



(11)

EP 4 266 705 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
25.10.2023 Bulletin 2023/43

(51) International Patent Classification (IPC):
H04R 25/00 ^(2006.01) **H04R 1/10** ^(2006.01)
H04R 3/00 ^(2006.01) **G10K 11/178** ^(2006.01)

(21) Application number: **22386020.6**

(52) Cooperative Patent Classification (CPC):
H04R 25/453; G10K 11/178; H04R 1/1016;
H04R 1/1083; H04R 1/1091; H04R 3/005;
H04R 25/505; H04R 2460/05; H04R 2460/13

(22) Date of filing: **20.04.2022**

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

(72) Inventors:
• **LIJZENG, Johannes**
1217 CM Hilversum (NL)
• **ANTONOPOULOS, Ilias Athanasios**
1217 CM Hilversum (NL)
• **ARENDS, Aernout Steven Ferdinand**
1217 CM Hilversum (NL)

(71) Applicant: **Absolute Audio Labs B.V.**
1217 CM Hilversum (NL)

(74) Representative: **Malamis, Alkisti-Irene**
Malamis & Associates
Attorneys at Law
8 Palaia Tatoiou Street
145 64 Kifissia, Athens (GR)

(54) **AUDIO PROCESSING METHOD FOR A WEARABLE AUTO DEVICE**

(57) The disclosure relates to a method of processing an audio signal in a wearable audio device. The method comprises recording a microphone signal with one or more microphones of the audio device, and further recording an accelerometer signal with one or more accelerometers of the audio device indicative of an acceleration of the audio device with respect to one or more axes; and determining an output audio signal based on the microphone signal(s) and the accelerometer signal(s).

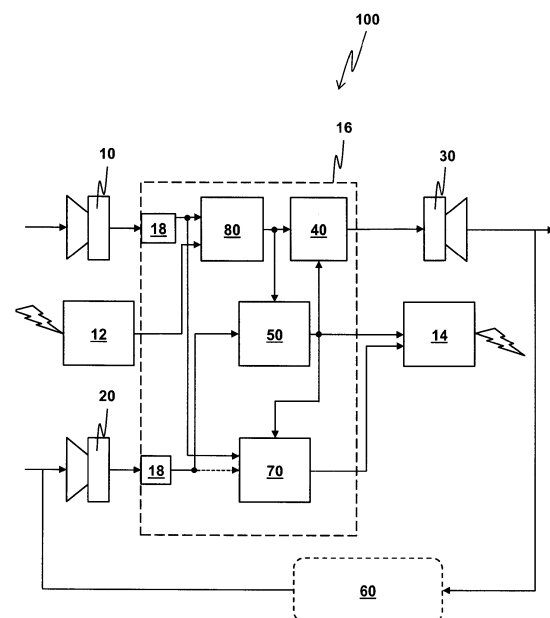


FIG 1

EP 4 266 705 A1

Description

FIELD

[0001] The invention relates to a wearable audio device, such as a hearing aid or earphone having an in-ear component, e.g. an earbud, for being inserted into the ear canal of a user and an audio signal processing method for such wearable audio device.

BACKGROUND

[0002] Wearable audio devices, such as hearing aids or earphones, can be used for providing an input audio signal to the wearer. Such input audio signal can be obtained from a microphone of the wearable audio device, or e.g. transmitted to the wearable audio device, e.g. wiredly or wirelessly. The wearable audio device can also be used for providing an output audio signal, e.g. for telephone conversation.

[0003] Hearing aids are developed to alleviate the effects of hearing loss in individuals. Modern hearing aids include a microphone for recording audio signals from the environment, and dedicated processing means for enhancing the audio signals so as to output an audio signal that is audible for the user. A variety of processing algorithms are employed in modern hearing aids, for example for noise reduction, audio compression, etc.

SUMMARY

[0004] It is aim to provide an improved processing method for processing audio signals in a wearable audio device.

[0005] Tissue conduction voice pickup is achieved by sensing the vibrations that the voice transmits through the user's body, predominantly through the skull, to the housing or other mechanical part of a wearable audio device (such as a hearing aid or earphone having an in-ear component, e.g. an earbud, for being inserted into the ear canal of a user) by an accelerometer of the wearable audio device. The tissue conduction can include bone conduction and/or soft tissue conduction. Herein, the term accelerometer is used to refer to any inertial sensor configured to detect vibration of the user's vocal chords, mouth or throat, based on vibrations in bones and tissue of the user's head. The accelerometer can e.g. be configured for determining acceleration along one or more, e.g. orthogonal, axis. The accelerometer can e.g. be a three-axis accelerometer. The wearable audio device however, typically also contains an additional input for obtaining an input audio signal. The input can e.g. be one or more microphones of the wearable audio device and/or a receiver for wiredly or wirelessly receiving an external input audio signal from an audio signal source, such as a mobile communications device. The wearable audio device also includes an output speaker for outputting sound representative of an input audio sig-

nal. A speaker signal can be provided to the speaker to be output by the speaker to the user. The speaker signal can be based on the input audio signal. The speaker, however, may also transmit vibrations via mechanical parts of the wearable audio device, such as the housing, to the accelerometer. This means that the accelerometer signal, recorded by the accelerometer, may not only include a tissue-conducted speech signal component originating from the user's own voice, but also a crosstalk signal component originating from the output speaker. This unwanted crosstalk component corrupts the accelerometer signal and can trouble the voice detection using the accelerometer signal. Herein the terms speech and speech signal are used interchangeably with voice and voice signal, respectively, and include speech, singing, humming and other sounds created using the user's mouth and/or throat.

[0006] Hereto, according to a first aspect, a method is provided of reducing crosstalk in tissue-conducted voice pickup in a wearable audio device having at least an accelerometer for voice pickup and an output speaker. The wearable audio device may for example be a hearing aid, or earphone having an in-ear component, e.g. an earbud, for being inserted into the ear canal of a user. The method comprises determining a characteristic crosstalk transfer function for audio signals being transferred from the output speaker to the accelerometer; determining, on the basis of a speaker signal to be output by the speaker and the crosstalk transfer function, an estimated crosstalk signal to be sensed by the accelerometer. The method comprises obtaining from the accelerometer an accelerometer signal representative of voice pickup, and determining a modified accelerometer signal by subtracting the estimated cross-talk signal from the accelerometer signal.

[0007] Hence, the accelerometer signal, recorded by the accelerometer, may for example include a crosstalk signal component and a tissue-conducted speech signal component, wherein an estimation of the tissue-conducted speech signal component may be estimated by subtracting the estimated crosstalk signal component from the recorded accelerometer signal.

[0008] The crosstalk transfer function represents an acoustic transfer characteristic of vibrations propagating from the output speaker to the accelerometer, particularly propagating through hardware components, such as a housing, of the wearable audio device. The accelerometer signal can be modified using the crosstalk transfer function to suppress a crosstalk signal originating from the output speaker being picked up by the accelerometer. The accelerometer can be filtered, e.g. in the time-domain or the frequency domain, using the cross-talk transfer function. The crosstalk transfer function may for example be represented as an impulse response function or a frequency response function. The crosstalk transfer function may be determined a priori, i.e. before the wearable audio device is being used by a user and/or while in use by the user. Optionally, a predetermined common

crosstalk transfer function is used for plurality of wearable audio devices, such as hearing aids or earphones, wherein the common crosstalk transfer function is used for cancelling a crosstalk component from the accelerometer signal in each wearable audio device. Optionally, an average crosstalk transfer function is determined over a plurality of wearable audio devices, wherein the average crosstalk transfer function is used for cancelling a crosstalk component from the accelerometer signal in each wearable audio device. The average crosstalk transfer function may for example be adaptively updated while in use by a user.

[0009] Optionally, the transfer function is device-specific and e.g. determined by outputting, with the output speaker of the wearable audio device, a predefined excitation signal, and recording a response thereof with the accelerometer. The excitation signal may for example be pulse, such as a short click-sound, or chirp signal outputted by the output audio device. The acoustic transfer characteristic may vary between wearable audio devices, and hence be device-specific, and may be determined for an accurate estimation of the crosstalk transfer function for a specific wearable audio device. Moreover, the acoustic transfer characteristic of the device may further vary depending on the user wearing the audio device, as the device and the user interact. Hence, the transfer function is optionally determined while the wearable audio device is worn by a user. The crosstalk transfer function may accordingly also be user specific.

[0010] Optionally, the excitation signal includes a maximum length sequence signal. The determining of the crosstalk transfer function can be particularly computationally efficient when using a maximum length sequence signal, i.e. a pseudo-random binary sequence, as the excitation signal. The crosstalk transfer function, e.g. an impulse response function, can for example be extracted by a computationally efficient deconvolution using a Hadamard transform of the accelerometer signal after outputting the maximum length sequence signal with the output speaker.

[0011] When the wearable audio device occludes the ear canal of the user, this is often experienced as unpleasant. Occlusion of the ear canal by wearable audio devices, in particular hearing aids or earphones having an in-ear component, e.g. earbuds or hearing aid receivers, for being inserted into the ear canal of a user, may however be beneficial for several reasons. For example, various audio signal processing methods are often more effective when the ear canal is occluded by the in-ear component, for example because the occlusion prevents acoustic bypass. Users may however experience such occlusion as uncomfortable, i.e. because their own voice is perceived differently compared to when the ear canal is open. The inventors realized that this phenomenon is two-fold. A first issue is that the occlusion blocks transmission of sounds from the surroundings of the user to the inner ear of the user. This first issue can be alleviated by recording surrounding sounds with one or more mi-

crophones and actively outputting the recorded sounds towards the inner ear of the user; effectively creating a "transparency mode" with the wearable audio device. A second issue is, however, that the occlusion prevents tissue-conducted audio signals to exit through the ear canal, as they otherwise would in absence of the occlusion.

[0012] Hereto, according to a second aspect, a method for actively reducing occlusion perception of a wearable audio device is provided. The method comprises obtaining an accelerometer signal with an accelerometer of the wearable audio device indicative of a tissue-conducted speech signal being generated by a user's own voice. The method comprises inverting the accelerometer signal and presenting the inverted accelerometer signal at a speaker of the wearable audio device to be outputted to the user.

[0013] The method can comprise obtaining an input audio signal. The input audio signal can be obtained with one or more microphones of the wearable audio device and/or wiredly or wirelessly received from an audio signal source. The method can comprise determining a modified input audio signal by adding the inverted accelerometer signal to the input audio signal. The modified input audio signal may be presented at the speaker of the wearable audio device and outputted to the user.

[0014] Hence, the inventors realized that the uncomfortable sense of occlusion can be reduced by presenting the inverted tissue-conducted speech signal, optionally together with the input audio signal, to be outputted by the wearable audio device towards the user. Hence, an anti-sound can be created cancelling the acoustic signal that comes from bone conduction in the ear canal. By presenting the inverted tissue-conducted speech signal, or effectively subtracting the tissue-conducted speech signal from the incoming sound, the user's sound perception is as if the tissue conducted speech signal exits the user's ear. The tissue-conducted speech signal is transmitted internally, through a body of the user, for example substantially by bone e.g. the skull of the user. The accelerometer may be arranged to sense vibrations of the user's tissue, and transduce these vibrations into an electronic signal. The accelerometer may for instance be arranged to sense vibrations of a mechanical part of the hearing aid, such as a housing, wherein said part is configured to be positioned in contact with the user's body, e.g. to a wall of the ear canal or skull. It will be appreciated that tissue-conducted audio signals may have different characteristics, such as frequency spectrum, compared to air-conducted audio signals in which sound waves propagate through the air.

[0015] Optionally, the method comprises presenting a filtered inverted accelerometer signal at the speaker. It is possible to filter the accelerometer signal before or after inverting. The filtering can e.g. include low pass filtering, e.g. for frequencies below 1 kHz, preferably below 500 Hz, more preferably between 50 and 400Hz. Low pass filtering may reduce the effect of phase differences

between the tissue conducted speech signal and the inverted accelerometer signal presented by the speaker. However, other forms of filtering are contemplated.

[0016] Optionally, the method comprises determining a magnitude spectrum of the tissue-conducted speech signal and subtracting said determined magnitude spectrum from a magnitude spectrum of the input audio signal. It has particularly been found that the frequency spectrum of the tissue-conducted speech signal equals, or at least is similar to, the frequency spectrum of an acoustic occlusion audio signal that would otherwise, i.e. with an open ear canal, be transmitted outward through the open ear canal.

[0017] Optionally, the method comprises obtaining from the accelerometer an accelerometer signal representative of voice pickup and modifying the obtained accelerometer signal by subtracting an estimated crosstalk signal from the accelerometer signal. This step may particularly be executed according to a method of reducing crosstalk in tissue-conducted voice pickup as described herein in the first aspect. The estimated cross talk signal to be subtracted from the accelerometer signal may be determined based on the speaker signal to be output by the speaker. The crosstalk signal originates from an output speaker of the wearable audio device and is conducted through the audio device's hardware components to the accelerometer. The accelerometer may sense vibrations that are originating from an output speaker via transmission through the wearable audio device, e.g. through the housing. This crosstalk signal may be considered as undesirable noise, and may therefore be cancelled from the accelerometer signal.

[0018] According to a third aspect, a method is provided of adjusting a relative contribution to an output audio signal of an accelerometer signal recorded with an accelerometer of a wearable audio device and a microphone signal recorded by one or more microphones of the wearable audio device. The wearable audio device can be a hearing aid or earphone having an in-ear component, e.g. an earbud, for being inserted into the ear canal of a user. The method comprises detecting whether or not a speech signal of a user wearing the audio device is present in the microphone signal and/or the accelerometer signal, determining a noise condition of the microphone signal when no speech signal is detected, and adjusting a contribution of the accelerometer signal relative to the microphone signal to the output audio signal based on the determined noise condition. For example, the output audio signal may be selected to be a combination of the microphone signal and the accelerometer signal according to an adjustable contribution ratio. The audio signal may for example be selected to correspond to either the microphone signal or the accelerometer signal, or a combination of the microphone signal and the accelerometer signal. It will be appreciated that the microphone signal and/or the accelerometer signal may have undergone processing, and that the output audio signal may accordingly correspond to, e.g. a processed

microphone signal and/or a processed accelerometer signal.

[0019] Optionally, the method comprises adjusting the contribution to the output audio signal of the microphone signal relative to the accelerometer signal based on a noise condition of the microphone signal. Microphones may generally provide better overall sound quality compared to accelerometers, but are also prone to being noisy. Speech intelligibility may therefore be enhanced in certain situations by reducing the contribution of the microphone signal and increasing the contribution of accelerometer signal in the output audio signal, for example when the microphone signal is very noisy. For example, with many, and/or high volume, background noises speech intelligibility may be improved by increasing the accelerometer signal contribution relative to the microphone signal in the output audio signal.

[0020] Optionally, the method comprises detecting whether or not a speech signal is present in the microphone signal and/or the accelerometer signal, and determining the noise condition of the microphone signal, e.g. only, when no speech signal is detected. For enhancing speech intelligibility, background noise may ideally be cancelled. An accurate determination of the background noise can be established in absence of a speech signal. A presence of a speech signal may be detected using various speech recognition methods. The recorded signal in absence of speech, is indicative of a background noise. This background noise may for example be filtered from the audio signal.

[0021] Optionally, the method comprises detecting whether or not a speech signal originating from a user's own voice is present in the accelerator signal, and preventing an adjustment of the relative contributions to the output audio signal of the microphone signal and the accelerometer signal in case a presence of a speech signal originating from the user's own voice is detected. It may be unwanted to adjust signal contributions to the output audio signal, while the user is talking. Hence, an adjustment setting may be frozen, for as long as the user is talking.

[0022] Optionally, the method comprises allowing adjustment of the relative contributions to the output audio signal of the microphone signal and the accelerometer signal only in case no speech signal originating from a user's own voice is detected for a predefined time period. Hence, after elapse of the predefined time period of no speech detection, the adjustment setting may be changed if required. The predefined time period may for example be at least an average pronunciation length of a phoneme, e.g. at least 0.15 seconds, preferably at least 0.25 seconds.

[0023] Optionally, a presence of a speech signal originating from the user's own voice in the accelerometer signal is ascertained, in case a loudness level of the accelerator signal exceeds a predefined loudness threshold level. The accelerometer signal may be filtered and/or cleaned up for increased accuracy of own voice detec-

tion. The accelerometer may be particularly sensitive to vibrations caused by a user's own speech, which may be transmitted through the user's body such as the skull. For example, a current magnitude of the accelerator signal exceeding a predefined magnitude threshold, may indicate that the user is currently talking. Conversely, a current magnitude of the accelerator signal being below the predefined magnitude threshold, may indicate that the user is not currently talking.

[0024] Optionally, the method includes subtracting an estimated crosstalk signal from the accelerometer signal prior to determining whether the accelerometer signal includes a speech signal originating from the user's own voice. This step may particularly be executed according to a method of reducing crosstalk in tissue-conducted voice pickup as described herein in the first aspect. The estimated cross talk signal to be subtracted from the accelerometer signal may be determined based on the speaker signal to be output by the speaker as described herein in the second aspect.

[0025] Optionally, the method includes increasing the relative contribution of the accelerometer signal, relative to the microphone signal, to the output audio signal, in case the noise condition of the microphone signal exceeds a predefined noise threshold. In case the microphone signal is very noisy, it may be desirable to increase the accelerometer signal relative to the microphone signal. However, if the user is currently speaking, such adjustment may be inhibited, and optionally executed after the user has stopped talking.

[0026] Optionally, the method includes increasing the relative contribution of the accelerometer signal, relative to the microphone signal, to the output audio signal, in case a loudness level of the microphone signal exceeds a predefined loudness threshold level. Loud sounds may be very uncomfortably perceived by the user, and furthermore microphone saturation is likely to occur. If sounds above the predefined loudness threshold level are detected, the microphone signal may be partially or entirely suppressed, or the microphone may be partly or entirely muted. This adjustment may be given precedence over other conditionals, and may for instance be executed regardless of the user talking.

[0027] Optionally, the method comprises eliminating the contribution of the microphone signal to the output audio signal entirely in case a loudness level of the microphone signal exceeds a predefined loudness threshold level. The microphone signal may for example be completely suppressed or the microphone may be completely muted, e.g. by switching off the microphone. The output audio signal may for example be selected to correspond to the accelerator signal in case the determined loudness level of the microphone signal exceeds the predetermined threshold level.

[0028] According to a further aspect, a wearable audio device, such as a hearing aid or earphone having an in-ear component, e.g. an earbud, for being inserted into the ear canal of a user, is provided configured for exe-

cuting a method as described herein.

[0029] The wearable audio device may for example comprise one or more microphones arranged for recording the microphone signal, an accelerometer arranged for recording the accelerometer signal, and a processing unit configured for receiving the microphone signal and the accelerometer signal and determining the output audio signal based on the microphone signal and the accelerometer signal. Optionally, the processing unit includes an active occlusion cancellation module for cancelling a user-perceived occlusion of its hearing canal. Optionally, the processing unit comprises a crosstalk cancellation module for canceling a crosstalk signal from the accelerometer signal. Optionally, the processing unit comprises a dynamic mixing module for receiving the microphone signal and the accelerometer signal, and outputting an audio signal being either one of the microphone signal and the accelerometer signal, or a mix of the microphone signal and accelerometer signal according to an adaptable contribution ratio.

[0030] A wearable audio device can be provided having at least an accelerometer for voice pickup and an output speaker. The wearable audio device can include a crosstalk cancellation module configured for determining a characteristic crosstalk transfer function for audio signals being transferred from the output speaker to the accelerometer; determining, on the basis of a speaker signal to be output by the speaker and the cross-talk transfer function, an estimated cross-talk signal to be sensed by the accelerometer; obtaining from the accelerometer an accelerometer signal representative of voice pickup; and determining a modified accelerometer signal by subtracting the estimated crosstalk signal from the accelerometer signal.

[0031] A wearable audio device can be provided having an accelerometer and a speaker. The wearable audio device can include an occlusion cancellation module configured for obtaining an accelerometer signal with an accelerometer of the wearable audio device indicative of a tissue-conducted speech signal being generated by a user's own voice; inverting the accelerometer signal; and presenting the inverted accelerometer signal at a speaker of the wearable audio device to be outputted to the user. Optionally, the occlusion cancellation module is configured for obtaining an input audio signal; determining a modified input audio signal by adding the inverted accelerometer signal to the input audio signal; and presenting the modified input audio signal at the speaker of the wearable audio device to be outputted to the user.

[0032] A wearable audio device can be provided having an accelerometer and one or more microphones. The wearable audio device can include a dynamic mixing module configured for detecting whether or not a speech signal is present in the microphone signal and/or the accelerometer signal; determining a noise condition of the microphone signal when no speech signal is detected; and adjusting a contribution of the accelerometer signal relative to the microphone signal to the output audio sig-

nal based on the determined noise condition.

[0033] It will be appreciated that any of the aspects, features and options described herein can be combined. It will particularly be appreciated that any of the aspects, features and options described in view of the methods apply equally to the wearable hearing device, and vice versa.

BRIEF DESCRIPTION OF THE DRAWINGS

[0034] Embodiments of the present invention will now be described in detail with reference to the accompanying drawings in which:

Figures 1 and 2 show schematic examples of a wearable audio device;

Figures 3-5 show exemplary flow charts.

DETAILED DESCRIPTION

[0035] Figure 1 shows a schematic example of a wearable audio device 100, such as a hearing aid or earphone, e.g. an in-earphone, on-earphone, or over-earphone, such as an earbud. Figure 2 shows an example of a wearable audio device 100, such as a hearing aid, an earphone, e.g. an in-earphone, on-earphone, or over-earphone. Figure 2 particularly shows an in-earphone for being inserted into a hearing canal of a user. In this example, the wearable audio device has an in-ear component for being inserted into the ear canal of a user. In this example, the wearable audio device 100 is arranged for receiving an input audio signal that is to be communicated to the user. The input audio signal can be retrieved by a first input transducer, such as a microphone 10. The wearable audio device 100 can include more than one microphone, although only one microphone is shown in Figure 1 for clarity. Alternatively, or additionally, the input audio signal can be wirelessly or wiredly received, e.g. from an audio signal source, such as a mobile communications device, using a receiver 12. In this example, the input audio signal can be received via an IEEE 802.15 connection, via Bluetooth, or BTLE. The wearable audio device 100 can be arranged for generating an output audio signal that is to be communicated to a third party, e.g. for telephone communication. The output audio signal can e.g. be transmitted using a transmitter 14. The output audio signal can be representative of speech of the user. The speech of the user can e.g. be retrieved by a second input transducer, such as one or more accelerometers 20.

[0036] The wearable audio device 100 also comprises one or more output transducers, including an output speaker 30, for outputting incoming sound representative of the input audio signal to the wearer of the audio device.

[0037] The wearable audio device 100 may particularly include an in-ear component for being inserted into a hearing canal of a user, such as an earbud, wherein the in-ear component may occlude the hearing canal. The

output speaker 30 may be arranged at an internal side of the in-ear component for transmitting the incoming sound substantially towards the inner ear of the user, while wearing the audio device. The accelerometer 20 may also be arranged at the in-ear component, and may for example be configured to sense vibrations, e.g. vibrations transmitted through the user's tissue, e.g. via bone conduction and/or soft tissue conduction. The microphone(s) 10 may be arranged for sensing sounds from an environment of the user, and may therefore also be arranged at an outward side of the in-ear component, but also at other positions such as at an optional behind-the-ear component of the wearable audio device. The wearable audio device 100 may include multiple microphones and/or multiple accelerometers, arranged at various positions.

[0038] The wearable audio device includes a processing unit 16 for processing the input audio signals and/or output audio signals.

[0039] The wearable audio device 100, e.g. the processing unit 16 may comprise an analog to digital converter 18 configured for receiving an analog audio signal, e.g. from the microphone 10 and/or the accelerometer 20, and converting the analog audio signal to a digital audio signal. The audio signal can be divided into multiple time frame signals. For example, the digital audio signal may be divided into successive time frame signals, either overlapping or non-overlapping, that are, at least partly, shifted in time. Each time frame signal may be represented by a finite number of samples, particularly 2^N samples wherein 2^N is a power of 2, such as 256, 512 or 1024.

[0040] The wearable audio device 100, e.g. the processing unit 16, here, comprises an active occlusion cancellation module 40, configured for cancelling a user-perceived occlusion of the hearing canal. Figure 3 shows an exemplary flow chart of a method performed by the occlusion cancellation module 40. The occlusion cancellation module 40 is in this example configured to receive 210 an input audio signal, e.g. from the microphone 10 or receiver 12 and accelerometer 20 and output 260 a modified audio signal for being outputted by the output speaker 30. As will be explained below, in this example the accelerometer signal is received via a crosstalk cancellation module 50. The user-perceived occlusion of the hearing canal by the in-ear component may be caused by two mechanisms. Firstly, the in-ear component blocks external air-conducted audio signals from entering the hearing canal. Passing through external sound signal such as received by the microphone 10 via a transparency mixer 80, e.g. amplified, may substantially alleviate this effect. Secondly, the in-ear component blocks an escape through the hearing canal of tissue-conducted audio signals generated by the user itself. This second effect may particularly be actively alleviated by the occlusion cancellation module 40.

[0041] The occlusion cancellation module 40 is particularly configured to obtain an accelerometer signal from the accelerometer 20 (here via crosstalk cancellation

module 50) indicative of a tissue-conducted speech signal being generated by a user's own voice. The occlusion cancellation module 40 is hence configured to receive 220 the accelerometer signal. Based on the accelerometer signal, a tissue-conducted speech signal is estimated 230 by the occlusion cancellation module 40. The tissue-conducted speech signal is generated by a user's own voice and conducted through the user's body tissue to the accelerometer. The tissue-conducted speech signal can e.g. be estimated to be proportional to, e.g. equal to, the accelerometer signal. The occlusion cancellation module 40 may invert 240 the accelerometer signal representative of speech and present 260 the inverted tissue-conducted speech signal at the speaker 30. This may be perceived by the user as if the tissue-conducted speech signal has escaped through the hearing canal, as it would in absence of any occlusion. The inverted tissue conducted speech signal may for example be filtered, e.g. in the frequency domain, by an occlusion cancellation filter, which occlusion cancellation filter is determined based on the accelerometer signal.

[0042] In case the wearable audio device 100 receives an input audio signal from the microphone(s) 10 and/or the receiver 12, such as via the transparency mixer 80, the inverted accelerometer signal representative of tissue-conducted speech may be added 250 to the input audio signal. Hence, a contribution of the tissue-conducted speech signal is suppressed in the sound provided to the user by the speaker 30, e.g. effectively by subtracting the estimated tissue-conducted speech signal from the input output audio signal.

[0043] In addition to tissue-conducted audio signals, such as originating from the user's own voice, the accelerometer 20 may also record a crosstalk signal originating from the output speaker and being transmitted through hardware components 60 of the wearable audio device 100 back towards the accelerometer 20. The accelerometer signal may therefore be corrupted with a crosstalk signal. For various applications, it may be desired to remove the crosstalk signal from the accelerometer signal. Particularly when performing occlusion cancellation it can be beneficial to also perform crosstalk cancellation. The wearable audio device 100, e.g. the processing unit 16, may hence, here further, comprise a crosstalk cancellation module 50. Figure 4 shows an exemplary flow chart of a method performed by the crosstalk cancellation module 50. The crosstalk cancellation module 50 is in this example arranged between the accelerometer 20 and the occlusion cancellation module 40. Hence, here, the accelerometer signal is first processed by the crosstalk cancellation module 50 wherein the resultant accelerometer signal is subsequently used by the occlusion cancellation module 40 for modifying the speaker signal to be output by the speaker. Hence, here, the accelerometer signal is first processed by the crosstalk cancellation module 50 wherein the resultant crosstalk-corrected accelerometer signal is subsequently used by the occlusion cancellation module 40 for mod-

ifying the speaker signal.

[0044] The crosstalk cancellation module 50 is particularly configured for determining 310 a device-specific crosstalk transfer function, characterizing a transfer of audio signals, e.g. vibrations, from the output speaker 30 to the accelerometer 20, e.g. via the wearable audio device hardware 60. A crosstalk component is, at least partially, cancelled 320 from the accelerometer signal using the determined device-specific crosstalk transfer function. The crosstalk transfer function of a device may for example be determined by outputting 311 an excitation signal with the output speaker 30, and recording 312 an accelerometer signal with the accelerometer 20. The device-specific crosstalk transfer function may be determined 310 in various ways from the known excitation signal and the observed response as recorded by the accelerometer 20. Various excitation signals may be used. A particularly computationally efficient determination of the device-specific crosstalk transfer function can be obtained by using a maximum length sequence (MLS) signal as the excitation signal. The device-specific crosstalk transfer function can subsequently be used, in various ways, for filtering the accelerometer signal so as to effectively subtract the crosstalk signal from the accelerometer signal. The crosstalk cancellation module 50 may receive a speaker signal, such as from the transparency mixer 80, to be outputted by the output speaker 30, and filter, e.g. with an adaptive filter, the accelerometer signal based on the speaker signal and the device-specific cross-talk transfer function. The crosstalk cancellation module may calculate, using the device-specific cross-talk transfer function and the received speaker signal, the predicted signal component that is predicted to be recorded at the accelerometer in view of the to be outputted audio signal, and subtract said signal component from the actually recorded accelerometer signal. The modified accelerometer signal may be provided to the transmitter 14. Alternatively, or additionally, the modified accelerometer signal may be provided to the occlusion cancellation module 40 as described above.

[0045] In this example, the wearable audio device 100 further comprises a dynamic mixing module 70 configured for receiving the microphone signal and the accelerometer signal, and generating an output audio signal, to be transmitted to a third party, being either one of the microphone signals and the accelerometer signal, or a mix of the microphone signal and accelerometer signal according to an adaptive mixing ratio, particularly to optimize speech intelligibility. The mixing ratio can be adaptively determined based various audio conditions in the microphone signal and/or the accelerometer signal. Figure 5 shows an exemplary flow chart of the dynamic mixing module 70. The mixing module 70 is arranged to detect 410 whether or not a loudness level of the microphone signal exceeds a predefined loudness threshold. If the microphone signal loudness exceeds the predefined loudness threshold (yes), adjustment of the relative contributions of the accelerometer signal and the micro-

phone signal is allowed 420. If the microphone signal loudness does not exceed the predefined loudness threshold (no), the mixing module 70 detects 430 whether or not a speech signal originating from the user's own voice is present, e.g. based on the accelerometer signal. If such speech signal is detected 430 (yes), adjustment of the relative contributions of the accelerometer signal and the microphone signal is prohibited 440. If such speech signal is not detected (no), adjustment of the relative contributions of the accelerometer signal and the microphone signal is allowed 420. Whether or not the user currently speaks may be detected using the accelerometer. It may for example be ascertained that the user is currently speaking if a magnitude of the accelerometer signal exceeds a predefined threshold.

[0046] The adjustment of the relative contributions of the accelerometer signal and the microphone signal to the output signal may be based on a determined noise level of the microphone signal. For example, a relative contribution of the accelerometer signal to the mixing module output signal may be increased, relative to the microphone signal contribution, when the microphone signal becomes more noisy. Regardless of the noise level, the mixing module may for example adjust the relative contributions in such a way, that the mixing module output signal only contains the accelerometer signal, in case a loudness level of the microphone signal exceeds a predetermined loudness threshold, e.g. regardless of the noise level and regardless of the user talking.

[0047] It will be appreciated that the mixing module 70 may be used in conjunction with the crosstalk cancellation module 50. In that case, the mixing module can receive the accelerometer signal from the crosstalk cancellation module 50. In case the mixing module 70 is used without using the crosstalk cancellation module 50, the mixing module can receive the accelerometer signal directly from the accelerometer 20 (or converter 18), as indicated by the dashed arrow in Figure 1.

[0048] Figure 1 shows an exemplary wearable audio device 100 comprising a particular combination of the occlusion cancellation module 40, the crosstalk cancelling module 50 and the mixing module 70, but it will be appreciated that the wearable audio device can include any combination of said modules. For example, the wearable audio device 100 may include only one, anyone, of said modules, or only two, any two, of said modules. Some or all of the occlusion cancellation module 40, the crosstalk cancelling module 50 and/or the mixing module 70 may be part of the processing unit 16 of the wearable audio device 100. For example, the wearable audio device can include the occlusion cancellation module 40. The wearable audio device can e.g. include the occlusion cancellation module while not including, or not using, the crosstalk cancellation module. It is possible to minimize crosstalk by other, such as mechanical, means. In another example, the wearable audio device can include the crosstalk cancelling module 50. The wearable audio device can e.g. include the crosstalk cancellation module

while not including, or not using, the occlusion cancellation module. In another example, the wearable audio device can include the mixing module 70. The wearable audio device can e.g. include the mixing module while not including, or not using, the occlusion cancellation module and/or the crosstalk cancellation module. In an example, the wearable audio device can include the occlusion cancellation module and the crosstalk cancelling module, e.g. while not including the mixing module.

[0049] Herein, the invention is described with reference to specific examples of embodiments of the invention. It will, however, be evident that various modifications and changes may be made therein, without departing from the essence of the invention. For the purpose of clarity and a concise description, features are described herein as part of the same or separate embodiments, however, alternative embodiments having combinations of all or some of the features described in these separate embodiments are also envisaged.

[0050] However, other modifications, variations, and alternatives are also possible. The specifications, drawings and examples are, accordingly, to be regarded in an illustrative sense rather than in a restrictive sense.

[0051] For the purpose of clarity and a concise description, features are described herein as part of the same or separate embodiments, however, it will be appreciated that the scope of the invention may include embodiments having combinations of all or some of the features described.

[0052] In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. The word 'comprising' does not exclude the presence of other features or steps than those listed in a claim. Furthermore, the words 'a' and 'an' shall not be construed as limited to 'only one', but instead are used to mean 'at least one', and do not exclude a plurality. The mere fact that certain measures are recited in mutually different claims does not indicate that a combination of these measures cannot be used to an advantage.

Claims

1. Method of reducing crosstalk in tissue-conduction voice pickup in a wearable audio device having at least an accelerometer for voice pickup and an output speaker, the method comprising:

determining a characteristic crosstalk transfer function for audio signals being transferred from the output speaker to the accelerometer, determining, on the basis of a speaker signal to be output by the speaker and the crosstalk transfer function, an estimated crosstalk signal to be sensed by the accelerometer, obtaining from the accelerometer an accelerometer signal representative of voice pickup, and determining a modified accelerometer signal by

- subtracting the estimated crosstalk signal from the accelerometer signal.
2. Method of claim 1, wherein the crosstalk transfer function is device-specific, and optionally determined, e.g. while the wearable audio device is worn by a user, by outputting, with the output speaker, a predefined excitation signal, and recording a response signal thereof with the accelerometer. 5
 3. Method of claim 2, wherein the excitation signal includes a maximum length sequence signal. 10
 4. Method of actively reducing occlusion perception of a wearable audio device, the method comprising: 15
 - obtaining an accelerometer signal with an accelerometer of the wearable audio device indicative of a tissue-conducted speech signal being generated by a user's own voice; 20
 - inverting the accelerometer signal; and
 - presenting the inverted accelerometer signal at a speaker of the wearable audio device to be outputted to the user.
 5. Method of claim 4, comprising: 25
 - obtaining an input audio signal;
 - determining a modified input audio signal by adding the inverted accelerometer signal to the input audio signal; and 30
 - presenting the modified input audio signal at the speaker of the wearable audio device to be outputted to the user.
 6. Method of claim 5, comprising determining a magnitude spectrum of the accelerometer signal and subtracting said determined magnitude spectrum from a magnitude spectrum of the input audio signal. 35
 7. Method of claim 4, 5 or 6, comprising obtaining from the accelerometer an accelerometer signal representative of voice pickup and modifying the obtained accelerometer signal by subtracting an estimated crosstalk signal from the accelerometer signal, wherein optionally the estimated crosstalk signal is determined according to any one of claims 1-3. 40
 8. Method of adjusting a relative contribution to an output audio signal of an accelerometer signal recorded with an accelerometer of a wearable audio device and a microphone signal recorded by one or more microphones of the wearable audio device, the method comprising 45
 - detecting whether or not a speech signal is present in the microphone signal and/or the accelerometer signal, 50
 - determining a noise condition of the microphone signal when no speech signal is detected, and adjusting a contribution of the accelerometer signal relative to the microphone signal to the output audio signal based on the determined noise condition.
 9. Method of claim 8, comprising detecting whether or not a speech signal originating from a user's own voice is present in the accelerator signal, and preventing an adjustment of the relative contributions to the output audio signal of the microphone signal and the accelerometer signal in case a presence of a speech signal originating from the user's own voice is detected.
 10. Method of claim 9, comprising allowing the adjustment of the relative contributions to the output audio signal of the microphone signal and the accelerometer signal only in case no speech signal originating from a user's own voice during a predefined time period is detected.
 11. Method of claim 9 or 10, wherein a presence of a speech signal originating from the user's own voice in the accelerometer signal is ascertained, in case a loudness level of the accelerator signal exceeds a predefined loudness threshold level.
 12. Method of any of claims 8-11, comprising increasing the relative contribution of the accelerometer signal, relative to the microphone signal, to the output audio signal, in case the noise condition of the microphone signal exceeds a predefined noise threshold.
 13. Method of any of claims 8-12, comprising increasing the relative contribution of the accelerometer signal, relative to the microphone signal, to the output audio signal, in case a loudness level of the microphone signal exceeds a predefined loudness threshold level.
 14. Method of claim 13, comprising eliminating the contribution of the microphone signal to the output audio signal entirely in case a loudness level of the microphone signal exceeds a predefined loudness threshold level.
 15. Wearable audio device, such as a hearing aid or earphone having an in-ear component, configured for executing a method of any preceding claim. 55

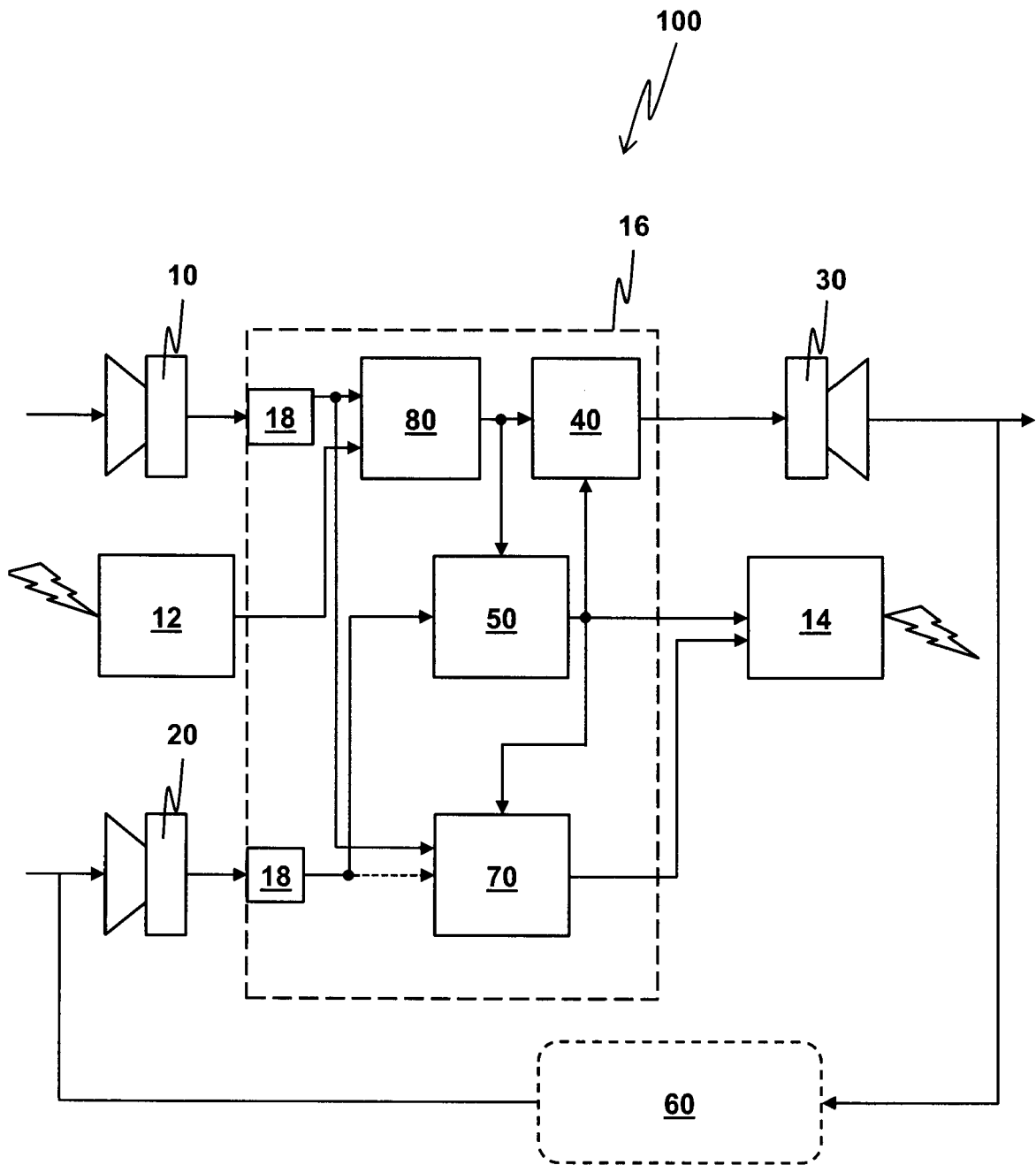


FIG 1

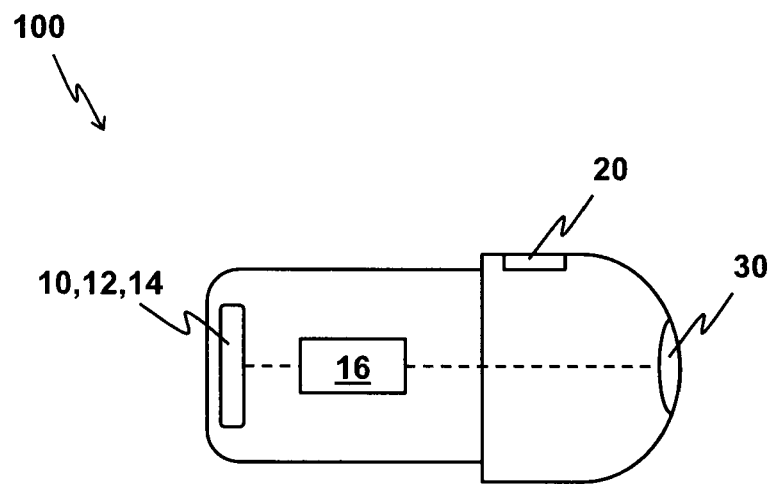


FIG 2

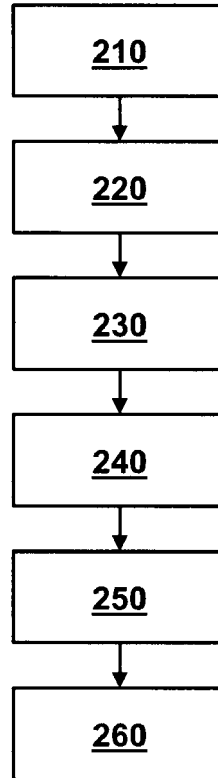


FIG 3

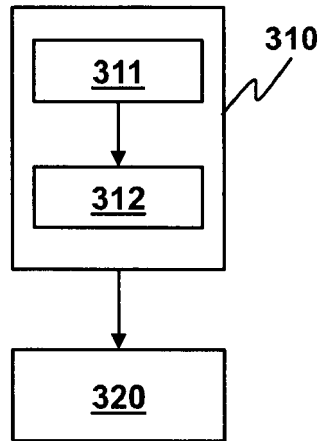


FIG 4

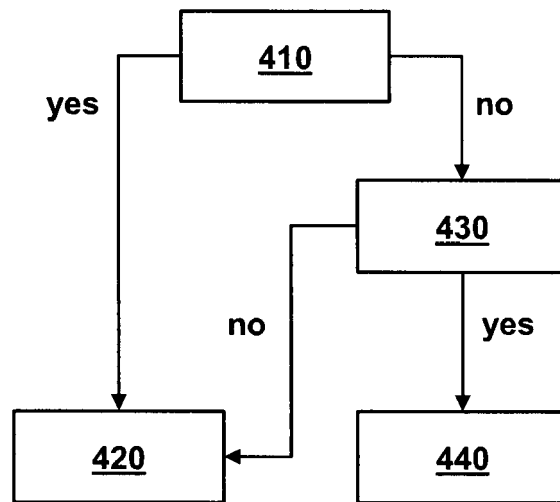


FIG 5



EUROPEAN SEARCH REPORT

Application Number

EP 22 38 6020

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	WO 2022/054008 A1 (COCHLEAR LTD [AU]) 17 March 2022 (2022-03-17) * abstract * * paragraphs [0032] - [0048] * * figures 3,3a,3b * -----	1-3,15	INV. H04R25/00 H04R1/10 H04R3/00 G10K11/178
X	WO 2016/185441 A1 (COCHLEAR LTD [AU]) 24 November 2016 (2016-11-24) * abstract * * paragraphs [0048] - [0073] * * figures 2,3,3a,3b * -----	1-3,15	
A	WO 2010/094812 A2 (PHONAK AG [CH]; HAMACHER VOLKMAR [CH]) 26 August 2010 (2010-08-26) * the whole document * -----	1-3,15	
X	EP 3 182 721 A1 (SONY MOBILE COMMUNICATIONS INC [JP]) 21 June 2017 (2017-06-21) * abstract * * paragraphs [0045] - [0058] * * paragraphs [0066] - [0068] * * figures 3,4,8,9 * -----	4-7,15	TECHNICAL FIELDS SEARCHED (IPC) H04R G10K
X	US 2020/020313 A1 (HUA THANH PHONG [US] ET AL) 16 January 2020 (2020-01-16) * abstract * * paragraphs [0033] - [0036] * * figures 7,8 * -----	4-7,15	
X	CN 113 645 534 A (GOERTEK TECH CO LTD) 12 November 2021 (2021-11-12) * the whole document * -----	4-7,15	
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 1 December 2022	Examiner Sucher, Ralph
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)



EUROPEAN SEARCH REPORT

Application Number

EP 22 38 6020

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 2022/068298 A1 (WAX ARIC J [US] ET AL) 3 March 2022 (2022-03-03) * abstract * * paragraphs [0037] - [0059] * * figures 3,4 *	8-15	
X	US 2014/270231 A1 (DUSAN SORIN V [US] ET AL) 18 September 2014 (2014-09-18) * abstract * * paragraphs [0041] - [0050] * * figures 4-6 *	8-15	
X	CN 111 683 319 A (BEIJING EDIFIER TECH CO LTD) 18 September 2020 (2020-09-18) * the whole document *	8-15	
The present search report has been drawn up for all claims			TECHNICAL FIELDS SEARCHED (IPC)
Place of search Munich			Date of completion of the search 1 December 2022
Examiner Sucher, Ralph			
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			



Application Number

EP 22 38 6020

5

10

15

20

25

30

35

40

45

50

55

CLAIMS INCURRING FEES

The present European patent application comprised at the time of filing claims for which payment was due.

☐ Only part of the claims have been paid within the prescribed time limit. The present European search report has been drawn up for those claims for which no payment was due and for those claims for which claims fees have been paid, namely claim(s):

☐ No claims fees have been paid within the prescribed time limit. The present European search report has been drawn up for those claims for which no payment was due.

LACK OF UNITY OF INVENTION

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

see sheet B

☒ All further search fees have been paid within the fixed time limit. The present European search report has been drawn up for all claims.

☐ As all searchable claims could be searched without effort justifying an additional fee, the Search Division did not invite payment of any additional fee.

☐ Only part of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the inventions in respect of which search fees have been paid, namely claims:

☐ None of the further search fees have been paid within the fixed time limit. The present European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims, namely claims:

☐ The present supplementary European search report has been drawn up for those parts of the European patent application which relate to the invention first mentioned in the claims (Rule 164 (1) EPC).

**LACK OF UNITY OF INVENTION
SHEET B**

Application Number

EP 22 38 6020

The Search Division considers that the present European patent application does not comply with the requirements of unity of invention and relates to several inventions or groups of inventions, namely:

1. claims: 1-3 (completely); 15 (partially)**method of reducing crosstalk in a wearable audio device**

2. claims: 4-7 (completely); 15 (partially)**method of reducing occlusion perception of a wearable audio device**

3. claims: 8-14 (completely); 15 (partially)**method of adjusting a relative contribution of an accelerometer signal and a microphone signal to an output signal of a wearable audio device**

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO.

EP 22 38 6020

01-12-2022

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.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2022054008 A1	17-03-2022	NONE	
WO 2016185441 A1	24-11-2016	CN 107925828 A	17-04-2018
		EP 3298802 A1	28-03-2018
		US 2016345107 A1	24-11-2016
		WO 2016185441 A1	24-11-2016
WO 2010094812 A2	26-08-2010	EP 2577996 A2	10-04-2013
		US 2013156202 A1	20-06-2013
		WO 2010094812 A2	26-08-2010
EP 3182721 A1	21-06-2017	CN 106888414 A	23-06-2017
		EP 3182721 A1	21-06-2017
		US 2017171679 A1	15-06-2017
US 2020020313 A1	16-01-2020	NONE	
CN 113645534 A	12-11-2021	NONE	
US 2022068298 A1	03-03-2022	NONE	
US 2014270231 A1	18-09-2014	NONE	
CN 111683319 A	18-09-2020	NONE	