



- (51) International Patent Classification:
G06F 3/01 (2006.01) *G06F 3/16* (2006.01)
- (21) International Application Number:
PCT/US2014/065621
- (22) International Filing Date:
14 November 2014 (14.11.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
14/140,293 24 December 2013 (24.12.2013) US
- (71) Applicant (for all designated States except US): **INTEL CORPORATION** [US/US]; 2200 Mission College Boulevard, Santa Clara, California 95054 (US).
- (72) Inventors; and
- (71) Applicants (for US only): **DADU, Saurabh** [US/US]; 15265 SW Firtree Drive, Tigard, Oregon 97223 (US). **KRISHNAMURTHY, Lakshman** [US/US]; 3640 NW Talamore Tr, Portland, Oregon 97229 (US). **SHAH, Saurin** [US/US]; 9241 NW Murdock Street, Portland, Oregon 97229 (US). **THARAPPEL, Francis M.** [US/US]; 5315 NW 135th Ave, Portland, Oregon 97229 (US). **KAR, Swarnendu** [IN/US]; 1314 NE Alex Way, Apt 186, Hillsboro, Oregon 97124 (US).
- (74) Agent: **JORDAN, B. Delano**; Jordan IP Law, LLC, c/o CPA Global, P.O. Box 52050, Minneapolis, Minnesota 55402 (US).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report (Art. 21(3))

(54) Title: PERFORMING AUTOMATED VOICE OPERATIONS BASED ON SENSOR DATA REFLECTING SOUND VIBRATION CONDITIONS AND MOTION CONDITIONS

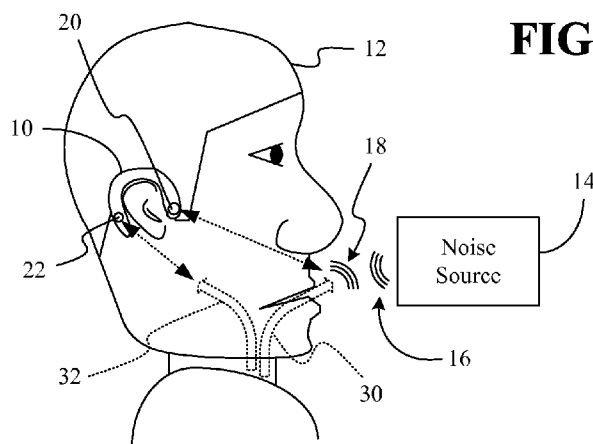
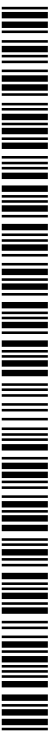


FIG. 1

(57) Abstract: Systems and methods may provide for determining a sound vibration condition of an ambient environment of a wearable device and determining a motion condition of the wearable device. In addition, one or more automated voice operations may be performed based at least in part on the sound vibration condition and the motion condition. In one example, two or more signals corresponding to the sound vibration condition and the motion condition may be combined.



PERFORMING AUTOMATED VOICE OPERATIONS BASED ON SENSOR DATA REFLECTING SOUND VIBRATION CONDITIONS AND MOTION CONDITIONS

TECHNICAL FIELD

5 Embodiments generally relate to wearable devices. More particularly, embodiments relate to performing automated voice operations based on sensor data reflecting sound vibration conditions and motion conditions associated with wearable devices.

BACKGROUND

10 Wearable devices such as Bluetooth (e.g., Institute of Electrical and Electronics Engineers/IEEE 802.15.1-2005, Wireless Personal Area Networks) headsets may be used to capture voice input from the wearer of the device in order to place and/or conduct telephone calls, control media playback, and so forth. While conventional wearable devices may be suitable under certain circumstances, there remains considerable room for improvement. For
15 example, the accuracy of voice controlled wearable devices may be limited with respect to the detection and/or evaluation of voice input, particularly in ambient environments having a relatively high level of audible noise. Moreover, low accuracy may lead to a poor user experience, higher power consumption and shorter battery life.

BRIEF DESCRIPTION OF THE DRAWINGS

20 The various advantages of the embodiments will become apparent to one skilled in the art by reading the following specification and appended claims, and by referencing the following drawings, in which:

 FIG. 1 is an illustration of an example of a wearable device in an ambient environment
25 containing noise according to an embodiment;

 FIG. 2 is a block diagram of an example of a condition detection model for automated voice operations according to an embodiment;

 FIG. 3 is a block diagram of an example of an automated voice operation pipeline according to an embodiment;

30 FIG. 4 is a block diagram of an example of a signal fusion solution according to an embodiment;

 FIG. 5 is a flowchart of an example of a method of controlling automated voice operations according to an embodiment; and

 FIG. 6 is a block diagram of an example of a wearable device according to an embodiment.

35

DESCRIPTION OF EMBODIMENTS

Turning now to FIG. 1, a wearable device 10 is shown, wherein the device 10 is worn by a user 12 who is speaking. More particularly, the vocal cords, voice box and/or larynx of the user 12 may move and/or vibrate in a manner that has a mouth and air related transfer function 30 resulting in audible sound waves and/or vibrations 18 in the air near the mouth of the user 12, wherein those sound vibrations 18 may be characterized as speech. In the illustrated example, the user 12 is speaking in an ambient environment containing a noise source 14 such as, for example, one or more other people, radios, televisions, machinery, etc. Thus, audible noise 16 from the source 14 may interfere/mix with the sound vibrations 18 while the user 12 is speaking. As a result, if an air-based (e.g., acoustic) microphone 20 of the wearable device 10 is configured to capture the sound vibrations 18 associated with the speech of the user 12, the audible noise 16 may degrade one or more audio signals generated by the air-based microphone 20.

The illustrated wearable device 10 also includes a contact-based component 22 to detect and/or determine additional conditions of the wearable device 10 to supplement the signals obtained from the air-based microphone 20. The contact-based component 22 may rely upon physical contact with the user 12 to detect and/or determine the additional conditions. In this regard, the jawbone, vocal cords, voice box and/or larynx of the user 12 may move and/or vibrate in a manner that has a head and bone conduction related transfer function 32 resulting in movement of the wearable device 10.

Thus, the contact-based component 22 may include, for example, an inertial sensor (e.g., gyroscope, accelerometer, etc.) and/or bone conduction microphone, that generates one or more signals to characterize the motion of the user's jawbone/face while speaking, physical vibrations of the user's vocal chords, voice box/larynx, etc., while speaking, and so forth. As will be discussed in greater detail, using the signals from the contact-based component 22 to supplement the signals obtained from the air-based microphone 20 may enable the wearable device 10 to achieve enhanced accuracy, a better user experience, lower power consumption and/or longer battery life relative to conventional wearable devices.

FIG. 2 shows a condition detection model 24 for automated voice operations 26. In the illustrated example, a speech source 28 such as the larynx of a user may have one or more mouth and air related transfer functions 30 as well as one or more head and bone conduction transfer functions 32. The result of the mouth and air related transfer functions 30 may mix with ambient noise 16 to form a result that is captured by the air-based microphone 20 (e.g., sound vibration condition detector). In addition, the result of the head and bone conduction transfer functions 32 may be captured by the contact-based component 22 (e.g., motion condition detector). The signals from the microphone 20 and the contact-based component 22 may be used to control the

automated voice operations 26. Of particular note is that the result of the head and bone conduction transfer functions 32 is not impacted by the ambient noise 16, in the illustrated example. The contact-based component 22 may therefore be a good complement to the air-based microphone 20, particularly in environments having relatively high levels of the ambient noise 16.

Although the dynamic range of the contact-based component 22 may be narrower than the dynamic range of the air-based microphone 20, the signals from the contact-based component 22 may still be useful in controlling the automated voice operations 26, particularly when used in combination with the signals from the air-based microphone 20. Indeed, in the case of an earpiece that is in direct contact of the user's ear, the signals from the contact-based component 22 may be highly reliable (e.g., much less spurious compared to a phone or other wearable devices).

Turning now to FIG. 3, a pipeline 34 of automated voice operations 26 is shown. In the illustrated example, a voice activity detection (VAD) stage 36 may determine whether an incoming audio signal contains voice content. Thus, the VAD stage 36 may be particularly useful if the wearable device is operated in an "always listening" mode in which the user may initiate voice interaction sessions purely with voice commands (e.g., hands-free). The VAD stage 36 may generally be used to reduce power consumption because further processing of the incoming audio signal may be avoided until a user actually speaks. The pipeline 34 may also include a keyword detection (KD) stage 38, which determines whether one or more key phrases (e.g., "Hello Jarvis") are contained in signals having voice content.

Additionally, a speaker verification (SV) stage 40 of the pipeline 34 may confirm whether an identified key phrase was spoken by a known/enrolled user (e.g., rather than a nearby individual, radio, television, etc.). The pipeline 34 may also include a voice interaction (VI) phase 42 that involves automated speech recognition (ASR) and personal assistant functionality. The personal assistant functionality might include, for example, searching for restaurants, answering questions, making reservations, etc., via a human-to-device interface (e.g., imitating a human-to-human interface). The voice interaction phase 42 may be conducted by a separate device (e.g., paired mobile phone, smart tablet, etc.), wherein the wearable device may sample the audio signal, pre-process the sampled audio signal and send it to the separate device.

In order to reduce the occurrence of false positives and/or false negatives at the various stages of the pipeline 34, a contact-based component such as the component 22 (FIG. 1), may be used to supplement a traditional air-based microphone such as the microphone 20 (FIG. 1). For example, if an inertial sensor is used as a contact-based component at the VAD stage 36, and the air-based microphone detects the presence of voice content in an incoming audio signal, the

audio signal might be passed to the VT stage 38 only if the inertial sensor detects a sufficient level of motion (e.g., due to movement of the user's jaw). Similarly, if a bone conduction microphone is used as a contact-based component at the VAD stage 36, and the air-based microphone detects the presence of voice content in an incoming audio signal, the audio signal
5 may be passed to the VT stage 38 only if the bone conduction microphone also detects activity (e.g., due to the presence of vibratory motion of the bone structure in the head of the user).

In addition, if a bone conduction microphone is used as a contact-based component at the VT stage 38, and the air-based microphone detects that one or more key phrases are contained in an audio signal having voice content, the SV stage 40 may be activated only if the bone
10 conduction microphone also detects that one or more key phrases are contained in the audio signal (e.g., due to a particular vibratory motion signature being present in the bone structure in the head of the user). Moreover, if a bone conduction microphone is used as a contact-based component at the SV stage 40, and the air-based microphone detects that an identified key phrase was spoken by a known/enrolled user, the voice interaction stage 42 may be activated only if the
15 bone conduction microphone also detects that the key phrase was spoken by the known/enrolled user (e.g., due to the particular vibratory motion signature being associated with the known/enrolled user). Simply put, source disambiguation may be achieved at the VAD stage 36, the VT stage 38 and/or the speaker verification stage 40 with a high level of accuracy.

In addition to using the benefits of source disambiguation, the word accuracy rates for the
20 voice interaction stage 42 may be further increased through multi-modal audio fusion. For example, since the modalities of both the air-based microphone and the bone conduction microphone may observe the same underlying phenomena (i.e., the user speaking), signals from the air-based microphone sensor and the bone conduction microphone may be combined/fused to form a composite signal that may be more accurate than either signal by itself. Techniques such
25 as, for example, traditional and/or adaptive filters may be used to estimate relative transfer functions such as the transfer functions 30, 32 (FIG. 1) and sensor uncertainties. The uncertainty characteristics of the two microphones may be used to derive optimal weights for the fusion.

Simply put, performing the automated voice operations based at least in part on the motion condition may increase accuracy and reduce power consumption in the wearable device. For
30 example, each stage of the pipeline 34 may be triggered by the previous stage and executed in a progressively higher power domain (e.g., either in different hardware components consuming more power, or the same hardware running at a higher frequency). Thus, the VAD stage 36 may be a simple hardware chip using very low power, where the KD and VT stages 38, 40, respectively, may operate on a small, low power microcontroller. The voice interaction stage
35 42, on the other hand, may be implemented primarily on a paired phone that consumes much

higher power than the wearable device.

FIG. 4 shows a signal fusion module 35 that may be used to combine a first signal 37 from an air-based microphone with a second signal 39 from a contact-based component. In the illustrated example, various filters 41 (e.g., high-pass, low-pass, adaptive) are applied to the signals 37, 39. In addition, noise estimation and error variance techniques may be used to optimize the filtered results and obtain fused audio 43 that may be provided to speaker verification and/or voice interaction stages of an automated voice operation pipeline.

Turning now to FIG. 5, a method 44 of controlling automated voice operations is shown. The method 44 may be implemented as a set of logic instructions stored in a machine- or computer-readable storage medium such as random access memory (RAM), read only memory (ROM), programmable ROM (PROM), firmware, flash memory, etc., in configurable logic such as, for example, programmable logic arrays (PLAs), field programmable gate arrays (FPGAs), complex programmable logic devices (CPLDs), in fixed-functionality logic hardware using circuit technology such as, for example, application specific integrated circuit (ASIC), complementary metal oxide semiconductor (CMOS) or transistor-transistor logic (TTL) technology, or any combination thereof.

Illustrated processing block 46 may determine a sound vibration condition of an ambient environment of a wearable device. As already noted, the sound vibration condition may be determined based on a signal associated with an air-based microphone of the wearable device. Additionally, a motion condition of the wearable device may be determined at block 48, wherein the motion condition may be determined based on a signal associated with an inertial sensor of the wearable device, a signal associated with a bone conduction microphone of the wearable device, and so forth.

Block 50 may perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition. In the case of an inertial sensor, at least one of the one or more automated voice operations may include a voice activity detection operation. In the case of a bone conduction microphone, at least one of the one or more automated voice operations might include a voice activity detection operation, a keyword detection operation, a speaker verification operation, a voice interaction operation, and so forth. Block 50 may also involve combining two or more signals corresponding to the sound vibration condition and the motion condition into a composite signal based on, for example, known transfer functions, sensor uncertainties, and so forth. Moreover, combining the signals may involve the use of one or more filters and/or relative weights.

FIG. 6 shows a wearable device 52. The wearable device 52 may include an earpiece (e.g., Bluetooth earpiece), headwear with a chinstrap (e.g., football helmet, hardhat, etc.), or other

device coming into contact with the head, neck, ear and/or face of a user. In the illustrated example, the wearable device 52 includes a battery 54 to supply power to the wearable device 52 and a processor 56 having an integrated memory controller (IMC) 58, which may communicate with system memory 60. The system memory 60 may include, for example, dynamic random access memory (DRAM).

The illustrated wearable device 52 also includes a input output (IO) module 62 that functions as a host device and may communicate with, for example, an inertial sensor 64, a bone conduction microphone 66, an air-based microphone 68, a speaker 74 and mass storage 70 (e.g., flash memory, etc.). The processor 56 may execute an operating system (OS, not shown), one or more communication applications (e.g., near field communication/NFC, Bluetooth, etc., not shown), and so forth.

The illustrated processor 56 also executes logic 72 that is configured to determine a sound vibration condition of an ambient environment containing the wearable device 52 based at least in part on a signal associated with the air-based microphone 68, determine a motion condition of the wearable device based on a signal from the inertial sensor 64 and/or bone conduction microphone 66, and perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition. In one example, the logic 72 combines/fuses two or more signals corresponding to the sound vibration condition and the motion condition into a composite signal. Thus, the illustrated logic 72 may implement one or more aspects of the method 44 (FIG. 5), already discussed. Moreover, the wearable device 52 may be readily substituted for the wearable device 10 (FIG. 1).

The logic 72 may alternatively be implemented external to the processor 56. Additionally, the processor 56 and the IO module 62 may be implemented together on the same semiconductor die as a system on chip (SoC).

Additional Notes and Examples:

Example 1 may include a wearable device comprising an air-based microphone and logic, implemented at least partly in fixed-functionality hardware, to determine a sound vibration condition of an ambient environment of the wearable device based at least in part on a signal associated with the air-based microphone, determine a motion condition of the wearable device, and perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition.

Example 2 may include the wearable device of Example 1, wherein the logic is to use a combination of two or more signals corresponding to the sound vibration condition and the motion condition to perform the one or more automated voice operations.

Example 3 may include the wearable device of any one of Examples 1 or 2, further

including an inertial sensor, wherein the motion condition is to be determined based at least in part on a signal associated with the inertial sensor.

Example 4 may include the wearable device of Example 3, wherein at least one of the one or more automated voice operations is to include a voice activity detection operation.

5 Example 5 may include the wearable device of any one of Examples 1 or 2, further including a bone conduction microphone, wherein the motion condition is to be determined based at least in part on a signal associated with the bone conduction microphone.

10 Example 6 may include the wearable device of Example 5, wherein at least one of the one or more automated voice operations is to include one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

15 Example 7 may include a method of controlling automated voice operations, comprising determining a sound vibration condition of an ambient environment of a wearable device, determining a motion condition of the wearable device and performing one or more automated voice operation based at least in part on the sound vibration condition and the motion condition.

Example 8 may include the method of Example 7, further including using a combination of two or more signals corresponding to the sound vibration condition and the motion condition.

20 Example 9 may include the method of Example 7, wherein the sound vibration condition is determined based at least in part on a signal associated with an air-based microphone of the wearable device.

Example 10 may include the method of Example 7, wherein the motion condition is determined based at least in part on a signal associated with an inertial sensor of the wearable device and at least one of the one or more automated voice operations includes a voice activity detection operation.

25 Example 11 may include the method of any one of Examples 7 to 10, wherein the motion condition is determined based at least in part on a signal associated with a bone conduction microphone of the wearable device, wherein at least one of the one or more automated voice operations includes one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

30 Example 12 may include at least one computer readable storage medium comprising a set of instructions which, if executed by a wearable device, cause the wearable device to determine a sound vibration condition of an ambient environment of the wearable device, determine a motion condition of the wearable device, and perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition.

35 Example 13 may include the at least one computer readable storage medium of Example

12, wherein the instructions, if executed, cause the wearable device to use a combination of two or more signals corresponding to the sound vibration condition and the motion condition.

Example 14 may include the at least one computer readable storage medium of Example 12, wherein the sound vibration condition is to be determined based at least in part on a signal associated with an air-based microphone of the wearable device.

Example 15 may include the at least one computer readable storage medium of any one of Examples 12 to 14, wherein the motion condition is to be determined based at least in part on a signal associated with an inertial sensor of the wearable device.

Example 16 may include the at least one computer readable storage medium of Example 15, wherein at least one of the one or more automated voice operations is to include a voice activity detection operation.

Example 17 may include the at least one computer readable storage medium of any one of Examples 12 to 14, wherein the motion condition is to be determined based at least in part on a signal associated with a bone conduction microphone of the wearable device.

Example 18 may include the at least one computer readable storage medium of Example 17, wherein at least one of the one or more automated voice operations is to include one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

Example 19 may include an apparatus to manage automated voice operations, comprising logic, implemented at least partly in fixed-functionality hardware, to determine a sound vibration condition of an ambient environment of a wearable device, determine a motion condition of the wearable device, and perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition.

Example 20 may include the apparatus of Example 19, wherein the logic is to use a combination of two or more signals corresponding to the sound vibration condition and the motion condition to perform the one or more automated voice operations.

Example 21 may include the apparatus of Example 19, wherein the sound vibration condition is to be determined based at least in part on a signal associated with an air-based microphone of the wearable device.

Example 22 may include the apparatus of any one of Examples 19 to 21, wherein the motion condition is to be determined based at least in part on a signal associated with an inertial sensor of the wearable device.

Example 23 may include the apparatus of Example 22, wherein at least one of the one or more automated voice operations is to include a voice activity detection operation.

Example 24 may include the apparatus of any one of Examples 19 to 21, wherein the

motion condition is to be determined based at least in part on a signal associated with a bone conduction microphone of the wearable device.

Example 25 may include the apparatus of Example 24, wherein at least one of the one or more automated voice operations is to include one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

Example 26 may include an apparatus to manage voice operations, comprising means for performing the method of any one of Examples 7 to 11.

Thus, techniques described herein may significantly improve voice activity detection operations even in environments with high levels of ambient noise (e.g., in the presence of group conversations, public areas, etc.). In addition, false acceptance rates and false rejection rates may be reduced in keyword detection operations and speaker verification operations. Moreover, improved security may be achieved with respect to speaker verification. Additionally, the word accuracy rates of automated speech recognition operations may be improved, which may in turn lead to improved dialogues with virtual personal assistants. Techniques may also enhance the user experience and reduce power consumption due to a reduction in errors and user retries.

Embodiments are applicable for use with all types of semiconductor integrated circuit ("IC") chips. Examples of these IC chips include but are not limited to processors, controllers, chipset components, programmable logic arrays (PLAs), memory chips, network chips, systems on chip (SoCs), SSD/NAND controller ASICs, and the like. In addition, in some of the drawings, signal conductor lines are represented with lines. Some may be different, to indicate more constituent signal paths, have a number label, to indicate a number of constituent signal paths, and/or have arrows at one or more ends, to indicate primary information flow direction. This, however, should not be construed in a limiting manner. Rather, such added detail may be used in connection with one or more exemplary embodiments to facilitate easier understanding of a circuit. Any represented signal lines, whether or not having additional information, may actually comprise one or more signals that may travel in multiple directions and may be implemented with any suitable type of signal scheme, e.g., digital or analog lines implemented with differential pairs, optical fiber lines, and/or single-ended lines.

Example sizes/models/values/ranges may have been given, although embodiments are not limited to the same. As manufacturing techniques (e.g., photolithography) mature over time, it is expected that devices of smaller size could be manufactured. In addition, well known power/ground connections to IC chips and other components may or may not be shown within the figures, for simplicity of illustration and discussion, and so as not to obscure certain aspects of the embodiments. Further, arrangements may be shown in block diagram form in order to

avoid obscuring embodiments, and also in view of the fact that specifics with respect to implementation of such block diagram arrangements are highly dependent upon the platform within which the embodiment is to be implemented, i.e., such specifics should be well within purview of one skilled in the art. Where specific details (e.g., circuits) are set forth in order to describe example embodiments, it should be apparent to one skilled in the art that embodiments can be practiced without, or with variation of, these specific details. The description is thus to be regarded as illustrative instead of limiting.

The term “coupled” may be used herein to refer to any type of relationship, direct or indirect, between the components in question, and may apply to electrical, mechanical, fluid, optical, electromagnetic, electromechanical or other connections. In addition, the terms “first”, “second”, etc. may be used herein only to facilitate discussion, and carry no particular temporal or chronological significance unless otherwise indicated.

Those skilled in the art will appreciate from the foregoing description that the broad techniques of the embodiments can be implemented in a variety of forms. Therefore, while the embodiments have been described in connection with particular examples thereof, the true scope of the embodiments should not be so limited since other modifications will become apparent to the skilled practitioner upon a study of the drawings, specification, and following claims.

CLAIMS

We claim:

1. A wearable device comprising:

an air-based microphone; and

logic, implemented at least partly in fixed-functionality hardware, to:

determine a sound vibration condition of an ambient environment of the wearable device based at least in part on a signal associated with the air-based microphone,

determine a motion condition of the wearable device, and

perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition.

2. The wearable device of claim 1, wherein the logic is to use a combination of two or more signals corresponding to the sound vibration condition and the motion condition to perform the one or more automated voice operations.

3. The wearable device of any one of claims 1 or 2, further including an inertial sensor, wherein the motion condition is to be determined based at least in part on a signal associated with the inertial sensor.

4. The wearable device of claim 3, wherein at least one of the one or more automated voice operations is to include a voice activity detection operation.

5. The wearable device of any one of claims 1 or 2, further including a bone conduction microphone, wherein the motion condition is to be determined based at least in part on a signal associated with the bone conduction microphone.

6. The wearable device of claim 5, wherein at least one of the one or more automated voice operations is to include one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

7. A method of operating a wearable device, comprising:

determining a sound vibration condition of an ambient environment of the wearable device;

determining a motion condition of the wearable device; and

performing one or more automated voice operations based at least in part on the sound

vibration condition and the motion condition.

8. The method of claim 7, further including using a combination of two or more signals corresponding to the sound vibration condition and the motion condition to perform the one or more automated voice operations.

9. The method of claim 7, wherein the sound vibration condition is determined based at least in part on a signal associated with an air-based microphone of the wearable device.

10. The method of claim 7, wherein the motion condition is determined based at least in part on a signal associated with an inertial sensor of the wearable device and at least one of the one or more automated voice operations includes a voice activity detection operation.

11. The method of any one of claims 7 to 10, wherein the motion condition is determined based at least in part on a signal associated with a bone conduction microphone of the wearable device, wherein at least one of the one or more automated voice operations includes one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

12. At least one computer readable storage medium comprising a set of instructions which, if executed by a wearable device, cause the wearable device to:
determine a sound vibration condition of an ambient environment of the wearable device;
determine a motion condition of the wearable device; and
perform one or more automated voice operations based at least in part on the sound vibration condition and the motion condition.

13. The at least one computer readable storage medium of claim 12, wherein the instructions, if executed, cause the wearable device to use a combination of two or more signals corresponding to the sound vibration condition and the motion condition to perform the one or more automated voice operations.

14. The at least one computer readable storage medium of claim 12, wherein the sound vibration condition is to be determined based at least in part on a signal associated with an air-based microphone of the wearable device.

15. The at least one computer readable storage medium of any one of claims 12 to 14, wherein the motion condition is to be determined based at least in part on a signal associated with an inertial sensor of the wearable device.

5 16. The at least one computer readable storage medium of claim 15, wherein at least one of the one or more automated voice operations is to include a voice activity detection operation.

10 17. The at least one computer readable storage medium of any one of claims 12 to 14, wherein the motion condition is to be determined based at least in part on a signal associated with a bone conduction microphone of the wearable device.

15 18. The at least one computer readable storage medium of claim 17, wherein at least one of the one or more automated voice operations is to include one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

19. An apparatus to manage voice operations, comprising:
logic, implemented at least partly in fixed-functionality hardware, to:
20 determine a sound vibration condition of an ambient environment of a wearable device,
determine a motion condition of the wearable device, and
perform one or more automated voice operations based at least in part on the
sound vibration condition and the motion condition.

25 20. The apparatus of claim 19, wherein the logic is to use a combination of two or more signals corresponding to the sound vibration condition and the motion condition to perform the one or more automated voice operations.

30 21. The apparatus of claim 19, wherein the sound vibration condition is to be determined based at least in part on a signal associated with an air-based microphone of the wearable device.

35 22. The apparatus of any one of claims 19 to 21, wherein the motion condition is to be determined based at least in part on a signal associated with an inertial sensor of the wearable

device.

23. The apparatus of claim 22, wherein at least one of the one or more automated voice operations is to include a voice activity detection operation.

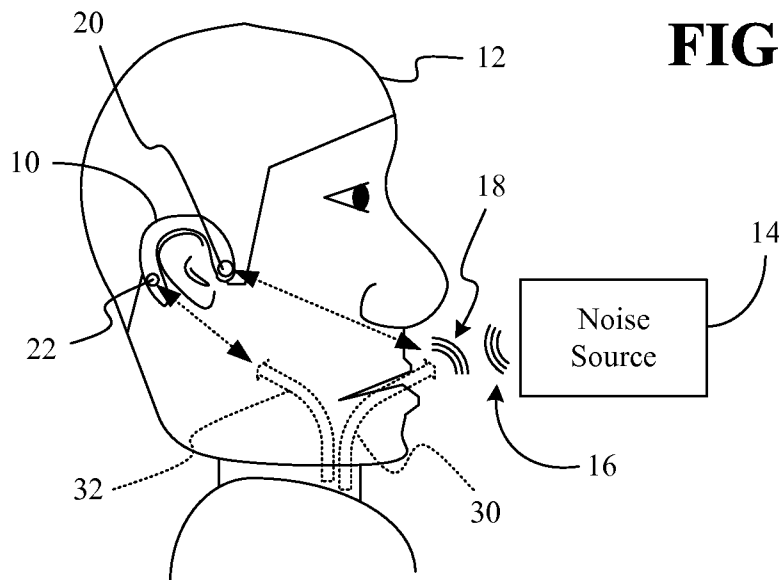
5

24. The apparatus of any one of claims 19 to 21, wherein the motion condition is to be determined based at least in part on a signal associated with a bone conduction microphone of the wearable device.

10

25. The apparatus of claim 24, wherein at least one of the one or more automated voice operations is to include one or more of a voice activity detection operation, a keyword detection operation, a speaker verification operation or a voice interaction operation.

FIG. 1



24

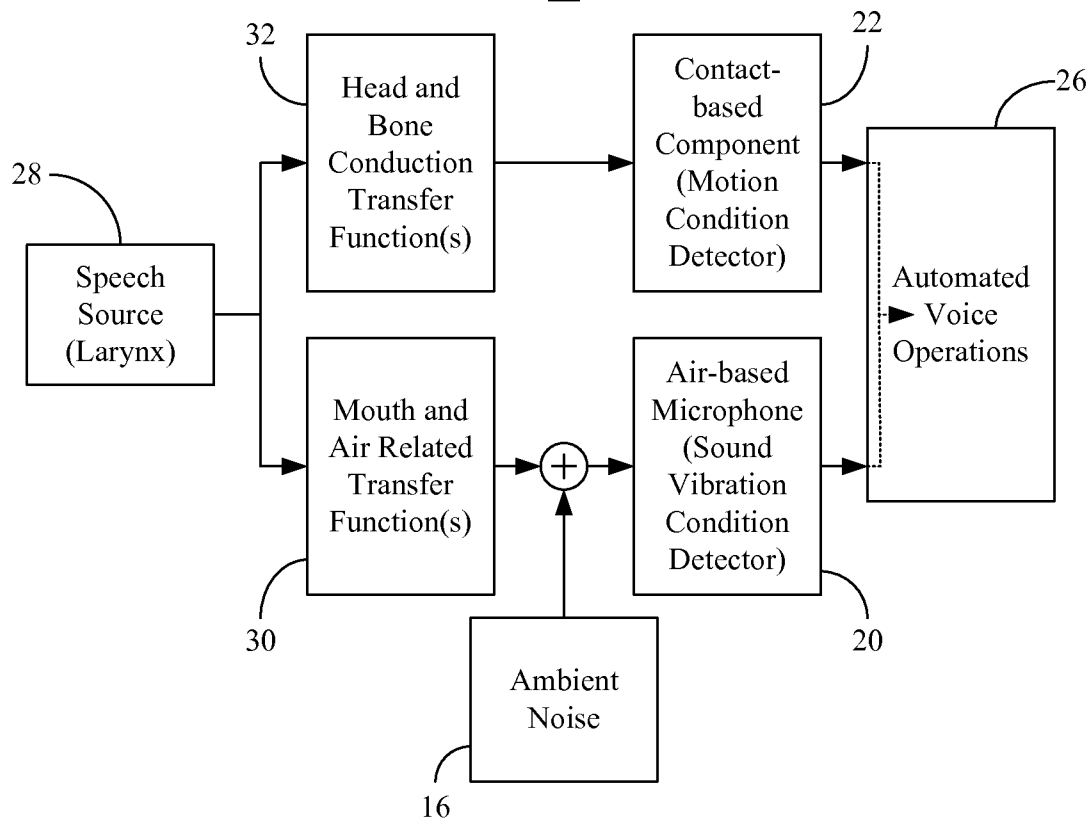
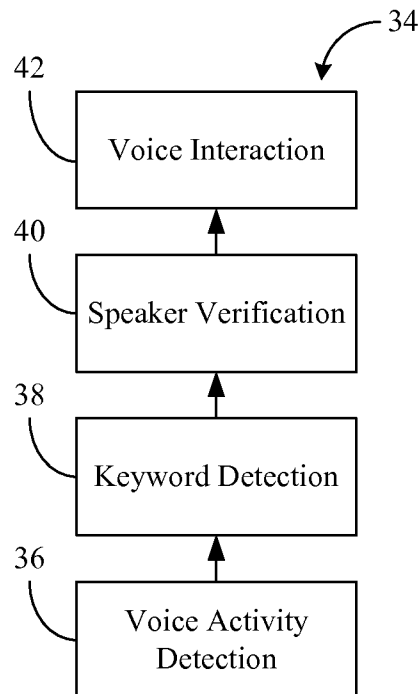
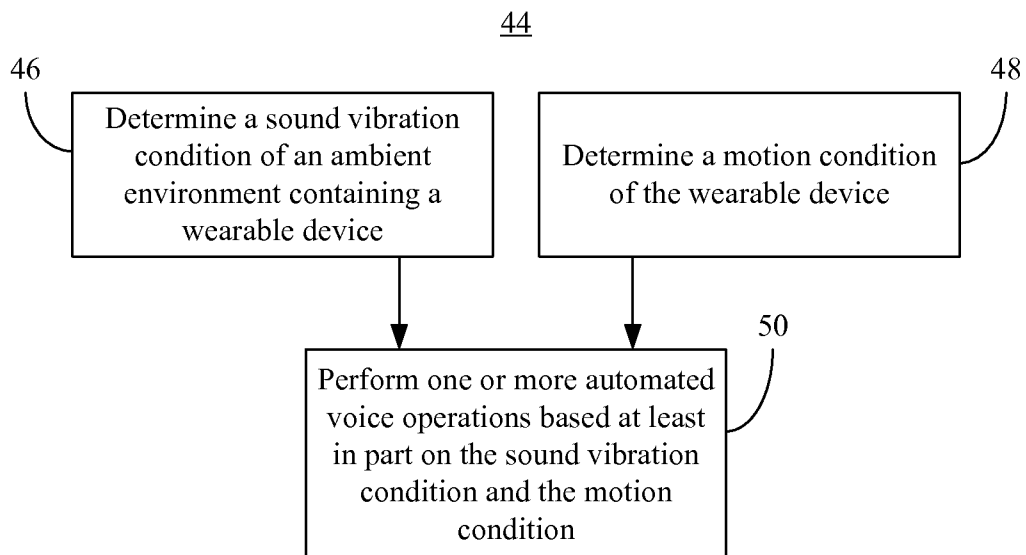
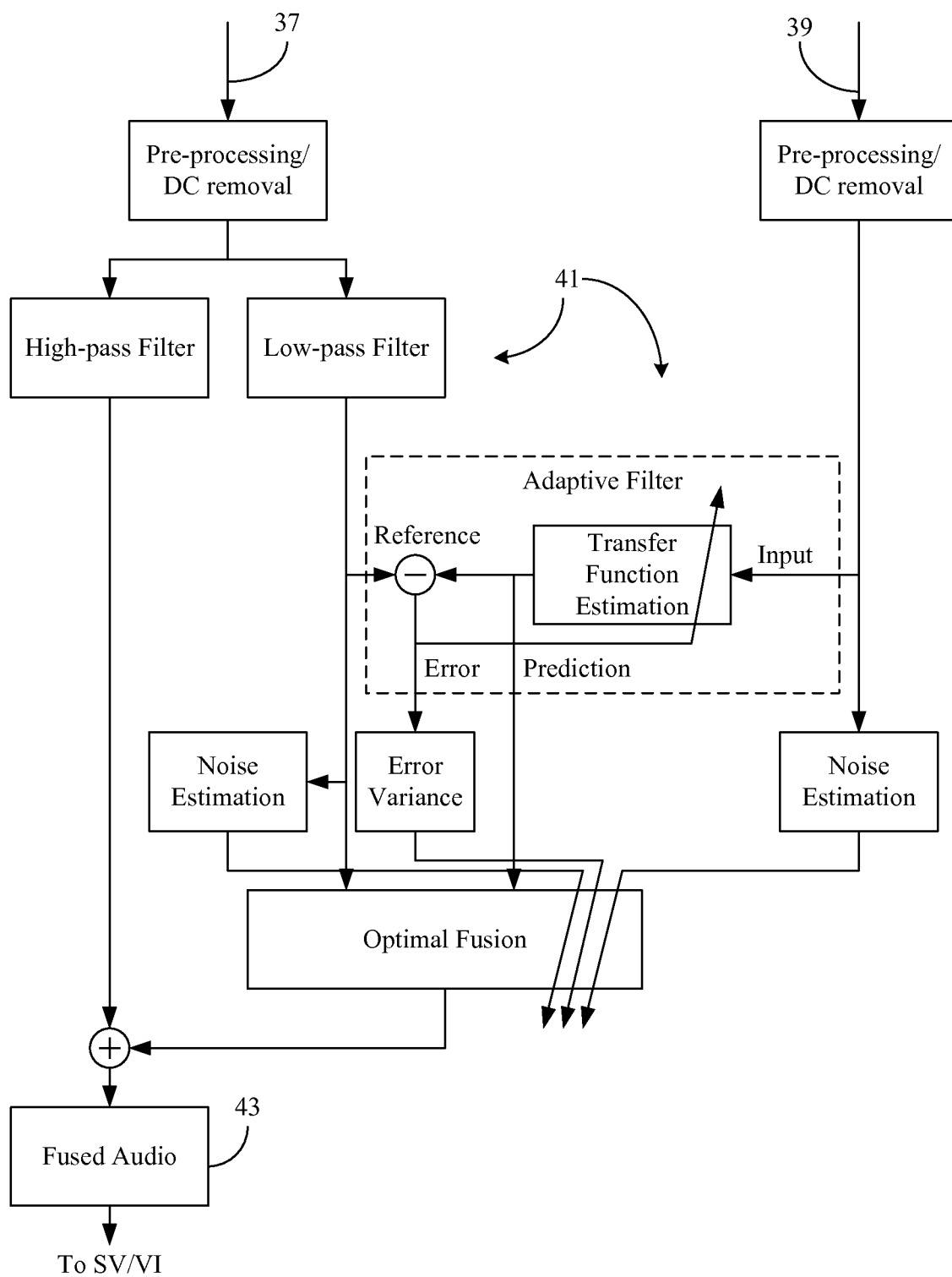


FIG. 2

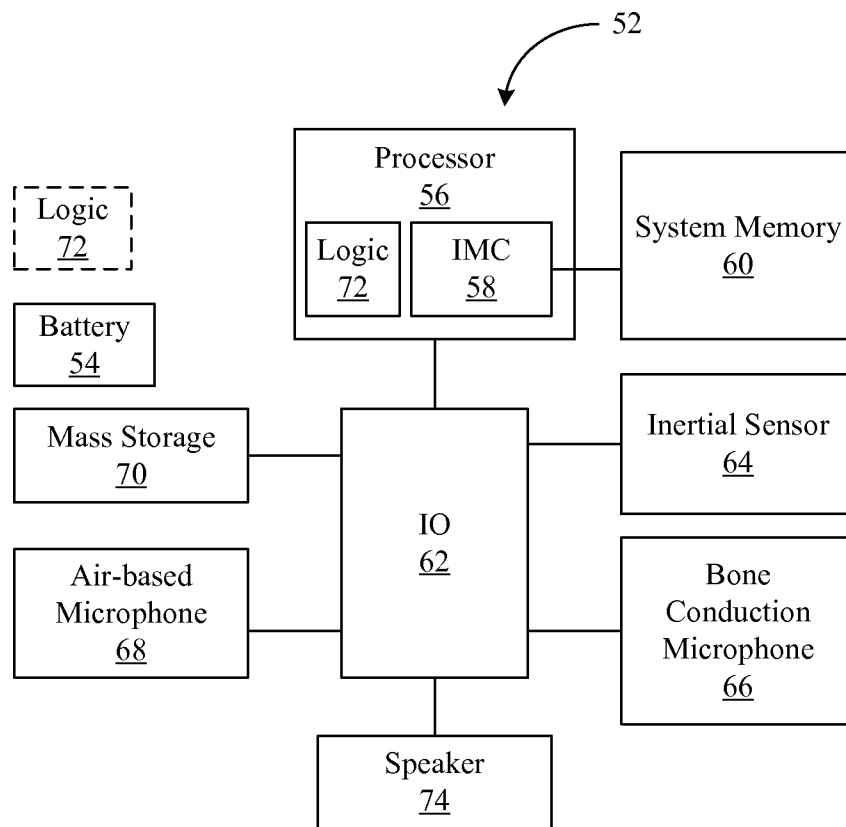
2/4

**FIG. 3****FIG. 5**

3/4

35**FIG. 4**

4/4

**FIG. 6**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/065621**A. CLASSIFICATION OF SUBJECT MATTER****G06F 3/01(2006.01)i, G06F 3/16(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 3/01; H04R 1/02; G09B 21/00; G01N 3/08; G10L 21/06; G10L 15/26; G01N 21/35; A61B 5/00; H04B 1/38; G10L 21/00; G06F 3/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & Keywords: wearable, microhpone, voice command, motion, ambient, noise, and similar terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012-0035924 A1 (JOHN NICHOLAS JITKOFF et al.) 09 February 2012 See paragraphs [0005], [0020]-[0048], [0076]; and figures 1A, 2B.	1-4, 7-10, 12-16, 19-23
Y		5-6, 11, 17-18, 24-25
Y	US 2009-0149722 A1 (ABOLFATHI AMIR et al.) 11 June 2009 See paragraph [0044]; and claim 2.	5-6, 11, 17-18, 24-25
A	US 2009-0129620 A1 (JUNICHI TAGAWA et al.) 21 May 2009 See paragraphs [0092]-[0095]; and claim 16.	1-25
A	US 2002-0194005 A1 (ROY J. LAHR) 19 December 2002 See paragraphs [0097]-[0123]; and figure 9.	1-25
A	US 2013-0119255 A1 (TIMOTHY DICKINSON et al.) 16 May 2013 See paragraphs [0142]-[0144]; and figure 21.	1-25

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

10 February 2015 (10.02.2015)

Date of mailing of the international search report

10 February 2015 (10.02.2015)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea

Facsimile No. ++82 42 472 3473

Authorized officer

BYUN, Sung Cheal

Telephone No. +82-42-481-8262



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/065621

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012-0035924 A1	09/02/2012	AU 2013-285618 A1 EP 2601601 A1 EP 2601601 A4 KR 10-2013-0101505 A US 2012-0035932 A1 US 8473289 B2 WO 2012-019028 A1	21/02/2013 12/06/2013 01/01/2014 13/09/2013 09/02/2012 25/06/2013 09/02/2012
US 2009-0149722 A1	11/06/2009	AU 2009-333769 A1 AU 333769 B2 CA 2707717 A1 CN 101919261 A CN 101919261 B EP 2227912 A1 EP 2227912 A4 GB 201011316 D0 GB 2470840 A GB 2470840 B JP 2011-512049 A JP 2013-208447 A US 2014-349597 A1 US 8795172 B2 WO 2009-073852 A1	11/06/2009 01/08/2013 11/06/2009 15/12/2010 10/09/2014 15/09/2010 12/11/2014 18/08/2010 08/12/2010 12/12/2012 14/04/2011 10/10/2013 27/11/2014 05/08/2014 11/06/2009
US 2009-0129620 A1	21/05/2009	CN 101390440 A CN 101390440 B JP 04931907 B2 US 8155345 B2 WO 2007-099908 A1	18/03/2009 10/10/2012 16/05/2012 10/04/2012 07/09/2007
US 2002-0194005 A1	19/12/2002	US 7082393 B2 WO 02-077972 A1	25/07/2006 03/10/2002
US 2013-0119255 A1	16/05/2013	US 8541745 B2	24/09/2013