



- (51) **International Patent Classification:**
H04R 1/10 (2006.01) *H04N 21/433* (2011.01)
- (21) **International Application Number:** PCT/CN2020/093172
- (22) **International Filing Date:** 29 May 2020 (29.05.2020)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (71) **Applicant:** NOKIA TECHNOLOGIES OY [FI/FI]; Karakaari 7, Espoo 02610 (FI).
- (71) **Applicant (for LC only):** NOKIA TECHNOLOGIES (BEIJING) CO., LTD. [CN/CN]; Unit 18, Room 1001, Level 10, East Tower, World Financial Center, No. 1, East Third Ring Middle Road, Chaoyang District, Beijing 100020 (CN).
- (72) **Inventors:** XIAO, Zhen; No. 301, Building 66, Dayouzhuang 100, Haidian District, Beijing 100091 (CN).
- (74) **Agent:** KING & WOOD MALLESONS; 20th Floor, East Tower, World Financial Centre, No. 1 Dongsanhuan Zhonglu, Chaoyang District, Beijing 100020 (CN).
- (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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,

(54) **Title:** NOISE CONTROL

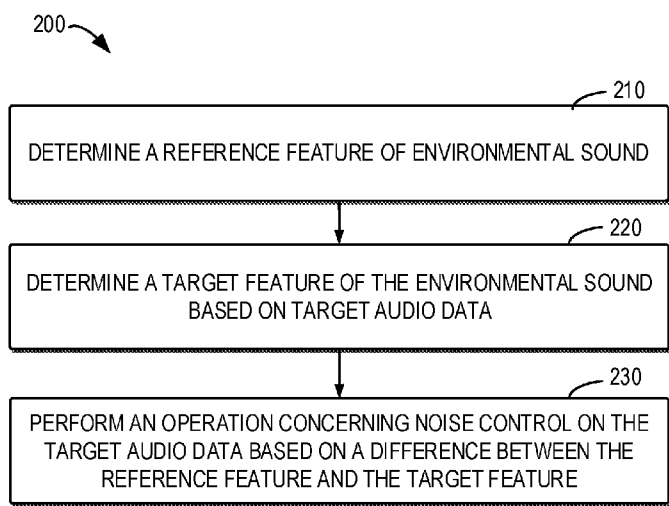


Fig. 2

(57) **Abstract:** Embodiments of the present disclosure relate to noise control. In a method, a reference feature of environmental sound is determined based on reference audio data captured from an environment within a plurality of reference time periods. A target feature of the environmental sound is determined based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods. An operation concerning noise control on the target audio data is performed based on a difference between the reference feature and the target feature.



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))*

NOISE CONTROL

FIELD

[0001] Embodiments of the present disclosure generally relate to an apparatus, method,
5 system, and computer readable storage medium for noise control.

BACKGROUND

[0002] Active noise control (ANC), also known as noise cancellation, or active noise
reduction (ANR) are methods for reducing unwanted sound by adding an additional sound.
10 The additional sound is specifically designed to cancel the unwanted sound. Currently, a
variety of headphones employ the ANC to reduce environmental noises. The existing
ANC headphones filter out or reduce noises based on frequencies of the noises only. As a
result, these headphones cannot distinguish between environmental noises and meaningful
sounds. For example, when a user wearing an ANC headphone is walking or running
15 outdoors, the ANC headphone may filter out the car horn. As another example, in an
office environment, the ANC headphone may filter out the voices of people speaking to the
user. However, the car horn and the voices of people speaking to the user are meaningful
sounds which should be heard by the user. Therefore, the existing ANC headphones
cannot intelligently pass such sounds to alert the user.

20

SUMMARY

[0003] In general, example embodiments of the present disclosure provide a solution for
noise control.

[0004] In a first aspect, there is provided an apparatus. The apparatus comprises at least
25 one processor; and at least one memory including computer program codes; the at least one
memory and the computer program codes are configured to, with the at least one processor,
cause the apparatus to determine a reference feature of environmental sound based on
reference audio data captured from an environment within a plurality of reference time
periods; determine a target feature of the environmental sound based on target audio data
30 captured from the environment within a target time period subsequent to the plurality of
reference time periods; and perform an operation concerning noise control on the target
audio data based on a difference between the reference feature and the target feature.

[0005] In a second aspect, there is provided a method. The method comprises: obtaining, by one or more processors, a reference feature of environmental sound, the reference feature determined based on reference audio data captured from an environment within a plurality of reference time periods; determining, by one or more processors, a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and performing, by one or more processors, an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

[0006] In a third aspect, there is provided an apparatus. The apparatus comprises means for determining a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods; means for determining a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and means for performing an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

[0007] In a fourth aspect, there is provided a system. The system one or more processors; and one or more computer-readable media having stored thereon instructions that are executable by the one or more processors to cause the system to: determine a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods; determine a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and perform an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

[0008] In a fifth aspect, there is provided a non-transitory computer readable medium comprising program instructions for causing an apparatus to perform at least the method according to the second aspect.

[0009] It is to be understood that the summary section is not intended to identify key or essential features of embodiments of the present disclosure, nor is it intended to be used to limit the scope of the present disclosure. Other features of the present disclosure will become easily comprehensible through the following description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Some example embodiments will now be described with reference to the accompanying drawings, where:

5 [0011] Fig. 1 illustrates an example scenario in which example embodiments of the present disclosure may be implemented;

[0012] Fig. 2 illustrates a flowchart of an example method according to some example embodiments of the present disclosure;

10 [0013] Fig. 3 illustrates a schematic diagram showing reference audio data and target audio data according to some example embodiments of the present disclosure;

[0014] Fig. 4 illustrates a schematic diagram of training an example auto-encoder according to some example embodiments of the present disclosure;

[0015] Fig. 5 illustrates a schematic diagram of using the example auto-encoder according to some example embodiments of the present disclosure;

15 [0016] Fig. 6 illustrates a schematic diagram of updating the example auto-encoder according to some example embodiments of the present disclosure; and

[0017] Fig. 7 illustrates a schematic block diagram of an example device suitable for implementing embodiments of the present disclosure.

20 [0018] Throughout the drawings, the same or similar reference numerals represent the same or similar element.

DETAILED DESCRIPTION

25 [0019] Principle of the present disclosure will now be described with reference to some example embodiments. It is to be understood that these embodiments are described only for the purpose of illustration and help those skilled in the art to understand and implement the present disclosure, without suggesting any limitation as to the scope of the disclosure. The disclosure described herein can be implemented in various manners other than the ones described below.

30 [0020] In the following description and claims, unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skills in the art to which this disclosure belongs.

[0021] References in the present disclosure to “one embodiment,” “an embodiment,” “an example embodiment,” and the like indicate that the embodiment described may include a particular feature, structure, or characteristic, but it is not necessary that every embodiment includes the particular feature, structure, or characteristic. Moreover, such phrases are not necessarily referring to the same embodiment. Further, when a particular feature, structure, or characteristic is described in connection with an embodiment, it is submitted that it is within the knowledge of one skilled in the art to affect such feature, structure, or characteristic in connection with other embodiments whether or not explicitly described.

[0022] It shall be understood that although the terms “first” and “second” etc. may be used herein to describe various elements, these elements should not be limited by these terms. These terms are only used to distinguish one element from another. For example, a first element could be termed a second element, and similarly, a second element could be termed a first element, without departing from the scope of example embodiments. As used herein, the term “and/or” includes any and all combinations of one or more of the listed terms.

[0023] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of example embodiments. As used herein, the singular forms “a,” “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “comprises,” “comprising,” “has,” “having,” “includes” and/or “including”, when used herein, specify the presence of stated features, elements, and/or components etc., but do not preclude the presence or addition of one or more other features, elements, components and/or combinations thereof.

[0024] As used in this application, the term “circuitry” may refer to one or more or all of the following:

(a) hardware-only circuit implementations (such as implementations in only analog and/or digital circuitry) and

(b) combinations of hardware circuits and software, such as (as applicable):

(i) a combination of analog and/or digital hardware circuit(s) with software/firmware and

(ii) any portions of hardware processor(s) with software (including digital signal processor(s)), software, and memory(ies) that work together to cause an apparatus, such as a mobile phone or server, to perform various functions) and

(c) hardware circuit(s) and or processor(s), such as a microprocessor(s) or a portion of a microprocessor(s), that requires software (e.g., firmware) for operation, but the software may not be present when it is not needed for operation.

5 [0025] This definition of circuitry applies to all uses of this term in this application, including in any claims. As a further example, as used in this application, the term circuitry also covers an implementation of merely a hardware circuit or processor (or multiple processors) or portion of a hardware circuit or processor and its (or their) accompanying software and/or firmware. The term circuitry also covers, for example and if applicable to the particular claim element, a baseband integrated circuit or processor integrated circuit for a
10 mobile device or a similar integrated circuit in server, a cellular network device, or other computing or network device.

[0026] As mentioned above, conventional ANC headphones cannot distinguish between environmental noises and meaningful sounds. For example, the car horn when walking outdoors and the voices of people saying to the user may be filtered out. As a result, these
15 meaningful sounds might not be heard by the user. Some solutions have been proposed to distinguish the meaningful sounds from background noises. However, these solutions are based on either a fixed dataset or predefined “novel” sounds and thus cannot be adapted to the varying ambient environment. For example, these solutions cannot recognize a sound which is not included in the dataset or is not predefined. Therefore, more intelligent ANC
20 is needed.

[0027] The inventors of the present disclosure have realized that people gradually adapt to a repetitive, mechanical noise after listening for a while. In other words, people turn a deaf ear to the repetitive, mechanical noise. Attention will not be attracted until an unusual sound appears. This function of the human brain may be imitated so as to perform active
25 noise control on repetitive, ordinary sounds and turn off or reduce active noise control on unusual sounds.

[0028] In order to at least partially solve the above and other potential problems, embodiments of the present disclosure provide a new solution for noise control. In this solution, a reference feature of environmental sound is obtained. The reference feature is
30 determined based on reference audio data captured from an environment within a plurality of reference time periods. A target feature of the environmental sound is determined based on target audio data captured from the environment within a target time period. The target time

period is subsequent to the plurality of reference time periods where the reference audio data is captured.

[0029] It will be appreciated from the following description that, according to example embodiments of the present invention, an operation concerning noise control on the target audio data can be performed based on a difference between the reference feature and the target feature. For example, if the target feature is significantly different from the reference feature, this means that the target audio data may represent a potentially novel sound from the environment. Accordingly, noise control on the target audio data may be disabled or reduced so as to pass the potentially novel sound to a user. If the target feature is similar to the reference feature, this means that the target audio data may only represent background sound from the environment. Accordingly, noise control on the target audio data may be enabled to cancel the background sound.

[0030] In this way, noise control on the audio data can be adapted to the environment. The repetitive sound from the environment can be cancelled and the novel sound from the environment can be passed to the user. Therefore, adaptive noise control can be achieved.

[0031] Some example embodiments of the present disclosure will be described in detail below.

Example Scenario

[0032] Fig. 1 illustrates an example scenario 100 in which example embodiments of the present disclosure may be implemented. In the example scenario 100, an apparatus 110 may continuously detect environmental sound 130 from an ambient environment where the apparatus 110 is located. Although not shown, the apparatus 110 may be configured with a function of noise control to cancel any noise in the environmental sound 130. For example, the apparatus 110 may comprise an ANC circuitry or any other suitable means for ANC.

[0033] At the apparatus 110, the detected environmental sound 130 may be converted or digitalized into audio data 150. As shown in Fig. 1, the audio data 150 may comprise a plurality of data samples. Each data sample may comprise audio data captured from the ambient environment within a time period and represent the environmental sound 130 within the time period. In other words, the audio data 150 may be divided into a plurality of sound slices, including the sound slices 101-103 for example. Each sound slice comprises the audio data captured within a time period. In the following, the terms

“sound slice” and “data sample” can be used interchangeably. In some example embodiments, each sound slice may have a same duration, such as, 10ms.

[0034] For each sound slice, the apparatus 110 may determine whether the sound slice comprises a potentially novel or meaningful sound (which may also be referred to as unusual sound). The apparatus 110 may disable the noise control (for example, ANC) or reduce the level of noise control on the sound slice comprising the unusual sound. As indicated by the curve 140, the level of noise control on the sound slices 102 and 103 may be reduced as compared to the sound slice 101. In this way, the unusual sound which is potentially meaningful to a user of the apparatus 110 can be heard by the user.

[0035] In some example embodiments, the solution for noise control may be implemented in a distributed manner. For example, as shown in Fig. 1, the apparatus 110 may communicate with one or more other apparatuses, including but not limited to one or more of a terminal device 120-1, a computer 120-2 and a server 120-3 in a cloud computing node. In the following, the terminal device 120-1, the computer 120-2 and the server 120-3 in the cloud computing node may be collectively referred to as a further apparatus 120 or an apparatus 120. In these example embodiments, the further apparatus 120 may process reference audio data and the apparatus 110 may receive information about the reference audio data from the further apparatus 120, as will be described below.

[0036] It is to be understood that the number of apparatuses and sound slices is only for the purpose of illustration without suggesting any limitations as to the scope of the present disclosure. Although the apparatus 110 is shown as a headphone in Fig. 1, this is only for the purpose of illustration. The solution for noise control according to embodiments of the present disclosure can be implemented in any suitable type of device and/or system where noise control is required.

Example Method

[0037] Now refer to Fig. 2. Fig. 2 illustrates a flowchart of an example method 200 according to some example embodiments of the present disclosure. The method 200 may be implemented at the apparatus 110 as shown in Fig. 1. Alternatively, the method 200 may be implemented across the apparatus 110 and the apparatus 120 as shown in Fig. 1. For the purpose of discussion, the method 200 will be described with reference to Figs. 1 and 3.

[0038] At block 210, a reference feature of environmental sound 130 is determined based

on reference audio data captured from an environment within a plurality of reference time periods. In some example embodiments, the reference feature may be determined at the apparatus 110.

[0039] In some example embodiments, the reference feature may be determined at the further apparatus 120, such as the terminal device 120-1, the computer 120-2 or the server 120-3 in the cloud computing node as shown in Fig. 1. The apparatus 110 may receive the reference feature from the further apparatus 120. In these example embodiments, the reference audio data may be captured by the apparatus 110 and communicated to the further apparatus 120. Alternatively, or in addition, if the further apparatus 120 is located in the same environment as the apparatus 110, the reference audio data may be captured by the further apparatus 120.

[0040] Now refer to Fig. 3. Fig. 3 illustrates a schematic diagram 300 showing reference audio data 310 and target audio data 305 according to some example embodiments of the present disclosure. As mentioned above, the environmental sound 130 may be converted or digitalized into audio data for further processing.

[0041] A sliding time window 350 may be used to locate the reference audio data 310. As shown in Fig. 3, the reference audio data 310 may comprise a plurality of data samples or sound slices. Fig. 3 shows sound slices 301-1, 301-2 and 301-3, which may be collectively referred to as a sound slice 301 or a reference sound slice 301. Each sound slice 301 may comprise audio data captured from the environment within a time period. The plurality of sound slices 301 in the sliding time window 350 may be collectively referred to as a reference sound sequence.

[0042] The duration of each sound slice 301 and the width of the sliding time window 350 may employ any suitable value. As an example rather than any limitation as to the scope of the present disclosure, the duration of each sound slice 301 may be 10 milliseconds and the width of the sliding time window 350 may be 2 seconds.

[0043] One or more features may be determined from the reference audio data 310 to determine the reference feature. The one or more features may include but not limited to frequency distribution, temporal variation, a distance and direction (for example, an angle) of a sound source, and semantics (if a speech is detected). For example, the feature of each sound slice 301 comprised in the reference audio data 310 may be extracted and the reference feature may be determined based on the features of the plurality of sound slices

301 in the reference audio data 310.

[0044] In some example embodiments, the reference feature may comprise a feature of a frequency of the environmental sound 130. For example, the highest frequency included in each sound slice 301 may be determined. Then, the reference feature may be determined as an average or a median of the highest frequencies of the plurality of sound slices 301.

[0045] Alternatively, or in addition, in some example embodiments, the reference feature may comprise a feature of a source of the environmental sound 130. As an example, a distance from the source of the environmental sound 130 to the apparatus 110 may be determined from each sound slice 301. Then, the reference feature may be determined based on the distances determined from the plurality of sound slices 301. As another example, a direction of the source of the environmental sound 130 relative to the apparatus 110 may be determined from each sound slice 301. Then, the reference feature may be determined based on the directions determined from the plurality of sound slices 301. As a further example, a type of a device producing the environmental sound 130 may be determined from each sound slice 301. Then, the reference feature may represent one or more types of devices producing the environmental sound 130.

[0046] Alternatively, or in addition, in some example embodiments, if a speech is detected in the environmental sound 130, the reference feature may comprise a feature of semantics of the environmental sound 130. For example, a semantic feature may be extracted from each sound slice 301 and the reference feature may be determined as an average of the semantic features of the plurality of sound slices 301.

[0047] Alternatively, or in addition, in some example embodiments, the reference feature may comprise a feature of a waveform of the environmental sound 130. The feature of the waveform may comprise a magnitude or a shape of the waveform for example. As an example, the magnitude of the waveform of each sound slice 301 may be determined. Then, the reference feature may be determined as an average or a median of the magnitudes of the plurality of sound slices 301.

[0048] The reference audio data 310 may represent various sounds, for example, the noise of an engine when commuting, the conversation between people in open offices, and the crisp sound of cup collisions in cafes. It may be difficult to find unified features of the various sounds in advance. Therefore, in some example embodiments, the reference

feature may be determined based on a neural network which is trained with the plurality of sound slices 301 of the reference audio data 310. In this case, a plurality of features of the plurality of sound slices 301 may be determined based on the trained neural network. Then, the reference feature may be determined based on the plurality of features of the plurality of sound slices 301.

[0049] In an example implementation, training of the neural network and determination of the reference feature may be both implemented at the apparatus 110. In another example implementation, the training of the neural network and the determination of the reference feature may be both implemented at the further apparatus 120. In this example implementation, the reference feature may be communicated from the further apparatus 120 to the apparatus 110.

[0050] In a further example implementation, the training of the neural network may be implemented at the further apparatus 120 and parameters of the trained neural network may be communicated from the further apparatus 120 to the apparatus 110. The apparatus 110 may determine the reference feature using the parameters of the trained neural network.

[0051] In some example embodiments, the neural network may be an auto-encoder. Such example embodiments will be detailed below with reference to Figs. 4-6.

[0052] Now still refer to Fig. 2. At block 220, a target feature of the environmental sound 130 is determined based on target audio data 305. The target audio data 305 is captured from the environment within a target time period and the target time period is subsequent to the plurality of reference time periods where the reference audio data 310 is captured. As shown in Fig. 3, the target sound slice 302 may comprise the target audio data 305 captured within the target time period. In some example embodiments, the target time period may be a current time period.

[0053] The determination of the target feature depends on the determination of the reference feature. As an example, if the reference feature is determined as the average or the median of the highest frequencies of the plurality of sound slices 301, the target feature may be determined as the highest frequency included in the target audio data 305. As another example, if the reference feature is determined based on the distances extracted from the plurality of sound slices 301, a distance from a sound source to the apparatus 110 may be extracted from the target audio data 305 as the target feature. As a further example, if the reference feature is determined based on the semantic features of the plurality

of sound slices 301, a semantic feature may be extracted from the target audio data 305 as the target feature.

[0054] In the example embodiments where the reference feature is determined based on the neural network, the target feature may be determined by applying the target audio data 305 to the neural network. Such example embodiments will be detailed below with reference to Figs. 4-6.

[0055] At block 230, an operation concerning noise control on the target audio data is performed based on a difference between the reference feature and the target feature. A difference threshold may be introduced to compare with the difference between the reference feature and the target feature. The difference threshold may be predetermined or configured at the apparatus 110. Alternatively, the difference threshold may be determined based on a distribution of the features of the plurality of sound slices 301 where those features are used to determine the reference feature. In an example implementation, if the reference feature is determined as the average or the median of the highest frequencies of the plurality of sound slices 301, the difference threshold may be determined based on a standard deviation of the highest frequencies of the plurality of sound slices 301.

[0056] If the difference between the reference feature and the target feature exceeds the difference threshold, this means that the target audio data 305 may represent a potentially meaningful or novel sound. In this case, the noise control on the target audio data 305 may be disabled. For example, the ANC circuitry in the apparatus 110 may be instructed or controlled so as not to perform ANC on the target audio data 305. Alternatively, the level of noise control on the target audio data 305 may be reduced. For example, the ANC circuitry in the apparatus 110 may be instructed or controlled to reduce the level of ANC on the target audio data 305.

[0057] As an example, the user wearing the apparatus 110 may be running or walking outdoors. The car horn which occurs suddenly may be quite different from the previous environmental sound. In this case, the target audio data 350 representing the car horn may not be filtered out as a noise.

[0058] If the difference between the reference feature and the target feature is below the difference threshold, this means that the target audio data 305 is unlikely to represent a potentially meaningful sound. In this case, the noise control on the target audio data 305 may be enabled. For example, the ANC circuitry in the apparatus 110 may be instructed

or controlled to perform ANC on the target audio data 305.

[0059] Continue with the above example of the car horn. The car horn may continue for a long time. For example, the duration of the car horn may exceed the duration of the sliding window 350. In this case, the car horn may be recognized as a noise and cancelled.

[0060] The human ear is adaptable to the environment. If the environment changes from a relatively quiet environment to a noisier environment, the novel sound from the external environment may initially attract attention of the user. However, after gradually adapting to the environmental sound, the brain of the user may enter a state of "turn-a-deaf-ear".

[0061] To achieve this adaptive function, the reference feature may be updated over time, for example as the sliding time window moves. In some example embodiments, a data sample may be selected from the plurality of data samples comprised in the reference audio data 310. The selected data sample is captured within a reference time period furthest from the target time period among the plurality of reference time periods. Then, the reference audio data may be updated by replacing the selected data sample with the target audio data and the reference feature may be updated based on the updated reference audio data.

[0062] Still refer to Fig. 3 to illustrate an example. As shown in Fig. 3, the sliding time window 350 may move over time. The sound slice 301-1 which is furthest from the sound slice 302 in time may be popped out of the sliding time window 350 and the sound slice 302 may be pushed into the sliding time window 350. As a result, the reference audio data 310 is updated to reference audio data 320. The reference feature of the environment sound 130 may be updated based on the reference audio data 320. The updated reference feature may be used to determine the operation concerning noise control on next target audio data 306.

[0063] The acts described above with respect to blocks 210, 220 and 230 may be performed on audio data captured within each time period. In this way, the sound which is potentially meaningful or novel to the user will not be cancelled and thus can be heard by the user.

[0064] The example method 200 has been described with reference to Fig. 2. Acts described above may be performed at a single physical device, such as an earphone. Alternatively, the acts may be performed across multiple physical devices. As an example,

some of the acts may be carried out at the earphone and some of the acts may be performed at a terminal device such as a mobile phone. As another example, some of the acts may be carried out at the earphone and some of the acts may be performed at a cloud computing node.

5 Example Auto-encoder

[0065] As mentioned above, in some example embodiments, the reference feature and the target feature may be determined based on a neural network. For example, the reference audio data may comprise a plurality of data samples, such as the plurality of sound slices 301. A plurality of features of the plurality of data samples may be determined based on the neural network which is trained with the plurality of data samples. Then, the reference feature may be determined based on the plurality of features of the plurality of data samples. Likewise, the target feature may be determined by applying the target audio data 305 to the neural network. The neural network may be based on any suitable machine learning technology. In some example embodiment, the neural network may comprise an auto-encoder.

Some example embodiments where the auto-encoder is used are now described with reference to Figs. 4-6. Fig. 4 illustrates a schematic diagram 400 of training an example auto-encoder 450 according to some example embodiments of the present disclosure. The simplest form of the auto-encoder is a feed-forward, non-recurrent neural network having an input layer, an output layer and one or more hidden layers connecting the input and output layers. As shown in Fig. 4, the example auto-encoder 450 comprises an input layer 451, an output layer 455 and three hidden layers 452-454. The number of nodes (which may be also referred to as neurons) of the output layer 455 is the same as that of the input layer 451.

[0066] The purpose of the auto-encoder 450 is to reconstruct its inputs. In other words, the auto-encoder 450 may be trained by minimizing the difference between the input and output of the auto-encoder 450. Therefore, the auto-encoder 450 is an unsupervised learning model.

[0067] The auto-encoder 450 may comprise two parts, an encoder and a decoder, which can be defined as following:

$$\phi : \mathcal{X} \rightarrow \mathcal{F} \quad (1)$$

$$\psi: \mathcal{F} \rightarrow \mathcal{X} \quad (2)$$

$$\phi, \psi = \arg \min_{\phi, \psi} \|X - (\psi \circ \phi)X\|^2 \quad (3)$$

where ϕ may represent the encoder and ψ may represent the decoder.

[0068] During training the auto-encoder 450, X represents each data sample or sound slice in the reference audio data 310. The input and target output of the auto-encoder 450 are the same data sample. All the data samples in the reference audio data 310 may be used to train the auto-encoder 450. For example, as shown in Fig. 4, the sound slice 301-1 is used to train the auto-encoder 450. The training of the auto-encoder 450 makes the output of the neural network as close to the target output as possible.

[0069] After training, the features of the sound slices 301 in the reference audio data 310 can be recorded in the auto-encoder. Generally, there is a bottleneck layer with fewer parameters in the auto-encoder. The bottleneck layer enables the auto-encoder to compress the data and obtain the characteristics of the data. As shown in Fig. 4, the hidden layer 450 with the fewest parameters is a bottleneck layer of the auto-encoder 450.

[0070] In some example embodiments, the training of the auto-encoder 450 may be implemented at the apparatus 110. In some example embodiments, the training of the auto-encoder 450 may be implemented at the further apparatus 120. In these example embodiments, parameters of the trained auto-encoder 450 may be communicated to the apparatus 110.

[0071] The auto-encoder 450 can be implemented by any suitable network structure, such as a simple shallow fully connected neural network (e.g., 1-5 hidden layers) or a more complex convolutional neural network. The network structure used may depend on the computing power provided by the hardware where the auto-encoder 450 is trained.

[0072] The reference feature may be determined based on the trained auto-encoder 450 and the sound slices 301 in the reference audio data 310. To determine the reference feature, the feature of each sound slice 301 may be determined first. It is assumed that X_i represents the i -th sound slice and L_i represents the feature of the i -th sound slice X_i , where i is a natural number.

[0073] The feature of each sound slice may be represented in a variety of ways. As an

example, the loss function for training the auto-encoder 450 may be used to calculate the feature. Accordingly, L_i may be defined as following:

$$L_i = \min_{\phi, \psi} \|X_i - (\psi \circ \phi)X_i\|^2 \quad (4)$$

where the value of L_i is tiny, since the auto-encoder 450 has been trained with the reference audio data 310.

[0074] The feature L_i of the i -th sound slice X_i may also be calculated in other ways. For example, the output of the hidden layer 453 which is the bottleneck layer may be used to calculate the feature L_i . The i -th sound slice X_i may be applied to the trained auto-encoder 450 and the sum of squared output of the hidden layer 453 may be calculated as the feature L_i .

[0075] The reference feature may be determined based on the plurality of features L_i of the sound slices 301 in the reference audio data 310. As an example, the reference feature may be determined as an average of the plurality of features L_i .

[0076] As another example, one or more statistic parameters of the plurality of features L_i may be used to determine the reference feature. The plurality of features L_i may have a certain distribution. Refer to Fig. 5. A plot 501 shows an example distribution of the plurality of features L_i . For example, the mean L_{mean} of the distribution of the plurality of features L_i can be used as the reference feature.

[0077] In addition, the difference threshold may be determined based on the standard deviation L_{std} of the distribution of the plurality of features L_i . For example, the difference threshold may be determined as $M * L_{std}$, where M may be a parameter configured in advance or given by a user of the apparatus 110.

[0078] The target feature may be determined by applying the target audio data 305 to the trained auto-encoder 450, as shown in Fig. 5. It is assumed that X_t represents the target audio data 305 and L_t represents the target feature. X_t may be applied to the trained auto-encoder 450 as input and the target feature L_t may be calculated accordingly. For example, the loss function as expressed in the equation (4) or the output of the hidden layer 453 may be used to calculate the target feature L_t .

[0079] If the difference between the reference feature L_{mean} and the target feature L_t exceeds the difference threshold, such as $M * L_{std}$, this means that the target audio data

305 is significantly different from the reference audio data 310. In this case, the noise control on the target audio data 305 may be disabled. Or, alternatively, the level of the noise control on the target audio data 305 may be reduced.

[0080] If the difference between the reference feature L_{mean} and the target feature L_t is below the difference threshold, such as $M * L_{std}$, this means that the target audio data 305 is similar to the reference audio data 310. In this case, the apparatus 110 may enable the noise control on the target audio data 305.

[0081] As mentioned above, to achieve the adaptive function, the reference feature may be updated over time. The auto-encoder 450 may be updated based on the target audio data 305 so as to update the reference feature. For example, the target audio data 305 may be added to the reference audio data and used to further train the auto-encoder 450. Since the auto-encoder 450 is further trained only with the newly added target audio data 305, the cost of training the auto-encoder 450 to update in real time is very low.

[0082] In some example embodiments, the determination of the target feature and the update of the auto-encoder 450 may be combined into a single process. In other words, the target audio data 305 may be applied to the auto-encoder 450 to determine the target feature as well as to further train the auto-encoder 450.

[0083] Refer to Fig. 6. As the sliding time window 350 moves forward over time, new sound slices may be used to update the auto-encoder 450 and thus the auto-encoder 450 may continuously learn the feature of new reference audio data 610. In this way, an adaptation to the external environment may be gradually achieved.

[0084] Example embodiments of the auto-encoder are described above with reference to Figs. 4-6. In these example embodiments, most-recent sound slices are used to continuously update the auto-encoder and learn the features of recent environmental sound. This approach uniquely mimics effort of human brain. In this way, a repetitive noise may be learnt and ignored, regardless of what kind of sound the noise is. Even if a noise is completely new and not present before in any dataset, the noise can also be learnt and filtered out in this way.

[0085] Moreover, the auto-encoder is an unsupervised learning model, which can automatically estimate the environment, without the need of manual intervention. Therefore, the apparatus 110 can operate in any unexpected environment with any distribution of background noises. The neural network of the auto-encoder can be simple or

gradually complicated. For a simpler auto-encoder, the required computing power is very low, and the auto-encoder can be trained in real time, for example at the apparatus 110.

[0086] In some example embodiments, an apparatus capable of performing the method 200 may comprise means for performing the respective steps of the method 200. The means
5 may be implemented in any suitable form. For example, the means may be implemented in a circuitry or software module.

[0087] In some example embodiments, the apparatus comprises means for determining a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods; means for determining a target
10 feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and means for performing an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

[0088] In some example embodiments, the means for performing the operation concerning noise control on the target audio data comprises: means for in accordance with a
15 determination that the difference exceeds a difference threshold, disabling the noise control on the target audio data.

[0089] In some example embodiments, the means for performing the operation concerning noise control on the target audio data comprises: means for in accordance with a
20 determination that the difference exceeds a difference threshold, reducing a level of the noise control on the target audio data.

[0090] In some example embodiments, the means for performing the operation concerning noise control on the target audio data comprises: means for in accordance with a
25 determination that the difference is below a difference threshold, enabling the noise control on the target audio data.

[0091] In some example embodiments, the means for determining the reference feature based on the reference audio data comprises: means for determining, based on a neural
30 network, a plurality of features of a plurality of data samples comprised in the reference audio data, each data sample captured within one of the plurality of reference time periods, the neural network trained with the plurality of data samples; and means for determining the reference feature based on the plurality of features of the plurality of data samples.

[0092] In some example embodiments, the neural network comprises an auto-encoder.

[0093] In some example embodiments, the means for determining the target feature comprises: means for determining the target feature by applying the target audio data to the neural network.

5 [0094] In some example embodiments, the apparatus further comprising: means for causing the neural network to be updated based on the target audio data, the updated neural network trained with the target audio data.

10 [0095] In some example embodiments, the apparatus further comprising: means for selecting a data sample from a plurality of data samples comprised in the reference audio data, each data sample captured within one of the plurality of reference time periods, the selected data sample captured within a reference time period furthest from the target time period among the plurality of reference time periods; means for updating the reference audio data by replacing the selected data sample with the target audio data; and means for updating the reference feature based on the updated reference audio data.

15 [0096] In some example embodiments, the reference feature comprises at least one of: a first feature of a frequency of the environmental sound, a second feature of a source of the environmental sound, a third feature of semantics of the environmental sound, or a fourth feature of a waveform of the environmental sound.

[0097] In some example embodiments, the apparatus comprises an earphone.

20 [0098] In some example embodiments, a system capable of performing the method 200 may comprise one or more processors and one or more computer-readable media. The one or more computer-readable media may have stored thereon instructions that are executable by the one or more processors to cause the system to carry out the acts described with respect to the example method 200.

25 [0099] In some example embodiments, the system capable of performing the method 200 may comprise distributed computing devices. For example, some of the acts may be carried out at a device such as an earphone and some of the acts may be carried out at one or more other devices such as a terminal device or a cloud computing node.

30 [00100] Fig. 7 illustrates a schematic block diagram of an example device 700 for implementing embodiments of the present disclosure. As shown, the device 700 comprises a central process unit (CPU) 701, which may execute various suitable actions and processing based on the computer program instructions stored in the read-only memory (ROM) 702 or computer program instructions loaded in the random-access memory (RAM) 703 from a

storage unit 708. The RAM 703 may also store all kinds of programs and data required by the operations of the device 700. CPU 701, ROM 702 and RAM 703 are connected to each other via a bus 704. The input/output (I/O) interface 705 is also connected to the bus 704.

5 [00101] A plurality of components in the device 700 is connected to the I/O interface 705, comprising: an input unit 706, such as keyboard, mouse and the like; an output unit 707, e.g., various kinds of display and loudspeakers etc.; a storage unit 708, such as memory disk, optical disk etc.; and a communication unit 709, such as network card, modem, wireless transceiver and the like. The communication unit 709 allows the device 700 to exchange information/data with other devices via the computer network, such as Internet, and/or
10 various telecommunication networks.

[00102] The above described various procedures and processing, such as method 200, may be executed by the processing unit 701. For example, in some embodiments, the method 200 may be implemented as computer software programs tangibly comprised in the machine-readable medium, such as storage unit 708. In some embodiments, the computer
15 program may be partially or fully loaded and/or mounted to the device 700 via ROM 702 and/or communication unit 709. When the computer program is loaded to RAM 703 and executed by the CPU 701, one or more actions of the above described method 200 may be executed.

[00103] Although not shown, the device 700 may comprise one or more units for detecting
20 and processing the environmental sound. The device 700 may also comprise one or more units for noise control, for example an ANC circuitry.

[00104] The present disclosure may be a method, apparatus, system and/or computer program product at any possible technical detail level of integration. The computer program product may comprise a computer-readable storage medium loaded with computer-readable
25 program instructions for executing various aspects of the present disclosure.

[00105] The computer-readable storage medium may be a tangible apparatus that maintains and stores instructions utilized by the instruction executing apparatuses. The computer-readable storage medium may be, but not limited to, such as electrical storage device, magnetic storage device, optical storage device, electromagnetic storage device,
30 semiconductor storage device or any appropriate combinations of the above. More concrete examples of the computer-readable storage medium (non-exhaustive list) comprise: portable computer disk, hard disk, random-access memory (RAM), read-only memory (ROM),

erasable programmable read-only memory (EPROM or flash), static random-access memory (SRAM), portable compact disk read-only memory (CD-ROM), digital versatile disk (DVD), memory stick, floppy disk, mechanical coding devices, punched card stored with instructions thereon, or a projection in a slot, and any appropriate combinations of the above. The computer-readable storage medium utilized here is not interpreted as transient signals per se, such as radio waves or freely propagated electromagnetic waves, electromagnetic waves propagated via waveguide or other transmission media (such as optical pulses via fiber-optic cables), or electric signals propagated via electric wires.

[00106] The described computer-readable program instruction herein may be downloaded from the computer-readable storage medium to each computing/processing device, or to an external computer or external storage via Internet, local area network, wide area network and/or wireless network. The network may comprise copper-transmitted cable, optical fiber transmission, wireless transmission, router, firewall, switch, network gate computer and/or edge server. The network adapter card or network interface in each computing/processing device receives computer-readable program instructions from the network and forwards the computer-readable program instructions for storage in the computer-readable storage medium of each computing/processing device.

[00107] The computer program instructions for executing operations of the present disclosure may be assembly instructions, instructions of instruction set architecture (ISA), machine instructions, machine-related instructions, microcodes, firmware instructions, state setting data, or source codes or target codes written in any combinations of one or more programming languages, wherein the programming languages comprise object-oriented programming languages, such as Smalltalk, C++ and the like, and traditional procedural programming languages, e.g., C language or similar programming languages. The computer-readable program instructions may be implemented fully on the user computer, partially on the user computer, as an independent software package, partially on the user computer and partially on the remote computer, or completely on the remote computer or server. In the case where remote computer is involved, the remote computer may be connected to the user computer via any type of networks, comprising local area network (LAN) and wide area network (WAN), or to the external computer (e.g., connected via Internet using the Internet service provider). In some embodiments, state information of the computer-readable program instructions is used to customize an electronic circuit, e.g., programmable logic circuit, field programmable gate array (FPGA) or programmable logic

array (PLA). The electronic circuit may execute computer-readable program instructions to implement various aspects of the present disclosure.

[00108] Each aspect of the present disclosure is disclosed here with reference to the flow chart and/or block diagram of method, apparatus (system) and computer program product according to embodiments of the present disclosure. It should be understood that each block of the flow chart and/or block diagram and combinations of each block in the flow chart and/or block diagram may be implemented by the computer-readable program instructions.

[00109] The computer-readable program instructions may be provided to the processing unit of general-purpose computer, dedicated computer or other programmable data processing apparatuses to manufacture a machine, such that the instructions that, when executed by the processing unit of the computer or other programmable data processing apparatuses, generate an apparatus for implementing functions/actions stipulated in one or more blocks in the flow chart and/or block diagram. The computer-readable program instructions may also be stored in the computer-readable storage medium and cause the computer, programmable data processing apparatus and/or other devices to work in a particular manner, such that the computer-readable medium stored with instructions contains an article of manufacture, comprising instructions for implementing various aspects of the functions/actions stipulated in one or more blocks of the flow chart and/or block diagram.

[00110] The computer-readable program instructions may also be loaded into computer, other programmable data processing apparatuses or other devices, so as to execute a series of operation steps on the computer, other programmable data processing apparatuses or other devices to generate a computer-implemented procedure. Therefore, the instructions executed on the computer, other programmable data processing apparatuses or other devices implement functions/actions stipulated in one or more blocks of the flow chart and/or block diagram.

[00111] The flow chart and block diagram in the drawings illustrate system architecture, functions and operations that may be implemented by device, method and computer program product according to multiple implementations of the present disclosure. In this regard, each block in the flow chart or block diagram may represent a module, a part of program segment or code, wherein the module and the part of program segment or code comprise one or more executable instructions for performing stipulated logic functions. In some alternative implementations, it should be noted that the functions indicated in the block may

also take place in an order different from the one indicated in the drawings. For example, two successive blocks may be in fact executed in parallel or sometimes in a reverse order dependent on the involved functions. It should also be noted that each block in the block diagram and/or flow chart and combinations of the blocks in the block diagram and/or flow chart may be implemented by a hardware-based system exclusive for executing stipulated functions or actions, or by a combination of dedicated hardware and computer instructions.

[00112] Various embodiments of the present disclosure have been described above and the above description is only exemplary rather than exhaustive and is not limited to the embodiments disclosed herein. Many modifications and alterations, without deviating from the scope and spirit of the explained various embodiments, are obvious for those skilled in the art. The selection of terms in the text aims to best explain principles and actual applications of each embodiment and technical improvements made to the technology in the market by each embodiment, or enable other ordinary skilled in the art to understand embodiments of the present disclosure.

WHAT IS CLAIMED IS:

1. An apparatus comprising:

at least one processor; and

at least one memory including computer program codes;

5 the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus at least to:

determine a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods;

10 determine a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and

perform an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

15 2. The apparatus of Claim 1, wherein the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

in accordance with a determination that the difference exceeds a difference threshold, disable the noise control on the target audio data.

20 3. The apparatus of Claim 1, wherein the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

in accordance with a determination that the difference exceeds a difference threshold, reduce a level of the noise control on the target audio data.

25 4. The apparatus of Claim 1, wherein the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

in accordance with a determination that the difference is below a difference threshold, enable the noise control on the target audio data.

30 5. The apparatus of Claim 1, wherein the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

determine, based on a neural network, a plurality of features of a plurality of data samples comprised in the reference audio data, each data sample captured within one of the

plurality of reference time periods, the neural network trained with the plurality of data samples; and

determine the reference feature based on the plurality of features of the plurality of data samples.

5

6. The apparatus of Claim 5, wherein the neural network comprises an auto-encoder.

7. The apparatus of Claim 5, wherein the at least one memory and the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

10

determine the target feature by applying the target audio data to the neural network.

8. The apparatus of Claim 5, wherein the at least one memory and the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

15

cause the neural network to be updated based on the target audio data, the updated neural network trained with the target audio data.

9. The apparatus of Claim 1, wherein the at least one memory and the at least one memory and the computer program codes are configured to, with the at least one processor, cause the apparatus to:

20

select a data sample from a plurality of data samples comprised in the reference audio data, each data sample captured within one of the plurality of reference time periods, the selected data sample captured within a reference time period furthest from the target time period among the plurality of reference time periods;

25

update the reference audio data by replacing the selected data sample with the target audio data; and

update the reference feature based on the updated reference audio data.

10. The apparatus of Claim 1, wherein the reference feature comprises at least one of:

30

a first feature of a frequency of the environmental sound,
a second feature of a source of the environmental sound,
a third feature of semantics of the environmental sound, or

a fourth feature of a waveform of the environmental sound.

11. The apparatus of Claim 1, wherein the apparatus comprises an earphone.

5 12. A method comprising:

determining, by one or more processors, a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods;

10 determining, by one or more processors, a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and

performing, by one or more processors, an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

15

13. The method of Claim 12, wherein performing the operation concerning noise control on the target audio data comprises:

in accordance with a determination that the difference exceeds a difference threshold, disabling, by one or more processors, the noise control on the target audio data.

20

14. The method of Claim 12, wherein performing the operation concerning noise control on the target audio data comprises:

in accordance with a determination that the difference exceeds a difference threshold, reducing, by one or more processors, a level of the noise control on the target audio data.

25

15. The method of Claim 12, wherein performing the operation concerning noise control on the target audio data comprises:

30 in accordance with a determination that the difference is below a difference threshold, enabling, by one or more processors, the noise control on the target audio data.

16. The method of Claim 12, wherein determining the reference feature based on the reference audio data comprises:

determining, by one or more processors based on a neural network, a plurality of

features of a plurality of data samples comprised in the reference audio data, each data sample captured within one of the plurality of reference time periods, the neural network trained with the plurality of data samples; and

5 determining, by one or more processors, the reference feature based on the plurality of features of the plurality of data samples.

17. The method of Claim 16, wherein the neural network comprises an auto-encoder.

10 18. The method of Claim 16, wherein determining the target feature comprises: determining, by one or more processors, the target feature by applying the target audio data to the neural network.

15 19. The method of Claim 16, further comprising: causing, by one or more processors, the neural network to be updated based on the target audio data, the updated neural network trained with the target audio data.

20 20. The method of Claim 12, further comprising: selecting, by one or more processors, a data sample from a plurality of data samples comprised in the reference audio data, each data sample captured within one of the plurality of reference time periods, the selected data sample captured within a reference time period furthest from the target time period among the plurality of reference time periods; updating, by one or more processors, the reference audio data by replacing the selected data sample with the target audio data; and
25 updating, by one or more processors, the reference feature based on the updated reference audio data.

21. The method of Claim 12, wherein the reference feature comprises at least one of: a first feature of a frequency of the environmental sound,
30 a second feature of a source of the environmental sound, a third feature of semantics of the environmental sound, or a fourth feature of a waveform of the environmental sound.

22. An apparatus comprising:

means for determining a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods;

5 means for determining a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and

means for performing an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

10 23. A system comprising:

one or more processors; and

one or more computer-readable media having stored thereon instructions that are executable by the one or more processors to cause the system to:

15 determine a reference feature of environmental sound based on reference audio data captured from an environment within a plurality of reference time periods;

determine a target feature of the environmental sound based on target audio data captured from the environment within a target time period subsequent to the plurality of reference time periods; and

20 perform an operation concerning noise control on the target audio data based on a difference between the reference feature and the target feature.

24. A non-transitory computer readable medium comprising program instructions for causing an apparatus to perform at least the method of any of claims 12-21.

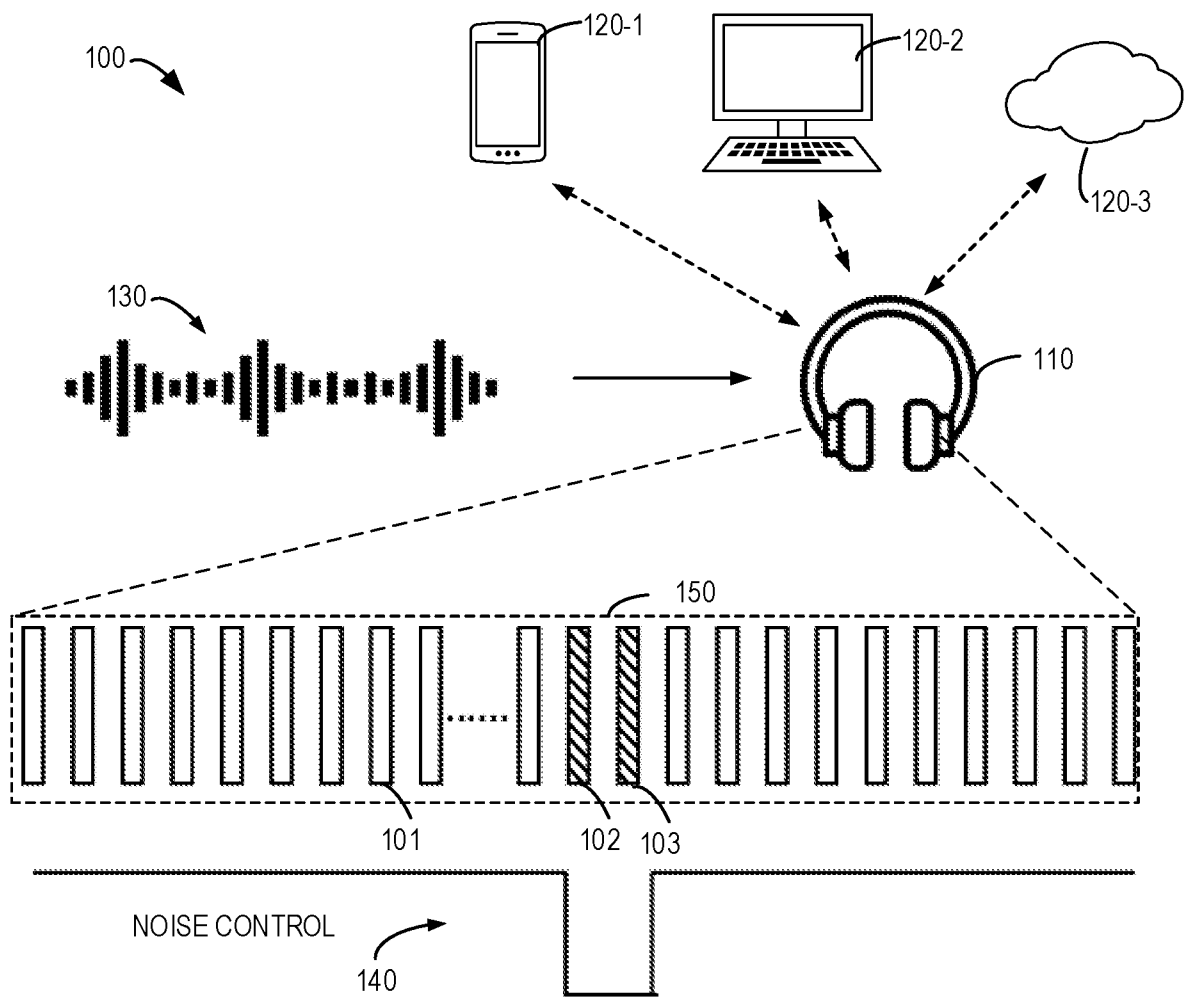


Fig. 1

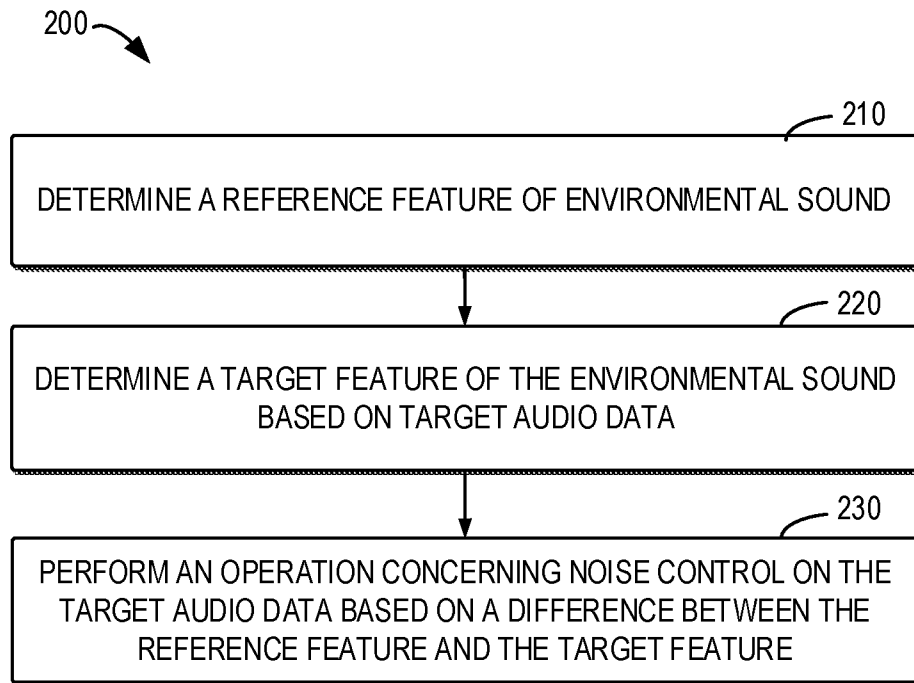


Fig. 2

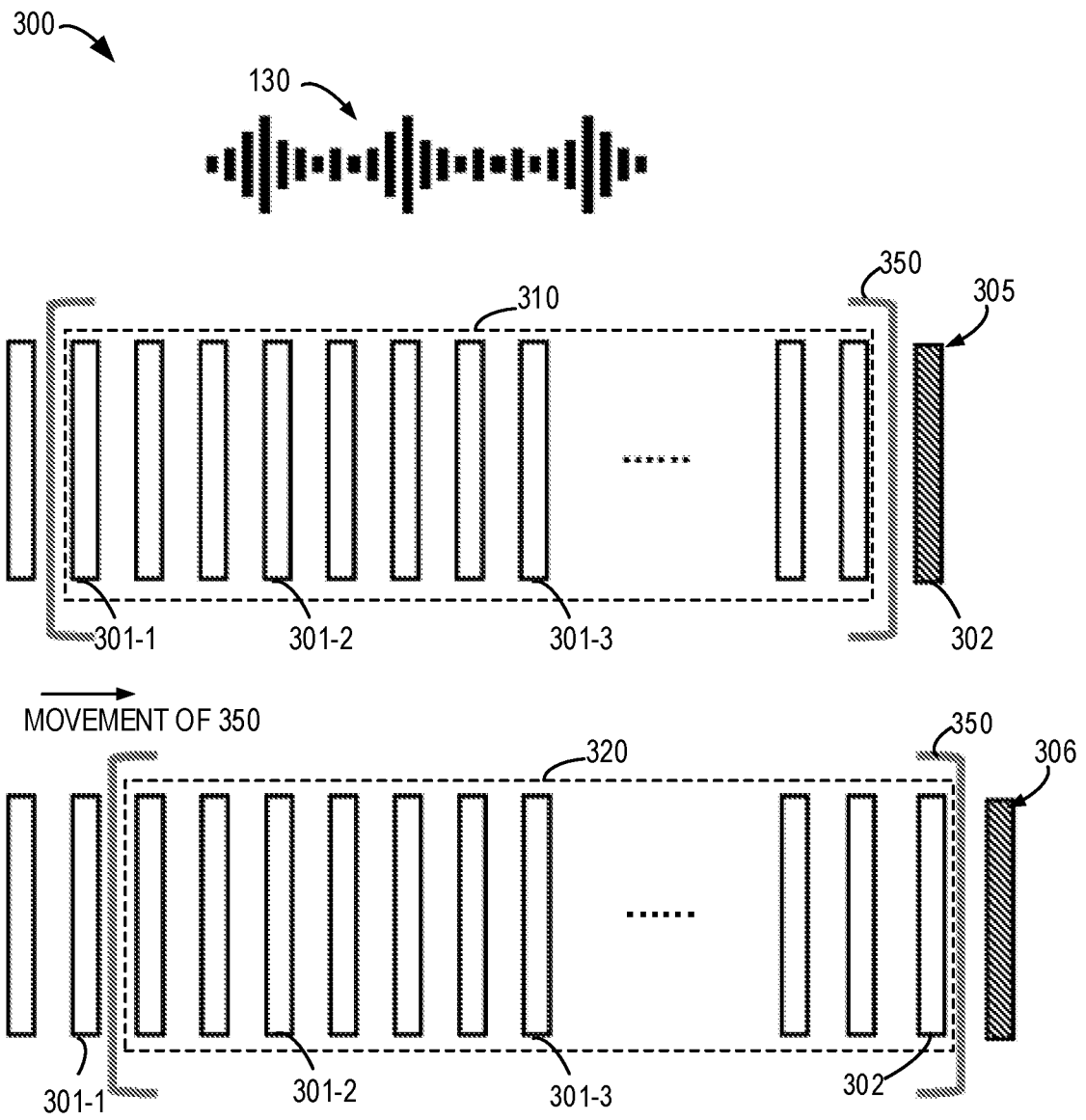


Fig. 3

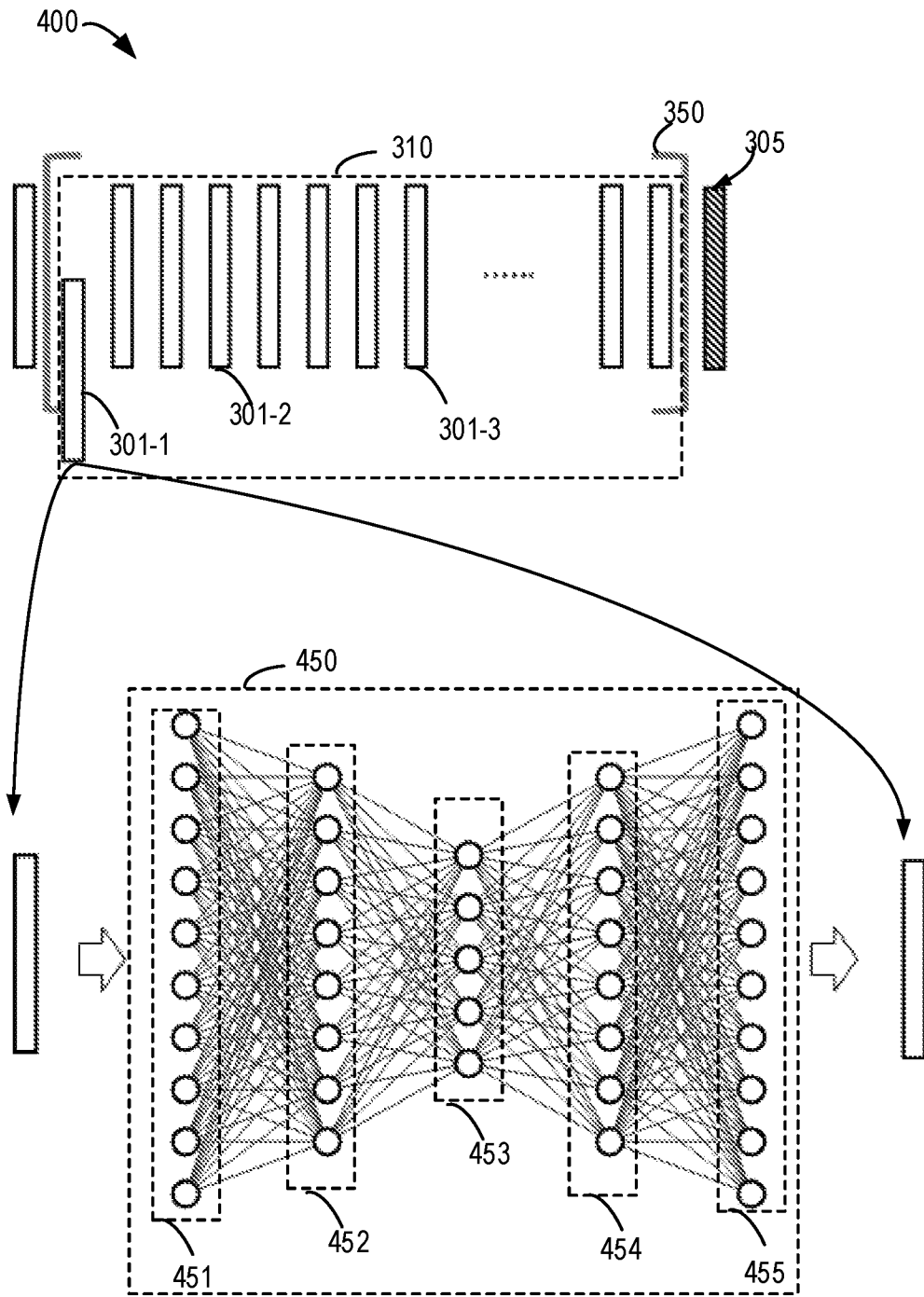


Fig. 4

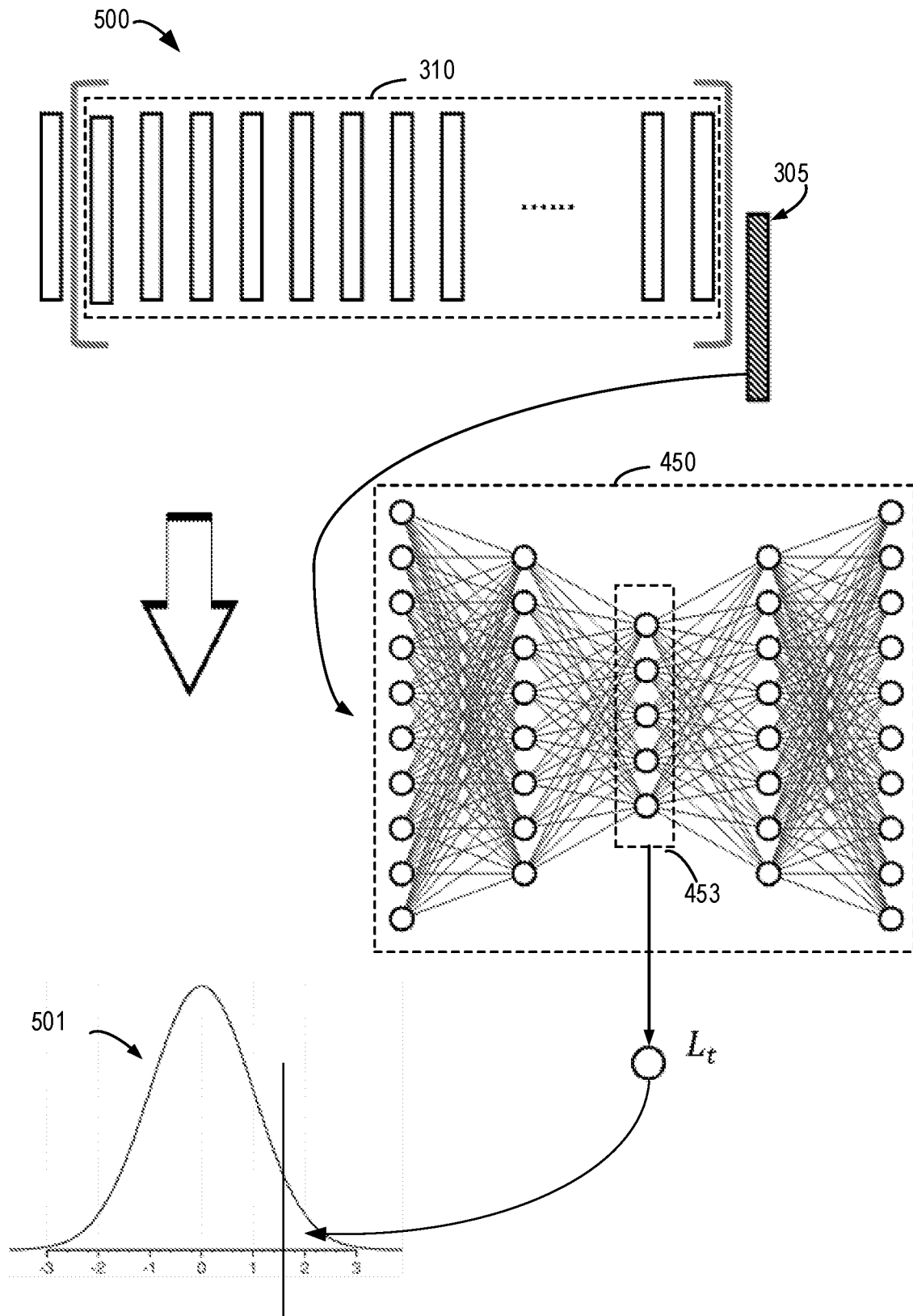


Fig. 5

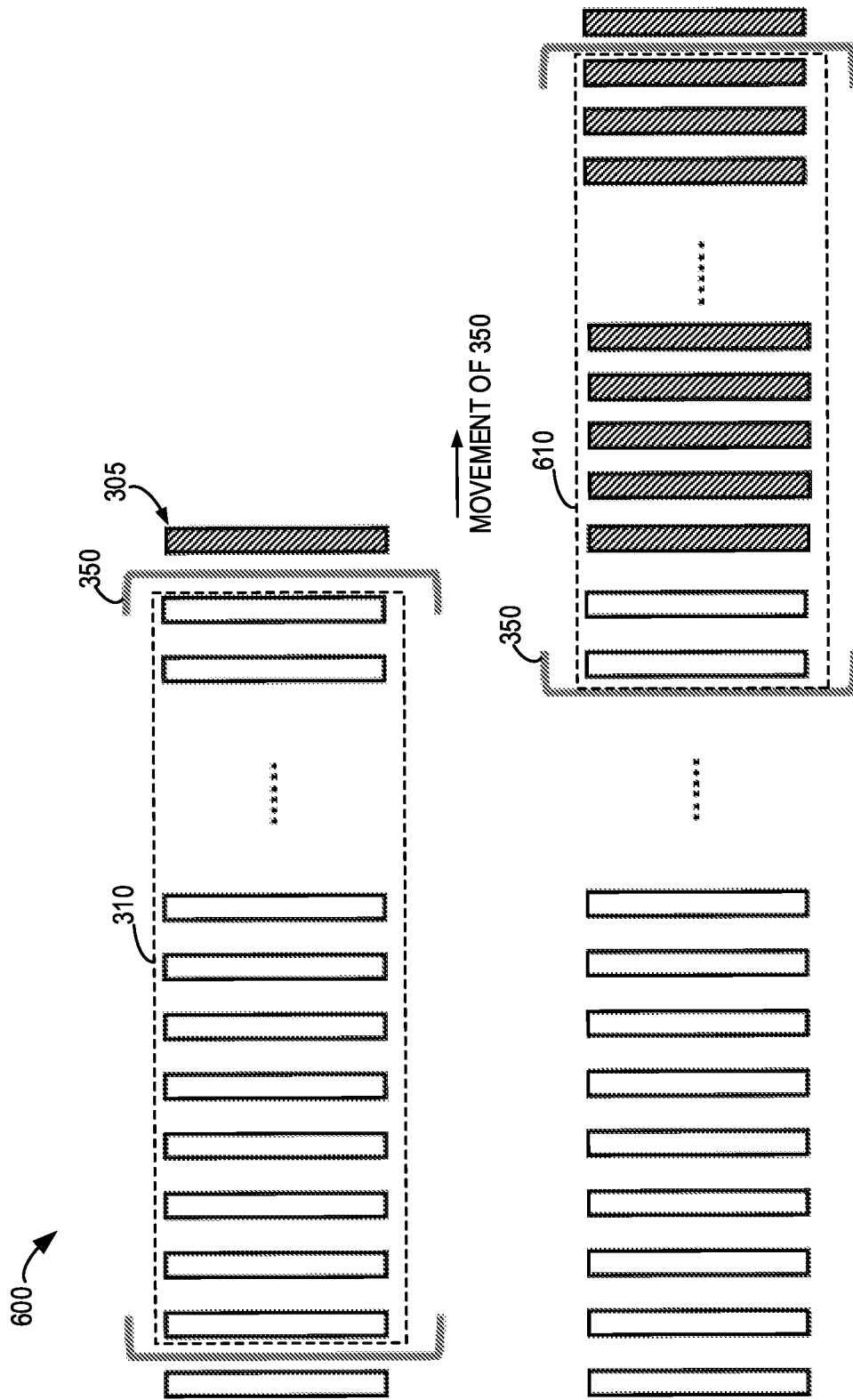


Fig. 6

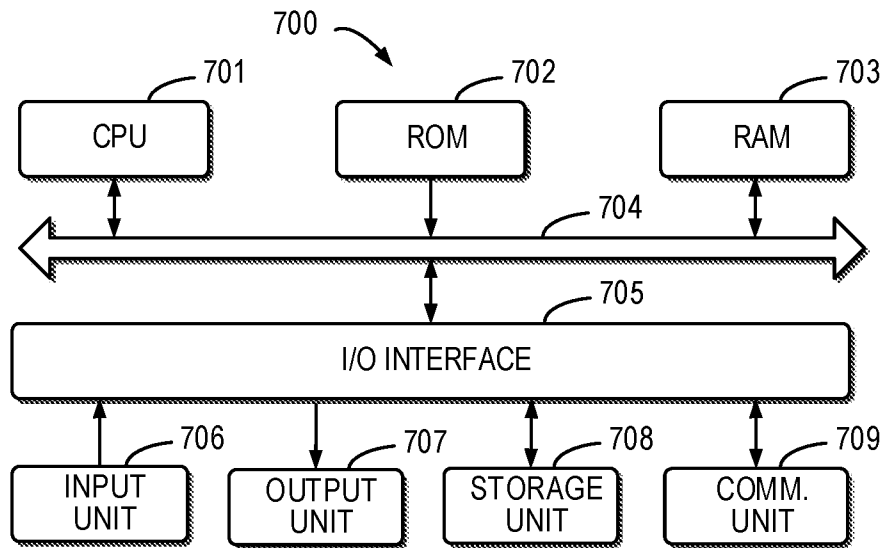


Fig. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/093172

A. CLASSIFICATION OF SUBJECT MATTER		
H04R 1/10(2006.01)i; H04N 21/433(2011.01)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)		
H04R H04N G11B		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
CNKI,CNPAT,WPLEPODOC: environment, sound, voice, audio, reference, feature, time, period, plurality, noise, difference, target, control		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 108429955 A (NANJING ZHUODA KECHUANG ACCELERATOR CO., LTD.) 21 August 2018 (2018-08-21) description, paragraphs [0046]-[0095]	1-24
X	CN 106792253 A (GUANGZHOU SHIYUAN ELECTRONICS TECHNOLOGY CO., LTD.) 31 May 2017 (2017-05-31) description, paragraphs [0022]-[0071]	1-24
X	CN 103137155 A (YINGSHUNYUAN SHANGHAI TECHNOLOGY CO., LTD. et al.) 05 June 2013 (2013-06-05) description, paragraphs [0022]-[0042]	1-24
A	CN 106060643 A (LETV HOLDING BEIJING CO., LTD. et al.) 26 October 2016 (2016-10-26) the whole document	1-24
A	WO 2016074438 A1 (ZTE CORPORATION) 19 May 2016 (2016-05-19) the whole document	1-24
A	US 2015117652 A1 (TOYOTA JIDOSHA KABUSHIKI KAISHA) 30 April 2015 (2015-04-30) the whole document	1-24
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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		Date of mailing of the international search report
08 February 2021		03 March 2021
Name and mailing address of the ISA/CN		Authorized officer
National Intellectual Property Administration, PRC 6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088 China		CHEN, Wenjing
Facsimile No. (86-10)62019451		Telephone No. 86-(10)-53961739

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/093172

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108429955	A	21 August 2018	None			
CN	106792253	A	31 May 2017	None			
CN	103137155	A	05 June 2013	None			
CN	106060643	A	26 October 2016	None			
WO	2016074438	A1	19 May 2016	CN	105657137	A	08 June 2016
US	2015117652	A1	30 April 2015	CN	104380378	A	25 February 2015
				EP	2858068	A1	08 April 2015
				JP	WO2013179464	A1	14 January 2016
				WO	2013179464	A1	05 December 2013