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

(11) **EP 3 624 465 A1**

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
18.03.2020 Bulletin 2020/12

(51) Int Cl.:
H04R 25/00 *(2006.01)* *G10L 21/0316* *(2013.01)*
G10L 21/0208 *(2013.01)*

(21) Application number: **18193769.9**

(22) Date of filing: **11.09.2018**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

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

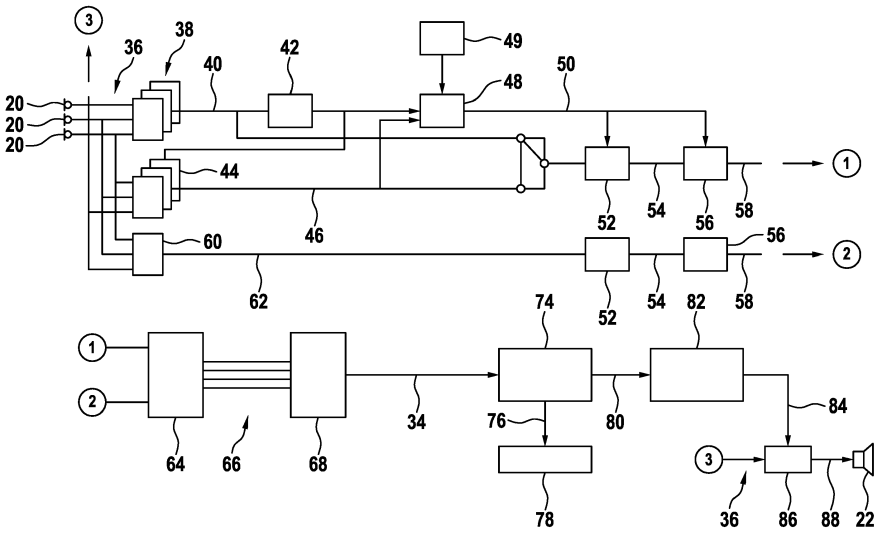
(72) Inventors:
- **Krueger, Harald**
  8910 Affoltern am Albis (CH)
- **Sigwanz, Ullrich**
  8634 Hombrechtikon (CH)
- **El Guindi, Nadim**
  8049 Zürich (CH)
- **Carstens, Mareike**
  8053 Zürich (CH)

(74) Representative: **Qip Patentanwälte**
Dr. Kuehn & Partner mbB
Goethestraße 8
80336 München (DE)

(54) HEARING DEVICE CONTROL WITH SEMANTIC CONTENT

(57) A method for directionally amplifying a sound signal (36) of a hearing device (12) comprises: receiving the sound signal (36) from a microphone (20) of the hearing device (12); extracting a user voice signal (62) and directional sound signals (40, 46) from the sound signal (36); determining a word sequence (54) from the user voice signal (62) and each directional sound signal (40, 46); determining a semantic representation (58) from each word sequence (54); identifying conversations (34) from the semantic representations (58), wherein each conversion (34) is associated with one or more directional sound signals (40, 46) and wherein each conversation (34) is identified by clustering semantic representations (58); and processing the sound signal (36), such that directional sound signals (40, 46) associated with one of the conversations (34) are amplified.

Fig. 3



EP 3 624 465 A1

Description

TECHNICAL FIELD

[0001] The invention relates to a method, a computer program and a computer-readable medium directionally amplifying a sound signal of a hearing device. Furthermore, the invention relates to a hearing system.

BACKGROUND

[0002] Hearing devices are generally small and complex devices. Hearing devices can include a processor, microphone, speaker, memory, housing, and other electronical and mechanical components. Some example hearing devices are Behind-The-Ear (BTE), Receiver-In-Canal (RIC), In-The-Ear (ITE), Completely-In-Canal (CIC), and Invisible-In-The-Canal (IIC) devices. Some hearing devices may compensate a hearing loss of a user.

[0003] In a situation with multiple voice sources it is typically not clear which voice sources the user of a hearing device wants to hear and which not, therefore it is difficult to optimize a directivity of the hearing device. As a rule, hearing devices steer the directivity in such situations either to the front (narrow or broad) or to a sector, where the sound sources are dominant.

[0004] US 6 157 727 A shows a hearing aid interconnected with a translation system.

DESCRIPTION

[0005] It is an objective of the present disclosure to support a user of a hearing device in situations with multiple voice sources. It is a further objective to better control the directivity of a hearing device in such situations.

[0006] These objectives are achieved by the subject-matter of the independent claims. Further exemplary embodiments are evident from the dependent claims and the following description.

[0007] A first aspect of the present disclosure relates to a method for directionally amplifying a sound signal of a hearing device. A hearing device may be a device worn by a user, for example in the ear or behind the ear. A hearing device may be a hearing aid adapted for compensating a hearing loss of the user.

[0008] According to an embodiment, the method comprises: receiving the sound signal from a microphone of the hearing device. The sound signal may be a digital signal. The sound signal may be composed of data packets encoding volume and/or frequencies of the sound signal over time. The sound signal may comprise sound data from more than one microphone of the hearing aid.

[0009] According to an embodiment, the method comprises: extracting directional sound signals and optionally a user voice signal from the sound signal. Such an extraction may be performed with spatial sound filters of the hearing device, which can extract sound from a specific direction from a sound signal. Such sound filters may comprise beam formers, which are adapted for amplifying sound from a specific direction based on sound data from several microphones.

[0010] A directional sound signal may be associated with a direction and/or a position, for example a position of a sound source and/or speaker contributing to the directional sound signal.

[0011] According to an embodiment, the method comprises: determining a word sequence from each directional sound signal and optionally the user voice signal. This may be performed with automatic speech recognition. For example, the hearing aid and/or an evaluation system may comprise an automatic speech recognition module, which translates the respective sound signal into a word sequence. A word sequence may be encoded as character string.

[0012] According to an embodiment, the method comprises: determining a semantic representation from each word sequence. A semantic representation may contain information about the semantic content of the word sequence.

[0013] A semantic content may refer to a specific situation, which is talked about in a conversation, and/or a topic of a conversation, such as weather, holydays, politics, job, etc. A semantic representation may encode the semantic content with one or more values and/or with one or more words representing the situation/topic.

[0014] For example, a semantic representation may contain semantic weights for the semantic content of the word sequence. These weights may be determined with automated natural language understanding. For example, the hearing aid and/or the evaluation system may comprise a natural language understanding module, which translates the word sequence into a semantic representation. A semantic weight may be a value that indicates the probability of a specific semantic content.

[0015] A semantic representation may be a vector of weights, for example output by the natural language understanding module. The natural language understanding module may be or may comprise a machine learning module, which identifies words belonging to the same conversation situation, such as weather, holyday, work, etc.

[0016] As a further example, the semantic representation may contain a count of specific words in the word sequence. The relative count of words also may be seen as a semantic weight for the words.

[0017] A semantic representation of a word sequence also may contain the substantives or specific substantiates extracted from the associated word sequence.

[0018] According to an embodiment, the method comprises: identifying conversations from the semantic representations, wherein each conversion is associated with one or more directional sound signals and wherein each conversation is identified by clustering semantic representations. For example, the hearing aid and/or the evaluation system may comprise a clustering module, which

identifies clusters in the semantic representations.

[0019] In an example, the semantic representations may be clustered by their semantic weights. Distances of semantic representations in a vector space of weights may be compared and semantic representations with a distance smaller than a threshold may be clustered.

[0020] As a further example, when the semantic representations comprise extract substantives of the word sequences, a set of substantive-pairs for each pair of sound sources and/or speakers may be compiled by pairing each substantive of the first sound source/speaker with each of the second sound source/speaker. A set of probabilities may be determined for each substantive-pair by looking up each pair in a dictionary. After that a conversation probability that the first sound source/speaker and the second sound source/speaker are in the same conversation may be determined from set of probabilities. The conversation probability may be compared with each other and the sound sources/speakers may be clustered based on this.

[0021] The dictionary of substantive-pairs may be determined from a large number of conversation transcriptions, for example in a big data approach. From the conversation transcriptions, substantives may be extracted and for each pair of substantives a probability that they occur in the same conversation may be determined. From these pairs, the dictionary of substantive-pairs and associated probabilities may be compiled.

[0022] It also is possible that conversations are identified based on question-response patterns, semantic relations of content between sound sources, etc.

[0023] A conversation may be a data structure having references to the directional sound signals and/or their semantic representations and optionally to the user voice signal and its semantic representation. A conversation also may have a semantic representation of its own, which is determined from the semantic representations of its directional sound signals and optionally the user voice signal. Furthermore, a conversation may be associated with a direction and/or position.

[0024] According to an embodiment, the method comprises: processing the sound signal, such that directional sound signals associated with one of the conversations are amplified. The directional sound signals associated with the selected conversation may be amplified stronger than other sound signals from other conversations. The conversion to be amplified may be selected automatically, for example as the conversation being associated with the user voice signal, and/or by the user, for example with a mobile device that is in communication with the hearing device.

[0025] With the method, directionality of the hearing device may be steered and/or controlled, such that all sound sources belonging to the same conversation are within the focus of directivity and/or sound sources which do not belong to the same conversation are suppressed. Semantic content may be used to decide whether multiple speech sources belong to the same conversation or not.

[0026] From the semantic representations and the directions and/or positions of the sound sources, which contribute to the directional sound signals, a conversation topology may be created. It may be decided, which conversation the user of the hearing device participates and the hearing aid may be controlled, such that sound sources belonging to this conversation are amplified, while other sound sources are suppressed.

[0027] Furthermore, it is possible that a current sound situation and/or environment, in which the user is situated, is determined from the semantic representations and/or the semantic content of the conversation to optimize the hearing aid processing. A program of the hearing device for processing a specific sound situation may be selected and/or started, when the specific sound situation is identified based on the semantic representations. For example, when one of the speakers around the user says "this is a really nice restaurant...", then a restaurant sound situation may be identified and/or a program for a restaurant sound situation may be started in the hearing aid.

[0028] According to an embodiment, the method comprises: outputting the processed sound signal by the hearing device. The hearing device may comprise a loudspeaker, which may output the processed sound signal into the ear of the user. It also may be possible that the hearing device comprises a cochlea implant as outputting device.

[0029] According to an embodiment, an environment of the user is divided into directional sectors and each directional sound signal is associated with one of the directional sectors. For example, the environment may be divided into quadrants around the user and/or in equally angled sectors around the user. Each directional sound signal may be extracted from the sound signal by amplifying sound with a direction from the corresponding directional sector. For example, a beam former with a direction and/or opening angle as the sector may be used for generating the corresponding directional sound signal.

[0030] In summary, the conversation topology may be analyzed sector based, i.e. the conversations may be analyzed per sector.

[0031] According to an embodiment, sound sources are identified in the sound signal. This may be performed with time-shifts between signal sound data from different microphones. A position of each sound source may be determined by analysis of spatial cues, directional classification and/or other signal processing techniques to enhance spatial cues in diffuse environments, such as onset or coherence driven analysis or direct to reverberation ratio analysis. Each directional sound signal may be associated with one of the sound sources. Each directional sound signal may be extracted from the sound signal by amplifying sound with a direction towards the corresponding sound source. Again, this may be performed with a beam former with a direction towards the

sound source and/or an opening angle solely covering the sound source.

[0032] In summary, the conversation topology may be analyzed source based, i.e. the conversations may be analyzed per source. Each sound source may be detected, located and analyzed individually.

[0033] According to an embodiment, speakers are identified in the sound signal. For example, a speaker may be identified with characteristics of his/her voice. Such characteristics may be extracted from the sound signal. A number and/or positions of speakers may be determined based on speaker analysis, such as fundamental frequency and harmonic structure analysis, male/female speech detection, distribution of spatial cues taking into account head movements, etc.

[0034] Each directional sound signal may be associated with one of the speakers. Each directional sound signal is extracted from the sound signal by amplifying sound with characteristics of the speaker. A speaker may be treated as a sound source and/or a beam former with a direction towards the speaker and/or an opening angle solely covering the speaker may generate the directional sound signal.

[0035] In summary, the conversation topology may be analyzed speaker based, i.e. the conversations may be analyzed per speaker. Each speaker may be detected, located and analyzed individually.

[0036] According to an embodiment, the sound signal is processed by adjusting at least one of a direction and a width of a beam former of the hearing device. In the end, when the conversation to be amplified is identified, only or nearly only sound from the sound sources from the conversation may be forwarded to the user. This may be performed with a beam former that is directed towards the conversation.

[0037] As already mentioned, according to an embodiment, a user voice signal is extracted from the sound signal and a word sequence is determined from the user voice signal.

[0038] It may be that the user voice signal is determined in another way. For example, the user voice signal may be determined via bone conduction sensors, ear canal sensors and/or additional acoustic microphones.

[0039] According to an embodiment, the method comprises: automatically selecting a conversation to be amplified by selecting a conversation, which has the highest concordance with the semantic representation of the user voice signal. For example, there may be only one conversation, which is associated with the user. In this case, this conversation may be selected.

[0040] In general, every conversation has been identified by clustering semantic representations. A conversation may have the highest concordance with a specific semantic representation, when the semantic representations, from which the conversation has been determined, have the lowest distances from the specific sematic representation in the space, where the clustering has been performed.

[0041] According to an embodiment, the method comprises: storing semantic representations of a user of the hearing aid over time, and automatically selecting a conversation to be amplified by selecting a conversation, which has the highest concordance with the stored semantic representations. The hearing system may learn preferences and/or interests of the user from semantic content detection for optimizing the selection of the conversation.

[0042] According to an embodiment, the method comprises: presenting a visualization of the conversations to the user. The conversation topology may be visualized and/or displayed on a mobile device of the user, such as a smartphone. For example, the sectors, sound sources and/or speakers may be shown by icons on a map of the environment of the user. Also, the conversations may be visualized in the map.

[0043] The user may then select one conversation or sound source. This selection of the user for a conversation to be amplified is then received in the hearing system.

[0044] The visualization of the conversation topology, for example with a smartphone, may also allow the user to assign names to the detected speakers (e.g. spouse, grandchild). Additionally, an information flow between speakers may be visualized in the conversation topology map, such as active, inactive, muted speakers, etc.

[0045] It may be that the visualization ins performed in a virtual reality and/or augmented reality. Such a virtual and/or augmented reality may be presented to the user with glasses and/or lenses with displays- A conversation selection then may be done with voice and/or gesture control and/or gaze analysis.

[0046] According to an embodiment, the method comprises: detecting head movements of a user of the hearing device. This may be performed with a sensor of the hearing device, such as an acceleration sensor and/or magnetic field sensor, i.e. a compass. Then, directions and/or positions associated with the sectors, sound sources, speakers and/or conversations may be updated based on the head movements.

[0047] According to an embodiment, the method comprises: receiving a remote sound signal from an additional microphone in the environment of the user of the hearing device. Such a microphone may be communicatively connected with hearing device and/or the evaluation system, for example with a mobile device of the user. In this case, at least the directional sound signals may be determined from the sound signal of the microphone of the hearing device and the remote sound signal.

[0048] It may be possible that the steps of the method associated with automatic speech recognition and/or natural language understanding are performed by the hearing device. However, it also may be possible that these steps are at least partially performed by an external evaluation system, for example by a mobile device of the user.

[0049] In this case, the data necessary for the steps performed in the evaluation system may be sent to the evaluation system and the results of the processing of

the evaluation system may be sent back to the hearing device. This may be performed via a wireless connection, such as Bluetooth™ or WiFi.

[0050] It also may be that at least parts of the evaluation systems are a server, which is connected via Internet with the hearing device and/or the mobile device. The analysis may be performed in the cloud based on raw microphone signals, i.e. the sound signal of the microphone(s) of the hearing device and optionally the remote sound signal from an additional, external microphone. It also is possible that the analysis may be performed in the cloud based on pre-processed sound signals.

[0051] According to an embodiment, wherein at least one of the sound signal, the user voice signal, the directional sound signals, the word sequences and the semantic representations are sent to the evaluation system and/or at least one of the user voice signal, the directional sound signals, the word sequences, the semantic representations and the conversations are determined by the evaluation system.

[0052] According to an embodiment, the hearing device is a hearing aid, wherein the sound signal is processed for compensating a hearing loss of the user. The sound signal, which has been processed for amplifying the selected conversation may be additionally processed for compensating a hearing loss of the user.

[0053] Further aspects of the present disclosure relate to a computer program for directionally amplifying a sound signal of a hearing device, which, when being executed by a processor, is adapted to carry out the steps of the method as described in the above and in the following as well as to a computer-readable medium, in which such a computer program is stored.

[0054] For example, the computer program may be executed in a processor of the hearing device, which hearing device, for example, may be carried by the person behind the ear. The computer-readable medium may be a memory of this hearing device. The computer program also may be executed by processors of the hearing device and/or the evaluation system. In this case, the computer-readable medium may be a memory of the hearing device and/or the evaluation system.

[0055] In general, a computer-readable medium may be a floppy disk, a hard disk, an USB (Universal Serial Bus) storage device, a RAM (Random Access Memory), a ROM (Read Only Memory), an EPROM (Erasable Programmable Read Only Memory) or a FLASH memory. A computer-readable medium may also be a data communication network, e.g. the Internet, which allows downloading a program code. The computer-readable medium may be a non-transitory or transitory medium.

[0056] A further aspect of the present disclosure relates to a hearing system for directionally amplifying a sound signal of a hearing device comprising the hearing device. The hearing system may be adapted for performing the method as described in the above and in the following.

[0057] Besides the hearing device, the hearing system may comprise an evaluation system, which at least performed some of the steps of the method. The evaluation system may comprise a mobile device carried by the user and/or a server connected via Internet to the hearing device and/or the mobile device.

[0058] The hearing device may send the sound signal to the mobile device, which may perform automatic speech recognition and/or natural language understanding. The mobile device also may send the sound signal to the server, which may perform automatic speech recognition and/or natural language recognition.

[0059] The hearing system may be adapted for detecting, which speakers participate in the same conversation and/or which ones do not, and based on this may optimize the hearing performance of the user of the hearing device. The hearing system may analyze the conversation topology and the semantic content of the conversation(s). It may analyze the semantic content of the voice of the user. It may cluster and/or compare semantic content and may detect in which conversation the user is participating. In the end, the hearing system may control and/or steer a directivity of the hearing device, such that speakers participating not in the same conversation as the user are suppressed.

[0060] It has to be understood that features of the method as described in the above and in the following may be features of the computer program, the computer-readable medium and the hearing system as described in the above and in the following, and vice versa.

[0061] These and other aspects of the present disclosure will be apparent from and elucidated with reference to the embodiments described hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0062] Below, embodiments of the present invention are described in more detail with reference to the attached drawings.

Fig. 1 schematically shows a hearing system.

Fig. 2 schematically shows a hearing situation, in which the system of Fig. 1 is used.

Fig. 3 schematically shows a modular configuration of the hearing system of Fig. 1.

[0063] The reference symbols used in the drawings, and their meanings, are listed in summary form in the list of reference symbols. In principle, identical parts are provided with the same reference symbols in the figures.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0064] Fig. 1 schematically shows a hearing system 10, which comprises a hearing device 12, a mobile device 14 and optionally a server 16.

[0065] The hearing device 12 may be a binaural hearing device 12, which has two components 18 for each

ear of a user. Each of the components 18 may be seen as a hearing device of its own.

[0066] Each of the components 18, which may be carried behind the ear or in the ear, may comprise one or more microphones 20 and one or more loudspeakers 22. Furthermore, each or one of the components 18 may comprise a sensor 23, which is adapted for measuring head movements of the user, such as an acceleration sensor.

[0067] The mobile device 14, which may be a smartphone, may be in data communication with the hearing device 12, for example via a wireless communication channel such as Bluetooth™. The mobile device 14 may have a display 24, on which a visualization of a conversation topology may be shown (see Fig. 2).

[0068] The mobile device 14 may be in data communication with the server 16, which may be a cloud server provided in a cloud computing facility remote from the hearing device 12 and/or the mobile device 14.

[0069] Furthermore, an additional microphone 26, which is situated in the environment around the user of the hearing device 12, may be in in data communication with the server 16.

[0070] Communication between the mobile device 14 and/or the additional microphone 26 and the server 16 may be established via Internet, for example via Bluetooth™ and/or a mobile phone communication network.

[0071] The mobile device 14 and the server 16 may be seen as an evaluation system 28 that may perform some of the steps of the method as described with respect to Fig. 3 externally to the hearing device 12.

[0072] Fig. 2 shows a diagram that may be displayed by the mobile device 14. It shows the user 30, further persons and/or speakers 32 in the environment of the user 30 and conversations 34 in which these speakers 32 participate. All these information may have been determined by the hearing system 10.

[0073] Fig. 3 shows a modular configuration of the hearing system 10. The modules described in the following also may be seen as method steps of a method that may be performed by the hearing system 10. It has to be noted that the modules described below may be implemented in software and/or may be part of the hearing device 12, the mobile device 14 and/or the server 16.

[0074] As shown in Fig. 3, the microphones 20 of the hearing device produce a sound signal 36, which may comprise sound data of all the microphones 20. The sound signal 36 may be seen as a multi-component sound signal.

[0075] A sector beam former module 38 receives the sound signal 36 and extracts sector sound signals 40 from the sound signal 36. An environment of the user 30 may be divided into directional sectors 39 (see Fig. 2), such as quadrants. For each of these sectors 39, the sector beam former module 38 may generate a sector sound signal 40. For example, the sector beam former module 38 may comprise a beam former for each sector 39, which direction and angle width is adjusted to the respective sector 39. Each sector sound signal 40 may be extracted from the sound signal 36 by amplifying sound with a direction from the corresponding directional sector 39, for example with a beam former.

[0076] It has to be noted that a plurality of sector sound signals 40 may be generated, but that only one of the signals 40 is shown in Fig. 3. Also, for much of the signals and/or data mentioned in the following, which are associated to speakers 32, sound sources, conversations 34, etc., only one line may be shown in Fig. 3.

[0077] The sector sound signals 40 may be received in a speaker detection module 42, which identifies speakers 32 and/or sound sources in the respective sector 39. For example, each speaker may be detected as a separate sound source. In general, a number of sound sources and positions/directions of these sound sources may be determined by the speaker detection module 42.

[0078] The directions and/or positions of the sound sources may be input into a speaker beam former module 44. The speaker beam former module 44 may comprise a plurality of beam formers, each of which may be adjusted to a direction and/or position of one of the sound sources. Each beam former of the speaker beam former module 44 then extracts a speaker voice signal 46 from the sound signal 36 by amplifying sound with a direction from the corresponding sound source 32.

[0079] It furthermore may be possible that the module 44 receives a remote sound signal from the additional microphone 26. The speaker voice signals 46 then may be determined from the sound signal 36 of the one or more microphones 20 of the hearing device 12 and the remote sound signal.

[0080] The speaker voice signals 46 and optionally the information from the speaker detection module 42 may be received by a speaker identification module 48, which extracts speaker characteristics 50 from the respective speaker sound signal 46.

[0081] It may be possible that characteristics of specific speakers 32, for example speakers known by the user, such as his spouse, his children, etc., are stored in a database 49. The speaker identification module 48 may identify a speaker 32 with the aid of speaker characteristics stored in the database 49. Also, the characteristics 50 extracted by the module 48 may be enriched with characteristics stored in the database 49 associated with an identified speaker 32.

[0082] It also may be possible that speakers 32 are identified in the sound signal 36 directly with the speaker identification module 48 and that the speaker voice signals 46 are extracted from the sound signal 36 by amplifying sound with characteristics 50 of the speaker 32. Also, these characteristics 50 may be identified by the speaker identification module 48 as described above.

[0083] In general, the sector sound signals 40 and the speaker voice signals 46 may be seen as directional sound signals, which are then further processed by automatic speech recognition and automatic natural language understanding.

[0084] The sector sound signals 40 and/or the speaker voice signals 46 are received by an automatic speech recognition module 52. As indicated in Fig. 3, either the sector sound signals 40 or the speaker voice signals 46 may be input into the automatic speech recognition module 52. For example, the user 30 may select one of these options.

[0085] The automatic speech recognition module 52 determines a word sequence 54 for the respective directional sound signal 40, 46, which is then input into a natural language understanding module 56, which determines a semantic representation 58. A semantic representation 58 may contain semantic weights for a semantic content of the word sequence 54. For example, the semantic representation may be a vector of weights and each weight indicates the probability of a specific conversation subject, such as holidays, work, family, etc.

[0086] Both the automatic speech recognition module 52 and the natural language understanding module 56 may be based on machine learning algorithm, which have to be trained for identifying words in a data stream containing spoken language and/or identifying semantic content in a word sequence.

[0087] The automatic speech recognition module 52 and/or the natural language understanding module 56 may use speaker characteristics 50 of the speaker 32 associated with the speaker voice signal during processing their input.

[0088] A further automatic speech recognition module 52 and a further natural language understanding module 56 process input from a user voice extractor 60. In particular, the user voice extractor 60 extracts a user voice signal 62 from the sound signal 36. The user voice signal 62 is then translated by the further automatic speech recognition module 52 into a word sequence 54. The further natural language understanding module 56 determines a semantic representation 58 from this word sequence.

[0089] In the end, semantic representations 58 for the user 30 and for several sectors 39 or speakers 32 have been generated. This information and/or data is then used to identify conversations 34 and to process the sound signal 36 in such a way that the conversion 34, in which the user 30 participates, is amplified.

[0090] The semantic representations 58 are then input into a comparator 64, which generates distance information 66 (which may be a single value) for each pair of semantic representations 58. For example, the distance information 66 may be or may comprise a distance in the vector space of weights of the semantic representations 58.

[0091] From the distance information, the clustering module 68 identifies conversations 34, which, for example, are sets of sound sources (such as the user, the speakers, the sectors), which have a low distance according to the distance information 66. It also may be that the clustering module 68 directly clusters semantic representations 58 by their semantic weights into conversations 34. Each conversation 34 may be associated with a sector 39, a sound source and/or a speaker 32 and optionally the user 30. Each conversation also may be associated with a semantic representation 58, which may be an average of the semantic representations 58, which define the cluster, the conversation 34 is based on.

[0092] The identified conversations 34 are input into a map module 74, which generates a conversation topology map 76. For this, the conversations 34 may be associated with the positions and/or directions of the sectors 39, sound sources and/or speakers 32 and optionally the user 30, the conversation is associated with.

[0093] It also may be possible that head movements of the user 30 are detected, for example with a sensor 23. The map module 74 then may update the conversation topology map 74, such as directions and/or positions associated with the conversations 34, based on the head movements. The conversation topology map 74 may be updated, when the user is moving and/or turning within his/her environment.

[0094] In general, it may be that the conversation topology map 74 is actualized over time, for example, when conversation subjects change and/or speakers 32 enter or leave conversations. The conversations 34 generated by the clustering module 68 may be identified with already existing conversations 34 in the conversation topology map 74, which are then updated accordingly.

[0095] The conversation topology map 76 may be visualized by a visualization module 78, which may generate an image that is presented on the display 24 of the mobile device 14 of the user 30. Such a visualization may look like the diagram of Fig. 2.

[0096] It may be possible that the user 30 can select one of the conversations 34 that are displayed and that this conversation 34 is then amplified by the hearing device 12. It also may be that the map module 74 automatically selects a conversation 34 to be amplified by selecting a conversation 34, which has the highest concordance with the semantic representation 58 of the user voice signal 62 and/or which is associated with the user 30.

[0097] After the selection, selection information 80, such as a direction, a position, an angle, etc. is input into a control module 82 of the hearing device 12, which controls and/or steers the sound processing of the hearing device 12. The control module 82 determines control parameters 84, which are provided to a signal processor 86, which processes the sound signal 36 and generates an output signal 88, which may be output by the loudspeaker 22 of the hearing device 12.

[0098] In general, the signal processor 86 may be adjusted with the control parameters 84, such that the directional sound signals 40, 46 associated with one of the conversations 34 are amplified. For example, this may be performed by adjusting a beam former such that its direction and opening angle (width) are directed towards all sound sources, sectors and/or speakers associated with the conversation 34. Additionally, it may be that the sound signal 36 is processed by the signal processor 86

for compensating a hearing loss of a user 30 of the hearing device 12.

[0099] While the invention has been illustrated and described in detail in the drawings and foregoing description, such illustration and description are to be considered illustrative or exemplary and not restrictive; the invention is not limited to the disclosed embodiments. Other variations to the disclosed embodiments can be understood and effected by those skilled in the art and practising the claimed invention, from a study of the drawings, the disclosure, and the appended claims. In the claims, the word "comprising" does not exclude other elements or steps, and the indefinite article "a" or "an" does not exclude a plurality. A single processor or controller or other unit may fulfill the functions of several items recited in the claims. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage. Any reference signs in the claims should not be construed as limiting the scope.

LIST OF REFERENCE SYMBOLS

[0100]

10 hearing system
12 hearing device
14 mobile device
16 server
18 component of hearing device
20 microphone
22 loudspeaker
23 movement sensor
24 display
26 additional microphone
28 evaluation system
30 user
32 person, speaker
34 conversation
36 sound signal
38 sector beam former module
39 sector
40 sector sound signal
42 speaker detection module
44 speaker beam former module
46 speaker voice signal
48 speaker identification module
49 database
50 speaker characteristics
52 automatic speech recognition module
54 word sequence
56 natural language understanding module
58 semantic representation
60 user voice extractor
62 user voice signal
64 comparator
66 distance information
68 clustering module
74 map module
76 conversation topology map
78 visualization module
80 selection information
82 control module
84 control parameters
86 signal processor
88 output signal

Claims

1. A method for directionally amplifying a sound signal (36) of a hearing device (12), the method comprising:

receiving the sound signal (36) from a microphone (20) of the hearing device (12);
extracting directional sound signals (40, 46) from the sound signal (36);
determining a word sequence (54) from each directional sound signal (40, 46);
determining a semantic representation (58) from each word sequence (54);
identifying conversations (34) from the semantic representations (58), wherein each conversion (34) is associated with one or more directional sound signals (40, 46) and wherein each conversation (34) is identified by clustering semantic representations (58);
processing the sound signal (36), such that directional sound signals (40, 46) associated with one of the conversations (34) are amplified.

2. The method of claim 1,
wherein an environment of a user (30) of the hearing device (12) is divided into directional sectors (39) and each directional sound signal (40) is associated with one of the directional sectors (39);
wherein each directional sound signal (40) is extracted from the sound signal (36) by amplifying sound with a direction from the corresponding directional sector (39).

3. The method of claim 1 or 2,
wherein sound sources (32) are identified in the sound signal (36);
wherein each directional sound signal (46) is associated with one of the sound sources (32);
wherein each directional sound signal (46) is extracted from the sound signal (36) by amplifying sound with a direction from the corresponding sound source (32).

4. The method of one of the previous claims,
wherein speakers (32) are identified in the sound signal (36);
wherein each directional sound signal (46) is associated with one of the speakers (32);

wherein each directional sound signal (46) is extracted from the sound signal (36) by amplifying sound with characteristics of the speaker (32).

5. The method of one of the previous claims, wherein the sound signal (36) is processed by adjusting at least one of a direction and a width of a beam former of the hearing device (12).

6. The method of one of the previous claims, wherein a user voice signal (62) is extracted from the sound signal (36) and a word sequence (54) is determined from the user voice signal (62).

7. The method of claim 6, further comprising:

automatically selecting a conversation (34) to be amplified by selecting a conversation (34), which has the highest concordance with the semantic representation (58) of the user voice signal (62).

8. The method of claim 6 or 7, further comprising:

storing semantic representations (58) of a user (30) of the hearing aid (12) over time;
automatically selecting a conversation (34) to be amplified by selecting a conversation (34), which has the highest concordance with the semantic representation (58) of the user voice signal (62) and with stored semantic representations (58).

9. The method of one of the previous claims, further comprising:

presenting a visualization of the conversations (34) to the user (30);
receiving a selection of the user (30) for a conversation (34) to be amplified.

10. The method of one of the previous claims, further comprising:

detecting head movements of a user (30) of the hearing device (12);
updating directions associated with the conversations (34) based on the head movements.

11. The method of one of the previous claims, further comprising:

receiving a remote sound signal from an additional microphone (26) in the environment of a user (30) of the hearing device (12);
wherein at least the directional sound signals (40, 46) are determined from the sound signal (36) of the microphone (20) of the hearing device (12) and the remote sound signal.

12. The method of one of the previous claims, wherein at least one of the sound signal (36), the directional sound signals (40, 46), the word sequences (54) and the semantic representations (58) are sent to an evaluation system (28);
wherein at least one of the directional sound signals (40, 46), the word sequences (54), the semantic representations (58) and the conversations (34) are determined by the evaluation system (28).

13. A computer program for directionally amplifying a sound signal of a hearing device, which, when being executed by a processor, is adapted to carry out the steps of the method of one of the previous claims.

14. A computer-readable medium, in which a computer program according to claim 13 is stored.

15. A hearing system (10) for directionally amplifying a sound signal (36) of a hearing device (12) comprising the hearing device (12) and being adapted for performing the method of one of claims 1 to 12.

Fig. 1

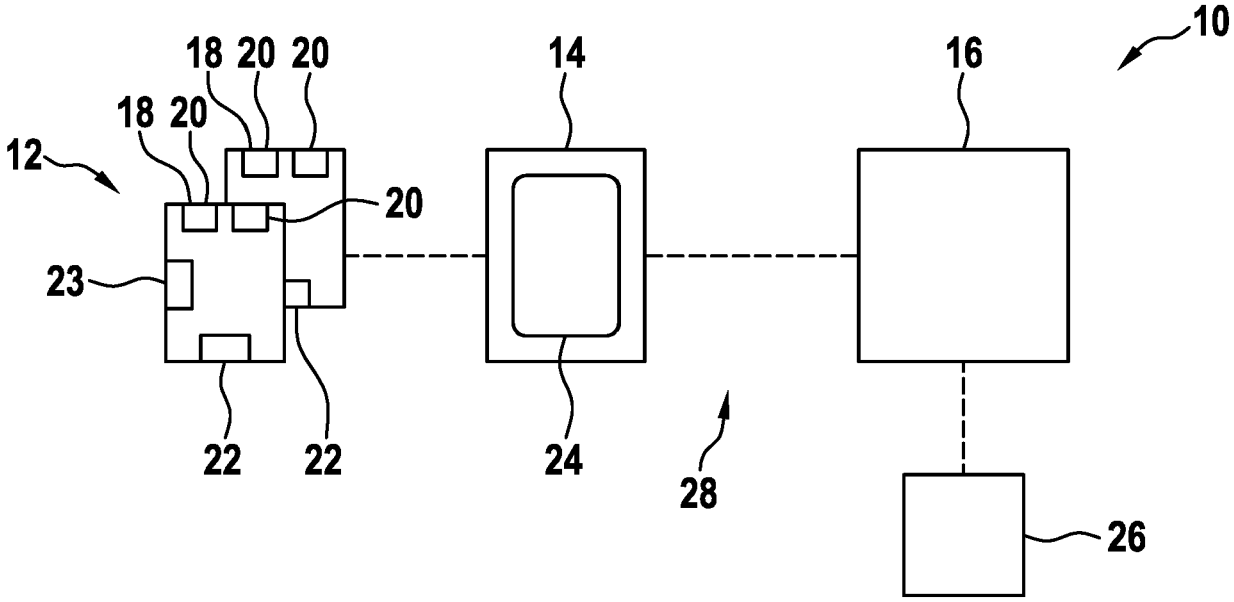
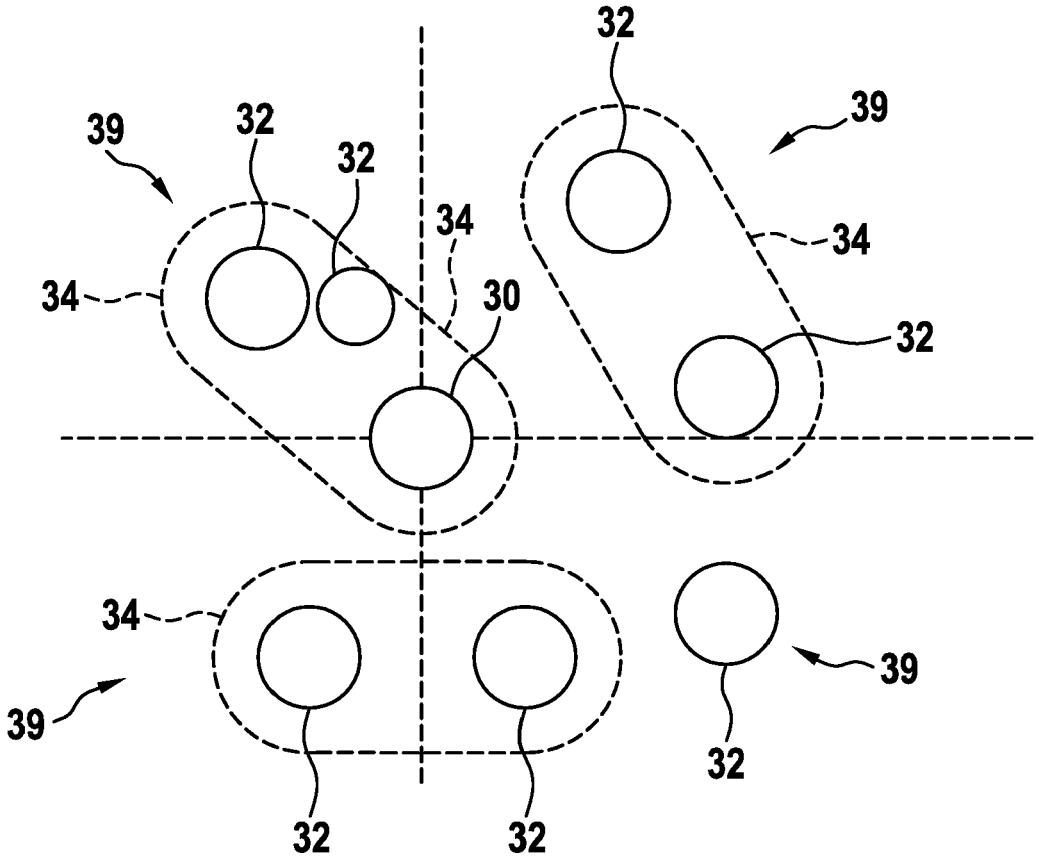

10
18
20
20
14
16
18
20
12
20
23
22
22
24
28
26

Fig. 2


32
39
39
32
32
34
34
34
30
32
34
39
39
32
32
32

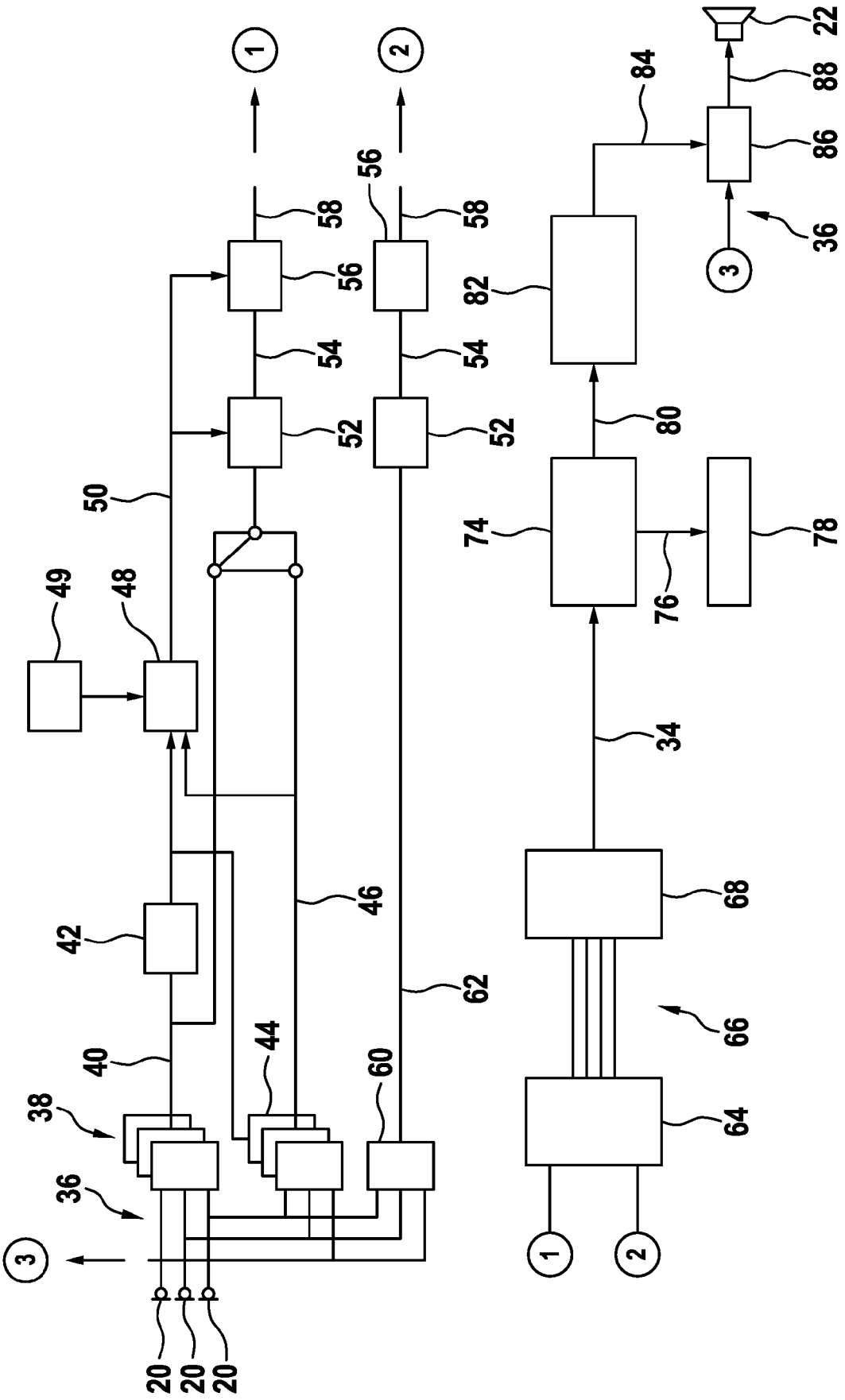

1
2
3
20
20
20
22
34
36
36
38
40
42
44
46
48
49
50
52
52
54
54
56
56
58
58
60
62
64
66
68
74
76
78
80
82
84
86
88

Fig. 3

Europäisches Patentamt
European Patent Office
Office européen des brevets

EUROPEAN SEARCH REPORT

Application Number
EP 18 19 3769

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	US 9 980 042 B1 (STAGES PCS LLC [US]; STAGES LLC [US]) 22 May 2018 (2018-05-22) * column 11, lines 16-43 * * column 14, lines 24-28 * -----	1-15	INV. H04R25/00 ADD. G10L21/0316
A	US 2018/033428 A1 (KIM LAE-HOON [US] ET AL) 1 February 2018 (2018-02-01) * paragraphs [0051] - [0059] * -----	1-15	G10L21/0208
A	EP 3 249 944 A1 (EM-TECH CO LTD [KR]) 29 November 2017 (2017-11-29) * paragraphs [0027] - [0030] * -----	1-15	
A	MOATTAR M H ET AL: "A review on speaker diarization systems and approaches", SPEECH COMMUNICATION, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 54, no. 10, 29 May 2012 (2012-05-29), pages 1065-1103, XP028449260, ISSN: 0167-6393, DOI: 10.1016/J.SPECOM.2012.05.002 [retrieved on 2012-06-06] * section "Acoustic beamforming"; page 1071 - page 1072 * -----	1-15	

TECHNICAL FIELDS SEARCHED (IPC)

H04R
G10L
H04M

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	11 February 2019	Rogala, Tomasz

CATEGORY OF CITED DOCUMENTS

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

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

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 18 19 3769

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.

11-02-2019

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 9980042 B1	22-05-2018	NONE	
US 2018033428 A1	01-02-2018	US 2018033428 A1	01-02-2018
		WO 2018022222 A1	01-02-2018
EP 3249944 A1	29-11-2017	CN 107438209 A	05-12-2017
		EP 3249944 A1	29-11-2017
		JP 2017211640 A	30-11-2017
		KR 101756674 B1	25-07-2017
		US 2017345408 A1	30-11-2017

EPO FORM P0459

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

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

- US 6157727 A **[0004]**