

(51) Int Cl.:  
**H04R 1/10** (2006.01)

(22) Date of filing: **29.11.2016**

(74) Representative: **Kransell & Wennborg KB**  
**P.O. Box 27834**  
**115 93 Stockholm (SE)**

(54) **SYSTEM AND METHOD FOR ENABLING COMMUNICATION OF AMBIENT SOUND AS AN AUDIO STREAM**

cation that the playback of the first audio stream should be altered. The method also comprises, in response to the obtained indication: altering the output of the first audio stream; by means of a microphone 5 of the communication device, recording a second audio stream; and outputting the second audio stream to the headphone for playback to the user.

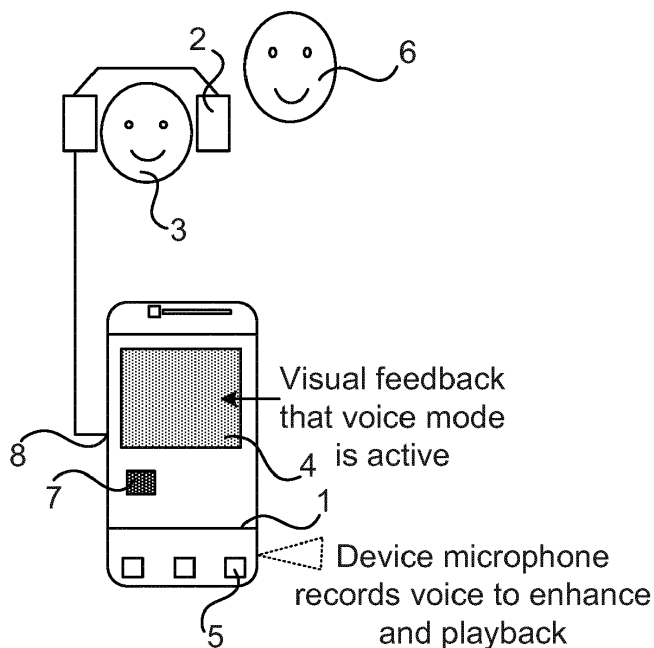


Fig. 1c

## Description

### TECHNICAL FIELD

[0001] The present disclosure relates to a communication device connected to a headphone and to enabling communication of ambient sound, including outputting an audio stream to the headphone for playback to a user of the communication device.

### BACKGROUND

[0002] When using headphones for e.g. listening to music, it is often desirable to shut out ambient sounds in order to improve the listening experience. There are also actively noise cancelling headphones on the market for further reduction of sound pollution when using the headphones. This implies that it may be difficult for a person using the headphones to hear another person trying to talk to him/her, unless the headphones are turned off or removed.

### SUMMARY

[0003] It is an objective of the present invention to improve verbal communication with a person wearing headphones, without the need to remove the headphones from the ears of said person.

[0004] According to an aspect of the present invention, there is provided a method performed by a communication device communicatively connected to a headphone (or headphones). The method comprises outputting a first audio stream to the headphone for playback to a user of the communication device. The method also comprises, via an interface of the communication device to its surroundings, obtaining an indication that the playback of the first audio stream should be altered. The method also comprises, in response to the obtained indication, altering the output of the first audio stream; by means of a microphone of the communication device, recording a second audio stream; and outputting the second audio stream to the headphone for playback to the user.

[0005] According to another aspect of the present invention, there is provided a computer program product comprising computer-executable components for causing a communication device to perform an embodiment of the method of the present disclosure when the computer-executable components are run on processing circuitry comprised in the communication device.

[0006] According to another aspect of the present invention, there is provided a communication device comprising processing circuitry, and storage storing instructions executable by said processing circuitry whereby said communication device is operative to output a first audio stream to a headphone for playback to a user of the communication device. The communication device is also operative to, via an interface of the communication device to its surroundings, obtain an indication that the

playback of the first audio stream should be altered. The communication device is also operative to, in response to the obtained indication: alter the output of the first audio stream; by means of a microphone of the communication device, record a second audio stream; and output the second audio stream to the headphone for playback to the user.

[0007] By altering the output of the first audio stream, e.g. discontinuing it, muting it, fading it out or reducing the volume of it, and using the microphone of the communication device (also called only device herein) for capturing and playing back, effectively amplifying, ambient sound (typically voice), the user of the communication device wearing the headphone(s) may better hear the ambient sound (via the microphone) without the need to remove the headphone(s). This may be called a voice mode of the device.

[0008] It is to be noted that any feature of any of the aspects may be applied to any other aspect, wherever appropriate. Likewise, any advantage of any of the aspects may apply to any of the other aspects. Other objectives, features and advantages of the enclosed embodiments will be apparent from the following detailed disclosure, from the attached dependent claims as well as from the drawings.

[0009] Generally, all terms used in the claims are to be interpreted according to their ordinary meaning in the technical field, unless explicitly defined otherwise herein. All references to "a/an/the element, apparatus, component, means, step, etc." are to be interpreted openly as referring to at least one instance of the element, apparatus, component, means, step, etc., unless explicitly stated otherwise. The steps of any method disclosed herein do not have to be performed in the exact order disclosed, unless explicitly stated. The use of "first", "second" etc. for different features/components of the present disclosure are only intended to distinguish the features/components from other similar features/components and not to impart any order or hierarchy to the features/components.

### BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Embodiments will be described, by way of example, with reference to the accompanying drawings, in which:

Fig 1a-d schematically illustrates some embodiments of the present invention.

Fig 2a-d schematically illustrates some other embodiments of the present invention.

Fig 3 is a schematic block diagram of an embodiment of a communication device of the present invention.

Fig 4 is a schematic illustration of an embodiment of a computer program product of the present invention.

tion.

Fig 5 is a schematic flow chart of embodiments of the method of the present invention.

## DETAILED DESCRIPTION

**[0011]** Embodiments will now be described more fully hereinafter with reference to the accompanying drawings, in which certain embodiments are shown. However, other embodiments in many different forms are possible within the scope of the present disclosure. Rather, the following embodiments are provided by way of example so that this disclosure will be thorough and complete, and will fully convey the scope of the disclosure to those skilled in the art. Like numbers refer to like elements throughout the description.

**[0012]** Figures 1a-d and 2a-d illustrate steps of some embodiments of the present invention. A communication device 1 is communicatively connected to a headphone or headphones 2 worn by a person 3 who is herein called a user of the communication device. The headphone(s) comprises speakers for playing back audio, e.g. music, to the user 3, and is e.g. arranged in, on or over an ear (or both ears) of the user.

**[0013]** The communication device 1 may e.g. be configured for wired power supply or comprise a battery. The communication device may e.g. be a radio device such as any device or user equipment (UE), mobile or stationary, enabled to communicate over a radio channel in a communication network, for instance but not limited to e.g. mobile phone, smartphone, media players, or any type of consumer electronic, for instance but not limited to television, radio, tablet computer, laptop, or personal computer (PC). The device 1 is communicatively connected, wired or wirelessly, to the headphone 2 via a headphone interface 8 for outputting an audio stream to the headphone 2 which may then be played back to the user by means of its speakers. In case of a wired headphone connection, the headphone interface may e.g. comprise a receiver for a headphone connector such as a 3.5 mm connector, a Lightning connector or a USB connector (e.g. a micro USB or USB-C). In case of a wireless headphone interface, the headphone interface may comprise a radio interface e.g. for Bluetooth, Local Area Network (LAN) or Wi-Fi, or Near-Field Communication (NFC). The device may also comprise a communication interface for a data connection e.g. to the Internet, which may be wired or wireless e.g. in accordance with a LAN or Third Generation Partnership Project (3GPP) communication standard. The device 1 also comprises a microphone interface 5, e.g. a microphone, and a User Interface (UI) 4 e.g. a Graphical UI (GUI) optionally comprising a touchscreen. Additionally or alternatively, the UI 4 may comprise mechanical buttons or keys.

**[0014]** In the situation shown in figures 1a and 2a, respectively, a user 3 listens to an audio stream (herein called a first audio stream) e.g. music or an audio book,

by means of the headphone 2 connected to the device 1. The first audio stream may be of a media file, or playlist of a plurality of media files, which is e.g. stored in a storage in the device 1 or streamed by the device from an external media server (being buffered in the device) and outputted to the headphone. This may be regarded as a starting situation for embodiments of the method of the present invention. The user may e.g. be working or travelling and uses the headphone and the first audio stream avoid disturbing ambient sounds.

**[0015]** In the example situation shown in figure 1b, the user 3 decides that he/she wants to, e.g. temporarily, hear ambient sound e.g. of what another person 6 is saying. The user 3 then uses the UI 4 to input a command to the device 1 to put the device in what is herein called voice mode, whereby the device receives an indication that the playback of the first audio stream should be altered in accordance with the voice mode. If the UI comprises a touchscreen, the user may e.g. input the command by making a touch gesture or by pressing a graphical element 7 of the GUI, which graphical element is associated with the voice mode and thus provides the indication to the device. The graphical element 7 may e.g. be presented by a software (SW) application (app) or widget running in the device, e.g. integrated in a media player in the device. The user may thus easily switch to voice mode by interaction via the UI 4.

**[0016]** Additionally or alternatively, in the example situation shown in figure 2b, the switching to voice mode may be initiated automatically, without the need for the user 3 to interact with the device 1 via the UI 4. In this situation, the device 1 detects a predefined sound by means of the microphone 5. The device has been pre-programmed to associate this sound with an indication that the device should be put in voice mode. The microphone may thus be active and, when the sound is detected, the device 1 is automatically put in voice mode. The sound may e.g. be a human voice. The human voice may have a volume which is above a predetermined threshold, e.g. a static threshold or a threshold which is relative to background noise in order to qualify as an indication for putting the device in voice mode. Additionally or alternatively, the human voice may have to speak a predetermined phrase, e.g. an activation word or phrase such as a name of the user 3. By this, an other person 6, or e.g. a speaker system in a train or plane, may automatically activate the voice mode without the user 3 having to see that the other person 6 is trying to make contact or without the other person having to speak loudly to be heard over the playback of the first audio stream. This may make it easier and less awkward to make contact with the user 3. For instance, if the user 3 is working while listening via headphones 2 it may be socially awkward to approach him/her which may require either entering the field of vision of the user 3, gesturing or tapping him/her or talking really loudly in order to get noticed and start a conversation.

**[0017]** Figures 1c and 2c, respectively, shows the sit-

uation after the device 1 has been put in voice mode, e.g. following any of the situations of figures 1b or 2b. The output of the first audio stream to the headphone 2 has been altered, e.g. such that the playback by means of the speakers in the headphone has been interrupted (stopped), muted, faded out, or reduced in volume, in order to allow the user 3 to hear ambient sound. The ambient sound is obtained/recorded by means of the microphone 5 and outputted to the headphone 2 as a second audio stream for playback to the user via the speakers. The ambient sound of the second audio stream typically comprises a human voice, and in some embodiments an audio filter (typically a digital audio filter) may be used to enhance the human voice and/or reduce noise before outputting the second audio stream to the headphone. In some embodiments, visual feedback to the user that the voice mode is active may be presented by means of the GUI 4. Thus, the user may hear another person (or a speaker system) via the microphone 5 in the device 1 and the speakers in the headphone 2, without the need for removing the headphone(s).

**[0018]** The device 1 may be kept in voice mode until the device, e.g. via an interface (e.g. UI 4 and/or microphone 5), obtains an indication that the playback of the first audio stream should be restored to as before the obtaining of the indication that the playback of the first audio stream should be altered. In response to the obtained indication that the playback of the first audio stream should be restored, the device 1 may discontinue the recording and outputting of the second audio stream, and alter the output of the first audio stream such that the playback of the first audio stream is restored to as it was before the obtaining of the indication that the playback of the first audio stream should be altered (e.g. as discussed in respect of figures 1a and 2a).

**[0019]** The situations shown in figures 1d and 2d, respectively, illustrates embodiments of the present invention after the playback of the first audio stream should be restored, similar to figures 1a and 2a. Depending on how the output of the first audio stream was altered, the first audio stream output may be similarly restored, e.g. resumed (started), unmuted, faded in, or increased in volume.

**[0020]** In figure 1d, where the indication that the playback of the first audio stream should be altered was obtained via the UI 4, the indication that the first audio stream should be restored may similarly be obtained via the UI 4, e.g. by making a touch gesture or by the user pressing the same, or a different, graphical element 7 of the GUI, or by releasing pressure on said graphical element 7 if the voice mode is only active while the user is continuously pressing the graphical element.

**[0021]** In figure 2d, where the indication that the playback of the first audio stream should be altered was obtained via the microphone 5, the indication that the first audio stream should be restored may similarly be obtained via the microphone 5, e.g. by detecting that the human voice is no longer heard, or is below a predeter-

mined volume threshold, during a predetermined time period.

**[0022]** Additionally or alternatively, the indication that the first audio stream should be restored may be obtained by the expiry of a timer which was activated when the device 1 was put in the voice mode.

**[0023]** Figure 3 schematically illustrates an embodiment of a communication device 1 of the present disclosure. The device 1 comprises processing circuitry 31 e.g. a central processing unit (CPU). The processing circuitry 31 may comprise one or a plurality of processing units in the form of microprocessor(s). However, other suitable devices with computing capabilities could be comprised in the processing circuitry 31, e.g. an application specific integrated circuit (ASIC), a field programmable gate array (FPGA) or a complex programmable logic device (CPLD). The processing circuitry 31 is configured to run one or several computer program(s) or software (SW) 41 (see also figure 4) stored in a storage 32 of one or several storage unit(s) e.g. a memory. The storage unit is regarded as a computer readable means 42 (see figure 4) as discussed herein and may e.g. be in the form of a Random Access Memory (RAM), a Flash memory or other solid state memory, or a hard disk, or be a combination thereof. The processing circuitry 31 may also be configured to store data in the storage 32, as needed. The SW 41 may comprise SW for making the device perform embodiments of the method of the present disclosure. The SW 41 may e.g. comprise app SW 33 which, when run on the processing circuitry 31 forms the app 34 by means of which the device 1 may perform at least a part of embodiments of the method. The device 1 also comprises the audio output/headphone interface 8, the microphone 5 and the UI 4 as previously discussed.

**[0024]** Figure 4 illustrates an embodiment of a computer program product 40. The computer program product 40 comprises a computer readable (e.g. nonvolatile and/or non-transitory) medium 42 comprising software/computer program 41 in the form of computer-executable components. The computer program 41 may be configured to cause a device 1, e.g. as discussed herein, to perform an embodiment of the method of the present disclosure. The computer program may be run on the processing circuitry 31 of the device 1 for causing it to perform the method. The computer program product 40 may e.g. be comprised in a storage unit or memory 32 comprised in the device 1 and associated with the processing circuitry 31. Alternatively, the computer program product 40 may be, or be part of, a separate, e.g. mobile, storage means/medium, such as a computer readable disc, e.g. CD or DVD or hard disc/drive, or a solid state storage medium, e.g. a RAM or Flash memory. Further examples of the storage medium can include, but are not limited to, any type of disk including floppy disks, optical discs, DVD, CD-ROMs, microdrive, and magneto-optical disks, ROMs, RAMs, EPROMs, EEPROMs, DRAMs, VRAMs, flash memory devices, magnetic or optical cards, nanosystems (including molecular

memory ICs), or any type of media or device suitable for storing instructions and/or data. Embodiments of the present disclosure may be conveniently implemented using one or more conventional general purpose or specialized digital computer, computing device, machine, or microprocessor, including one or more processors, memory and/or computer readable storage media programmed according to the teachings of the present disclosure. Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will be apparent to those skilled in the software art.

**[0025]** Figure 5 is a schematic flow chart of some embodiments of the method of the present invention. The method is performed by a communication device 1 communicatively connected to a headphone 2. The method comprises outputting S1 a first audio stream to the headphone for playback to a user 3 of the communication device. The method also comprises, via an interface (e.g. UI 4 and/or microphone 5) to the surroundings of the device 1, obtaining S2 an indication that the playback of the first audio stream should be altered. The method also comprises, in response to the obtained S2 indication: altering S3 the output of the first audio stream; by means of the microphone 5 of the communication device, recording S4 a second audio stream; and outputting S5 the second audio stream to the headphone for playback to the user.

**[0026]** In some embodiments, the method may further comprise, via the interface 4 and/or 5, obtaining S6 an indication that the playback of the first audio stream should be restored to as before the obtaining S2 of the indication that the playback of the first audio stream should be altered. The method may also comprise, in response to the obtained S6 indication that the playback of the first audio stream should be restored: discontinuing S7 the recording S4 and outputting S5 of the second audio stream; and altering S8 the output of the first audio stream such that the playback of the first audio stream is restored.

**[0027]** In some embodiments, the first audio stream is of a media file stored in the communication device 1 or streamed from a media server.

**[0028]** In some embodiments, the interface 4 comprises a touchscreen of a GUI, and the indication is obtained S2 by detecting an input via the touchscreen corresponding to the user 3 pressing a graphical element of the GUI associated with the indication that the playback of the first audio stream should be altered.

**[0029]** In some embodiments, the interface comprises the microphone 5, and the indication is obtained S2 by via the microphone detecting sound which the communication device 1 has been preprogrammed to associate with the indication that the playback of the first audio stream should be altered. In some embodiments, the sound comprises a human voice. In some embodiments, the detected human voice sound has a volume above a predetermined threshold. In some embodiments, the de-

tected human voice sound corresponds to a predetermined phrase.

**[0030]** In some embodiments, the recording S4 of the second audio stream comprises using an audio filter to reduce noise in the second audio stream.

**[0031]** In some embodiments, the method is performed at least partly by means of a software application 34 running on the communication device 1.

**[0032]** In some embodiments, the communication device is a mobile phone, e.g. a smartphone.

**[0033]** In some embodiments, the interface of the device 1 comprises a touchscreen of a UI 4 e.g. GUI, or the interface comprises a microphone 5.

**[0034]** The present disclosure has mainly been described above with reference to a few embodiments. However, as is readily appreciated by a person skilled in the art, other embodiments than the ones disclosed above are equally possible within the scope of the present disclosure, as defined by the appended claims.

## Claims

1. A method performed by a communication device (1) communicatively connected to a headphone (2), the method comprising:

outputting (S1) a first audio stream to the headphone for playback to a user (3) of the communication device;  
via an interface (4;5) of the communication device to its surroundings, obtaining (S2) an indication that the playback of the first audio stream should be altered; and  
in response to the obtained (S2) indication:

altering (S3) the output of the first audio stream,  
by means of a microphone (5) of the communication device, recording (S4) a second audio stream, and  
outputting (S5) the second audio stream to the headphone for playback to the user.

2. The method of claim 1, wherein the first audio stream is of a media file stored in the communication device (1) or streamed from a media server.
3. The method of claim 1 or 2, wherein the interface (4) comprises a touchscreen of a GUI, and wherein the indication is obtained (S2) by detecting an input via the touchscreen corresponding to the user (3) pressing a graphical element of the GUI associated with the indication.
4. The method of claim 1 or 2, wherein the interface comprises the microphone (5), and wherein the indication is obtained (S2) by via the microphone de-

tecting sound which the communication device (1) has been preprogrammed to associate with the indication.

5. The method of claim 4, wherein the sound comprises a human voice. 5

6. The method of claim 5, wherein the detected voice sound has a volume above a predetermined threshold. 10

7. The method of claim 5 or 6, wherein the detected voice sound corresponds to a predetermined phrase. 15

8. The method of any preceding claim, wherein the recording (S4) of the second audio stream comprises using an audio filter to reduce noise in the second audio stream. 20

9. The method of any preceding claim, further comprising:

via the interface (4;5), obtaining (S6) an indication that the playback of the first audio stream should be restored to as before the obtaining (S2) of the indication that the playback of the first audio stream should be altered; and in response to the obtained (S6) indication that the playback of the first audio stream should be restored: 25 30

discontinuing (S7) the recording (S4) and outputting (S5) of the second audio stream, and altering (S8) the output of the first audio stream such that the playback of the first audio stream is restored. 35

10. The method of any preceding claim, wherein the method is performed by means of a software application (34) running on the communication device (1). 40

11. A computer program product (40) comprising computer-executable components (41) for causing a communication device (1) to perform the method of any one of claims 1-9 when the computer-executable components are run on processing circuitry (31) comprised in the communication device. 45 50

12. A communication device (1) comprising:

processing circuitry (31); and storage (32) storing instructions (41) executable by said processing circuitry whereby said communication device is operative to: 55

output a first audio stream to a headphone

(2) for playback to a user (3) of the communication device; via an interface (4;5) of the communication device to its surroundings, obtain an indication that the playback of the first audio stream should be altered; and in response to the obtained indication:

alter the output of the first audio stream, by means of a microphone (5) of the communication device, record a second audio stream, and output the second audio stream to the headphone for playback to the user.

13. The communication device of claim 12, wherein the communication device is a mobile phone, e.g. a smartphone.

14. The communication device of claim 11 or 12, wherein the interface (4) comprises a touchscreen of a GUI or wherein the interface comprises the microphone (5).

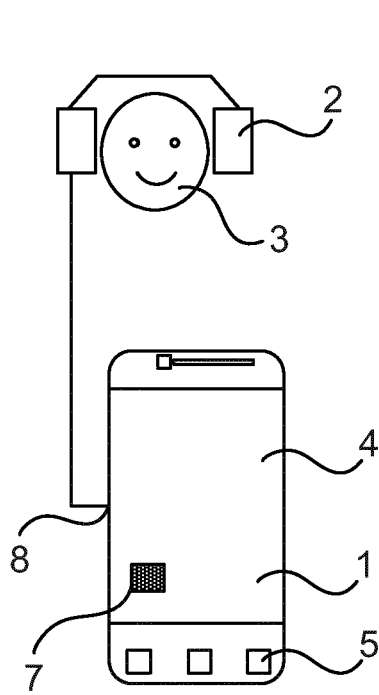


Fig. 1a

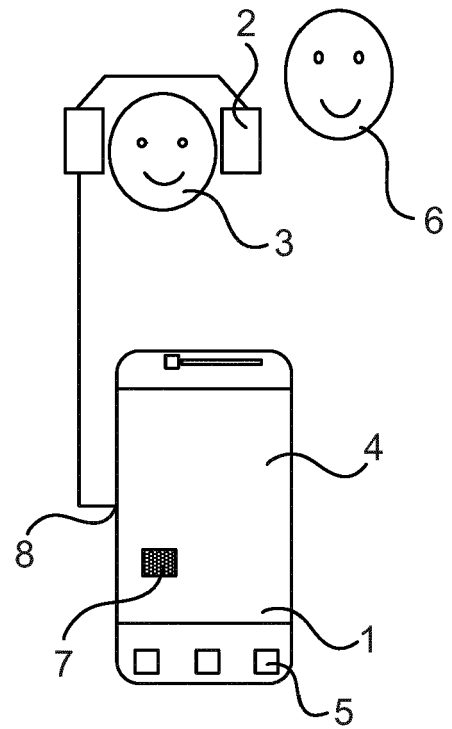


Fig. 1b

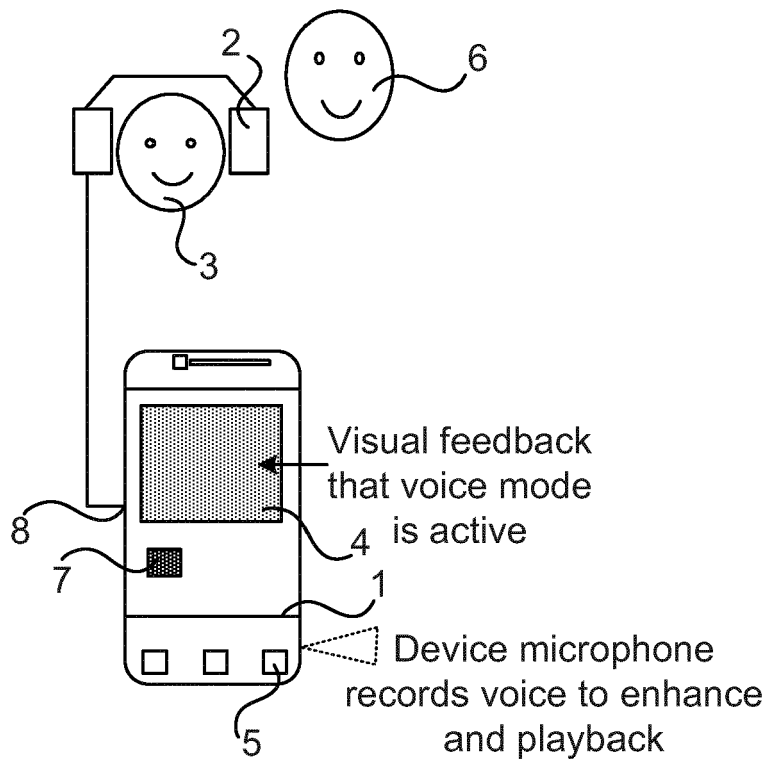


Fig. 1c

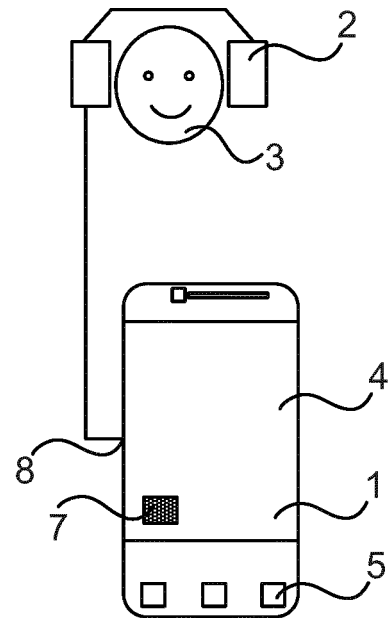


Fig. 1d

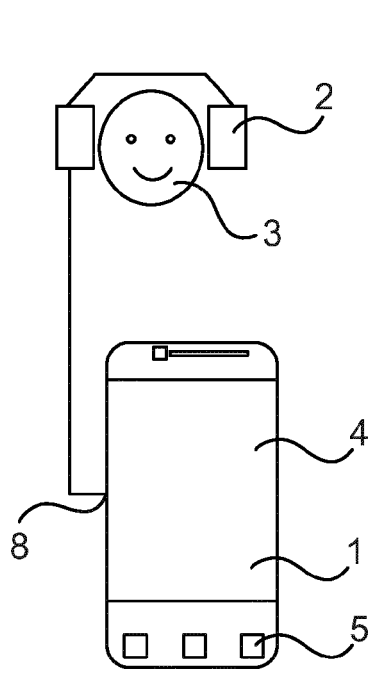


Fig. 2a

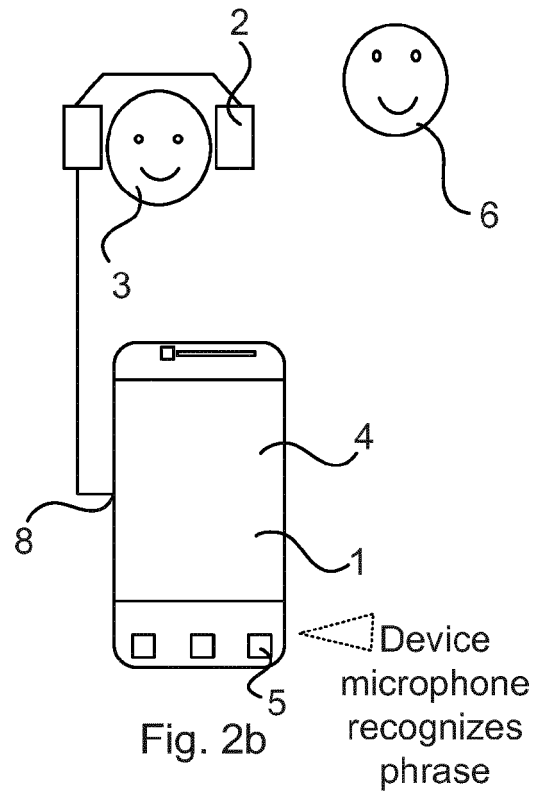


Fig. 2b

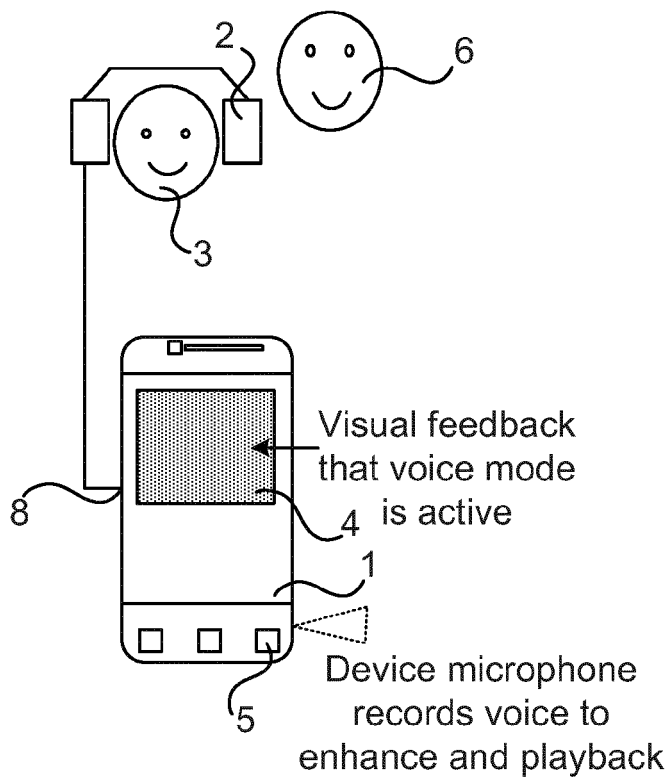


Fig. 2c

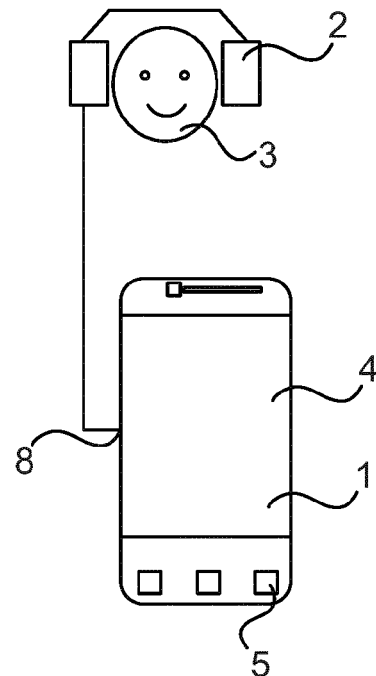


Fig. 2d



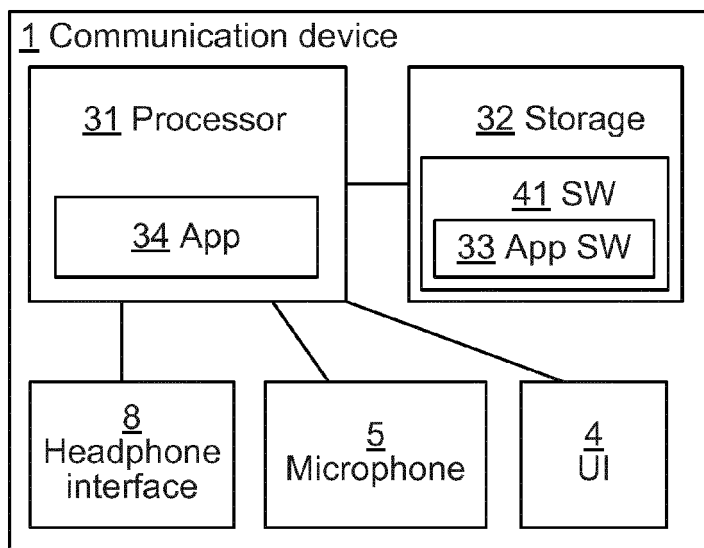


Fig. 3

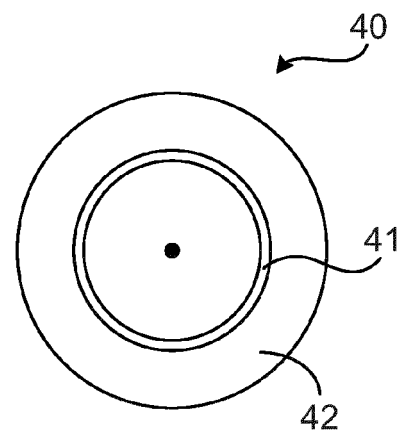


Fig. 4

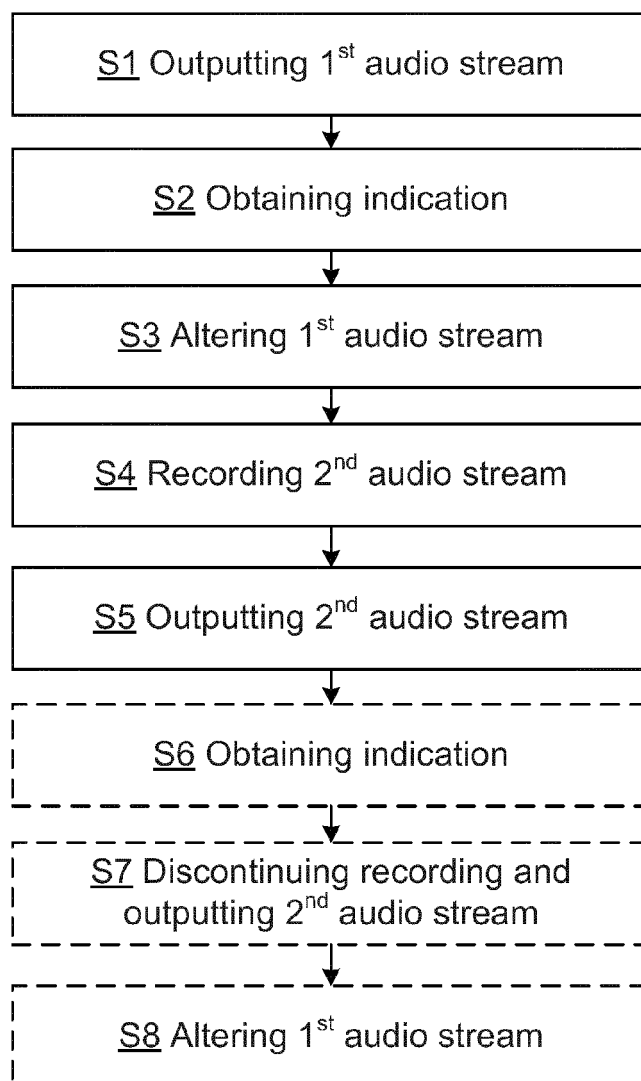


Fig. 5



## EUROPEAN SEARCH REPORT

Application Number  
EP 16 20 1116

5

10

15

20

25

30

35

40

45

50

55

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 2015/170645 A1 (DI CENSO DAVIDE [US] ET AL) 18 June 2015 (2015-06-18) * paragraphs [0025] - [0037]; figures 2,3,4a,8,9 *	1-14	INV. H04R1/10
X	US 2015/222977 A1 (ANGEL JR DONALD EARL [US]) 6 August 2015 (2015-08-06) * paragraphs [0036] - [0042]; figures 1,3,4 *	1-14	
			TECHNICAL FIELDS SEARCHED (IPC)
			H04R
The present search report has been drawn up for all claims			
Place of search <b>The Hague</b>		Date of completion of the search <b>23 March 2017</b>	Examiner <b>Vázquez del Real, S</b>
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 03/02 (P04C01)

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

EP 16 20 1116

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.

23-03-2017

10

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2015170645 A1	18-06-2015	CN 105814913 A	27-07-2016
		EP 3081011 A1	19-10-2016
		US 2015170645 A1	18-06-2015
		WO 2015088984 A1	18-06-2015
-----			
US 2015222977 A1	06-08-2015	NONE	
-----			

15

20

25

30

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