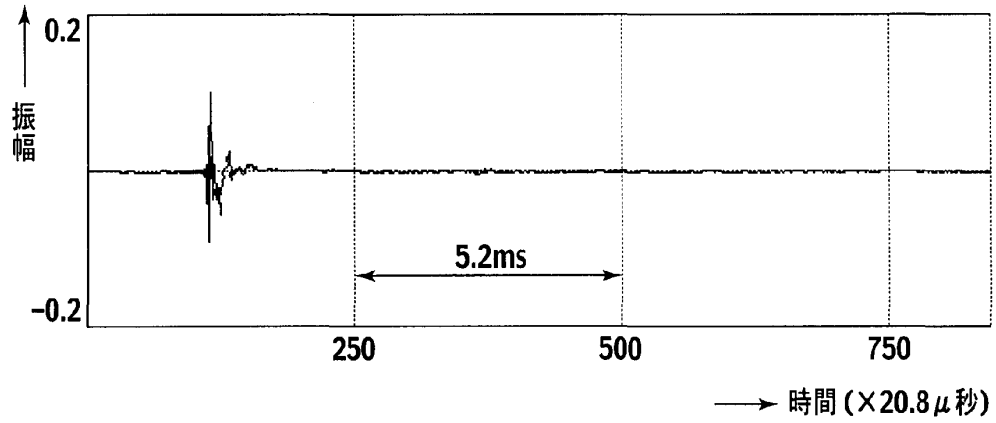
(A) インパルス応答波形



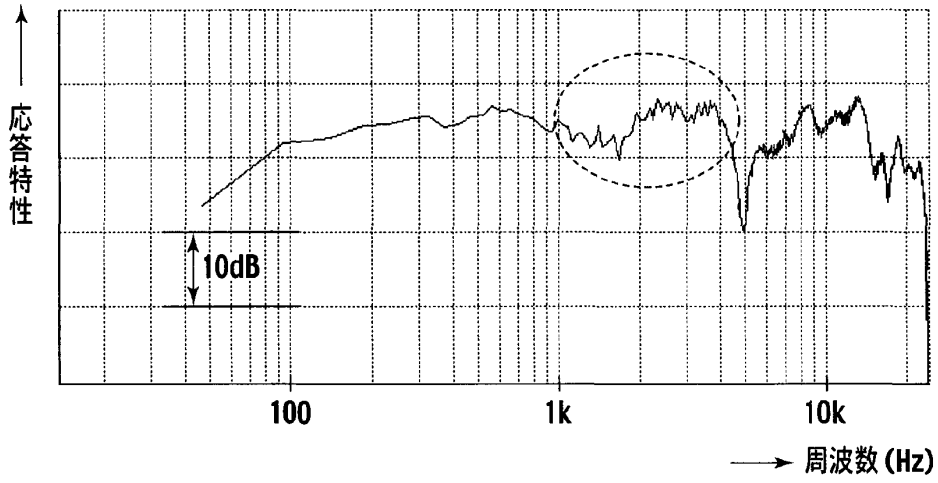
(B) 周波数応答特性

頭部伝達関数hl (t) D=1m

[図37]



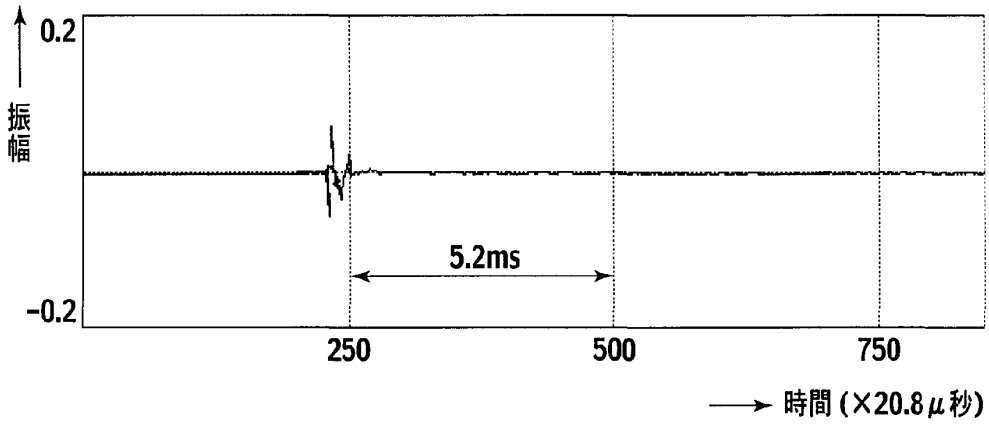
(A) インパルス応答波形



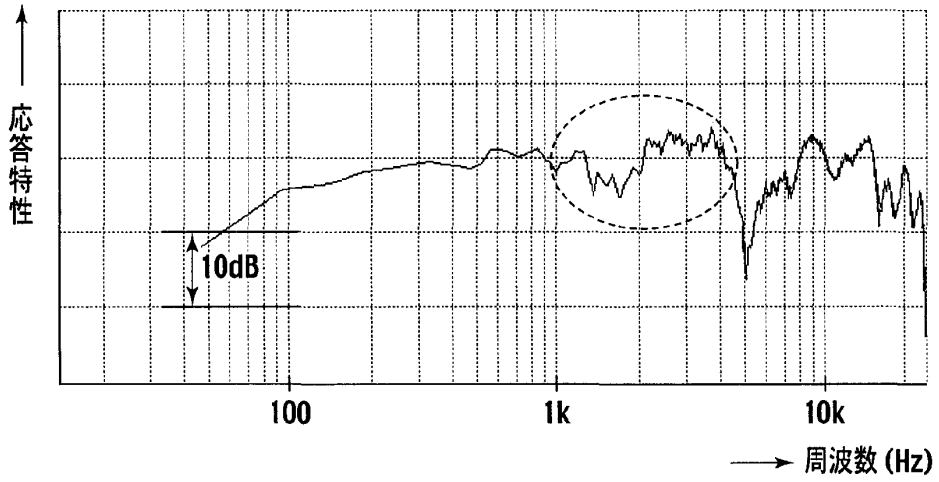
(B) 周波数応答特性

頭部伝達関数hr (t) D=1m

[図38]



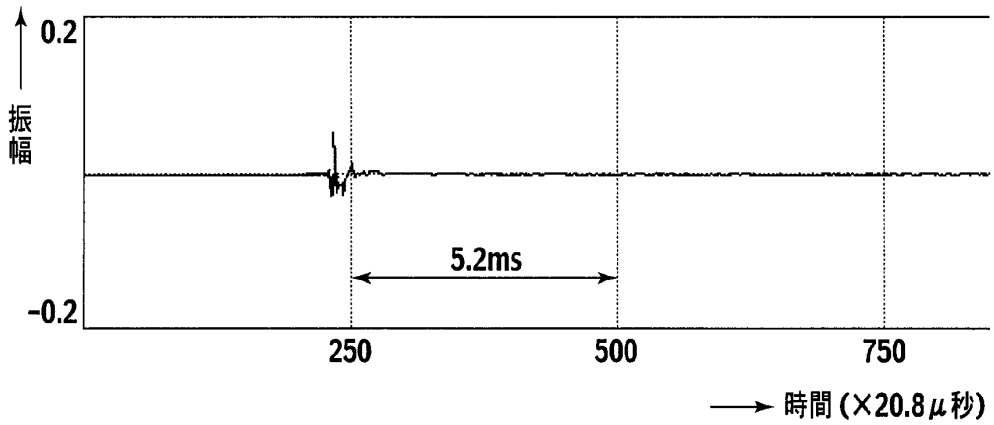
(A) インパルス応答波形



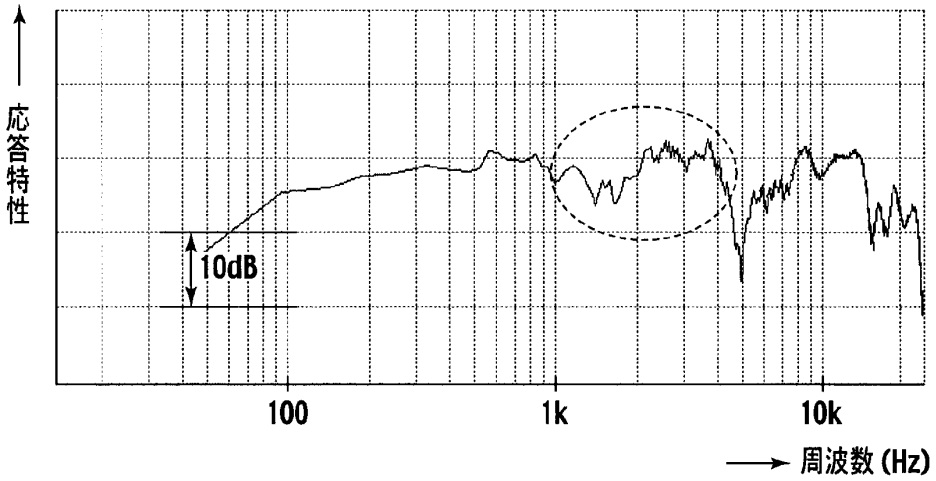
(B) 周波数応答特性

頭部伝達関数hl (t) D=2m

[図39]



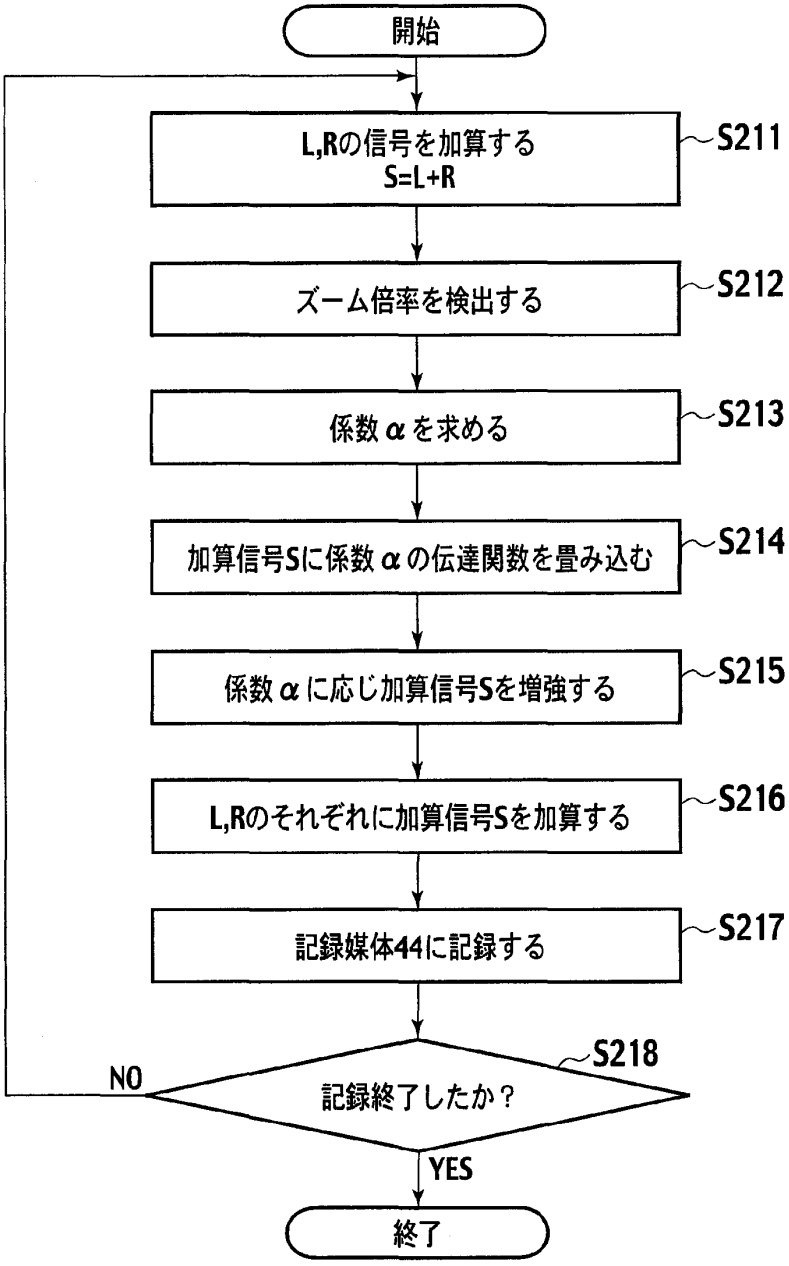
(A) インパルス応答波形



(B) 周波数応答特性

頭部伝達関数hr (t) D=2m

[図40]


開始
L,Rの信号を加算する
S=L+R
S211
ズーム倍率を検出する
S212
係数αを求める
S213
加算信号Sに係数αの伝達関数を畳み込む
S214
係数αに応じ加算信号Sを増強する
S215
L,Rのそれぞれに加算信号Sを加算する
S216
記録媒体44に記録する
S217
記録終了したか？
S218
NO
YES
終了

[図41]

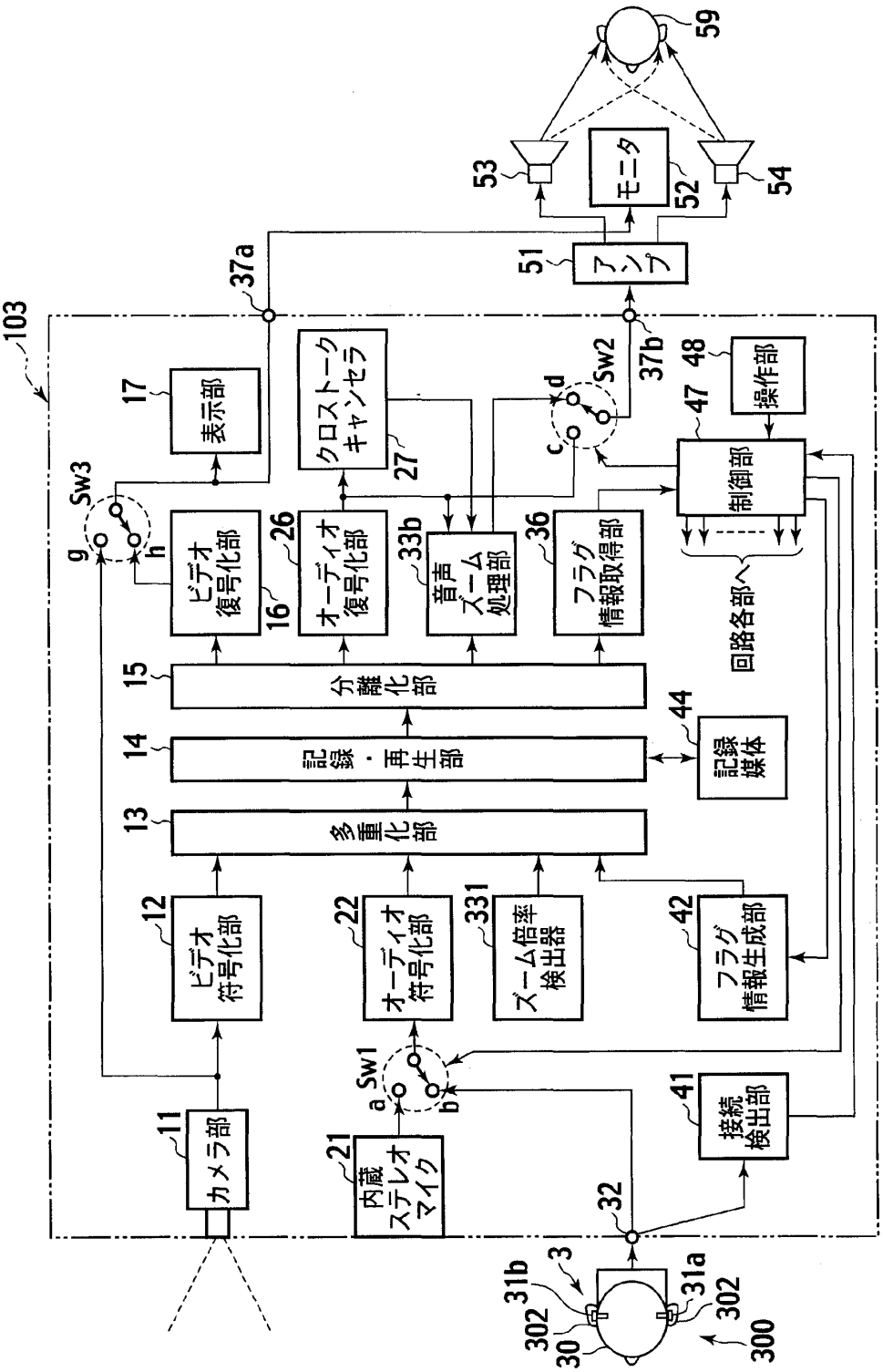
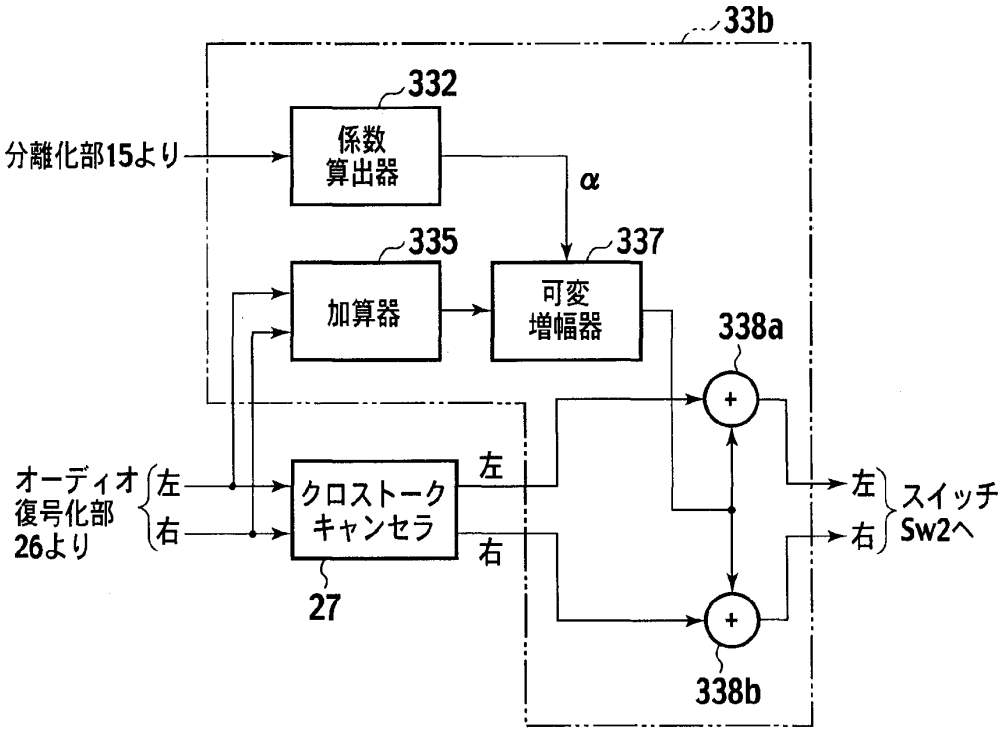
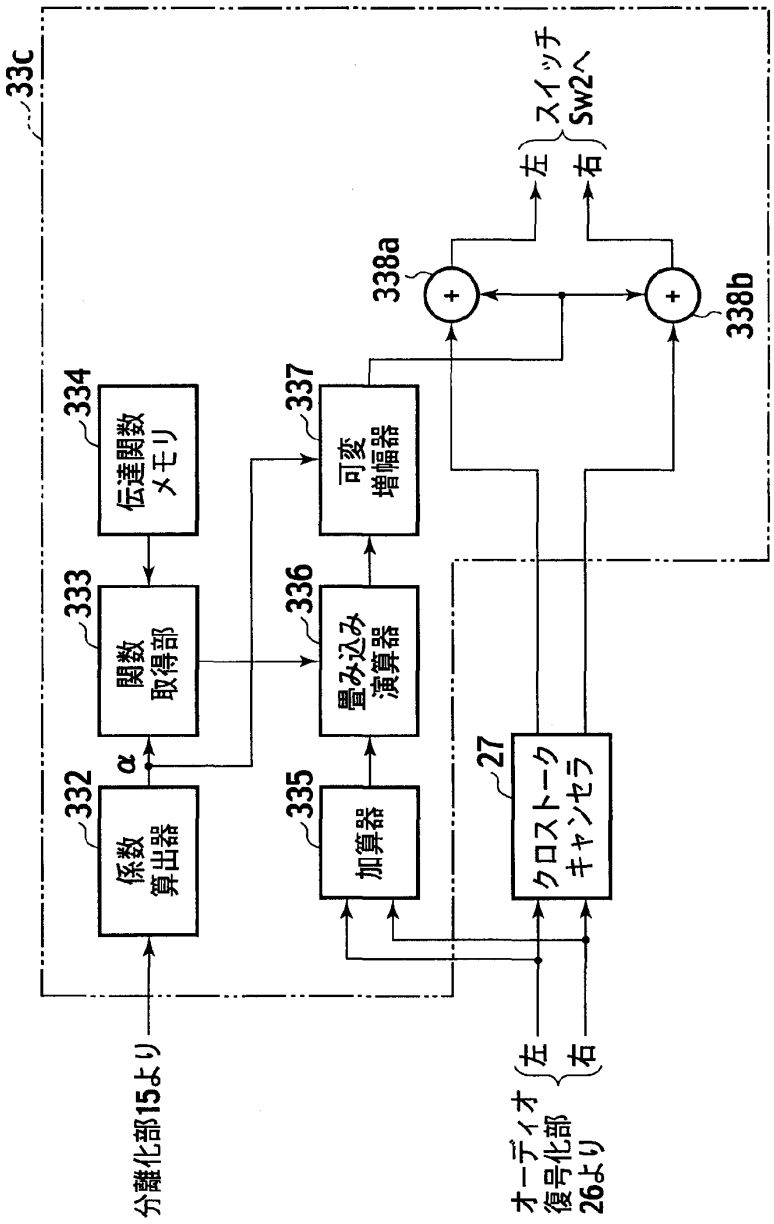


Exhibit 1004
Page 0576 of 1001

[図42]


33b
332
係数
算出器
分離化部15より
α
335
337
加算器
可変
増幅器
338a
オーディオ
復号化部
26より
左
右
クロストーク
キャンセラ
27
左
右
左
右
スイッチ
Sw2へ
338b

[図43]


33c
334
伝達関数
メモリ
333
関数
取得部
332
係数
算出器
α
337
可変
増幅器
336
畳み込み
演算器
335
加算器
27
クロストーク
キャンセラ
338a
338b
左
右
スイッチ
Sw2へ
分離化部15より
オーディオ
復号化部
26より
左
右

[図44]

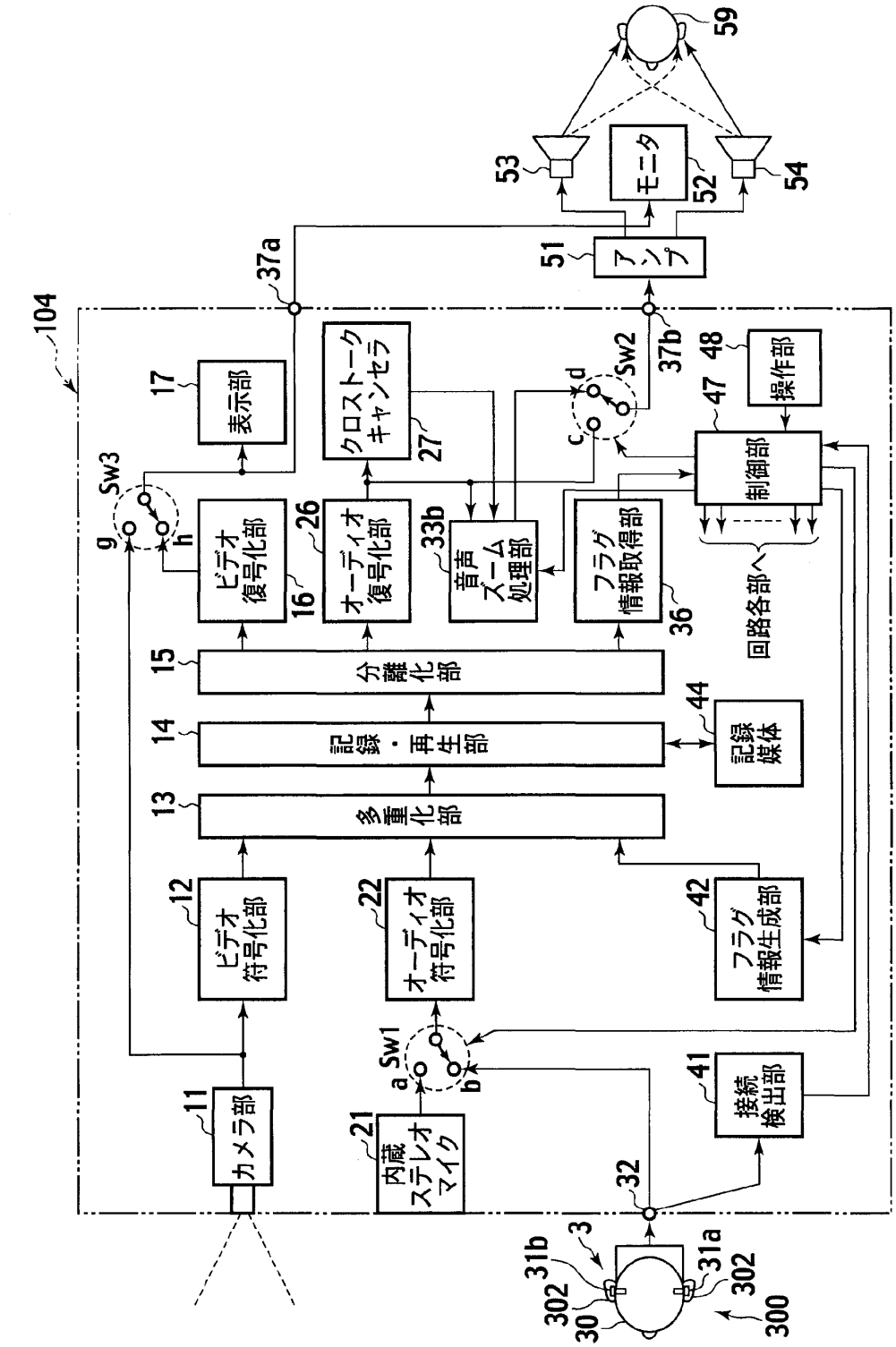
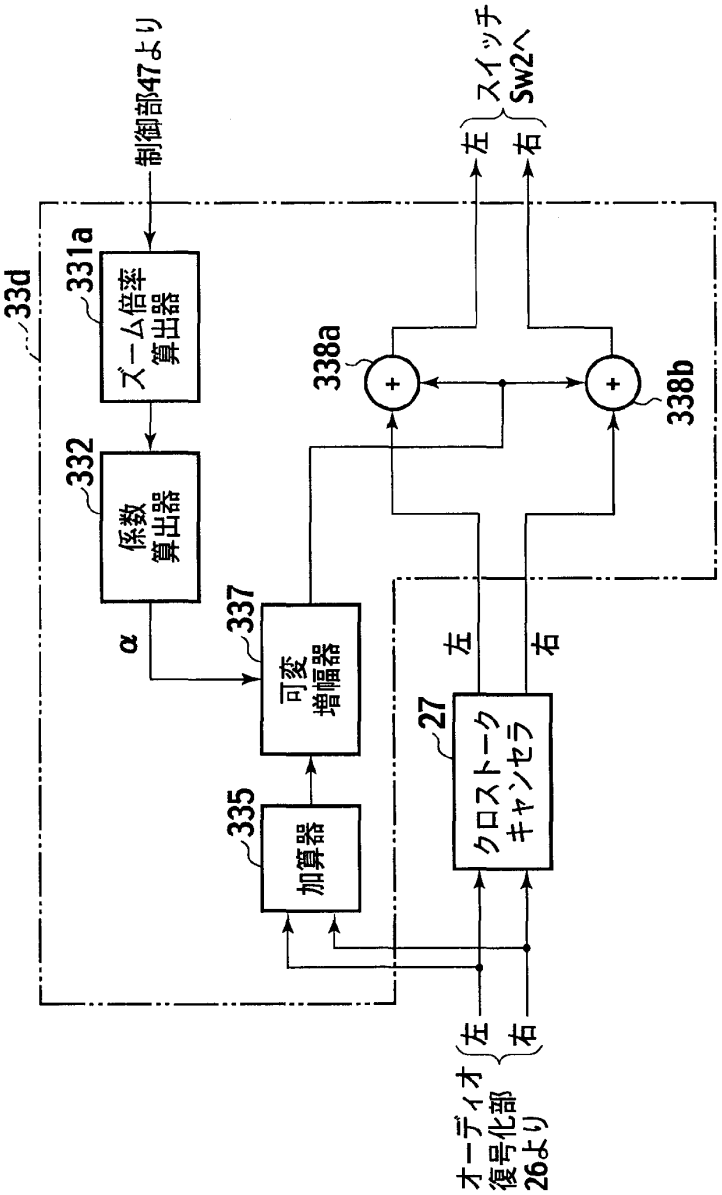
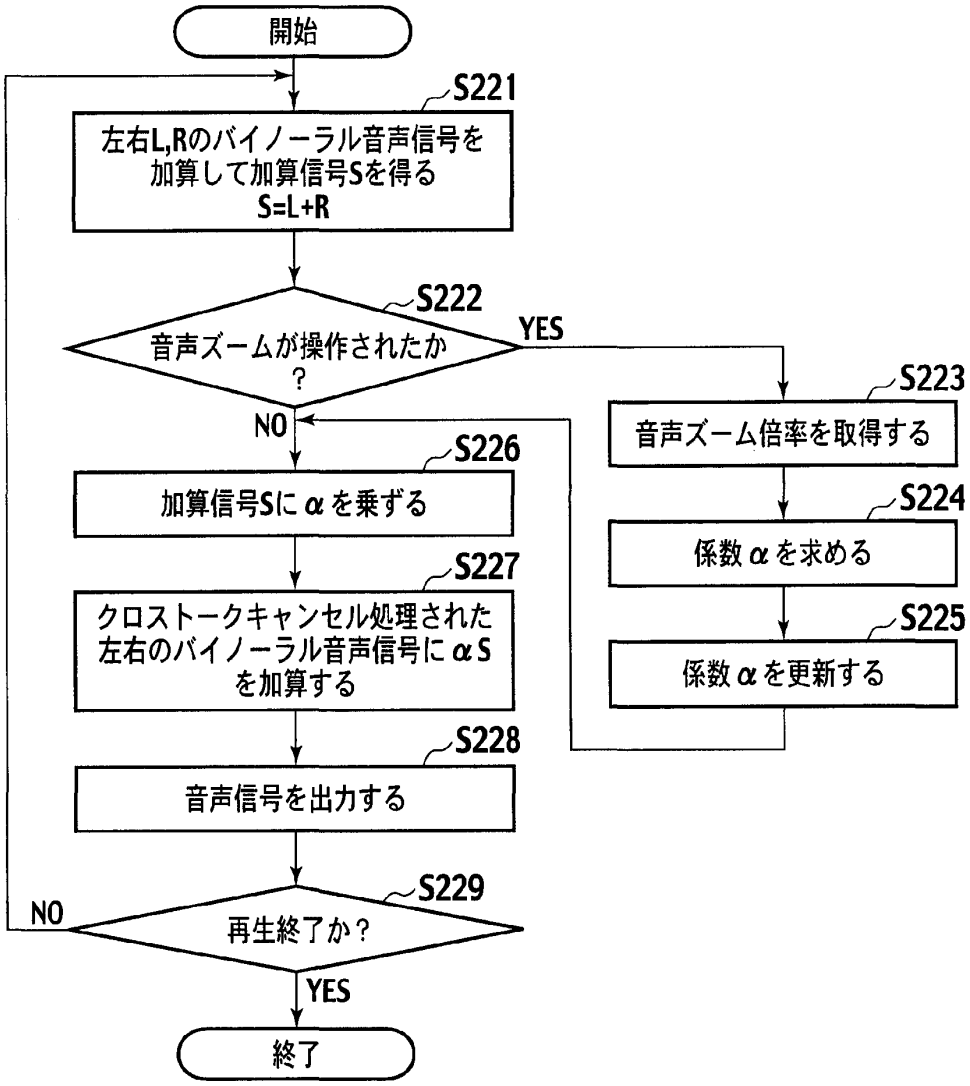


Exhibit 1004
Page 0579 of 1001

[図45]


制御部47より
33d
331a
ズーム倍率算出器
332
係数算出器
α
337
可変増幅器
335
加算器
338a
338b
27
クロストークキャンセラ
左
右
スイッチSw2へ
オーディオ復号化部26より

[図46]


開始
S221
左右L,Rのバイノーラル音声信号を
加算して加算信号Sを得る
S=L+R
S222
音声ズームが操作されたか
?
YES
NO
S223
音声ズーム倍率を取得する
S224
係数αを求める
S225
係数αを更新する
S226
加算信号Sにαを乗ずる
S227
クロストークキャンセル処理された
左右のバイノーラル音声信号にαS
を加算する
S228
音声信号を出力する
S229
再生終了か?
NO
YES
終了

[図47]

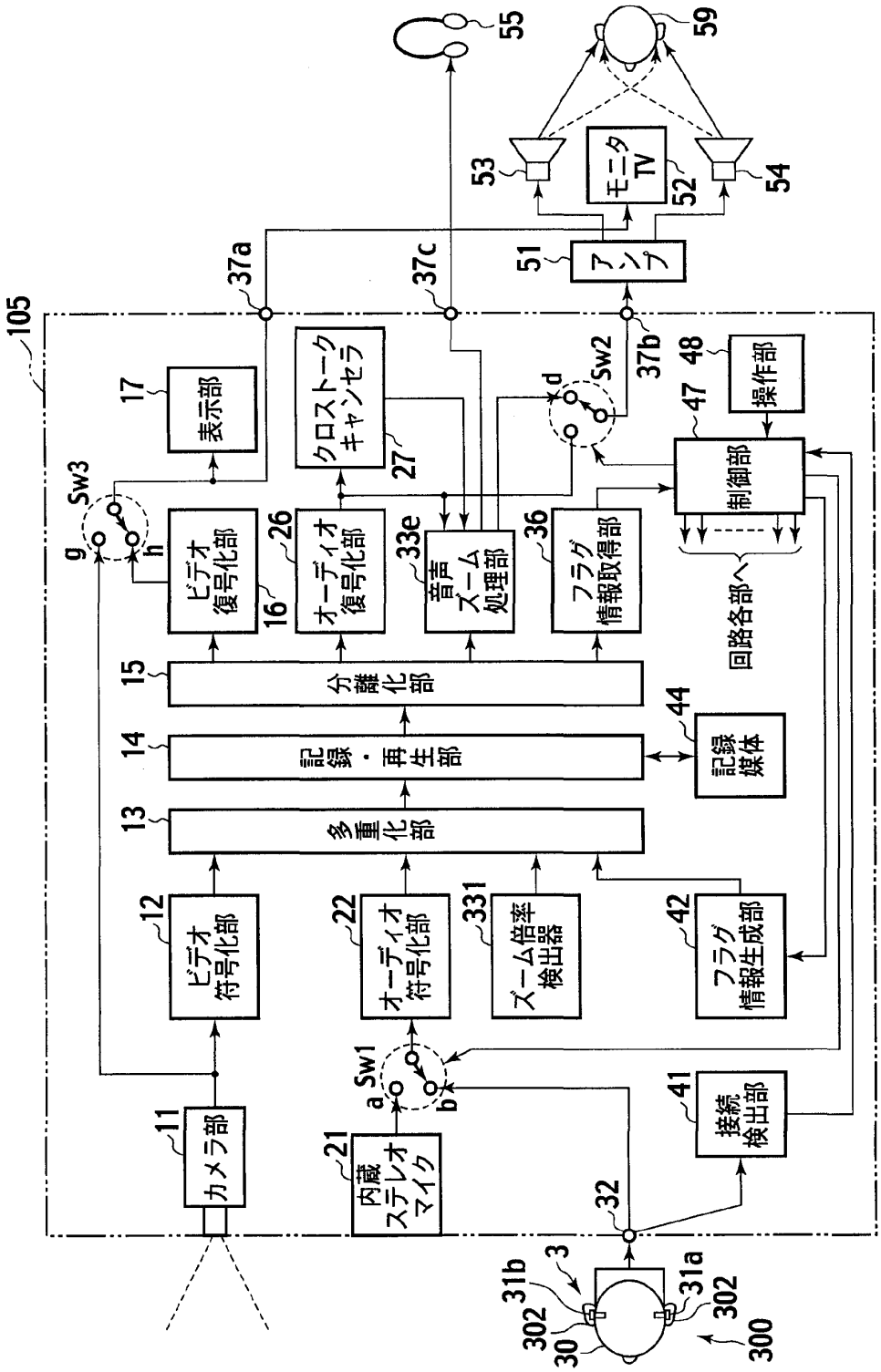
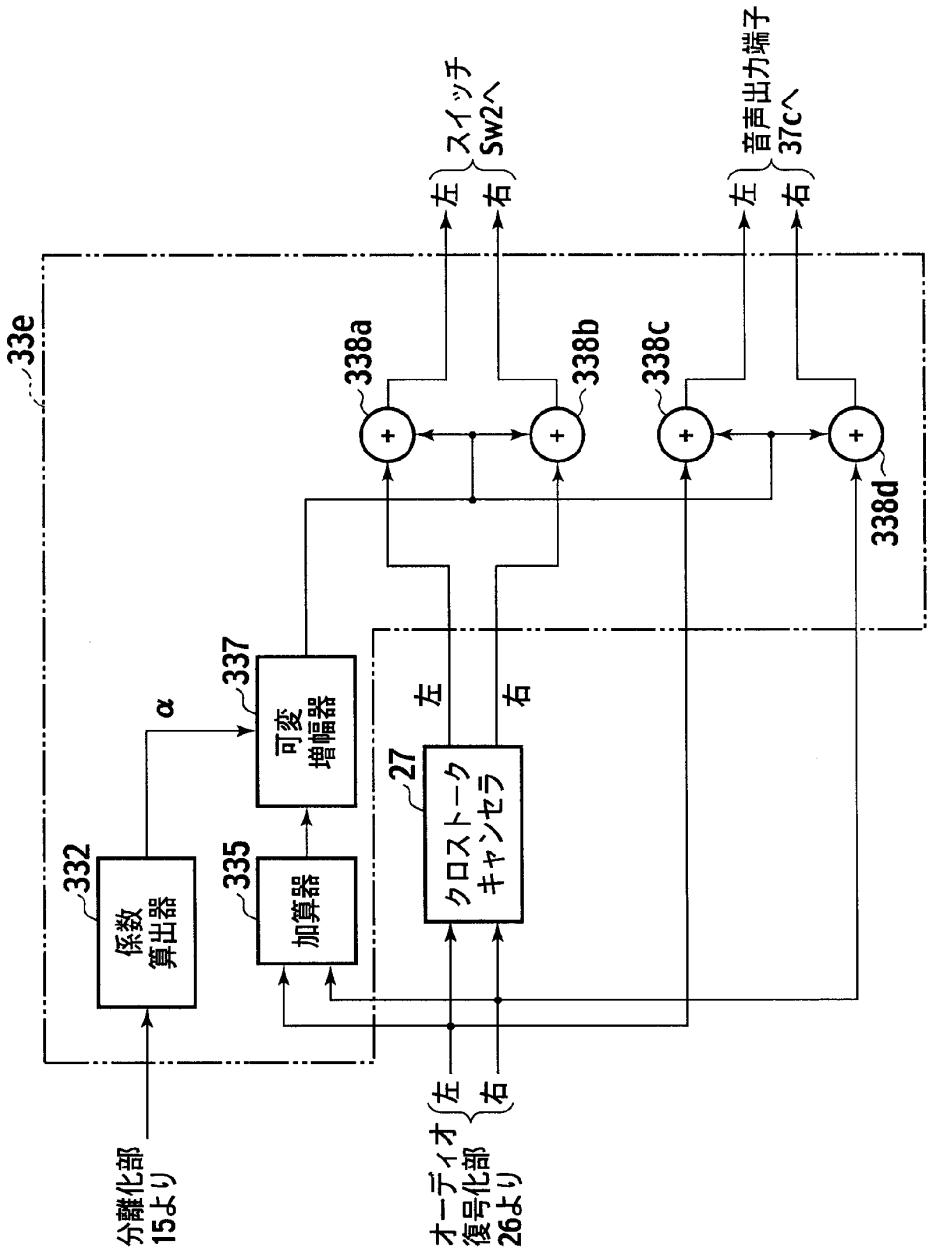
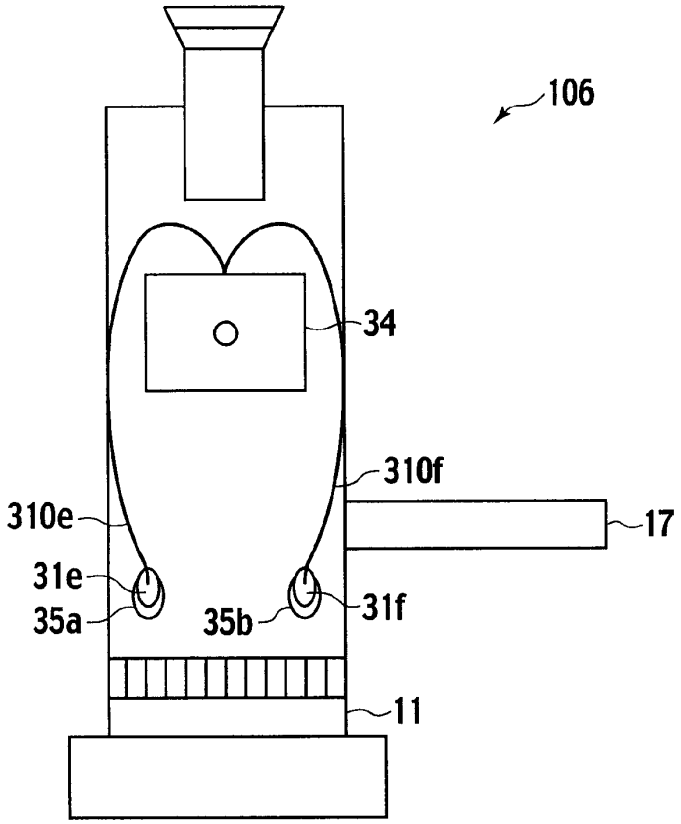


Exhibit 1004
Page 0582 of 1001

[図48]


33e
338a
338b
338c
338d
左
右
スイッチ
Sw2へ
音声出力端子
37cへ
337
可変
増幅器
α
27
クロストーク
キャンセラ
332
係数
算出器
335
加算器
分離化部
15より
オーディオ
復号化部
26より

[図49]



[図50]

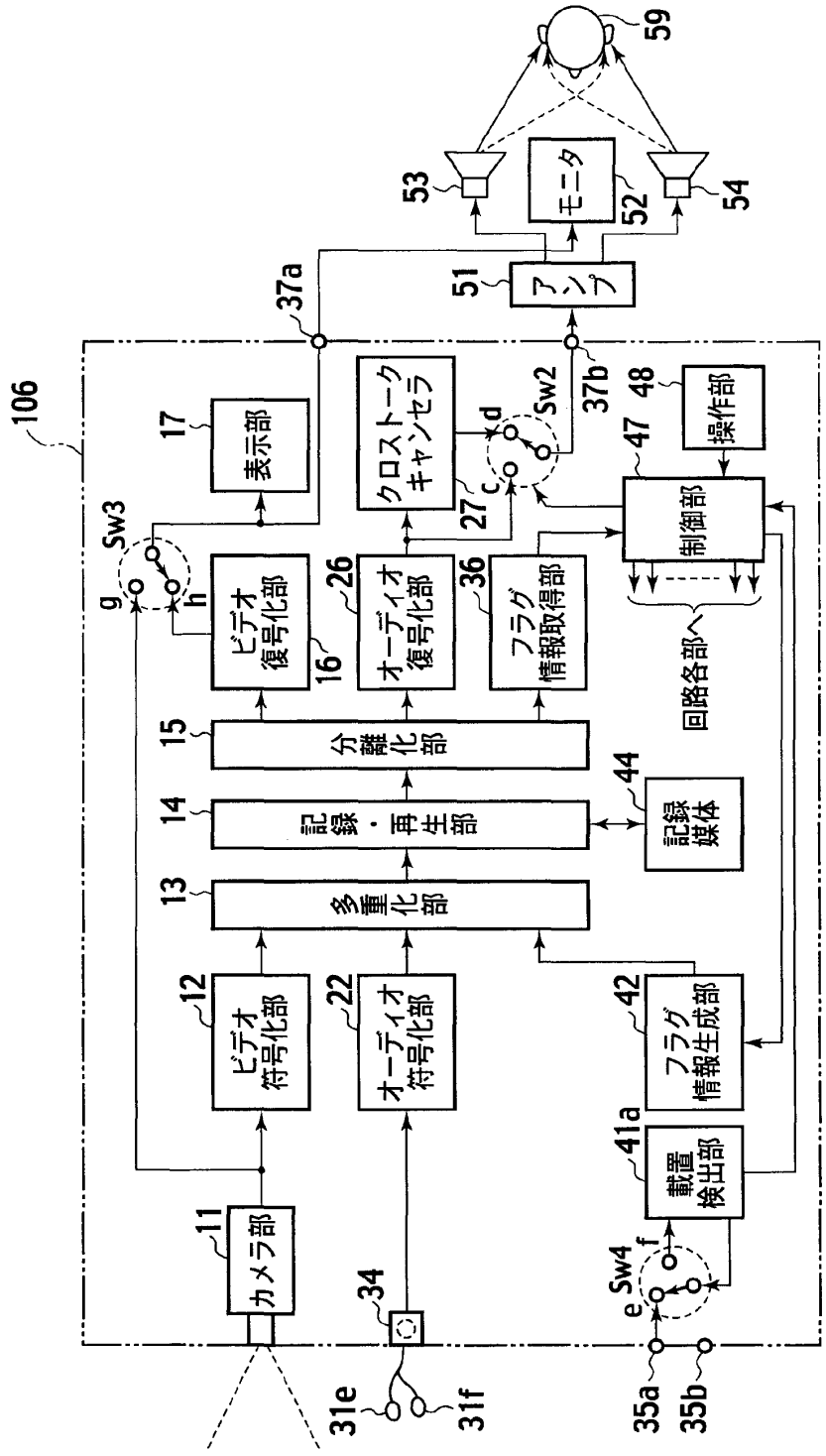
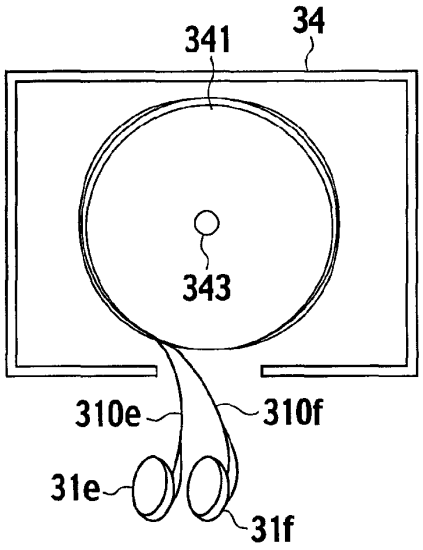


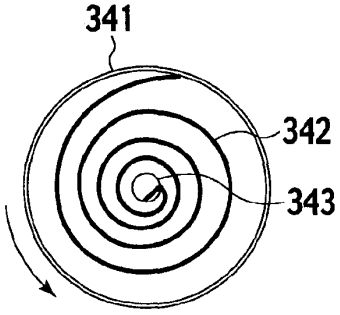
Exhibit 1004
Page 0585 of 1001

[図51]

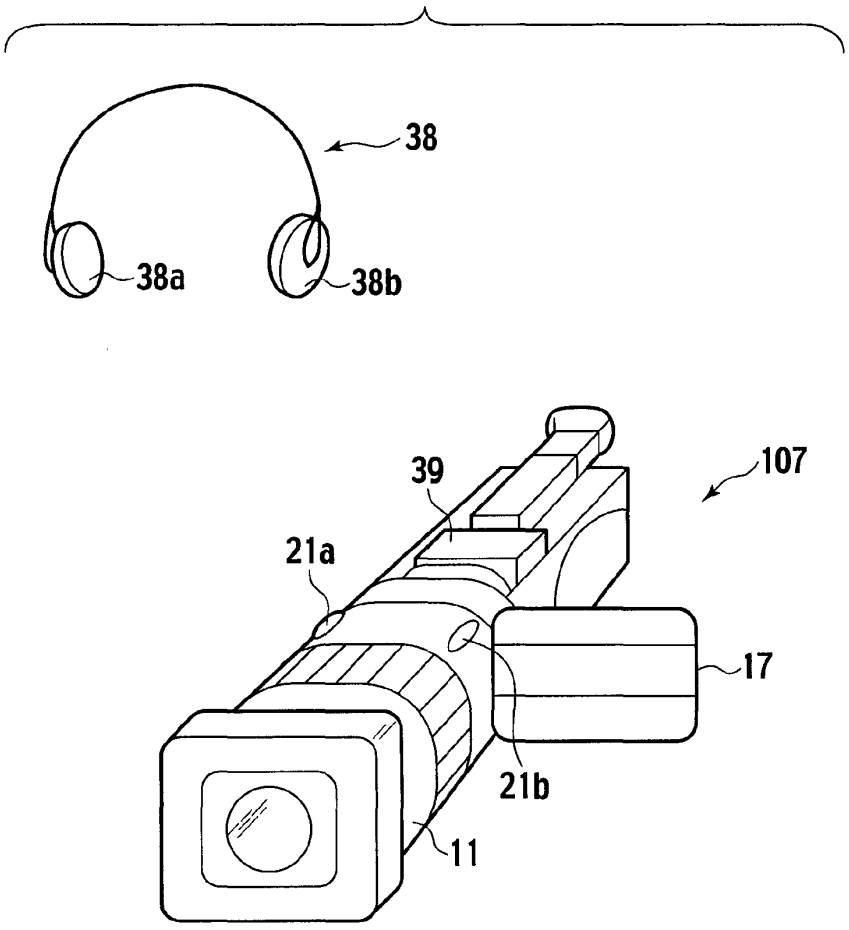
(A)



(B)



[図52]



[図53]

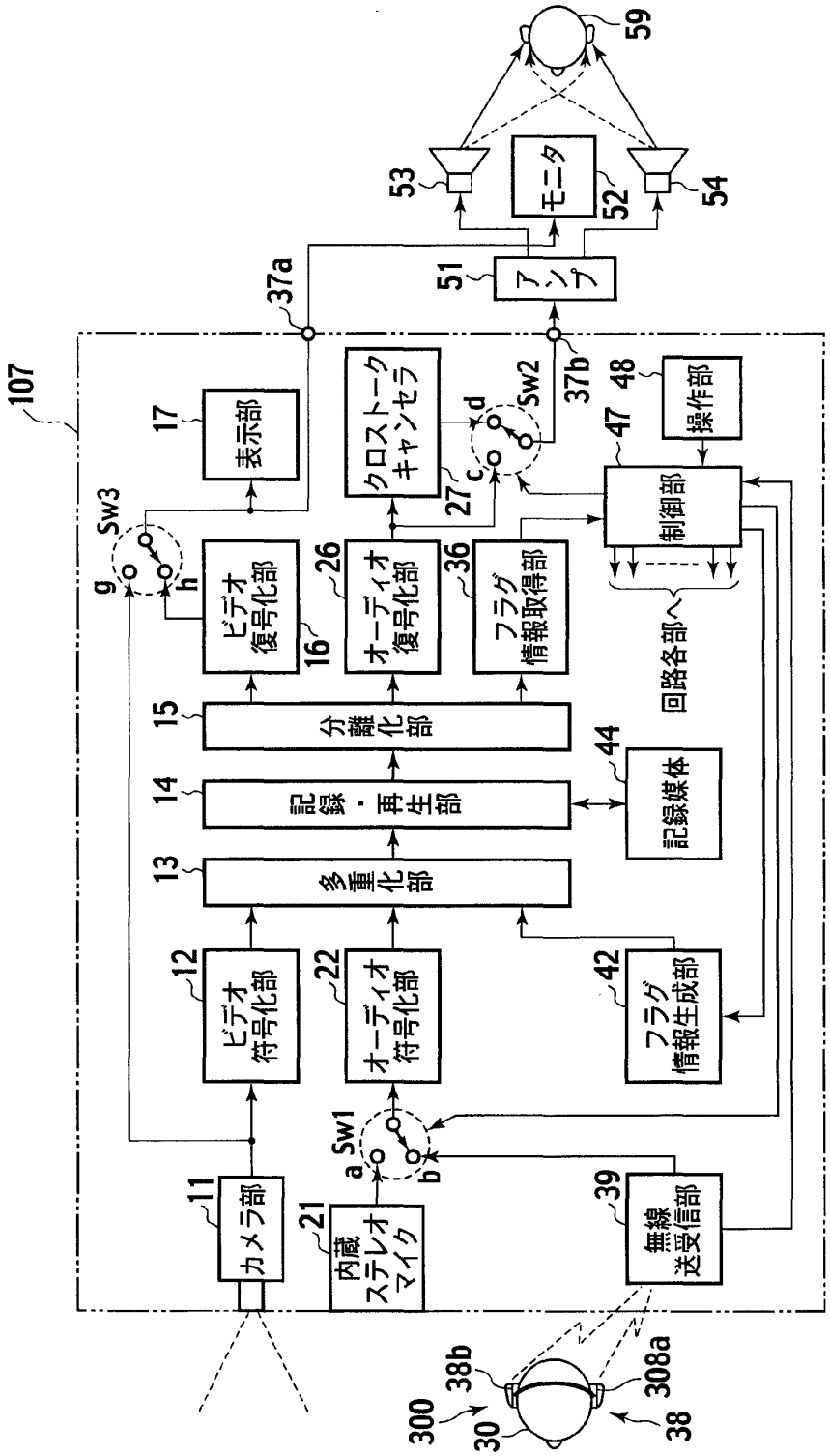
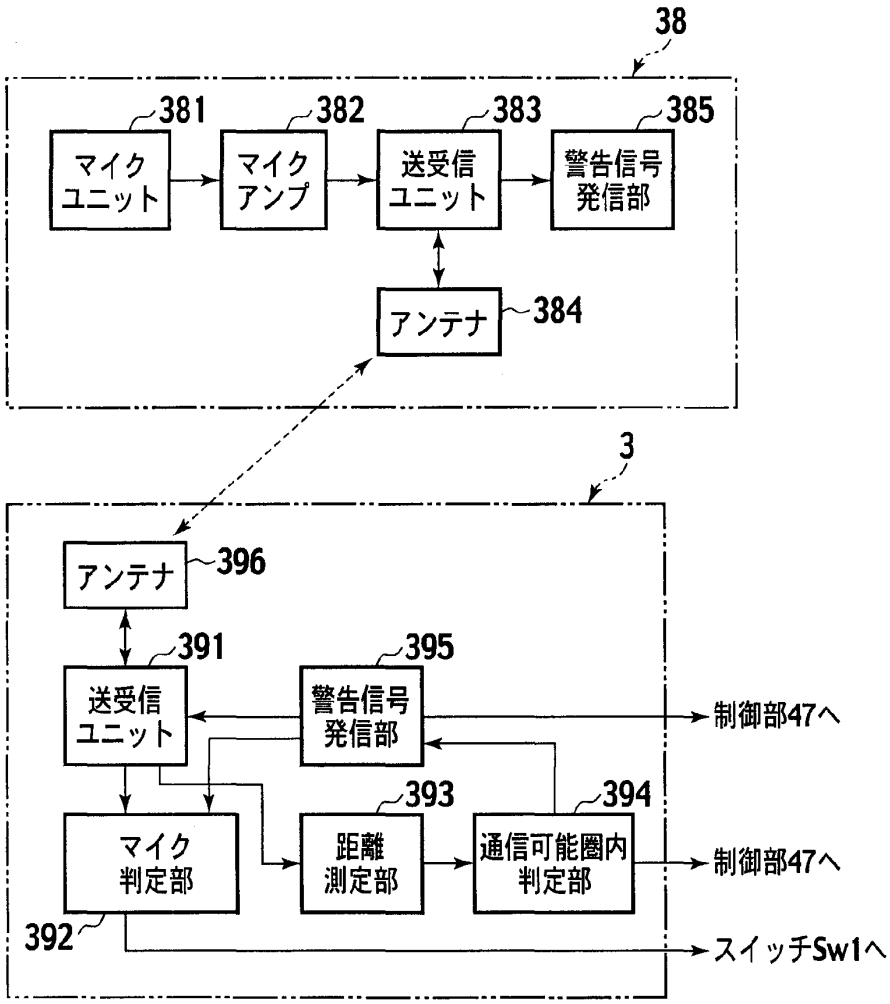
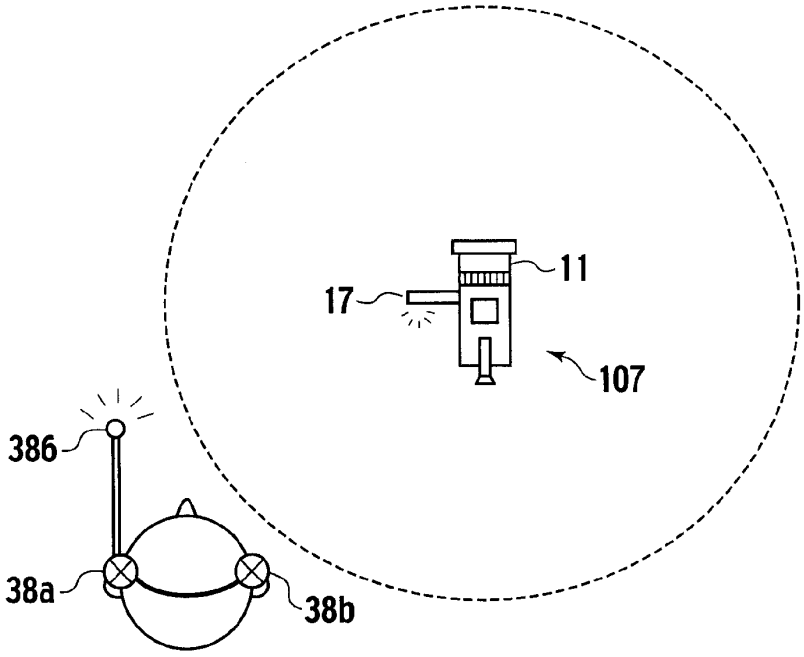


Exhibit 1004
Page 0588 of 1001

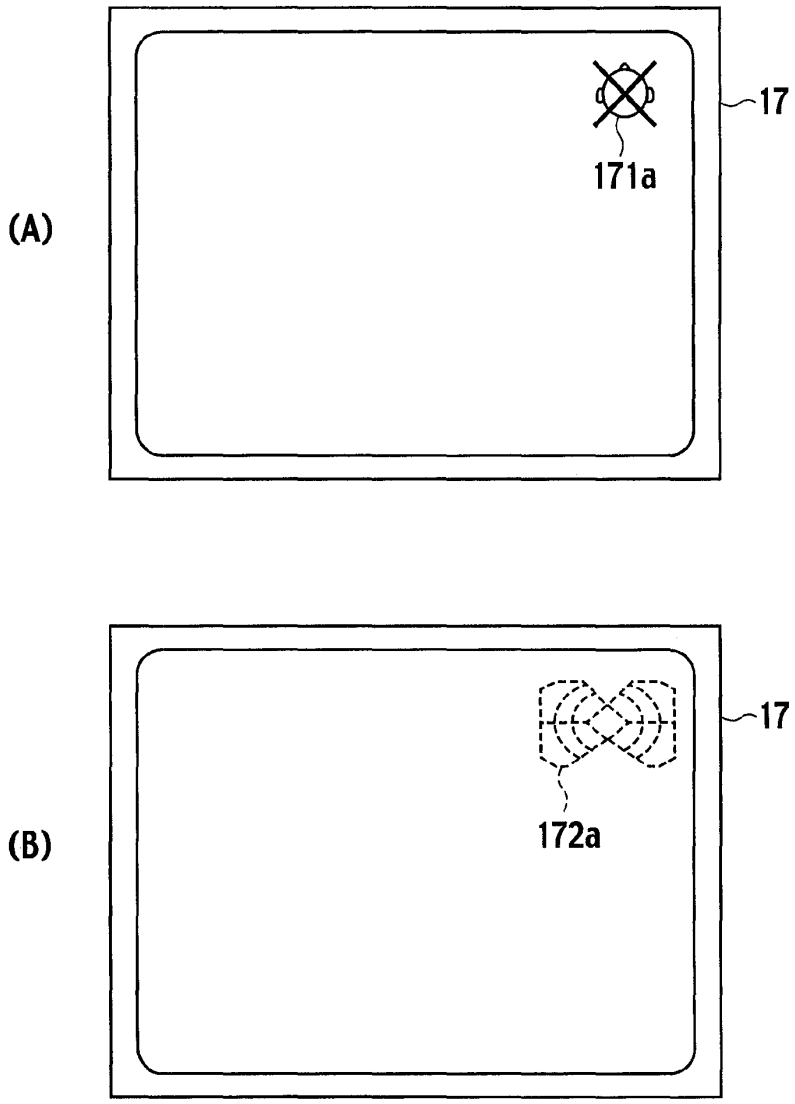
[図54]


38
381
マイク
ユニット
382
マイク
アンプ
383
送受信
ユニット
385
警告信号
発信部
アンテナ
384
3
アンテナ
396
391
送受信
ユニット
395
警告信号
発信部
制御部47へ
393
マイク
判定部
394
距離
測定部
通信可能圏内
判定部
制御部47へ
392
スイッチSw1へ

[図55]



[図56]



[図57]

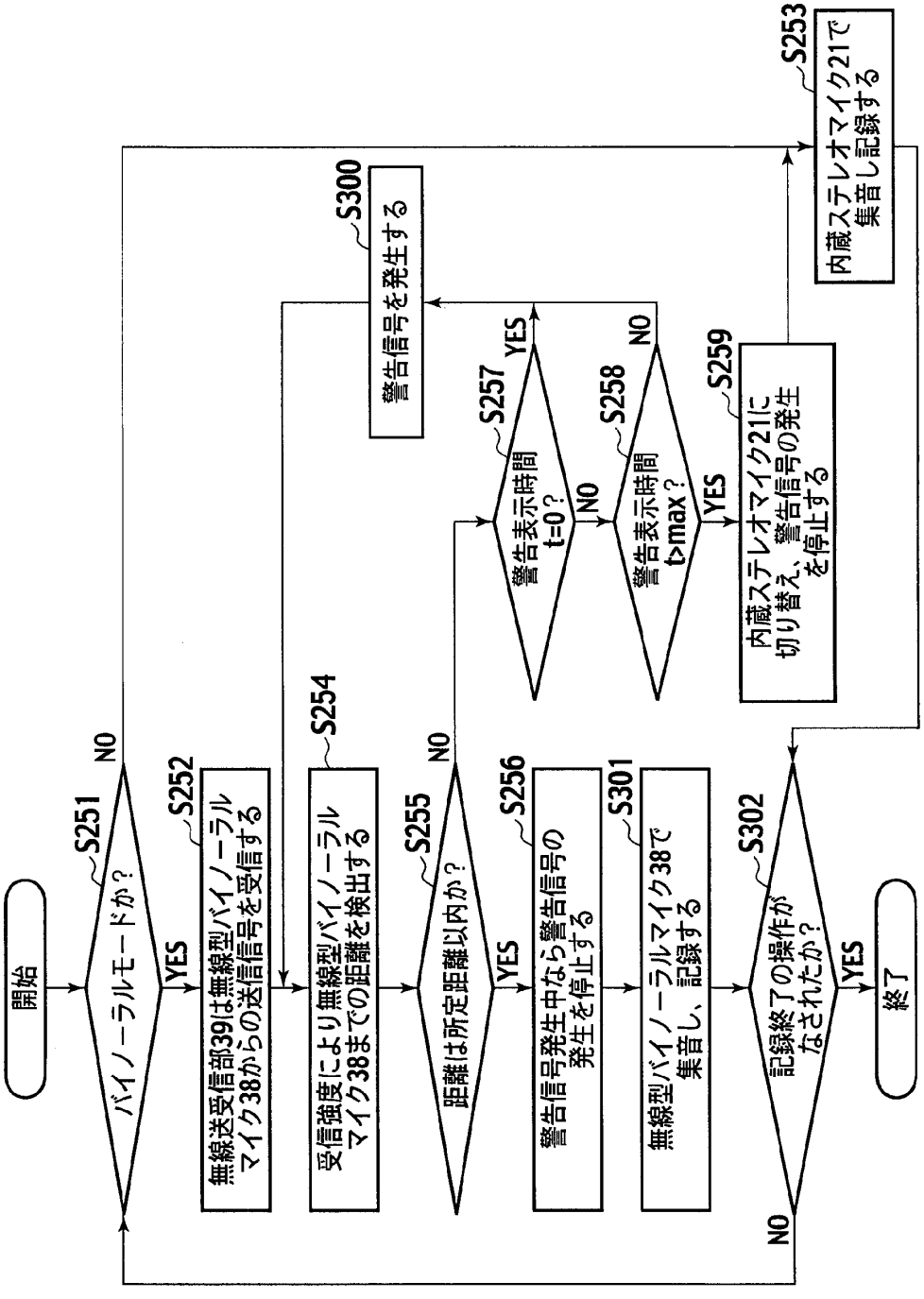


Exhibit 1004
Page 0592 of 1001

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2005/021247

A. CLASSIFICATION OF SUBJECT MATTER
H04S5/02(2006.01), ***H04S1/00***(2006.01), ***H04S7/00***(2006.01)

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)
H04S5/02(2006.01), ***H04S1/00***(2006.01), ***H04S7/00***(2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2006
Kokai Jitsuyo Shinan Koho 1971-2006 Toroku Jitsuyo Shinan Koho 1994-2006

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.
Y	JP 5-244683 A (Sony Corp.), 21 September, 1993 (21.09.93), Par. Nos. [0008] to [0013], [0028] to [0030] (Family: none)	1-8
Y	JP 11-296998 A (Pioneer Electronic Corp.), 29 October, 1999 (29.10.99), Par. Nos. [0272] to [0281] & EP 952578 A2 & EP 1104928 A2 & CN 1439960 A	1-8
Y	JP 2004-222131 A (Matsushita Electric Industrial Co., Ltd.), 05 August, 2004 (05.08.04), Par. Nos. [0012] to [0016] (Family: none)	1-6

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
03 February, 2006 (03.02.06)	21 February, 2006 (21.02.06)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

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

INTERNATIONAL SEARCH REPORT

international application No.
PCT/JP2005/021247

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	JP 7-312712 A (Sanyo Electric Co., Ltd.), 28 November, 1995 (28.11.95), Par. Nos. [0006] to [0008] (Family: none)	5 1-4,6-8
Y A	JP 7-336793 A (Matsushita Electric Industrial Co., Ltd.), 22 December, 1995 (22.12.95), Par. Nos. [0002] to [0009] (Family: none)	5 1-4,6
Y	JP 2002-165300 A (Sharp Corp.), 07 June, 2002 (07.06.02), Par. Nos. [0006] to [0012] (Family: none)	1-8
Y A	JP 10-70789 A (Central Research Laboratories Ltd.), 10 March, 1998 (10.03.98), Par. No. [0008] & US 5917916 A & GB 9610394 A & EP 808075 A	7-8 1-6

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

国際調査報告	国際出願番号 PCT／JP2005／021247

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. *H04S5/02(2006.01), H04S1/00(2006.01), H04S7/00(2006.01)*

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. *H04S5/02(2006.01), H04S1/00(2006.01), H04S7/00(2006.01)*

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922－1996年
日本国公開実用新案公報	1971－2006年
日本国実用新案登録公報	1996－2006年
日本国登録実用新案公報	1994－2006年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー＊	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Y	JP 5-244683 A（ソニー株式会社）1993.09.21, 【0008】-【0013】, 【0028】-【0030】（ファミリ無し）	1-8
Y	JP 11-296998 A（パイオニア株式会社）1999.10.29, 【0272】-【0281】 & EP 952578 A2 & EP 1104928 A2 & CN 1439960 A	1-8
Y	JP 2004-222131 A（松下電器産業株式会社）2004.08.05, 【0012】-【0016】（ファミリ無し）	1-6

☑ C欄の続きにも文献が列挙されている。　☐ パテントファミリーに関する別紙を参照。

＊ 引用文献のカテゴリー
「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
03．02．2006	21．02．2006

国際調査機関の名称及びあて先	特許庁審査官（権限のある職員）	5Z	9072
日本国特許庁（ISA／JP） 郵便番号100－8915 東京都千代田区霞が関三丁目4番3号	井出 和水 電話番号 03－3581－1101 内線 3541		

様式PCT／ISA／210（第2ページ）（2005年4月）

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求の範囲の番号
Y A	JP 7-312712 A（三洋電機株式会社）1995.11.28, 【0006】-【0008】 （ファミリ無し）	5 1-4,6-8
Y A	JP 7-336793 A（松下電器産業株式会社）1995.12.22, 【0002】-【0009】 （ファミリ無し）	5 1-4,6
Y	JP 2002-165300 A（シャープ株式会社）2002.06.07, 【0006】-【0012】 （ファミリ無し）	1-8
Y A	JP 10-70789 A（セントラル リサーチ ラボラトリーズ リミティド） 1998.03.10, 【0008】 & US 5917916 A & GB 9610394 A & EP 808075 A	7-8 1-6

Europäisches Patentamt
European Patent Office
Office européen des brevets

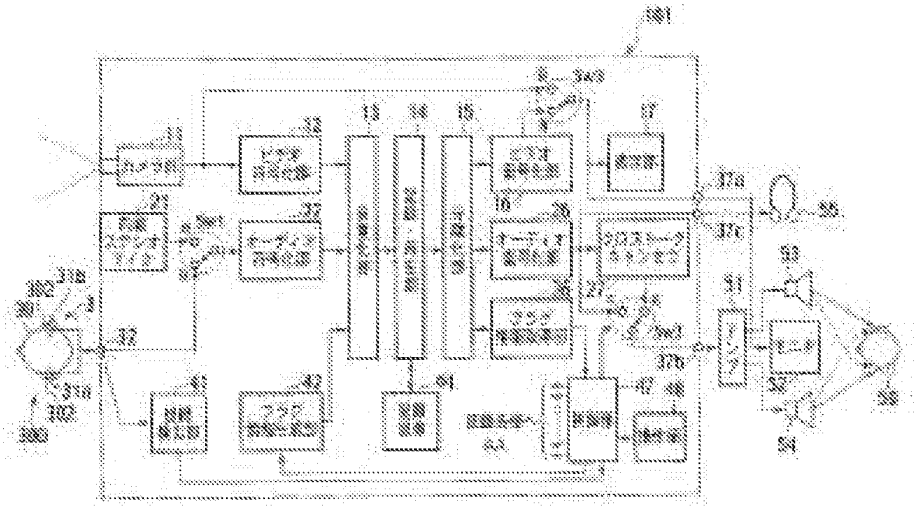
Patent Translate
Powered by EPO and Google

Notice

This translation is machine-generated. It cannot be guaranteed that it is intelligible, accurate, complete, reliable or fit for specific purposes. Critical decisions, such as commercially relevant or financial decisions, should not be based on machine-translation output.

ABSTRACT JPWO2006054698A1

The video / audio recording / playback device (101) includes a built-in stereo microphone (21a, 21b) and an external microphone connection terminal (32). A binaural microphone (3) to be worn on the ear of the photographer (300) is connected to the external microphone connection terminal (32). When the ambient sound is collected by the binaural microphone (3), the audio signal to be recorded on the recording medium is switched from the audio signal from the built-in stereo microphones (21a, 21b) to the binaural audio signal from the binaural microphone (3). The photographer (300) wears a binaural microphone (3 (31a.31b)) in his ear and collects ambient sounds of the photographer (300) including sounds emitted from the subject. The camera unit (11) captures a subject. The binaural audio signal, the captured video signal, and the binaural flag signal are recorded on the recording medium.



11 CAMERA PART
21 INTERNALLY MOUNTED STEREO MICROPHONES
41 CONNECTION DETECTING PART
12 VIDEO ENCODING PART
22 AUDIO ENCODING PART
42 FLAG INFORMATION GENERATING PART
13 MULTIPLEXING PART
14 RECORDING/REPRODUCING PART
44 RECORDING MEDIUM
15 SEPARATING PART
16 VIDEO DECODING PART
26 AUDIO DECODING PART
36 FLAG INFORMATION ACQUIRING PART
AA TO VARIOUS CIRCUIT COMPONENTS
17 DISPLAY PART
27 CROSSTALK CANCELLER
47 CONTROL PART
48 OPERATING PART
51 AMPLIFIER
52 MONITOR

WO2006054698A1 VIDEO/AUDIO RECORDING APPARATUS AND METHOD, AND VIDEO/AUDIO REPRODUCING APPARATUS AND METHOD

Data originating from sources other than the EPO may not be accurate, complete, or up to date.

The wording below is an initial machine translation of the original publication. To generate a version using the latest translation technology, go to the original language text and use Patent Translate.

Specification

Video / audio recording apparatus and method, audio / video reproduction apparatus and method Technical field

[0001] TECHNICAL FIELD The present invention relates to a video and audio recording apparatus and method for recording a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer including a sound emitted from a subject, and an image and sound recording apparatus A video and audio reproducing apparatus and method for reproducing a signal and an audio signal, and more particularly to a video and audio recording apparatus and method, and a video and audio reproducing apparatus and method capable of reproducing realistic sound with a captured image .

BACKGROUND ART

[0002] Video / audio recording / reproducing apparatuses (1, so-called video cameras) are widely used to record the image signal of the subject and the sound signal collected by the photographer's surroundings including the sound emitted from the subject. This type of video / audio recording / reproducing apparatus has a stereo mic for recording sounds in a stereo manner. In recent years, the video / audio recording / reproducing apparatus has been further miniaturized, and there is a problem that it is difficult to record a realistic sound in a stereo microphone mounted on a compact video / audio recording / reproducing apparatus. Therefore, there is a demand for a video / audio recording / reproducing apparatus capable of recording a more realistic sound.

[0003] In International Publication No. 96Z10884 pamphlet records the binaural collected sound field of the sound field around the photographer together with the image signal of the object picked up by disposing the ear structure on both sides of the main body of the video and audio recording and reproducing apparatus Audio / video recording / reproducing apparatus capable of recording / reproducing video / audio data.

[0004] In the video / audio recording / reproducing apparatus in which the binaural microphone disclosed in the above-mentioned document is attached, the lateral width of the main unit of the apparatus, that is, the distance between the left and right microphones is not equal to that of the human head, I can not record sensible sounds. In recent years, the main body of video / audio recording / reproducing apparatus which has been commercialized has been miniaturized in accordance with the advancement of high density recording technology, digital signal recording technology and video compression processing technology so that video / audio recording / reproducing Even if a microphone intended for binaural sound collection is mounted on the main body of the apparatus, the effect expected can not be obtained. Further, since the shape of the apparatus is largely different from the human head, it is considered difficult to obtain the effect disclosed in the above document.

Disclosure of invention

[0005] SUMMARY OF THE INVENTION The present invention has been made in view of the above problems, and it is an object of the present invention to provide a video and audio recording apparatus and method capable of reproducing realistic voices with images captured irrespective of the size and shape of the apparatus main body, It is an object of the present invention to provide a video and audio reproducing apparatus and method

[0006] A video and audio recording apparatus and method capable of reproducing a more realistic sound in conjunction with zooming-up of a subject when zooming up the subject, and a video and audio reproducing apparatus and method It is aimed to provide.

[0007] Furthermore, even in the case where an arbitrary photographer shoots and watches an arbitrary viewer, as in the case where the viewer who views the reproduction signal and the photographer who captured the binaural

sound are different from each other, It is an object of the present invention to provide a video and audio reproducing apparatus and method capable of listening to a sensitive voice.

[0008] In order to solve the problem of the above-described conventional technique, the present invention provides an image recording apparatus for recording an image signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer including a sound emitted from the subject A voice recording apparatus comprising: a camera unit for picking up an image of the object; and a microphone for picking up the ambient sound, the microphone other than a power binaural microphone using a binaural microphone mounted on the ear of the photographer is switched A switching unit, a video processing unit for processing the video signal outputted from the force melting unit, an audio processing unit for processing a sound signal outputted from a microphone for collecting the ambient sound, And when switching to the binaural microphone as the ambient sound, the sound mode for collecting the ambient sound is set as a binaural mode, A flag information generating section for generating a binaural flag signal indicating binaural mode, a flag information generating section for generating a binaural flag signal indicating binaural mode, a flag information generating section for generating a binaural flag signal indicating binaural mode, And a recording unit for recording the video and audio data on the recording medium.

[0009] According to the present invention, it is possible to reproduce a sound with realistic sensation together with a captured image irrespective of the size and shape of the apparatus main body. In addition, when photographing by zooming up the subject, it is possible to reproduce sound with more realistic interlocking with the zoom-up of the subject. Furthermore, even in the case where an arbitrary photographer shoots and watches an arbitrary viewer, as in the case where the viewer viewing the reproduction signal and the photographer who captures the binaural are different from each other, there is almost no sense of incompatibility I can listen to a sound with a sense

In this configuration, as the built-in microphone built in the video and audio recording apparatus, an external microphone connection terminal, and an external microphone connected to the external microphone connection terminal, a microphone other than the binaural microphone power binaural microphone , A connection detection section for detecting whether or not an external microphone is connected to the external microphone connection terminal, and a connection detection section for outputting, as an audio signal to be supplied to the audio processing section, A switch for switching between an audio signal output from said external microphone and an audio signal output from said external microphone, and said external microphone is set as a binaural microphone by said setting unit, and an external microphone is connected to said external microphone connection terminal by said connection detection unit And a control unit for setting the audio mode to the binaural mode when it is detected that it is connected Wherein the control unit performs switching control of the switch so that the switch supplies the audio signal output from the external microphone to the audio processing unit when the audio mode is the binaural mode, It is preferable that the generation unit controls the flag information generation unit to generate the binaural flag signal.

[0010] The display control unit includes a display unit for displaying the video signal output from the camera unit, and a display control unit for displaying a binaural mark indicating that the binaural mode is displayed on the display unit when the audio mode is the binaural mode .

[0011] Furthermore, the camera section has a zoom function of enlarging and capturing the subject, and includes a sound zoom processing section that amplifies the sound signal collected by the binalolic microphone according to the enlargement ratio in the camera section It is preferable.

[0012] Furthermore, the camera unit has a zoom function for enlarging and capturing the object, and the sound source of the sound signal collected by the binaural microphone is virtually brought close to the listener to form a virtual sound source A transfer function memory for storing a head related transfer function of the head sound source in association with a plurality of distances between the virtual sound source and the listener and a plurality of head related transfer functions stored in the transfer function memory, A convolution arithmetic unit for convoluting the head related transfer functions acquired by the function acquiring unit with respect to the sound signals collected by the binaural microphone, And a sound voice zoom processing unit having a sound processing unit.

[0013] In order to solve the problem of the above-described conventional technique, the present invention provides an image recording apparatus for recording an image signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer including a sound emitted from the subject An audio recording method comprising: an image capturing step of capturing an image of the object; and switching

whether to use a microphone other than a power binaural microphone that uses a binaural microphone to be worn by the photographer's ear as a microphone that collects the ambient sound A video processing step of processing a video signal obtained by imaging the subject; an audio processing step of processing an audio signal outputted from a microphone that collects the ambient sound; When switching to the binaural microphone as a microphone for collecting the ambient sound, A flag information generating step of generating a binaural flag signal indicating that the binaural mode is set as a binary mode as a binary mode; And a recording step of recording the image signal and the noisy ralph lag signal on a recording medium.

[0014] In order to solve the problem of the above-described conventional technique, the present invention is characterized in that a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of the photographer including the sound emitted from the subject force are recorded A reproducing unit for reproducing a recording signal recorded on the recording medium by a video / audio reproducing device for reproducing a recording medium; and a recording unit for reproducing the recording signal reproduced by the reproducing unit on the basis of the video signal and the audio signal A video processing unit for processing the video signal separated by the separating unit, an audio processing unit for processing the audio signal separated by the separating unit, When a binaural flag signal indicating that a binaural microphone to be worn by the photographer's ear is used as the microphone for collecting the ambient sound is recorded on the recording medium, the binaural flag A flag information acquiring unit that acquires a lag signal from the speaker, a crosstalk signal that is generated when the flag information acquiring unit acquires the binaural flag signal, when the sound signal processed by the audio processing unit is sounded by a speaker And a crosstalk canceller for processing the audio signal so as to cancel the speech signal, wherein the crosstalk canceller has a filter convolving filter characteristics obtained in advance with respect to the speech signal, Wherein the filter characteristic is a filter characteristic based on a head related transfer function measured using a sound signal collected by a pair of microphones mounted on the structure.

[0015] A video and audio reproduction method for reproducing a recording medium on which a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer including a sound emitted from the subject are recorded, A separating step of separating the video signal and the audio signal from the recording signal reproduced in the reproducing step; a separating step of separating the video signal and the audio signal from the recording signal reproduced in the reproducing step; An audio processing step of processing the audio signal separated by the separating step; an audio processing step of processing the audio signal separated by the separating step as a microphone for collecting the ambient sound on the recording medium; When a binaural flag signal indicating that a binaural microphone to be mounted on the camera is used is recorded, the binaural flag signal is taken A flag information acquisition step of acquiring a binaural flag signal at the flag information acquisition step, canceling a crosstalk signal generated when the sound signal processed by the audio processing step is sounded by a speaker And a crosstalk cancellation step of processing the audio signal so as to process the audio signal, wherein the crosstalk cancellation step is a step of convolving a filter characteristic obtained in advance with respect to the audio signal, Characterized in that the filter characteristic is a filter characteristic based on a head related transfer function measured by using a sound signal collected by a pair of microphones mounted on a type structure.

Brief Description of the Drawings

[0016] FIG. 1 is an external perspective view showing a video / audio recording / reproducing apparatus according to a first embodiment of the present invention.

FIG. 2 is a diagram showing a state in which a subject is photographed using the video / audio recording / reproducing apparatus according to the first embodiment of the present invention.

FIG. 3 is a block diagram showing an internal configuration example of a video / audio recording / reproducing apparatus according to a first embodiment of the present invention.

FIG. 4 is a diagram showing a display screen for performing initial setting of audio mode in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 5 is a diagram showing a display example of a binaural rakemake in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 6 is a diagram showing a modified example of the binaural rakemake used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 7 is a diagram showing a modified example of binaural rakemake used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

[FIG. 8] FIG. 8 is a diagram showing a modified example of a binaural mike used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 9 is a diagram showing an example of a description format of a binaural Ralph Lag signal in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

[FIG. 10] FIG. 10 is a diagram showing another example of the description format of the binauraflag signal in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

[FIG. 11] FIG. 11 is a diagram showing still another example of the description format of the binauraflag lag signal in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 12 is a flowchart for explaining a recording operation in the video / audio recording / reproducing apparatus according to the first embodiment of the present invention.

FIG. 13 is a flowchart for explaining a playback operation in the video / audio recording / playback apparatus according to the first embodiment of the present invention.

[FIG. 14] FIG. 14 is a block diagram showing a configuration example of a crosstalk canceler used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

[FIG. 15] FIG. 15 is a diagram showing a head related transfer function measuring device for obtaining a head related transfer function characteristic used in a crosstalk canceler used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 16 is a view showing a cylindrical structure to which a microphone unit used in the head related transfer function measuring apparatus shown in FIG. 15 is attached and a dummy head microphone for comparison. 17] FIG. 17 is a waveform diagram showing an impulse response waveform measured by the head related transfer function measuring apparatus shown in FIG. 15.

(18) FIG. 18 is a waveform diagram showing frequency characteristics measured by the head related transfer function measuring apparatus shown in FIG.

FIG. 19 is a waveform diagram showing an impulse response waveform measured using a dummy head microphone.

FIG. 20 is a waveform diagram showing frequency characteristics measured using a dummy head microphone

FIG. 21 is a diagram for explaining a crosstalk cancellation characteristic based on a filter characteristic based on a head related transfer function measured using a cylindrical structure having a microphone unit attached thereto.

FIG. 22 is a diagram for explaining crosstalk canceling characteristics by filter characteristics based on head related transfer functions measured using a dummy head microphone.

FIG. 23 is a diagram for explaining crosstalk canceling characteristics by filter characteristics based on a head related transfer function measured using a cylindrical structure having a microphone unit attached thereto.

[FIG. 24] FIG. 24 is a diagram for explaining crosstalk cancellation characteristics by filter characteristics based on head related transfer functions measured using a dummy head microphone.

FIG. 25 is a block diagram showing another configuration example of a crosstalk canceler used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

[FIG. 26] FIG. 26 is a block diagram showing still another configuration example of the crosstalk canceler used in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 27 is a flowchart showing a playback operation on a headphone in the video / audio recording / reproducing apparatus according to each embodiment of the present invention.

FIG. 28 is a block diagram showing an internal configuration example of a video / audio recording / reproducing apparatus according to a second embodiment of the present invention.

FIG. 29 is a block diagram showing a configuration example of a voice zoom processing unit in the video / audio recording / reproducing apparatus according to the second embodiment of the present invention. [FIG. 30] FIG. 30 is a flowchart for explaining the operation in the case of performing the voice zoom processing in the video / audio recording / reproducing apparatus according to the second embodiment of the present invention.

[FIG. 31] FIG. 31 is a block diagram showing another configuration example of a voice zoom processing unit in the video / audio recording / reproducing apparatus according to the second embodiment of the present invention.

32 is a diagram showing a head related transfer function measuring device for finding a head related transfer function used in another configuration example of the audio zoom processing unit in FIG. 31.

FIG. 33 is a cross-sectional view showing a dummy head microphone used in the head related transfer function measuring apparatus of FIG. 32.

[FIG. 34] FIG. 34 is a diagram showing the characteristics of a head related transfer function obtained by measuring with the head-related transfer function measuring apparatus of FIG. 32.

FIG. 35 is a diagram showing characteristics of a head related transfer function obtained by measuring with the head-related transfer function measuring apparatus of FIG. 32.

36 is a diagram showing the characteristics of the head-related transfer function obtained by measuring with the head-related transfer function measuring apparatus of FIG. 32. FIG.

37 is a diagram showing characteristics of a head-related transfer function obtained by measuring with the head-related transfer function measuring apparatus of FIG. 32. FIG.

38 is a diagram showing characteristics of a head related transfer function obtained by measuring with the head related transfer function measuring apparatus of FIG. 32.

39 is a diagram showing the characteristics of the head-related transfer function obtained by measuring with the head-related transfer function measuring apparatus of FIG. 32. FIG.

FIG. 40 is a block diagram showing the operation of the audio-video recording / reproducing apparatus according to the second embodiment of the present invention in the case of performing audio zoom processing using another configuration example of the audio zoom processing section shown in FIG. 31 7 is a flow chart for explaining.

[FIG. 41] FIG. 41 is a block diagram showing an internal configuration example of a video / audio recording / reproducing apparatus according to a third embodiment of the present invention.

FIG. 42 is a block diagram showing a configuration example of a voice zoom processing unit in the video / audio recording / reproducing apparatus according to the third embodiment of the present invention.

FIG. 43 is a block diagram showing another configuration example of a voice zoom processing unit in the video / audio recording / reproducing apparatus according to the third embodiment of the present invention.

FIG. 44 is a block diagram showing an internal configuration example of the video / audio recording / reproducing apparatus according to the fourth embodiment of the present invention.

FIG. 45 is a block diagram showing a configuration example of a voice zoom processing unit in the video / audio recording / reproducing apparatus according to the fourth embodiment of the present invention.

FIG. 46 is a flowchart for explaining the audio sound zoom process in the video / audio recording / reproducing apparatus according to the fourth embodiment of the present invention.

FIG. 47 is a block diagram showing an internal configuration example of a video / audio recording / reproducing apparatus according to a fifth embodiment of the present invention.

[FIG. 48] FIG. 48 is a block diagram showing a configuration example of a voice zoom processing unit in the video / audio recording / reproducing apparatus according to the fifth embodiment of the present invention.

49 is an external perspective view showing a video / audio recording / reproducing apparatus according to a sixth embodiment of the present invention. FIG.

FIG. 50 is a block diagram showing an internal configuration example of a video / audio recording / reproducing apparatus according to a sixth embodiment of the present invention.

51 is a plan view showing an example of the structure of a cord housing section in a video / audio recording / reproducing apparatus according to a sixth embodiment of the present invention. FIG.

FIG. 52 is an external perspective view showing a video / audio recording / reproducing apparatus according to a seventh embodiment of the present invention.

FIG. 53 is a block diagram showing an internal configuration example of the video / audio recording / reproducing apparatus according to the seventh embodiment of the present invention.

[FIG. 54] FIG. 54 is a block diagram showing a specific configuration example of a wireless binaural microphone and a radio transmission / reception unit in the video / audio recording / reproducing apparatus according to the seventh embodiment of the present invention.

FIG. 55 is a diagram for explaining a warning when the wireless binaural microphone crosses the communicable area in the video / audio recording / reproducing apparatus according to the seventh embodiment of the present invention.

56 is a view showing an example of a warning mark displayed on the display unit when the wireless binaural microphone exceeds the communicable area in the video / audio recording / reproducing apparatus according to the seventh embodiment of the present invention .

57 is a flowchart for explaining the operation of the video / audio recording / reproducing apparatus according to the seventh embodiment of the present invention.

BEST MODE FOR CARRYING OUT THE INVENTION

[0017] DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS Hereinafter, embodiments of the video and audio recording apparatus and method, the video and audio reproducing apparatus and method of the present invention will be described with reference to the drawings.

[0018] **⟨⟨** First embodiment **⟩⟩**

FIG. 1 is a perspective view showing an external configuration example of the video / audio recording / reproducing apparatus 101 according to the first embodiment.

[0019] The video / audio recording / reproducing apparatus 101 shown in FIG. 1 includes a camera section 11, a display section 17, internal stereo microphones 21 a and 21 b, and an external microphone connection terminal 32. An earphone type binaural microphone 3 having right and left microphones 31a and 31b of omnidirectionality is detachably connected to the external microphone connection terminal 32. [Here, a state where the binaural microphone 3 is connected to the external microphone connection terminal 32 is illustrated. The microphones 31a and 31b have diaphragms therein. As will be described in detail later, the video / audio recording / reproducing apparatus 101 selectively performs photography (sound recording) using the built-in

stereo microphones 21 a and 21 b and photography (binaural recording) using the binaural microphone 3 . Note that shooting means not only capturing an image of a subject but also capturing the image of the subject and collecting the ambient sound of the photographer including the sound emitted from the subject May occur.

[0020] FIG. 2 shows a state where the photographer 300 shoots a subject (not shown) using the video / audio recording / reproducing apparatus 101. When taking a subject while collecting sound using the binaural microphone 3, as shown in FIG. 2, the photographer 300 attaches left and right microphones 31 a and 31 b to the left and right ears 302. As a result, the ambient sound of the photographer 300 including the sound emitted from the subject with the binaural acoustic characteristics given by the positional relationship between the head 30 of the photographer 300 and the microphones 31a and 31b is collected. While viewing the monitor image of the object displayed on the display unit 17, the photographer 300 photographs the subject with the camera unit 11 and collects the surrounding sound with the binaural microphone 3. As will be described in detail later, the video signal picked up by the camera section 11 and the audio signal picked up by the binaural microphone 3 are recorded on a recording medium, not shown here. Also, as will be described in detail later, the video signal recorded on the recording medium is reproduced together with a voice having a realistic feeling as if the viewer is viewing in the same shooting environment as the photographer 300.

[0021] FIG. 3 is a block diagram showing a concrete internal configuration example of the video / audio recording / reproducing apparatus 101.

[0022] The video / audio recording / reproducing apparatus 101 includes a camera unit 11, a video encoding unit 12, a multiplexing unit 13, a recording / reproducing unit 14, a demultiplexing unit 15, a video decoding unit 16, a display unit 17, a built-in stereo microphone 21 An audio decoding unit 26, a crosstalk canceller 27, an external microphone connection terminal 32, a flag information acquisition unit 36, a video output terminal 37a, an audio output terminal 37b, a connection A detection unit 41, a flag information generation unit 42, a recording medium 44, a control unit 47, an operation unit 48, and switches Sw 1, Sw 2, Sw 3. The recording medium 44 may be a removable recording medium such as a disc-like recording medium or a tape cassette, and may be a recording medium such as a hard disc that is mounted in the video / audio recording / reproducing apparatus 101 in advance and is V Even if it is there! /,.

[0023] A monitor 52 such as a television receiver is connected to the video output terminal 37 a. Speakers 53 and 54 are connected to the audio output terminal 37 b via an amplifier 51. The sound produced by the speakers 53 and 54 is listened to by the viewer 59. In FIG. 3, for the sake of convenience, for the sake of convenience, it is shown that the photographer 300 and the viewer 59 are illustrated simultaneously. Normally, photographing by the photographer 300 and watching of the reproduced image / sound by the viewer 59 are separately performed Yes.

[0024] ⟨Recording Operation⟩

The recording operation in the video / audio recording / reproducing apparatus 101 will be described.

[0025] First, when the photographer 300 performs an operation to display the initial setting image (window) of the sound mode by the operation unit 48, the control unit 47 displays the initial setting image 170 as shown in FIG. 4 on the display unit 17 as an example . Both the binaural microphone 3 as described in FIGS. 1 and 2 and a normal external microphone can be connected to the external microphone connection terminal 32 as an external microphone. When the photographer 300 collects sounds using the noisel microphone 3 by a predetermined operation by the operation unit 48, it selects "binaural" as shown in FIG. 4, and collects sounds using a normal external microphone Select "Normal" if you want. The control unit 47 serves as a setting unit for setting as an external microphone connected to the external microphone connection terminal 32 whether to use a microphone other than the innovative microphone 3 or the binaural microphone. When "binaural" is selected as the external microphone input and the connection detecting section 41 detects that the plug of the external microphone is inserted in the external microphone connecting terminal 32, the control section 47 controls the video / audio recording / reproducing apparatus 101 controls each part of the circuit so as to perform a recording operation corresponding to photographing using the binaural microphone 3. The sound mode in which the binaural microphone 3 is used to sound ambient sounds and the collected sound signals are recorded is referred to as binaural mode. The ambient sound is collected using the built-in stereo microphone 21 or an ordinary external microphone, and the sound mode in which the collected sound signal is recorded is set to the normal mode.

[0026] If the plug of the binaural microphone 3 is formed exclusively for the binaural microphone 3 different from the usual external microphone and the external microphone connection terminal 32 is used as a connection terminal dedicated to the binaural microphone 3, the initial setting of the above-mentioned sound mode can be omitted can do.

[0027] Specifically, in FIG. 3, when detecting that the external microphone is connected to the external microphone connection terminal 32, the connection detection unit 41 supplies a detection signal to the control unit 47. When a detection signal indicating that an external microphone is connected is input while the external microphone is set to "binaural", the control unit 47 switches the switch Swl to a terminal a for inputting a sound signal from the built-in stereo microphone 21 To the terminal b for inputting the audio signal from the binaural microphone 3. As a result, the audio signal from the binaural microphone 3 is supplied to the audio coding section 22. The switch Swl serves as a switching unit for switching whether to use a microphone other than a force binaural microphone that uses a no noise mic to be attached to the ear 30 2 of the photographer 300 as a microphone for collecting ambient sounds.

[0028] In addition, the control unit 47 controls to generate and output flag information (binaural flag signal) indicating that the flag information generation unit 42 is in binaural mode. The binaural flag signal is input to the multiplexing unit 13.

[0029] When the binaural mode is set in this manner, the control unit 47 preferably displays a mark indicating that the binaural mode is set on the display unit 17. An example of the mark is shown in FIG. A mark 171 shown in FIG. 5A is a mark schematically showing the photographer 300 using the binaural microphone 3, and a mark 172 shown in FIG. 5B schematically shows a sphygmomanic reproduction of binaural sound It is a mark shown. As the mark indicating the binaural mode, any of the marks in FIGS. 5A and 5B may be used, but other marks may be used. If these marks are displayed on the display section 17 after being superimposed on the image photographed by the camera section 11 after the above-described initial setting or when the binaural microphone 3 is connected to the external microphone connection terminal 32 Yo With such a mark display, the photographer 300 can confirm whether the audio mode is the force set to the binaural mode when using the binaural microphone 3. When the audio mode is the binaural mode, the control unit 47 is a display control unit for displaying a binaural mark (marks 171, 172) indicating that it is in the binaural mode on the display unit 17.

[0030] The viewer 300 wearing the left and right microphones 3 la, 3 lb of the binaural microphone 3 on the left and right ears 302 images the subject with the camera unit 11. The video signal output from the camera unit 11 is input to the video encoding unit (video processing unit) 12 and the terminal g of the switch Sw 3. When photographing (recording) is being performed by the video / audio recording / reproducing apparatus 101, the switch Sw 3 is switched to the terminal g, the video signal from the camera section 11 is supplied to the display section 17, An image is displayed. At the same time, due to the positional relationship between the head 30 of the photographer 300 and the microphones 31a, 31b, the microphone 31a, 31 obtains a sound signal with a noisy sound whose subject is in the median plane direction. This audio signal is input to the audio coding section (audio processing section) 22 via the switch Swl.

[0031] In the case where the recording medium 44 is a DV cassette, the video encoding section 12 AZD converts the input video signal and encodes it according to the DV compression method to generate an encoded video signal. The audio coding section 22 AZD converts the input audio signal and rearranges the data positions of the uncompressed audio signal by shuffling to generate a coded audio signal.

[0032] The multiplexing unit 13 time division multiplexes the input coded image signal, the coded voice signal, and the binaural flag signal according to the signal format defined by the consumer digital VCR specification to generate a multiplexed signal . The multiplexed signal output from the multiplexing unit 13 is input to the recording / reproducing unit 14. The recording / reproducing unit 14 records the input multiplexed signal on the recording medium 44 in accordance with the recording format defined by the consumer digital VCR specification. The specific recording method of the Binotto-Ralf-Lag signal will be described in detail later. [0033] Here, a modified example of the binaural microphone 3 will be described.

[0034] ⟨Modified Example of Binaural Microphone 3⟩

FIG. 6 (A) shows a microphone 31 c as a first modification of the microphones 31 a and 31 b, and FIG. 6 (B) shows a microphone 31 d as a second modification of the microphones 31 a and 31 b. A microphone 31 c shown in FIG. 6 (A) is provided to be connected to an upper part of the microphone holding part 312 and a microphone holding part 312 to be inserted into the ear 302 of the photographer 300, and houses a microphone unit such as a

diaphragm And a microphone housing part 311 for making it! By making the binaural microphone 3 a microphone 31 c structured to be separated into the microphone storage part 311 and the microphone holding part 312, even if the photographer 300 mounts the binaural microphone 3, it is possible to listen to the sound arriving from the outside satisfactorily it can. The microphone 31 d shown in FIG. 6 (B) is provided with a microphone storage part 311 connected to the lower part of the microphone holding part 312. Even microphone 3 Id has the same effect as microphone 31 c.

[0035] FIG. 7 is a perspective view showing a specific configuration example of the microphone holding unit 312. Here, the microphone 31 c of FIG. 6 (A) in which the microphone housing portion 311 is provided on the top of the microphone holding portion 312 is shown. The microphone holding portion 312 shown in FIG. 7 (A) has a holding portion main body 312 a formed with a tapered sound passage hole 313 a whose diameter decreases toward the inside of the ear 302. The microphone holding portion 312 shown in FIG. 7 (B) is formed by forming the holding portion main body 312 a circular columnar sound hole 313 b. The holder main body 312a of FIG. 7 (A) is easy to insert into the ear 302 of the photographer 300, and the holding part main body 312b of FIG. 7 (B) has the feature that the attenuation of external sound when worn is small.

[0036] FIG. 8 shows an example in which the external shape of the microphone holding part 312 of the microphone 31 c in FIG. 6 (A) is made different. 8A shows a microphone holding portion 312 having a large outer shape, FIG. 8B shows a microphone holding portion 312 having an outer shape, and FIG. 8C shows a microphone holding portion 312 having a small outer shape. In this way, if a plurality of sizes are prepared as the microphone holding unit 312, the photographer 300 can select one matching the size of his ear 302. 8A to 8C, the shape and size of the microphone housing portion 311 (microphone unit) are the same, and the sensitivity and response frequency characteristics of the microphone unit are also the same.

[0037] ⟨Recording of binaural flag signal⟩

In order to distinguish between the binaural sound collected by the photographer 300 wearing the binaural microphone 3 and the stereo sound collected by the built-in stereo microphone 21, binaural sound and a binaural sound signal are recorded on the recording medium 44 at the time of binaural sound collection Record. The binno-lull-lag signal is generated by the flag information generating unit 42.

[0038] First, a specific recording method of the binaural flag signal when the recording medium 44 is a DV cassette will be described.

[0039] FIG. 9 shows a data format for recording audio data in the DV cassette.

0When recording audio data of ~ 89 bytes, the 0th and 1st bytes are synchronous codes, the 2nd to 4th by ID (identification) code, the 5th to 9th bytes are audio auxiliary data (AUX) 10 to 81 record audio data, and 82 to 89 record inner code parity data for error data detection and correction. For example, the flag information generating unit 42 outputs 1 as a binaural flag signal in binaural mode, and outputs 0 as long as it is not in binaural mode (in normal mode). The multiplexing unit 13 generates the data format signal shown in FIG.

[0040] Next, a specific recording method of the binaural flag signal in the case where the recording medium 44 is a disk for recording will be described. As a disk for recording, there are a disk for recording and reproducing using red laser light such as DVD-RAM, DVD-RW, DVD-R and a blue laser light such as Blu-ray Disc and HD-DVD There is a disk to record and reproduce. Here we will explain the case of superimposing the binauraflag lag signal using the DVD video standard that can be applied to these recording discs in general.

[0041] As a first method of superimposing the Noino-Ralph lag signal using the DVD video standard, a method of superimposing a binaural flag signal on a DVD-Video zone in the DVD video standard can be considered.

[0042] As shown in Fig. 10, the Volume space of the DVD standard is powerful with Volume and Film structure, DVD - Video zone and DVD others zone. The DVD-Video zone includes V MG (Video Manager) and VTS (Video Title Set) # 1 to DVTS # n. n is a predetermined integer of 2 or more. Each VTS includes control data (Control Data) and VOBS (Video Object Set). The VOBS includes a plurality of VOBs (Video Objects). The VOBS includes a plurality of CELLs. CELL includes a plurality of VOBU (Video Object Unit). VOBU includes navigation pack (NV_PACK), audio pack (A_PACK), video pack (V_PAC K). In the present embodiment, a data pack (D_PACK) including a binaural flag signal is provided in the VOBU.

[0043] The data pack (D_PACK) includes pack header, packet header, substream ID, audio frame information, audio data information, and noino ralph lag signal. Noino Ralph Lag signal will also be multiple audio frame layer power.

[0044] In this manner, information including a binaural flag signal in a format conforming to the DVD video standard is packed as a data pack (D_PACK) and MPEG multiplexed. This makes it possible to know which audio frame portion of the audio signal is a binalaural audio signal and which audio frame portion is a normal stereo sound while maintaining compatibility with the DVD video standard. Therefore, it is possible to easily determine in which audio frame portion the crosstalk cancellation process should be performed.

[0045] As a second method of superimposing the Noino-Ralph lag signal using the DVD video standard, a method of superimposing a binaural flag signal on DVD other zone in the DVD video standard can be considered. The DVD other zone is an area for recording auxiliary data related to data of the main body of video and audio, and it is also a user data recording area.

[0046] In the present embodiment, the data structure of the DVD other zone which is the user data recording area is structured similar to the data structure of the DVD-Video zone so as to be apportioned by comparing FIG. 10 with FIG. 11. As shown in FIG. 11, the DVD other zone also includes VMG, VTS, VOBS, VOB, CELL, and VOBU. In order to distinguish these pieces of information in the DVD other zone shown in FIG. 11 from that of FIG. 10, the name given with D is added to the head.

[0047] As shown in FIG. 11, the DVD other zone includes DVMG and DVTS # 1 to DVTS # n.

Each DVTS includes DVTSI (Video Title Set Information) and DVOBS. The DV OBS includes a plurality of DVOBs. DVOBS contains multiple DCELLs. DCELL includes a plurality of DVOBUs. The DVOBU includes a plurality of audio frame layers. The audio frame layer is an area for recording audio frame data such as coding parameters of audio signals, and a part of this audio frame layer is set as the recording area of the binaural flag signal.

[0048] By describing the binaural flag signal in the DVD other zone in the format conforming to the DVD video standard as described above, audio signals (binaural audio signal or normal stereo sound) included in the DVD-Video zone and binaural flag signal Are associated with each other. It is possible to know which audio frame portion of the audio signal is a binaural sound signal and which audio frame portion is a normal stereo sound while maintaining compatibility with the DVD video standard. Therefore, it is possible to easily determine in which audio frame portion the crosstalk cancellation process should be performed.

[0049] In the cases of FIGS. 10 and 11, for example, an audio signal picked up while photographing is started by operating the start button and the stop button is operated to end shooting is set as one or a plurality of audio frame layers, Add audio mode information to each audio frame layer. Then, the binaural flag signal is described in the voice mode information and managed as a binaural information packet. By managing the Noino Ralph lag signal as a binaural information packet, it is possible to easily acquire voice mode information for each audio frame. Thus, even when the binaural audio signal and the normal stereo audio signal are mixedly recorded at the time of reproducing the recording medium 44, the crosstalk canceller 27 Can be switched appropriately. Recording of the audio mode information is required to be performed at least for each operation of starting photographing, and it is more preferable to record at every predetermined time.

[0050] Also, even in the case of using, for example, a semiconductor memory as the recording medium 44, by defining the recording area of the binaural flag signal and specifying the sound mode of the sound signal to be recorded in the same manner as described above, it is possible to record binaural It is possible to identify the audio signal and reproduce the audio signal while appropriately switching the operation / non-operation of the crosstalk canceler 27.

[0051] It is also conceivable to describe the binaural flag signal in the user data of the multiplexing layer of the encoding method such as MPEG. As an example, it is assumed that a cellular phone having a communication function of video and audio is used, and a case where a video signal photographed together with a voice signal collected by the binaural microphone 3 from a cellular phone on the transmitting side is transmitted to the mobile phone on the receiving side . Even in such a case, the binaural flag signal can be transmitted from the transmitting-side cellular phone to the receiving-side cellular phone. By transmitting the binaural flag signal to the audio signal and transmitting it, it becomes possible to reproduce a sound with realism in the case of binaural sound. In this case, a binaural flag signal is described in a predetermined portion of the video and audio

packet data to be transmitted by the cellular phone and transmitted as packet data. For example, when using the transmission method specified by MPEG-4, it is possible to transmit a no-no-ruff-lag signal as shown in FIG. 9 in the user data recording area of the elementary stream. When a transport stream specified by the MPEG -4 standard is used, a binaural flag signal can be described in a private data area (private_data_type).

[0052] Also, when transmitting video data and audio data as fileed data in the form of attached files, binaural flag signals are written in the file header portion.

[0053] The recording operation in the video / audio recording / reproducing apparatus 101 will be described further using the flow chart shown in FIG.

[0054] First, in step S151, the control unit 47 determines whether the external microphone connected to the initial setting force external microphone connection terminal 32 described in FIG. 4 is set in the binaural microphone 3 or not. If it is determined in step S 151 that the initial setting is binaural (YES), the process proceeds to step S 152. If it is not determined that the binaural setting is made (NO), the control unit 47 switches the switch Swl to the terminal a. In step S154, the video / audio recording / reproducing apparatus 101 transmits the audio signal from the built-in stereo microphone 21 get. In step S152, the control unit 47 determines whether or not the connection detecting unit 41 is detecting the insertion of the plug of the external microphone into the external microphone connecting terminal 32. [In step S 152, if it is determined that an external microphone is connected to the external microphone connection terminal 32 (YES), the control unit 47 switches the switch Swl to the terminal b, and in step S153, the video / audio recording / reproducing apparatus 101 acquires an audio signal from the no-noal microphone 3. If it is not determined in step S 152 that an external microphone is connected to the external microphone connection terminal 32 (NO), the control unit 47 switches the switch Swl to the terminal a, and in step S 154, The reproduction apparatus 101 acquires a sound signal from the built-in stereo microphone 21.

[0055] Next, in step S155, the video signal from the camera section 11 is temporarily stored in a memory (not shown) in the video encoding section 12, and the audio signal from the binaural microphone 3 or the built-in stereo microphone 21 is temporarily stored in the audio encoding section In a memory (not shown) in the unit 22. At step SI 56, the video encoder 12 encodes the video signal, and the audio encoder 22 encodes the audio signal. In step S157, the coded video signal is temporarily stored in a buffer (not shown) in the video encoding section 12, and the coded audio signal is temporarily stored in a buffer (not shown) in the audio coding section 22. Then, Further, in step S158, the flag information generating unit 42 generates a binaural flag signal based on an instruction from the control unit 47 in the binaural mode.

[0056] Then, in step S159, the multiplexing unit 13 multiplexes the encoded video signal, the encoded audio signal, and the binaural flag signal, and in step S 160, generates the packetized stream signal. In step S 161, the recording / reproducing unit 14 records a stream signal in the medium recording 44. In step S162, the video encoder 12 and the audio encoder 22 determine whether there is a video signal and a sound signal to be encoded, and it is determined that there is still a video signal and a sound signal to be encoded (YES), the process returns to step S152, and the above operation is repeated. If it is not determined in step S162 that there is still a video signal and an audio signal to be coded (No), the process is terminated.

[0057] ⟨Playback Operation⟩

Returning to FIG. 3, the playback operation in the video / audio recording / playback apparatus 101 will be described. In FIG. 3, when a reproduction button (not shown) of the operation unit 48 is operated, the recording / reproduction unit 14 reproduces the multiplexed signal which is the recording signal recorded on the recording medium 44 under the control of the control unit 47. The multiplexed signal reproduced by the recording and reproducing unit 14 is supplied to a demultiplexing unit 15 which separates the input multiplexed signal into a coded video signal, .

[0058] The coded video signal is supplied to a video decoding section (video processing section) 16, the coded audio signal is supplied to an audio decoding section (audio processing section) 26, and the noino ralphag signal is transmitted as flag information And is supplied to the acquisition unit 36. The video decoding section 16 decodes the input coded image signal to obtain a video signal. In accordance with the operation of the reproduction button, the switch Sw 3 is switched to the terminal h under the control of the control unit 47. The video signal outputted from the video decoding section 16 is displayed on the display section 17 and also supplied to the monitor 52 via the video output terminal 37a for display. The audio decoding section 26 decodes

the input coded voice signal to obtain a voice signal. The audio signal is supplied to the crosstalk canceller 27 and the terminal c of the switch Sw 2.

[0059] The crosstalk canceler 27 generates a first tone signal that is generated when a binaural collected sound signal is reproduced by the speakers 53 and 54 and a signal sounded from the left speaker 54 is listened to by the right ear of the listener 59 Generates a signal for canceling the crosstalk component and the second crosstalk component that the signal generated by the right speaker 53 is listened to by the left ear of the listener 59 and adds it to the audio signal And generates a crosstalk processed signal. The flag information acquisition unit 36 holds the binaural flag signal obtained by the separation unit 15. The control unit 47 switches over the switch Sw 2 depending on whether the flag information acquisition unit 36 holds the binaural flag signal or not. That is, if the flag information acquiring unit 36 holds the binaural flag signal, if the switch Sw 2 is connected to the terminal d and the crosstalk canceling signal of the crosstalk canceller 27 power is also output to the sound output terminal 37 b If it is not controlled and the switch Sw 2 is connected to the terminal c and the audio signal which is not crosstalk processed from the audio decoding unit 26 is transmitted to the audio output terminal 37 b So as to output.

[0060] The sound signal output from the sound output terminal 37 b and amplified by the amplifier 51 is sounded from the left and right speech forces 53, 54. When the audio signal output from the audio output terminal 37 b is the crosstalk processed signal output from the crosstalk canceler 27, the viewer 59 views the video displayed on the monitor 52, and the photographer 300 captures an image It is possible to listen to the surrounding sound of the photographer 300 at the time of going as a sound with realism. At this time, since the crosstalk canceller 27 cancels the crosstalk component by using the head-related transfer function which will be described in detail later, even if the photographer 300 and the viewer 59 are different from each other, an arbitrary photographer 300 So that even when an arbitrary viewer 59 views it, he / she can listen to a sound with realistic feeling with little sense of discomfort.

[0061] The reproduction operation in the video / audio recording / reproducing apparatus 101 will be described further using the flow chart shown in FIG.

[0062] Referring to FIG. 13, in step S181, the recording and reproduction unit 14 reproduces the recording medium 44 and obtains a stream signal based on the multiplexed signal. In step S182, the recording and reproduction unit 14 decodes the stream signal to generate a packet signal. In step S183, the demultiplexing unit 15 separates the packet signal into a video signal, an audio signal, and a no-no-rural signal. In step S 184, the video decoding unit 16 decodes the video signal, and the audio decoding unit 26 decodes the audio signal. In step S185, the video decoding unit 16 and the audio decoding unit 26 temporarily store the decoded video signal and audio signal in a buffer (not shown). In step S186, the flag information acquiring unit 36 acquires the binaural flag signal.

[0063] Then, in step S187, on the basis of the binauraf luggage signal acquired by the flag information acquiring unit 36, the control unit 47 determines whether the reproduced audio signal is a force binaural audio signal which is a normal stereo audio signal judge. If it is determined in step S 187 that the signal is a binaural sound signal (YES), the process proceeds to step S 188. If it is determined in step S187 that the sound signal is not a binaural sound signal (NO), the process proceeds to step S 191, the control unit 47 switches the switch Sw 2 to the terminal c, and synchronizes the video signal and the audio signal .

[0064] If it is a binaural audio signal, in step S188, the control unit 47 switches the switch Sw2 to the terminal d, thereby setting the crosstalk cancellation processing by the crosstalk canceler 27 to an effective state. In step S189, the control unit 47 controls each unit of the circuit so as to synchronously reproduce the crosstalk canceled audio signal by the video signal and the crosstalk canceller 27. If it is determined in step S 190 that there is still a video signal and an audio signal to be reproduced (YES), the process returns to step S 182 and the above operation is repeated. If it is not determined that the video signal and the audio signal to be reproduced still exist in step S 190 (NO), the processing is terminated.

[0065] **〈クロストークキャンセル〉**

Here, the specific configuration and operation of the crosstalk canceler 27 will be described with reference to FIG. 14. As shown in FIG. 14, the crosstalk canceller 27 is configured to include finishers 272 a to 272 d, calo calculators 274 a and 274 b, and filters 275 a and 275 b.

[0066] In FIG. 14, the left channel signal P (t) of the binaural sound signal is filtered by the filter 2

L

72a, 272 b [This signal, right channel signal P (t) voice letter 272 c, 272 d 【This will be done. The filters 272a to 272d are sequentially assigned to the head related transfer functions h (t), h (t), h (t), h (t)

The filter characteristic (filter coefficient) created on the basis of rs lo ro is stored. The filters 272 a and 272 d have filter characteristics equivalent to the head related transfer functions h (t) and h (t), and the filters 272 b and 2

rs Is

72c has filter characteristics equivalent to those obtained by inverting the head related transfer functions h (t) and h (t). Feces

lo ro

For convenience, the filter characteristics of the filters 272a to 272d are expressed as h (t), - h (t), - h (t), h (t)

Write rs lo ro Is. Filters 272 a to 272 d are configured to add input signals P (t) and P (t)

L R

And outputs it.

[0067] The power calculator 274 a calculates the signal output from the filters 272 a and 272 c, and the filter 275 a gives a filter characteristic of d (t) to the added signal. Calo calculator 274 b adds the signals output from filters 272 b and 27 2 d, and filter 275 b gives a filter characteristic of d (t) to the added signal.

[0068] Here, the filter characteristic d (t) stored in the filters 275 a and 275 b is the following equation (1).

[0069] d (t) = {h (t) X h (t) - h (t) X h (t)} _ 1,

Is rs lo ro

The signals output from the filters 275a and 275b are crosstalk processed signals subjected to the crosstalk cancellation processing so that the crosstalk is canceled when sound is generated from the speakers 53 and. The crosstalk processed signal output from the filters 275 a and 275 b is amplified by the left channel amplifier 5 la and the right channel amplifier 5 lb of the amplifier 51, respectively, and is sounded from the speech forces 53 and 54.

[0070] A signal (voice) generated from the speaker 53 is listened to with the left ear of the viewer 59, and a part of the sounded signal is received as a first crosstalk signal indicated by a broken line at the right ear of the viewer 59 Listen. The crosstalk canceler 27 generates a first crosstalk cancellation signal for canceling the first crosstalk signal listened to by the right ear of the viewer 59 and causes the speaker 54 to generate sound. The first crosstalk cancellation signal cancels (attenuates) the first crosstalk signal by the first crosstalk cancellation signal. Similarly, a signal (voice) generated from the speaker 54 is listened to by the right ear of the viewer 59 and a part of the sounded signal is received as a second crosstalk signal indicated by a broken line on the left side of the viewer 59 Listen with your ears. The crosstalk canceller 27 generates a second crosstalk cancellation signal for canceling the second crosstalk signal to be listened to by the left ear of the viewer 59 and causes the speaker 53 to generate sound. The second crosstalk signal is canceled (attenuated) by the second crosstalk cancellation signal. As a result, the viewer 59 listens to the audio signal P (t) whose crosstalk is canceled with the left ear, the crosstalk is canceled with the right ear

1

Listens to the voice signal P (t).

[0071] ⟨Measurement of Head Related Transfer Function⟩

With reference to FIG. 15, a head related transfer function measuring device 6 for obtaining head related transfer function characteristics stored in the filters 272 a to 272 d, 275 a and 275 b will be described. As shown in FIG. 15, the head-related transfer function measuring device 6 includes a personal computer 61, an amplifier 62, a speaker 63, microphone units 65 a and 65 b, a cylindrical structure 65 e, and amplifiers 66 a and 66 b.

[0072] How to measure the head related transfer function!

[0073] First, a personal computer 61 generates a measurement signal composed of impulse sound, for example. The measurement signal is amplified by the amplifier 62. Measurement signals generated from the left speaker 63 are received by the left and right microphone units 65a and 65b. The right and left signals obtained by receiving the sound are amplified by the amplifiers 66 a and 66 b and input to the personal computer 61. This input signal is the head related transfer function h (t), h (t) of the sound received from the speaker 63 and received by the left and right microphone units 65a and 65b attached to the cylindrical structure 65e is there. Head trance

Is lo

The reaching function h (t) indicates that the signal generated from the left speaker 63 is the left microphone unit 65 a

Is

The head related transfer function h (t) is a characteristic to be listened to and received by the left speaker 63

lo

Is the characteristic of the crosstalk component listened to by the microphone unit 65 b on the right side.

[0074] Likewise, the measurement signal generated from the right speaker 64 is received by the left and right microphone units 65a and 65b. The right and left signals obtained by receiving the sound are amplified by the amplifiers 66 a and 66 b and input to the personal computer 61. The personal computer 61 compares the generated measurement signal with the received sound signal so that the sound generated from the speaker 64 is received by the left and right microphone units 65a and 65b attached to the cylindrical structure 65e Head transfer functions h (t) and h (t) are obtained. The head related transfer function h (t) is obtained from the right speaker 64

rs ro rs

The head related transfer function h (t) is a characteristic in which the sound generated by the right speaker 64 is a crosstalk component listened to by the left microphone unit 65a It is a characteristic.

[0075] The cylindrical structure 65 e will be described with reference to FIG. 16. 16A is a top view of the cylindrical structure 65e, FIG. 16B is a perspective view of the cylindrical structure 65e, and FIG. 16C is a cross-sectional view of a so-called dummy head microphone shown for comparison .

[0076] As shown in FIGS. 16 (A) and 16 (B), microphone units 65 a and 65 b are mounted on the surface of the cylindrical structure 65 e in a state of being separated by 180 °. As shown in the drawing, the microphone units 65a and 65b do not have an auricle and external ear canal. The diaphragm (not shown) of the microphone units 65a, 65b is disposed at a position substantially coincident with the surface of the cylindrical structure 65e. On the other hand, the dummy head microphone 69 shown in FIG. 16 (C) has pinna shaped members 692 a, 692 b and ear canals 693 a, 693 b on both sides of the artificial head 691, and a microphone unit 694 a, and 694 b, respectively. The microphone units 694 a and 694 b are provided at positions corresponding to the position of the eardrum of a person, and collect sound signals approximate to sounds to be heard by a person.

[0077] The sound receiving characteristics of the microphone units 65a and 65b attached to the cylindrical structure 65e shown in FIGS. 16 (A) and 16 (B) are influenced by characteristic differences given by auricles and external auditory canals different in size and shape from person to person And it is a characteristic that can measure the head related transfer function. Although sound waves generated from the speakers 63, 64 are shielded by the cylindrical structure 65e, they are diffracted along the cylindrical structure 65e and reach the microphone units 65a, 65b. The microphones 65a and 65b measure the characteristics of sound waves directly

emitted from the speakers 63 and 64 and reaching by diffracting along the cylindrical structure 65e. By using the cylindrical structure 65e, it is possible to obtain a head related transfer function having an average head shielding characteristic. Therefore, even when various viewers who have different sizes and shapes of the heads and have different head shielding characteristics listen to the binaural sound signal, it is possible to receive speech with realistic feeling with almost no discomfort.

[0078] 17A to 17D are diagrams showing the head related transfer function h (t (t) of the cylindrical structure 65e measured by the acoustic signal transfer characteristic measuring device 6, the impulse sound generated by the acoustic signal transfer characteristic measuring device 6 described above,), h (t), h (t), h (t) are convoluted with each other

Is lo rs ro

. FIG. 17 (D) shows the filter characteristic d (t) shown in the above formula (1). In Figs. 17 (A) - 17 (E), the vertical axis shows the amplitude of the signal voltage normalized with the prescribed output voltage, and the horizontal axis shows the number of samples when the measured signal is sampled at 48 kHz The time shown is shown.

[0079] 18 (A) to 18 (E) show the frequency characteristics obtained by Fourier analysis of the signals shown in FIGS. 17 (A) to 17 (E). In FIGS. 18 (A) to (E), the frequency positions of 100 Hz, 1 kHz, and 10 kHz are indicated by the broken vertical lines, and the gain difference between the horizontal lines of the two broken lines in the response characteristics of the vertical axis is 10 dB.

[0080] In the filters 272a to 272d of FIG. 14, the head related transfer functions h (t)

rs

, h (t), h (t), h (t) are given. As mentioned above, Phil lo ro Is

Filter characteristics corresponding to the head related transfer functions h (t) and h (t) are given to the tapes 272 a and 272 d

rs Is

, Filter characteristics corresponding to the head related transfer functions h (t), h (t) whose polarity is reversed - h (t) and - lo ro lo h (t) are given to the filters 272 b and 272 c.

ro

[0081] For comparison, the characteristics measured using the dummy head microphone 69 shown in FIG. 16 (C) instead of the microphone units 65a and 65b attached to the cylindrical structure 65e are shown in FIGS. 19 and 20. FIG. 19 shows characteristics obtained by measurement similarly to FIG. 17. As can be understood by comparing FIG. 17 with FIG. 19, the impulse response waveform measured by using the microphone units 65 a, 65 b attached to the cylindrical structure 65 e is the impulse response measured by using the dummy head microphone 69 This waveform is closer to the impulse sound which is the input signal for measurement than the waveform.

[0082] FIG. 20 shows the frequency response characteristics measured using the dummy head microphone 69. The characteristics obtained by using the microphone units 65a and 65b attached to the cylindrical structure 65e are smaller than the distortion force S of the frequency characteristic and close to flat . Respective characteristics shown in FIGS. 20 (A) to 20 (E) are enhanced or attenuated in the response characteristics at a frequency of 1.5 to 7 kHz. The characteristics shown in FIGS. 18 (A) to 18 (E) are smaller in degree of enhancement and attenuation of response characteristics. This is because there is no disturbance in characteristics caused by the auricle or the external auditory meatus when the microphone units 65a and 65b attached with the cylindrical structure 65e are used. In the dummy head microphone 69, a part of the sound waves emitted from the speakers 63, 64 is reflected by the auricle, and the reflected sound wave and the directly arriving sound wave are synthesized and synthesized in phase and synthesized in reverse phase It attenuates. In addition, it is strengthened or attenuated at a specific frequency due to the influence of external ear canal resonance, antiresonance, etc. By using the microphone units 65 a and 65 b attached to the cylindrical structure 65 e, it is possible to suppress adverse effects as in the case of using the dummy head microphone 69.

[0083] Filter characteristics based on the head related transfer function measured by using the microphone units 65a and 65b attached to the cylindrical structure 65e are added to the filters 272a to 272d and the filters 275a and 275b of the crosstalk canceler 27) And a filter characteristic based on the head-related transfer function measured using the dummy head microphone 69 (second condition), and a comparative viewing test was conducted by a plurality of viewers . A small, small microphone was inserted into the ear canal of the viewer, and the sound receiving characteristic was measured as the sound received by the small microphone is the sound heard by the viewer.

[0084] FIG. 21 shows the characteristics measured by a certain viewer under the first condition. In FIG. 21, (A) shows an impulse signal on the left side input signal P (t), a silent input signal P (t) on the right side

L R

This is the waveform of the impulse response signal which is sounded from the speakers 53 and 54 as the number and received by the compact microphone worn by the viewer's left ear. (B) is the waveform of the crosstalk component sounded with the same condition as (A) and received by a compact microphone worn by the viewer's right ear. In FIG. 21 (A), a large impulse response waveform is obtained, and in FIG. 21 (B), a small level waveform is obtained. FIG. 21 (C) shows the frequency analysis result of the response waveform, Ca is the response characteristic obtained by frequency analysis of the response waveform of (A), and Cb is the response characteristic obtained by frequency analysis of the response waveform of (B). Crosstalk cancellation effect exceeding 20 dB at 100 Hz to 2 kHz is obtained.

[0085] Further, in FIG. 21, (D) shows the case where the left input signal P (t) is no signal, the right input signal P

L

(t) as an impulse signal from the speakers 53 and 54, and worn on the left ear of the viewer

R

It is the waveform of the crosstalk component received by the small microphone. (E) is the waveform of the impulse response received by a compact microphone worn on the right ear of the viewer, which is pronounced under the same condition as (D). In FIG. 21 (D), a waveform of a small level is obtained, and in FIG. 21 (E) a large level of impulse response waveform is obtained. FIG. 21 (F) is a frequency analysis result of the response waveform, F d is a response characteristic obtained by frequency analysis of the response waveform of (D), and Fe is a response characteristic obtained by frequency analysis of the response waveform of (E) . A crosstalk cancellation effect of about 16 dB is obtained at 100 Hz to 2 kHz. FIG. 22 shows the characteristics measured by the same viewer as in FIG. 21 under the second condition. The measurement conditions are the same as in FIG. The crosstalk canceling effect of FIG. 22 (C) is about 14 dB, and the crosstalk canceling effect of FIG. 22 (F) is about l ldB. Under the second condition, the effect is inferior compared to the first condition.

[0087] Fig. 23 shows the characteristics measured with the same measurement conditions as in Fig. 21 under the first condition with different viewers from Fig. 21 and Fig. 22. The crosstalk canceling effect of FIG. 23C is about 22 dB, and the crosstalk canceling effect of FIG. 23F is about 18 dB. Even different viewers have good effects.

[0088] FIG. 24 shows the characteristics measured under the same conditions as in FIG. 22 under the second condition with the same viewers as in FIG. 23. The crosstalk canceling effect of FIG. 24 (C) is about 14 dB, and the crosstalk canceling effect of FIG. 24 (F) is about 10 dB. Even for different viewers, it can be seen that the effect is inferior to the first condition under the second condition. Furthermore, similar measurements were performed by different viewers, and it was confirmed that the same effect as above can be obtained for each of the first condition and the second condition.

[0089] From the above measurement results, the filter characteristics given to the filters 272 a - 272 d and the filters 275 a, 275 b of the crosstalk canceller 27 are more improved than the filter characteristics based on the head related transfer function measured by using the dummy head microphone 69 Direction of filter characteristics based on head related transfer functions measured using microphones 65a and 65b attached to form structure 65e Crosstalk component sounded from the left speaker and listened to by the right ear and right

side It is f * i ¾ that the effect of canceling the crosstalk component sounded by the speaker of the left ear listened to is high.

[0090] In the filter characteristics based on the head related transfer function measured using the microphone units 65 a and 65 b attached to the cylindrical structure 65 e, disturbance of the frequency characteristics in the high frequency range is small. That is, by using the cylindrical structure 65e, it is possible to reduce the frequency characteristic being greatly reduced or increased at a specific frequency portion, and it is possible to minimize the degradation of the sound quality as much as possible. As a result, the viewer can listen to the sound with realistic feeling with little sense of incongruity.

[0091] Filter characteristics based on head related transfer functions measured using the microphone units 65a and 65b attached to the cylindrical structure 65e with filter characteristics given to the filters 272a to 272d and the filters 275a and 275b of the crosstalk canceller 27 , The crosstalk cancellation is performed near the entrance of the ear canal of the viewer 59 which is the structure when listening to the binaural sound signal. Therefore, the cancellation of the crosstalk component is effectively performed for a plurality of viewers 59 having different acoustic characteristics of the auricle and the ear canal.

[0092] The cylindrical structure 65e is not limited to a complete cylindrical shape, but may be a slightly cylindrical deformed cylindrical shape! /,. It is preferable not to have a concave-convex portion that gives a change in response characteristics such as the auricle or the external auditory canal, etc. In a state where the microphone units 65 a, 65 b are attached to the cylindrical structure 65 e, the response characteristics are disturbed As much as possible, it is preferable,

[0093] The crosstalk canceler 27 is not limited to the configuration shown in FIG. 14, but may be a band division type crosstalk canceller which can further reduce the opposite phase feeling occurring in the low frequency band. The band division type crosstalk canceller divides a binaural sound signal input as a signal of the entire frequency band into a signal of a low frequency band and a signal of a mid high frequency band and converts the binaural sound signal into a binaural sound signal of a middle to high frequency band Crosstalk cancellation processing is performed only for the crosstalk.

[0094] FIG. 25 shows a band division type crosstalk canceler 27a, and its configuration and operation will be described. Constituent parts having the same functions as those of the crosstalk canceller 27 shown in FIG. 14 are denoted by the same reference numerals, and description thereof is omitted.

[0095] As shown in FIG. 25, the crosstalk canceller 27a is different from the crosstalk canceler 27 shown in FIG. 14 in that it includes low-pass filters (LPF) 271a and 271d, high-pass filters (HPF) 271b and 271c, delay units 273a and 273b, Gain control amplifiers (GC) 276 a - 276 d, calo calculators 2 77 a, 277 b Different in that they are being caught!

[0096] Among the binaural audio signals input to the crosstalk canceller 27 a, the left channel signal P (t) is input to the LPF 271 a and the HPF 271 b, the right channel signal P (t) is input to the LP

L R

F 271 d and HPF 271 c, and is divided into a low frequency band and a middle high frequency band. : The cut-off frequency of the LPFs 271a, 271d and the HPFs 271b, 271c is set to about 100-200 Hz. The middle and high frequency band signals output from the HPFs 271 b and 271 c are circuit portions composed of the finer switches 272 a to 272 d configured in the same manner as the crosstalk canceler 27, the calorific calculators 274 a and 274 b, and the focus switches 275 a and 275 b Crosstalk cancellation processing is performed. Signals in the middle to high frequency band subjected to the crosstalk cancellation processing are input to the gain control amplifiers 276 b and 276 c, and the gain is adjusted.

[0098] The signals in the low frequency band output from the LPFs 271a and 271d are input to the delay units 273a and 273b and are delayed by substantially the same time as the time required for the crosstalk cancellation process for the signals in the middle to high frequency band. The signals in the low frequency band output from the delay units 273 a and 273 b are input to the gain control amplifiers 276 a and 276 d, and the gain is adjusted so as to eliminate the level difference from the signals in the middle to high frequency band.

[0099] The power calculators 277 a and 277 b add the low frequency band signal output from the gain control amplifiers 276 a to 276 d and the middle high frequency band signal. The signals output from the adders 277a

and 277b are crosstalk processed signals subjected to crosstalk cancellation processing only for signals in the middle to high frequency band. The crosstalk processed signal output from the adders 277 a and 277 b is amplified by the left channel amplifier 5 la and the right channel amplifier 51 b of the amplifier 51, respectively, and is sounded from the speakers 53 and 54.

[0100] In the configuration shown in FIG. 25, crosstalk cancellation processing is not applied to signals in the low frequency band, so that it is reproduced as a signal having no antiphase feeling in the low frequency band.

[0101] In the crosstalk canceler 27 shown in FIG. 14, the crosstalk cancellation effect by the crosstalk canceller 27 is sufficiently obtained at less than or equal to 10 Hz as described above with reference to FIGS. 21 to 24. The low frequency band below lOOHz has less influence on the localization of the sound source than the V band. Without canceling the crosstalk, the signal is audibly listened to with a sense of incompatibility as a signal of opposite phase.

[0102] In the crosstalk canceller 27a shown in FIG. 25, crosstalk cancellation is not performed in a low frequency band of 100 to 200 Hz or less so as to realize a crosstalk canceller such that a reverse phase signal does not occur in a low frequency band Do it.

[0103] FIG. 26 shows a frequency band division type crosstalk canceler 27 b having a filter construction method different from that of FIG. 25, and its configuration and operation will be described. Constituent parts having the same functions as those of the crosstalk canceller 27 a shown in FIG. 25 are denoted by the same reference numerals, and description thereof is omitted.

[0104] The crosstalk canceller 27 b shown in FIG. 26 differs from the crosstalk canceler 27 a in FIG. 25 in that filters 278 a and 278 b and filters 279 a and 279 b are provided in place of the filters 272 a to 272 d and the filters 275 a and 275 b The way of wiring is also different. The crosstalk canceller 27 a is configured with a filter characteristic as a feedforward FIR (finite impulse response) type, whereas the crosstalk canceller 27 b uses a feedback type FIR type filter characteristic .

[0105] In FIG. 26, cross-track cancellation processing is performed by the FIR type filters 278 a, 278 b, 279 a, 279 b and adders 274 a, 274 b for signals in the middle to high frequency band output from the HPFs 271 b, 271 c. The filter characteristics obtained by the head-related transfer function measuring device 6 are shown in the illustration in FIG. 7 in the winding positions 278 a, 278 b, 279 a, 279 b. The winding patterns 278 a, 278 b, 279 a, 279 b The respective filter characteristics are convoluted with the input signal and output. According to the crosstalk canceller 27 b, when the discomfort is reduced as if the reverse phase signal does not occur in the low frequency band, the same action and effect as the crosstalk canceller 27 a can be obtained. The crosstalk canceller 27 b shown in FIG. 26 can reduce the order of the filter used than the crosstalk canceller 27 a shown in FIG. 25, and the configuration can be simplified. In place of FIR type filter, IIR type (infinite impulse response) infinite impulse response filter may be used.

[0106] In the configuration shown in FIG. 3, when the force control unit 47 having the crosstalk canceler 27 (or 27 a, 27 b) configured separately from the control unit 47 is a microprocessor having a DSP (Digital signal processor) , The control unit 47 may execute the functions of the crosstalk canceler 27, 27 a, 27 b. In addition, the crosstalk canceler 27, 27 a, 27 b is not limited to hardware, and can be realized by software.

[0107] ⟨Headphone playback⟩

Meanwhile, in the video / audio recording / reproducing apparatus 101 shown in FIG. 3, the audio signal outputted from the audio decoding unit 26 can be listened to even by headphones. When listening to the nooraur audio signal with headphones, the above-mentioned crosstalk components do not occur V,. When listening to the crosstalk processed signal processed by the crosstalk canceller 27 with headphones, the opposite phase component of the binaural sound signal is listened to by the left and right ears. Since the reverse phase component is an acoustic signal component not in nature and it is not preferable for listening, crosstalk cancellation processing is not performed when listening to binaural sound signals even with headphones.

[0108] Therefore, as shown in FIG. 3, the audio signal outputted from the audio decoding section 26 is outputted to the audio output terminal 37 c without going through the crosstalk canceler 27. The audio signal output from the audio output terminal 37 c is supplied to the headphone 55. As described above, the viewer 59 can listen to the crosstalk processed signal outputted from the crosstalk canceler 27 by the speech strength 53, 54, or the

audience 59 can listen to the sound signal not processed by the crosstalk canceler 27 as the headphone 55 You can also listen to.

[0109] A reproduction procedure of the binaural audio signal at the headphone 55 will be described with reference to FIG. 27. The same processes as those in the flowchart shown in FIG. 13 are denoted by the same reference numerals, and description thereof is omitted.

[0110] In FIG. 27, steps S181 to S186 are as described with reference to FIG. 13, and in step S192 it is determined whether or not reproduction is performed by the headphone 55. FIG. It is sufficient to detect by a connection detecting unit whether or not a plug is inserted into the sound output terminal 37c which is a connection terminal of the headphone 55. FIG. If it is determined in step S192 that the headphone is to be reproduced (YES), in step S193, a binaural audio signal which is not subjected to crosstalk cancellation processing in synchronization with the video signal is reproduced by the headphone 55, and in step S190 Transfer. Similarly, when the reproduced audio signal is not binaural audio signal but normal stereo sound, the audio signal outputted from the audio decoding section 26 may be similarly supplied to the headphone 55.

[0111] If it is not determined in step S192 that the headphone is to be reproduced (NO), the process of steps S187 to S190 is performed as in FIG. Unlike the reproduction process by the headphone 55 in step S193, the process in step S189 is a process of reproducing the binaural audio signal by the speakers 53,

[0112] ⟨⟨ Second Embodiment ⟩⟩

The photographer 300 attaches the binaural microphone 3 to the left and right ears 302 and picks up the subject and records on the recording medium 44 while collecting the sound, so that the viewer 59 can recognize the entirety when the photographer 300 shoots You can listen to the ambient sound in the direction. However, an image shot by a normal video / audio recording / reproducing apparatus (video camera) is an image in the range of about 60 degrees in front of the camera. In the case of zoom photographing, photographing is performed with a narrower viewing angle. Therefore, in the case of performing zoom shooting, it is desirable to enhance the sound produced in the vicinity of the zoomed up subject and collect the sound. In the second embodiment, in addition to the zoom-up of the subject, the sound in the vicinity of the subject is enhanced, and collected and recorded.

[0113] FIG. 28 shows the video / audio recording / reproducing apparatus 102 provided with the audio processing unit of audio which is the second embodiment, and its configuration and operation will be described. Components having the same functions as those of the video / audio recording / reproducing apparatus 101 of the first embodiment shown in FIG. 3 are denoted by the same reference numerals, and description thereof is omitted. The video / audio recording / reproducing apparatus 102 is different from the video / audio recording / reproducing apparatus 101 in that it includes an audio zoom processing section 33. In FIG. 28, illustration of the audio output terminal 37 c which is the connection terminal of the headphone 55 and the headphone 55 is omitted.

[0114] In FIG. 28, the audio signal from the binaural microphone 3 input via the external microphone connection terminal 32 is input to the audio zoom processing unit 33. The camera unit 11 includes a plurality of lenses (not shown), and has a zooming function of zooming up and down the subject by moving one or a plurality of lenses and varying the distance between the lenses. When a zoom-up operation is performed by the operation unit 48, a zoom-up control signal is input from the control unit 47 to the camera unit 11, the subject is zoomed in and imaged, and the zoom-up control signal is subjected to audio zoom processing The audio signal inputted and input to the unit 33 is subjected to audio zoom processing.

[0115] On the basis of the zoom-up control signal, the audio zoom processing unit 33 augments the audio signal collected on the median plane of the photographer 300 including a lot of audio signals to be emitted from the binaural sound power of the subject, and performs zoom-in Thereby generating an audio signal. The zoom-up audio signal is input to the audio coding section 22 via the switch Swl. The image signal of the subject picked up by zooming up is encoded by the video encoding section 12 and the zoom up audio signal is encoded by the audio encoding section 22 and is recorded on the recording medium 44 as in the first embodiment It is recorded. FIG. 29 is a specific configuration example of the audio zoom processing unit 33. As shown in FIG. 29, the audio zoom processing unit 33 includes a zoom magnification detector 331, a coefficient calculator 332, an adder 335, a variable amplifier 337, and adders 338 a and 338 b.

[0117] In FIG. 29, the zoom magnification detector 331 detects the zoom magnification based on the zoom-up control signal supplied from the control unit 47. The coefficient calculator 332 calculates a coefficient a representing the degree of enhancement of the sound emitted from the vicinity of the subject according to the detected zoom magnification. The adder 355 adds the binaural sound signals of the right and left channels input from the external microphone connection terminal 32. The variable amplifier 337 amplifies the output signal of the adder 335 according to the coefficient α from the coefficient calculator 332. The adders 338 a and 338 b add the left and right channel binaural sound signals to the output signal of the variable amplifier 337, respectively. The diaphragms in the microphones 31a and 31b are substantially parallel in a state in which the microphones 31a and 31b of the binaural microphone 3 are attached to the left and right ears 302 of the photographer 300. Since the sound from the left and right direction of the photographer 300 sometimes includes components having phases opposite to each other, as a result of the addition of the left and right channels in the adder 335, the sound from the left and right directions is partially canceled out and attenuated. Accordingly, from the audio zoom processing unit 33, a zoom-up audio signal in which the sound collected on the median plane of the photographer 300 is enhanced is output.

[0118] With reference to FIG. 30, the operation of the video / audio recording / reproducing apparatus 102 including the operation in the audio zoom processing unit 33 will be further described. 30, in step S201, the adder 335 adds binaural sound signals of left and right channels from the binaural microphone 3 to obtain an addition signal S. In step S undefined 202, the zoom magnification detector 331 detects the zoom magnification obtained by the control unit 47 according to the operation of the operation unit 48. The zoom magnification may be obtained from the relationship between the applied voltage to the motor for driving the lens of the camera unit 1 1 and the driving time. In step S203, the coefficient calculator 332 calculates a coefficient oc representing the enhancement degree according to the zoom magnification. In step S 204, the variable amplifier 337 multiplies the output signal of the adder 335 by the coefficient α to obtain a S. In step S 205, the calorie calculators 338 a and 338 b add the binaural sound signals of the left and right chambers and the output signal (S) of the variable amplifier 337. Then, in step S 206, the zoom-up audio signal is recorded on the recording medium 44. In step S207, the control unit 47 determines whether the recording has ended. If it has not ended (NO), the control unit 47 repeats the force of step S201. If recording has ended in step S207 (YES), the processing in the audio zoom processing unit 33 is ended.

[0119] FIG. 31 shows an audio zoom processing unit 33 a showing another configuration example of the audio zoom processing unit 33. If a head related transfer function that gives the effect that the sound source approaches the listener with respect to the sound signal of the median plane is convoluted, the sound source also receives a power in conjunction with the zoom-up of the subject in the camera unit 11 It is felt as if it approached the listener, and it is possible to receive and listen to the audio signal having the realistic feeling. The audio zoom processing unit 33 a shown in FIG. 31 is configured to convolve a head related transfer function that provides an effect of bringing the sound source closer to the sound signal of the median plane.

[0120] The audio zoom processing unit 33 a shown in FIG. 31 differs from the audio zoom processing unit 33 of FIG. 29 in that a function acquiring unit 333, a transfer function memory 334, and a convolution operator 336 are added.

[0121] In FIG. 31, a transfer function memory 334 stores a head related transfer function for forming a virtual sound source in which a sound source is virtually positioned in close proximity. The head related transfer function is a function to obtain the listening characteristics of the sound emitted from the virtual sound source and the listening characteristics are determined by the distance between the virtual sound source and the listener.

[0122] The function acquiring unit 333 acquires from the transfer function memory 334 a head related transfer function corresponding to a sound localization position assumed according to the coefficient oc calculated by the coefficient calculator 332. Note that the coefficient α in FIG. 29 and the coefficient α in FIG. 31 and the other figures are not necessarily the same. For the sake of convenience, the same reference numerals are attached. The convolution arithmetic unit 336 convolves the head related transfer function acquired by the function acquiring unit 333 with the addition signal of the binaural sound generated by the adder 335. The variable amplifier 337 amplifies the added signal convoluted with the head-related transfer function according to the coefficient α obtained by the coefficient calculator 33 2. Calo calculators 338 a and 338 b add the left and right channel's neural audio signals and the variable amplifier 337 output signals. Here, it is sufficiently effective only to locate the force sound source having the variable amplifier 337 virtually in close proximity, and it is also possible to omit the variable amplifier 337. The coefficient α for selecting the head-related transfer function by the function acquiring unit 333 and the coefficient determining the degree of amplification in the

variable amplifier 337 may be set as different coefficients. [0123] With reference to FIG. 32, A method of measuring the head related transfer function for forming a virtual sound source will be explained.

[0124] The head-related transfer function measuring apparatus 6 a shown in FIG. 32 is configured to include a personal computer 61, an amplifier 62, a speaker 63, amplifiers 66 a and 66 b, and a dummy head microphone **68. Dummy one head microphone 68 í, artificial head 681** 【Install the microphone units 684 a, 684 b! The cranial transfer function measuring device 6a uses a dummy head microphone 68 instead of the microphone units 65a and 65b attached to the cylindrical structure 65e, as compared with the head related transfer function measuring device 6 shown in FIG. 15 Except that the point and the speaker are arranged on the median plane of the dummy head microphone 68 as only one of the left and right speakers 63 and 64 (in this case, the speaker 63 is used)

[0125] FIG. 33 shows a sectional view of the dummy head microphone 68. In the dummy head microphone 68, micro switches 684 a and 684 b are disposed in the vicinity of entrances of the auditory canals 683 a and 683 b of the artificial head 681 having the pinna shaped members 682 a and 682 b. The dummy head microphone 69 shown in FIG. 16 (C) is obtained by disposing the microphone units 694 a, 694 b at positions corresponding to the position of the eardrum of the person in the back of the ear canals 693 a, 693 b, and the dummy head microphone 68 , The dummy head microphone 69 is different from the dummy head microphone 69 in that the micro units 684 a and 684 b are arranged near the entrances of the auditory canals 683 a and 683 b. Normally, the dummy head microphone refers to one in which the microphone units 694 a and 694 b are arranged at positions corresponding to the position of the eardrum of the person at the back of the ear canals 693 a and 693 b as shown in FIG. 16 (C) 33 in which the microphone units 684a, 684b are disposed in the vicinity of the population of the auditory canals 683a, 683b of the artificial head 681 having the 1S auricular shapes 682a, 682b is also referred to as a dummy head microphone for convenience.

[0126] By using the dummy head microphone 68, it is possible to collect sound from the speaker 63 as binaural sound excluding the influence by the auditory canal 683 a, 683 b.

[0127] Returning to FIG. 32, the personal computer 61 generates a measurement signal composed of impulse sound, for example. The measurement signal is amplified by the amplifier 62. Measurement signals generated from the speaker 63 are received by the left and right microphone units 684 a, 684 b of the dummy head microphone 68. The left and right signals obtained by receiving the sound are amplified by the amplifiers 66a and 66b and input to the personal computer 6 1. The personal computer 61 compares the generated measurement signal with the received signal to determine the head related transfer functions h (t), h (t) of the dummy head microphone 68. Head

1 r

The part transfer function h (t) is the head related transfer obtained by the sound receiving signal at the left microphone unit 684a

1

The function, head related transfer function h (t) is a head related transfer function obtained from a sound receiving signal at the right microphone unit 684 b. The distance D between the speaker 63 and the dummy head microphone 68 is set to, for example, 0.5 m, 1 m, 2 m, etc., and head related transfer functions at each distance D are sequentially found.

[0128] 34 to 39 show characteristics of the head related transfer function obtained by measuring with the head-related transfer function measuring device 6 a shown in FIG. 32.

[0129] The impulse response waveform shown in FIG. 34 (A) is a waveform received by the left microphone unit 684 a when the distance D between the speaker 63 and the dummy head microphone 68 is 50 cm. The vertical axis is the normalized amplitude (voltage). The horizontal axis is time, and it is indicated by the number of sample points of the signal whose sampling frequency is 48 kHz. FIG. 34 (B) is a frequency response characteristic obtained by Fourier analysis of the impulse response waveform shown in FIG. 34 (A) by the personal computer 61. The horizontal axis is the frequency (Hz), and the vertical axis is the response characteristic.

[0130] FIG. 35 (A) shows the impulse response waveform received by the right microphone unit 684 b when the distance D is 50 cm. FIG. 35 (B) is a frequency response characteristic obtained by Fourier analysis of the impulse response waveform shown in FIG. 35 (A). The measurement conditions are the same as in FIG. 34.

[0131] Similarly, FIG. 36 (A) shows an impulse response waveform received by the left microphone unit 684 a when the distance D is 1 m, and FIG. 36 (B) shows its frequency response characteristic.

[0132] FIG. 37 (A) shows the impulse response waveform received by the right microphone unit 684 b when the distance D is 1 m, and FIG. 37 (B) its frequency response characteristic.

[0133] FIG. 38 (A) shows the impulse response waveform received by the left microphone unit 684 a when the distance D is 2 m, and FIG. 38 (B) its frequency response characteristic.

[0134] FIG. 39 (A) shows the impulse response waveform received by the right microphone unit 684 b when the distance D is 2 m, and FIG. 39 (B) its frequency response characteristic.

[0135] Comparing these characteristics, the amplitude of the waveform decreases as the impulse response waveform shown in (A) of FIGS. 34 to 39 increases the value of the distance D to 0.5 m, 1 m, 2 m. In the frequency response characteristic shown in (B) of FIG. 34 to FIG. 39, when the distance D is 0.5 m, regular concavo-convex characteristics at intervals of about 400 Hz in the frequency portion of lkHz to 4 kHz enclosed by the ellipse of a broken line Dip), and when the distance D is lm, it becomes somewhat irregular asperity characteristics, and when the distance D is 2 m, it is a combination of a plurality of concavo-convex characteristics with different frequency intervals. In addition, when the distances D are the same, substantially the same characteristics are obtained on the left and right.

[0136] The personal computer 61 compares the measurement signal of the generated impulse sound with the waveform of the impulse response signal input from the amplifiers 66a and 66b to obtain head related transfer characteristics for each distance D. The head related transfer characteristic obtained for each distance D is a characteristic for causing the listener to listen to the sound signal of the virtual sound source force by locating the sound source virtually at the position of the distance D. In this embodiment, the distances D are set to 0.5 m, 1 m, and 2 m, but even if the respective distances D may be determined to be narrower than 0.5 m even if the respective characteristics are obtained at more distances Good.

[0137] The head related transfer characteristics obtained as described above are stored in the transfer function memory 334 of FIG. Which of the stored transfer functions is to be used for the zoom magnification detected by the zoom magnification detector 331 is determined as an example by setting the distance to the subject measured **by the automatic focus measuring function Use the coefficient α obtained by dividing by the zoom** magnification. For example, when the distance to the subject is 10 m and the zoom magnification is 5, the **coefficient is 2. When the distance to the subject is 10 m and the zoom magnification is 10, the coefficient α is 1, and when the zoom magnification is 20, the coefficient α is 0.5.**

[0138] With reference to FIG. 40, the operation of the video / audio recording / reproducing apparatus 102 including the operation at the audio zoom processing unit 33 a will be further described. 40, in step S211, the calorifier 335 adds binaural sound signals of left and right channels from the binaural microphone 3 to obtain a calorie calculation signal S. In step S undefined 212, the zoom magnification detector 331 detects the zoom magnification obtained by the control unit 47 according to the operation of the operation unit 48. In step S213, the coefficient calculator 332 determines which one of the plurality of transfer functions stored in the transfer **function memory 334 is selected according to the zoom magnification, and determines the coefficient α** representing the degree of enhancement in the variable amplifier 337 . As the coefficient OC, the number obtained by dividing the distance to the subject by the zoom magnification may be used as it is, or the value generated based on the number obtained by dividing the distance to the subject by the zoom magnification Even if it is! /,.

[0139] In step S 214, the function acquiring unit 333 acquires a transfer function from the transfer function memory 334 according to the coefficient oc, and the convolution operator 336 convolutes the transfer function with the addition signal output from the adder 335. In step S215, the variable amplifier 337 multiplies the output **signal of the convolution arithmetic unit 336 by the coefficient α to enhance it. In step S 216, the adders 338 a** and 338 b add the binaural sound signals of the left and right channels and the output signal of the variable amplifier 337. Then, in step S 216, the zoom-up audio signal is recorded on the recording medium 44. In step S218, the control unit 47 determines the force at which recording has ended, and if not completed (NO), repeats

the force of step S211. If the recording has ended (YES) in step S218, the processing in the audio zoom processing section 33a is ended.

[0140] ⟨⟨ Third embodiment ⟩⟩

In the second embodiment, the zoom-up process of audio is performed on the recording side, but in the third embodiment, the zoom-up process of audio is performed on the reproduction side. In the video / audio recording / reproducing apparatus 103 according to the third embodiment shown in FIG. 41, the same reference numerals are given to components having the same functions as the video / audio recording / reproducing apparatus 101 of the first embodiment shown in FIG. 3 And its explanation will be omitted. Compared with the video / audio recording / reproducing apparatus 101, the video / audio recording / reproducing apparatus 103 has the audio zoom processing section 33 b at the subsequent stage of the separation section 15, and the zoom magnification detector 331 is provided at the front stage of the multiplexing section 13 The difference is that it is. In FIG. 41, illustration of the audio output terminal 37 c which is a connection terminal of the headphone 55 and the headphone 55 is omitted.

[0141] The operation of the video / audio recording / reproducing apparatus 103 will be described. The operation unit 48 is operated, and the lens drive signal of the camera unit 11 generated by the control unit 47 is supplied to the camera unit 11 and the zoom magnification detector 331. The zoom magnification detector 331 detects the zoom magnification by analyzing the zoom direction, the zoom speed, and the lens driving time of the lens driving signal. The zoom magnification information indicating the detected zoom magnification is supplied to the multiplexing unit 13. The multiplexing unit 13 also multiplexes the zoom magnification information together with the encoded video signal, the encoded audio signal, and the binaural flag signal. The recording / reproducing unit 14 records the multiplexed signal including the zoom magnification information on the recording medium 44.

[0142] The recording / reproducing unit 14 reproduces the multiplexed signal recorded on the recording medium 44, and the demultiplexing unit 15 separates the zoom magnification information together with the coded video signal, the encoded voice signal and the binaural flag signal . The zoom magnification information is input to the audio zoom processing unit 33 b.

[0143] FIG. 42 is a specific configuration example of the audio zoom processing unit 33 b. As shown in FIG. 42, the audio zoom processing unit 33 b differs from the audio zoom processing unit 33 of FIG. 29 in that the zoom magnification detector 33 1 is eliminated and the signals input to the adders 338 a and 338 b are But is different from the output signal of the crosstalk canceller 27.

[0144] In FIG. 42, the coefficient calculator 332 calculates a coefficient α used when the input signal is amplified by the variable amplifier 337 using the zoom magnification information separated and input by the demultiplexer 15. The adder 335 adds the noisy speech signal input from the audio decoding unit 26. The variable amplifier 337 amplifies the output signal from the adder 335 according to the coefficient α input from the coefficient calculator 332. The adders 338 a and 338 b add the output signal from the crosstalk canceller 27 and the amplified signal from the variable amplifier 337.

[0145] Meanwhile, there are cases where zoom processing in the camera unit 11 is executed by computation processing by software using, for example, DSP. When the audio zoom processing is performed at the time of reproduction, it is unnecessary to secure the signal processing time of the DSP necessary for the zoom processing. Therefore, it is possible to sufficiently perform signal processing such as optimization processing of a video signal photographed by a DSP at the time of recording, encoding of a video signal, recording control or the like. By performing the audio zoom processing at the time of reproduction, it is possible to avoid the possibility of occurrence of a problem that the calculation man-hours at the DSP are assigned to the calculation processing for the zoom processing and the operation time for recording is insufficient Show.

[0146] 43 shows an audio zoom processing unit 33 c configured to convolve a head related transfer function that gives the proximity effect to the audio signal of the median plane to the audio zoom processing unit 33 b of FIG. 42 ing. The audio zoom processing unit 33 c differs from the audio zoom processing unit 33 b in that a function acquiring unit 333, a transfer function memory 334, and a convolution operator 336 are added. The functions of the function acquiring unit 333, the transfer function memory 334, and the convolution computing unit 336 are the same as those in FIG. 31, and description thereof is omitted.

[0147] ⟨⟨ Fourth Embodiment ⟩⟩

The video / audio recording / reproducing apparatus 104 according to the fourth embodiment shown in FIG. 44 is configured so that the audio zoom processing is performed manually by an external operation at the time of reproducing the recording medium 44. That is, when the zoom magnification information is not recorded on the recording medium 44, the viewer 59 performs the zoom up processing of the audio while watching the video signal reproduced by the motor 52. The zoom-up process of the audio manually performed by the viewer 59 is referred to as manual audio zoom process.

[0148] Compared with the video / audio recording / reproducing apparatus 103, the video / audio recording / reproducing apparatus 104 shown in FIG. 44 omits the zoom magnification detector 331, has the audio zoom processing section 33 d instead of the audio zoom processing section 33 b, Different in that! /.

[0149] When the viewer 59 instructs the operation unit 48 to perform manual audio zooming, the control unit 47 inputs the zoom-up control signal to the audio zoom processing unit 33 d. The audio zoom processing unit 33 d performs zoom-up processing on the binaural audio signal decoded by the audio decoding unit 26 according to the zoom-up control signal.

[0150] FIG. 45 is a specific configuration example of the audio zoom processing unit 33 d. As shown in FIG. 45, the audio zoom processing unit 33 d is provided with a zoom magnification detector 331 a to which the zoom-up control signal from the control unit 47 is input, in comparison with the zoom processing unit 33 b in FIG. 42, and coefficient calculation But differs in that it inputs the zoom magnification information generated by the zoom magnification detector 33 la instead of the zoom magnification information from the separation section 15 to the unit 332. The other parts are the same operations as those of the zoom processing section 33 b, and the description thereof will be omitted.

[0151] The manual audio zoom process in the fourth embodiment will be further described with reference to FIG. 46. In FIG. 46, in step S221, the adder 335 adds the right and left signals of the binalolar audio signal obtained by reproduction and obtains the addition signal S. In step S undefined 222, the control unit 47 determines whether or not the audio zoom magnification is varied by the operation unit 48. If the audio zoom magnification is varied in step S222 (YES), the process goes to step S223, and if not changed (NO), the process goes to step S226.

[0152] When the audio zoom magnification is varied, the zoom magnification detector 33 la calculates the zoom magnification based on the zoom-up control signal in step S223. In step S224, the coefficient calculator 332 calculates a coefficient (X) based on the zoom magnification input from the zoom magnification detector 331a, and the coefficient (X is a head for locating the sound source in front of the viewer But may also include the transfer function characteristic In step S225, the value of the coefficient α is updated to the newly calculated value.

[0153] In step S226, the variable amplifier 337 multiplies the added signal S by the coefficient α to generate a S. In the case where the steps S223 to S225 have not passed, the coefficient α is a value before the audio zoom magnification is varied. In step S227, the adders 338a and 338b add the signal "S" to the binaural sound signal subjected to the crosstalk cancellation processing by the crosstalk canceller 27. Then, In step S228, the audio signal obtained in step S227 is output via the switch Sw2 and the audio output terminal 37b. In step S229, the control unit 47 determines whether or not the reproduction has ended. If it is not completed (NO), the control unit 47 returns to step S22 1, and if it is completed (YES), ends the processing.

[0154] ⟨⟨ Fifth Embodiment ⟩⟩

The video / audio recording / reproducing apparatus 105 of the fifth embodiment shown in FIG. 47 shows a configuration suitable for listening to the zoom-up audio signal with the headphone 55. The video / audio recording / reproducing apparatus 105 shown in FIG. 47 is different from the video / audio recording / reproducing apparatus 103 shown in FIG. 41 in that it has an audio zoom processing section 33 e instead of the audio zoom processing section 33 b, and the audio / 33 e are supplied to the headphone 55 via the audio output terminal 37 c. The crosstalk cancellation processing is not performed on the headphone 55 by the crosstalk canceller 27, and the binaural audio signal subjected to the zoom-up processing by the audio zoom processing unit 33 e is supplied.

[0155] FIG. 48 is a specific configuration example of the audio zoom processing unit 33 e. The audio zoom processing unit 33 e differs from the audio zoom processing unit 33 b of FIG. 42 in that adders 338 c and 338 d

are added. The adders 338 c and 338 d add the zoom-up sound signal output from the variable amplifier 337 to the binaural sound signal obtained by decoding by the audio decoding section 26. The headphone listening audio signal added by the adders 338 c and 338 d is supplied to the headphone 55 via the audio output terminal 37 c.

[0156] The effect of reproducing the zoom-up audio signal according to the second to fifth embodiments described above is as follows.

[0157] When the zoom magnification is set to be small, the signal added by the adder 335 is not increased by the variable amplifier 337, so that the viewer 59 sets the image signal displayed on the monitor 52 and the photographing And the audio signal having realistic sensation reproduced by 360 degrees audio signal surrounding the person 300 with the speakers 53 and 54. Even when the viewing angle is set wide, the viewing angle is about 60 degrees. The viewer 59 sometimes recognizes that it is a so-called blank speech sound in which the sound generated by the subject force displayed on the monitor 52 is insufficient due to the difference between the viewing angle of the image and the angle range when collecting the sound signal . On the other hand, as described above, since the audio signal subjected to the audio zoom-up process is enhanced for the photographer 300 by adding the signal component from the median plane to the binaural sound signal, the blank area is compensated Sound signal. Therefore, it is possible to prevent the viewer 59 from recognizing that the viewer 59 is a hollow speech sound, and the viewer 59 can listen to a sound having a realistic feeling with less sense of discomfort as compared with the first embodiment.

[0158] ⟨⟨ Sixth Embodiment ⟩⟩

The video / audio recording / reproducing apparatus 106 according to the sixth embodiment shown in FIGS. 49 and 50 is not provided with built-in stereo microphone 21 and binaural microphone 3 separately as in the first to fifth embodiments Shows a configuration in which a normal stereophonic sound collecting microphone and a binaural microphone are shared. 49 is a plan view showing an external configuration example of the video / audio recording / reproducing apparatus 106 in the sixth embodiment, and FIG. 50 is a block diagram showing a concrete internal configuration example of the video / audio recording / reproducing apparatus 106. In FIGS. 49 and 50, components having the same functions as those in FIGS. 1 and 3 are denoted by the same reference numerals, and description thereof is omitted.

[0159] As shown in FIG. 49, the video / audio recording / reproducing apparatus 106 houses the microphone placing sections 35 a, 35 b for placing the microphones 31 e, 31 f on the top surface and the microphone cords 310 e, 310 f connected to the microphones 31 e, 31 f And a cord storage section 34.

[0160] In FIG. 49, when sound is collected by normal stereo sound using the microphones 31 e, 31 f, the photographer 300 places the microphones 3 le, 3 If on the microphone placing parts 35 a, 35 b and collects sound. In the case of collecting binaural sounds, the photographer 300 pulls out the microphone cords 310e and 310f from the cord storage unit 34, attaches the microphones 31e and 31f to the ear 302, and collects sounds. The video / audio recording / reproducing apparatus 106 has a projection-like detecting section (not shown) for detecting the microphones 31 e, 31 f placed on the microphone placing sections 35 a, 35 b. With the on-Z off of the switch (corresponding to the switch Sw 4 in FIG. 50) that operates in conjunction with the protrusion detector, the video / audio recording / reproducing device 106 places the microphones 3 le, 31 f on the microphone placing portions 35 a, 35 b Whether or not the force is being detected. The detection of whether or not the force placed on the microphone placement portions 35a and 35b is detected is not limited to this, and for example, the magnetic field generating also the permanent magnet force built in the microphones 31e and 31f may be detected by a Hall element or a magneto- It is also possible to detect using.

[0161] Specifically, in FIG. 50, if the microphones 31 e and 31 f are not placed on the microphone placing parts 35 a and 35 b, the switch Sw 4 is connected to the terminal e and turned off, the microphones 31 e and 31 f are placed If it is connected to the terminal f, it is turned on. The placement detection section 41a detects whether or not the microphone 31e, 31f is placed on the microphone placing sections 35a, 35b, depending on whether the switch Sw 4 is connected to the force terminal f connected to the terminal e . The detection signal from the placement detection section 41 a is supplied to the control section 47.

[0162] When the placement detection unit 41a detects that the microphones 31e and 31f are placed, the microphones 31e and 3 If perform normal stereo sound collection, and the control unit 47 controls the video / audio recording / reproducing apparatus 106 controls each part of the circuit so as to perform a recording operation corresponding to photographing in the normal mode. In this case, the roles of the microphones 31 e and 31 f are equivalent to the built-in stereo microphone 21 in FIG. 3, and the flag information generating unit

42 does not generate the binaural flag signal indicating that it is in the binaural mode. On the other hand, when the placing / detecting unit 41 a detects that the microphones 31 e and 31 f are not placed, the photographer 300 wears the microphones 31 e and 3 If on the ears 302 to collect binaural sounds in binaural mode , The control unit 47 controls each unit of the circuit so that the video / audio recording / reproducing apparatus 106 performs the recording operation corresponding to the imaging in the binaural mode. In this case, the flag information generating unit 42 generates a no-noal flag signal under the control of the control unit 47.

[0163] In the sixth embodiment, the microphones 31 e and 31 f, the microphone placing parts 35 a and 35 b, the placing and detecting part 41 a, and the control part 47 as a whole, as microphones for collecting ambient sounds, a bike to be worn by the photographer's ears It acts as a switching unit that switches whether to use a microphone other than a binaural microphone that uses a nor- mal microphone.

[0164] Here, an example of the structure of the cord storage section 34 will be described with reference to FIG. 51. 51 (A) shows the internal structure of the cord storage portion 34 with respect to the upper surface side force, and the microphone cords 310 e and 310 f are wound around the reel 341 having the rotation shaft 343. 51 (B) shows an internal structure of the cord storage portion 34 as viewed from the lower surface side, and a spiral spring 342 is provided inside the reel 341. Instead of or in addition to detecting whether or not the microphones 31e and 31f are placed on the microphone placing parts 35a and 35b, by detecting the rotation angle of the reel 341, it is possible to determine whether or not the binaural mode You can also detect.

[0165] ⟨⟨ Seventh embodiment ⟩⟩

The video / audio recording / reproducing apparatus 107 of the seventh embodiment shown in FIG. 52 shows a configuration using a wireless type binaural microphone that wirelessly transmits the sound signal obtained by collecting the sound signal to the main body of the apparatus. In FIG. 52, components having the same functions as those in FIG. 1 are denoted by the same reference numerals, and description thereof is omitted.

[0166] In FIG. 52, the video / audio recording / reproducing apparatus 107 includes a radio transmission / reception section 39 instead of the external microphone connection terminal 32 of FIG. 1, and performs voice recording using a radio type binaural microphone 38 instead of the noise noisy microphone 3 . The photographer 300 attaches the wireless binaural microphone 38 wirelessly coupled to the main body of the apparatus to the head and inserts the right and left microphones 38 a and 38 b into the ear 302 to collect the sound so that it is not interrupted by the microphone code You can shoot. It is also possible for two persons, the photographer 300 and a sound collector (not shown) to take photographs.

[0167] With reference to FIG. 53, the internal configuration of the video / audio recording / reproducing apparatus 107 will be described. In FIG. 53, components having the same functions as those in FIG. 3 are denoted by the same reference numerals, and description thereof is omitted. 53 is different from the video / audio recording / reproducing apparatus 10 1 shown in FIG. 3 in that a wireless transmitting / receiving unit 39 is provided instead of the external microphone connecting terminal 32 and the connection detecting unit 41 Different in point!

[0168] When the wireless transmitting / receiving unit 39 receives the binaural sound signal from the wireless type binaural microphone 38, the control unit 47 determines that the switch Swl is connected to the terminal b And causes the binaural audio signal to be supplied from the wireless type binaural microphone 38 to the audio encoding section 22. In this case, the control unit 47 controls the flag information generating unit 42 to generate a binaural flag signal. Further, if it is determined that the wireless main binaural microphone 38 also exceeds the predetermined distance, the control unit 47 connects the switch Swl to the terminal a and transmits the stereo sound signal from the built-in stereo microphone 21 to the terminal a And supplies it to the audio encoding section 22. In this case, the flag information generating unit 42 does not generate the binaural flag signal.

[0169] FIG. 54 shows an example of the internal configuration of the wireless binaural microphone 38 and the wireless transceiver 39, and their operation will be described.

[0170] As shown in FIG. 54, the microphone 38 a of the wireless type binaural microphone 38 includes a microphone unit 38 1, a microphone amplifier 382, a transmission / reception unit 383, an antenna 384, and a warning signal transmission unit 385. The force microphone 38 b (not shown) has the same configuration as that of the microphone 38 a except that it does not include the warning signal transmitting unit 385. The wireless transceiver 39 includes a transmission / reception unit 391, a microphone determination unit 392, a distance

measurement unit 393, a communication coverage area determination unit 394, a warning signal transmission unit 395, and an antenna 396.

[0171] The microphone unit 381 of the microphone 38 a (38 b) generates a binaural sound signal. The microphone amplifier 382 amplifies the noisy audio signal from the microphone unit 381. The transmission / reception unit 383 modulates the amplified binaural sound signal from the microphone amplifier 382 according to a predetermined modulation scheme, and transmits the modulated binaural sound signal via the antenna 384. The warning signal transmitting unit 385 generates an alarm signal based on the warning signal generated by the warning signal transmitting unit 395 of the wireless transmitting / receiving unit 39 (described later) and transmitted via the transmitting / receiving unit 391 and the transmitting / receiving unit 383.

[0172] The antenna 396 of the wireless transceiver 39 receives the modulated signals transmitted from the right and left microphones 38 a and 38 b. The transmitting / receiving unit 391 demodulates the received modulated signal to generate a binaural sound signal and measures the intensity of the received power of the modulated signal. The distance measuring unit 393 estimates the distance from the wireless transmitting / receiving unit 39 to the wireless binaural microphone 38 based on the measured intensity of the received power. The communicable range inside determination unit 394 determines whether the estimated distance is a force within a communicable predetermined distance. The determination result of the communication coverage area determination unit 394 is supplied to the control unit 47. If the estimated distance is within a predetermined communication distance, the control unit 47 connects the switch Swl to the terminal b, and controls the flag information generation unit 42 to generate a binaural flag signal. The control unit 47 connects the switch Swl to the terminal "a" when the estimated distance exceeds a predetermined distance capable of communicating.

[0173] The warning signal transmitting unit 395 generates a warning signal when it is determined that the distance estimated by the communication range in-coverage determining unit 394 exceeds a predetermined distance that can be communicated. The warning signal is supplied to the control unit 47. The control unit 47 generates a warning mark and supplies it to the display unit 17 to display a predetermined warning mark on the display unit 17. When the warning signal transmitting unit 395 does not generate the warning signal, the microphone judging unit 392 judges that the binaural sound signal is normally obtained, and the binaural sound signal obtained by demodulating by the transmitting / And supplies it to the audio encoder 22 via the switch Swl.

[0174] FIG. 55 shows an example of a warning display performed by the wireless type binaural microphone 38 and the video / audio recording / reproducing apparatus 107. The microphone 38 a is provided with a rod-like member having a light emitting diode (LED) 386 attached to its distal end portion. A warning signal generated by the warning signal transmitting unit 385 is supplied to the LED 386, and the LED 386 blinks (or lights) based on the warning signal. In addition to blinking LED 386, or LED 386 may flash warning sound instead of blinking! In this case, it is possible to lower the level of the warning sound or to set the frequency of the warning sound to a low frequency of several tens of Hz or less, for example, so that the warning sound is not heard by the microphone unit 381 Preferable,.

[0175] The warning signal transmitter 395 generates a warning signal when it is determined that the wireless binaural microphone 38 has crossed the communicable area indicated by the dashed circle. As shown in FIG. 55, if the wireless type binaural microphone 38 exceeds the communicable area, a predetermined warning mark is displayed on the display unit 17.

[0176] FIG. 56 shows an example of a warning mark displayed on the display unit 17. The warning mark 171 a shown in FIG. 56 (A) is displayed with the X mark superimposed on the no-noal mark 171 shown in FIG. 5 (A). The warning mark 172a shown in FIG. 56 (B) is obtained by graying out the mark 172 shown in FIG. 5 (B). Any of the marks in FIGS. 56 (A) and 56 (B) may be used as the warning mark, and other marks may be used as a matter of course. If it is assumed that the received power at the wireless transmission / reception unit 39 is below the reception limit despite the warning display, the control unit 47 switches from the wireless binaural microphone 38 to the built-in stereo microphone 21.

[0177] The operation of the video / audio recording / reproducing apparatus 107 will be described further using FIG. 57. In FIG. 57, in step S251, the control unit 47 determines whether it is in binaural mode. If it is not determined in step S251 that the binaural mode is set (NO), in step S253, the recording and reproducing unit 14 collects the sound by the built-in stereo microphone 21 and records a normal stereo sound signal on the recording medium 44. FIG. If the binaural mode is determined in step S251 (YES), in step S252, the wireless transceiver 39 receives the transmission signal from the wireless binaural microphone 38. In step S undefined

254, the distance measuring unit 393 detects the distance from the wireless transmitting / receiving unit 39 to the wireless binaural microphone 38 based on the strength (receiving intensity) of the received power. In step S255, the communication coverage area determination unit 394 determines whether or not the detected distance is within a predetermined distance.

[0178] If it is not determined in step S255 that the distance is within the predetermined distance (NO), the control unit 47 determines in step S257 whether or not the warning display time t is 0 (non-present) and determines whether or not the warning display time If t is 0, the warning signal transmitting unit 395 is controlled to generate a warning signal in step S300. After the warning signal is generated, the process returns to step S254 to repeat the same process. In step S 257, if the warning display time t is not 0 (NO), in step S 258, the control unit 47 determines whether the warning display time t exceeds a predetermined maximum time t max or not and determines the maximum If the time tmax has not been exceeded (NO), the process proceeds to step S300, and the process returns to step S254 in the same manner. If it exceeds the maximum time tmax (YES), in step S259, the control unit 47 controls the switch Swl so as to switch from the wireless type binaural lull microphone 38 to the built-in stereo microphone 21, and stops the generation of the warning signal So as to control the warning signal transmitter 395. Then, the process proceeds to step S25 3.

[0179] If it is determined in step S255 that the distance is within the predetermined distance (YES), in step S256, if the warning signal transmitting unit 395 is generating a warning signal, the control unit 47 stops generating the warning signal As shown in FIG. Then, in step S 301, the recording 'playback unit 14 collects the sound by the wireless type binaural microphone 38 and records the binaural sound signal on the recording medium 44. In step S302, the control unit 47 determines whether or not an operation for ending the recording has been performed. If the operation for ending the recording has not been performed (NO), the control unit 47 returns to step S251, and if the recording end operation has been performed (YES), the process is terminated.

Industrial applicability

[0180] The present invention can be applied not only to consumer video cameras but also to professional video cameras to video and audio recording and reproducing apparatuses that need to reproduce realistic sounds together with the captured images. It can also be applied to digital cameras and mobile phones with movie shooting functions. Furthermore, although the present invention is suitable for use in a video / audio recording / reproducing apparatus for recording and reproducing both a video signal and an audio signal, it is sufficiently effective even when used in an audio recording / reproducing apparatus for recording / reproducing only an audio signal .

WO2006054698A1 VIDEO/AUDIO RECORDING APPARATUS AND METHOD, AND VIDEO/AUDIO REPRODUCING APPARATUS AND METHOD

Data originating from sources other than the EPO may not be accurate, complete, or up to date.

The wording below is an initial machine translation of the original publication. To generate a version using the latest translation technology, go to the original language text and use Patent Translate.

The scope of the claims

[1] (101, 102, 103, 104, 105, 106) for recording a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer (300) including a sound emitted from the subject, 107) 【こっ!/ヽて,

A camera unit (11) for imaging the object,

A switch unit (Sw1) for switching whether to use a microphone other than the binaural microphone (3) using the binaural microphone (3) worn on the ear of the photographer (300) as the ambient sound)When,

A video processing section (12) for processing the video signal outputted from the camera section (11), an audio processing section (22) for processing the audio signal outputted from the microphone which sounds the ambient sound,

A sound mode when the ambient sound is turned on as a binaural mode when switching to the binaural microphone (3) as a microphone for collecting the ambient sound by the switching unit (Swl), and that the sound mode is a noorrhard mode A flag information generating section (42) for generating a binaural flag signal indicating a binaural flag

A recording unit (14) for recording the video signal processed by the video processing unit (12), the audio signal processed by the audio processing unit (22), and the binaural flag signal on a recording medium,

(101, 102, 103, 104, 105, 107) characterized in that it comprises:

[2] A built-in microphone (21) incorporated in the video and audio recording apparatus (101, 102, 103, 104, 105, 107)

An external microphone connection terminal (32)

A setting section (48, 47) for setting as the external microphone to be connected to the external microphone connection terminal (32) whether to set the binaural lull mike (3) or a microphone other than the binaural microphone; a setting section A connection detection section (41) for detecting whether or not an external microphone is connected to the connection section

A switch (Swl) for switching between the audio signal output from the built-in microphone (21) and the audio signal output from the external microphone as audio signals to be supplied to the audio processing unit (22)

It is detected that the external microphone is set to the binaural microphone (3) by the setting section (47, 48) and the external microphone is connected to the external microphone connection terminal (32) by the connection detection section (41) , And a control unit (47) for setting the audio mode to the binaural mode,

When the voice mode is the binaural mode, the control unit (47) controls the switch (Sw1) so that the switch (Sw1) supplies the audio signal output from the external microphone to the audio processing unit (22)), And the flag information generating section (42) controls the flag information generating section (42) so as to generate the binaural flag signal as described in the first aspect of the invention Audio recording apparatus (101, 102, 103, 104, 105, 107).

[3] A display unit (17) for displaying a video signal output from the camera unit (11)

Characterized in that it further comprises a display control unit (47) for displaying a binaural mark indicating that the binaural mode is displayed on the display unit (17) when the audio mode is the inno informal mode. The video and audio recording apparatus (101, 102, 103, 104, 105, 107) according to 1.

[4] The camera unit (11) has a zoom function for enlarging and capturing the object,

A sound zoom processing unit (33) for amplifying the sound signal collected by the binaural microphone (3) according to the enlargement ratio in the camera unit (11). (102) according to the first aspect of the invention.

[5] The camera unit (101) has a zoom function for enlarging and capturing the object,

A head related transfer function for forming a virtual sound source by virtually bringing a sound source of an audio signal collected by the binaural microphone (3) close to a listener corresponds to a plurality of distances between the virtual sound source and the listener A transfer function memory (334) for storing and storing a plurality of head related transfer functions stored in the transfer function memory (33 4) selectively according to the enlargement ratio in the camera section (11) A convolution operator (336) for convolving the head related transfer functions acquired by the function acquiring unit (333) with respect to a sound signal collected by the binaural microphone (3), a function acquiring unit (333) (102, 103) according to claim 1, further comprising an audio zoom processing section (33a, 33c) having a plurality of audio signal processing sections (33a, 33b)

[6] A video and audio recording method for recording a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer (300) including a sound emitted from the subject,

An imaging step (S155) of imaging the subject,

A switching step (S 151) of switching whether to use a microphone other than a binaural microphone that uses a binaural microphone (3) worn on the ear of the photographer (300) as the ambient sound,

A video processing step (S156) of processing a video signal obtained by imaging the subject, an audio processing step (S 156) of processing a sound signal output from a microphone for collecting the ambient sound,

Wherein when switching to the binaural rhythmic (3) as a microphone for collecting the ambient sound at the switching step (S 151), a sound mode at the time of collecting the ambient sound is set as a binaural mode, a noiselal mode A flag information generating step (S 158) for generating a binaural flag signal indicating that the binaural flag signal is present,

A recording step (S161) of recording in the recording medium (44) the video signal processed in the video processing step (S156), the audio signal processed in the audio processing step (S156) and the binaural flag signal)When,

And a video / audio recording method.

[7] An audio / video reproducing apparatus (101) for reproducing a recording medium (44) on which a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer (300) including a sound emitted from the subject are recorded , 102, 103, 104, 105, 106, 107) 【Ko!

A reproduction unit (14) for reproducing a recording signal recorded on the recording medium,

A separation unit (15) for separating the video signal and the audio signal from the recording signal reproduced by the reproduction unit (14)

A video processing unit (16) for processing the video signal separated by the separation unit (15), an audio processing unit (26) for processing the audio signal separated by the separation unit (15), ,

When a binaural flag signal indicating that a binaural microphone worn on the ear of the photographer (300) is used as the microphone for collecting the ambient sound on the recording medium (44), the binaural flag A flag information acquisition unit (36) for acquiring a flag signal, a flag information acquisition unit (36) for acquiring a flag signal, a flag information acquisition unit (36) for acquiring a flag signal, And a crosstalk

canceller (27) for processing the audio signal so as to cancel out a crosstalk signal generated when sound is generated by the sound source (53, 54)

The crosstalk canceller (27) has filters (272 a - 272 d) for convolving the filter characteristics obtained in advance with respect to the audio signal, and the filter characteristic is obtained by combining the measurement signal with a pair of microphones (101, 102, 103, 104, 105, 106, 107) characterized in that it is a filter characteristic based on a head related transfer function measured using collected sound signals.

A video and audio reproducing method for reproducing a recording medium (44) in which a video signal obtained by imaging a subject and an audio signal obtained by collecting ambient sound of a photographer (300) including sound emitted from the subject are recorded,

A reproducing step (S 181) of reproducing a recording signal recorded on the recording medium (44); a separating step of separating the video signal and the audio signal from the recording signal reproduced in the reproducing step (S183)

A video processing step (S 184) for processing the video signal separated in the separating step (S 183)

An audio processing step (S 184) for processing the audio signal separated by the separating step (S 183)

A binaural flag signal indicating that a binaural microphone (3) to be worn on the ear of the photographer (300) is used as the microphone for collecting the ambient sound is recorded on the recording medium (44) A step (S 186) of acquiring the binaural flag signal when acquiring the binaural flag signal in the flag information obtaining step (S 186), and a step of acquiring the binaural flag signal And a crosstalk canceling step (S 188) of processing the audio signal so as to cancel a crosstalk signal generated when the sound signal processed by the speaker (53, 54) is sounded,

The crosstalk cancellation step (S188) is a step of convolving the filter characteristics obtained in advance with respect to the voice signal, and the filter characteristic is a sound obtained by converting the measurement signal into a sound collected by a pair of microphones mounted on the cylindrical structure Wherein the filter characteristic is a filter characteristic based on a head related transfer function measured using a signal.

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

REVISED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 August 2007 (16.08.2007)

PCT

(10) International Publication Number
WO 2007/092660 A1

(51) International Patent Classification:
H04N 7/18 (2006.01) *G06F 19/00* (2006.01)

(21) International Application Number:
PCT/US2007/060620

(22) International Filing Date: 17 January 2007 (17.01.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/765,628 6 February 2006 (06.02.2006) US

(71) Applicant *(for all designated States except US)*: KONINKLUKE PHILIPS ELECTRONICS, N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: MOBERLY, David [US/US]; 3899 Magnolia Drive, Palo Alto, California 94306 (US).

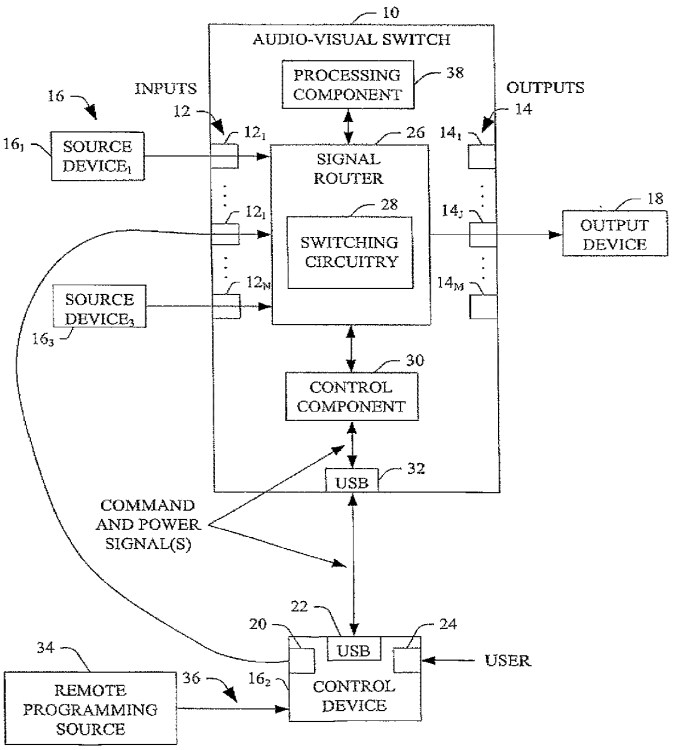
(74) Common Representative: KONINKLIJKE PHILIPS ELECTRONICS, N.V.; c/o Douglas B. McKnight, 595 Miner Road, Cleveland, Ohio 44143 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: USB-ENABLED AUDIO-VIDEO SWITCH



(57) Abstract: A telemedicine system (8) includes an audio-video switch (10) that receives at least two different signals and conveys at least one of the at least two different signals to an output device (18) that presents it to a user. A control device (162) monitors a position of the audio-video switch (10) and automatically controls the position to select which of the at least two different signals is passed through the audio-video switch (10) to the output device (18) in response to a signal from a remote source (34). Signals received from the remote source (34) include medically related audio video message signals, which are input to the audio-video switch (10), and a control signal that causes the audio-video switch (10) to select between passing the medically related audio-video message signals and other received signals to the output device (18).

WO 2007/092660 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).

Declaration under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(H))*

Published:

— *with international search report*

(88) Date of publication of the supplementary international search report: 30 October 2008

(15) Information about Correction:
see Notice of 30 October 2008

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/060620

A CLASSIFICATION OF SUBJECT MATTER
INV. H04N7/18 G06F19/00

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)
H04N G06F

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	EP 1 126 715 A2 (KATZ RONALD ALAN [US] TELEBUYER LLC [US]) 22 August 2001 (2001-08-22) page 3, paragraph 13 page 9, paragraph 70 page 16, paragraph 134 - page 17, paragraph 151; figures 1-5	1-26
X	US 2005/052527 A1 (REMY CHRISTOPHE [US] ET AL) 10 March 2005 (2005-03-10) page 3, paragraph 30 - paragraph 32; figure 3 page 5, paragraph 41 - paragraph 42; figures 7,8	1-26

-/-

[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 prionty claim(s) or which fs 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 pπonty 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 docu¬ments, 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 Apri l 2007	13/04/2007

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

Form PCT/ISA/210 (second sheet) (April 200B)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/060620

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 717 609 B2 (SUGIMOTO HIDEO [JP] ET AL) 6 April 2004 (2004-04-06) column 7, line 19 - column 8, line 42 abstract; figures 1-5	1-26
X	US 5 553 609 A (CHEN YAOBIN [US] ET AL) 10 September 1996 (1996-09-10) column 9, line 57 - column 12, line 53; figures 5-9	1,23,26
A	EP 1 343 326 A1 (MACROSYSTEM DIGITAL VIDEO A6 [DE]). 10 September 2003 (2003-09-10) abstract; figure 1 column 6, line 22 - line 58	1-26

Form PCT/ISAS10 (continuation of second sheet) (April 200B)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2007/060620

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1126715 A2	22-08-2001	NONE	
US 2005052527 A1	10-03-2005	US 2006119701 A1	08-06-2006
US 6717609 B2	06-04-2004	DE 10101064 A1	19-07-2001
		JP 2001197485 A	19-07-2001
		US 2003197781 A1	23-10-2003
US 5553609 A	10-09-1996	AU 4980596 A	27-08-1996
		WO 9624284 A1	15-08-1996
EP 1343326 A1	10-09-2003	AT 275799 T	15-09-2004
		DE 50200955 D1	14-10-2004
		US 2003193577 A1	16-10-2003

Note: This international search report was established in addition to the report duly established by the competent International Searching Authority specified by the applicant. It is published for information only and has no legal status for the purposes of the PCT procedure (for example, in the computation of time limits).

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US07/60620

A. CLASSIFICATION OF SUBJECT MATTER
IPC: **H04N 7/18**(2006.01),**9/47**(2006.01)

USPC: 348/61
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. : 348/61, 65, 74; 600/522

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,553,609 (CHEN et al.) 10 September 1996, entire document	1-3, 5, 7-11, 13-24
---		---
Y		4, 6, 12, 25-26
Y	US 2008/005421 A1 (CHANG et al.) 03 January 2008; entire document	4, 6, 12, 25-26

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
14 July 2008 (14.07.2008)

Date of mailing of the international search report
29 AUG 2008

Name and mailing address of the ISA/US
Mail Stop PCT, Attn ISA/US
Commissioner for Patents
P O **Box 1450**
Alexandria, Virginia 22313-1450
Facsimile No. (571) 273-3201

Authorized officer
Mehrdad Dastouri
Telephone No. 571-272-2600

Form PCT/ISA/210 (second sheet) (April 2007)

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 May 2008 (02.05.2008)

PCT

(10) International Publication Number
WO 2008/050583 A1

(51) International Patent Classification:
H04R 3/02 (2006.01) *H04R 1/02* (2006.01)
H04M 1/62 (2006.01) *H04R 1/28* (2006.01)

(21) International Application Number:
PCT/JP2007/069131

(22) International Filing Date:
21 September 2007 (21.09.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2006-291843 26 October 2006 (26.10.2006) JP
2006-291844 26 October 2006 (26.10.2006) JP

(71) Applicant *(for all designated States except US)*: **MATSUSHITA ELECTRIC WORKS, LTD.** [JP/JP]; 1048, Oaza-Kadoma, Kadoma-shi, Osaka, 5718686 (JP).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KITADA, Kosaku** [JP/JP]; c/o Matsushita Electric Works, Ltd., 1048, Oaza-Kadoma, Kadoma-shi, Osaka, 5718686 (JP). **KIMOTO, Shinya** [JP/JP]; c/o Matsushita Electric Works, Ltd., 1048, Oaza-Kadoma, Kadoma-shi, Osaka, 5718686 (JP). **AKASAKA, Osamu** [JP/JP]; c/o Matsushita Electric Works, Ltd., 1048, Oaza-Kadoma, Kadoma-shi, Osaka, 5718686 (JP). **YOSHIDA, Keiichi** [JP/JP]; c/o Matsushita Electric Works, Ltd., 1048, Oaza-Kadoma, Kadoma-shi, Osaka, 5718686 (JP). **ARIKAWA, Yasushi** [JP/JP]; c/o Matsushita Electric Works, Ltd., 1048, Oaza-Kadoma, Kadoma-shi, Osaka, 5718686 (JP).

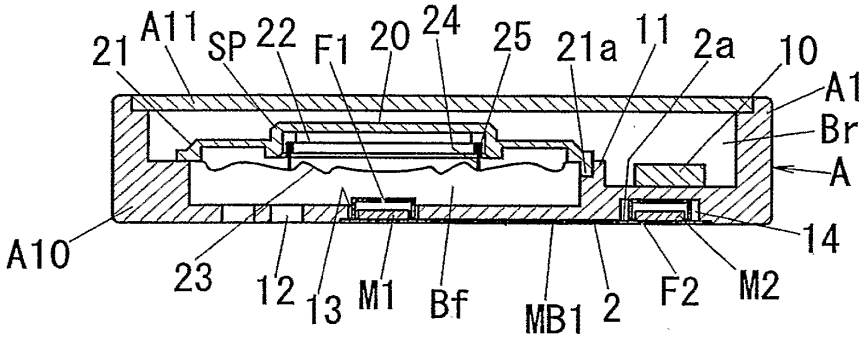
(74) Agents: **NISHIKAWA, Yoshikiyo** et al.; Hokuto Patent Attorneys Office Umeda-Daiichiseimei Bldg. 5th Floor, 12-17, Umeda 1-chome, Kita-ku, Osaka-shi, Osaka 5300001 (JP).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) **Title:** INTERCOM DEVICE AND WIRING SYSTEM USING THE SAME



(57) **Abstract:** A compact intercom device is provided, which efficiently prevents howling phenomenon. This intercom device includes a housing; a speaker accommodated in the housing such that a front air chamber is defined at a front side of the speaker and a rear air chamber is defined as a rear side of the speaker, a first microphone disposed such that its sound receiving surface faces the front air chamber, a second microphone disposed such that its sound receiving surface faces outside of the housing; and a signal processing unit. When an audio output of the speaker is picked up by the second microphone, the signal processing unit removes a signal component corresponding to the audio output of the speaker from an output signal of the second microphone by use of an output signal of the first microphone.

WO 2008/050583 A1

DESCRIPTION

INTERCOM DEVICE AND WIRING SYSTEM USING THE SAME

TECHNICAL FIELD

The present invention relates to an intercom device suitable for an interphone system or the like, and a wiring system using the same.

BACKGROUND ART

Conventionally, an interphone system has been widely used as short-distance conversation means between separated rooms or between an entrance and a room of a building or a house. A typical interphone system for household use is mainly formed with an indoor intercom device installed on a wall surface of the room, and an outdoor intercom device installed on a wall surface of the entrance, and connected to the indoor intercom device through a transmission line. For example, the outdoor intercom device has a microphone for receiving a visitor's voice, and a speaker for outputting a dweller's voice. By operating the outdoor intercom device, the visitor can have a conversation with the dweller using the indoor intercom device.

In this kind of intercom device, since the speaker is located near the microphone, there is an inconvenience that howling phenomenon easily occurs when the audio output of the speaker is picked up by the microphone. To solve this inconvenience, various measures or methods for preventing the howling phenomenon have been proposed.

For example, Japanese Patent Publication No. 2607257 discloses a loudspeaker communication device **300**, which is suitable for a handsfree telephone. As shown in FIG. 35A, this communication device **300** has a speaker **SP**, and a pair of microphones (**M1**, **M2**). In addition, to prevent the howling phenomenon caused when an audio output of the speaker **SP** is picked up by the microphone (**M1**, **M2**), the communication device **300** also includes a pair of amplifying circuits (**310**, **320**), a level-control amplifying circuit **330**, a

delay circuit **340**, and a differential amplifying circuit **350**, as shown in FIG. 35B. The delay circuit **340** delays an output of the microphone **M1** located closer to the speaker **SP** by a delay time, which is determined according to sound velocity and a difference between a distance **d1** between the microphone **M1** and the speaker **SP** and a distance **d2** between the microphone **M2** and the speaker **SP**. The level-control amplifying circuit **330** controls the output level of the microphone **M2** such that an amplification factor is **d2/d1**. In other words, the output level of the microphone **M2** is controlled so as to be substantially equal to the output level of the microphone **M1** with respect to the signal component corresponding to the audio output of the speaker **SP**. The differential amplifying circuit **350** receives an output of the delay circuit **340** and an output of the level-control amplifying circuit **330**, and provides an output signal corresponding to a difference between these outputs.

According to this communication device, the signal component corresponding to the audio output of the speaker **SP** is cancelled from the output of the microphone (**M1**, **M2**) by the differential amplifying circuit **350**. Therefore, it is possible to achieve a comfortable communication, while preventing the howling phenomenon.

However, as shown in FIG. 35A, since each of the microphones (M1, M2) is spaced from the speaker SP by a relatively large distance, there is another problem that the device itself becomes large in size. In addition, since the intercom device is usually installed on the wall surface or the like, it is strongly desired that the intercom device has a thin thickness. On the other hand, when the speaker is accommodated in an excessively thin housing, a volume of a rear air chamber, which is defined as a space of a rear side of the speaker, is reduced, so that a radiating sound pressure of the speaker lowers and a minimum resonance frequency of the speaker shifts to a higher frequency level. As a result, there is a fear that deteriorations in sound quality and electro-acoustic conversion efficiency of the speaker occur.

Thus, the conventional intercom device still has plenty of room for

improvement in downsizing of the device itself, sound quality, electro-acoustic conversion efficiency as well as the prevention of howling phenomenon.

SUMMARY OF THE INVENTION

Therefore, a primary concern of the present invention is to provide a compact intercom device, which has the capability of preventing howling phenomenon, and achieving an improvement in sound quality and efficiency.

That is, the intercom device of the present invention is characterized by comprising:

a housing;

a speaker accommodated in the housing such that a front air chamber is defined as a space enclosed by a front surface of the speaker and an inner surface of the housing, and a rear air chamber is defined as a space enclosed by a rear surface of the speaker and an inner surface of the housing;

a first microphone disposed such that its sound receiving surface faces the front air chamber;

a second microphone disposed such that its sound receiving surface faces outside of the housing; and

a signal processing unit configured to, when an audio output of the speaker is picked up by the second microphone, remove a signal component corresponding to the audio output of the speaker from an output signal of the second microphone by use of an output signal of the first microphone.

According to the invention, since the first microphone easily and efficiently collects the audio information output from the speaker, howling phenomenon can be efficiently prevented by the signal processing unit. In addition, a distance between the first microphone and the speaker becomes short, thereby allowing downsizing of the intercom device. Furthermore, since the sound receiving surface of the second microphone faces the outside of the housing, an acoustic coupling between the speaker and the second microphone reduces, and the second microphone becomes hard to pick up the audio output of the

speaker. As a result, the second microphone can be located closer to the speaker. Thus, it is possible to provide a compact intercom device capable of preventing the howling phenomenon.

From the viewpoint of more efficiently preventing the howling phenomenon, it is preferred that the rear air chamber is the space hermetically-sealed by the rear surface of the speaker and the inner surface of the housing. In addition, it is preferred that the first microphone is arranged such that the sound receiving surface of the first microphone faces a diaphragm of the speaker. Moreover, when the first microphone and the second microphone are mounted on a same circuit board, it is preferred that the circuit board is mounted on an outside surface of the housing such that the sound receiving surface of the first microphone faces the front air chamber and the sound receiving surface of the second microphone faces the outside of the housing.

On the other hand, from the viewpoint of further downsizing the intercom device, the second microphone is preferably located adjacent to the speaker. In addition, it is preferred that the first and second microphones are mounted on a single board.

From the viewpoint of improving sound quality and electro-acoustic conversion efficiency of the speaker, it is preferred that the intercom device has an acoustic tube having an opening communicated with the rear air chamber at its one end, and a closed end at the other end. Alternatively, the intercom device may have a plurality of acoustic tubes having different lengths, each of which has an opening communicated with the rear air chamber at its one end, and a closed end at the other end. In addition, it is preferred that the acoustic tube has a tube length of an odd multiple of one quarter of a wavelength determined in consideration of sonic velocity and an intended frequency where an increase in sound pressure level of the audio output of the speaker is needed.

Furthermore, it is preferred that the acoustic tube has a resonant frequency between a frequency equivalent to a lowest resonant frequency of the speaker that is supposed to be mounted on an infinite baffle and a lowest

resonant frequency of the speaker mounted on the rear air chamber with the absence of the acoustic tube. The acoustic tube can be formed along an inner surface of the rear air chamber. If necessary, a sound absorbing material may be disposed at the opening of the acoustic tube.

A further concern of the present invention is to provide a wiring system using the intercom device described above. That is, the wiring system comprises at least one transmission line installed in a building structure to transmit electric power and information signals, and a plurality of base units each adapted in use to be mounted in a wall surface of the building structure, and connected to the transmission line. The intercom device has a first connector detachably connectable to a second connector formed in each of the base units. This wiring system is characterized in that when the first connector is connected to the second connector of one of the base units, the intercom device functions such that the information signals provided through the transmission line are output as audio information from the speaker unit, and an output of the signal processing unit is sent out through the transmission line.

By the way, in the intercom device installed on the wall surface of the house, one provided with a liquid crystal display or the like for displaying not only the audio information but also visual information such as images comes into practical use recently. Once such an intercom device is installed on the wall surface, not only electric wiring work but also mounting/repair work to the wall surface are required in a case where a layout of the intercom device has to be changed. However, it is not easy for a general user himself/herself to do the work. In addition, a function of the conventional intercom device is completed by itself, and a new-type intercom device is required in a case where it is desired to add another function. In such a case, not only a purchase of the new-type intercom device but also the above-described mounting/repair work becomes necessary, so that this is a great burden in cost for the user. The wiring system of the present invention is effect. According to the wiring system of the present

invention, since the intercom device is detachably connectable to the base unit, there advantages that the wiring system allows great flexibility for the change of layout of the intercom device, and makes it easy to replace the intercom device without cumbersome repairing work. Thus, the present invention provides the wiring system capable of solving the problems described above.

As a preferred embodiment of the wiring system of the present invention, the wiring system further comprises a function unit adapted in use to be connected to the transmission line through one of the base units, the function unit having a third connector detachably connectable to one of the second connector of the base unit, and an extended connector formed in the intercom device. When the third connector of the function unit is connected to the second connector of the base unit, or when the first connector of the intercom device is connected to the second connector of the base unit, and the third connector of the function unit is connected to the extended connector of the intercom device, the function unit provides at least one of functions for supplying electric power provided through the transmission line, outputting information signals provided through the transmission line, and sending out information signals through the transmission line. At this time, it is particularly preferred that the first connector and the third connector are stylized in shape and arrangement, and the second connector and the extended connector are stylized in shape and arrangement corresponding to the stylization of the first connector and the third connector.

As another preferred embodiment of the wiring system of the present invention, the wiring system further comprises a function unit adapted in use to be connected to the transmission line through one of the base units. The function unit has a third connector detachably connectable to the second connector of the base unit and an extended connector detachably connectable to the first connector of the intercom device. When the third connector of the function unit is connected to the second connector of the base unit, and the extended connector of the function unit is connected to the first connector of the

intercom device, the function unit provides at least one of functions for supplying electric power provided from the transmission line through the base unit, outputting information signals provided from the transmission line through the base unit, and sending out information signals to the transmission line through the base unit, and the intercom device operates such that the information signals provided from the transmission line through the base unit and the function unit are output as audio information from the speaker unit, and the output of the signal processing unit is sent out to the transmission line through the base unit and the function unit. At this time, it is particularly preferred that the first connector and the third connector are stylized in shape and arrangement, and the second connector and the extended connector are stylized in shape and arrangement corresponding to the stylization of the first connector and the third connector.

In the wiring system described above, it is preferred that the first and second connectors provide an electric power transmission between the base unit and the intercom device by means of electromagnetic coupling, and/or the first and second connectors provide a signal transmission between the base unit and the intercom device by means of optical coupling.

Further characteristics of the present invention and advantages brought thereby will be clearly understood from the best mode for carrying out the invention described below.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a perspective view of an intercom device according to a first embodiment of the present invention;

FIG. 2 is a cross-sectional view of the intercom device;

FIG. 3 is a circuit diagram of a voice processing unit of the intercom device;

FIG. 4 is a cross-sectional view of a microphone board of the intercom device;

FIG. 5 is a cross-sectional view of a bare chip employed in the microphone board;

FIG. 6A is a simplified plan view of the microphone board, and FIG. 6B is a simplified circuit diagram of the microphone board;

FIG. 7 is a circuit diagram of an impedance converting circuit of the microphone board;

FIG. 8 is a circuit diagram of a signal processing unit of the intercom device;

FIG. 9A and FIG. 9B are diagrams showing signal waveforms in the signal processing unit;

FIG. 10A and FIG. 10B are diagrams showing further signal waveforms in the signal processing unit;

FIG. 11A and FIG. 11B are diagrams showing another signal waveforms in the signal processing unit;

FIGS. 12A to 12C are diagrams showing still another signal waveforms in the signal processing unit;

FIG. 13 is a cross-sectional view of an intercom device according to a second embodiment of the present invention;

FIGS. 14A to 14C are cross-sectional and plan views of another housing of the intercom device according to the second embodiment;

FIG. 15 is a graph showing frequency characteristics of the cancelled amount of voice components of the speaker;

FIG. 16 is a graph showing frequency characteristics of the radiating sound pressure of the speaker;

FIG. 17 is a cross-sectional view of an intercom device using an ideal baffle plate;

FIG. 18 is a graph showing sound pressure characteristics of the intercom device having an acoustic tube and the intercom device having no acoustic tube;

FIG. 19 is a simplified cross-sectional view showing a modification of the housing having the acoustic tube;

FIG. 20 is a simplified cross-sectional view of a housing having a plurality of acoustic tubes;

FIG. 21 is a plan view of a housing having a sound absorbing material;

FIG. 22 is a schematic diagram of a wiring system using an intercom device according to a third embodiment of the present invention;

FIG. 23 is a schematic circuit diagram of a base unit of the wiring system;

FIG. 24 is a rear perspective view of the base unit;

FIG. 25A is a perspective view showing a gate housing and a main housing of the base unit, and FIG. 25B is a plan view of a module port of the gate housing;

FIG. 26 is a plan view of an attachment plate for mounting the base unit on a switch box;

FIG. 27 is a schematic circuit diagram of the function unit;

FIG. 28 is a perspective view of the intercom device for the wiring system;

FIG. 29 is a schematic circuit diagram of the intercom device for the wiring system;

FIG. 30 is a perspective view of the intercom device connected to the function unit or the base unit;

FIG. 31 is an exploded perspective view of the base unit;

FIG. 32A is a front view of the function unit, FIG. 31B is a side view of the function unit, FIG. 32C is an exploded side view of the function unit, and FIG. 32D is a perspective view of a joining member;

FIG. 33 is a perspective view illustrating how to mechanically couple the function unit to the base unit;

FIG. 34 is a schematic circuit diagram of an intercom device for a power line carrier type wiring system; and

FIG. 35A is a schematic perspective view of a conventional loudspeaker communication device, and FIG. 35B is a partial block diagram of the loudspeaker communication device.

BEST MODE FOR CARRYING OUT THE INVENTION

Referring to the attached drawings, an intercom device of the present invention and a wiring system using the intercom device are explained in detail

according to preferred embodiments.

(First Embodiment)

As shown in FIGS. 1 and 2, an intercom device **A** according to a first embodiment of the present invention includes a housing **A1**, a speaker **SP**, a microphone board **MB1**, and a voice processing unit **10**. The housing **A1** consists of a body **A10** having an opening at its back face, and a cover **A11** for closing the opening of the body **A10**, and accommodates the speaker **SP** and the voice processing unit **10** therein. The housing **A10** has a plurality of sound holes **12** formed in a front face of the body **A10**. The speaker **SP** accommodated in the housing **A1** provides a voice output through the sound holes **12** to outside of the housing **A1**. The microphone board **MB1** has a sound hole **F2** that picks up a voice emitted from a user of the intercom device **A**. In operation, when a conversation switch **SW1** is operated to the ON position, the intercom device **A** sends the picked-up voice to another intercom device **A** and receives signals from another intercom device **A** via an information line **Ls**.

As shown in FIG. 2, the speaker **SP** includes a cylindrical yoke **20** that is open at one end, and a round support **21** that extends radially outward from the open end of the yoke **20**. The yoke **20** is formed of an iron-containing material, such as a cold rolling steel plate (SPCC, SPCEN) or an electromagnetic soft iron (SUY), and has a thickness of about 0.8 mm. A columnar permanent magnet **22** (having a residual flux density of, for example, 1.39 T to 1.43 T) formed of, for example, neodymium is placed in the yoke **20**. An edge surface of the round support **21** is fixed on an edge portion on an outer peripheral side of a dome-shaped diaphragm **23**.

The diaphragm **23** is formed of a thermoplastic material, such as PET (PolyEthylene Terephthalate) or PEI (Polyetherimide), and has a thickness of, for example, 12 μm to 50 μm. A cylindrical bobbin **24** is fixed to the back of the diaphragm **23**, and a voice coil **25** is provided at the rear end of the bobbin **24**. The voice coil **25** is located at the open end of the yoke **20**, and the bobbin **24**

and the voice coil **25** are allowed to freely move back and forth (in the vertical direction in FIG. 2) in the vicinity of the open end of the yoke **20**. The voice coil **25** is formed by winding a polyurethane copper wire (having a diameter of 0.05 mm, for example) around a paper sleeve of craft paper.

When the polyurethane copper wire of the voice coil **25** receives a voice signal, an electromagnetic power is generated in the voice coil **25** by electric current of the voice signal and a magnetic field of the permanent magnet **22**, thereby to vibrate the bobbin **24**, along with the diaphragm **23**, forward and backward. As a result, a sound corresponding to the voice signal is emitted from the diaphragm **23**. That is to say, the speaker **SP** of this embodiment is an electrodynamic speaker. The speaker has, for example, a diameter of 20 mm to 25 mm, and a thickness of about 4.5 mm.

A rib **11** is formed on an inner surface of the front wall of the housing **A1**, which is opposed to the diaphragm **23**, and is in contact with an end face of a convex portion **21a** that protrudes frontward from an outer peripheral portion of the round support **21**. Thus, the speaker **SP** is fixed to the housing **A1** such that the diaphragm **23** faces the inner surface of the front wall of the housing **A1**.

With the speaker **SP** thus fixed within the housing **A1**, a front air chamber **Bf** is formed as a space enclosed by the inner surface of the front wall of the housing **A1** and a front surface (on which the diaphragm **23** is located) of the speaker **SP**, and a rear air chamber **Br** is formed as a space enclosed by an inner surface of the rear and side walls of the housing **A1** and the rear surface (on which the yoke **20** is located) of the speaker **SP**. The front air chamber **Bf** communicates with outside of the housing **A1** through the sound holes **12** formed in the front wall of the housing **A1**. The rear air chamber **Br** is insulated from (or does not communicate with) the front air chamber **Bf** since the end portion of the support **21** and the rib **11** are attached firmly to each other. Furthermore, the cover **A11** is tightly fitted in the opening formed in the rear wall of the body **A10**, so that the rear air chamber **Br** provides a

hermetically-sealed space that is insulated also from the outside of the housing **A1**. A gasket may be provided between the end portion (or the convex portion **21a**) of the support **21** and the rib **11** so as to increase the degree of adhesion between the support **21** and the rib **11**.

As shown in FIG. 3, the voice processing unit **10** is formed by an integral circuit (IC), which includes a communication unit **10a**, echo canceling units (**10b**, **10c**), an amplification unit **10d** and a signal processing unit **10e**. In operation, the communication unit **10a** receives a voice signal transmitted from an intercom device **A** installed, for example, in another room through the information line **Ls**. The received voice signal is transmitted through the echo canceling unit **10b** to the amplification unit **10d**, in which the signal is amplified, and then output from the speaker **SP**. When the conversation switch **SW1** is operated to the ON position, the voice communication function of the intercom device **A** becomes available. In this condition, voice signals respectively received from a microphone **M1** (first microphone) and a microphone **M2** (second microphone) on the microphone board **MB1** are processed by the signal processing unit **10e**. In the signal processing unit **10e**, each received signal is processed in the manner as described later, and the resulting signal passes through the echo canceling unit **10c**. Then, upon reception of the resulting signal from the echo canceling unit **10c**, the communication unit **10a** transmits the resulting signal through the information line **Ls** to the intercom device **A** installed in, for example, another room. In this manner, the intercom device **A** serves as an interphone that permits two-way voice communications between the rooms. An electric power for the intercom device **A** may be supplied from a wall socket disposed in the vicinity of the location where the intercom device **A** is installed. Alternatively, the electric power may be supplied through the information line **Ls**, or may be supplied from an internal battery accommodated in the intercom device **A**.

As shown in FIG. 4, the microphone board **MB1** has a module board **2** and the microphones **M1** and **M2**, which are mounted on a (top) surface **2a** of the

module board **2**. The microphone **M1** is mainly formed with a bare chip **BC1**, an IC chip **Ka1**, and a shield case **SC1**. Each of the bare chip **BC1**, the IC chip **Ka1**, and a wiring pattern (not shown) on the module board **2** are connected to each other with wires **W** (wire bonding). The shield case **SC1** covers the bare chip **BC1** and the IC chip **Ka1**. The microphone **M2** has the same construction as that of the microphone **M1**. That is to say, the microphone **M2** is formed with a bare chip **BC2**, an IC chip **Ka2**, and a shield case **SC2** mounted on the module board **2** that covers the bare chip **BC2** and the IC chip **Ka2**. The microphones (**M1**, **M2**) are preferred to be formed on a silicon substrate by a micro structure manufacturing process. That is, it is preferred that the microphones (**M1**, **M2**) are IC chips of Micro Electro Mechanical System (MEMS).

Each of the bare chips **BC1** and **BC2**, which will be generally referred to as "bare chip **BC**", provides a condenser-type silicon microphone, as shown in FIG. 5. More specifically, the bare chip BC includes a silicon thin film **1d** that is formed on a (top) surface of a silicon substrate **1b** so as to cover a cavity **1c** formed through the silicon substrate **1b**. An electrode **1f** is formed above the silicon thin film **1d** with an air gap **1e** interposed therebetween, and two pads **1g** are provided on the surface of the silicon substrate **1b** for producing an output in the form of a voice signal. When the silicon thin film **1d** is vibrated by a sound wave from outside, such as a human voice, a capacitance between the silicon thin film **1d** and the electrode **1f** changes, namely, the amount of electric charge changes. With the change in the electric charge amount, electric current proportional to the received sound wave flows from the two pads **1g**. Each of these bare chips **BC1** and **BC2** is fixed by die-bonding the silicon substrate **1b** on the module board **2**, as shown in FIG. 4. The silicon thin film **1d** of the bare chip **BC2** faces the sound hole **F2** formed in the module board **2**.

As shown in FIG. 4, a sound receiving surface of the microphone **M1** is positioned on a surface having the sound hole **F1** of the shield case **SC1**. On the other hand, a sound receiving surface of the microphone **M2** is positioned

on a surface having the sound hole **F2** of the module board **2**. Thus, the sound receiving surface of the microphone **M1** is to receive sound from an upper side of FIG. 4, and the sound receiving surface of the microphone **M2** is to receive sound from an opposite side, i.e., the lower side of FIG. 4. In brief, as shown in FIG. 2, the module board **2** is mounted to the housing **A1** such that the sound receiving surface of the microphone **M1** faces the diaphragm **23** of the speaker **SP**, and the sound receiving surface of the microphone **M2** faces outside of the housing **A1**. Since both of the microphones (**M1**, **M2**) are mounted on the same surface **2a** of the module board **2**, the thickness of the microphone board **MB1** can be advantageously reduced.

FIG. 6A is a simplified plan view of the module board **2**. The module board **2** has a substantially T-shape, which is composed of a rectangular portion **2f**, on which the microphone **M1** is mounted, a rectangular portion **2g**, on which the microphone **M2** is mounted, and a connecting portion **2h** that connects the rectangular portions **2f** and **2g**. The rectangular portion **2g** is formed to be larger than the rectangular portion **2f**. The rectangular portion **2g** has a negative power supply pad **P1**, a positive power supply pad **P2**, and output pads (**P3**, **P4**) on an edge portion of the rectangular portion **2g** such that each of the pads (**P1**, **P2**, **P3**, **P4**) forms a line along an end face of the edge portion of the rectangular portion **2g**.

As shown in FIG. 6B, the negative power supply pad **P1** is connected to a negative terminal of an external power supply, and the positive power supply pad **P2** is connected to a positive terminal of the external power supply. The electric power received from the pads (**P1**, **P2**) is supplied to the microphones (**M1**, **M2**) through the wiring patterns on the module board **2**. A voice signal corresponding to the voice picked up by the microphone **M1** is output from the output pad **P3** via the wiring pattern on the module board **2**. On the other hand, a voice signal corresponding to the voice picked up by the microphone **M2** is output from the output pad **P4** via the wiring pattern on the module board **2**. The negative power supply pad **P1** also serves as a ground pad that establishes

a ground for the voice signals delivered from the output pads (**P3**, **P4**).

As described above, electric powers needed to activate the microphones (**M1**, **M2**) are supplied from the power supply via the common negative and positive power supply pads (**P1**, **P2**). The negative power supply pad **P1** also serves as the ground pad for the outputs of microphones (**M1**, **M2**). Therefore, the number of pads used in the microphone board **MB1** can be reduced, and the construction of the microphone board **MB1** can be simplified.

Next, an operation of the microphone board MB1 will be explained. First, electric currents flowing from the bare chips **BC1** and **BC2** corresponding to the received sound signal are subjected to impedance-conversion, and converted to voltage signals by the IC chips **Ka1** and **Ka2**. Then, the voltage signals are output as voice signals from the output pads (**P3**, **P4**), respectively.

For example, each of the IC chip **Ka1** and the IC chip **Ka2**, which will be generally referred to as "IC chip **Ka**", has a circuit configuration, as shown in FIG. 7. The IC chip **Ka** has a constant-voltage circuit **Kb** formed of a chip IC that converts a power supply voltage +V (e.g., 5V) supplied from the power supply pads (**P1**, **P2**) to a constant voltage Vr (e.g., 12V). The constant-voltage circuit **Kb** applies the constant voltage Vr to a series circuit of a resistance **R11** and the bare chip **BC**. A connection midpoint of the resistance **R11** and the bare chip **BC** is connected to a gate terminal of a junction-type J-FET **S11** through a capacitor **C11**. A drain terminal of the J-FET **S11** is connected to an operating voltage +V, and a source terminal is connected through a resistance **R12** to the negative terminal of the power supply. Herein, the J-FET **S11** is intended for converting electrical impedance, and the voltage of the source terminal of the J-FET **S11** produces an output as the voice signal. It is to be understood that the impedance converting circuit of the IC chip **Ka** is not limited to the above-described configuration, and may be replaced by a circuit having a function of a source follower circuit by an operation amplifier. For example, an amplifying circuit for the voice signal may be provided in the IC chip **Ka**, if necessary.

In the microphone board **MB1**, on which signal transmission and supply of electric power are performed via the wiring patterns on the module board **2** as described above, the signal lines and the power-supply lines can be arranged with high efficiency, and attached to the outer surface of the housing **A1**, as shown in FIG. 1. In this embodiment, as shown in FIG. 1, the surface **2a** of the module board **2** is attached along an outer surface of the front wall of the housing **A1**, and the microphone **M1** is inserted in an opening **13** formed in the front wall of the housing **A1** such that the sound receiving surface of the microphone **M1** faces the front air chamber **Bf**. The diaphragm **23** faces the microphone **M1** through the sound hole **F1** formed in the shield case **SC1**. With this arrangement, the microphone **M1** has a high directivity to the voice emitted from the speaker **SP** via the sound hole **F1**, so that the voice emitted from the speaker **SP** can be efficiently picked up by the microphone **M1**. The microphone **M2** fits in a recess **14** formed in the front wall of the housing **A1**, and faces the outside (or the front) of the housing **A1** through the sound hole **F2** formed in the module board **2**. In this regard, a direction where the second microphone **M2** receives a voice from the outside of the housing **A1** is substantially equal to the direction of emitting a voice from the speaker **SP**. Thus, the microphone **M2** has a high directivity to the voice of a talker who is located in front of the intercom device **A**. When **X1**, **X2** represent distances from the center of the speaker **SP** to the centers of microphones (**M1**, **M2**), respectively, the relationship between the distances (**X1**, **X2**) is expressed as **X1** $<$**X2**.

Since the rear air chamber **Br** facing the rear surface of the speaker **SP** is hermetically-sealed in the housing **A1** as shown in FIG. 2, the voice emitted from the rear surface of the speaker **SP** (rear surface of the diaphragm **23**) is less likely to leak from the rear air chamber **Br**, so that an acoustic coupling between the speaker **SP** and the microphone **M2** is lowered. BY the way, the voice emitted from the rear surface of the speaker **SP** is reversed in phase with respect to the voice emitted from the front surface of the speaker **SP** (front surface of the

diaphragm **23**). Therefore, if the voice emitted from the rear surface of the speaker **SP** goes around to the front of the speaker **SP**, these voices emitted from the front and rear surfaces of the speaker **SP** are cancelled with each other. As a result, the radiating sound pressure of the speaker **SP** is lowered, and it becomes difficult for the talker located in front of the intercom device **A** to hear the voice (sound) emitted from the speaker **SP**. However, since the intercom device of this embodiment is formed to reduce the leaking of the voice emitted from the rear surface of the speaker **SP** to the outside of the housing **A1**, as described above, it is possible to prevent such a reduction in the radiating sound pressure, which is caused by the voice emitted from the rear surface of the speaker **SP** going around to the front of the speaker **SP**.

As shown in FIG. 2, since the recess **14**, in which the microphone **M2** is inserted, is an insulated space not communicated with the rear air chamber **Br**, the microphone **M2** is even less likely to pick up the voice (or sound) emitted from the speaker **SP**. Therefore, the acoustic coupling between the speaker **SP** and the microphone **M2** can be further reduced. With the above arrangement, the microphones (**M1**, **M2**) are adapted to separately pick up the voice emitted from the speaker **SP** and the voice of (generated by) the talker, respectively.

Furthermore, if the microphone board **MB1** is placed in the housing **A1**, it is difficult to maintain the spatial insulation between the front air chamber **Bf** and the rear air chamber **Br**. In this embodiment, however, the microphone board **MB1** is attached to the outer surface of the housing **A1**. Therefore, it is possible to maintain the spatial insulation between the front air chamber **Bf** and the rear air chamber **Br**.

As described below, the intercom device of this embodiment has a configuration to prevent a howling phenomenon caused when the microphones (**M1**, **M2**) pick up the voice emitted from the speaker **SP**.

As shown in FIG. 8, the signal processing unit **10e** of the voice processing unit **10** includes an amplifying circuit **30** for amplifying the output of the microphone **M1**, a band-pass filter **31** for removing noises of frequencies other

than a voice-band (300 to 4000 Hz) from the output of the amplifying circuit **30**, and a delaying circuit **32** for delaying the output of the band-pass filter **31**. The signal processing unit **10e** further includes an amplifying circuit **33** for amplifying the output of the microphone **M2** while reversing the phase of the output of the microphone **M2** by 180 degrees, a band-pass filter **34** for removing noises of frequencies other than the voice-band (300 to 4000 Hz) from the output of the amplifying circuit **33**, and an adder circuit **35** for adding the outputs of the delaying circuit **32** and the band-pass filter **34**.

Referring to FIG. 8 to FIG. 12C, operations of the signal processing unit **10e** will be explained below. First, an operation of the signal processing unit **10e** when the voice emitted from the speaker **SP** is picked up by the microphones (**M1**, **M2**) will be explained. FIGS. 9A and 9B, 10A and 10B, 11A and 11B, and 12A to 12C show voice signal waveforms that appear at respective units of the signal processing unit **10e** when the microphones (**M1**, **M2**) respectively pick up the voice emitted from the speaker **SP**. As mentioned above, the distance **X1** from the center of the speaker **SP** to the center of the microphone **M1** is smaller than the distance **X2** from the center of the speaker **SP** to the center of the microphone **M2**. Accordingly, as shown in FIGS. 9A and 9B, when the microphones (**M1**, **M2**) pick up the voice emitted from the speaker **SP**, the amplitude of the output **Y21** of the microphone **M2** is smaller than that of the output **Y11** of the microphone **M1** due to the difference between the distances (**X1**, **X2**). In addition, the phase of the output **Y21** of the microphone **M2** is delayed with respect to that of the output **Y11** of the microphone **M1** by a delay time **Td** (=(**X2**-**X1**)/**Vs**, where **Vs** is sonic velocity), which is proportional to the difference between the distances (**X2**-**X1**).

The amplifying circuit **30** amplifies the output **Y11** to produce an output **Y12**, and the amplifying circuit **33** amplifies the output **Y21** while reversing the phase thereof by 180 degrees, to produce an output **Y22**. At this time, each of the amplifying circuits (**30**, **33**) performs level adjustment based on the difference between the distances (**X2**-**X1**), so that the output levels of the

microphones (**M1**, **M2**) relative to the voice emitted from the speaker **SP** become equal to each other, as shown in FIGS. 10A and 10B. In this embodiment, since an amplitude rate of the amplifying circuit **30** is set to substantially 1, the amplifying circuit **30** may be eliminated from the signal processing unit **10e**.

As shown in FIGS. 11A and 11B, the band-pass filters (**31**, **34**) remove frequencies other than the voice band from the outputs **Y12** and **Y22**, to produce outputs **Y13** and **Y23**, respectively.

The delaying circuit **32** can be formed by, for example, a time delay device or a CR phase delaying circuit. The delaying circuit **32** delays the output of the microphone **M1**, which is located closer to the speaker **SP**, by the above-described delay time **Td**, thereby producing an output **Y14** whose phase coincides with the phase of the output **Y23** of the band-pass filter **34**. Thus, it is possible to reduce noises contained in a voice signal to be transmitted.

Through the amplifying and delaying processes described above, a voice component emitted from the speaker **SP**, which is contained in the output **Y14**, and the voice component emitted from the speaker **SP**, which is contained in the output **Y23**, have the same amplitude and the reversed phase relative to each other, as shown in FIG. 12A and FIG. 12B. The adder circuit **35** adds the output **Y14** and the output **Y23** to cancel the voice component of the speaker **SP** from the output of the microphone **M2**. Thus, it is possible to produce an output **Ya** shown in FIG. 12C, in which the voice component from the speaker **SP** is removed or reduced. As described above, the signal processing unit **10e** removes the voice component of the speaker **SP** from the output signal **Y21** of the microphone **M2** by using the output signal **Y11** of the microphone **M1**.

On the other hand, an operation of the signal processing unit **10e** when the voice emitted from the talker **H** in front of the intercom device **A** is picked up by the microphones (**M1**, **M2**) will be described below. Since the microphone **M2** has a high directivity to the voice from the talker **H** located in front of the intercom device **A**, the amplitude of the output **Y21** of the microphone **M2** is larger than that of the output **Y11** of the microphone **M1**. Furthermore, since

the signal from the microphone **M1** is not amplified while the signal from the microphone **M2** is amplified in the amplifying circuit **33**, the voice component from the talker **H** contained in the output **Y23** is sufficiently larger than the voice component from the talker **H** contained in the output **Y14**. Thus, even after the above-described adding process is performed in the adder circuit **35**, the signal corresponding to the voice generated by the talker **H** remains in the output **Ya** with a sufficiently large amplitude.

In the manner described above, the voice component from the speaker **SP** is reduced in the output **Ya** of the adder circuit **35**, whereas the voice component generated by the talker **H** in front of the intercom device **A** remains in the output **Ya**. Thus, a relative difference in the output **Ya** between the voice component from the talker **H**, which is desired to be saved, and the voice component from the speaker **SP**, which is desired to be reduced, becomes larger. That is to say, even when the voice from the talker **H** and the voice from the speaker **SP** are simultaneously generated, only the voice component from the speaker **SP** is reduced while the voice component from the talker **H** remains sufficiently large. Accordingly, it is possible to prevent the howling phenomenon caused when the microphones (**M1**, **M2**) pick up the voice output of the speaker **SP**.

Since the microphone **M1** is disposed in the front air chamber **Bf** so as to face the diaphragm **23**, as shown in FIG. 2, the microphone **M1** efficiently picks up the voice emitted from the speaker **SP**. Therefore, the signal processing unit **10e** can efficiently achieve the howling phenomenon prevention function. Furthermore, since the sound receiving surface of the microphone **M2** faces outside of the housing **A1**, and the output direction of the speaker **SP** and the directivity of the microphone **M2** substantially conform to each other, the acoustic coupling between the speaker **SP** and the microphone **M2** is reduced. Thus, the microphone **M2** is less likely to pick up the voice emitted from the speaker **SP**, and can be placed close to the speaker **SP**, that is to say, close to the front air chamber **Bf**, which makes it possible to downsize the intercom device **A**.

On the other hand, the signal processing unit **10e** transmits the voice signal to the echo canceling unit **10c** through a digital-analog converter (not shown), and the echo canceling unit (**10b**, **10c**) shown in FIG. 3 perform the process as described below to further efficiently prevent the howling phenomenon.

First, the echo canceling unit **10c** retrieves the output of the echo canceling unit **10b** as a reference signal, and performs a suitable computation on the output of the signal processing unit **10e**, thereby canceling a voice signal, which goes around from the speaker **SP** to the microphones (**M1**, **M2**). On the other hand, the echo canceling unit **10b** also retrieves the output of the echo canceling unit **10c** as the reference signal, and performs a suitable computation on the output of the communication unit **10a**, thereby canceling a voice signal, which goes around from a speaker of an intercom device installed in another room to its microphones.

More specifically, the echo canceling units (**10b**, **10c**) control the amount of loss at variable loss means (not shown) provided in a loop circuit consisting of the speaker **SP**, the microphones (**M1**, **M2**), the signal processing unit **10e**, the echo canceling unit **10c**, the communication unit **10a**, the echo canceling unit **10b**, the amplifying unit **10d** and the speaker **SP**, so as to prevent the howling phenomenon by making the loop gain less than 1. Herein, one of the transmitted voice signal and the received voice signal, which has a smaller signal level, is regarded as not important, and a transmission loss of a variable loss circuit, which is inserted in a transmission path of the signal having the smaller level, is increased. In this embodiment, the intercom devices transmit signals via the information line **Ls**. However, the transmission of signals is not limited to a wired line. A non-wired transmission manner such as radio wave, optical communication system, and infrared radiation may be alternatively employed.

(Second Embodiment)

As shown in FIG. 13, an intercom device **A'** according to a second

embodiment of the present invention includes an acoustic tube **40**, which is provided in the rear air chamber **Br** having a small volume, so as to improve the sound quality and electro-acoustic conversion efficiency of the speaker **SP**. The intercom device **A'** according to the second embodiment will be explained below.

The acoustic tube **40** has an opening at its one end and a closed end at the other end, and is provided in the rear air chamber **Br** to be along an inner wall surface(s) of the rear air chamber **Br**. By setting the resonance frequency of the acoustic tube **40** to an appropriate level, the lowest resonance frequency **fo** of the speaker **SP**, which is provided in the intercom device **A'**, is shifted to a lower frequency level, and the sound pressure level of the speaker **SP** increases, resulting in improvements of the sound quality and electro-acoustic conversion efficiency of the speaker **SP**. It is noted that except for the sound tube **40**, the construction of the intercom device **A'** shown in FIG. 13 is the same as that of the intercom device **A** shown in FIG. 2. Accordingly, details of the intercom device **A'** will be omitted herein.

The lowest resonance frequency **fo** of the speaker **SP** is generally determined by an equivalent mass (diaphragm, coil, additional mass of air) **Mo** of a vibration system of the speaker, stiffness **So** of an edge or the like supporting the vibration system, and stiffness **Sc** of air in the rear air chamber **Br**. That is, it is represented by an equation:

$$fo = \{1/(2\pi)\} \cdot \sqrt{\{(So+Sc)/Mo\}}$$

FIG. 14A through FIG. 14C show another housing **A2** of the intercom device **A'** according to a modification of the present embodiment. The housing **A2** accommodates the acoustic tube **40** therein. Effects of the sound tube **40** are explained below referring to the housing **A2** as an example. The housing **A2** essentially consists of a body **A20** having opening at its rear surface, and a cover **A21** for closing the opening of the body **A20**. In FIG. 14B, only the body **A20** is illustrated. As in the intercom device **A** of the first embodiment, the housing **A2** is capable of accommodating a speaker **SP**, a microphone board **MB1**, a

conversation switch **SW1**, and a voice processing unit **10**. The acoustic tube **40** is provided in the body **A20**.

The housing **A2** has two spaces **A2a** and **A2b** therein, which are insulated from each other. The space **A2a** has a lateral dimension of 30 mm, a vertical dimension of 40 mm, and a thickness of 8 mm, to accommodate the speaker **SP** therein. In the space **A2a**, columnar bosses **51** are formed at regularly spaced four locations on the inner surface of the front wall of the body **A20**, and screw holes **52** are respectively formed in the bosses **51**. The speaker **SP** is fixed to the inner surface of the front wall of the body **A20** by using mounting screws, which are thread engageable with the screw holes **52**. With the construction described above, a front air chamber **Bf** and a rear air chamber **Br** similar to those of the first embodiment (FIG. 2) are formed. The front air chamber **Br** is provided by a space enclosed by the inner surface of the front wall of the housing **A2** and the front surface side (diaphragm **23** side) of the speaker **SP**. The rear air chamber **Br** is provided by a space enclosed by the inner surface of the rear and side walls of the housing **A2** and the rear surface side (yoke **20** side) of the speaker **SP**. A plurality of sound holes **12** are formed in the front wall of the housing **A2**.

As shown in FIGS. 14A through 14C, the acoustic tube **40**, which has a rectangular shape in cross section, is formed along the inner walls of the rear air chamber **Br** at the back of the speaker **SP** over substantially three-quarter of the circumference of the rear air chamber **Br**. The acoustic tube **40** has an opening at one end **40a** and a closed end at the other end **40b**. The overall length **L** of the acoustic tube **40** is set to an odd multiple of one quarter of a wavelength, which is determined by the sonic velocity and a frequency **f** of which sound pressure level is desired to be increased. In this embodiment, the acoustic tube **40** has a length of one quarter of the wavelength determined by the frequency where an increase in sound pressure level of the audio output of the speaker **SP** is needed and the sonic velocity. In detail, since an open-end correction is performed on the acoustic tube **40**, where sonic velocity is set to C

and an open-end correction value is set to σ (=0.8**d**, wherein **d** is an open-end diameter of the acoustic tube **40**), the overall length is determined from an equation described below:

$$f = [C / \{4 (L+\sigma)\}]$$

In this embodiment, the overall length L of the acoustic tube **40** is set 95.2 mm and a cross-sectional area of the acoustic tube **40** is set to 4.2 mm^2 (equivalent to $\phi = 2.3$). By forming the acoustic tube **40** along the inner walls of the rear air chamber **Br**, it is possible to minimize a volume reduction of the rear air chamber **Br** caused by the formation of the acoustic tube **40**. In addition, with the acoustic tube **40** thus provided, the sound pressure level of the speaker **SP** can be increased in a desired frequency band (at around 700 Hz in this embodiment). As a result, it is possible to improve the sound quality and the electro-acoustic conversion efficiency of the speaker **SP**.

By utilizing the fact that input impedance becomes extremely small at the resonance frequency (i.e., a frequency at which the overall length of the tube is equal to an odd-number multiple of an approximately quarter (1/4) wavelength) of a closed tube, the acoustic tube **40** can reduce the sound leakage from the rear air chamber **Br** to the outside of the housing. That is, the sound waves radiated from the rear surface of the speaker **SP** is reflected toward the rear surface side of the speaker **SP**.

Referring to FIG. 15 through FIG. 18, effects of the acoustic tube **40** provided in the housing **A2** will be explained below in comparison with the intercom device **A** of the first embodiment, which has no acoustic tube in the housing **A1**, and an intercom device having an infinite baffle, which will be explained in detail below.

FIG. 15 shows frequency characteristics of the cancelled amount of the voice component from the speaker **SP** in the case of using the signal processing unit **10e** of intercom device **A** of the first embodiment, and FIG. 16 shows frequency characteristics of the radiating sound pressure measured in front of the speaker **SP**. FIG. 15 and FIG. 16 also show results obtained in the case

where the housing has no acoustic tube as in the first embodiment, and the case where an ideal baffle plate is used as the housing.

FIG. 17 shows a baffle plate **C** having an infinite size as an example of the above-mentioned ideal baffle plate, but the baffle plate **C** is not realistic for mounting a real speaker thereon. Thus, in this embodiment, a standard baffle as specified in JIS C 5532 is used to measure a frequency equivalent to a lowest resonant frequency of the speaker **SP** that is supposed to be mounted on the baffle plate **C**. The lowest resonant frequency of the speaker **SP** mounted on the standard baffle is equivalent to that of the speaker **SP** that is supposed to be mounted on the baffle plate **C** (an ideal baffle plate). The cancelled amount (**Y1a** in FIG. 15) in the case where the speaker **SP** and the microphone board **MB1** are mounted on the standard baffle to form an intercom device is kept at 10 dB or more in a frequency band of 100 Hz to 10000 Hz, and the radiating sound pressure characteristics (**Y2a** in FIG. 16) in the same case indicate that the lowest resonance frequency **fo1** of the speaker **SP** is 600 Hz, which is an equally ideal characteristic as provided by the speaker **SP** alone. A standard closed box as specified in JIS C 5532 may be used to measure a frequency equivalent to the lowest resonant frequency of the speaker **SP** that is supposed to be mounted on the baffle plate **C** having an infinite size (an ideal baffle plate) as shown in Fig. 17.

In the case where a rear air chamber having no acoustic tube **40** is used, as in the rear air chamber **Br** of the first embodiment, the cancelled amount (**Y1b** in FIG. 15) of the voice component from the speaker **SP** is kept at 10 dB or more in the frequency band of 100 Hz to 10000 Hz. Thus, substantially the same cancelled amount as that of the case where the standard baffle is used can be obtained. However, in the case where the rear air chamber **Br** is formed as an enclosed space, radiating sound pressure characteristics comparable to those of the case where the standard baffle is used can be obtained when the volume of the rear air chamber **Br** is sufficiently larger. On the other hand, when the volume of the rear air chamber **Br** is small, the mechanical equivalent stiffness

of the rear air chamber **Br** increases, and the lowest resonance frequency of the speaker **SP** becomes higher. As higher the lowest resonance frequency, the radiating sound pressure reduces, so that the sound quality and the electro-acoustic conversion efficiency of the speaker **SP** deteriorate. For example, the radiating sound pressure characteristics (**Y2b** in FIG. 16) of the case where the rear air chamber **Br** of the first embodiment is used indicate that the lowest resonance frequency **fo2** of the speaker **SP** is 1200 Hz, which is noticeably higher than the lowest resonance frequency obtained when the standard baffle is used. Furthermore, the sound pressure level in a frequency band below 800 Hz is lower by about 5 to 20 dB than that of the case where the standard baffle is used. It means that the sound quality and the electro-acoustic conversion efficiency of the speaker **SP** are deteriorated in this frequency band.

The radiating sound pressure characteristics (**Y2c** in FIG. 16) of the case where the volume of the rear air chamber **Br** is made three times larger than that of the rear air chamber **Br** of the first embodiment indicate that the lowest resonance frequency **fo3** of the speaker **SP** is 800 Hz, and the sound quality of the speaker **SP** is improved as compared with the case of using the housing **A1** of the first embodiment. However, the overall size of the housing increases with the increase in volume of the rear air chamber **Br**, which makes it difficult to downsize the intercom device. Thus, in this embodiment, by providing the acoustic tube **40** in the rear air chamber **Br** having a small volume, the sound quality and the electro-acoustic conversion efficiency of the speaker **SP** are improved, while making the intercom device smaller.

The overall length **L** of the acoustic tube **40** is set so that the resonance frequency of the acoustic tube **40** is located between the lowest resonance frequency **fo2** (=1200 Hz) of the case where the rear air chamber **Br** having no acoustic tube **40** is used, and the lowest resonance frequency **fo1** (=600 Hz) of the case where the standard baffle is used. By setting the overall length **L** of the acoustic tube **40**, as described above, the lowest resonance frequency **fo5** of

the speaker **SP** in the case where the housing **A2** having the acoustic tube **40** is used is set to about 700 Hz between **fo1** and **fo2**.

FIG. 18 shows frequency characteristics of the radiating sound pressure of the intercom device **A'**, which is provided with the housing **A2** and the sound tube **40** of the present embodiment, and the radiating sound pressure of the intercom device **A** of the first embodiment, which has no acoustic tube. The characteristics **Y3a** of the intercom device **A'** in FIG. 18 indicate that the sound pressure level in a frequency band not larger than 800 Hz is increased by about 5 to 10 dB, as compared with the characteristics **Y3b** of the intercom device **A**, and the lowest resonance frequency is reduced from **fo4** (=1200 Hz) to **fo5** (= about 700 Hz). That is, improvements in the speaker efficiency and the sound quality can be achieved. For example, the sound pressure level at around 700 Hz is increased by about 10 dB, and the improvement of the speaker efficiency brought by the increase of the sound pressure level corresponds to a 70% reduction in input current and a 90% reduction in input power of the speaker **SP**. Thus, the sound pressure level of the speaker **SP** increases around the lowest resonance frequency of the speaker **SP**. Therefore, by setting the overall length of the acoustic tube **40** to an appropriate length, and shifting the lowest resonance frequency of the speaker **SP** to a lower frequency side, it is possible to achieve improvements in sound quality and electro-acoustic conversion efficiency of the speaker **SP**.

As another preferred acoustic tube of the present invention, an acoustic tube **41** that communicates with the rear air chamber **Br** may be provided at the outside of a housing **A3**. For example, the acoustic tube **41** can be obtained by attaching a tubular member to into a hole (not shown) formed in the housing **A3**, as schematically shown in FIG. 19. When tubular member is made of urethane or the like, the overall length **L** of the acoustic tube **41** can be easily changed, if desired.

As a further modification of this embodiment, two or more acoustic tubes may be formed in the rear air chamber **Br**. As an example, three acoustic tubes

42, **43**, **44** can be formed in the rear air chamber **Br** of a housing **A4**, as schematically shown in FIG. 20. In this example, the acoustic tubes (**42**, **43**, **44**) have different lengths from each other. This means that they have mutually different resonance frequencies. Accordingly, the sound pressure level of the speaker **SP** is increased at a plurality of frequencies. By setting the overall lengths of the respective acoustic tubes (**42**, **43**, **44**) to appropriate values, it is possible to provide desired sound quality and electro-acoustic conversion efficiency of the speaker **SP**.

As shown by a dotted line in FIG. 21, a sound absorbing material **45** such as a nonwoven fabric may be provided at or in the vicinity of the open end **40a** of the acoustic tube **40**. By use of the sound absorbing material **45**, it is possible to provide an enhanced sound absorbing effect, and perform a fine adjustment of the resonance frequency of the acoustic tube **40**.

(Third Embodiment)

The present embodiment provides a wiring system using an intercom device **101** of the present invention. That is, this wiring system enables power and signal transmissions between electric devices spaced away from each other in a building structure.

As shown in FIG. 22, the wiring system comprises a power line **Lp** and an information line **Ls** installed in the building structure, which are connected to a commercial power source **AC** and the Internet network **NT** through a distribution board, switch boxes **102** embedded in wall surfaces of the building structure, base units **103** mounted on the switch boxes, and connected to the power line **Lp** and the information line **Ls**, and function units **104**, each of which is detachably connected to a desired one of the base units **103**, and has at least one of functions for supplying electric power from the power line **Lp**, outputting information from the information line **Ls**, and inputting information into the information line **Ls** in a connected state with the desired base unit. An intercom device **101** of the present invention, which is preferably provided by

the intercom device (**A**, **A'**) of the first or second embodiment described above, can be regarded as one of the function units **104**. In the present description, the wall surfaces of the building structure are not limited to surfaces of sidewalls standing between adjacent rooms. That is, the meaning of the wall surfaces includes outdoor and indoor wall surfaces, and the indoor wall surfaces comprise ceiling and floor surfaces as well as the sidewall surfaces. In the drawings, **MB** designates a main circuit breaker, **BB** designates a branch circuit breaker, and **GW** designates a gateway (e.g., a router or a built-in hub).

As shown in FIGS. 23 and 24, each of the base units **103** is formed with a gate housing **130** having terminals (**131a**, **132a**, **131b**, **132b**) connected to the power supply line **Lp** and the information line **Ls**, and a main housing **135** detachably connected to the function unit **104**. Theses housings can be made of a synthetic resin having electrical insulating characteristics (e.g., a non-crystalline plastic such as ABS resin). The terminals (**131b**, **132b**) are used to extend the wirings. The gate housing **130** and the main housing **135** have a pair of a module port **134** and a module connector **142**, which are detachably connected to each other to simultaneously establish both of supplying the electric power from the gate housing **130** to the main housing **135**, and making a signal transmission therebetween. In addition, when the function unit **104** has the module connector **142**, it can be connected to the gate housing **130** in place of the main housing **135**. If necessary, the gate housing **130** and main housing **135** may be integrally formed.

Circuit components of the base unit **103** are designed to transmit the electric power and the information signal to the function unit **104**. For example, the base unit **103** is provided with an AC/AC converter **160**, DC power section **161**, transceiver section **162**, E/O converter **163**, O/E converter **165**, and a function portion **167**.

The AC/AC converter **160** converts commercial AC voltage to a lower AC voltage having an increased frequency, and applies the lower AC voltage to a coil

172. The DC power section **161** generates an operating voltage of the internal circuit components from a stable DC voltage obtained by rectifying and smoothing the lower AC voltage. The transceiver section **162** transmits and receives the information signal for enabling the mutual communication through the information line **Ls**. The E/O converter **163** converts the information signal received from the information line **Ls** to an optical signal, and outputs the optical signal though a light emitting device (LED) **164**. On the other hand, the O/E converter **165** receives the optical signal provided from the outside, e.g., the function unit **104** by a light receiving device (PD) **166**, and converting the received optical signal into the information signal to transmit it to the transceiver section **162**. The base unit **103** shown in FIG. 23 has an outlet tap as the function portion **167**. Alternatively, a sensor device or a controller may be formed as the function portion **167**.

The module port **134** formed at a front surface of the gate housing **130** is composed of an electric power port **134a** for supplying the electric power and an information signal port **134b** for accessing the information line **Ls**, as shown in FIG. 25B. Arrangement and shapes of the electric power port **134a** and the information signal port **134b** are standardized (normalized or stylized) in the wiring system of this embodiment. For example, as shown in FIG. 25B, each of the electric power port **134a** and the information signal port **134b** is configured in a substantially rectangular shape such that they are arranged in parallel to each other.

On the other hand, the module connector **142** formed at a rear surface of the main housing **135** is composed of an electric power connector **142a** and an information signal connector **142b**, as shown in FIG. 25A. Arrangement and shapes of the electric power connector **142a** and the information signal connector **142b** are standardized (normalized or stylized) in the wiring system of this embodiment. For example, as shown in FIG. 25A, each of the electric power connector **142a** and the information signal connector **142b** is configured

in a substantially rectangular shape such that they are arranged in parallel to each other.

In this embodiment, the module port **134** has a guide portion **133** such as a ring-like wall or a ring-like groove extending around the electric power port **134a** and the information signal port **134b**. This guide portion **133** is formed to be engageable to an engaging portion **145** such as a ring-like wall of the main housing **135**, which is formed around the electric power connector **142a** and the information signal connector **142b**. By simply engaging the engaging portion **145** to the guide portion **133**, the electric power connector **142a** and the information signal connector **142b** can be simultaneously connected to the electric power port **134a** and the information signal port **134b**. Therefore, there advantages that the main housing **135** can be easily attached to or removed from the gate housing **130** with improved connection reliability. A module connector having the same structure described above may be formed in the function unit **104**. The module port **134** and the module connector **142** may be formed by female and male connectors.

In the present embodiment, the gate housing **130** is directly fixed to the switch box **102**. If necessary, the gate housing **130** may be fixed to the switch box **102** through an attachment member **175**, as shown in FIG. 26. For example, this attachment member **175** is formed by a rectangular frame **176** with a window hole **177**. In this case, the gate housing **130** is mounted on the attachment member **175** by fitting a front portion of the gate housing into the window hole **177**, and then engaging hooks (not shown) formed at both sides of the gate housing to the rectangular frame **176**. The attachment member **175** mounting the gate housing **130** thereon is fixed to the switch box **102** by use of fixing screws. Alternatively, the gate housing **130** may be directly fixed to the wall surface by use of exclusive clamps without using the switch box **102**.

The function unit **104** is designed to provide a desired function by using the electric power supplied to the function unit **104** through the base unit **103** or

using the mutual communication of the information signal between the function unit **104** and the information line **Ls** through the base unit **103**. For example, when the function unit **104** is connected to the base unit **103** mounted in the wall surface at a high position near the ceiling, it preferably has a receptacle function of receiving a plug with hook of a lighting apparatus, security function such as a motion sensor, temperature sensor, and monitoring camera, or a sound function such as speaker. In addition, when the function unit **104** is connected to the base unit **103** mounted in the wall surface at a middle height, at which the function unit **104** can be easily operated by the user, it preferably has a switch function of turning on/off the lighting apparatus, control function for electric appliances such as air-conditioning equipments, a display function such as liquid crystal display or a timer function. The intercom device of the present invention is preferably connected as one of the function units **104** to the base unit **103** at the middle height. Moreover, when the function unit **104** is connected to the base unit **103** mounted in the wall surface at a low position near the floor, it preferably has a receptacle function for receiving a plug of an electric appliance such as electric vacuum cleaner, the sound function such as speaker, or a footlight function.

Specifically, as shown in FIG. 27, when a function section **181** of the function unit **104** is formed by a switch, operation data obtained by operating the switch is transmitted to a processing section **188** such as CPU through an I/O interface **189**. Then the processed data is sent to, for example, an infrared remote controller (not shown) through a transceiver section **187**, so that an electric appliance to be controlled is turned on/off by receiving a remote control signal emitted from the infrared remote controller. In addition, when the function section **181** is formed by a sensor, data detected by the sensor is transmitted as the information signal to the information line **Ls**, and then informed to the user by a required communicator.

The coil **172** wound around a core **170** of the base unit **103** shown in FIG.

23 is used as a power transmission means for supplying electric power to the function unit **104** in a noncontact manner. That is, the coil **172** of the base unit **103** provides an electromagnetic coupling portion that works as a first side of a transformer. On the other hand, as shown in FIG. 27, the function unit **104** has an electromagnetic coupling portion comprised of a coil **182** wound around a core **180**, which works as a second side of the transformer. Therefore, by forming electromagnetic coupling between the base unit **103** and the function unit **104**, a low AC voltage can be induced in the coil **182** of the function unit **104** to supply of electric power from the base unit **103** to the function unit **104**. Similarly, an electromagnetic coupling portion **X2** is formed at the opposite side of the electromagnetic coupling portion **X1**, and used to make a connection between the function units. In this embodiment, since the low AC voltage having the higher frequency than the commercial AC voltage is obtained by the AC/AC converter **160**, there is an advantage that the electromagnetic coupling portions used as the transformer can be downsized.

In addition, the light emitting device (LED) **164** of the base unit **103** shown in FIG. 23 is used to transmit an optical signal as the information signal to the function unit **104** in a noncontact manner. On the other hand, as shown in FIG. 27, a light receiving device (PD) **186** is disposed in the function unit 10 such that the light emitting device **164** is in a face-to-face relation with the light receiving device **186** when the function unit **104** is connected to the base unit **103**. Similarly, to transmit the optical signal as the information signal from the function unit **104** to the base unit **103**, the function unit **104** has a light emitting device (LED) **184**, which is disposed in the face-to-face relation with the light receiving element (PD) **166** when the function unit **104** is connected to the base unit **103**. Thus, each of the base unit **103** and the function unit **104** has the pair of the E/O converter (**163**, **183**) and the O/E converter (**165**, **185**) as an optical coupling portion to enable the mutual communication of the information signal therebetween in a noncontact manner.

As shown in FIG. 23, the electromagnetic coupling portion **X** used to supply the electric power and the optical coupling portion **Y** used for the mutual communication of the information signal are disposed at a side surface of the base unit **103** to be spaced from each other by a constant distance. Arrangements and shapes of the electromagnetic coupling portion **X** and the optical coupling portion **Y** are standardized (or stylized) to make the base unit **103** shareable among the function units **104**. In addition, the pair of the electromagnetic coupling portion **X** and the optical coupling portion **Y** may be formed at each of both sides of the function unit **104**, as shown in FIG. 27. That is, the optical coupling portion **Y1** formed at one side (e.g., left side) of the function unit **104** is composed of the light receiving device **186** located at the upper side and the light emitting device **184** located at the lower side, and the optical coupling portion **Y2** formed at the opposite side (e.g., right side) of the function unit **104** is composed of a light emitting device **194** located at the upper side and a light receiving device **196** located at the lower side. In this case, since the base unit **103** is connected to the one side of the function unit **104**, and another function unit or the intercom device of the present invention can be connected to the other side of the function unit **104**, there is an advantage of further improving function expandability of the wiring system.

The intercom device **101** of this embodiment can be connected to the base unit **103**, as shown in FIG. 28. In addition, internal components of the intercom device **101** of this embodiment are shown in FIG. 29. As understood from the comparison with the function unit **104** of FIG. 27, the intercom device **101** of FIG. 29 is substantially equivalent to a function unit having the components of the intercom device shown in FIG. 2 as the function section **181**.

Since this intercom device **101** has the stylized connectors as well as the function unit **104** and base unit **103** described above, it can be used in a desired combination with the base unit **103** and the function unit **104**. For example, as shown by the arrow ① in FIG. 30, the intercom device **101** can be

connected to the function unit **104** having the timer function, which is connected to the base unit **103**. Alternatively, as shown by the arrow ② in FIG 30, the intercom device **101** may be connected at its one side to the base unit **103**. At this time, the function unit **104** having the timer function can be connected to the other side of the intercom device **101**. Thus, there is a remarkable advantage that a layout of the function units **104** having various functions and the intercom device **101** can be changed with a high degree of freedom according to the user's needs.

In addition, when the intercom device **101** is operated in conjunction with at least one functions unit **104**, the wiring system of this embodiment can provide a higher order function in addition to the communication function. For example, when the intercom device **101** is connected to one of the base units **103**, and the function unit **104** having a sensor function (e.g., motion sensor or fire sensor) is connected to another one of the base units **103**, which is disposed at a proper location, the intercom device **101** can output an alarm sound by receiving alarm signal from the function unit. Therefore, the intercom device **101** of this case has an alarm generating function for disaster or crime prevention as well as the communication function. In addition, since it is not needed to individually prepare a speaker for generating the alarm sound, there is a further advantage of achieving an improvement in cost performance of the wiring system as a whole.

Next, means for mechanically coupling between the base unit **103** and the function unit **104** (or the intercom device **101**) is explained.

For example, the base unit **103** shown in FIG. 31 has a cosmetic cover **112**, which is detachably attached to a front surface of the gate housing **135**, and horizontal guide rails **114** formed at upper and lower portions of the gate housing. The numeral **115** designates a stopper wall formed in a substantially center position in the longitudinal direction of the guide rail **114**. On the other hand, as shown in FIGS. 32A to 32C, the function unit **104** (or the intercom

device **101**) has concave portions **128** for accommodating joining members **150** therein, which are formed at both of its upper and lower sides, horizontal guide rails **124** extending in the concave portions **128**, and cover members **126** pivotally supported at upper and lower ends of the function unit **104**. In addition, as shown in FIG. 32D, the joining member **150** has a groove **152**, in which the horizontal guide rail (**114**, **124**) can be slidably fitted.

In this case, a cosmetic cover **112** is firstly removed from the gate housing **135** of the base unit **103**. Then, the electromagnetic coupling portion **X** and the optical coupling portion **Y** of the base unit **103** are connected to the electromagnetic coupling portion **X1** and the optical coupling portion **Y1** of the function unit **104** (or the intercom device **101**). Next, the cover member **126** is opened, and the joining member **150** is slid along the guide rail **124**, as shown by an arrow in FIG. 33. The slide movement of the joining member **150** is performed until the joining member **150** contacts the stopper wall **115**. As a result, the joining member **150** is engaged to the guide rail **114** of the base unit **103** over about half length of the joining member **150**. On the other hand, the remaining portion of the joining member **150** is still engaged to the guide rail **124** of the function unit **104** (or the intercom device **101**). Thus, by engaging the joining member **150** to both of the base unit **103** and the function unit **104** the base unit **103** can be mechanically coupled with the function unit **104** (or the intercom device **101**). Finally, the cosmetic cover **112** is attached to the base unit **103**, and the cover member **126** is closed.

According to this coupling means, since the joining member **150** is held between the cosmetic cover **112** and the gate housing **135**, it is possible to prevent accidental falling of the joining member **150**, and obtain the stable mechanical connection therebetween without spoiling the beauty of them. In addition, since the joining member **150** is always accommodated in the concave portion **128** of the function unit **104** (or the intercom device **101**), there is no worry about loss of the joining member **150**. This coupling means is also usual

in the case of mechanically coupling between the function units **104**, or between the function unit **104** and the intercom device **101**.

As an information-signal transmitting method available in the wiring system, a baseband transmission or a broadband transmission can be used. In addition, the protocol is not limited to a specific one. For example, sound and visual information signals may be transmitted and received according to JT-H232 packet to make the mutual communication between the intercom devices. In a control system, it is also preferred to use a routing protocol for a broadcast or a unicast that controlling can be performed at a control ratio of 1:1 or 1:N according to operation data. Alternatively, the protocol used between the base units may be different from the protocol used in the function unit or the intercom device connected to the base unit. In this case, a protocol conversion is preferably performed by the base unit.

In addition, the wiring system of the above embodiment is essential to use both of the information line and the power line installed in the building structure. As a modification of this embodiment, a power-line carrier communication may be used in the wiring system.

That is, the wiring system according to this modification is essential to install only the power line in the building structure. Therefore, the base unit is connected to only the power line. On the other hand, the base unit, the function unit and/or the intercom device need to have a transceiver unit for transmitting and receiving information signals in the power-line carrier communication manner. For example, when the transceiver unit is formed in the base unit, the electric power carrying the information signals thereon is separated to the electric power transmission and the information signal transmission by the base unit. Therefore, the same function unit and the intercom device used in the wiring system of the above embodiment are also available in this power-line carrier type wiring system.

Alternatively, the transceiver unit may be formed in the intercom device **101**, as shown in FIG. 34. In this case, the intercom device is detachably

connected to the base unit and/or the function unit by a power-line carrier type connector **Z**. In addition, the intercom device also comprises a PLC modem **210** for receiving the information signals carried by the power line, and transmitting information signals through the power line, a processing section **220** connected to the PLC modem, an I/O interface **230** between the processing section, and a function section **240** for providing the intercom function. Similar circuit components may be formed in the function unit and/or the base unit. As a modulation method for the power-line carrier communication, for example, it is possible to use a broad spectrum diffusion model, multi carrier model or an orthogonal Frequency division multiplexing (OFDM) model.

INDUSTRIAL APPLICABILITY

Thus, according to the present invention, the first microphone disposed in the front air chamber efficiently receives the voice output of the speaker, and the output of the first microphone is effectively used in the signal processing unit to remove the signal component corresponding to the voice output of the speaker from the output of the second microphone. Therefore, it is possible to improve the howling prevention effect. In addition, the intercom device can be downsized because a distance between the first microphone and the speaker becomes short. Furthermore, the intercom device is suitable for the wiring system with excellent function expandability, as described in the above embodiments.

Therefore, the present invention is expected to further increase the demand of the intercom device, and especially as one of the components of the wiring system of the next generation.

CLAIMS:

1. An intercom device comprising:
a housing;
a speaker accommodated in said housing such that a front air chamber is defined as a space enclosed by a front surface of said speaker and an inner surface of said housing, and a rear air chamber is defined as a space enclosed by a rear surface of said speaker and an inner surface of said housing;
a first microphone disposed such that its sound receiving surface faces the front air chamber;
a second microphone disposed such that its sound receiving surface faces outside of said housing; and
a signal processing unit configured to, when an audio output of said speaker is picked up by said second microphone, remove a signal component corresponding to the audio output of said speaker from an output signal of said second microphone by use of an output signal of said first microphone.

2. The intercom device according to claim 1, wherein the rear air chamber is the space hermetically-sealed by the rear surface of said speaker and the inner surface of said housing.

3. The intercom device according to claim 1, wherein said first microphone is arranged such that the sound receiving surface of said first microphone faces a diaphragm of said speaker.

4. The intercom device according to claim 1, wherein said second microphone is located adjacent to said speaker.

5. The intercom device according to claim 1, wherein said first microphone and said second microphone are mounted on a same circuit board.

6. The intercom device according to claim 5, wherein the circuit board is mounted on an outside surface of said housing such that the sound receiving surface of said first microphone faces the front air chamber and the sound receiving surface of said second microphone faces the outside of said housing.

7. The intercom device according to claim 1, further comprising an acoustic tube having an opening communicated with the rear air chamber at its one end, and a closed end at the other end.

8. The intercom device according to claim 7, wherein said acoustic tube has a tube length of an odd multiple of one quarter of a wavelength determined in consideration of sonic velocity and an intended frequency where an increase in sound pressure level of the audio output of said speaker is needed.

9. The intercom device according to claim 1, further comprising a plurality of acoustic tubes having different lengths, each of which has an opening communicated with the rear air chamber at its one end, and a closed end at the other end.

10. The intercom device according to claim 7, wherein said acoustic tube has a resonant frequency between a frequency equivalent to a lowest resonant

frequency of said speaker that is supposed to be mounted on an infinite baffle and a lowest resonant frequency of said speaker mounted on the rear air chamber with the absence of said acoustic tube.

11. The intercom device according to claim 7, wherein said acoustic tube is formed along an inner surface of the rear air chamber.

12. The intercom device according to claim 7, further comprising a sound absorbing material disposed at the opening of said acoustic tube.

13. A wiring system using the intercom device as set forth in claim 1, comprising at least one transmission line installed in a building structure to transmit electric power and information signals, and a plurality of base units each adapted in use to be mounted in a wall surface of said building structure, and connected to said transmission line;

wherein the intercom device has a first connector detachably connectable to a second connector formed in each of said base units,

when said first connector is connected to said second connector of one of said base units, the intercom device functions such that the information signals provided through said transmission line are output as audio information from said speaker unit, and an output of said signal processing unit is sent out through said transmission line.

14. The wiring system as set forth in claim 13, further comprising a function unit adapted in use to be connected to said transmission line through one of said base units, said function unit having a third connector detachably

connectable to one of said second connector of said base unit, and an extended connector formed in the intercom device,

when said third connector of said function unit is connected to said second connector of said base unit, or when said first connector of the intercom device is connected to said second connector of said base unit, and said third connector of said function unit is connected to said extended connector of the intercom device, said function unit provides at least one of functions for supplying electric power provided through said transmission line, outputting information signals provided through said transmission line, and sending out information signals through said transmission line.

15. The wiring system as set forth in claim 14, wherein said first connector and said third connector are stylized in shape and arrangement, and said second connector and said extended connector are stylized in shape and arrangement corresponding to the stylization of said first connector and said third connector.

16. The wiring system as set forth in claim 13, further comprising a function unit adapted in use to be connected to said transmission line through one of said base units,

said function unit has a third connector detachably connectable to said second connector of said base unit and an extended connector detachably connectable to said first connector of said intercom device,

when said third connector of said function unit is connected to said second connector of said base unit, and said extended connector of said function unit is connected to said first connector of the intercom device, said function unit provides at least one of functions for supplying electric power provided from said transmission line through said base unit, outputting information signals provided from said transmission line through said base unit, and sending out

information signals to said transmission line through said base unit, and the intercom device operates such that the information signals provided from said transmission line through said base unit and said function unit are output as audio information from said speaker unit, and the output of said signal processing unit is sent out to said transmission line through said base unit and said function unit.

17. The wiring system as set forth in claim 16, wherein said first connector and said third connector are stylized in shape and arrangement, and said second connector and said extended connector are stylized in shape and arrangement corresponding to the stylization of said first connector and said third connector.

18. The wiring system as set forth in claim 13, wherein said first and second connectors provide an electric power transmission between said base unit and the intercom device by means of electromagnetic coupling.

19. The wiring system as set forth in claim 13, wherein said first and second connectors provide a signal transmission between said base unit and the intercom device by means of optical coupling.

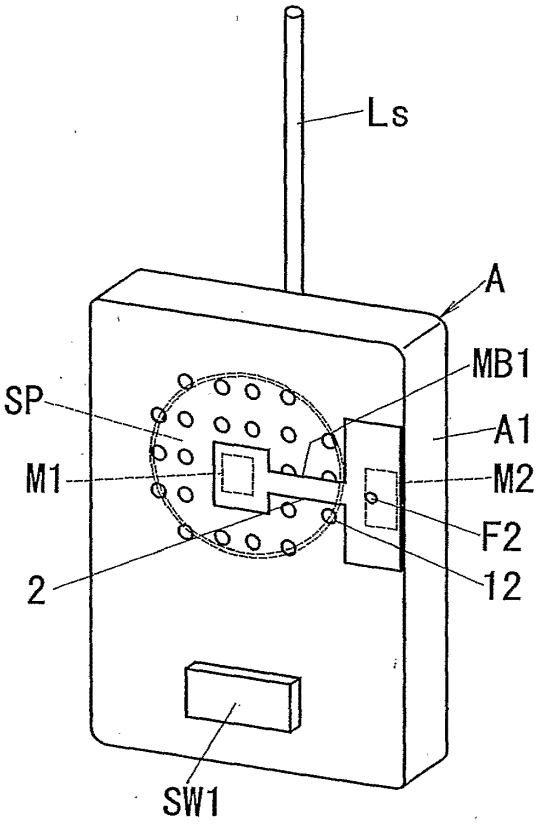


FIG. 1

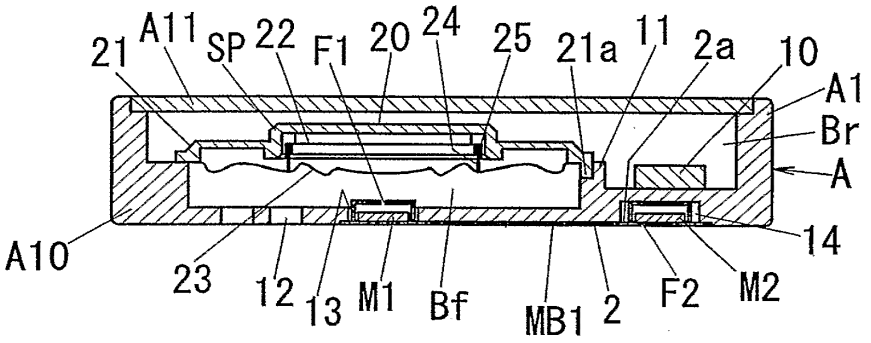

21
A11
SP
22
F1
20
24
25
21a
11
2a
10
A1
Br
A
14
A10
23
12
13
M1
Bf
MB1
2
F2
M2

FIG. 2

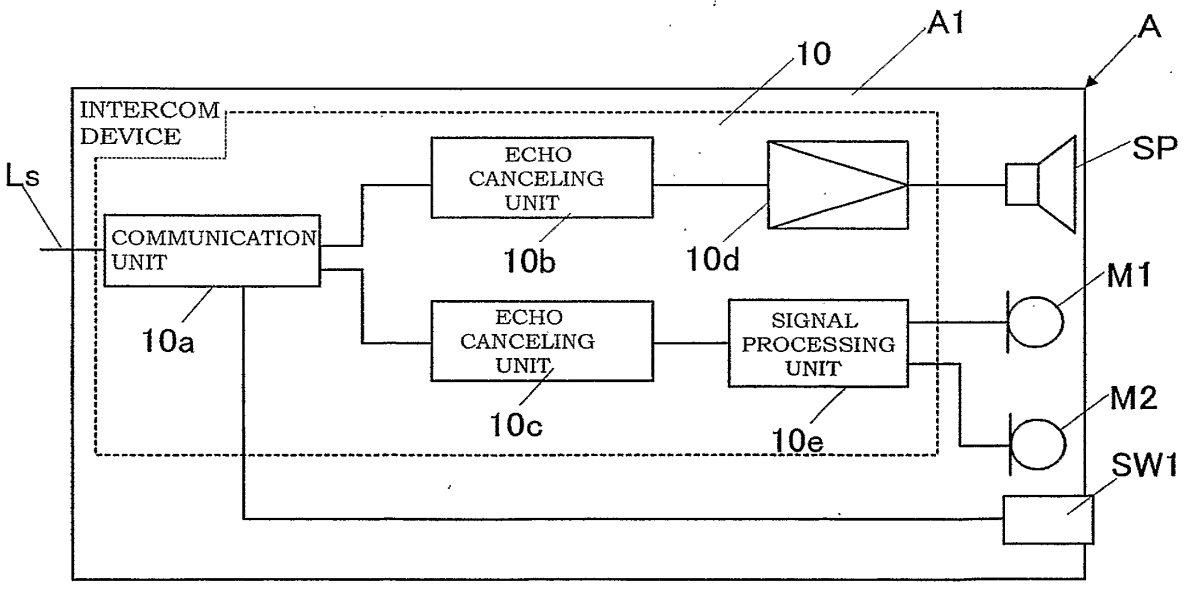

A1
A
10
INTERCOM DEVICE
Ls
COMMUNICATION UNIT
10a
ECHO CANCELING UNIT
10b
10d
SP
ECHO CANCELING UNIT
10c
SIGNAL PROCESSING UNIT
10e
M1
M2
SW1

FIG. 3

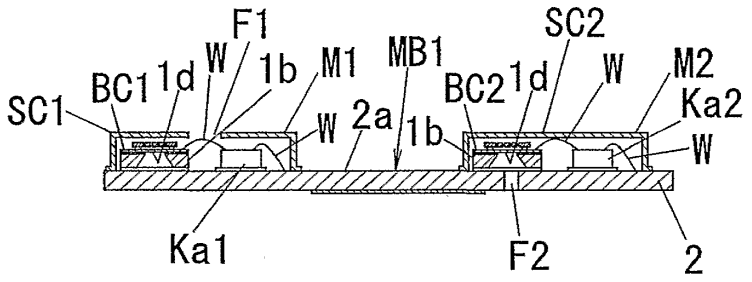

F1
W
1b
M1
MB1
SC2
W
M2
SC1
BC1
1d
W
2a
1b
BC2
1d
Ka2
W
Ka1
F2
2

FIG. 4

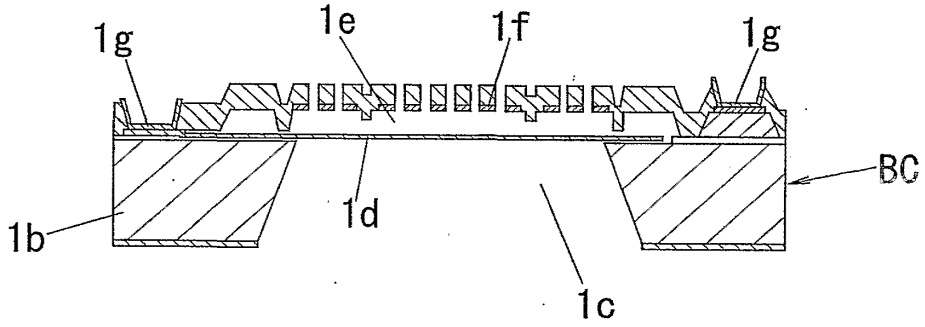

1g
1e
1f
1g
BC
1b
1d
1c

FIG. 5

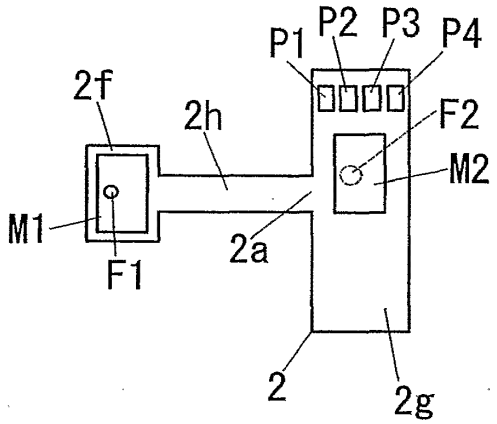


FIG. 6A

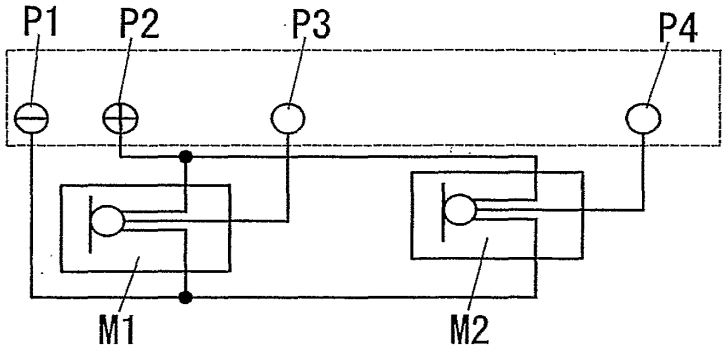


FIG. 6B

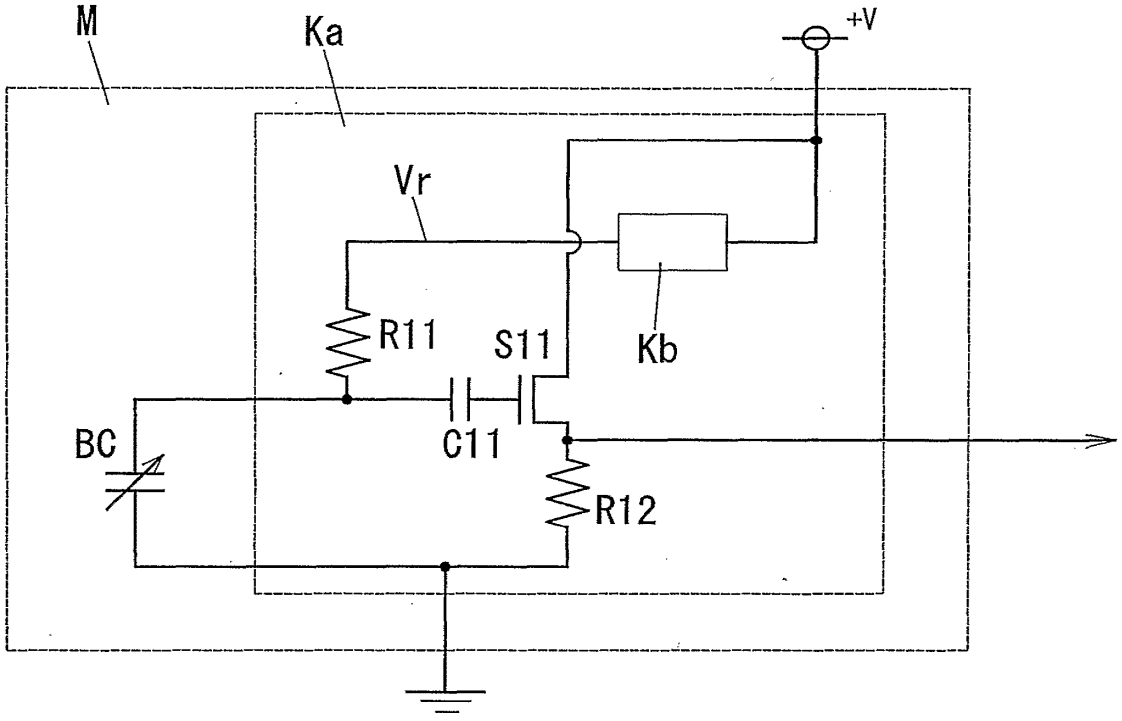


FIG. 7

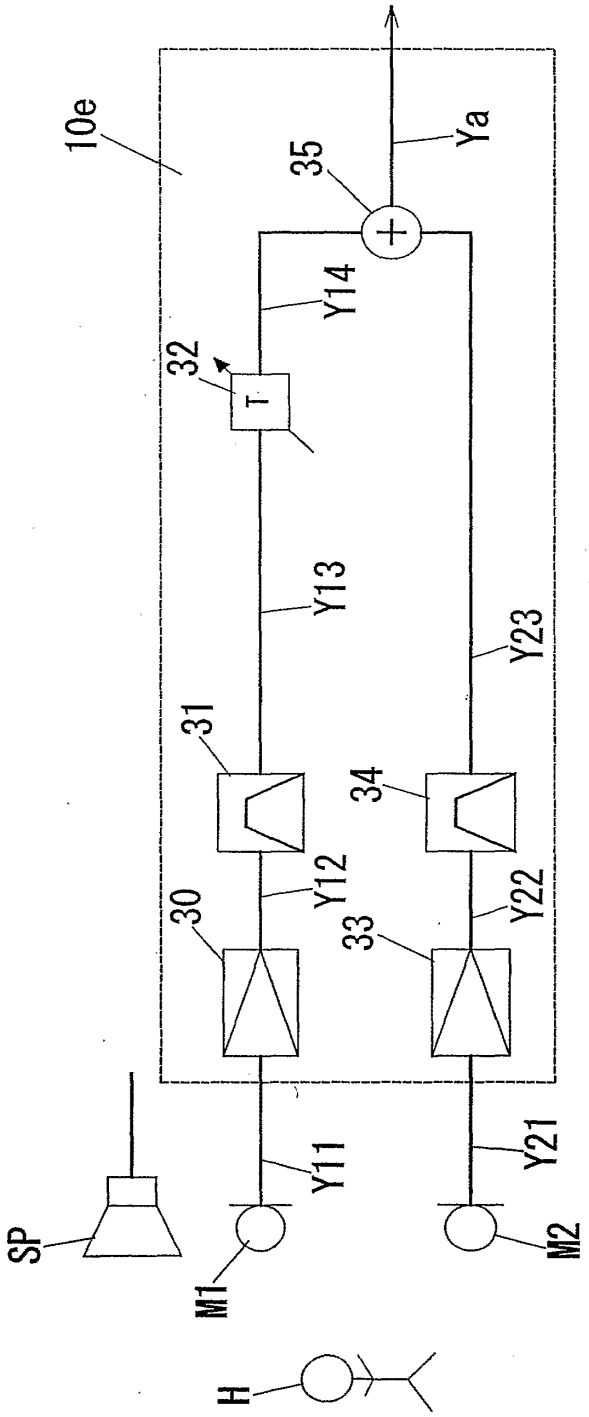


FIG. 8

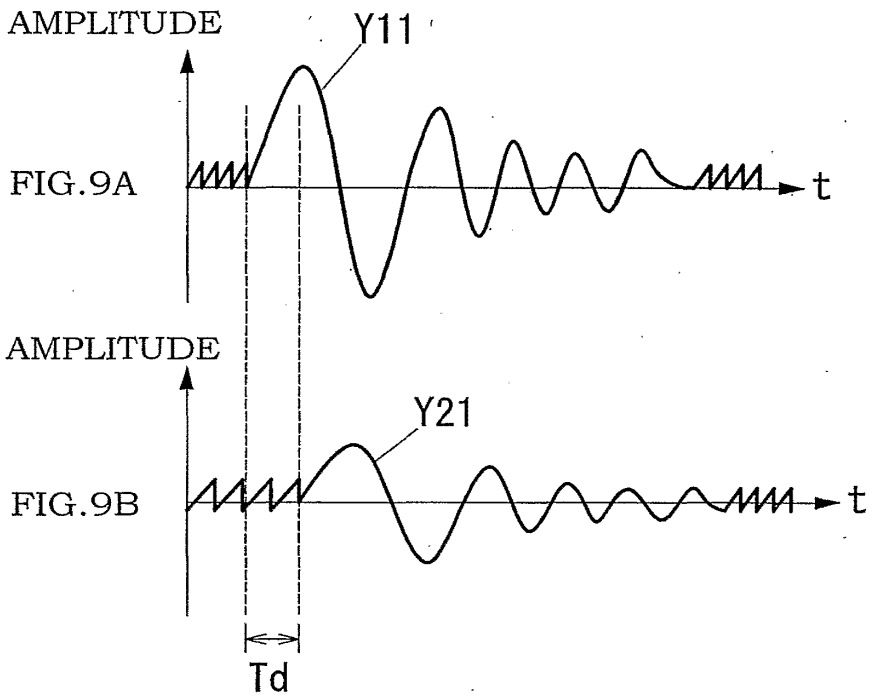


FIG.9A

FIG.9B

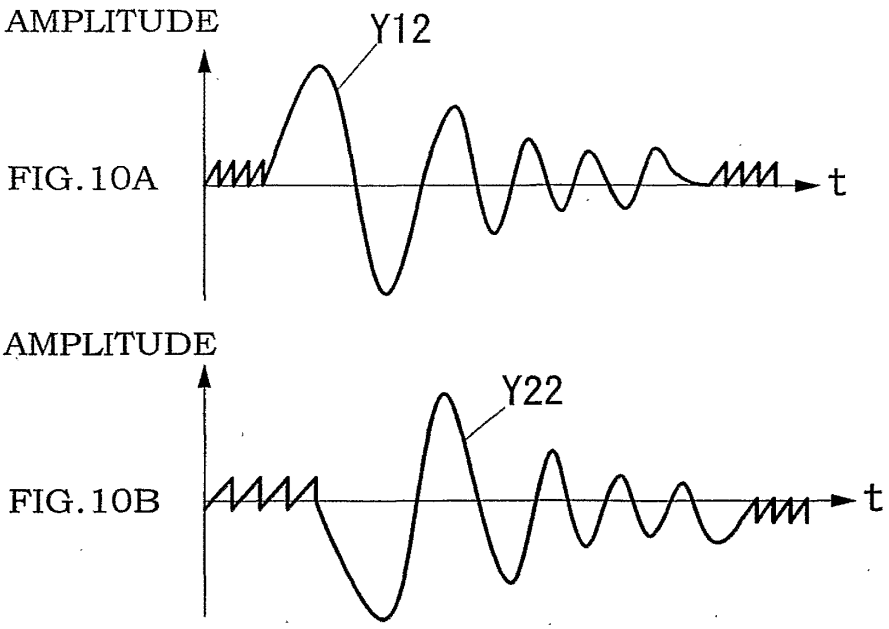


FIG.10A

FIG.10B

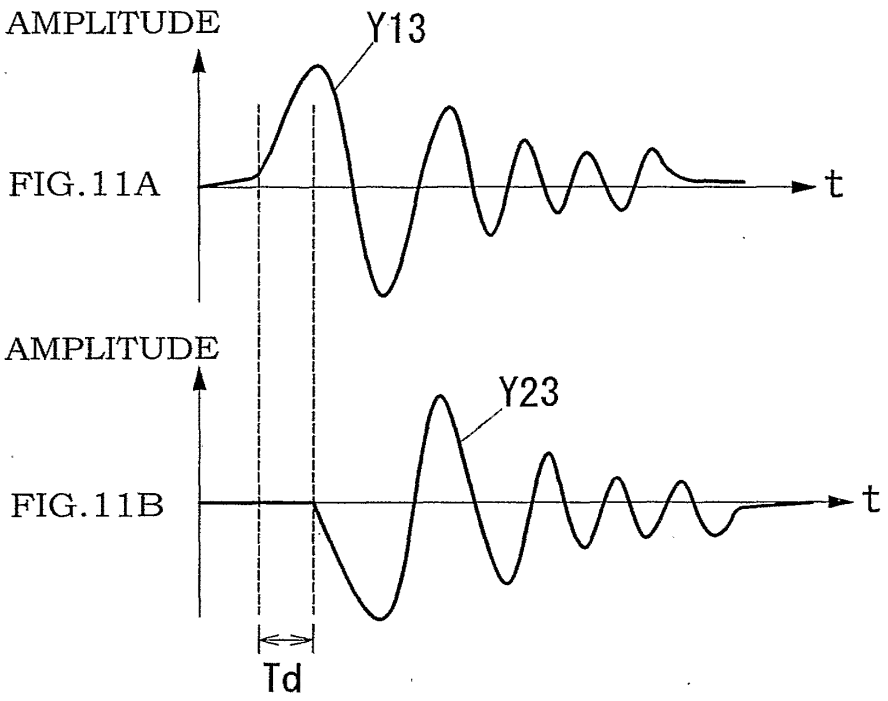


FIG.11A

FIG.11B

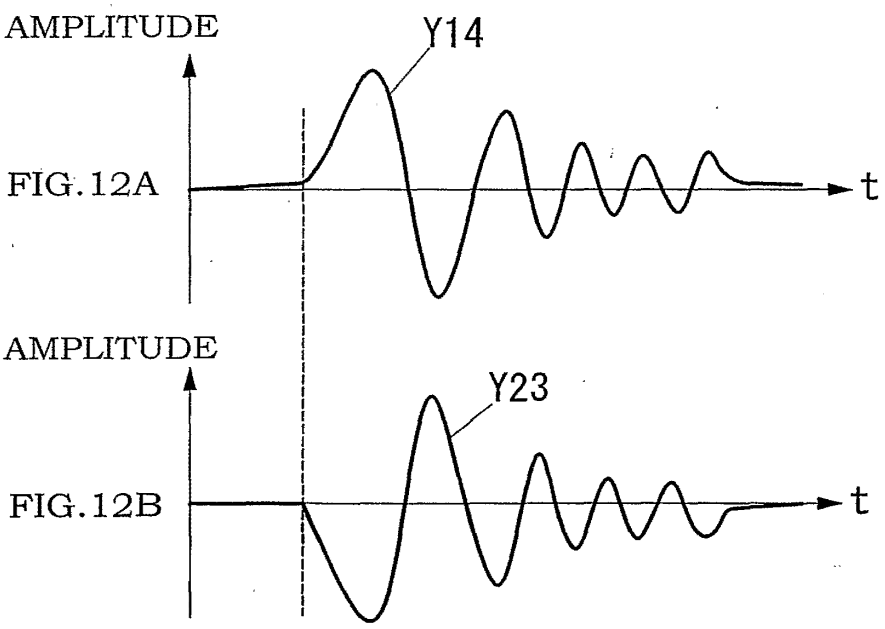


FIG.12A

FIG.12B

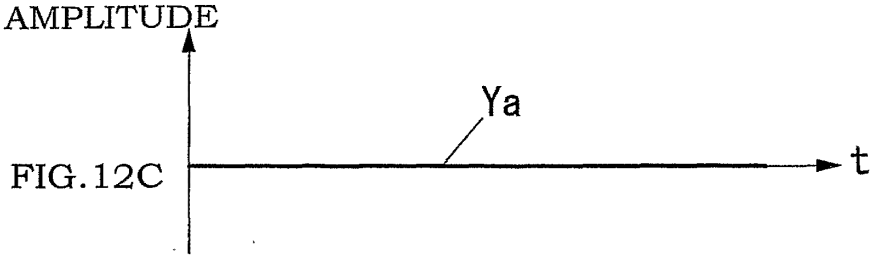
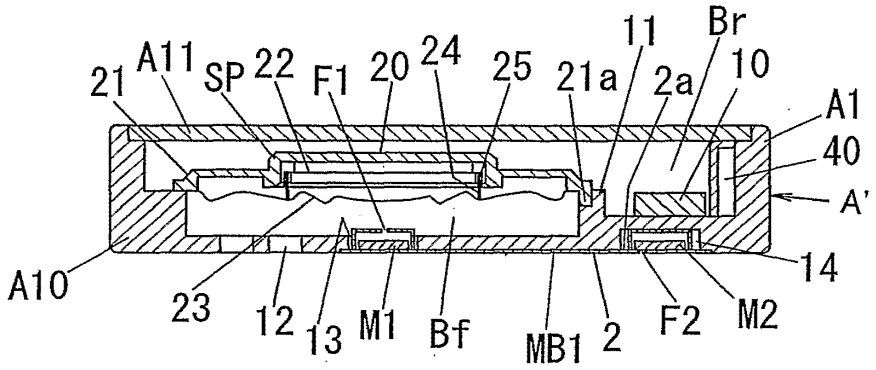


FIG.12C

FIG. 13



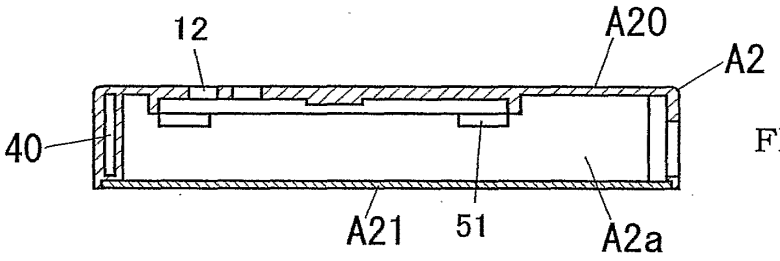
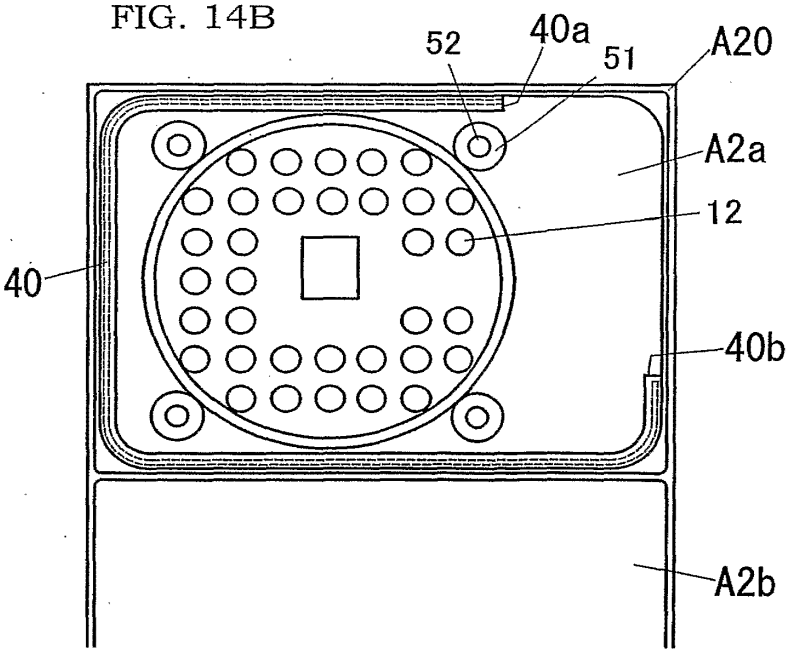


FIG. 14A

FIG. 14B



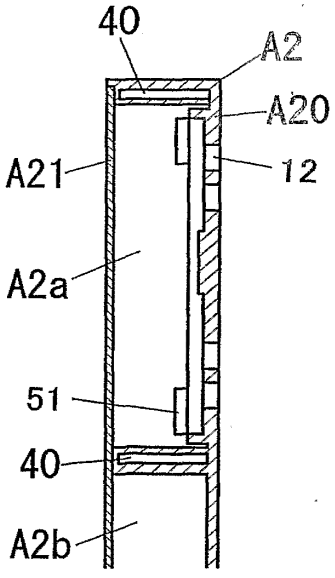


FIG. 14C

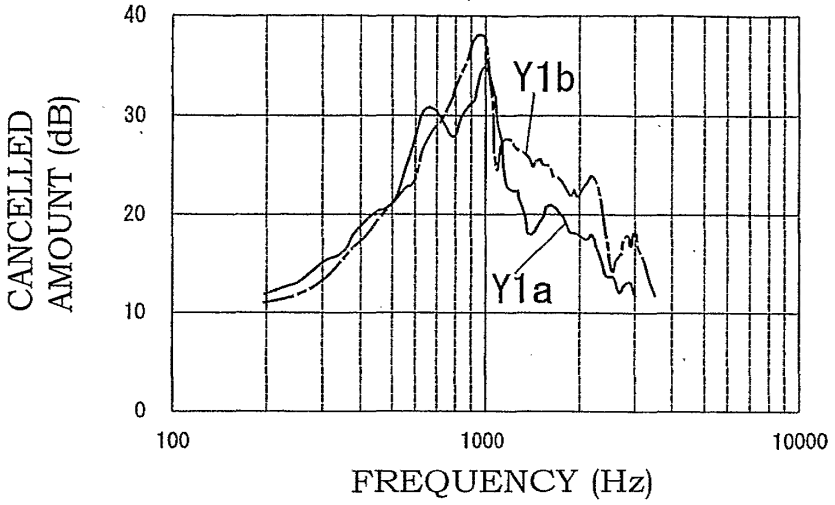


FIG. 15

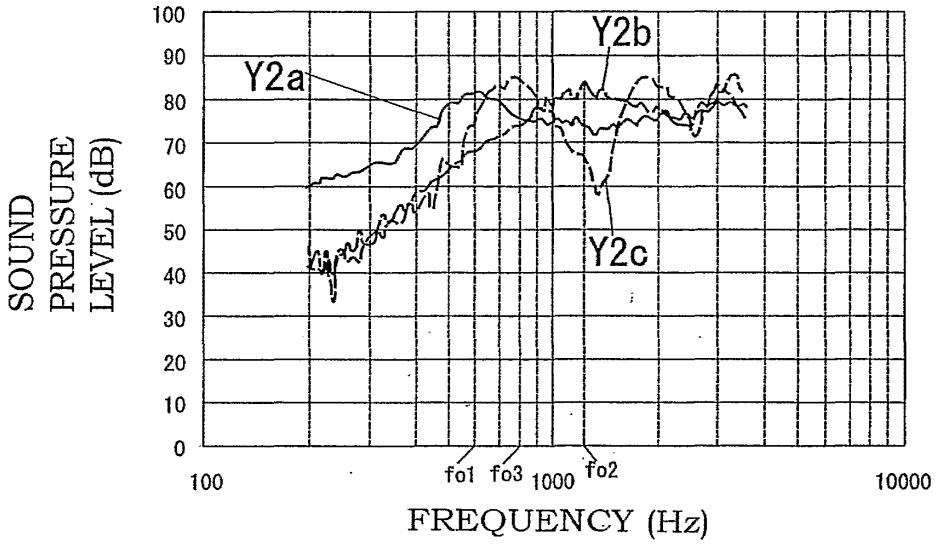


FIG. 16

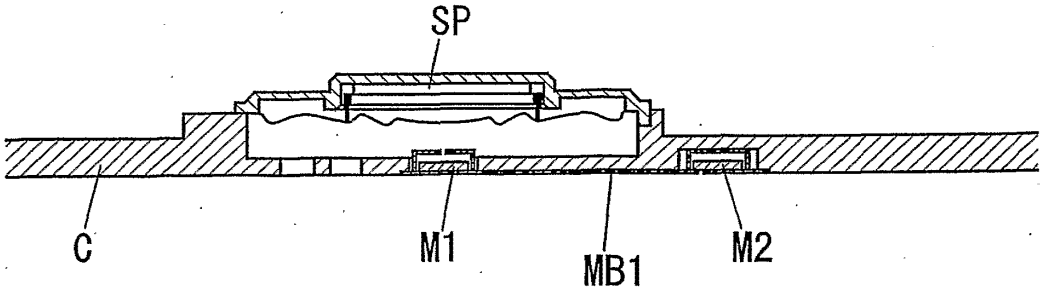


FIG. 17

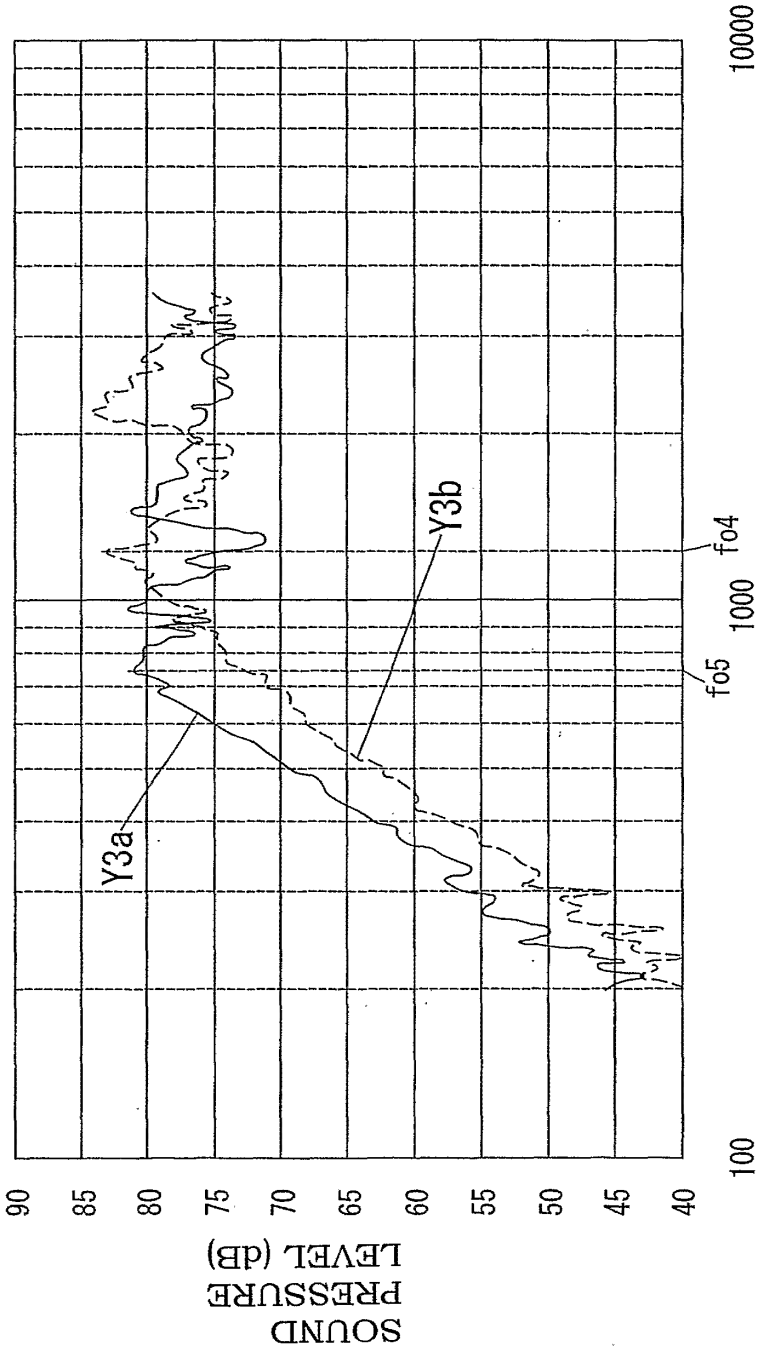

SOUND PRESSURE LEVEL (dB)
90
85
80
75
70
65
60
55
50
45
40
100
1000
10000
FREQUENCY (Hz)
fo4
fo5
Y3a
Y3b

FIG. 18

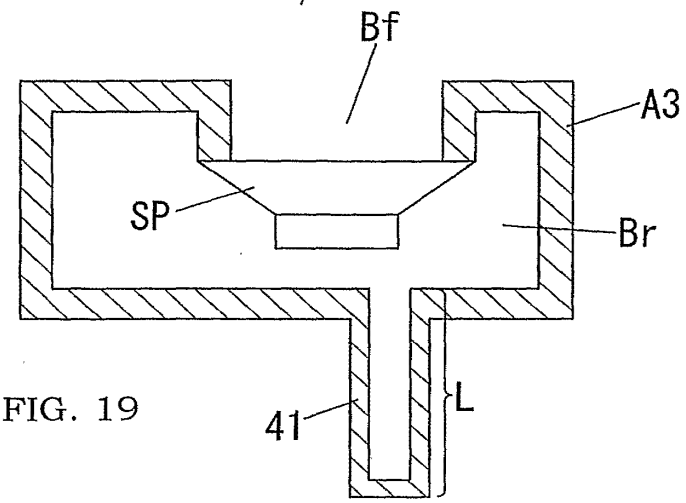


FIG. 19

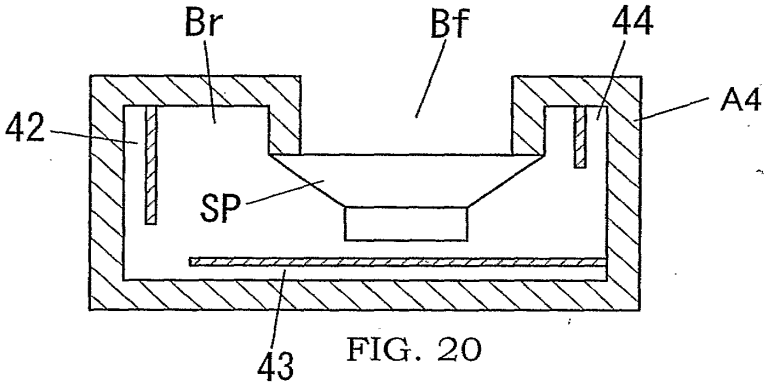


FIG. 20

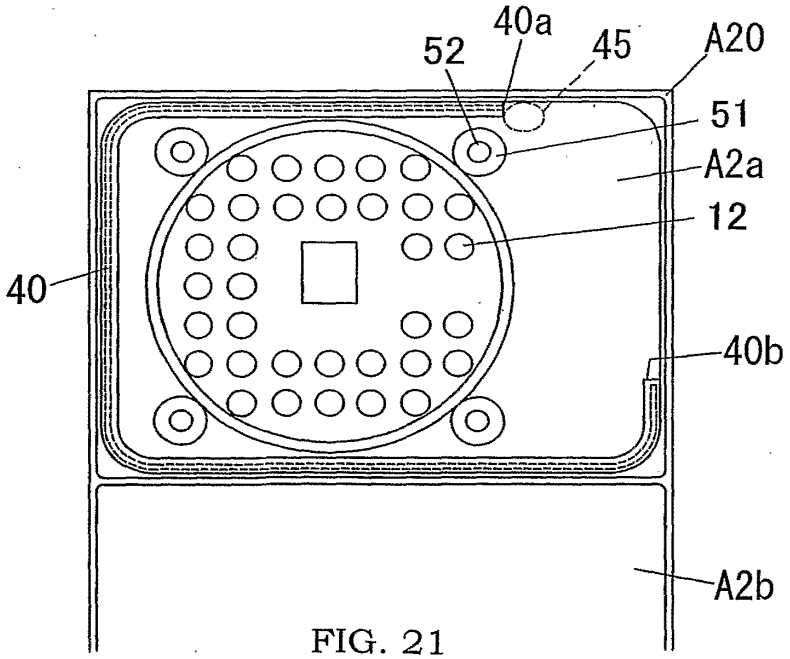


FIG. 21

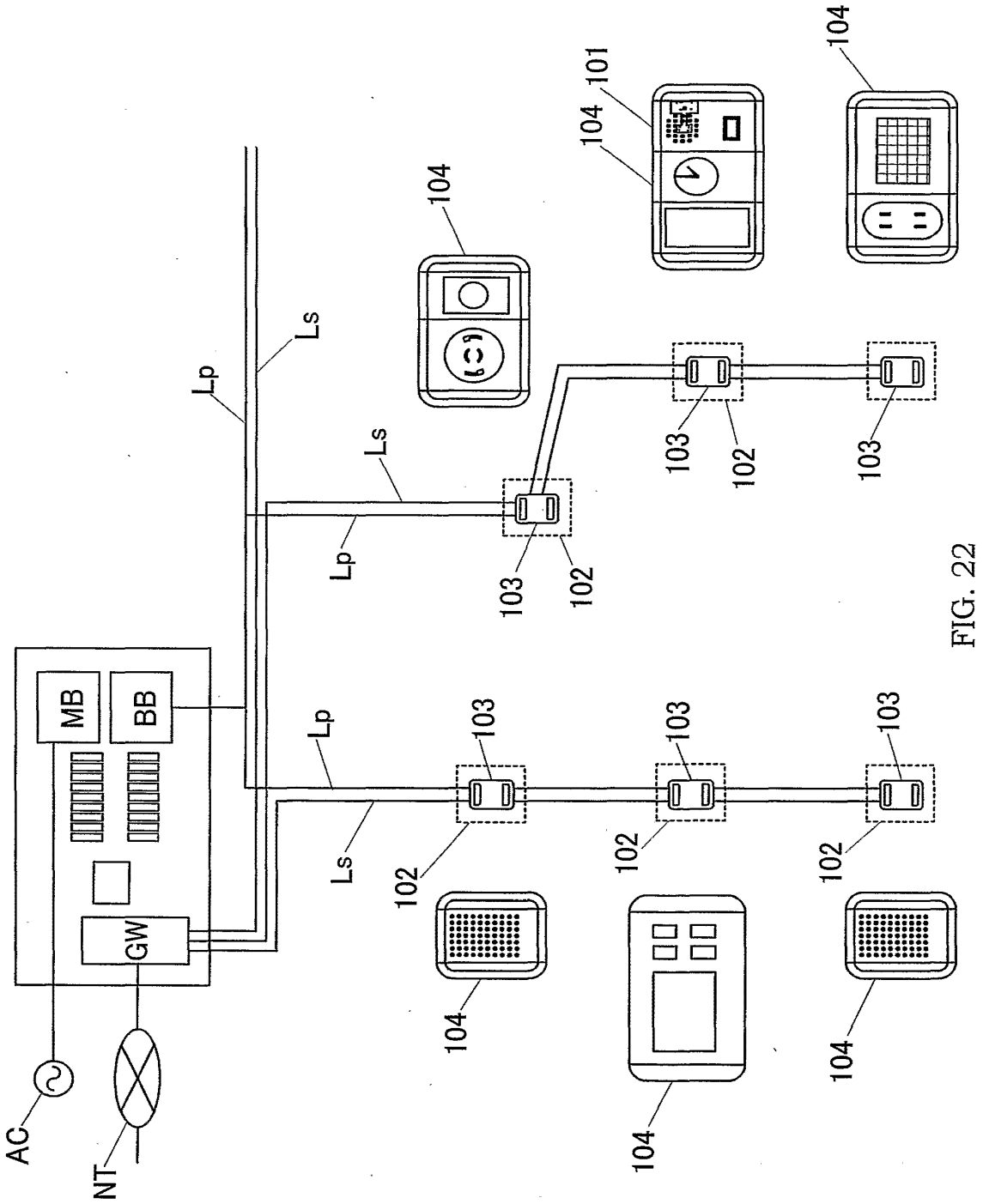


FIG. 22

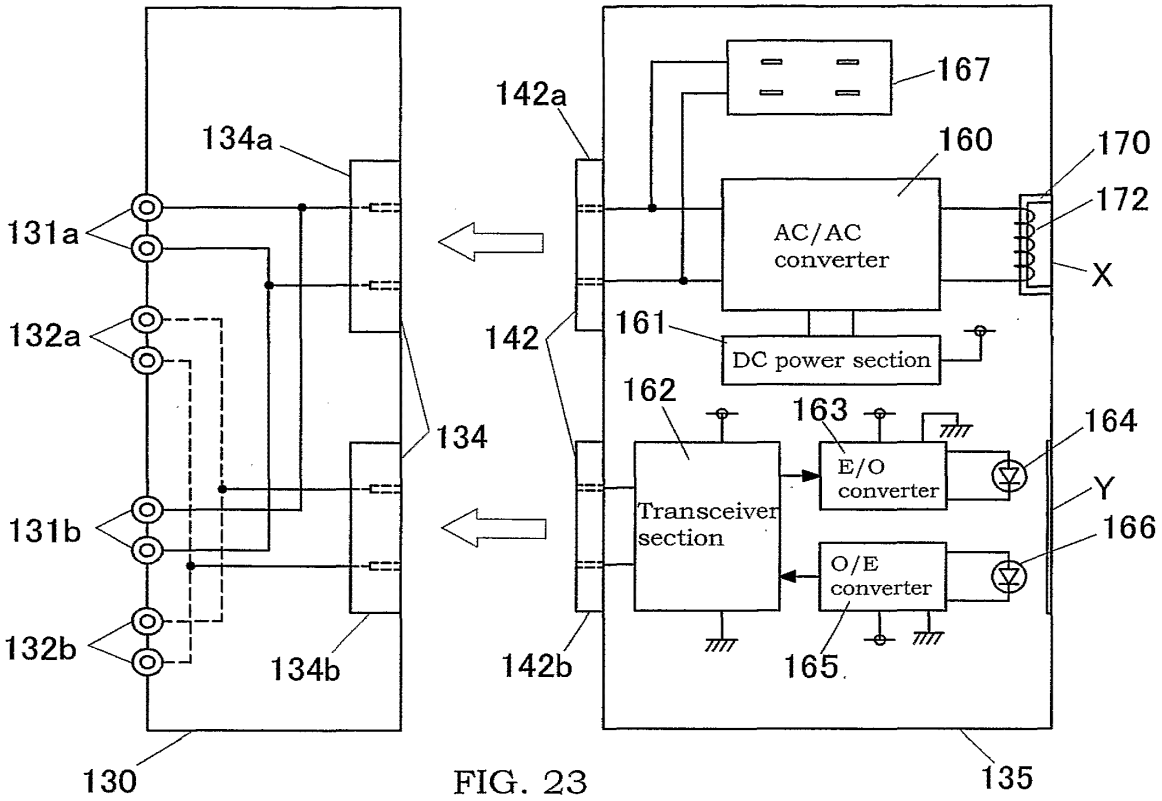

134a
131a
132a
131b
132b
134
134b
130
142a
167
160
170
172
AC/AC converter
X
142
161
DC power section
162
163
164
E/O converter
Y
Transceiver section
166
O/E converter
142b
165
135

FIG. 23

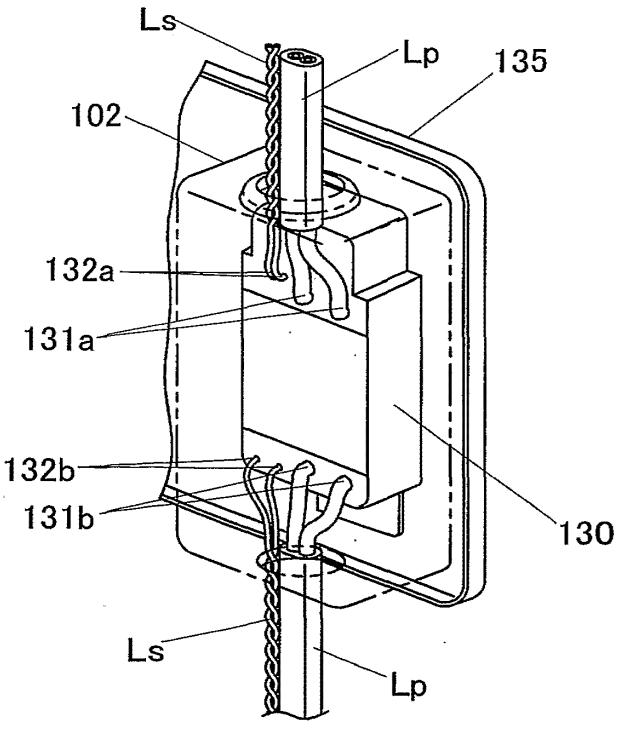

Ls
Lp
135
102
132a
131a
132b
131b
130
Ls
Lp

FIG. 24

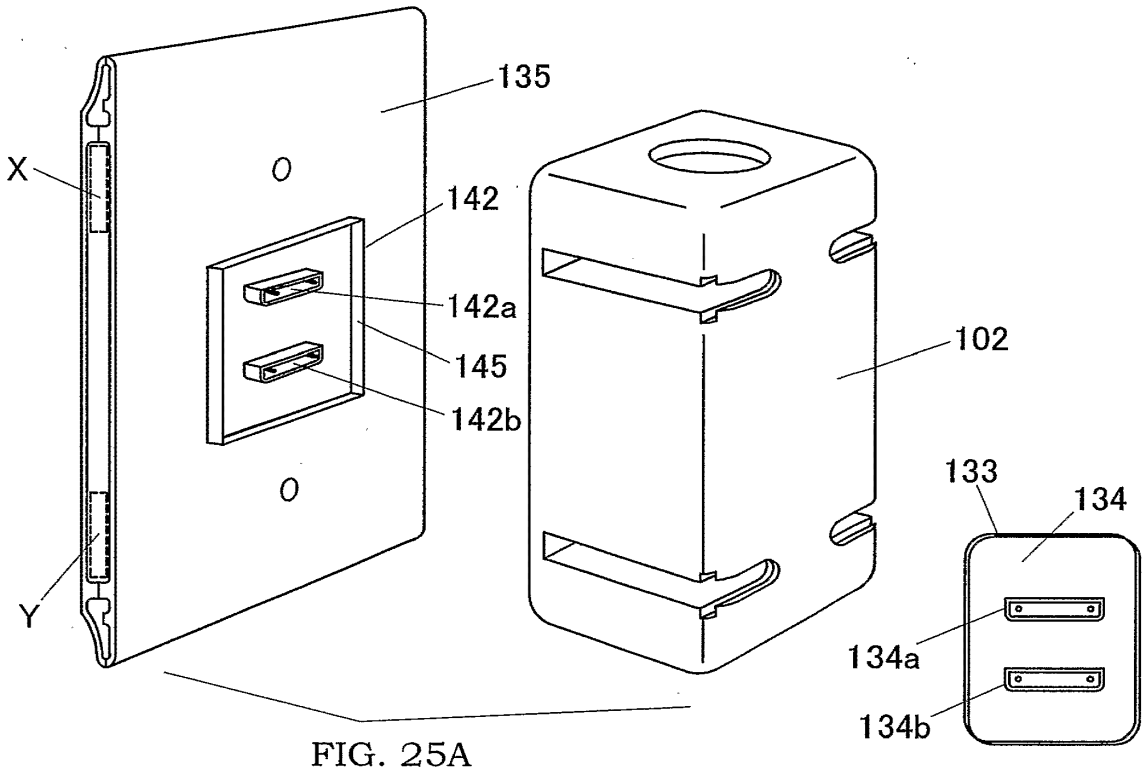


FIG. 25A

FIG. 25B

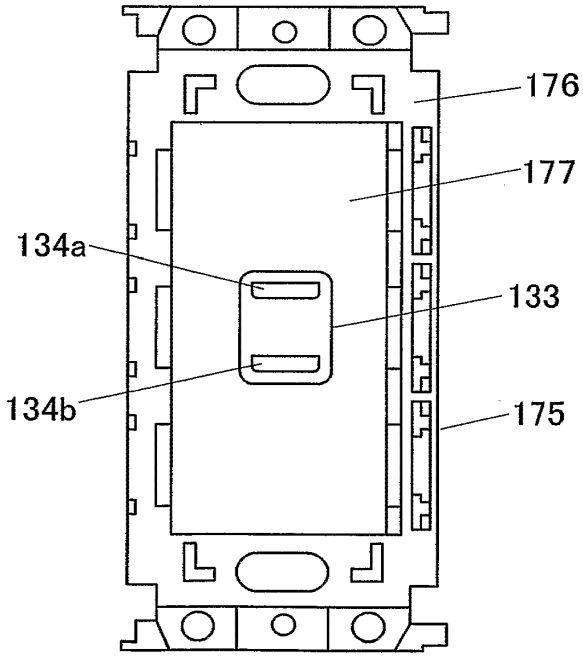


FIG. 26

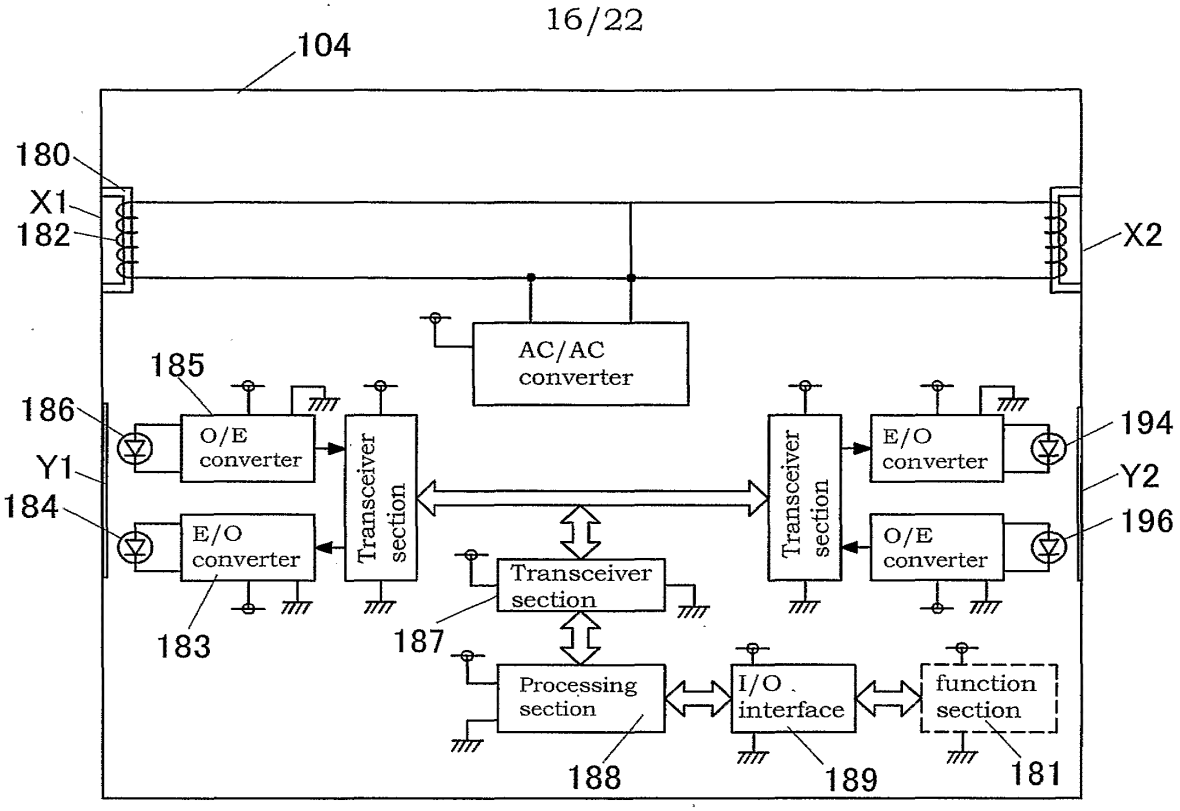

104
180
X1
182
X2
AC/AC converter
185
186
O/E converter
Transceiver section
Transceiver section
E/O converter
194
Y1
Y2
184
E/O converter
O/E converter
196
Transceiver section
183
187
Processing section
I/O interface
function section
188
189
181

FIG. 27

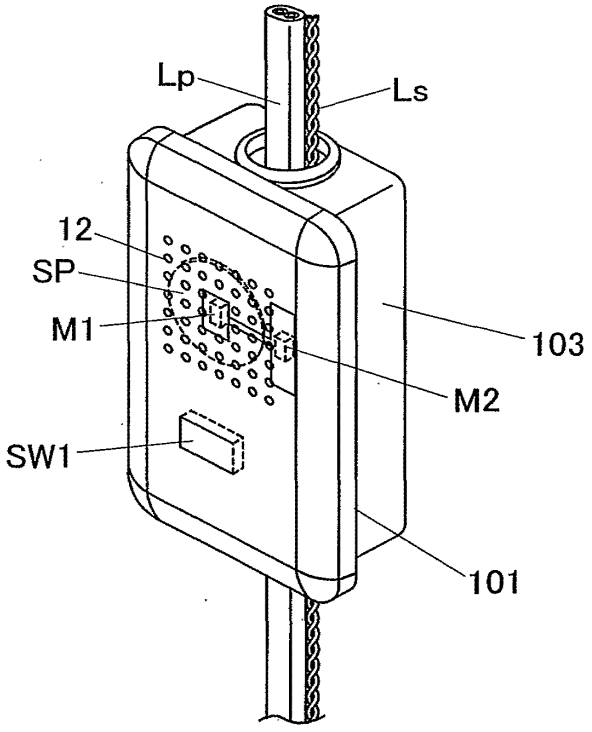

Lp
Ls
12
SP
M1
103
M2
SW1
101

FIG. 28

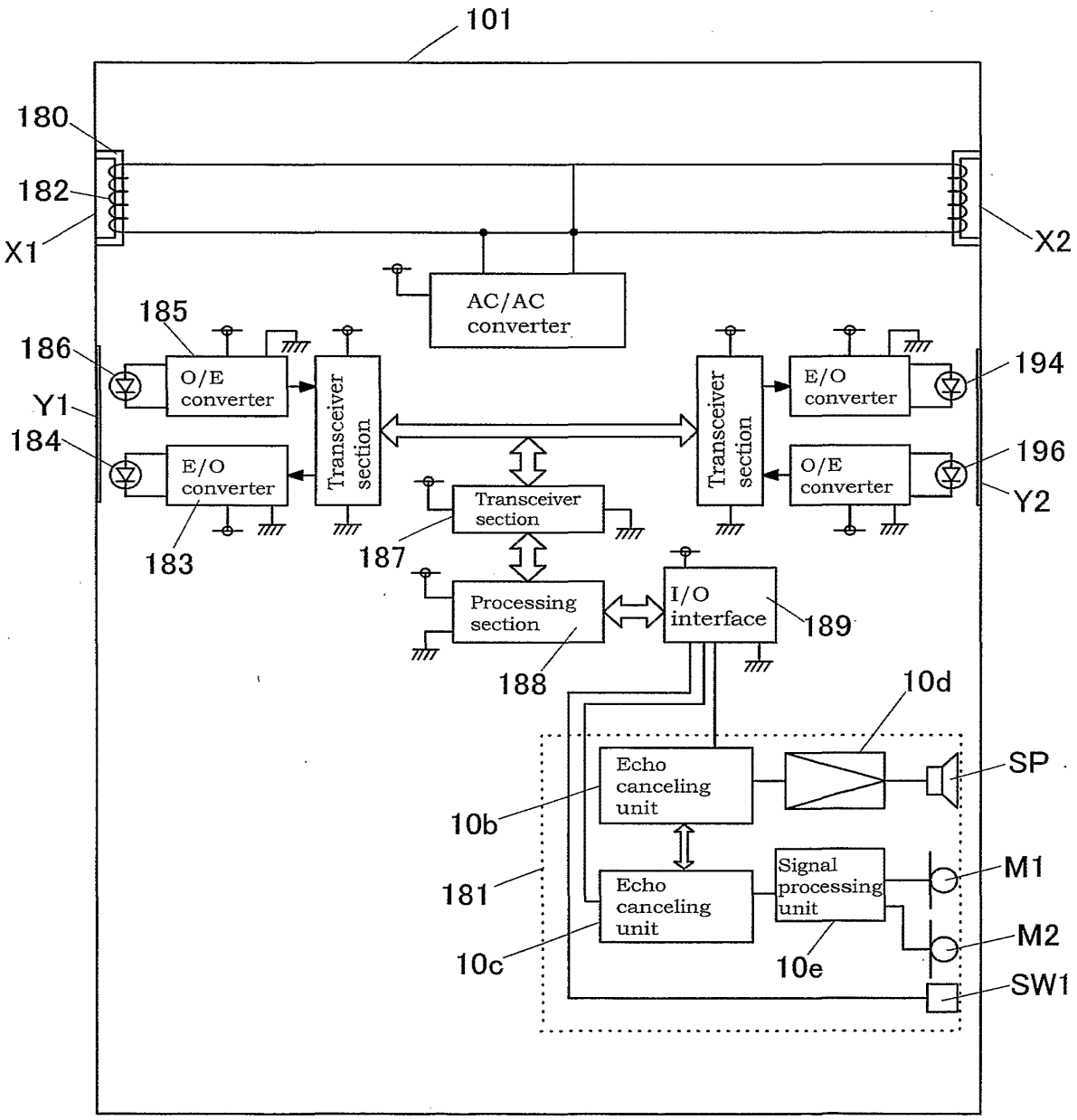

101
180
182
X1
X2
AC/AC converter
185
186
O/E converter
Transceiver section
Y1
184
E/O converter
183
Transceiver section
187
Transceiver section
E/O converter
194
O/E converter
196
Y2
Processing section
I/O interface
189
188
10d
Echo canceling unit
SP
10b
181
Echo canceling unit
Signal processing unit
M1
M2
10c
10e
SW1

FIG. 29

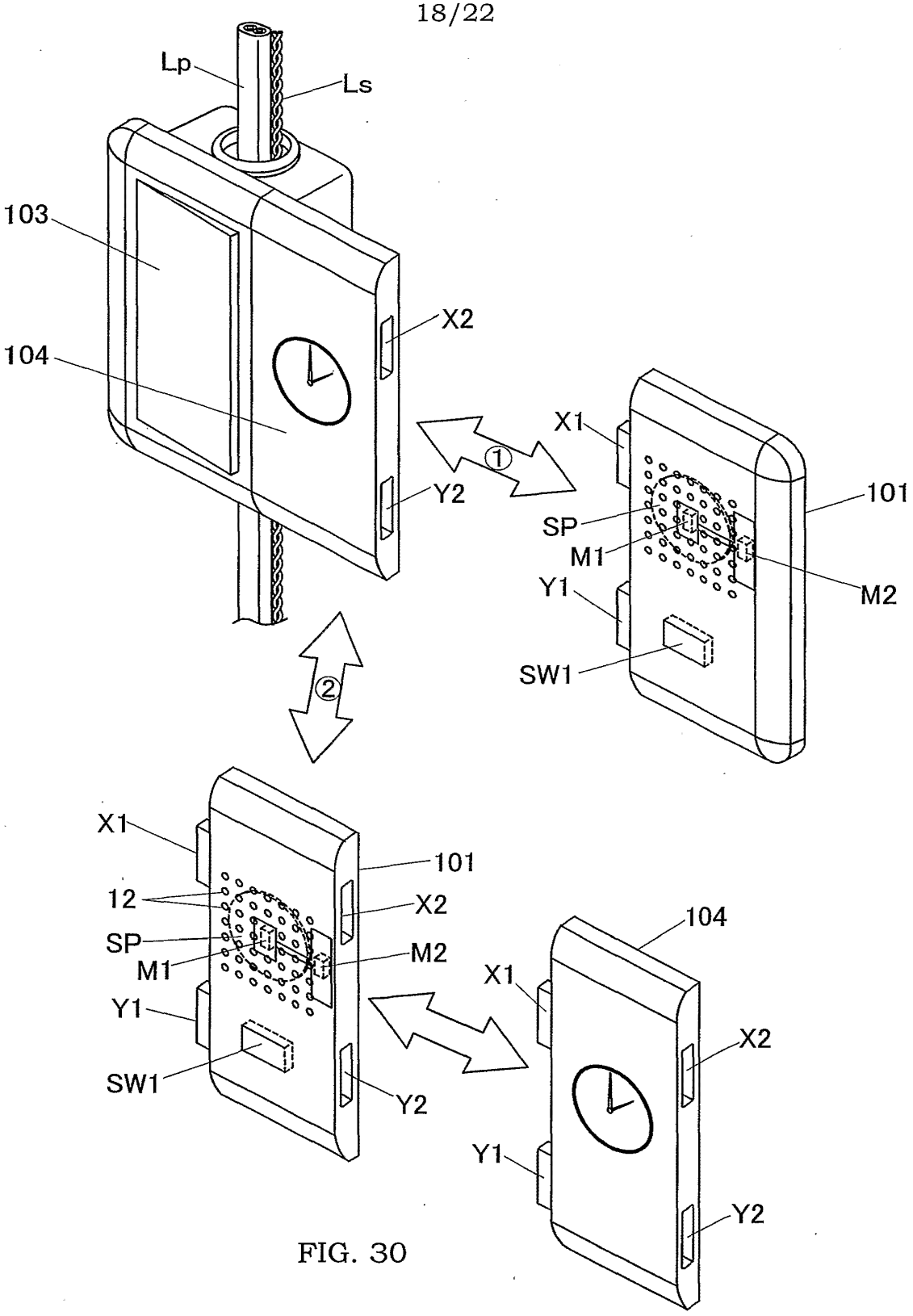

Lp
Ls
103
104
X2
Y2
①
X1
101
SP
M1
M2
Y1
SW1
②
X1
101
12
X2
SP
M1
M2
Y1
SW1
Y2
104
X1
X2
Y1
Y2

FIG. 30

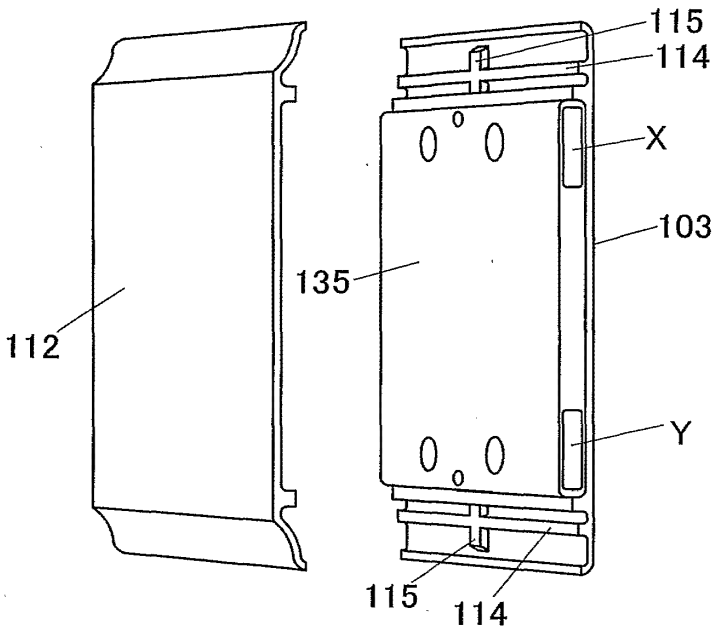


FIG. 31

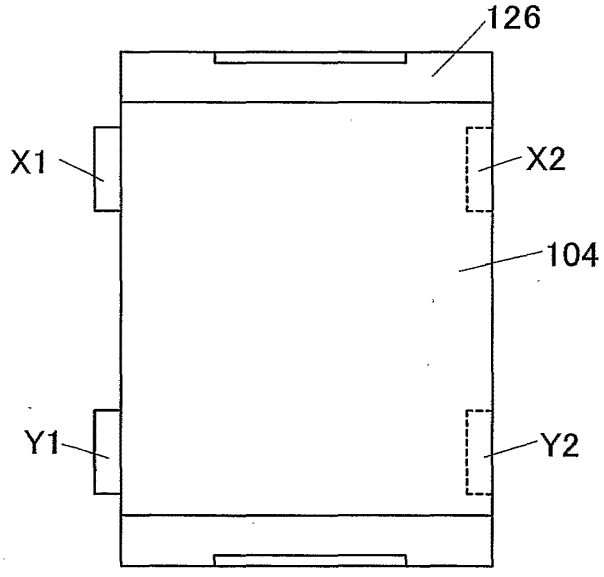


FIG. 32A

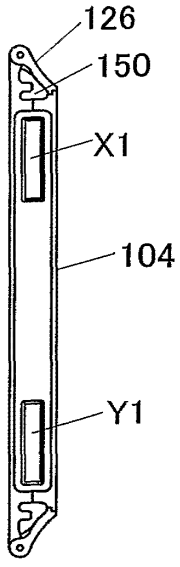


FIG. 32B

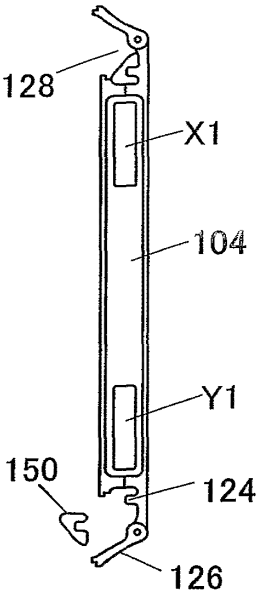


FIG. 32C

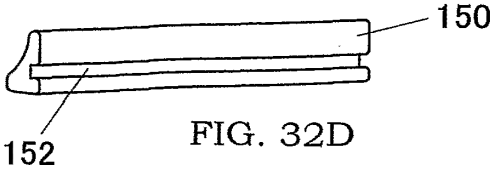


FIG. 32D

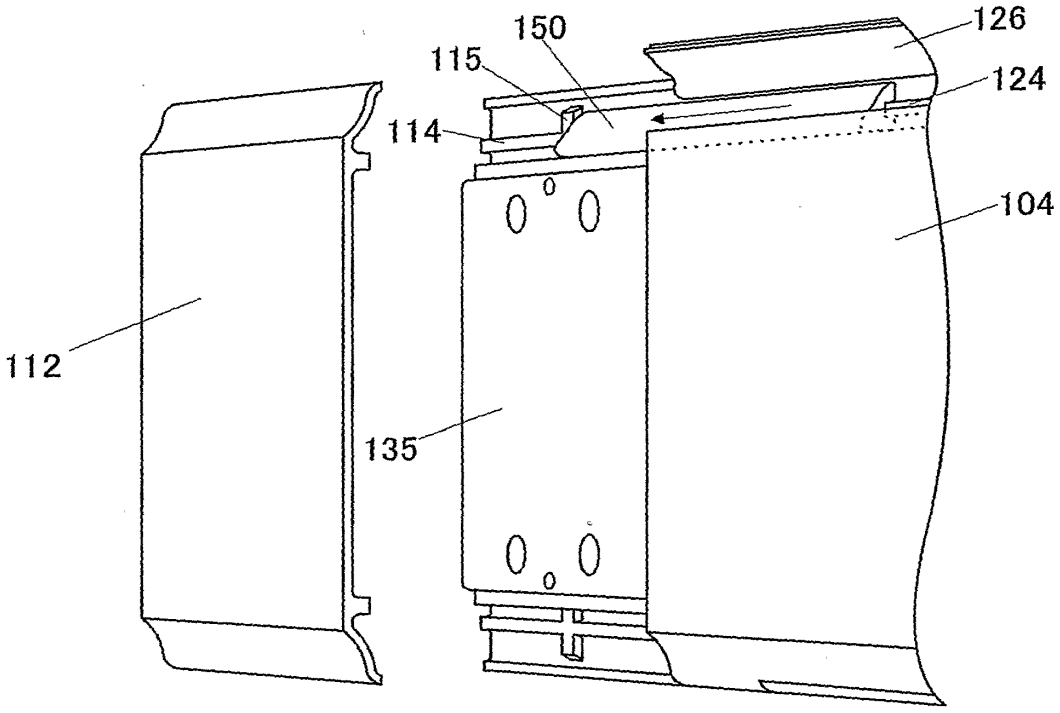


FIG. 33

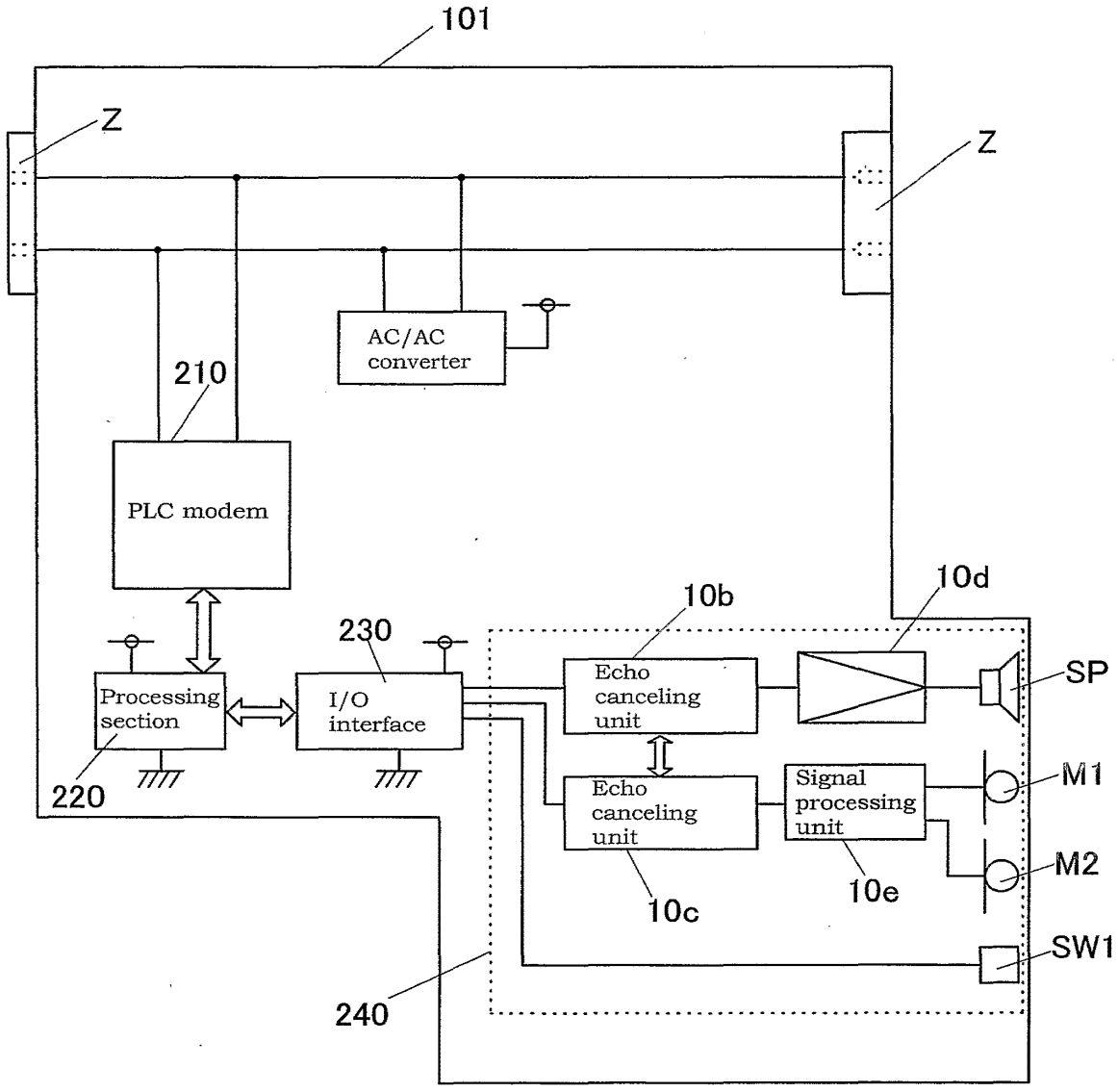

101
Z
Z
AC/AC converter
210
PLC modem
10b
10d
230
Processing section
I/O interface
Echo canceling unit
SP
220
Echo canceling unit
Signal processing unit
M1
M2
10c
10e
SW1
240

FIG. 34

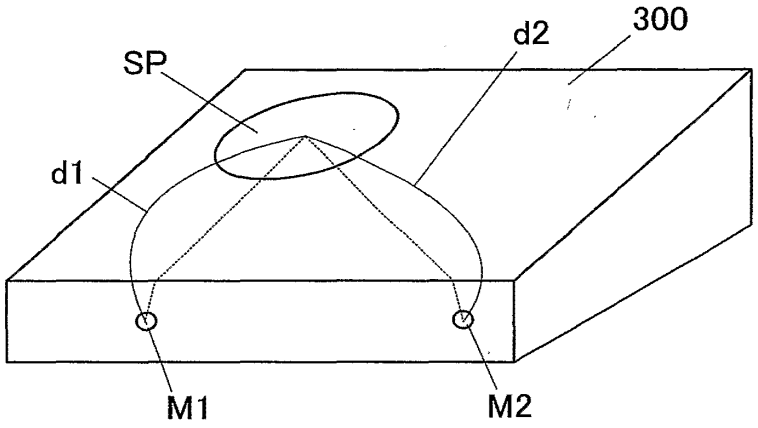


FIG. 35A

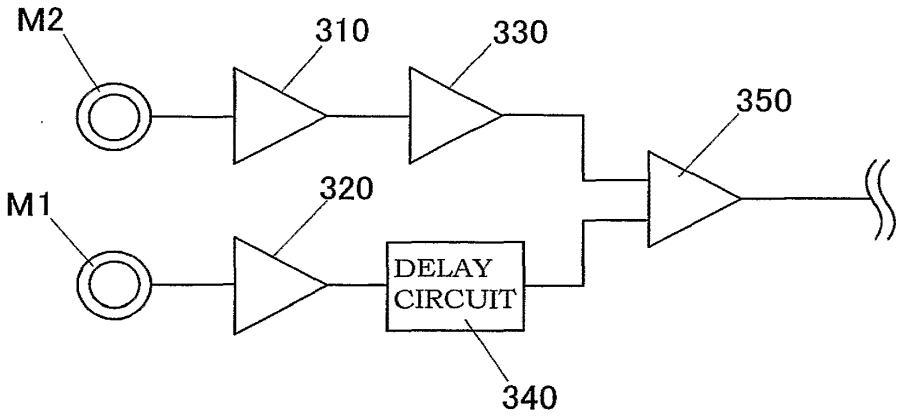


FIG. 35B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2007/069131

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04R3/02(2006.01)i, H04M1/62(2006.01)i, H04R1/02(2006.01)i, H04R1/28(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H04R3/02, H04M1/62, H04R1/02, H04R1/28

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2007
Registered utility model specifications of Japan 1996-2007
Published registered utility model applications of Japan 1994-2007

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.
A	JP 63-135100 A (BIBA KABUSHIKI KAISHA) 1988.06.07, Full Text; Fig.1 (No Family)	1-19
A	JP 2003-143677 A (MITSUBISHI ELECTRIC CORPORATION) 2003.05.16, Full Text; Fig.2 (No Family)	1-19
A	JP 01-253352 A (MATSUSHITA ELECTRIC WORKS,LTD.) 1989.10.09, Full Text; Fig.2 (No Family)	1-19

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25.12.2007	08.01.2008

Name and mailing address of the ISA/JP	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Hiroshi Ono 5Z 3654 Telephone No. +81-3-3581-1101 Ext. 3541

Form PCT/ISA/210 (second sheet) (April 2007)

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 February 2009 (19.02.2009)

PCT

(10) International Publication Number
WO 2009/023784 A1

(51) International Patent Classification:
H04R 25/00 (2006.01)

(21) International Application Number:
PCT/US2008/073189

(22) International Filing Date: 14 August 2008 (14.08.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/955,848 14 August 2007 (14.08.2007) US
60/968,237 27 August 2007 (27.08.2007) US

(71) Applicant *(for all designated States except US)*: **PERSONICS HOLDINGS INC.** [US/US]; Suite 510, 5200 Town Center Circle Tower II, Boca Raton, FL 33486 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **USHER, John** [GB/CA]; 55 de Bresoles, Apt. 400, Montreal, Québec H2Y 1V7 (CA). **BOILLOT, Marc** [US/US]; 9110 NW 11th Court, Plantation, FL 33322 (US).

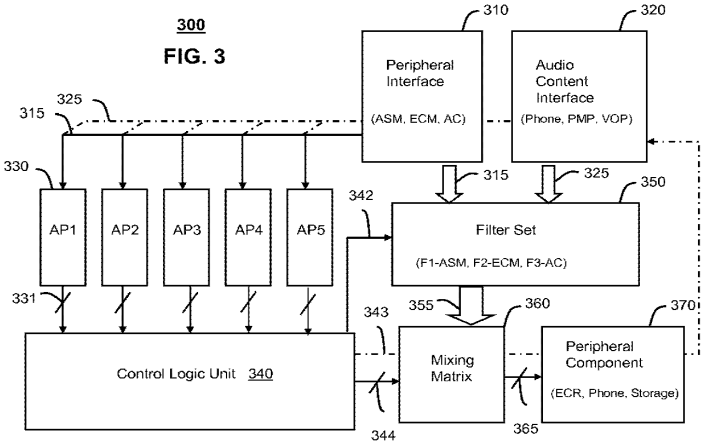
(74) Agents: **KURTZ, Richard** et al.; Greenberg Traurig LLC, Suite 1000, 2101 L Street NW, Washington, DC 20037 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(54) **Title:** METHOD AND DEVICE FOR LINKING MATRIX CONTROL OF AN EARPIECE II



(57) **Abstract:** A device (100) and method (500) for configuring audio delivery on an earpiece is provided. The earpiece can include a peripheral interface (310) configured to receive a plurality of sound signals and direct the plurality of sound signals to a plurality of audio processing modules (330) that produce audio control information (331) responsive to an analysis of the sound signals, a logic control unit (340) to receive the audio control information from the plurality of audio processing modules and generate configuration data (342), a filter set (350) to process the plurality of sound signals in accordance with the configuration data to produce filtered sound signals (355), and a mixing matrix (360) to mix the plurality of filtered sound signals in accordance with the configuration data to produce output sound signals (365) and route the output sound signals to at least one peripheral component (370). Other embodiments are disclosed.

WO 2009/023784 A1

METHOD AND DEVICE FOR LINKING MATRIX CONTROL OF AN EARPIECE II

FIELD

[0001] The present invention relates to a method of controlling an earpiece device, and more particularly, though not exclusively, a method of configuring and managing audio input and output on an earpiece.

BACKGROUND

[0002] Headphones or earpieces can be used for music enjoyment or voice communication. Use of these devices has steadily been increasing, and more products are expanding functionality to support audio delivery to headphones and earpieces.

[0003] The earpieces and associated products are becoming more intelligent as more communication features become available. However, the earpieces themselves have limited resources and must make efficient use of the many audio input and output configurations required to manage and support audio delivery.

[0004] A need therefore exists for improving audio configuration paths of headphones or earpieces.

SUMMARY

[0005] In a first embodiment, an earpiece, can include a peripheral interface configured to receive a plurality of sound signals and direct the plurality of sound signals to a plurality of audio processing modules that produce audio control information responsive to an analysis of the sound signals, a logic control unit operatively coupled to the plurality of audio processing modules to receive the audio control information from the plurality of audio processing modules and generate configuration data, a filter set operatively coupled to the peripheral interface to process the plurality of sound signals in accordance with the configuration data to produce filtered sound signals, and a mixing matrix operatively coupled to the control logic unit and filter set to mix the plurality of filtered sound signals in accordance with the audio control information to produce output sound signals and route the output sound signals to at least one peripheral component.

[0006] The peripheral component can be an Ear Canal Receiver (ECR), a phone, a portable communication device, or a data storage. The peripheral interface can include at

least one Ambient Sound Microphone (ASM) configured to convert an ambient sound to an ambient sound signal, and at least one Ear Canal Microphone (ECM) configured to convert an internal sound from an ear canal of a user to an internal sound signal. The earpiece can further include an audio content interface configured to receive a plurality of audio streams and direct the plurality of audio streams to the plurality of audio processing modules. The audio content interface can receive an audio stream from a phone, a media player, or a portable communication device. The audio content interface can mix the plurality of audio streams based on a user context that is one among an incoming call, a music session, or a voice mail.

[0007] The audio control information can include filter data for processing the plurality of sound signals, audio control data for assigning a priority to the plurality of sound signals, and router data for mixing the plurality of filtered signals according to the priority. The priority can be event driven responsive to detecting a sound signature, a background noise condition, a battery life indication, a manual interaction, or a voice recognition command.

[0008] In a second embodiment an earpiece can include a peripheral interface configured to receive a plurality of sound signals and direct the plurality of sound signals to a plurality of audio processing modules, an audio content interface configured to receive a plurality of audio streams and also direct the plurality of audio streams to the plurality of audio processing modules, at least one signal analysis module operatively coupled to the peripheral interface and audio content interface to provide a shared analysis of the sound signals and the audio streams for the plurality of audio processing modules, a logic control unit operatively coupled to the plurality of audio processing modules and the at least one signal analysis module to receive the shared analysis and audio control information to generate configuration data, a filter set operatively coupled to the peripheral interface to process the plurality of sound signals in accordance with the configuration data to produce filtered sound signals, and a mixing matrix operatively coupled to the control logic unit and filter set to mix the plurality of filtered sound signals in accordance with the audio control information to produce output sound signals and route the output sound signals to at least one peripheral component.

[0009] The shared analysis can include spectral analysis, spectral band energy level analysis, spectral envelope analysis, voice activity detection analysis, and cross-correlation analysis. A separate signal analysis module can be provided for components of the peripheral interface and the audio content interface that is shared among the plurality of

audio processing modules. The peripheral interface can include at least one Ambient Sound Microphone (ASM) coupled to an ASM signal analysis module configured to analyze an ambient sound signal, and at least one Ear Canal Microphone (ECM) coupled to an ECM signal analysis module configured to analyze an internal sound from an ear canal of a user. The audio content interface can be coupled to an audio content (AC) signal analysis module and configured to analyze an audio stream from a phone, a media player, or a portable communication device. The logic control unit can produce a linkage matrix of mixing gains that are applied to the plurality of filtered sound signals to produce the plurality of output sound signals to each of a respective peripheral output device.

[0010] In a third embodiment, a method for configuring audio delivery on an earpiece can include the steps of receiving at least one sound signal and at least one audio stream, performing an analysis of the at least one sound signal and the at least one audio stream, presenting the analysis, the at least one sound signal, and the at least one audio stream to a plurality of audio processing modules that generate configuration data responsive to the receiving. The method can include filtering the at least one sound signal and at least one audio stream according to the configuration data to produce filtered sound signals, mixing the filtered signals according to the configuration data to produce output sound signals, and routing the output sound signals to at least one peripheral component. The at least one sound signal can be an ambient sound signal, an ear canal sound signal. The at least one audio stream can be received from a phone, a media player, or a portable communication device. A linkage matrix of mixing gains can be generated based on the configuration data that are applied to the plurality of filtered sound signals for producing the plurality of output sound signals.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] FIG. 1 is a pictorial diagram of an earpiece in accordance with an exemplary embodiment;

[0012] FIG. 2 is a block diagram of the earpiece in accordance with an exemplary embodiment;

[0013] FIG. 3 is an exemplary schematic of a software system for the earpiece in accordance with an exemplary embodiment;

[0014] FIG. 4 is a more detailed exemplary schematic of the software system of FIG. 3 for the earpiece in accordance with an exemplary embodiment t;

[0015] FIG. 5 is a flowchart of a method for generating filter configuration data from a plurality of processing modules in accordance with an exemplary embodiment in accordance with an exemplary embodiment;

[0016] FIG. 6 is a flowchart of a method for generating and applying filter configuration data from a plurality of processing modules in accordance with an exemplary embodiment;

[0017] FIG. 7 is an exemplary schematic for configuring audio input and output via a mixing matrix in accordance with an exemplary embodiment;

[0018] FIG. 8 is another exemplary schematic of a software system for the earpiece providing separate analysis and re-synthesis modules for peripheral inputs in accordance with an exemplary embodiment; and

[0019] FIG. 9 is a more detailed schematic for analysis and re-synthesis for a particular peripheral input in accordance with an exemplary embodiment.

DETAILED DESCRIPTION

[0020] The following description of at least one exemplary embodiment is merely illustrative in nature and is in no way intended to limit the invention, its application, or uses.

[0021] Processes, techniques, apparatus, and materials as known by one of ordinary skill in the relevant art may not be discussed in detail but are intended to be part of the enabling description where appropriate, for example the fabrication and use of transducers. Additionally in at least one exemplary embodiment the sampling rate of the transducers can be varied to pick up pulses of sound, for example less than 50 milliseconds.

[0022] In all of the examples illustrated and discussed herein, any specific values, for example the sound pressure level change, should be interpreted to be illustrative only and non-limiting. Thus, other examples of the exemplary embodiments could have different values.

[0023] Note that similar reference numerals and letters refer to similar items in the following figures, and thus once an item is defined in one figure, it may not be discussed for following figures.

[0024] Note that herein when referring to correcting or preventing an error or damage (*e.g.,* hearing damage), a reduction of the damage or error and/or a correction of the damage or error are intended.

[0025] At least one exemplary embodiment of the invention is directed to an earpiece for background noise mitigation. Reference is made to **FIG. 1** in which an earpiece device,

generally indicated as earpiece 100, is constructed in accordance with at least one exemplary embodiment. As illustrated, earpiece 100 depicts an electro-acoustical assembly 113 for an in-the-ear acoustic assembly, as it would typically be placed in the ear canal 131 of ear 117 of user 135 The earpiece 100 can be an in the ear earpiece, behind the ear earpiece, receiver in the ear, open-fit device, or any other suitable earpiece type. The earpiece 100 can be partially or fully occluded in the ear canal, and is suitable for use with users having healthy or abnormal auditory functioning.

[0026] Earpiece 100 includes an Ambient Sound Microphone (ASM) 111 to capture ambient sound, an Ear Canal Receiver (ECR) 125 to deliver audio to an ear canal 131, and an Ear Canal Microphone (ECM) 123 to capture internal sounds within the ear canal and also assess a sound exposure level within the ear canal. The earpiece 100 can partially or fully occlude the ear canal 131 to provide various degrees of acoustic isolation. The assembly is designed to be inserted into the users ear canal 131, and to form an acoustic seal with the ear canal walls 129 at a location 127 between the entrance to the ear canal and the tympanic membrane (or ear drum) 133 Such a seal is typically achieved by means of a soft and compliant housing of assembly 113. Such a seal is pertinent to the performance of the system in that it creates a closed cavity 131 of approximately 5cc between the in–ear assembly 113 and the tympanic membrane 133. As a result of this seal, the ECR (speaker) 125 is able to generate a full range bass response when reproducing sounds for the user. This seal also serves to significantly reduce the sound pressure level at the user's eardrum resulting from the sound field at the entrance to the ear canal. This seal is also the basis for the sound isolating performance of the electro-acoustic assembly.

[0027] Located adjacent to the ECR 125, is the ECM 123, which is acoustically coupled to the (closed) ear canal cavity 131. One of its functions is that of measuring the sound pressure level in the ear canal cavity 131 as a part of testing the hearing acuity of the user as well as confirming the integrity of the acoustic seal and the working condition of itself and the ECR. The ECM 123 can also be used for capturing voice that is resonant within the ear canal when the user is speaking to permit voice communication.

[0028] The ASM 111 is housed in an ear seal 113 and monitors sound pressure at the entrance to the occluded or partially occluded ear canal. The ASM 111 can also be used to capture the user's voice externally for permitting voice communication. All transducers shown can receive or transmit audio signals to a processor 121 that undertakes audio

signal processing and provides a transceiver for audio or voice via the wired or wireless communication path 119.

[0029] The earpiece 100 can actively monitor a sound pressure level both inside and outside an ear canal and enhance spatial and timbral sound quality while maintaining supervision to ensure safes sound reproduction levels. The earpiece 100 in various embodiments can conduct listening tests, filter sounds in the environment, monitor warning sounds in the environment, present notification based on identified warning sounds, maintain constant audio content to ambient sound levels, and filter sound in accordance with a Personalized Hearing Level (PHL).

[0030] Referring to **FIG. 2**, a block diagram of the earpiece 100 in accordance with an exemplary embodiment is shown. As illustrated, the earpiece 100 can include a processor 206 operatively coupled to the ASM 111, ECR 125, and ECM 123 via one or more Analog to Digital Converters (ADC) 202 and Digital to Analog Converters (DAC) 203. The processor 206 can measure ambient sounds in the environment received at the ASM 111 and internal sounds captured at the ECM 130. Ambient sounds correspond to sounds within the environment such as the sound of traffic noise, street noise, conversation babble, or any other acoustic sound.

[0031] Ambient sounds measured by the ASM 111 can also correspond to industrial sounds present in an industrial setting, such as, factory noise, lifting vehicles, automobiles, and robots. The processor 206 can monitor the ambient sound captured by the ASM 110 for sounds in the environment, such as an abrupt high energy sound corresponding to an on-set of a warning sound (e.g., bell, emergency vehicle, security system, etc.), siren (e.g., police car, ambulance, etc.) , voice (e.g., “help”, “stop”, ”police”, etc.), or specific noise type (e.g., breaking glass, gunshot, etc.).

[0032] Internal sounds measured by the ECM 123 can correspond to sounds contained within the ear canal 131 such as spoken voice or audio content delivered by way of the ECR 125. The internal sounds can include residual background noise related to ambient sounds in the environment; for example, high level sounds that leak around the ear seal 127 and enter the ear canal 131. The processor 206 can monitor internal sounds captured by the ECM 123 and analyze the internal sounds. The processor 206 can also adjust a mixing between the ambient sound signals measured at the ASM 111 and the internal sound signals measured at the ECM 123, for example, responsive to assessing ambient background noise conditions.

[0033] The processor 206 can utilize computing technologies such as a microprocessor, Application Specific Integrated Chip (ASIC), and/or digital signal processor (DSP) with associated storage memory 208 such a Flash, ROM, RAM, SRAM, DRAM or other like technologies for controlling operations of the earpiece device 100. The memory 208 can store program instructions for execution on the processor 206 as well as captured audio processing data.

[0034] The memory 208 can also store program instructions for execution on the processor 206. The memory 208 can be off-chip and external to the processor 208, and include a data buffer to temporarily capture the ambient sound and the internal sound, and a storage memory to save from the data buffer the recent portion of the history in a compressed format responsive to a directive by the processor. The data buffer can be a circular buffer that temporarily stores audio sound at a current time point to a previous time point. It should also be noted that the data buffer can in one configuration reside on the processor 206 to provide high speed data access. The storage memory can be non-volatile memory such as SRAM to store captured or compressed audio data.

[0035] The memory 208 can be a machine-readable medium. The term "machine-readable medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed, and/or associated caches and external memory) that store the one or more sets of instructions. The term "machine-readable medium" shall also be taken to include any medium that is capable of storing, encoding or carrying a set of instructions for execution by the machine and that cause the machine to perform any one or more of the methodologies of the present disclosure.

[0036] The term "machine-readable medium" shall accordingly be taken to include, but not be limited to: solid-state memories such as a memory card or other package that houses one or more read-only (non-volatile) memories, random access memories, or other re-writable (volatile) memories; and/or magneto-optical or optical medium; and carrier wave signals such as a signal embodying computer instructions in a transmission medium. Accordingly, the disclosure is considered to include any one or more of a machine-readable medium or a distribution medium, as listed herein and including art-recognized equivalents and successor media, in which the software implementations herein are stored.

[0037] The earpiece 100 can include an audio interface 212 operatively coupled to the processor 206 to receive audio content, for example from a media player or cell phone, and deliver the audio content to the processor 206. The processor 206 responsive to detecting

ambient sounds can adjust the audio content and pass the ambient sounds directly to the ear canal. For instance, the processor 206 can lower a volume of the audio content played out the ECR 125 responsive to detecting an acute sound for transmitting the ambient sound to the ear canal. The processor 206 can also actively monitor the sound exposure level inside the ear canal via the ECM 123 and adjust the audio content to within a safe and subjectively optimized listening level range.

[0038] The earpiece 100 can further include a transceiver 204 that can support singly or in combination any number of wireless access technologies including without limitation Bluetooth™, Wireless Fidelity (WiFi), Worldwide Interoperability for Microwave Access (WiMAX), and/or other short or long range communication protocols. The transceiver 204 can also provide support for dynamic downloading over-the-air to the earpiece 100. It should be noted also that next generation access technologies can also be applied to the present disclosure.

[0039] The power supply 210 can utilize common power management technologies such as replaceable batteries, supply regulation technologies, and charging system technologies for supplying energy to the components of the earpiece 100 and to facilitate portable applications. A motor (not shown) can be a single supply motor driver coupled to the power supply 210 to improve sensory input via haptic vibration. As an example, the processor 206 can direct the motor to vibrate responsive to an action, such as a detection of a warning sound or an incoming voice call.

[0040] The earpiece 100 can further represent a single operational device or a family of devices configured in a master-slave arrangement, for example, a mobile device and an earpiece. In the latter embodiment, the components of the earpiece 100 can be reused in different form factors for the master and slave devices.

[0041] **FIG. 3** is an exemplary schematic of a software system 300 for operating the earpiece 100. The software system 300 can reside at least in part or whole on the processor 206, the memory 208, and/or any associated machine readable storage medium operated on by the processor 206 (see FIG. 2). Generally stated, the software system 300 by way of the processor 206 can manage a configuration of audio input and output (paths) to the earpiece 100 to support audio delivery.

[0042] As illustrated, the software system 300 includes a peripheral interface 310 configured to manage a plurality of sound signals 315 and direct the plurality of sound signals 315 to a plurality of audio processing modules 330. The sound signals can be an

ambient sound signal measured by the ASM 111 or an internal sound signal from measured by the ECM 123. The software system 300 can include an audio content interface 320 configured to manage a plurality of audio streams 325 and direct the plurality of audio streams 325 to the plurality of audio processing modules 330. The audio stream can be a voice signal from a Phone, a music signal from a personal media player (PMP), or an audio signal provided by the earpiece (e.g., loopback signal). The plurality of sound signals 315 and the plurality of audio streams 325 are also passed to a filter set 350 as shown by the wide arrows.

[0043] The plurality of audio processing modules 330 can produce audio control information 331 responsive to an analysis of the sound signals. The audio control information 331 is provided to the logic control unit 340 for configuring filtering and mixing operations on the plurality of sound signals 315 and the plurality of audio streams 325 at a filter set 350. The logic control 340 unit is operatively coupled to the plurality of audio processing modules 330 to receive the audio control information 331 from the plurality of audio processing modules 330 and generate configuration data. The configuration data can include filter data 342 for processing the plurality of sound signals, audio control data 343 for assigning a priority to the plurality of sound signals, and router data 344 for mixing a plurality of filtered signals 355 according to the priority

[0044] The software system 300 includes the filter set 350 operatively coupled to the peripheral interface to process the plurality of sound signals 315 in accordance with the configuration data 342 for generating the filtered sound signals 355. The filter set 350 also processes the plurality of audio streams 325 in accordance with the configuration data 344 to produce filtered sound signals 355. The filtered sound signals 355 are passed to a mixing matrix 360.

[0045] The mixing matrix 360 is operatively coupled to the control logic unit 340 and filter set 350 to mix the plurality of filtered sound signals 355 in accordance with the configuration data 344 and generate output sound signals 365 that are routed to at least one output peripheral component 370. The peripheral component 370 can be the ECR 125, a phone, or a data storage. The mixing matrix 360 can include individual gains that are applied to the plurality of filtered sound signals 355 for producing the output sound signals.

[0046] **FIG. 4** is a more detailed schematic of the software system 300 of FIG. 3. System components of FIG. 3 will be referred to when describing components of the detailed schematic.

[0047] The peripheral interface 310 (see FIG. 3) can include the Ambient Sound Microphone (ASM) 111 configured to convert an ambient sound to an ambient sound signal, and the Ear Canal Microphone (ECM) 123 configured to convert an internal sound from an ear canal of a user to an internal sound signal. The peripheral interface 310 receives the ASM sound signals from the external environment including background noise, warning sounds, and the ECM sound signal including the voice of a wearer of the earpiece 100 and any audio content playing out the ECR 125. The components (e.g., ASM-R, ECM-R) of the peripheral interface 310 can contain more than the components shown, for instance, a secondary ASM for sound localization or noise suppression. The peripheral interface 310 can also include components (ASM-L, ECM-L) from a second earpiece such as a L-left or R-right earpiece. In this case, the software system 300 can reside on one of the earpieces, although resources can be shared if the software system 300 is enabled on both earpieces.

[0048] The audio content interface 320 (see FIG. 3) can receive audio streams, for example, from a phone 421, a personal media player (PMP) 422, a portable communication device (e.g., VOP) 423, or a component 411 of the earpiece. The local component 411 (e.g., ECR 125) can permit audio feedback for allowing the user to hear audio, for example, loop back when the user is speaking on the phone, or playing comfort noise during non-speech intervals. The audio content interface (AC) 420 can mix the plurality of audio streams to produce a single audio stream delivered to the processing modules (AP1-AP5). In one arrangement, the AC 420 can perform audio content mixing based on a user context.

[0049] The user context identifies an operation or mode of the earpiece, such as, an incoming call, a music session, or a voice mail. For example, if the user is listening to music and an incoming call is detected, the earpiece 100 by way of the software system 300 (see FIG. 3) further discussed herein can lower the volume of the music relative to the incoming call ring tone to permit a hearing of the incoming call (e.g., ring tone). As another example, if the user is using the earpiece in a transparent "safe" mode while listening to music, then any harmful ambient sounds (e.g., loud jackhammer, bus noises) can be attenuated with the music, for instance, by reducing the ASM to ECM pass through levels. Similarly, important warning sounds captured at the ASM 111 can be elevated in the mix with respect to other sounds by increasing the ratio of ASM 111 to ECM 123 sound levels. The AC 420 can also mix the sounds in accordance with manual intervention, for example, if the user

adjusts a volume level manually by way of a user interface or directly mixes the audio streams.

[0050] Referring to both FIGS. 3 and 4, the sound signals from peripheral interface 310 and the mixed sound signals from the audio content interface 320 can be passed through a signal analysis module 410 operatively coupled to the peripheral interface 310 and the audio content interface 320 to provide a shared analysis of the sound signals and the audio streams. The sound signals and the shared analysis can then be sent to the plurality of audio processing modules (AP1 – AP5).The shared analysis can include spectral analysis, spectral band energy level analysis, spectral envelope analysis, voice activity detection analysis, and cross-correlation analysis.

[0051] The shared analysis module 410 permits a sharing of processing resources among the processing modules (AP1 – AP5) to minimize individual resource use of the processing modules from performing a common analysis. For instance, instead of each processing module performing an FFT analysis, the shared module 410 can perform an FFT and share output results with the modules. In another arrangement, output results from the shared analysis can be conveyed to the processing modules based on frequency band requirements or trigger events and thresholds. For instance, AP1 may register as an event listener for ASM signals that exceed a certain level in a frequency band; AP2 may register as an event listener to receive ECM signals that match a particular masking profile, and so on.

[0052] Alternatively, each processing module can interpret the output results individually and make their own respective determination as how to process the sound signals. For instance, each processing module (AP1 – AP5) can selectively process the plurality of sound signals in accordance with its own requirements. For example, AP1 can be a sound detection module to identify one or more sounds in the user's environment, for example, in accordance with the teachings presented in U.S. Patent Application No11/966,457 filed on December 28, 2007 entitled "Method and Device for Sound Signature Detection" herein incorporated by reference in its entirety. AP 2 can be a sound exposure monitoring module to assess safe listening levels, for instance, in accordance with the teachings presented in U.S. provisional patent application No. 60/887,165 filed on January 30, 2007 entitled "Sound Pressure Level Monitoring and Notification System" herein incorporated by reference in its entirety. AP 3 can be a sound enhancement module or voice control module, for instance, in accordance with the teachings presented in U.S. Provisional

Application No. 60/911,691 filed on April 13, 2007 entitled "Method and Device for Voice Operated Control", and U.S. Provisional Application No. 60/885,917 filed on January 22, 2007 entitled "Method and Device for Acute Sound Detection and Reproduction" herein incorporated by reference in its entirety. AP 4 can be a sound correction module to modify a sound signal based on a sound level exposure within safe listening levels, for instance, in accordance with the teachings presented in U.S. Provisional Application No. 60/866,420 filed on November 18, 2006, entitled "Method and Device for Personalized Hearing" the entire disclosure of which is incorporated herein by reference.

[0053] Each of the processing modules (AP1 – AP5) generates separate audio control information 331 in view of the audio signals that can be used by the logic control unit 340 to make an informed decision related to filtering and mixing the sound signals and audio streams. The audio control information 331 includes filter control data 342, audio control data 343, and router control data 344. The filter control data 342 can include filter coefficients that each processing module (AP) identifies as being significant to their intended function. The audio control data 343 can control a mixing of the plurality of audio streams at the AC 420 based on an established priority in view of the processing module decisions (e.g., warning signal detected, sound exposure level exceeded, echo feedback condition). The router control data 344 can include mixing gains (g1-g2) to amplify or attenuate the filtered sound streams.

[0054] The logic control unit 340 responsive to receiving the audio control information 331 generates a priority that is used for generating the configuration data (341,342 and 343) used to filter the sound signals and mix the filtered sound signals. The priority controls how the filtered signals are filtered at the filter set 350 (e.g. F1, F2, F3, see also FIG. 3). As illustrated, F1 is a first filter for the ambient sound signal, and F2 is a second filter for the internal sound signal in the ear canal. The logic control unit 340 varies the level of filtering at both F1 and F2 in accordance with the filter configuration data 342 as established by the determined priority. The filter coefficients for the F1 filter and F2 filter are provided on a frame-by-frame basis in real-time with the filter configuration data 342 from the logic control unit 340. The priority also controls how the audio steams are mixed in the audio control unit 320 (e.g., see VOP, Phone, PMP, see also FIG. 3) based on the audio control data, and how the filtered signals are mixed in the mixing matrix 360 (e.g., g1-g9, see also FIG. 3).

[0055] For example, an AP1 module configured for warning sound detection upon identifying a warning sound in an ASM sound signal, can generate filter coefficients to

amplify one or a group of frequency bands of the warning sound (e.g., horn or siren). The AP1 module can consider the other frequency bands as don't cares since they are not part of the warning sound. AP1 can also raise or flag a priority level that a warning sound has been detected. Similarly, AP3 module for speech enhancement may determine from a simultaneous ECM audio stream that the wearer is speaking and generate filter coefficients that accentuate certain portions of the sound spectrum. AP3 can raise a flag or priority level based on an energy level of the spoken voice. AP4 may determine an echo condition feedback that could potentially damage the users hearing and generate filter coefficients to null out the feedback. AP4 can raise a flag or priority level indicating an immediate danger condition. AP2 may determine that a sound level exposure is being exceeded and generate audio control data to turn volume down on a media player. The logic control unit 340 upon receiving the audio control information and corresponding priorities from the respective processing modules (AP1-AP5) can then determine the appropriate filter configuration given the filter coefficients and the priority.

[0056] Notably, the logic control unit 340 evaluates the audio control information from all the processing modules (AP1-APN) individually, and then collectively as a whole, for prioritizing the configuration data. For instance, warning sounds that are given a higher priority over voice from the ECM or music from the AC, will be mixed according to the priority. Thus, the output sound signal generated by the mixing matrix 360 emphasizes the warning sound, followed by the voice, followed by the audio content. If voice commands are given a higher priority over music, than the mixing matrix 360 can reduce the music levels from the AC module when an AP module detects voice. Again, the priority can be established manually (e.g., via user interface) or automatically (e.g., user context, or presence). As another example, the mixing matrix 360 can increase ASM to ECM pass through upon detection of a warning sound, while simultaneously lowering a phone volume if the user is checking voice mail messages. As yet another example, the mixing matrix 360 can reduce ASM to ECM pass and simultaneously lower a music level if an incoming phone call is detected. The priority can also be event driven responsive to detecting a sound signature, a background noise condition, a battery life indication, a manual interaction, or a voice recognition command.

[0057] In the foregoing, a detailed description for one exemplary operation of audio input/output configuration of the software system 400 shown in FIG. 4 is provided. It should

be noted that other configurations apply, for instance, using more or less than the number of mixing coefficients (g1-g9) or filters (F1-F9) shown.

[0058] As illustrated, The AC control unit 420 mixes audio content (e.g., phone, VOP, PMP) and outputs an AC signal (or pair of stereo channels if two earphone devices are used). The mixing is controlled accordance with configuration data 343 received from the control logic unit 340 and also automatically by the AC control unit 420. The AC signal generated from the AC 420 is then filtered by audio content filter F3. The internal ear-canal audio signal from the ECM 123 is filtered by ECM filter F1 205, and the ambient audio signal from the ASM is filtered by ASM filter F2. All three filters (F1, F2 and F3) are updated and controlled by the logic control unit 340 via filter configuration data 342.

[0059] The filtered audio signals 355 are then passed to the mixing matrix 360. The 9 mixing coefficients (g1-g9) of the mixing matrix 360 are controlled and updated by the control logic unit 340 via configuration data 344. Each mixing coefficient (e.g., g1-g9) can be a time-varying positive number, negative number, or zero. As part of the mixing, the output sound signals are routed to the three peripheral output components (Phone 421, ECR 125, and Storage 430). The mixing gains can be normalized so as to permit balanced audio delivery. The peripheral output components comprise the ear canal receiver ECR 125 for delivering audio to the ear canal, the output signal to the phone 421 (i.e. the signal that is transmitted to another remote individual), and the audio storage device 430 (e.g. hard drive on a PMP).

[0060] The filter configuration data 342 for the 2 filters, the AC configuration data 343 for the AC control unit, and the routing configuration data 344 for the mixing matrix 360 is generated responsive to receiving the audio control information from the plurality of processing modules (AP1-AP5). Recall, each module (e.g., modules 1-5, or more) receives as its inputs at least one of the unfiltered audio input signals (e.g. the ECM and ASM signals) or audio sound streams (e.g., music signal from PMP). As illustrated, each module analyzes the input audio signal and generates 3 output types: coefficients for the router (e.g., g1-g9), filter coefficients for the 3 signal filters (e.g., F1,F2 and F3), and audio control signals for the AC control unit. Each of these output types can have "don't care" states-which can be represented by a reserved number value (e.g., -1).

[0061] The control logic unit 340 prioritizes the sound signals and audio streams for each signal type, for instance, it can prioritize the router coefficient values (e.g., mixing gains g1-g9) higher for one AP module than the other. In some cases, the control logic can

maximally prioritize the signal type from one module so that the output router coefficients, for example, from one module is copied directly to the output of the control logic unit. In other cases, the output signals of the same type from different modules can be combined by a weighted addition. Similarly, the filter coefficients for a particular filter (F1 – F3) can be generated by a weighted addition (frequency domain) or convolution (time domain).

[0062] **FIG. 5** is a flowchart of a method 500 for generating filter configuration data in accordance with an exemplary embodiment. The method 500 be practiced with more or less than the number of steps shown and is not limited to the order shown. To describe the method 500, reference will be made to components of FIG 2, although it is understood that the method 500 can be implemented in any other manner using other suitable components. The method 500 can be implemented in a single earpiece, a pair of earpieces, headphones, or other suitable headset audio delivery device.

[0063] In particular, method 500 describes one exemplary arrangement in which the logic circuit 340 generates filter configuration data from the audio content received by individual AP modules to produce a combination filter output. In the exemplary embodiment, only one filter of the three filters shown in FIG. 4 is explained (e.g. generation of F1 coefficients for the ECM audio path). In other exemplary embodiments, the filter output for method 300 includes all 3 filters (F1, F2 and F3) , for instance, the ECM sound signal, the ASM sound signal and the AC sound signal.

[0064] The method 500 can start at step **502**. At step **504**, the logic control unit 340 can receive N filter vectors from the N processing modules **506** (AP1-APN). Each processing module generates a filter vector that is a set of filter coefficients (time or frequency domain). Each of the N processing modules (AP1-APN) also outputs an associated module priority.

[0065] At step **508**, the logic control unit 340 assesses the priorities raised by each processing module and evaluates the priorities against a corresponding reference module priority at step **510**. If a processing module does not assign a particular priority value, then a default priority value is retrieved from the reference module priority list. At step **512**, the logic control unit 340 identifies the processing module with the highest priority. If a first processing module returns the same module priority value as a second processing module, then the logic control unit 340 assigns a unique module priority value to both the second and first modules, for example, based on a user context.

[0066] Upon selecting the processing module with the highest priority, the logic control unit 340 proceeds to determine if any configuration data is a “don’t care” as shown in step

514. A "don't care" value is a value that does not affect the filter state of the respective processing module. The filter configuration data 342 generated by each processing module can include filter values of a time domain filter vector (e.g. for an FIR type filter) of frequency domain filter (e.g., bi-quads) with "don't care" value; the logic control unit 340 can assign a reserved numerical value of symbol for "don't care" filter states. A "dont care" flag can be associated with each filter output from each module.

[0067] At step **516**, the logic control unit 340 can copy the configuration data (e.g., filter data, router data, audio content data) for "care" states to internal memory for output to the filter set 350, mixing matrix 360, and audio control 420 (see FIG. 4). The logic control unit 340 can continue to update the configuration data for "care" states as it is retrieved from each of the respective processing modules according to the priorities. Starting with the next highest priority module, as shown in step **520**, the filter values for this high priority module are then checked at step **514**to see if they correspond to a "don't care" state in the method loop. The method loop continues until the end of the module priority list as shown in step **518**.

[0068] If the logic control unit 340 identifies a "don't care" state is present, the filter values for that state is not updated; else, the filter values for this module are copied to the output and no further update of the filters is conducted until new filter vectors are received from all modules as shown in step **522**. In such regard, the logic control unit 340 output filter states using filter values directly from the processing modules based strictly on priority. If all module filter values are "don't care" states, then the filter output is the same as a previous filter output. The method 500 continues back at step **504** to evaluate the priorities of each module in real-time as results are received. The logic control unit 340 thus continually updates the filter values in view of priority for the respective filters (F1, F2 and F3) to process audio signals and audio streams in real time.

[0069] **FIG. 6** is a flowchart of a method 600 for generating filter configuration data where filter data is collectively combined based on priority. The method 600 is an extension of method 500 of FIG. 5 and can practiced with more or less than the number of steps shown and is not limited to the order shown. Briefly, FIG. 6 presents an embodiment wherein processing modules can have a same priority value. In such a case when two modules have the same priority value, and neither contain (don't care" filter states, then the filter values from each of these two module with the same priority value are combined. The

combination can be a simple summation of filter weights (e.g. for frequency domain filters) or a convolution of the two sequences (e.g. for time domain weights).

[0070] The method 600 can start at step **602**. At step **604**, the logic control unit 340 can receive N filter vectors from the N processing modules **606** (AP1-APN). Each processing module generates a filter vector that is a set of filter coefficients (time or frequency domain). Each of the N processing modules (AP1-APN) also outputs an associated module priority. At step **608**, the logic control unit 340 assesses the priorities raised by each processing module. If a processing module does not assign a particular priority value, then a default priority value can be retrieved from a reference module priority list.

[0071] At step **610**, the logic control unit 340 identifies the processing module with the highest priority. The logic control unit 340 at step **618** then examines the audio control information from each of the N processing modules to identify "don't care" states. Filter values with "don't care" are not considered in generating output filter configuration data 342 or router configuration data 344 (see FIG. 4). Accordingly, the method 600 continues to get the next most important priority modules at step **620** until the end of module priority list is reached at step **622**.

[0072] Filter values other than "don't care" are however used in generating a combined filter output; that is, a filter output representative of the collective priority of all the processing modules (AP1-APN) based on their processing results. The combination also depends on whether the module is uniquely defined; for instance, it has precedence over combination with other filter vales. Thus at step **614**, the logic control unit 340 determines whether a processing module with a "care" filter value is uniquely defined. If yes, then the filter values for that unique model are copied to the output (e.g., router configuration, filter configuration data, audio configuration data) at step **616**. If however, the module is not unique, then the logic control unit 340 at step **612** combines the filter value with filter data from modules of the same priority. These filter values are then copied to the output at step **616** if no other modules eclipse their collective priority, or until a module with higher priority is identified. Upon evaluating all new filter from all the modules at step **624**, the method 600 can return back to step 604 to continue to retrieve the filter vectors in real-time.

[0073] FIG. 7 is an exemplary schematic 700 for configuring audio input and output via a mixing matrix in accordance with an exemplary embodiment. Reference will be made to FIG. 4 when describing the schematic 700. For instance, the filtered ECM signal 702 is the output of the ECM filter (F1) of FIG. 4. The mixing gains g1 (704), g4 (704) and g7 (707)

correspond to the first column of the mixing matrix 360 as shown in FIG. 4. The phone out 710, ECR 712, and audio storage 714 correspond to the output peripheral components (421, 125, and 430) of the mixing matrix 360.

[0074] The mixing gains (g1-g9) of the mixing matrix 360 can be applied to the filtered audio signals 355 (see FIG. 3) before they are routed to the corresponding peripheral output components (phone out 710, ECR 712, and audio storage 714). FIG. 7 shows a signal path diagram for the filtered ECM input signal 702 to the 3 peripheral output components of the matrix 260. The mixing gains can be normalized coefficients whose total contribution across a column of the mixing matrix is unity. For example, mixing gains of g1=0.2, g4=0.3, and g7=0.5 are applied to the filtered ECM signal 702 to produce phone out 710, ECR out 125, and audio storage out 714. The coefficient values can be time variant for each new input audio sample into the mixing matrix 360, or can remain constant for a block of input samples.

[0075] In one exemplary embodiment, the 9 coefficients (g1-g9) for each of the mixing gains in mixing matrix 360 can be generated using a similar logic process as described in FIG. 4 and FIG. 4, except the inputs to the logic system are not filter vectors but rather 9 coefficient values from each module. The priority value assigned to each module, or the priority value automatically generated by that module, can be a different value for the filter vectors than for the 9 coefficient values. Moreover, 9 coefficients are shown since there are 3 audio paths (ASM, ECM, and AC) and 3 output peripheral components (Phone, ECR, data storage). The number of mixing gains and structure of the mixing matrix 360 can thus be a function of the number of audio input and audio output paths.

[0076] FIG. 8 is another exemplary schematic 800 of the software system 300 providing separate analysis and re-synthesis modules for peripheral inputs in accordance with an exemplary embodiment. As illustrated, the schematic 800 combines the three audio filters F1, F2 and F3 with the shared analysis module 410 of the software system shown in FIG 4. Analysis/Re-synthesis unit 810 performs shared analysis of the ECM sound signal for processing modules (AP1-AP3). Analysis/Re-synthesis unit 820 performs shared analysis of the ASN sound signal for processing modules (AP1-AP3). Analysis/Re-synthesis unit 830 performs shared analysis of the Audio Content (AC) sound signal for processing modules (AP1-AP3).

[0077] The combined analysis/resysnthesis units 810, 820 and 830 filter the 3 input audio signals; the ECM signal; the ASM signal, and Audio Content (AC) signal. The AC

signal comprises at least one of; an earcon signal (e.g. a "low battery" auditory warning cue); a signal from a mobile telephone; and a stereo signal from a PMP (e.g. portable DVD player, music audio player etc). The output signal of the analysis/resynthesis units 810, 820 and 830 are passed to the mixing columns (g1, g4, g7; g2, g5, g8; and g3, g6, g9) of the mixing matrix 360. The mixing gains g1-g9 are controlled via the router configuration data from the logic control unit 340, which in turn were determined from the audio control information generated by the processing modules (AP1-AP3).

[0078] As illustrated the mixing matrix 360 can associate a phone with a first row of a linkage matrix, associates an ECR with a second row of a linkage matrix, associate an audio storage with a third row of a linkage matrix, associate an ECM with a first column of a linkage matrix, associate an ASM with a second column of a linkage matrix, and associate an AC controller with a third column of a linkage matrix. The mixing matrix can modify the output sound signal of the phone using values in the first row of the linkage matrix, modify the output sound signal of the ECR using values in the second row of the linkage matrix, and modify the output sound signal sent to the audio storage using values in the third row of the linkage matrix.

[0079] Recall each processing module has two or more output signals: filter gain coefficients which control the resynthesis of the signals in the analysis/resythesis units; 9 coefficients for the router, and audio configuration data for the audio control. The processing modules (AP1-AP3) create AC Control signals to control the routing of the input AC signals using AC control unit (420, see FIG. 4). The processing modules are configured and their internal state can be queried by the user interface 850 (e.g. a computer connected to the DSP unit via USB).

[0080] FIG. 9 is a more detailed schematic 900 of an analysis and re-synthesis unit for a particular peripheral input in accordance with an exemplary embodiment. For instance, the analysis and re-synthesis unit 910 corresponds to the ASM sound signal path of the software system 800 shown in FIG. 8. That is, the analysis and re-synthesis unit 910 can receive audio input 901 from the ASM 111. Notably, the schematic 900 can be applied for producing filtered output for also the ECM 123 sound signals and AC 420 sound signals (see FIG. 4) or other signals.

[0081] As illustrated in FIG. 9, the analysis and re-synthesis unit 910 incorporates a filter-bank 912 (e.g., cascade of band-pass filters (BPF)) structure for decomposing the ASM sound signal into a plurality of filter-banks. The filter-bank can be based on a linear

frequency scale or a non-linear frequency scale, such as, an octave, 1/3 octave, critical band, equivalent rectangular band (ERB), or melody (mel)-frequency scale. Each output of each filter-bank can then be presented to each of the N processing modules (AP1 920-AP2 930). As previously indicated, the processing modules 920 and 930 generate audio control information 331 (see FIG. 3) that the logic control unit 340 uses in generating filter configuration data 342 (see FIG. 4). In the embodiment shown, the filter configuration data 342 is used by the analysis and re-synthesis unit 910 to scale the filter-bank output signals. For instance, the filter-bank outputs are gain scaled to amplify or attenuate the individual filter-banks with the filter coefficients in the filter configuration data 342. The filter-bank output signals can then be summed to generate a composite filtered sound signal. The filtered sound signal is then presented to the mixing matrix (see FIG. 8) for routing and mixing to the one or more output peripheral components (e.g., Phone, ECR, data storage).

[0082] In the following, a detailed description for operation of the analysis and re-synthesis unit 910 for use with an audio input signal is provided. The input signal 901 is called “audio input 1”, and can be any one of the exemplary three signals described previously (i.e. ECM, ASM or AC signal).The audio input signal is split and processed by the different band-pass-filters 912 (i.e. with different centre frequencies). In one arrangement, the response of each respective band pass filter 912 is such that if the filter outputs are summed, then the power spectral density is the same as the original audio input signal. The band pass filters 912 can be time-domain FIR type filters or IIR type filters, or the analysis could be with a frequency domain FFT. The output signal of each band-pass filter is fed to at least one module (for the sake of clarity, 2 modules are shown in the exemplary embodiment of figure 8). Each module then creates an output signal to control the re-sythesis of the band-pass filtered input signals. This output signal is a vector of weights to control the gains of the filtered signals (i.e. the number of gains is equal to the number of filter used in the analysis). The logic control unit 340 determines a single set of gain coefficients for the gains (i.e. G1, G2, G3 and G4 in the exemplary embodiment of figure 8). After the first set of band pass filtered signals have been multiplied with the gains (G1-G4) to create a second set of amplified (or attenuated) band-pass filtered signal, the second set of signals are summed with the summing unit to create a single filtered ASM output sound signal, which form one of the three exemplary inputs to the previously described mixing matrix 360.

[0083] While the present invention has been described with reference to exemplary embodiments, it is to be understood that the invention is not limited to the disclosed exemplary embodiments. The scope of the following claims is to be accorded the broadest interpretation so as to encompass all modifications, equivalent structures and functions of the relevant exemplary embodiments. Thus, the description of the invention is merely exemplary in nature and, thus, variations that do not depart from the gist of the invention are intended to be within the scope of the exemplary embodiments of the present invention. Such variations are not to be regarded as a departure from the spirit and scope of the present invention.

CLAIMS

What is claimed is,

1. An earpiece, comprising:

a peripheral interface configured to manage a plurality of sound signals and direct the plurality of sound signals to a plurality of audio processing modules that produce audio control information responsive to an analysis of the plurality of sound signals;

a logic control unit operatively coupled to the plurality of audio processing modules to receive the audio control information from the plurality of audio processing modules and generate configuration data;

a filter set operatively coupled to the peripheral interface to process the plurality of sound signals in accordance with the configuration data to produce filtered sound signals; and

a mixing matrix operatively coupled to the control logic unit and filter set to mix the plurality of filtered sound signals in accordance with the configuration data to produce output sound signals and route the output sound signals to at least one peripheral component.

2. The earpiece of claim 1, wherein the peripheral component is an Ear Canal Receiver (ECR), a phone, a portable communication device, or a storage.

3. The earpiece of claim 1, wherein the peripheral interface comprises:

at least one Ambient Sound Microphone (ASM) configured to convert an ambient sound to an ambient sound signal; and

at least one Ear Canal Microphone (ECM) configured to convert an internal sound from an ear canal of a user to an internal sound signal.

4. The earpiece of claim 1, further comprising:

an audio content interface configured to manage a plurality of audio streams and direct the plurality of audio streams to the plurality of audio processing modules.

5. The earpiece of claim 4, wherein the audio content interface receives an audio stream from a phone, a media player, or a portable communication device.

6. The earpiece of claim 4, wherein the audio content interface mixes the plurality of audio streams based on a user context that is one among an incoming call, a music session, or a voice mail.

7. The earpiece of claim 1, wherein the configuration data comprises filter data for processing the plurality of sound signals, audio control data for assigning a priority to the plurality of sound signals, and router data for mixing the plurality of filtered signals according to the priority

8. The earpiece of claim 4, wherein the priority is event driven responsive to detecting a sound signature, a background noise condition, a battery life indication, a manual interaction, or a voice recognition command.

9. The earpiece of claim 3, wherein the filter set comprises:

a first filter for the ambient sound signal; and

a second filter for the internal sound signal,

where the filter coefficients for the first filter and second filter are provided on a frame-by-frame basis in real-time with the configuration data from the logic control unit.

10. An earpiece, comprising:

a peripheral interface configured to manage a plurality of sound signals and direct the plurality of sound signals to a plurality of audio processing modules;

an audio content interface configured to manage a plurality of audio streams and direct the plurality of audio streams to the plurality of audio processing modules;

at least one signal analysis module operatively coupled to the peripheral interface and audio content interface to provide a shared analysis of the sound signals and the audio streams for the plurality of audio processing modules;

a logic control unit operatively coupled to the plurality of audio processing modules and the at least one signal analysis module to receive the shared analysis and audio control information to generate configuration data;

a filter set operatively coupled to the peripheral interface to process the plurality of sound signals in accordance with the configuration data to produce filtered sound signals; and

a mixing matrix operatively coupled to the control logic unit and filter set to mix the plurality of filtered sound signals in accordance with the audio control information to produce output sound signals and route the output sound signals to at least one peripheral component.

11. The earpiece of claim 10, wherein the shared analysis comprises spectral analysis, spectral band energy level analysis, spectral envelope analysis, voice activity detection analysis, and cross-correlation analysis.

12. The earpiece of claim 10, wherein a separate signal analysis module is provided for components of the peripheral interface and the audio content interface and is shared among the plurality of audio processing modules.

13. The earpiece of claim 10, wherein the peripheral interface comprises

at least one Ambient Sound Microphone (ASM) coupled to an ASM signal analysis module configured to analyze an ambient sound signal; and

at least one Ear Canal Microphone (ECM) coupled to an ECM signal analysis module configured to analyze an internal sound from an ear canal of a user.

14. The earpiece of claim 10, wherein the audio content interface is coupled to an audio content (AC) signal analysis module and configured to analyze an audio stream from a phone, a media player, or a portable communication device.

15. The earpiece of claim 10, wherein the logic control unit produces a linkage matrix of mixing gains that the mixing matrix applies to the plurality of filtered sound signals for producing the plurality of output sound signals.

16. The earpiece of claim 10, wherein the mixing matrix:

associates a phone with a first row of a linkage matrix;

associates an ECR with a second row of a linkage matrix;

associates an audio storage with a third row of a linkage matrix;

associates an ECM with a first column of a linkage matrix;

associates an ASM with a second column of a linkage matrix; and

associates an AC controller with a third column of a linkage matrix.

17. The earpiece of claim 16, wherein the mixing matrix:

modifies the output sound signal of the phone using values in the first row of the linkage matrix;

modifies the output sound signal of the ECR using values in the second row of the linkage matrix; and

modifies the output sound signal sent to the audio storage using values in the third row of the linkage matrix.

18. A method for configuring audio delivery on an earpiece, the method comprising the steps of:

receiving at least one sound signal and at least one audio stream;

performing an analysis of at least one sound signal and the at least one audio stream;

presenting the analysis, the at least one sound signal, and the at least one audio stream to a plurality of audio processing modules that generate configuration data responsive to the receiving;

filtering the at least one sound signal and at least one audio stream according to the configuration data to produce filtered sound signals;

mixing the filtered signals according to the configuration data to produce output sound signals; and

routing the output sound signals to at least one peripheral component.

19. The method of claim 18, wherein the at least one sound signal is an ambient sound signal, an ear canal sound signal, and the at least one audio stream is received from a phone, a media player, or a portable communication device.

20. The method of claim 18, comprising generating a linkage matrix of mixing gains based on the configuration data that are applied to the plurality of filtered sound signals for producing the plurality of output sound signals.

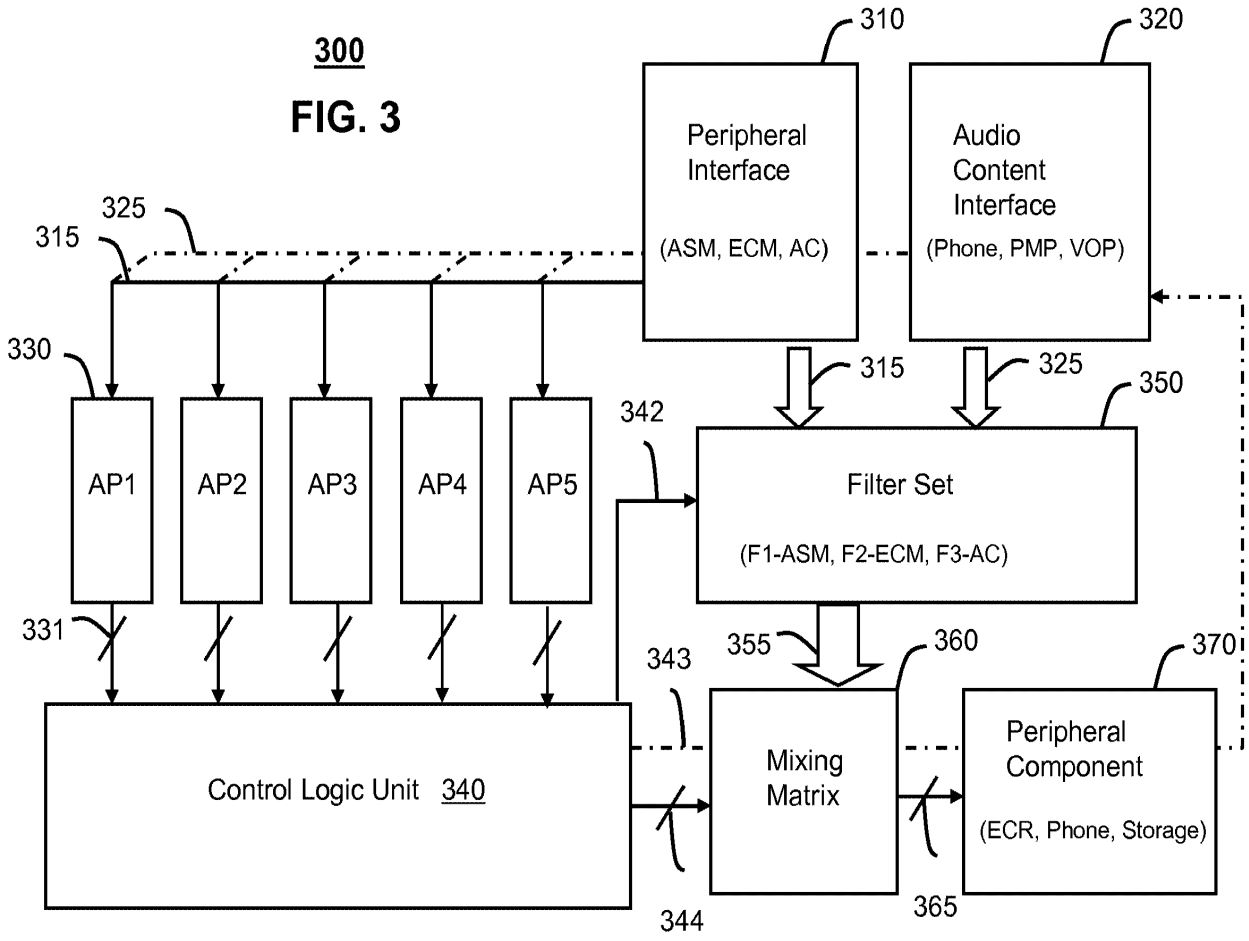
FIG. 1

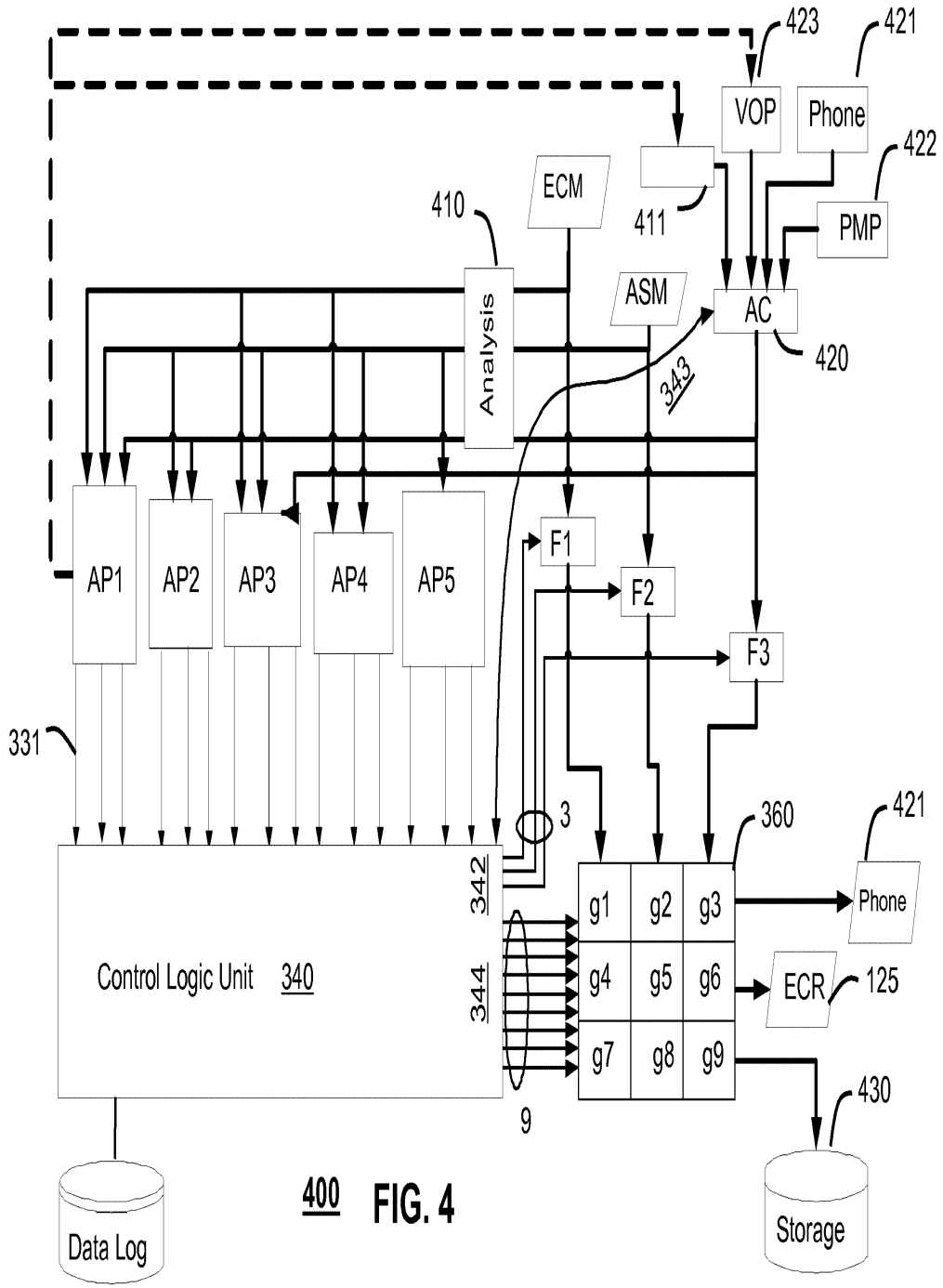
Ambient Sound
111
ASM
ADC 202
Transceiver 204
Processor 206
Memory 208
Power Supply 210
Audio Interface 212
Audio Content (e.g., music, cell phone, voice mail)
ADC 202
123
ECM
DAC 203
Audio
ECR
125
100

FIG. 2

300

FIG. 3


310
320
Peripheral Interface
(ASM, ECM, AC)
Audio Content Interface
(Phone, PMP, VOP)
325
315
330
315
325
350
342
AP1
AP2
AP3
AP4
AP5
Filter Set
(F1-ASM, F2-ECM, F3-AC)
331
355
360
370
343
Control Logic Unit 340
Mixing Matrix
Peripheral Component
(ECR, Phone, Storage)
344
365


423
421
VOP
Phone
422
ECM
411
410
PMP
ASM
AC
Analysis
420
343
AP1
AP2
AP3
AP4
AP5
F1
F2
F3
331
3
360
421
342
g1
g2
g3
Phone
Control Logic Unit 340
344
g4
g5
g6
ECR
125
g7
g8
g9
9
430
Data Log
Storage

400 FIG. 4

Start
502
506
Module 1
Module 2
Module N
Get N filter vectors from N modules
504
508
510
Get module priority list
Reference Module priority list
Get highest priority module number (M)
512
M
514
516
Module M filter "Don't Care"?
N
Copy Module 1 filters to output of Logic Circuit
Y
520
Get next most important module (module M)
518
End of Module Priority list?
N
Y
522
New filters from all modules?
NO
Y

500

FIG. 5

600
FIG. 6
Start
602
606
Module 1
604
Get N filter vectors from N modules
Module N
608
Get module priority list
610
Get highest priority module number(M)
612
Combine filter from the modules with same priorities
N
616
620
618
614
Get next most important priority module(s)
Module M filter "Don't Care?
N
Unique module priority value?
Y
Copy Module1 filters to output of Logic Circuit
Y
622
End of Module Priority list?
N
Y
624
New filters from all modules?
NO

701
g1
704
g4
706
g7
702
Filtered ECM signal
X
710
PhoneOut
712
ECR
714
Audio storeage
700

FIG. 7

VOP
Phone
ECM
AC
PMP
ASM
Analysis/ Resynthesis 810
Analysis/ Resynthesis 820
Analysis/ Resynthesis 830
123
AP1
AP2
AP3
User Interface 850
Control Logic Unit
123
g1
g2
g3
g4
g5
g6
g7
g8
g9
Phone
ECR
Storage
800

FIG. 8

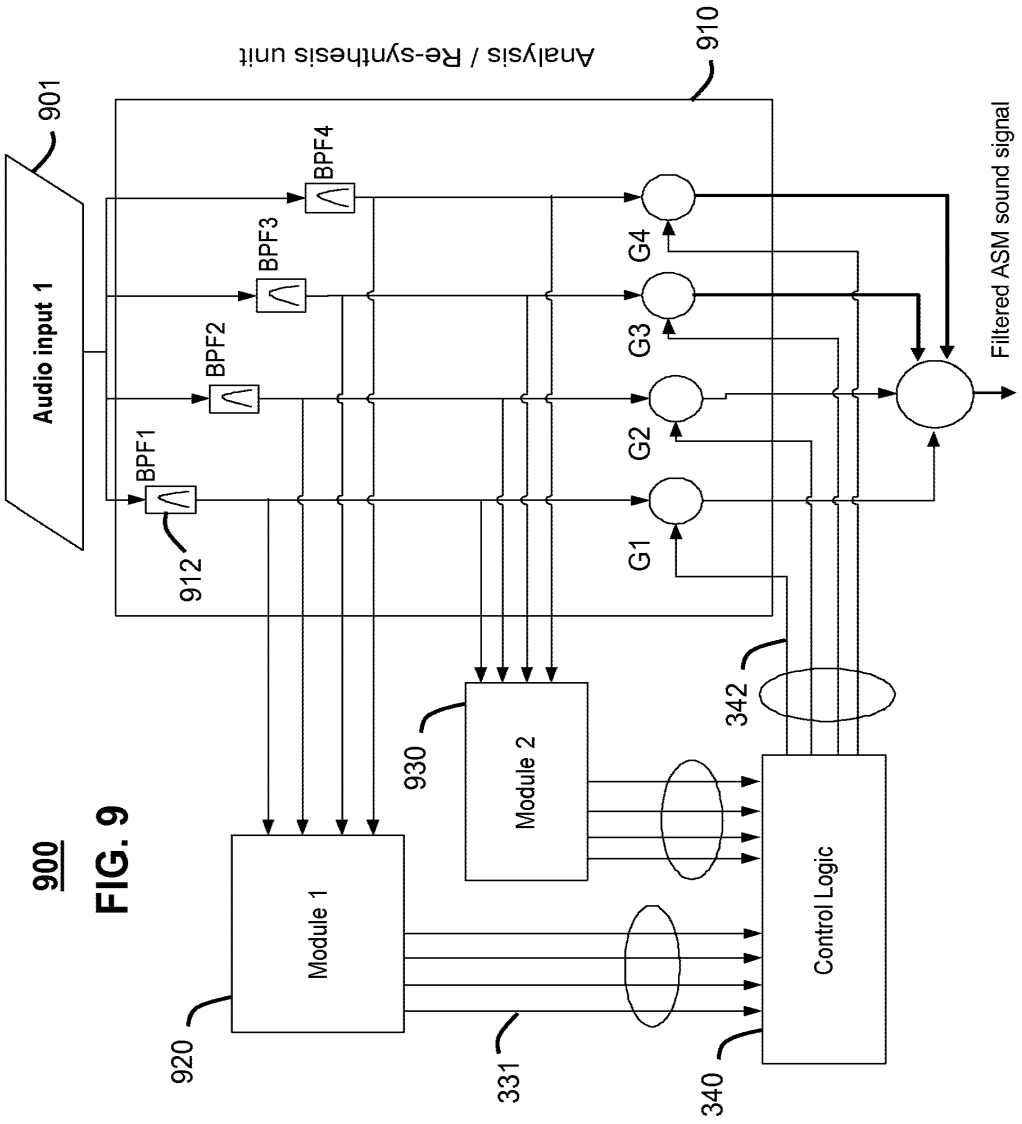


FIG. 9

Exhibit 1004
Page 0737 of 1001

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 08/73189

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - H04R 25/00 (2008.04)
USPC - 381/320
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
USPC: 381/320

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 381/312-331 (see terms below)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST(PGPB,USPT,USOC,EPAB,JPAB), GOOGLE SCHOLAR
terms: mix, matrix, linkage, spectral analysis, ambient, environmental, signal, sound, ear canal, meatus, in-the-ear, earpiece, earplug, earbud, priority, cellular phone.

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 2005/0058313 A1 (Victorian et al.) 17 March 2005 (17.03.2005), fig. 1, 4a, para [0004]-[0005], [0030]-[0032], [0042]-[0052], [0063].	1-8, 10, 12-20 ---------- 9, 11
Y	US 6,754,359 B1 (Svean et al.) 22 June 2004 (22.06.2004), fig. 1-2, col 4, ln 39 - col 5, ln 3, col 7, ln 26 - col 8, ln 17.	9
Y	US 2006/0262944 A1 (Rasmussen et al.) 23 November 2006 (23.11.2006), fig. 1 and 4, para [0008], [0021]-[0023].	11

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 October 2008 (02.10.2008)	22 OCT 2008
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (April 2007)

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 July 2012 (19.07.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/097150 A1

(51) International Patent Classification:
H04B 1/00 (2006.01)

(21) International Application Number:
PCT/US2012/021077

(22) International Filing Date:
12 January 2012 (12.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/432,016 12 January 2011 (12.01.2011) US

(71) Applicant *(for all designated States except US)*: **PERSONICS HOLDINGS, INC.** [US/US]; 4800 North Federal Highway, Suite A205, Boca Raton, FL 33431 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **USHER, John** [GB/GB]; 1 Lion Villa, Causeway, Beer, Devon EX123LB (GB). **GOLDSTEIN, Steven, W.** [US/US]; 6390 Maddalena Place, Delray Beach, FL 33446 (US). **CASALI, John, G.** [US/US]; 385 Emerald Boulevard, Christiansburg, VA 24073 (US).

(74) Agents: **SPADT, Jonathan, H.** et al.; RatnerPrestia, P.O. Box 980, Valley Forge, PA 19482 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

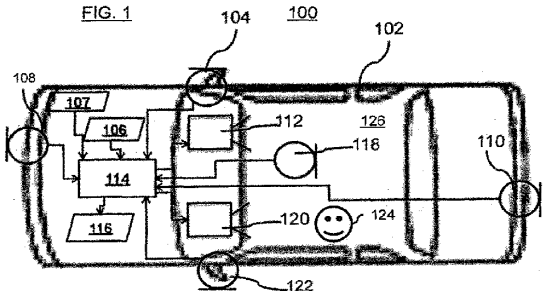
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) **Title**: AUTOMOTIVE SOUND RECOGNITION SYSTEM FOR ENHANCED SITUATION AWARENESS

FIG. 1



(57) **Abstract**: Sound recognition systems for a vehicle and methods for increasing auditory situation awareness in a vehicle are provided. A sound recognition system includes at least one ambient sound microphone (ASM), at least one vehicle cabin receiver (VCR) and a processor. The ASM is disposed on the vehicle and configured to capture ambient sound external to the vehicle. The VCR is configured to deliver audio content to a vehicle cabin of the vehicle. The processor coupled to the at least one ASM and the at least one VCR. The processor is configured to detect at least one sound signature in the ambient sound and to adjust the audio content delivered to the vehicle cabin based on the detected at least one sound signature.

WO 2012/097150 A1

AUTOMOTIVE SOUND RECOGNITION SYSTEM FOR ENHANCED SITUATION AWARENESS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is related to and claims the benefit of U.S. Provisional Application No. 61/432,016 entitled "AUTOMOTIVE SOUND RECOGNITION SYSTEM FOR ENHANCED SITUATION AWARENESS" filed on January 12, 2011, the contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

[0002] The present invention relates to a device that monitors sound directed to a vehicle cabin, and more particularly, though not exclusively, to an audio system and method that detects ambient warning sounds and adjusts audio delivered to a vehicle cabin based on the detected warning sounds to enhance auditory situation awareness.

BACKGROUND OF THE INVENTION

[0003] People that use audio systems in vehicles generally do so for either music enjoyment or voice communication. The user is generally immersed in the audio experience when using such devices and is acoustically isolated within a sealed vehicle cabin. Background noises in the external vehicle environment (e.g., road, engine, wind and traffic noise) can contend with the acoustic sounds produced from these devices. As the background noise levels change, the user may need to adjust the volume to listen to their music over the background noise. Alternatively, the level of reproduced audio may be automatically increased, for example, by audio systems that increase the audio level as the vehicle velocity increases (i.e., to compensate for the rise in noise level from road, engine and aerodynamic noise). One example of such an automatic gain control system is described in US patent No. 5,081,682.

SUMMARY OF THE INVENTION

[0004] Aspects of the present invention relate to a sound recognition system for a vehicle. A sound recognition system includes at least one ambient sound microphone (ASM), at least one vehicle cabin receiver (VCR) and a processor. The ASM is disposed on the vehicle and configured to capture ambient sound external to the vehicle. The VCR is configured to deliver audio content to a vehicle cabin of the vehicle. The processor coupled to the at least one ASM and the at least one VCR. The processor is configured to detect at least one sound signature in the ambient sound and to adjust the audio content delivered to the vehicle cabin based on the detected at least one sound signature.

[0005] Aspects of the present invention also relate to methods for increasing auditory situation awareness in a vehicle. The method includes capturing ambient sound external to the vehicle from at least one ambient sound microphone (ASM) disposed on the vehicle; monitoring the ambient sound for a target sound by detecting a sound signature

corresponding to the target sound in the ambient sound; and adjusting a delivery of audio content by at least one vehicle cabin receiver (VCR) to a vehicle cabin of the vehicle based on the target sound.

[0006] Aspects of the present invention further relate to methods for sound signature detection for a vehicle. The method includes capturing ambient sound external to the vehicle from at least one ambient sound microphone (ASM) disposed on the vehicle; and receiving a directive to learn a sound signature within the ambient sound. A voice command or an indication from a user is received and is used to initiate the steps of capturing and learning.

[0007] Aspects of the present invention also relate to methods for personalized listening in a vehicle. The method includes capturing ambient sound external to the vehicle from at least one ambient sound microphone (ASM) disposed on the vehicle; detecting a sound signature within the ambient sound that is associated with a warning sound; and mixing the warning sound with audio content delivered to the vehicle cabin via at least one vehicle cabin receiver (VCR) in accordance with a priority of the warning sound and a personalized hearing level (PHL).

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The invention may be understood from the following detailed description when read in connection with the accompanying drawing. It is emphasized, according to common practice, that various features of the drawings may not be drawn to scale. On the contrary, the dimensions of the various features may be arbitrarily expanded or reduced for clarity. Moreover, in the drawing, common numerical references are used to represent like features. Included in the drawing are the following figures:

[0009] FIG. 1 is a pictorial diagram of a vehicle including an exemplary automotive sound recognition system for enhanced situation awareness in accordance with an embodiment of the present invention;

[0010] FIG. 2 is a block diagram of the system shown in FIG. 1 in accordance with an exemplary embodiment of the present invention;

[0011] FIG. 3 is a flowchart of an exemplary method for ambient sound monitoring and warning detection in accordance with an embodiment of the present invention;

[0012] FIG. 4 illustrates various system modes in accordance with an exemplary embodiment of the present invention;

[0013] FIG. 5 is a flowchart of an exemplary method for sound signature detection in accordance with an embodiment of the present invention;

[0014] FIG. 6 is a flowchart of an exemplary method for managing audio delivery based on detected sound signatures in accordance with an embodiment of the present invention;

[0015] FIG. 7 is a flowchart of an exemplary method for sound signature detection in accordance with an embodiment of the present invention;

[0016] FIG. 8 is a pictorial diagram for mixing ambient sounds and warning sounds with audio content in accordance with an exemplary embodiment of the present invention; and

[0017] FIG. 9 is a flowchart of an exemplary method for updating the sound signature detection library dependent on the vehicle location in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0018] The following description of at least one exemplary embodiment is merely illustrative in nature and is in no way intended to limit the invention, its application, or uses.

[0019] Processes, techniques, apparatus, and materials as known by one of ordinary skill in the relevant art may not be discussed in detail but are intended to be part of the enabling description where appropriate, for example the fabrication and use of transducers. Additionally in at least one exemplary embodiment the sampling rate of the transducers can be varied to pick up pulses of sound, for example less than 50 milliseconds.

[0020] In all of the examples illustrated and discussed herein, any specific values, for example the sound pressure level change, should be interpreted to be illustrative only and non-limiting. Thus, other examples of the exemplary embodiments could have different values.

[0021] Note that herein when referring to correcting or preventing an error or damage (e.g., hearing damage), a reduction of the damage or error and/or a correction of the damage or error are intended.

[0022] Automotive vehicle operators are often auditorially removed from their external ambient environment. Ambient sound cues such as from oncoming emergency (and non-emergency) vehicle sound alerts are often not heard by the vehicle operator due to acoustic isolation of the vehicle cabin and internal cabin noise from engine and road noise and, especially, due to loud music and speech reproduction levels in the vehicle cabin.

[0023] Accordingly, background noises in the external vehicle environment contend with the acoustic sounds produced from the vehicle audio system. The vehicle operator, thus, becomes auditorially disassociated with their ambient environment, thereby increasing the danger of accidents from collisions with oncoming vehicles. A need therefore exists for improving the auditory situation awareness of vehicle operators to automatically alert the operator to ambient warning alerts.