Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2400104 A1 2001/03/29

(21) 2 400 104

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2001/01/05
(87) Date publication PCT/PCT Publication Date: 2001/03/29
(85) Entrée phase nationale/National Entry: 2002/08/13
(86) N° demande PCT/PCT Application No.: CH 2001/000008
(87) N° publication PCT/PCT Publication No.: 2001/020965

(51) Cl.Int.^{7}/Int.Cl.^{7} H04R 25/00

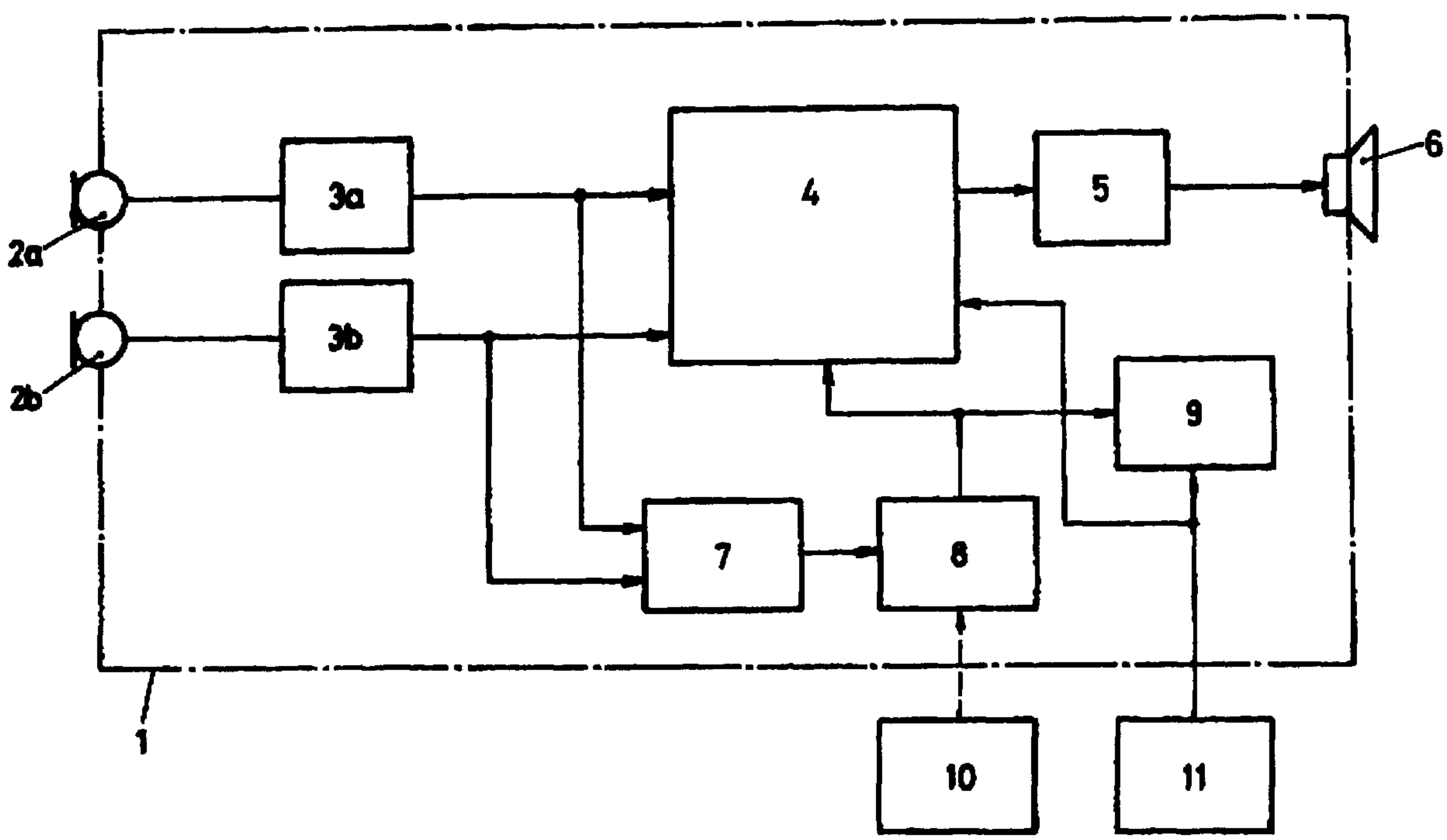
(71) Demandeur/Applicant:
PHONAK AG, SZ

(72) Inventeurs/Inventors:
ALLEGRO, SILVIA, CH;
BUCHLER, MICHAEL, CH

(74) Agent: ROBIC

(54) Titre : PROCEDE DE DETERMINATION D'UNE SITUATION D'ENVIRONNEMENT ACOUSTIQUE MOMENTANEE, UTILISATION DE CE PROCEDE, ET PROTHESE AUDITIVE

(54) Title: METHOD FOR DETERMINING A CURRENT ACOUSTIC ENVIRONMENT, USE OF SAID METHOD AND A HEARING-AID



(57) Abrégé/Abstract:
The invention relates to a method for determining a current acoustic environment. The method is characterised in that specific characteristics are extracted from an acoustic signal which has been recorded using at least one microphone (2a, 2b) and that the current acoustic environment is determined in an identification phase, on the basis of the extracted characteristics. According to the invention, at least auditory-based characteristics are determined in the extraction phase. The invention also relates to the use of said method and to a hearing-aid.

CA 02400104 2002-08-13

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
29. März 2001 (29.03.2001)

PCT

(10) Internationale Veröffentlichungsnummer
WO 01/20965 A3

(51) Internationale Patentklassifikation[7]**:** **H04R 25/00**

(21) Internationales Aktenzeichen: PCT/CH01/00008

(22) Internationales Anmeldedatum: 5. Januar 2001 (05.01.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(71) Anmelder *(für alle Bestimmungsstaaten mit Ausnahme von US)*: **PHONAK AG** [CH/CH]; Laubisrütistrasse 28, CH-8712 Stäfa (CH).

(72) Erfinder; und
(75) Erfinder/Anmelder *(nur für US)*: **ALLEGRO, Silvia** [CH/CH]; Pfannenstielstrasse 3, CH-8618 Oetwil am See (CH). **BÜCHLER, Michael** [CH/CH]; Lindenbachstrasse 50, CH-8006 Zürich (CH).

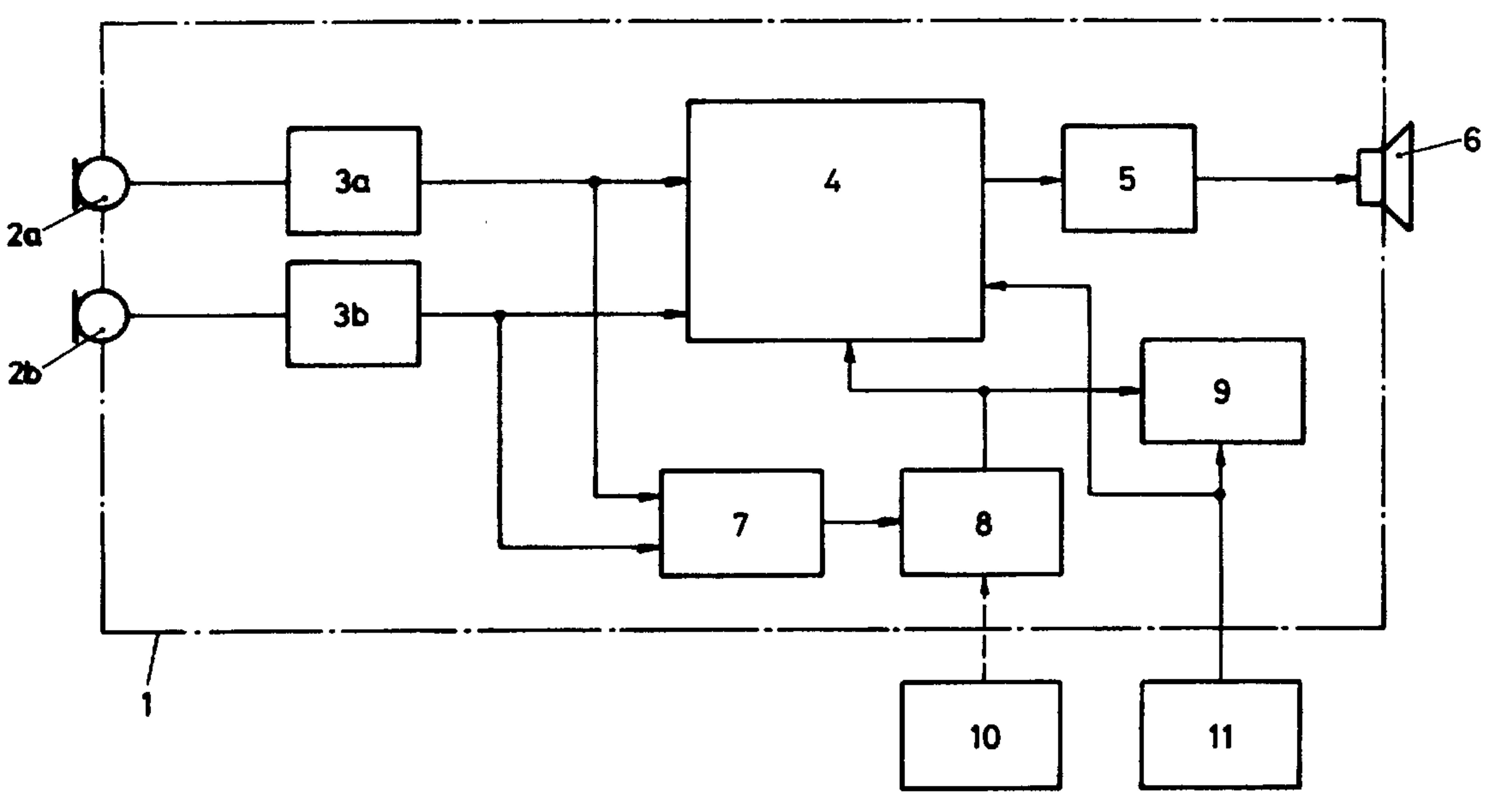
(74) Anwalt: RIGLING, Peter; Troesch Scheidegger Werner AG, Schwäntenmos 14, CH-8126 Zumikon-Zürich (CH).

(81) Bestimmungsstaaten *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR DETERMINING A CURRENT ACOUSTIC ENVIRONMENT, USE OF SAID METHOD AND A HEARING-AID

(54) Bezeichnung: VERFAHREN ZUR BESTIMMUNG EINER MOMENTANEN AKUSTISCHEN UMGEBUNGSSITUATION, ANWENDUNG DES VERFHARENS UND EIN HÖRGERÄT



(57) Abstract: The invention relates to a method for determining a current acoustic environment. The method is characterised in that specific characteristics are extracted from an acoustic signal which has been recorded using at least one microphone (2a, 2b) and that the current acoustic environment is determined in an identification phase, on the basis of the extracted characteristics. According to the invention, at least auditory-based characteristics are determined in the extraction phase. The invention also relates to the use of said method and to a hearing-aid.

(57) Zusammenfassung: Die Erfindung betrifft zunächst ein Verfahren zur Bestimmung einer momentanen akustischen Umgebungssituation, wobei das Verfahren darin besteht, dass in einer Extraktionsphase charakteristische Merkmale aus einem mit mindestens einem Mikrophon (2a, 2b) aufgenommenen akustischen Signal extrahiert

[Fortsetzung auf der nächsten Seite]

WO 01/20965 A3

WO 01/20965 A3

MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Bestimmungsstaaten *(regional)*: ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:
— *mit internationalem Recherchenbericht*
— *vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen*
— *auf Antrag des Anmelders, vor Ablauf der nach Artikel 21 Absatz 2 Buchstabe a geltenden Frist*

(88) Veröffentlichungsdatum des internationalen Recherchenberichts: 11. April 2002

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

werden, und dass in einer Identifikationsphase die momentane akustische Umgebungssituation aufgrund der extrahierten Merkmale ermittelt wird. Erfindungsgemäss werden in der Extraktionsphase zumindest auditorisch-basierte Merkmale ermittelt. Des weiteren ist eine Verwendung des erfindungsgemässen Verfahrens sowie ein Hörgerät angegeben.

METHOD FOR DETERMINING A CURRENT ACOUSTIC ENVIRONMENT, USE OF SAID METHOD AND A HEARING-AID

The present invention is related to a method for determining a momentary acoustic situation, a use of the method for hearing devices and a hearing device.

Modern-day hearing devices, when employing different hearing programs - typically two to a maximum of three such hearing programs - permit their adaptation to varying acoustic situations. The idea is to optimize the effectiveness of the hearing device for its user in all situations.

The hearing program can be selected either via a remote control or by means of a switch on the hearing device itself. For many users, however, having to switch program settings is a nuisance, or difficult, or even impossible. Nor is it always easy even for experienced wearers of hearing devices to determine at what point in time which program is most comfortable and offers optimal speech discrimination. An automatic recognition of the acoustic situation and corresponding automatic switching of the hearing program in the hearing device is therefore desirable.

At present, there exist several different approaches to the automatic classification of acoustic situations. All of the methods concerned involve the extraction of different features from the input signal which may come from one or several microphones in the hearing device. Based on these features, a pattern-recognition device employing a particular algorithm makes a determination as to the attribution of the analyzed signal to a specific acoustic situation. These various existing methods differ from one another both in terms of the features on the basis of which they define the acoustic situation (signal analysis) and with regard to the pattern-recognition device which serves to classify these features (signal identification).

For the extraction of features in audio signals, J.M. Kates in his article entitled "Classification of Background Noises for Hearing-Aid Applications " (1995, Journal of the Acoustic Society of America 97(1), pp 461 - 469), suggested an analysis of time-related sound-level fluctuations and of the spectrum. Furthermore, an analysis of the amplitude histogram is proposed in the European patent having the number EP-B1-0 732 036 in order to reach the same goal. Finally, the extraction of features has been investigated and implemented based on an analysis of different modulation frequencies. In this connection, reference is made to the two papers by Ostendorf et. al. entitled "Empirical Classification of Different Acoustic Signals and of Speech by Means of a Modulation-Frequency Analysis" (1997, DAGA 97, pp 608 - 609), and "Classification of

Acoustic Signals Based on the Analysis of Modulation Spectra for Application in Digital Hearing Devices" (1998, DAGA 98, pp 402 - 403). A similar approach is disclosed in an article by Edwards et. al. entitled "Signal-processing algorithms for a new software-based, digital hearing device" (1998, The Hearing Journal 51, pp 44 - 52). Other possible features include the sound-level itself of the zero-crossing rate as described e.g. in the article by H.L. Hirsch entitled "Statistical Signal Characterization" (Artech House 1992). It is evident that the features used to date for the analysis of audio signals are strictly technically based.

It is fundamentally possible to use prior art pattern identification methods for sound classification purposes. Particularly suitable pattern recognition systems are the so-called distance classifiers, Bayes classifiers, fuzzy logic systems or neural networks. Details of the first two of the methods mentioned are contained in the publication entitled "Pattern Classification and Scene Analysis" by Richard O. Duda and Peter E. Hart (John Wiley & Sons, 1973). For information on neural networks, reference is made to the treatise by Christopher M. Bishop, entitled "Neural Networks for Pattern Recognition" (1995, Oxford University Press). Reference is also made to the following publications: Ostendorf et. al., "Classification or Acoustic Signals Based on the Analysis of Modulation Spectra for Application in Digital Hearing Devices" (Zeitschrift für Audiologie (Journal of Audiology), pp 148

- 150); F. Feldbusch, "Sound Recognition Using Neural Networks" (1998, Journal of Audiology, pp 30 - 36); European patent application having publication number EP-A1-0 814 636; and US patent having publication number US-5 604 812. Yet all of the pattern-recognition methods mentioned are deficient in one respect in that they merely model static properties of the sound categories of interest.

The known methods for sound classification, involving feature extraction and patter recognition, have the drawback that, although unambiguous and solid identification of speech signals is basically possible, a number of different acoustic situations cannot be satisfactorily classified, or not at all. While these known methods permit a distinction between pure speech signals and "non-speech" signals - meaning all other acoustic situations -, it is not enough for selecting an optimal hearing program for a momentary acoustic situation. As a result thereof, the number of possible hearing programs is either limited to those two automatically recognizable acoustic situations or either the hearing device wearer himself has to recognize the acoustic situations that are not covered while manually selecting the appropriate hearing program.

It is therefore the object of the present invention to introduce first of all a method for determining a momentary acoustic situation, which method, compared to prior art methods, is substantially more reliable and more precise.

This is accomplished by measures specified in claim 1. Advantageous embodiments of the present invention, a use of the method as well as a hearing device are specified in additional claims.

The invention is based on an extraction of signal features with the subsequent separation of different sound sources as well as an identification of different sounds. Instead of or besides technically-based features, auditory-based features are taken into account in the signal analyzing process for the feature extraction. These auditory-based features are being determined by the method of the Auditory Scene Analysis (ASA). In a further embodiment of the method according to the present invention, a grouping of the features using the Gestalt principles is employed in a context-dependent or context-independent manner. The identification or classification, respectively, of the audio signals is performed by using, in a preferred embodiment, Hidden Markov Models (HMM) applied on the extracted features. The present invention has the advantage that the number of recognizable sound categories and therewith the number of hearing programs is increased. As a result thereof, the performance of the sound classification and therewith the comfort is increased for the user of the hearing device.

The following will explain the present invention in more detail by way of an example with reference to a drawing. The only figure is a functional block diagram of a hearing device in which the method according to the present invention has been implemented.

In the only figure, a hearing device is identified by 1. For the purpose of the following description, the term "hearing device" is intended to include hearing aids as used to compensate for the hearing impairment of a person, but also all other acoustic communication systems such as e.g. radio transceivers.

The hearing device 1 incorporates in conventional fashion two electro-acoustic converters 2a, 2b and 6, these being one or several microphones 2a, 2b and a speaker 6, also referred to as receiver. A main component of the hearing device 1 is a transmission unit identified by 4 in which transmission unit 4, in case of a hearing aid, signal modification takes place in adaptation to the requirements of the user of the hearing device 1. However, the operations performed in the transmission unit 4 are not only a function of the nature of a specific purpose of the hearing device 1 but are also, and particularly, a function of the momentary acoustic situation. For this reason, there have already been hearing aids on the market where the wearer can manually switch between different hearing programs tailored to specific acoustic situations. There also exist hearing aids capable of automatically recognizing the acoustic situation. In this connection, reference is again made to the European patents EP-B1- 0 732 036 and EP-A1-0 814 636 and to the US patent 5 604 812, as well as to the "Claro Autoselect" brochure by Phonak Hearing Systems (28148 (GB)/0300, 1999).

In addition to the aforementioned components - such as microphones 2a, 2b, the transmission unit 4 and the receiver 6 - the hearing device 1 contains a signal analyzing unit 7 and a signal identifying unit 8. If the hearing device 1 is based on digital technology, one or several analog-to-digital converters 3a, 3b are arranged in between the microphones 2a, 2b and the transmission unit 4 and one digital-to-analog converter 5 is provided in between the transmission unit 4 and the receiver 6. While a digital implementation of the invention is preferred, it is equally possible to use analog components throughout. In this case, of course, the converters 3a, 3b and 5 are not needed.

The signal analyzing unit 7 receives the same input signal as the transmission unit 4. Finally, the signal identifying unit 8, which is connected to the output of the signal analyzing unit 7, is connected to the transmission unit 4 and to a control unit 9.

A training unit is identified by 10, which training unit 10 serves to establish in off-line operation the parameters required in the signal identifying unit 8 for the classification process.

By means of a user input unit 11, the user can override the settings of the transmission unit 4 and the control unit 9 as established by the signal analyzing unit 7 and the signal identifying unit 8.

The method according to the present invention is explained as follows:

It is essentially based on the fact that characteristic features are extracted from an acoustic signal in an extraction phase, whereby, instead of or in addition to the technically-based features - such as the above mentioned zero-crossing rates, time-related sound-level fluctuations, different modulation frequencies, the sound level itself, the spectral peak, the amplitude distribution, etc. - auditory-based features are employed as well. These auditory-based features are determined by means of an Auditory Scene Analysis (ASA) and include in particular the loudness, the spectral pattern (timbre), the harmonic structure (pitch), common build-up and decay times (on-/offsets), coherent amplitude modulations, coherent frequency modulations, coherent frequency transitions, binaural effects, etc. Detailed descriptions of Auditory Scene Analysis can be found e.g. in the articles by A. Bregman, "Auditory Scene Analysis" (MIT Press, 1990) and

W.A. Yost, "Fundamentals of Hearing - An Introduction (Academic Press, 1977). The individual auditory-based features are described, inter alia, by A. Yost and S. Sheft in "Auditory Perception" (published in the "Human Psychophysics" by W.A. Yost, A.N. Popper and R.R. Fay, Springer 1993), by W.M. Hartmann in "Pitch, Periodicity, and Auditory Organization" (Journal of the Acoustical Society of America, 100(6), pp 3491 - 3502, 1996), and by D.K. Mellinger and B.M. Mont-Reynaud in "Scene Analysis" (published in "Auditory Computation" by H.L. Hawkins, T.A. McMullen, A.N. Popper and R.R. Fay, Springer 1996).

As an example for the use of auditory-based features in signal analysis, the characterization of the tonality of the acoustic signal is now mentioned by analyzing the harmonic structure, which is particularly useful for the identification of tonal signals such as speech and music.

A further embodiment of the method according to the present invention, a grouping of the features is furthermore employed in the signal analyzing unit 7 by using of Gestalt principles. This process applies the principles of the Gestalt theory, by which such qualitative properties as continuity, proximity, similarity, common fate, unity, good continuation and others are examined, to the auditory-based or perhaps technically-based features for the creation of auditory objects. This grouping - and, for that matter, the extraction of features in the extraction phase - can take place in context-independent fashion, i.e. without any enhancement by additional knowledge (so-called "primitive" grouping), or in context-dependent fashion in the sense of human auditory perception employing additional information or hypotheses regarding the signal content (so-called "schema-based" grouping). This means that the context-dependent grouping is adapted to any given acoustic situation. For a detailed explanation of the principles of the Gestalt theory and of the grouping process employing Gestalt analysis, substitutional reference is made to the publication entitled "Perception Psychology" by E.B. Goldstein (Spektrum Akademischer Verlage, 1997), "Neural Fundamentals of Gestalt Perception" by A.K. Engel and W. Singer (Spektrum der Wissenschaft, 1998, pp 66 - 73), and "Auditory Scene Analysis" by A. Bregman, (MIT Press, 1990).

The advantage of applying this grouping process lies in the fact that it allows further differentiation of the features of the input signals. In particular, signal segments are identifiable, which originate in different sound sources. The extracted features can thus be mapped to specific individual sound sources, providing additional information on these sources and, hence, on the current auditory situation.

The second aspect of the method according to the present invention as described here relates to pattern recognition, i.e. the signal identification that takes place during the identification phase. The preferred form of implementation of the method according to the present invention employs the Hidden Markov Model (HMM) method in the signal identifying unit 8 for the automatic classification of the acoustic situation. This also permits the use of time changes of the computed characteristics for the classification process. Accordingly, it is possible to also take into account dynamic and not only static properties of the acoustic situation and the sound categories. Equally possible is a combination of HMMs with other classifiers such as multi-stage recognition processes for identifying the acoustic situation.

The output signal of the signal identifying unit 8 thus contains information on the nature of the acoustic surroundings (the acoustic situation). This information is fed to the transmission unit 4 in which the program or set of parameters is selected most suitable for the identified acoustic situation for the transmission. At the same time, the information gathered in the signal identifying unit 8 is fed to the control unit 9 for further actions whereby, depending on the situation, any given action, such as an acoustic signal, can be triggered.

If Hidden Markov Models are being used in the identification phase, it will require a complex process for establishing the parameters needed for the classification. This parameter determination is therefore best done in the off-line mode, individually for each category at a time. The actual identification of various acoustic situations requires very little memory space and computational capacity. It is therefore recommended that a training unit 10 be provided which has enough computing power for parameter determination and which can be connected via appropriate means to the hearing device 1 for data transfer purposes. The connecting means mentioned may be simple wires with suitable plugs.

The method according to the present invention thus makes it possible to select from among numerous available settings and automatically pollable actions the one best suited without the need for the user of the device to make the selection. This makes the device significantly more comfortable for the user since upon the recognition of a new acoustic situation it promptly and automatically selects the right program or function in the hearing device 1.

The user of hearing devices often want to switch off the automatic recognition of the acoustic situation and corresponding automatic program selection, described above. For this purpose a user input unit 11 is provided by means of which it is possible to override the automatic response or program selection. The user input 11 may be in the form of a switch on the hearing device 1 or a remote control which the user can operate. There are also other options which offer themselves, for instance a voice-activated user input device.

CLAIMS:

1. Method for determining momentary acoustic, said method consists in that

- characteristic features are extracted from an acoustic signal captured by at least one microphone (2a, 2b) during an extraction phase, and

- said momentary acoustic situation is determined on the basis of the extracted features during an identification phase,

by determining at least auditory-based features during the extraction phase.

2. Method according to claim 1, characterized in that, for the identification of the characteristic features during the extraction phase, Auditory Scene Analysis (ASA) techniques are employed.

3. Method according to claim 1 or 2, characterized in that ASA-(Auditory Scene Analysis) methods are being used to determine the characteristic features during the extraction phase.

4. Method according to one of the claims 1 to 3, characterized in that one or several of the following auditory-based features are identified during the extraction of said characteristic features: Loudness, spectral pattern, harmonic structure, common on- and offsets, coherent amplitude modulations, coherent frequency modulations, coherent frequency transitions and binaural effects.

5. Method according to one of the preceding claims, characterized in that any other suitable features are identified in addition to the auditory-based features.

6. Method according to one of the preceding claims, characterized in that, to create auditory objects, the auditory-based and any other features are grouped along the principles of the Gestalt theory.

7. Method according to claim 6, characterized in that the extraction of features and/or the grouping of the features are/is performed either in context-independent or in context-dependent fashion in the sense of human auditory perception, based upon additional information or hypotheses relative to the signal content and providing an adaptation to the respective acoustic situation.

8. Method according to one of the preceding claims, characterized in that, during the identification phase, data are accessed which were acquired in an off-line training phase.

9. Method according to one of the preceding claims, characterized in that the extraction phase and the identification phase take place in continuous fashion or at regular or irregular time intervals.

10. Use of the method according to one of the claims 1 to 9 to adjust a hearing device (9) to a momentary acoustic situation.

11. Use according to claim 10, characterized in that, on the basis of a detected momentary acoustic situation, a program or a transmission function between at least one microphone (2a, 2b) and a receiver (6) in the hearing device (1) is selected.

12. Use according to claim 9 or 10, characterized in that, in response to a detected momentary acoustic situation, any other function is triggered in the hearing device (1).

13. Use of the method according to one of the claims 1 to 9 to recognize speech.

14. Hearing device (1) with a transmission unit (4) which is, on its input side, operationally connected to at least one microphone (2a, 2b) and, on its output side, to a receiver (6), characterized in that the input signal of the transmission unit (4) is simultaneously fed to a signal analyzing unit (7) for the extraction of at least auditory-based features, and that the signal analyzing unit (7) is operationally connected to a signal identifying unit (8) in which the momentary acoustic situation is determined, and that the signal identifying unit (8) is operationally connected to the transmission unit (4) to adjust a program or a transfer function.

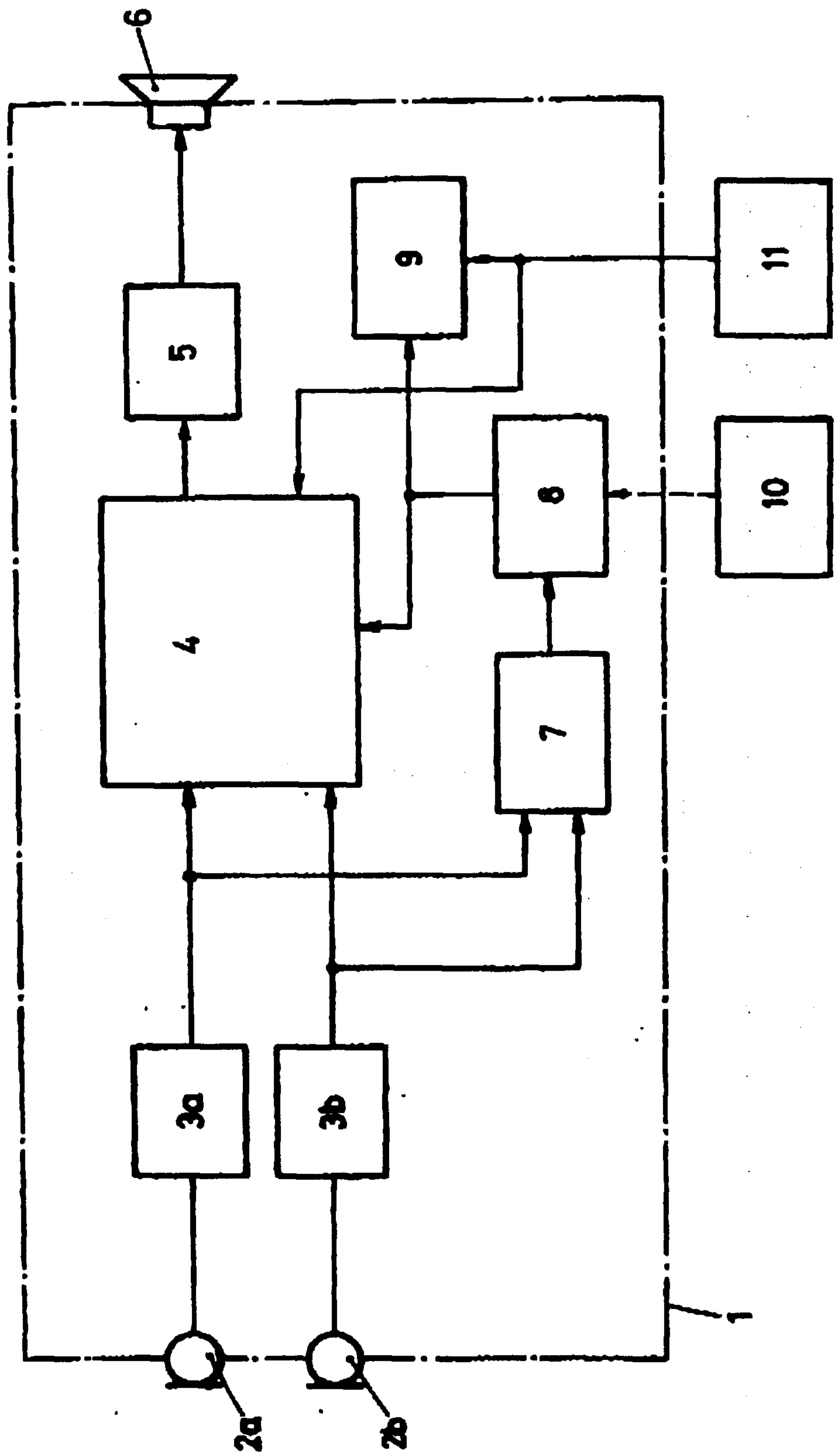
15. Hearing device (1) according to claim 14, characterized in that a user input unit (11) is provided which is operationally connected to the transmission unit (4).

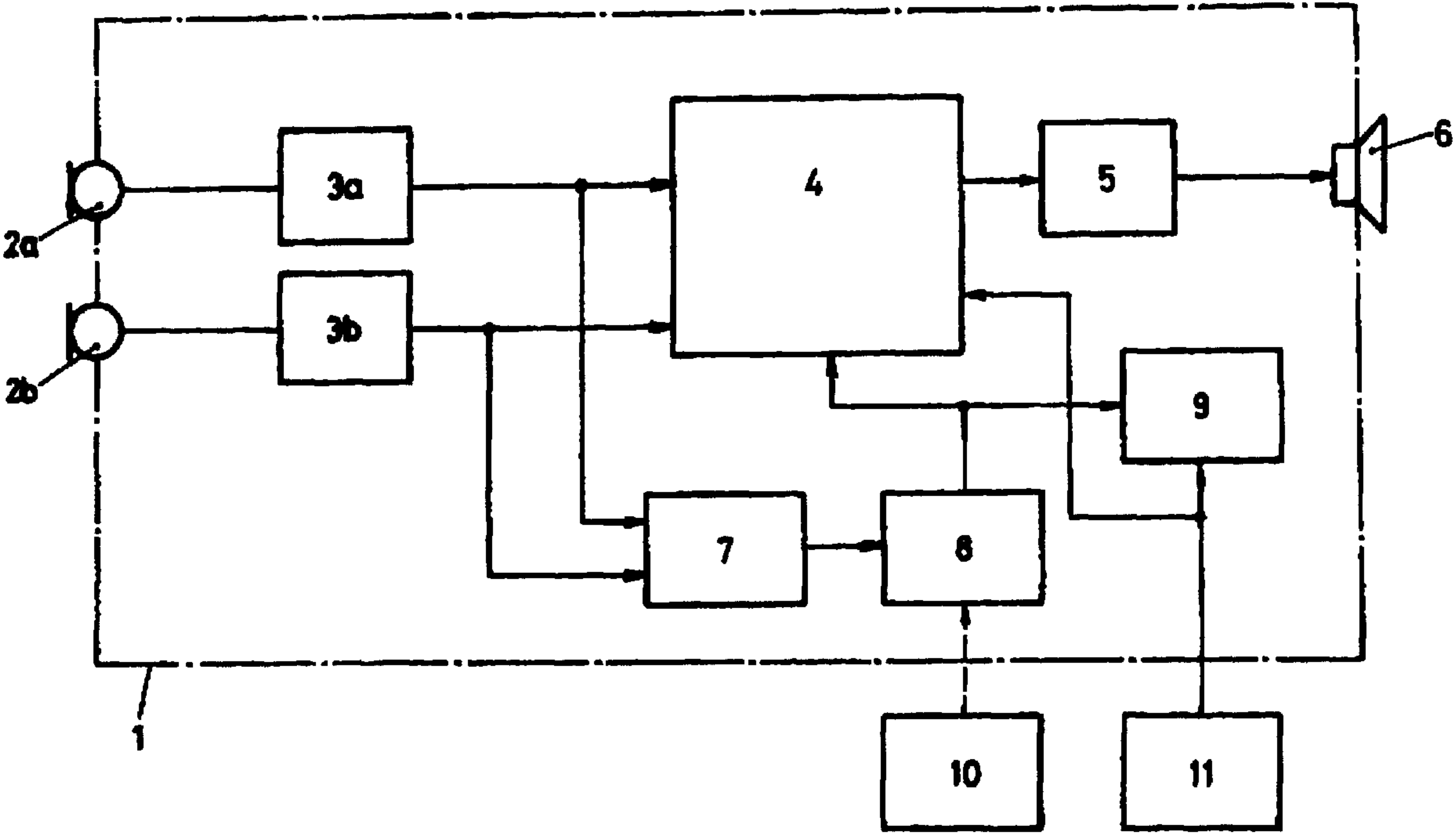
16. Hearing device (1) according to claim 14 or 15, characterized in that a control unit (9) is provided and that the signal identifying unit (8) is operationally connected to said control unit (9).

17. Hearing device (1) according to claim 15 or 16, characterized in that the user input unit (11) is operationally connected to the control unit (9).

18. Hearing device (1) according to one of the claims 14 to 17, characterized in that any means serving to transfer

parameters from a training unit (10) to the signal identifying unit (8) are provided.


6
5
9
11
8
10
4
7
3a
3b
2a
2b
1


2a
2b
3a
3b
4
5
6
7
8
9
10
11
1