



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
12.02.2025 Bulletin 2025/07

(51) International Patent Classification (IPC):
H04R 1/10 (2006.01) **H04R 25/00** (2006.01)

(21) Application number: **23190449.1**

(52) Cooperative Patent Classification (CPC):
H04R 1/1041; H04R 1/1083; H04R 25/507;
H04R 25/43; H04R 2225/41; H04R 2225/43;
H04R 2225/61; H04R 2430/01

(22) Date of filing: **09.08.2023**

(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 ME MK MT NL
NO PL PT RO RS SE SI SK SM TR**
Designated Extension States:
BA
Designated Validation States:
KH MA MD TN

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

(72) Inventors:
• **Müller, Stephan**
8610 Uster (CH)
• **Cornelisse, Leonard**
Waterloo, Ontario, N2J 4Y9 (CA)

(54) **OPERATING A HEARING DEVICE FOR CLASSIFYING AN AUDIO SIGNAL**

(57) The disclosure relates to a method of operating a hearing device configured to be worn at an ear of a user, the method comprising; receiving an audio signal (311); classifying the audio signal (311) by attributing at least one class (321, 323, 324, 325) from a plurality of pre-determined classes to the audio signal (311), wherein different audio processing instructions (331, 333, 334, 335) are associated with different classes (321, 323, 324, 325); modifying the audio signal (311) by applying the audio processing instruction (331, 333, 334, 335) associated with the class (321, 323, 324, 325) attributed to the audio signal (331); and controlling an output transducer (117) included in the hearing device to generate a sound output according to the modified audio signal (311). The disclosure further relates to a hearing device configured

to perform the method.

To avoid undesired trade-offs between benefits and negative consequences for the user when applying different audio processing instructions (331, 333, 334, 335), the disclosure proposes that the audio processing instruction (331, 333, 334, 335) associated with at least one of said classes (321, 323, 324, 325) includes an inhibition instruction which, when executed, inhibits applying the audio processing instruction (331, 333, 334, 335) associated with at least another one of said classes (321, 323, 324, 325), the method further comprising executing the inhibition instruction when the audio processing instruction (331, 333, 334, 335) including the inhibition instruction is applied

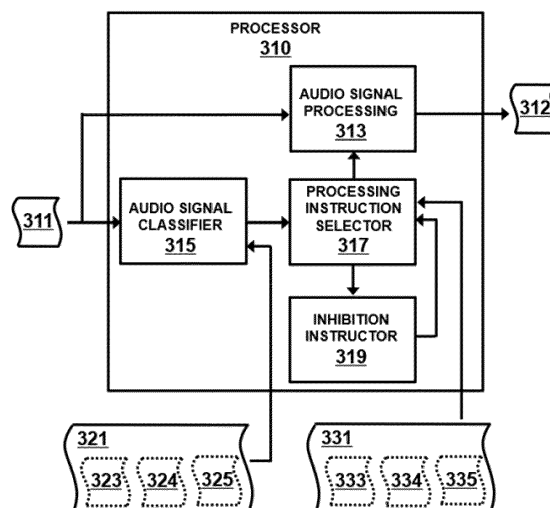


Fig. 4

Description

TECHNICAL FIELD

[0001] The disclosure relates to method of operating a hearing device configured to be worn at an ear of a user, according to the preamble of claim 1. The disclosure further relates to a computer-readable medium, according to the preamble of claim 14, and to a hearing device, according to the preamble of claim 15.

BACKGROUND

[0002] Hearing devices may be used to improve the hearing capability or communication capability of a user, for instance by compensating a hearing loss of a hearing-impaired user, in which case the hearing device is commonly referred to as a hearing instrument such as a hearing aid, or hearing prosthesis. A hearing device may also be used to output sound based on an audio signal which may be communicated by a wire or wirelessly to the hearing device. A hearing device may also be used to reproduce a sound in a user's ear canal detected by an input transducer such as a microphone or a microphone array. The reproduced sound may be amplified to account for a hearing loss, such as in a hearing instrument, or may be output without accounting for a hearing loss, for instance to provide for a faithful reproduction of detected ambient sound and/or to add audio features of an augmented reality in the reproduced ambient sound, such as in a hearable. A hearing device may also provide for a situational enhancement of an acoustic scene, e.g. beamforming and/or active noise cancelling (ANC), with or without amplification of the reproduced sound. A hearing device may also be implemented as a hearing protection device, such as an earplug, configured to protect the user's hearing. Different types of hearing devices configured to be worn at an ear include earbuds, earphones, hearables, and hearing instruments such as receiver-in-the-canal (RIC) hearing aids, behind-the-ear (BTE) hearing aids, in-the-ear (ITE) hearing aids, invisible-in-the-canal (IIC) hearing aids, completely-in-the-canal (CIC) hearing aids, cochlear implant systems configured to provide electrical stimulation representative of audio content to a user, a bimodal hearing system configured to provide both amplification and electrical stimulation representative of audio content to a user, or any other suitable hearing prostheses. A hearing system comprising two hearing devices configured to be worn at different ears of the user is sometimes also referred to as a binaural hearing device. A hearing system may also comprise a hearing device, e.g., a single monaural hearing device or a binaural hearing device, and a user device, e.g., a smartphone and/or a smartwatch, communicatively coupled to the hearing device.

[0003] Hearing devices are often employed in conjunction with communication devices, such as smartphones or tablets, for instance when listening to sound data

processed by the communication device and/or during a phone conversation operated by the communication device. More recently, communication devices have been integrated with hearing devices such that the hearing devices at least partially comprise the functionality of those communication devices. A hearing system may comprise, for instance, a hearing device and a communication device.

[0004] In recent times, some hearing devices are also increasingly equipped with different sensor types. Traditionally, those sensors often include an input transducer to detect a sound, e.g., a sound detector such as a microphone or a microphone array. An amplified and/or signal processed version of the detected sound may then be outputted to the user by an output transducer, e.g., a receiver, loudspeaker, or electrodes to provide electrical stimulation representative of the outputted signal. In an effort to provide the user with even more information about himself and/or the ambient environment, various other sensor types are progressively implemented, in particular sensors which are not directly related to the sound reproduction and/or amplification function of the hearing device. Those sensors include inertial sensors, such as accelerometers, allowing to monitor the user's movements. Physiological sensors, such as optical sensors and bioelectric sensors, are mostly employed for monitoring the user's health.

[0005] Modern hearing devices provide several features that aim to facilitate speech intelligibility, improve sound quality, reduce noise level, etc. Many of such sound cleaning features are designed to benefit the hearing device user's hearing performance in very specific situations. In order to activate the functionalities only in the situations where benefit can be expected, an automatic steering system is often implemented which activates sound cleaning features depending on a combination of, e.g., an acoustic environment classification, a physical activity classification, a directional classification, etc.

[0006] To provide for the acoustic environment classification, hearing devices have been equipped with a sound classifier to classify an ambient sound. An input transducer can provide an audio signal representative of the ambient sound. The sound classifier can classify the audio signal allowing to identify different listening situations by determining a characteristic from the audio signal and assigning the audio signal to at least one relevant class from a plurality of predetermined classes depending on the characteristic. Usually, the sound classification does not directly modify a sound output of the hearing device. Instead, different audio processing instructions are stored in a memory of the hearing device specifying different audio processing parameters for a processing of the audio signal, wherein the different classes are each associated with one of the different audio processing instructions. After assigning the audio signal to one or more classes, the one or more associated audio processing instructions are executed. The audio processing

parameters specified by the audio processing instructions can then provide a processing of the audio signal customized for the particular listening situation corresponding to the at least one class identified by the classifier. The different listening situations may comprise, for instance, different classes of listening conditions and/or different classes of sounds. For example, the different classes may comprise speech and/or nonspeech and/or music and/or traffic noise and/or other ambient noise.

[0007] The classification may be based on a statistical evaluation of the audio signal, as disclosed in EP 3 036 915 B1. More recently, machine learning (ML) algorithms have been employed to classify the ambient sound. The classifier can be implemented by an artificial intelligence (AI) chip which may be configured to classify the audio signal by at least one deep neural network (DNN). The classifier may comprise a sound source separator configured to separate sound generated by different sound sources, for instance a conversation partner, passengers passing by the user, vehicles moving in the vicinity of the user such as cars, airborne traffic such as a helicopter, a sound scene in a restaurant, a sound scene including road traffic, a sound scene during public transport, a sound scene in a home environment, and/or the like. Examples of such a sound source separator are disclosed in international patent application Nos. PCT/EP 2020/051 734 and PCT/EP 2020/051 735, and in German patent application No. DE 2019 206 743.3.

[0008] Some sound cleaning features, however, also introduce side effects that might even counteract the intended benefits of the audio processing in a current situation. Similar to the intended benefits, the occurrence of side effects is also depending on the situation that the user is currently in. Automatic steering algorithms aim to activate features if one or more activation criteria are fulfilled by the classifier(s). However, that only exposes one side of the model, since at the same time, as beneficial use case criteria are detected, criteria that would indicate occurrence of negative side effects of a feature may also be present. These negative criteria are not generally considered or exposed within the steering system, or they are built into assumptions when the system is optimized.

[0009] An approach that simply goes from classification to steering of adaptive features includes assumptions about the trade-offs, e.g., based on the upfront classification. To illustrate, if the output of the classification is Speech in Noise or Conversation in a Crowd, then the applied solution is determined based on the settings for the predefined situation or class. Although the settings of individual features may be modified via an individual fitting of the hearing device according to the specific needs of the user, e.g., for specific situations and/or classes, the important dimension of trade-offs between the positive and negative consequences is not apparent to the user. A challenge with situational classification is that the situations are defined generically. There can still be considerable variability in terms of other relevant

perceptual dimensions such as, e.g., an overall level and/or a target signal to noise ratio (SNR). One approach could be to increase the number of situations or available classes which, however, would quickly lead to a very large number of classes.

[0010] Another approach would be to mix different features associated with different classes. To this end, a mixed mode classifier has been proposed in EP 1 858 292 B1. The mixed mode classifier can attribute one, two or more classes to the audio signal, wherein the different features in the form of audio processing instructions associated with the different classes can be mixed in dependence of class similarity factors. The class similarity factors are indicative of a similarity of the current acoustic environment with a respective predetermined acoustic environment associated with the different classes. The mixing of the different audio processing instructions may imply, e.g., a linear combination of base parameter sets representing the audio processing instructions associated with the different classes, or other non-linear ways of mixing the audio processing instructions. The different audio processing instructions may be provided as sub-functions, which can be included into a transfer function used by the signal processing circuit according to the desired mixing of the audio processing instructions. For example, audio processing instructions, e.g., in the form of the base parameter sets, related to a beamformer and/or a gain model (i.e., an amplification characteristic) may be mixed depending on whether or to which degree the audio signal is attributed, e.g., by the class similarity factors, to one or more of the classes music and/or speech in noise and/or speech.

[0011] EP 2 201 793 B1 discloses a classifier configured for an automatic adaption of the audio processing instructions associated with the different classes depending on adjustments performed by the user. Adjustment data indicative of the user adjustments can be logged, e.g., stored in a storage unit, and evaluated to learn correction data for correcting the audio processing instructions. In a mixed mode classifier, for a current sound environment and depending on the adjustment data, an offset can be learned for the mixed base parameter sets representing the audio processing instructions associated with the different classes. For the purpose of learning, correction data may be separately provided for different classes.

[0012] Such a mixed mode classifying, however, makes it challenging to modify hearing device settings in a uniquely mixed situation. Another approach would be to use a multi-label approach in which the presence or absence of certain characteristics leads to specific settings. However, such a multi-label approach would have the disadvantage that there would be no clear relationship between the specific situations and the settings actuated by the instrument. What is missing is an intermediate step, which can help translate from unique situations to a steering of the adaptive features without the need for a large number of predefined situations or

classes.

SUMMARY

[0013] It is an object of the present disclosure to avoid at least one of the above mentioned disadvantages and to account, when the audio processing instructions associated with one or more classes representative for a current situation are applied, for an additional dimension of the possible trade-offs between benefits and negative consequences for the user. It is another object to provide for an audio processing in different situations which mimics the way a human auditory system operates. It is a further object to also provide for an enhanced safety of the user when modifying the audio signal depending on a current situational classification. It is a further object to provide a hearing device which is configured to operate in such a manner.

[0014] At least one of these objects can be achieved by a method of operating a hearing device configured to be worn at an ear of a user comprising the features of claim 1 and/or a computer-readable medium comprising the features of claim 14 and/or a hearing device comprising the features of claim 15. Advantageous embodiments of the invention are defined by the dependent claims and the following description.

[0015] Accordingly, the present disclosure proposes a method of operating a hearing device configured to be worn at an ear of a user, the method comprising

- receiving an audio signal;
- classifying the audio signal by attributing at least one class from a plurality of predetermined classes to the audio signal, wherein different audio processing instructions are associated with different classes;
- modifying the audio signal by applying the audio processing instruction associated with the class attributed to the audio signal; and
- controlling an output transducer included in the hearing device to generate a sound output according to the modified audio signal, wherein the audio processing instruction associated with at least one of said classes includes an inhibition instruction which, when executed, inhibits applying the audio processing instruction associated with at least another one of said classes, the method further comprising
- executing the inhibition instruction when the audio processing instruction including the inhibition instruction is applied.

[0016] In this way, an excitation path of an audio processing instruction to be applied in a current situation, e.g., in accordance with a class attributed to the audio signal, can be inextricably linked to an inhibition path of

another audio processing instruction which may also be applicable in the current situation e.g., in accordance with another class attributed to the audio signal. In particular, a negative impact of applying the audio processing instruction associated with the other class on the current audio processing can thus be effectively inhibited, and an intended benefit of the currently applied audio processing instruction, in accordance with the excitation path, can be fully exploited. Playing off the individual benefits of different audio processing instruction against one another can thus be effectively avoided, and according trade-offs for the user's sound perception can be circumvented. Implementing the at least one excitation and inhibition path for the steering of the audio processing in such a manner can allow to mimic the human audio perception as triggered by neuropsychological mechanisms of the human brain in which an activity of one brain region may be excited, i.e. activated, at the expense of another brain region which may be inhibited, i.e. deactivated. Moreover, since in modern hearing aids multiple classifiers are often digesting a multitude of sensory inputs, these classifiers may accordingly be used to generate inhibitory and excitatory information.

[0017] Independently, the present disclosure also proposes a non-transitory computer-readable medium storing instructions that, when executed by a processor, cause a hearing device to perform operations of the method.

[0018] Independently, the present disclosure also proposes a hearing device configured to be worn an ear of a user, the hearing device comprising

- an input transducer configured to provide an audio signal indicative of a sound detected in the environment of the user;
- a processor configured to
 - classify the audio signal by attributing at least one class from a plurality of predetermined classes to the audio signal, wherein different audio processing instructions are associated with different classes; and
 - modify the audio signal by applying the audio processing instruction associated with the class attributed to the audio signal; and
- an output transducer configured to generate a sound output according to the modified audio signal, wherein the audio processing instruction associated with at least one of said classes includes an inhibition instruction which, when executed, inhibits applying the audio processing instruction associated with at least another one of said classes, wherein the processor is further configured to
- execute the inhibition instruction when the audio

processing instruction including the inhibition instruction is applied.

[0019] Subsequently, additional features of some implementations of the method of operating a hearing device and/or the computer-readable medium and/or the hearing device are described. Each of those features can be provided solely or in combination with at least another feature. The features can be correspondingly provided in some implementations of the method and/or the hearing device.

[0020] In some implementations, the audio processing instructions comprise

- an audio processing instruction associated with a class representative of a speech in front of the user and/or noise from the side or back of the user contained in the audio signal, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction associated with at least another class representative of noise in front of the user and/or speech from the side or back of the user contained in the audio signal; and/or
- an audio processing instruction associated with a class representative of a static noise contained in the audio signal, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction associated with at least another class representative of a modulated noise in the audio signal and/or an inhibition instruction which inhibits applying another audio processing instruction providing for a speech enhancement; and/or
- an audio processing instruction associated with a class representative of a music contained in the audio signal, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction associated with at least another class representative of a speech, e.g., speech in a complex scenario, in the audio signal; and/or
- an audio processing instruction associated with a class representative of a speech and/or noise contained in the audio signal, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction associated with at least another class representative of a traffic noise contained in the audio signal; and/or
- an audio processing instruction associated with a class representative of a speech present in a soft sound environment contained in the audio signal, wherein the audio processing instruction includes an

inhibition instruction which inhibits applying another audio processing instruction associated with at least another class representative of a modulated noise contained in the audio signal.

[0021] In some implementations, the audio processing instructions comprise

- an audio processing instruction providing for noise cancelling, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction providing for beamforming; and/or
- an audio processing instruction providing for speech enhancement, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction providing for noise cancelling; and/or
- an audio processing instruction providing for speech enhancement, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction providing for music enhancement; and/or
- an audio processing instruction providing for noise cancelling, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction providing for speech enhancement.

[0022] In some implementations, the audio processing instructions comprise a first audio processing instruction associated with a first class and a second audio processing instruction associated with a second class, wherein the first audio processing instruction includes a first inhibition instruction which, when executed, inhibits applying the second audio processing instruction and/or the second audio processing instruction includes a second inhibition instruction which, when executed, inhibits applying the first audio processing instruction, the method further comprising

- determining a priority measure indicative of whether the first audio processing instruction or the second audio processing instruction has a higher priority to be applied; and
- applying, depending on the priority measure, one of the first audio processing instruction and second audio processing instruction.

[0023] In some implementations, depending on the priority measure, the first or second inhibition instruction included in the first or second audio processing instruction may be executed. Thus, when the first or second inhibition instruction is executed, the other of the first

audio processing instruction and the second audio processing instruction is inhibited. E.g., the one of the first audio processing instruction and second audio processing instruction for which a higher priority has been determined may be applied and/or the one of the first audio processing instruction and second audio processing instruction for which a lower priority has been determined may be inhibited.

[0024] In some implementations, the priority measure is determined based on the audio signal. In some implementations, the priority measure is indicative of whether the first class or the second class is dominantly represented in the audio signal. In some implementations, the determining the priority measure based on the audio signal comprises at least one of

- determining a signal to noise ratio (SNR) in the audio signal, wherein the priority measure is indicative of the SNR;
- determining a presence of a content in the audio signal, e.g., a speech content, a music content, etc., wherein the priority measure is indicative of the presence of the content;
- determining a presence of a sound emitted by at least one acoustic object in the environment of the user in the audio signal, wherein the priority measure is indicative of the presence of the sound emitted by the acoustic object;
- evaluating the audio signal in a psychoacoustic model, wherein the priority measure is indicative of a deviation of the audio signal from the psychoacoustic model;
- evaluating the audio signal with regard to spatial cues indicative of a difference of a sound detected on a different position at the user, wherein the priority measure is indicative of the spatial cues; and
- determining an amount of a temporal dispersion of an impulse in the audio signal, wherein the priority measure is indicative of the temporal dispersion.

[0025] In some implementations, the priority measure is indicative of whether the audio signal is representative of an acoustic situation in which applying the first audio processing instruction or the second audio processing instruction poses a larger threat to the user, wherein the other of the first audio processing instruction or the second audio processing instruction posing a smaller threat to the user is determined to have the higher priority.

[0026] In some implementations, the method further comprises

- receiving, from a sensor, sensor data indicative of a displacement of the hearing device and/or indicative

of a property of the user and/or an ambient environment of the user, wherein the priority measure is determined based on the sensor data.

[0027] In some implementations, the sensor comprises

a displacement sensor configured to provide at least part of the sensor data as displacement data indicative of a displacement of the hearing device; and/or

a location sensor configured to provide at least part of the sensor data as location data indicative of a current location of the user; and/or

a physiological sensor configured to provide at least part of the sensor data as physiological data indicative of a physiological property of the user; and/or

an environmental sensor configured to provide at least part of the sensor data as environmental data indicative of a property of the environment of the user.

[0028] In some implementations, the priority measure is indicative of whether the sensor data is representative of a situation in which applying the first audio processing instruction or the second audio processing instruction poses a larger threat to the user, wherein the other of the first audio processing instruction or the second audio processing instruction posing a smaller threat to the user is determined to have the higher priority.

[0029] In some implementations, the method further comprises

- monitoring the priority measure when the one of the first audio processing instruction and second audio processing instruction is applied; and, when the priority measure is indicative of a higher priority of the other of the first audio processing instruction and second audio processing instruction,
- applying the other of the first audio processing instruction and second audio processing instruction. In particular, the first or second inhibition instruction included in the other of the first or second audio processing instruction may then be executed.

[0030] In some implementations, the audio processing instructions provide for at least one of an enhancement of a speech content of a single talker in the audio signal; an enhancement of a speech content of a plurality of talkers in the audio signal; a reproduction of sound emitted by an acoustic object in the environment of the user encoded in the audio signal; a reproduction of sound emitted by a plurality of acoustic objects in the environment of the user encoded in the audio signal; a reduction and/or canceling of noise and/or reverberations in the audio signal; a preservation of acoustic cues contained in the audio

signal; a suppression of noise in the audio signal; an improvement of a signal to noise ratio in the audio signal; a spatial resolution of sound encoded in the audio signal depending on a direction of arrival (DOA) of the sound and/or depending on a location of at least one acoustic object emitting the sound in the environment of the user; a directivity of an audio content in the audio signal, which may be provided by a beamforming, or a preservation of an omnidirectional audio content in the audio signal; an amplification of sound encoded in the audio signal adapted to an individual hearing loss of the user; and an enhancement of music content in the audio signal.

[0031] In some implementations, the audio processing algorithms comprise at least one of a gain model (GM); a noise cancelling (NC) algorithm; a wind noise cancelling (WNC) algorithm; a reverberation cancelling (RevC) algorithm; a feedback cancelling (FC) algorithm; a speech enhancement (SE) algorithm; an impulse noise cancelling (INC) algorithm; an acoustic object separation (AOS) algorithm; a binaural synchronization (BS) algorithm; and a beamforming (BF) algorithm.

[0032] In some implementations, the audio signal is indicative of a sound in the ambient environment of the user. In some implementations, the audio signal is received from an input transducer, e.g., a microphone or a microphone array, included in the hearing device. In some implementations, the audio signal is received by an audio signal receiver included in the hearing device, e.g., via radio frequency (RF) communication. In some implementations, the audio signal is received from a remote microphone, e.g., a table microphone and/or a clip-on microphone.

BRIEF DESCRIPTION OF THE DRAWINGS

[0033] Reference will now be made in detail to embodiments, examples of which are illustrated in the accompanying drawings. The drawings illustrate various embodiments and are a part of the specification. The illustrated embodiments are merely examples and do not limit the scope of the disclosure. Throughout the drawings, identical or similar reference numbers designate identical or similar elements. In the drawings:

- Fig. 1 schematically illustrates an exemplary hearing device;
- Fig. 2 schematically illustrates an exemplary sensor unit comprising one or more sensors which may be implemented in the hearing device illustrated in Fig. 1;
- Fig. 3 schematically illustrates an embodiment of the hearing device illustrated in Fig. 1 as a RIC hearing aid;
- Fig. 4 schematically illustrates an exemplary algorithm of processing an audio signal ac-

cording to principles described herein; and

Figs. 5, 6 schematically illustrate some exemplary methods of processing an audio signal according to principles described herein.

DETAILED DESCRIPTION OF THE DRAWINGS

[0034] FIG. 1 illustrates an exemplary hearing device 110 configured to be worn at an ear of a user. Hearing device 110 may be implemented by any type of hearing device configured to enable or enhance hearing or a listening experience of a user wearing hearing device 110. For example, hearing device 110 may be implemented by a hearing aid configured to provide an amplified version of audio content to a user, a sound processor included in a cochlear implant system configured to provide electrical stimulation representative of audio content to a user, a sound processor included in a bimodal hearing system configured to provide both amplification and electrical stimulation representative of audio content to a user, or any other suitable hearing prosthesis, or an earbud or an earphone or a hearable.

[0035] Different types of hearing device 110 can also be distinguished by the position at which they are worn at the ear. Some hearing devices, such as behind-the-ear (BTE) hearing aids and receiver-in-the-canal (RIC) hearing aids, typically comprise an earpiece configured to be at least partially inserted into an ear canal of the ear, and an additional housing configured to be worn at a wearing position outside the ear canal, in particular behind the ear of the user. Some other hearing devices, as for instance earbuds, earphones, hearables, in-the-ear (ITE) hearing aids, invisible-in-the-canal (IIC) hearing aids, and completely-in-the-canal (CIC) hearing aids, commonly comprise such an earpiece to be worn at least partially inside the ear canal without an additional housing for wearing at the different ear position.

[0036] As shown, hearing device 110 includes a processor 112 communicatively coupled to a memory 113, an input transducer 115, and an output transducer 117. Hearing device 110 may include additional or alternative components as may serve a particular implementation. Input transducer 115 may be implemented by any suitable device configured to detect sound in the environment of the user and to provide an input audio signal indicative of the detected sound, e.g., a microphone or a microphone array. Output transducer 117 may be implemented by any suitable audio transducer configured to output an output audio signal to the user, for instance a receiver of a hearing aid, an output electrode of a cochlear implant system, or a loudspeaker of an earbud.

[0037] Processor 112 is configured to receive, from input transducer 115, an input audio signal indicative of a sound detected in the environment of the user; to classify the input audio signal by attributing at least one class from a plurality of predetermined classes to the input audio signal, wherein different audio processing

instructions are associated with different classes; to modify the input audio signal by applying the audio processing instruction associated with the class attributed to the audio signal; and to control output transducer 117 to generate sound output according to the modified audio signal. The audio processing instruction associated with at least one of said classes includes an inhibition instruction which, when executed, inhibits applying the audio processing instruction associated with at least another one of said classes, wherein processor 112 is further configured to execute the inhibition instruction when the audio processing instruction including the inhibition instruction is applied. These and other operations, which may be performed by processor 112, are described in more detail in the description that follows.

[0038] Memory 113 may be implemented by any suitable type of storage medium and is configured to maintain, e.g. store, data controlled by processor 112, in particular data generated, accessed, modified and/or otherwise used by processor 112. For example, memory 113 may be configured to store instructions used by processor 112 to process the input audio signal received from input transducer 115, e.g., audio processing instructions in the form of one or more audio processing programs. The audio processing programs may comprise different audio processing instructions of modifying the input audio signal received from input transducer 115. For instance, the audio processing instructions may include algorithms providing a gain model, noise cleaning, noise cancelling, wind noise cancelling, reverberation cancelling, narrowband coupling, beamforming, in particular static and/or adaptive beamforming, and/or the like.

[0039] As another example, memory 113 may be configured to store instructions used by processor 112 to classify the input audio signal received from input transducer 115 by attributing at least one class from a plurality of predetermined sound classes to the input audio signal. Exemplary classes may include, but are not limited to, low ambient noise, high ambient noise, traffic noise, music, machine noise, babble noise, public area noise, background noise, speech, nonspeech, speech in quiet, speech in babble, speech in noise, speech from the user, speech from a significant other, background speech, speech from multiple sources, quiet indoor, quiet outdoor, speech in a car, speech in traffic, speech in a reverberating environment, speech in wind noise, speech in a lounge, car noise, applause, music, e.g. classical music, and/or the like. In some instances, the different audio processing instructions can be associated with different classes. Further, one or more audio processing instructions associated with a respective class may include an inhibition instruction which, when executed, inhibits applying the audio processing instruction associated with at least another class. As another example, memory 113 may be configured to store instructions used by processor 112 to determine a priority measure indicative of whether which of different processing instructions attributed to a respective class has a higher priority to be

applied.

[0040] Memory 113 may comprise a non-volatile memory from which the maintained data may be retrieved even after having been power cycled, for instance a flash memory and/or a read only memory (ROM) chip such as an electrically erasable programmable ROM (EEPROM). A non-transitory computer-readable medium may thus be implemented by memory 113. Memory 113 may further comprise a volatile memory, for instance a static or dynamic random access memory (RAM).

[0041] As illustrated, hearing device 110 may further comprise a communication port 119. Communication port 119 may be implemented by any suitable data transmitter and/or data receiver and/or data transducer configured to exchange data with another device. For instance, the other device may be another hearing device configured to be worn at the other ear of the user than hearing device 110 and/or a communication device such as a smartphone, smartwatch, tablet and/or the like. Communication port 119 may be configured for wired and/or wireless data communication. For instance, data may be communicated in accordance with a Bluetooth™ protocol and/or by any other type of radio frequency (RF) communication.

[0042] As illustrated, hearing device 110 may also comprise at least one further sensor 125 communicatively coupled to processor 112 in addition to input transducer 115. A sensor unit 120 may comprise input transducer 115 and the at least one further sensor 125. Some examples of a sensor which may be implemented in sensor unit 120 in place of sensor 125 are illustrated in Fig. 2.

[0043] As illustrated in FIG. 2, sensor unit 120 may include at least one environmental sensor configured to provide environmental data indicative of a property of the environment of the user in addition to input transducer 115, for example a barometric sensor 131 and/or an ambient temperature sensor 132. Sensor unit 120 may include at least one physiological sensor configured to provide physiological data indicative of a physiological property of the user, for example an optical sensor 133 and/or a bioelectric sensor 134 and/or a body temperature sensor 135. Optical sensor 133 may be configured to emit the light at a wavelength absorbable by an analyte contained in blood such that the physiological sensor data comprises information about the blood flowing through tissue at the ear. E.g., optical sensor 133 can be configured as a photoplethysmography (PPG) sensor such that the physiological sensor data comprises PPG data, e.g. a PPG waveform. Bioelectric sensor 134 may be implemented as a skin impedance sensor and/or an electrocardiogram (ECG) sensor and/or an electroencephalogram (EEG) sensor and/or an electrooculography (EOG) sensor.

[0044] Sensor unit 120 may include a displacement sensor 136 configured to provide displacement data indicative of a displacement of hearing device 110, which may then also indicate a movement of the user, for

example an accelerometer and/or a gyroscope and/or a magnetometer. Sensor unit 120 may include a user interface 137 configured to provide interaction data indicative of an interaction of the user with hearing device 110, e.g., a touch sensor and/or a push button. Sensor unit 120 may include at least one location sensor 138 configured to provide location data indicative of a current location of the user, for instance a GPS sensor. Sensor unit 120 may include at least one clock 139 configured to provide time data indicative of a current time. Context data may be defined as data indicative of a local and/or temporal context of the data provided by other sensors 115, 131 - 137. Context data may comprise the location data and/or the time data provided by location sensor 138 and/or clock 139. Context data may also be received from an external device via communication port 119, e.g., from a communication device. E.g., one or more of sensors 115, 131 - 137 may then be included in the communication device. Sensor unit 120 may include further sensors providing sensor data indicative of a property of the user and/or the environment and/or the context.

[0045] FIG. 3 illustrates an exemplary implementation of hearing device 110 as a RIC hearing aid 210. RIC hearing aid 210 comprises a BTE part 220 configured to be worn at an ear at a wearing position behind the ear, and an ITE part 240 configured to be worn at the ear at a wearing position at least partially inside an ear canal of the ear. BTE part 220 comprises a BTE housing 221 configured to be worn behind the ear. BTE housing 221 accommodates processor 112 communicatively coupled to input transducer 115 and may also include further sensor 125, which may include any of sensors 115, 131 - 139. BTE part 220 further includes a battery 227 as a power source. ITE part 240 is an earpiece comprising an ITE housing 241 at least partially insertable in the ear canal. ITE housing 241 accommodates output transducer 117 and may also include another sensor 241, which may include any of sensors 115, 131 - 139. Sensor unit 120 of exemplary RIC hearing aid 210 thus comprises input transducer 115 and sensors 125, 245. BTE part 220 and ITE part 240 are interconnected by a cable 251. Processor 112 is communicatively coupled to output transducer 117 and sensor 245 of ITE part 240 via cable 251 and cable connectors 252, 253 provided at BTE housing 221 and ITE housing 241.

[0046] FIG. 4 illustrates a functional block diagram of an exemplary audio signal processing algorithm that may be executed by a processor 310. For instance, processor 310 may comprise processor 112 of hearing device 110 and/or another processor communicatively coupled to processor 112. As shown, the algorithm is configured to be applied to an input audio signal 311 indicative of a sound detected in the environment of the user, which may be provided by input transducer 115. After a processing of audio signal 311, the algorithm provides a modified audio signal based on which an output audio signal 312 can be outputted by output transducer 117.

[0047] The algorithm comprises an audio signal pro-

cessing module 313, an audio signal classification module 315, a processing instruction selection module 317, and an inhibition instruction module 319. Input audio signal 311 is received by audio signal classification module 315. Audio signal classification module 315 is configured to classify the audio signal 311 by attributing at least one class from a plurality of predetermined classes to the audio signal 311. To this end, audio signal classification module 315 may comprise an audio signal analyzer module configured to analyze audio signal 311 to determine a characteristic of audio signal 311. For instance, the audio signal analyzer may be configured to provide a feature vector from audio signal 311 and/or to identify at least one signal feature in audio signal 311. Exemplary characteristics and/or features include, but are not limited to, a mean-squared signal power, a standard deviation of a signal envelope, a mel-frequency cepstrum (MFC), a mel-frequency cepstrum coefficient (MFCC), a delta mel-frequency cepstrum coefficient (delta MFCC), a spectral centroid such as a power spectrum centroid, a standard deviation of the centroid, a spectral entropy such as a power spectrum entropy, a zero crossing rate (ZCR), a standard deviation of the ZCR, a broadband envelope correlation lag and/or peak, and a four-band envelope correlation lag and/or peak. For example, the audio signal analyzer may determine the characteristic from audio signal 311 using one or more algorithms that identify and/or use zero crossing rates, amplitude histograms, auto correlation functions, spectral analysis, amplitude modulation spectrums, spectral centroids, slopes, roll-offs, auto correlation functions, and/or the like. In some instances, the characteristic determined from audio signal 311 is characteristic of an ambient noise in an environment of the user, for instance a noise level, and/or a speech, for instance a speech level. The audio signal analyzer may be configured to divide audio signal 311 into a number of segments and to determine the characteristic from a particular segment, for instance by extracting at least one signal feature from the segment. The extracted feature may be processed to assign the audio signal to the corresponding class.

[0048] Audio signal classification module 315 can attribute, e.g., depending on the characteristics and/or features determined from audio signal 311 by the audio signal analyzer, at least one sound class from a plurality of predetermined classes to audio signal 311. E.g., the characteristics and/or signal features may be processed to assign audio signal 311 to one or more corresponding classes. The classes may represent a specific content in the audio signal. Exemplary classes include, but are not limited to, low ambient noise, high ambient noise, traffic noise, music, machine noise, babble noise, public area noise, background noise, speech, nonspeech, speech in quiet, speech in babble, speech in noise, speech from the user, speech from a significant other, background speech, speech from multiple sources, speech from multiple sources, quiet indoor, quiet outdoor, speech in a car, speech in traffic, speech in a reverberating enviro-

onment, speech in wind noise, speech in a lounge, car noise, applause, music, e.g. classical music, and/or the like. To this end, information about the plurality of predetermined classes 323, 324, 325 may be stored in a database 321 and accessed by audio signal classification module 315. E.g., the information may comprise different patterns associated with each class 323 - 325, wherein it is determined whether audio signal 311, in particular the characteristics and/or features determined from audio signal 311, matches, at least to a certain extent, the respective pattern such that the respective class 323 - 325 can be attributed to the audio signal 311. In particular, a probability may be determined whether the respective pattern associated with the respective class 323 - 325 matches the characteristics and/or features determined from audio signal 311, wherein the respective class 323 - 325 may be attributed to audio signal 311 when the probability exceeds a threshold.

[0049] The one or more classes 323 - 325 attributed to audio signal 311 can then be employed by processing instructions selection module 317 to select audio processing instructions 333, 334, 335 which are associated with the one or more classes 323 - 325 attributed to audio signal 311. In particular, each of audio processing instructions 333, 334, 335 may be associated with at least one respective class 323, 324, 325, or a plurality of respective classes 323 - 325. For example, audio processing instructions 333, 334, 335 may be stored in a database 331 and accessed by processing instructions selection module 317. For instance, audio processing instructions 333 - 335 may be implemented as different audio processing programs and/or different audio processing programs which can be executed by audio signal processing module 313. For instance, audio processing instructions 323 - 325 may include instructions executable by processor 310 providing for at least one gain model (GM), noise cancelling (NC), wind noise cancelling (WNC), reverberation cancelling (RevC), narrowband coupling, feedback cancelling (FC), speech enhancement (SE), noise cleaning, beamforming (BF), in particular static and/or adaptive beamforming, and/or the like. E.g., at least one of audio processing instructions 323 - 325 may implement a beamforming in a rear direction of the user and at least another one of audio processing instructions 323 - 325 may implement a beamforming in a front direction of the user.

[0050] In some instances, e.g., when audio signal classification module 315 is implemented as a mixed-mode classifier, at least one of classes 323 - 325, e.g., two or more classes 323 - 325, can be attributed to the audio signal 311. For instance, when a probability that the respective pattern associated with a plurality of classes 323 - 325 matches audio signal 311 is determined to exceed a respective threshold, the plurality of classes 323 - 325 may be attributed to audio signal 311. Processing instructions selection module 317 can then be configured to select the audio processing instructions 333 - 335 associated with the plurality of classes 323 - 325

attributed to audio signal 311. For instance, the different audio processing instructions 333 - 335 associated with the different classes 323 - 325 may be mixed when executed by audio signal processing module 313, e.g., in dependence of class similarity factors indicative of a similarity of the current acoustic environment with a respective predetermined acoustic environment associated with the different classes, as disclosed in EP 1 858 292 B1 and EP 2 201 793 B1.

[0051] One or more audio processing instructions 333 - 335 can include one or more inhibition instructions which, when executed, inhibit applying the audio processing instruction 333 - 335 associated with at least another one of said classes 323 - 325. To this end, inhibition instruction module 319 may receive, from processing instruction selection module 317, information about which of audio processing instructions 333 - 335 have been selected to be currently executed by audio signal processing module 313 and/or information about one or more inhibition instructions included in one or more audio processing instructions 333 - 335 which have been selected to be currently executed by audio signal processing module 313. Depending on this information, inhibition instruction module 319 can then instruct processing instruction selection module 317 to inhibit the selection of the audio processing instruction 333 - 335 associated with at least another one of said classes 323 - 325 in order not to be applied by audio signal processing module 313.

[0052] To illustrate, during applying, by audio signal processing module 313, the at least one audio processing instruction 333 - 335 associated with at least one class 323 - 325 currently attributed to audio signal 311, audio signal classification module 315 may attribute at least another class to audio signal 311, e.g., when the current acoustic environment may have changed. Inhibition instruction module 319 may then verify whether the at least one audio processing instruction 333 - 335 currently applied includes at least one inhibition instruction which would inhibit applying the audio processing instruction 333 - 335 associated with the at least other class 323 - 325. In a case in which the at least one audio processing instruction 333 - 335 currently applied would include such an inhibition instruction, inhibition instruction module 319 can instruct processing instruction selection module 317 to inhibit the selection of the audio processing instruction 333 - 335 associated with the at least other class 323 - 325. In a case in which the at least one audio processing instruction 333 - 335 currently applied would not include such an inhibition instruction, audio processing instruction 333 - 335 can proceed to select the audio processing instruction 333 - 335 associated with the at least other class 323 - 325 such that it is applied by audio signal processing module 313, e.g., in addition to the at least one audio processing instruction 333 - 335 currently applied by audio signal processing module 313, in particular when audio signal classification module 315 would also currently attribute the at least one class to audio signal 311 which is associated with the at least one audio

processing instruction 333 - 335 currently applied by audio signal processing module 313.

[0053] In this way, an intended benefit of the at least one audio processing instruction 333 - 335 currently applied by audio signal processing module 313 can be effectively preserved and an undesired application of the at least one other audio processing instruction 333 - 335 associated with the at least other class newly attributed to audio signal 311 can be effectively inhibited, e.g., in order to avoid a negative impact on the user's hearing perception by applying the at least one other audio processing instruction 333 - 335 in addition to the at least one currently applied audio processing instruction 333 - 335. Such an operational mode may be similar to neurological processes of the human brain in which an activity of one brain region may be excited, i.e. activated, at the expense of another brain region which may then be inhibited, i.e. deactivated.

[0054] In some implementations, when the at least one audio processing instruction 333 - 335 currently applied would include an inhibition instruction which would inhibit applying the at least one other audio processing instruction 333 - 335 associated with the at least other class which has been newly attributed to audio signal 311 by audio signal classification module 315, processing instructions selection module 317 can be configured to determine a priority measure indicative of whether the at least one audio processing instruction 333 - 335 currently applied or the at least one other audio processing instruction 333 - 335 associated with the at least other class which has been newly attributed to audio signal 311 would have a higher priority to be applied. Depending on the priority measure, one of the currently applied audio processing instruction 333 - 335 and the audio processing instruction 333 - 335 associated with class 323 - 325 newly attributed to audio signal 311 which is determined to have the higher priority can then be selected by processing instructions selection module 317 to be applied by audio signal processing module 313. Accordingly, audio signal processing module 313 can then be further controlled not to apply the other of the currently applied audio processing instruction 333 - 335 and the audio processing instruction 333 - 335 associated with class 323 - 325 newly attributed to audio signal 311, which is determined to have the lower priority.

[0055] To illustrate, under certain circumstances, which may be defined by the determined priority measure, it can be preferred to cancel the inhibition instruction included in a first audio processing instruction 333 - 335 associated with a first class 323 - 325 attributed to audio signal 311 in favor of a second audio processing instruction 333 - 335 associated with a second class 323 - 325 which is also attributed to audio signal 311. Accordingly, processing instructions selection module 317 can then select the second audio processing instruction 333 - 335 to be applied by audio signal processing module 313 at the expense of applying the first audio processing instruction 333 - 335. For example, even if the first audio

processing instruction 333 - 335 associated with the first class 323 - 325 attributed to audio signal 311 would be currently applied by audio signal processing module 313 and would include an inhibition instruction inhibiting applying the second audio processing instruction 333 - 335 associated with the second class 323 - 325, which would be newly attributed to audio signal 311, the priority measure may be determined such that the second audio processing instruction 333 - 335 has a higher priority and then be selected by processing instructions selection module 317 in order to be applied by audio signal processing module 313.

[0056] As another example, in a situation in which the first and second class 323 - 325 would be newly attributed to audio signal 311, wherein the first audio processing instruction 333 - 335 would include an inhibition instruction which would inhibit applying the second audio processing instruction 333 - 335 and/or the second audio processing instruction 333 - 335 would include an inhibition instruction which would inhibit applying the first audio processing instruction 333 - 335, the priority measure may be determined by processing instructions selection module 317 in order to decide which one of the first and second audio processing instruction 333 - 335 shall be selected in order to be applied by audio signal processing module 313. In this way, a possible deadlock between applying one of the first and second audio processing instruction 333 - 335 may be resolved. Additionally or alternatively, the decision between applying the first or second audio processing instruction 333 - 335 based on the priority measure can ensure that additional aspects or circumstances are taken into account in the decision.

[0057] In some instances, the priority measure can be determined based on audio signal 311. For instance, the priority measure may then be indicative of whether the first class or the second class is dominantly represented in the audio signal. Determining the priority measure in such a way may comprise, e.g., determining the priority measure may then comprise determining a signal to noise ratio (SNR) in audio signal 311, wherein the priority measure is indicative of the SNR; and/or determining a presence of a content in audio signal 311, e.g., a speech content, a music content, etc., wherein the priority measure is indicative of the presence of the content; and/or determining a presence of a sound emitted by at least one acoustic object in the environment of the user audio signal 311, wherein the priority measure is indicative of the presence of the sound emitted by the acoustic object; and/or evaluating audio signal 311 in a psychoacoustic model, wherein the priority measure is indicative of a deviation of audio signal 311 from the psychoacoustic model; and/or evaluating audio signal 311 with regard to spatial cues indicative of a difference of a sound detected on a different position at the user, wherein the priority measure is indicative of the spatial cues; and/or determining an amount of a temporal dispersion of an impulse in audio signal 311, wherein the priority measure is indicative of the temporal dispersion.

[0058] In some instances, the priority measure can be determined to be indicative of whether audio signal 311 is representative of an acoustic situation in which applying the first audio processing instruction 333 - 335 or the second audio processing instruction 333 - 335 poses a larger threat to the user, wherein the other of the first audio processing instruction 333 - 335 or the second audio processing instruction 333 - 335 posing a smaller threat to the user is determined to have the higher priority. For instance, when audio signal 311 would be representative of an acoustic situation in which the user is rather likely encountering a traffic environment, the one of the first or second audio processing instruction 333 - 335 would be determined to have a higher priority which would allow the user to cope easier with the traffic environment. E.g., an audio processing instruction 333 - 335 providing for a beamforming in a looking direction of the user or an audio processing instruction 333 - 335 providing for an omnidirectional sound reproduction would then be determined to have a higher priority as compared to an audio processing instruction 333 - 335 providing for a beamforming in a back direction of the user. As another example, when audio signal 311 would be representative of an acoustic situation in which the user is rather likely involved in an emergency situation, e.g., when a sound of sirens or an alarm are determined to be present in audio signal 311, the one of the first or second audio processing instruction 333 - 335 would be determined to have a higher priority which would allow the user to cope easier with the emergency situation. E.g., an audio processing instruction 333 - 335 providing for a rather aggressive noise suppression may also suppress sound contributions from potentially dangerous sound sources in audio signal 311. Accordingly, a less aggressive noise suppression algorithm may be determined to have a higher priority.

[0059] In some instances, the priority measure can be determined based on sensor data. E.g., the sensor data may be provided by any of sensors 131 - 139 described above in conjunction with Fig. 2. In particular, the sensor data may comprise environmental sensor data, which may be provided, e.g., by input transducer 115, as described above, and/or barometric sensor 131 and/or ambient temperature sensor 132. The sensor data may also comprise user data indicative of a property detected on the user at the location of hearing device 110, 210, which may be provided by a user sensor 133 - 135, 137. E.g., the user sensor may be implemented as a physiological sensor 133 - 135 configured to provide the user data as physiological data indicative of a physiological property of the user. The user sensor may also be implemented as a user interface 137 configured to provide the user data as user interface data indicative of an interaction of the user with user interface 137. The sensor data may also comprise displacement data indicative of a displacement of hearing device 110, 210, which may be provided by displacement sensor 136. E.g., displacement sensor 136 may be implemented as an acceler-

ometer and/or a gyroscope and/or a magnetometer and/or the like. The sensor data may also comprise context data which may be defined as data indicative of a local and/or temporal context of the data provided by other sensors 115, 131 - 137. The context data may comprise location data and/or time data provided by location sensor 138 and/or clock 139. Sensor data may also be received from an external device via communication port 119, e.g., from a communication device.

[0060] In some implementations, processing instruction selection module 317 may determine the priority measure to be indicative of whether the sensor data is representative of a situation in which applying the first audio processing instruction or the second audio processing instruction poses a larger threat to the user, wherein the other of the first audio processing instruction or the second audio processing instruction posing a smaller threat to the user is determined to have the higher priority. To illustrate, a movement pattern contained in displacement data provided by displacement sensor 136 may indicate whether the user is moving, e.g., running or walking. In such a situation, which would for instance include the user moving in a traffic environment, the one of the first or second audio processing instruction 333 - 335 would be determined to have a lower priority which would restrict the user's hearing ability in his looking direction such as, e.g., an attenuation of sound propagating toward the user's head or a reproduction of an audio signal received from a streaming server from a remote location, for instance from a table microphone. In consequence, such an audio processing instruction 333 - 335 would be inhibited in favor of another audio processing instruction 333 - 335 which would provide for an uncompromised reproduction of sound arriving from the user's looking direction.

[0061] As another example, physiological data provided by physiological sensor 133 - 135 may indicate whether the user is in a medical emergency situation, e.g., suffering from a cardiovascular disease. In such a situation, which would for instance include a cardiac infarction, the one of the first or second audio processing instruction 333 - 335 would be determined to have a lower priority which would restrict the user's hearing ability to communicate with another person, such as an emergency personnel. In consequence, such an audio processing instruction 333 - 335 would be inhibited in favor of another audio processing instruction 333 - 335 which would allow the user to clearly understand the speech of a person in his environment and/or his own speech during talking.

[0062] In some implementations, one or more of audio processing instructions 333 - 335 provide for an enhancement of a speech content of a single talker in input audio signal 311 and/or an enhancement of a speech content of a plurality of talkers in the input audio signal 311 and/or a reproduction of sound emitted by an acoustic object in the environment of the user encoded in the input audio signal 311 and/or a reproduction of sound

emitted by a plurality of acoustic objects in the environment of the user encoded in the input audio signal 311 and/or a reduction and/or cancelling of noise and/or reverberations in the input audio signal 311 and/or a preservation of acoustic cues contained in the input audio signal 311 and/or a suppression of noise in the input audio signal 311 and/or an improvement of a signal to noise ratio (SNR) of sound encoded in the input audio signal 311 and/or a spatial resolution of sound encoded in the input audio signal 311 depending on a direction of arrival (DOA) of the sound and/or depending on a location of at least one acoustic object emitting the sound in the environment of the user and/or a directivity of an audio content in the input audio signal 311 provided by a beamforming or a preservation of an omnidirectional audio content in the input audio signal 311 and/or an amplification of sound encoded in the input audio signal 311 adapted to an individual hearing loss of the user and/or an enhancement of music content in the input audio signal 311. For instance, audio processing algorithms 333 - 335 may comprise at least one of a gain model (GM); a noise cancelling (NC) algorithm; a wind noise cancelling (WNC) algorithm; a reverberation cancelling (RevC) algorithm; a feedback cancelling (FC) algorithm; a speech enhancement (SE) algorithm; an impulse noise cancelling (INC) algorithm; an acoustic object separation (AOS) algorithm; a binaural synchronization (BS) algorithm; and a beamforming (BF) algorithm.

[0063] In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction associated with at least one of classes 323 - 325 representative of a speech in front of the user and/or noise from the side or back of the user contained in audio signal 311, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 associated with at least another class 323 - 325 representative of noise in front of the user and/or speech from the side or back of the user contained in audio signal 311. In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction associated with at least one of classes 323 - 325 representative of a static noise contained in audio signal 311, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 associated with at least another class 323 - 325 representative of a modulated noise in audio signal 311 and/or an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 providing for a speech enhancement.

[0064] In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction associated with at least one of classes 323 - 325 representative of a music contained in the audio signal, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 associated with

at least another class 323 - 325 representative of a speech, e.g., speech in a complex scenario, in audio signal 311. In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction associated with at least one of classes 323 - 325 representative of a speech and/or noise contained in audio signal 311, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 associated with at least another class 333 - 335 representative of a traffic noise contained in audio signal 311. In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction associated with at least one of classes 323 - 325 representative of a speech present in a soft sound environment contained in audio signal 311, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 associated with at least another class 323 - 325 representative of a modulated noise contained in audio signal 311.

[0065] In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction providing for noise cancelling, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 providing for beamforming. In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction providing for speech enhancement, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 providing for noise cancelling. In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction providing for speech enhancement, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 providing for music enhancement. In some implementations, audio processing instructions 333 - 335 comprise an audio processing instruction providing for noise cancelling, wherein the audio processing instruction includes an inhibition instruction which inhibits applying another audio processing instruction 333 - 335 providing for speech enhancement.

[0066] In some implementations, the priority measure is monitored when the one of the first audio processing instruction 333 - 335 and second audio processing instruction 333 - 335 is applied. E.g., the priority measure may be monitored continuously or in predetermined time intervals. During the monitoring, when the priority measure would be indicative of a higher priority of the other of the first audio processing instruction 333 - 335 and second audio processing instruction 333 - 335, the other of the first audio processing instruction 333 - 335 and second audio processing instruction 333 - 335 can be applied. In particular, in such a case, the first or second inhibition instruction included in the other of the first or

second audio processing instruction 333 - 335 may be executed. In this way, the priority measure may be continually taken into account, e.g., during applying of any of audio processing instructions 333 - 335.

[0067] FIG. 5 illustrates a block flow diagram for an exemplary method of processing input audio signal 311. The method may be executed by processor 310 of hearing device 110, 210 and/or another processor communicatively coupled to processor 310. At operation S12, audio signal 311, which may be provided by input transducer 115, is received. Further, at S12, audio signal 311 is classified by attributing at least one of classes 323 - 325 to audio signal 311. At operation S13, at least one of audio processing instructions 333 - 335 associated with the at least one class 323 - 325 attributed to audio signal 311 is selected. This may imply verifying whether the at least one audio processing instruction 333 - 335 associated with the at least one class 323 - 325 includes an inhibition instruction.

[0068] In a case in which at least one of audio processing instructions 333 - 335 attributed to at least one of classes 323 - 325 includes such an inhibition instruction which would inhibit applying at least another one of audio processing instructions 333 - 335, which may be attributed to at least another one of classes 323 - 325, the inhibition instruction is executed. Thus, applying the at least other audio processing instruction 333 - 335 to audio signal 311 at operation S14 is inhibited. Further, at operation S14, the audio processing instruction 333 - 335 including the inhibition instruction can be applied to audio signal 311. This may imply receiving audio signal 311 again at S24, e.g., an updated version of audio signal 311, to which the audio processing instruction 333 - 335 including the inhibition instruction is applied. In particular, in such a case, the at least other audio processing instruction 333 - 335 may not comprise an inhibition instruction which would inhibit applying the audio processing instruction 333 - 335 including the inhibition instruction, which is applied to audio signal 311 at operation S14.

[0069] In another case in which a first audio processing instructions 333 - 335 attributed to at least one of classes 323 - 325 would include such an inhibition instruction which would inhibit applying a second audio processing instruction 333 - 335 attributed to at least another one of classes 323 - 325, and the second audio processing instruction 333 - 335 would also include an inhibition instruction which would inhibit applying the first audio processing instruction 333 - 335, a priority measure may be determined at S13. The priority measure may be indicative of whether the first audio processing instruction 333 - 335 or the second audio processing instruction 333 - 335 would have a higher priority to be applied. Accordingly, at operation S14, the audio processing instruction 333 - 335 which has been determined to have the higher priority can be applied to audio signal 311 at operation S14. Moreover, applying the audio processing instruction 333 - 335 which has been determined to have the lower priority to audio signal 311 at operation S14 can be

inhibited.

[0070] In another case, when none of audio processing instructions 333 - 335 attributed to at least one of classes 323 - 325 would include an inhibition instruction which would inhibit applying another audio processing instruction 333 - 335 which have been attributed to at least another one of classes 323 - 325, all the audio processing instructions 333 - 335 which have been attributed to classes 323 - 325 may be applied to audio signal 311 at operation S14. In particular, in such a mode of operation, the audio processing instructions 333 - 335 may be applied to audio signal 311 by mixing the audio processing instructions 333 - 335, e.g., in accordance with the working principle of a mixed-mode classifier described above.

[0071] FIG. 6 illustrates a block flow diagram of an exemplary implementation of the method illustrated in FIG. 5. After receiving audio signal 311, audio signal 311 is classified at operation S22. During classifying of audio signal 311, only a single class 323 may be attributed to audio signal 311. Subsequently, at operation S23, audio processing instruction 333 associated with class 323 attributed to audio signal 311 is selected. Subsequently, at operation S24, audio processing instruction 333 is applied to audio signal 311. In particular, at S24, audio signal 311 may be received again, e.g., an updated version of audio signal 311, to which audio processing instruction 333 is applied.

[0072] At operation S32, e.g., after receiving audio signal 311 again, audio signal 311 is again classified. During classifying, a second class 324 may be attributed to the newly received audio signal 311, e.g., in addition to the first class 323 which already has been attributed to audio signal 311 at S22. Subsequently, at operation S33, the first audio processing instruction 333 associated with the first class 323 attributed to audio signal 311 and a second audio processing instruction 334 associated with the second class 324 attributed to audio signal 311 is selected. Accordingly, it is verified, at S33, whether the first audio processing instruction 333 would include an inhibition instruction which would inhibit applying the second audio processing instruction 334. Further, it is verified, at S33, whether the second audio processing instruction 334 would include an inhibition instruction which would inhibit applying the first audio processing instruction 333. In the illustrated example, the second audio processing instruction 334 includes an inhibition instruction inhibiting applying the first audio processing instruction 333 to audio signal 311. The first audio processing instruction 333, however, does not include an inhibition instruction which would inhibit applying the second audio processing instruction 334 to audio signal 311. Accordingly, at operation S34, e.g., after receiving audio signal 311 again, the inhibition instruction included in the second audio processing instruction 334 is executed such that applying the first audio processing instruction 333 to audio signal 311 is inhibited. However, in place of applying the first audio processing instruction

333, the second audio processing instruction 334 is applied to audio signal 311.

[0073] While the principles of the disclosure have been described above in connection with specific devices and methods, it is to be clearly understood that this description is made only by way of example and not as limitation on the scope of the invention. The above described preferred embodiments are intended to illustrate the principles of the invention, but not to limit the scope of the invention. Various other embodiments and modifications to those preferred embodiments may be made by those skilled in the art without departing from the scope of the present invention that is solely defined by the 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 fulfil 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.

Claims

1. A method of operating a hearing device configured to be worn at an ear of a user, the method comprising

- receiving an audio signal (311);
- classifying the audio signal (311) by attributing at least one class (321, 323, 324, 325) from a plurality of predetermined classes to the audio signal (311), wherein different audio processing instructions (331, 333, 334, 335) are associated with different classes (321, 323, 324, 325);
- modifying the audio signal (311) by applying the audio processing instruction (331, 333, 334, 335) associated with the class (321, 323, 324, 325) attributed to the audio signal (311); and
- controlling an output transducer (117) included in the hearing device to generate a sound output according to the modified audio signal (311),

characterized in that the audio processing instruction (331, 333, 334, 335) associated with at least one of said classes (321, 323, 324, 325) includes an inhibition instruction which, when executed, inhibits applying the audio processing instruction (331, 333, 334, 335) associated with at least another one of said classes (321, 323, 324, 325), the method further comprising

- executing the inhibition instruction when the audio processing instruction (331, 333, 334, 335) including the inhibition instruction is applied.

2. The method of claim 1, wherein the audio processing instructions (331, 333, 334, 335) comprise

- an audio processing instruction associated with a class (321, 323, 324, 325) representative of a speech in front of the user and/or noise from the side or back of the user contained in the audio signal (311), wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) associated with at least another class (321, 323, 324, 325) representative of noise in front of the user and/or speech from the side or back of the user contained in the audio signal; and/or
- an audio processing instruction associated with a class (321, 323, 324, 325) representative of a static noise contained in the audio signal (311), wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) associated with at least another class (321, 323, 324, 325) representative of a modulated noise in the audio signal (311) and/or an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) providing for a speech enhancement; and/or
- an audio processing instruction associated with a class (321, 323, 324, 325) representative of a music contained in the audio signal (311), wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) associated with at least another class (321, 323, 324, 325) representative of a speech in the audio signal (311); and/or
- an audio processing instruction associated with a class (321, 323, 324, 325) representative of a speech and/or noise contained in the audio signal (311), wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) associated with at least another class (321, 323, 324, 325) representative of a traffic noise contained in the audio signal (311); and/or
- an audio processing instruction associated with a class (321, 323, 324, 325) representative of a speech present in a soft sound environment contained in the audio signal (311), wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) associated with at least another class (321, 323, 324, 325) re-

presentative of a modulated noise contained in the audio signal (311).

3. The method of claim 1 or 2, wherein the audio processing instructions (331, 333, 334, 335) comprise

- an audio processing (331, 333, 334, 335) instruction providing for noise cancelling, wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) providing for beamforming; and/or

- an audio processing instruction (331, 333, 334, 335) providing for speech enhancement, wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) providing for noise cancelling; and/or

- an audio processing instruction (331, 333, 334, 335) providing for speech enhancement, wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) providing for music enhancement; and/or

- an audio processing instruction (331, 333, 334, 335) providing for noise cancelling, wherein the audio processing instruction (331, 333, 334, 335) includes an inhibition instruction which inhibits applying another audio processing instruction (331, 333, 334, 335) providing for speech enhancement.

4. The method of any of the preceding claims, wherein the audio processing instructions (331, 333, 334, 335) comprise a first audio processing instruction (331, 333, 334, 335) associated with a first class (321, 323, 324, 325) and a second audio processing instruction (331, 333, 334, 335) associated with a second class (321, 323, 324, 325), wherein the first audio processing instruction (331, 333, 334, 335) includes a first inhibition instruction which, when executed, inhibits applying the second audio processing instruction (331, 333, 334, 335) and/or the second audio processing instruction (331, 333, 334, 335) includes a second inhibition instruction which, when executed, inhibits applying the first audio processing instruction (331, 333, 334, 335), the method further comprising

- determining a priority measure indicative of whether the first audio processing instruction (331, 333, 334, 335) or the second audio processing instruction (331, 333, 334, 335) has a higher priority to be applied; and

- applying, depending on the priority measure, one of the first audio processing instruction (331, 333, 334, 335) and second audio processing instruction (331, 333, 334, 335).

5. The method of claim 4, wherein the priority measure is determined based on the audio signal (311).

6. The method of claim 4 or 5, wherein the priority measure is indicative of whether the first class (321, 323, 324, 325) or the second class (321, 323, 324, 325) is dominantly represented in the audio signal (311).

7. The method of any of claims 4 to 6, wherein the determining the priority measure based on the audio signal (311) comprises at least one of

- determining a signal to noise ratio (SNR) in the audio signal (311), wherein the priority measure is indicative of the SNR;

- determining a presence of a content in the audio signal (311), wherein the priority measure is indicative of the presence of the content;

- determining a presence of a sound emitted by at least one acoustic object in the environment of the user in the audio signal (311), wherein the priority measure is indicative of the presence of the sound emitted by the acoustic object;

- evaluating the audio signal (311) in a psychoacoustic model, wherein the priority measure is indicative of a deviation of the audio signal (311) from the psychoacoustic model;

- evaluating the audio signal (311) with regard to spatial cues indicative of a difference of a sound detected on different positions at the user, wherein the priority measure is indicative of the spatial cues; and

- determining an amount of a temporal dispersion of an impulse in the audio signal (311), wherein the priority measure is indicative of the temporal dispersion.

8. The method of any of claims 4 to 7, wherein the priority measure is indicative of whether the audio signal is representative of an acoustic situation in which applying the first audio processing instruction (331, 333, 334, 335) or the second audio processing instruction (331, 333, 334, 335) poses a larger threat to the user, wherein the other of the first audio processing instruction (331, 333, 334, 335) or the second audio processing instruction (331, 333, 334, 335) posing a smaller threat to the user is determined to have the higher priority.

9. The method of any of claims 4 to 8, further comprising

- receiving, from a sensor (120, 115, 125, 131 - 139), sensor data indicative of a displacement of the hearing device and/or indicative of a property of the user and/or an ambient environment of the user, 5
wherein the priority measure is determined based on the sensor data.
- 10.** The method of claim 9, wherein the sensor (115, 120, 125, 131 - 139) comprises 10
- a displacement sensor (136) configured to provide at least part of the sensor data as displacement data indicative of a displacement of the hearing device; and/or 15
- a location sensor (138) configured to provide at least part of the sensor data as location data indicative of a current location of the user; and/or
- a physiological sensor (133, 134, 135) configured to provide at least part of the sensor data as physiological data indicative of a physiological property of the user; and/or 20
- an environmental sensor (115, 131, 132) configured to provide at least part of the sensor data as environmental data indicative of a property of the environment of the user. 25
- 11.** The method of claim 10, wherein the priority measure is indicative of whether the sensor data is representative of a situation in which applying the first audio processing instruction (331, 333, 334, 335) or the second audio processing instruction (331, 333, 334, 335) poses a larger threat to the user, wherein the other of the first audio processing instruction (331, 333, 334, 335) or the second audio processing instruction (331, 333, 334, 335) posing a smaller threat to the user is determined to have the higher priority. 30
- 12.** The method of any of any of claims 4 to 11, further comprising 40
- monitoring the priority measure when the one of the first audio processing instruction (331, 333, 334, 335) and second audio processing instruction (331, 333, 334, 335) is applied; and, when the priority measure is indicative of a higher priority of the other of the first audio processing instruction (331, 333, 334, 335) and second audio processing instruction (331, 333, 334, 335), 45
- applying the other of the first audio processing instruction (331, 333, 334, 335) and second audio processing instruction (331, 333, 334, 335). 50
- 13.** The method of any of the preceding claims, wherein said audio processing instructions (331, 333, 334, 335) provide for at least one of 55
- an enhancement of a speech content of a single talker in the audio signal (311);
- an enhancement of a speech content of a plurality of talkers in the audio signal (311);
- a reproduction of sound emitted by an acoustic object in the environment of the user encoded in the audio signal (311);
- a reproduction of sound emitted by a plurality of acoustic objects in the environment of the user encoded in the audio signal (311);
- a reduction and/or cancelling of noise and/or reverberations in the audio signal (311);
- a preservation of acoustic cues contained in the audio signal (311);
- a suppression of noise in the audio signal (311);
- an improvement of a signal to noise ratio (SNR) in the audio signal (311);
- a spatial resolution of sound encoded in the audio signal (311) depending on a direction of arrival (DOA) of the sound and/or depending on a location of at least one acoustic object emitting the sound in the environment of the user;
- a directivity of an audio content in the audio signal (311) or a preservation of an omnidirectional audio content in the audio signal (311);
- an amplification of sound encoded in the audio signal (311) adapted to an individual hearing loss of the user; and
- an enhancement of music content in the audio signal (311).
- 14.** A computer-readable medium storing instructions that, when executed by a processor included in a hearing device, cause the processor to perform the method according to any of the preceding claims.
- 15.** A hearing device configured to be worn at an ear of a user, the hearing device comprising
- an input transducer (115) configured to provide an audio signal (311) indicative of a sound detected in the environment of the user;
- a processor (112, 310) configured to
- classify the audio signal (311) by attributing at least one class (321, 323, 324, 325) from a plurality of predetermined classes to the audio signal (311), wherein different audio processing instructions (331, 333, 334, 335) are associated with different classes (321, 323, 324, 325); and
- modify the audio signal (311) by applying the audio processing instruction (331, 333, 334, 335) associated with the class (321, 323, 324, 325) attributed to the audio signal (331); and
- an output transducer (117) configured to gen-

erate a sound output according to the modified audio signal (311), **characterized in that** the audio processing instruction (331, 333, 334, 335) associated with at least one of said classes (321, 323, 324, 325) includes an inhibition instruction which, when executed, inhibits applying the audio processing instruction (331, 333, 334, 335) associated with at least another one of said classes (321, 323, 324, 325), wherein the processor (112, 310) is further configured to

- execute the inhibition instruction when the audio processing instruction (331, 333, 334, 335) including the inhibition instruction is applied.

20

25

30

35

40

45

50

55

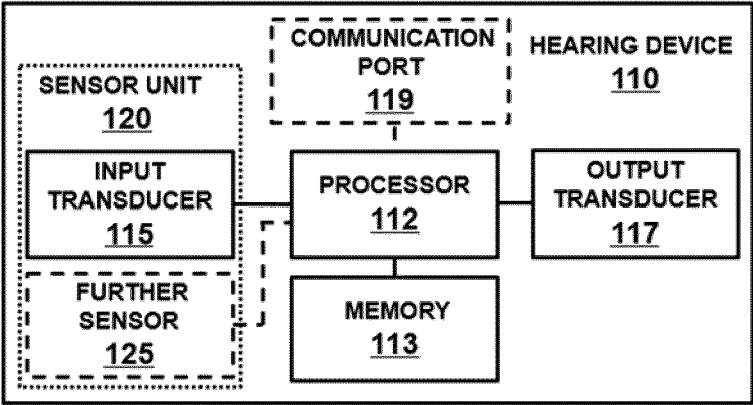


Fig. 1

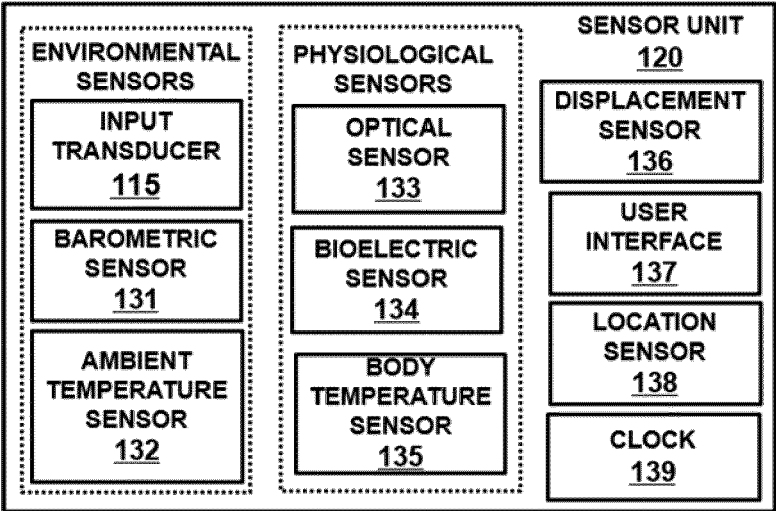


Fig. 2

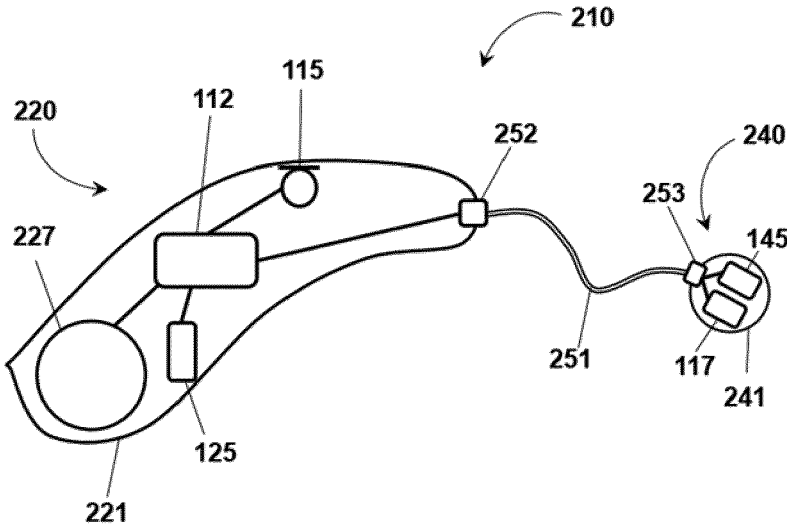


Fig. 3

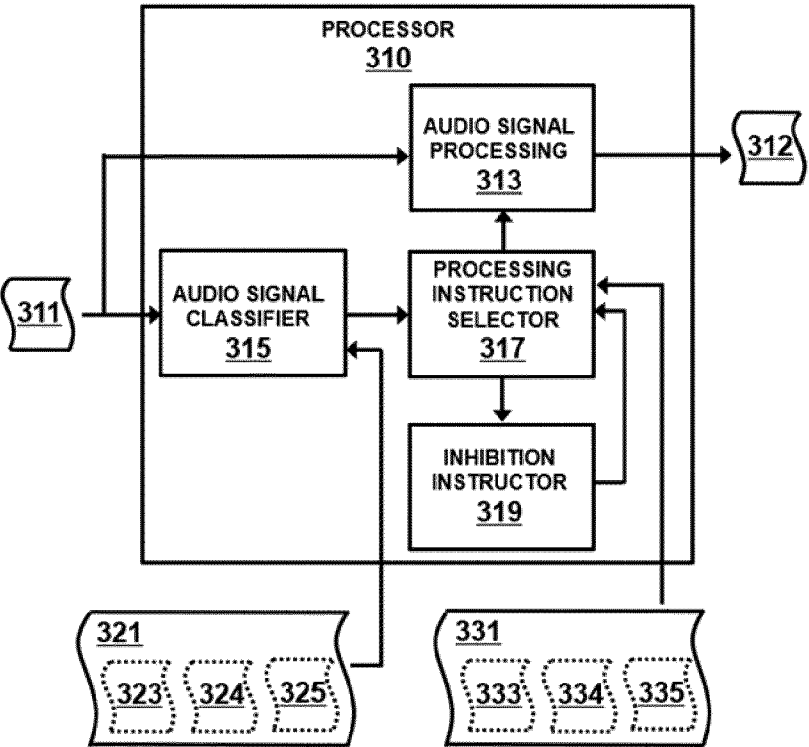


Fig. 4

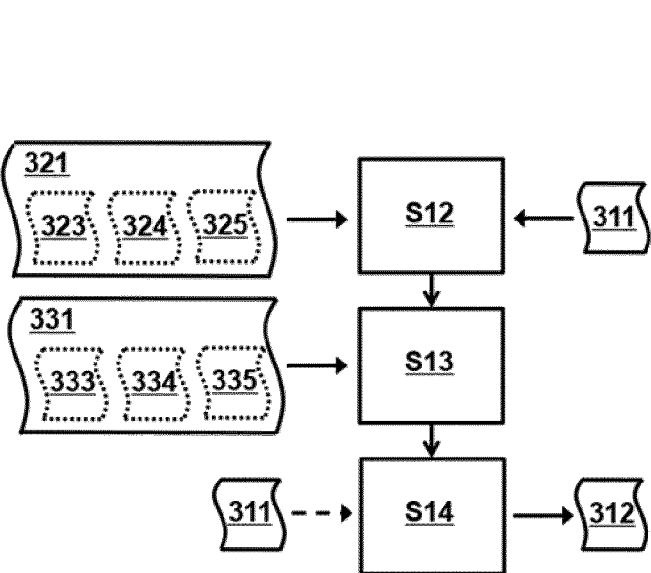


Fig. 5

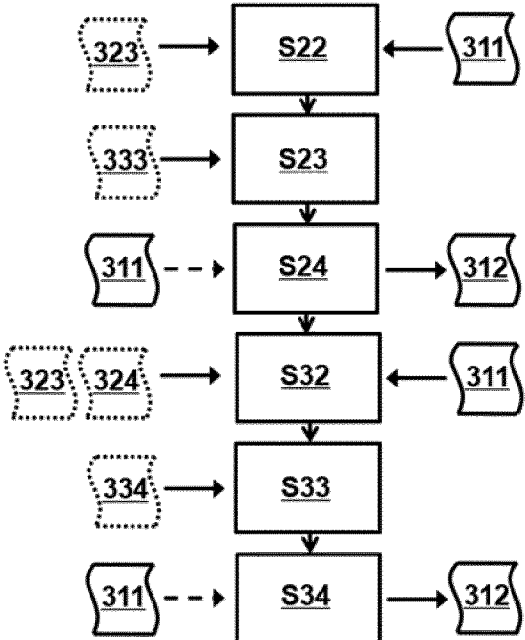


Fig. 6



EUROPEAN SEARCH REPORT

Application Number

EP 23 19 0449

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	US 2021/233539 A1 (WEXLER YONATAN [IL] ET AL) 29 July 2021 (2021-07-29) * the whole document * -----	1-15	INV. H04R1/10 H04R25/00
			TECHNICAL FIELDS SEARCHED (IPC)
			H04R
The present search report has been drawn up for all claims			
Place of search		Date of completion of the search	Examiner
The Hague		18 January 2024	Fachado Romano, A
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 23 19 0449

5

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.

18-01-2024

10

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2021233539 A1	29-07-2021	CN 113196803 A	30-07-2021
		CN 113747330 A	03-12-2021
		EP 3868128 A2	25-08-2021
		EP 3901739 A1	27-10-2021
		EP 3901740 A1	27-10-2021
		EP 3905007 A1	03-11-2021
		US 2020296521 A1	17-09-2020
		US 2021160624 A1	27-05-2021
		US 2021160625 A1	27-05-2021
		US 2021168516 A1	03-06-2021
		US 2021168517 A1	03-06-2021
		US 2021233539 A1	29-07-2021
		US 2021235201 A1	29-07-2021
		US 2021235202 A1	29-07-2021
		US 2021235207 A1	29-07-2021
		US 2021235208 A1	29-07-2021
		US 2021258703 A1	19-08-2021
		US 2022021988 A1	20-01-2022
		US 2022377467 A1	24-11-2022
		US 2023080418 A1	16-03-2023
WO 2020079485 A2	23-04-2020		

35

40

45

50

55

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

- EP 3036915 B1 [0007]
- EP 2020051734 W [0007]
- EP 2020051735 W [0007]
- DE 2019206743 [0007]
- EP 1858292 B1 [0010] [0050]
- EP 2201793 B1 [0011] [0050]