



(51) International Patent Classification:

G10L 21/0208 (2013.01) H03G 3/32 (2006.01)
H04M 3/00 (2024.01) G10L 25/60 (2013.01)
H04M 9/08 (2006.01) G10L 21/0364 (2013.01)

(21) International Application Number:

PCT/US2024/036469

(22) International Filing Date:

01 July 2024 (01.07.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/512,424 07 July 2023 (07.07.2023) US
63/635,570 17 April 2024 (17.04.2024) US

(71) Applicant: **DOLBY LABORATORIES LICENSING CORPORATION** [US/US]; 1275 Market Street, San Francisco, California 94103 (US).

(72) Inventors: **WANG, Ning**; c/o Dolby Laboratories, Inc., 1275 Market Street, San Francisco, California 94103 (US). **THASARATHAN, Shanush Prema**; c/o Dolby Laboratories, Inc., 1275 Market Street, San Francisco, California 94103 (US). **CHO, Byung Hoon**; c/o Dolby Laboratories,

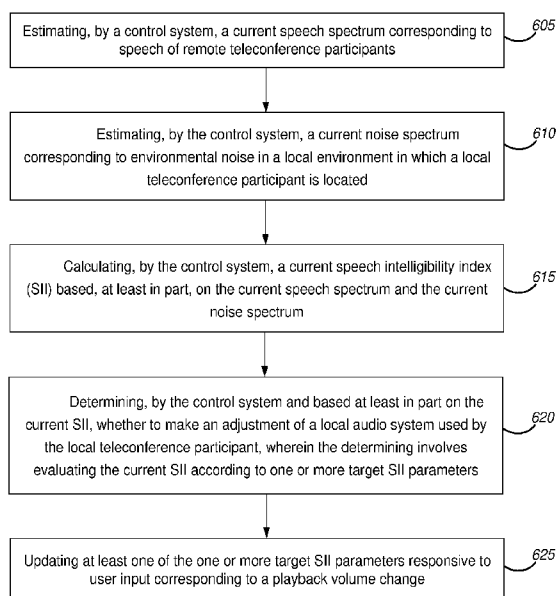
Inc., 1275 Market Street, San Francisco, California 94103 (US).

(74) Agent: **ANDERSEN, Robert L.** et al.; Dolby Laboratories Licensing Corporation, Intellectual Property Group, 1275 Market Street, San Francisco, California 94103 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,

(54) Title: ENVIRONMENTAL NOISE COMPENSATION IN TELECONFERENCING



600

Figure 6

(57) Abstract: A method of compensating for environmental noise during a teleconference may involve: estimating, by a control system, a current speech spectrum corresponding to speech of remote teleconference participants; estimating, by the control system, a current noise spectrum corresponding to environmental noise in a local environment in which a local teleconference participant is located; calculating, by the control system, a current speech intelligibility index (SII) based, at least in part, on the current speech spectrum and the current noise spectrum; determining, by the control system and based at least in part on the current SII, whether to make an adjustment of a local audio system used by the local teleconference participant, wherein the determining involves evaluating the current SII according to one or more target SII parameters; and updating at least one of the one or more target SII parameters responsive to user input corresponding to a playback volume change.

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))*

ENVIRONMENTAL NOISE COMPENSATION IN TELECONFERENCING

CROSS REFERENCE TO RELATED APPLICATIONS

- 5 **[0001]** The present application claims priority to U.S. Provisional Patent Application No. 63/512,424, filed on July 7, 2023, and U.S. Provisional Patent Application No. 63/635,570, filed on April 17, 2024, all of which are incorporated by reference in their entirety.

TECHNICAL FIELD

- 10 **[0002]** This disclosure pertains to devices, systems and methods for environmental noise compensation (ENC), particularly ENC in the teleconferencing context. As used herein, the term “teleconferencing” encompasses both audio/video teleconferencing and audio teleconferencing.

BACKGROUND

- 15 **[0003]** Teleconferencing has become an important part of modern life. The ability to communicate clearly while teleconferencing is based mainly on speech intelligibility, which in turn is based in part on the presence or absence of noise in the audio signal(s). Although existing devices, systems and methods for speech intelligibility estimation, noise estimation and ENC provide benefits, improved
20 systems and methods would be desirable.

SUMMARY

- [0004]** At least some aspects of the present disclosure may be implemented via methods. Some such methods involve compensating for environmental noise during a teleconference. For example, some methods may involve estimating, by
25 a control system, a current speech spectrum corresponding to speech of remote teleconference participants and estimating, by the control system, a current noise spectrum corresponding to environmental noise in a local environment in which a local teleconference participant is located. Some methods may involve calculating, by the control system, a current speech intelligibility index (SII)
30 based, at least in part, on the current speech spectrum and the current noise spectrum. Some methods may involve determining, by the control system and

based at least in part on the current SII, whether to make an adjustment of a local audio system used by the local teleconference participant. Some methods may involve updating, by the control system, at least one of the one or more target SII parameters responsive to user input corresponding to a playback volume change.

- 5 [0005] The determining may involve evaluating the current SII according to one or more target SII parameters. In some examples, the determining may involve determining whether the current SII is within a target SII range. Some methods may involve adjusting, by the control system, at least a portion of the local audio system responsive to determining that the adjustment should be made.
- 10 [0006] Some methods may involve determining a confidence value corresponding to the current noise spectrum. The confidence value may, for example, indicate a likelihood of a current input audio frame corresponding mainly to ambient noise. According to some examples, the confidence value may be a broadband confidence value. Determining whether to make the adjustment
- 15 of the local audio system may be based, at least in part, on the broadband confidence value. Some methods may involve determining a band-based confidence value for each frequency band of a plurality of frequency bands. In some examples, determining whether to make the adjustment of the local audio system may involve determining whether to update one or more noise statistics
- 20 for each frequency band based, at least in part, on the band-based confidence value. Some methods may involve determining whether to update one or more noise statistics based, at least in part, on the confidence value.
- [0007] In some examples, estimating the current noise spectrum may involve estimating an echo coupling gain corresponding to local loudspeaker playback
- 25 captured by a local microphone system. In some such examples, estimating the echo coupling gain may involve determining a maximum band loudspeaker reference power for each frequency band of a current audio frame and previous N-1 audio frames, where N is an integer corresponding to a number of audio frames in a delay line. In some such examples, estimating the echo coupling gain
- 30 may involve tracking a minimum power for each frequency band of an input microphone signal and estimating an echo coupling gain for each frequency band of the input microphone signal based, at least in part, on the maximum band loudspeaker reference power and the minimum power. Some methods may

involve determining whether to update an estimated echo coupling gain based, at least in part, on whether a current estimated echo coupling gain value has changed more than a threshold amount from a most recent estimated echo coupling gain value, on whether an estimated echo coupling gain has been
5 updated within a threshold time interval, or both.

[0008] Some methods may involve, responsive to determining that the adjustment should be made, adjusting at least a portion of the local audio system to maintain the current SII within a target SII range. Some methods may involve, responsive to determining that the adjustment should be made, adjusting at least a
10 portion of the local audio system to maintain the current SII within a target SII range. According to some examples, maintaining the current SII within a target SII range may involve increasing a loudspeaker playback volume of one or more audio frames until the current SII is greater than a first target SII and less than a high target SII. The first target SII may, in some examples, be greater than a
15 median target SII and less than the high target SII. In some examples, maintaining the current SII within a target SII range may involve decreasing a loudspeaker playback volume of one or more audio frames until the current SII is less than a second target SII and greater than a low target SII. According to some examples, the second target SII may be less than a median target SII and greater
20 than the low target SII.

[0009] Some or all of the operations, functions and/or methods described herein may be performed by one or more devices according to instructions (e.g., software) stored on one or more non-transitory media. Such non-transitory media may include memory devices such as those described herein, including but not
25 limited to random access memory (RAM) devices, read-only memory (ROM) devices, etc. Accordingly, some innovative aspects of the subject matter described in this disclosure can be implemented in a non-transitory medium having software stored thereon.

[0010] At least some aspects of the present disclosure may be implemented via
30 apparatus. For example, one or more devices may be capable of performing, at least in part, the methods disclosed herein. In some implementations, an apparatus may include an interface system and a control system. The control system may include one or more general purpose single- or multi-chip

processors, digital signal processors (DSPs), application specific integrated circuits (ASICs), field programmable gate arrays (FPGAs) or other programmable logic devices, discrete gates or transistor logic, discrete hardware components, or combinations thereof. In some examples, the apparatus may be one of the above-referenced audio devices. However, in some implementations the apparatus may be another type of device, such as a mobile device, a laptop, a server, etc.

[0011] Details of one or more implementations of the subject matter described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages will become apparent from the description, the drawings, and the claims. Note that the relative dimensions of the following figures may not be drawn to scale.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Figure 1A shows examples of sound sources that may be captured by a local microphone during a teleconference.

[0013] Figure 1B is a block diagram that shows examples of components of an apparatus capable of implementing various aspects of this disclosure.

[0014] Figure 1C is a flow diagram that outlines an example of a method that may be performed by an apparatus or system such as those disclosed herein.

[0015] Figure 2A shows example blocks of a novel noise estimator according to some examples.

[0016] Figure 2B shows blocks of the main noise estimator of Figure 2A according to some disclosed implementations.

[0017] Figure 3A shows blocks of the echo coupling gain estimator of Figure 2B according to some disclosed implementations.

[0018] Figure 3B shows a plot of estimated SII over time according to one example.

[0019] Figure 4 shows a plot of estimated SII over time according to another example.

[0020] Figure 5 shows another plot of estimated SII over time.

[0021] Figure 6 is a flow diagram that outlines another example of a method that may be performed by an apparatus or system such as those disclosed herein.

[0022] Like reference numbers and designations in the various drawings indicate like elements.

DETAILED DESCRIPTION

[0023] Figure 1A shows examples of sound sources that may be captured by a local microphone during a teleconference. In these examples, the local speech 102 of a local teleconference participant, the environmental noise 104—also referred to herein as “background noise” or “ambient noise”—and the loudspeaker playback sounds 106—also referred to herein as “echo”—from the loudspeaker 108 are captured by microphone 112. During a teleconference, the speech of remote teleconference participants—also referred to herein as “far-end speech”—is played back by the loudspeaker 108. The term “remote teleconference participants” refers to teleconference participants who are in locations other than that of the local teleconference participant.

[0024] Figure 1A also shows an example of a noise estimator 114, examples of which are disclosed herein. The noise estimator 114 may, for example, be implemented via an instance of the control system 110 that is described with reference to Figure 1B.

[0025] The speech of remote teleconference participants and the local speech 102 of a local teleconference participant are usually time divided. Instances of “double talk,” during which the far-end speech and local speech 102 overlap in time, are atypical. Most of the time, double talk only happens when teleconference participants want to interrupt. To prevent recaptured far-end speech being transmitted back to the far end, echo management is normally employed in the signal chain, and the echo signal 106 is often mostly suppressed.

[0026] Reliable noise estimation and noise reduction techniques can help meeting participants to better understand what is being said by other teleconference participants. Various improved noise estimation and noise reduction techniques are disclosed herein.

[0027] Figure 1B is a block diagram that shows examples of components of an apparatus capable of implementing various aspects of this disclosure. As with other figures provided herein, the types and numbers of elements shown in Figure 1B are merely provided by way of example. Other implementations may include

more elements, fewer elements, different types of elements, or combinations thereof. According to some examples, the apparatus 101 may be, or may include, a device that is configured for performing at least some of the methods disclosed herein, such as a smart audio device, a laptop computer, a cellular telephone, a
5 tablet device, a smart home hub, etc. In some such implementations the apparatus 101 may be, or may include, a server that is configured for performing at least some of the methods disclosed herein.

[0028] In this example, the apparatus 101 includes an interface system 105 and a control system 110. In some implementations, the control system 110 may be
10 configured for performing, at least in part, the methods disclosed herein. The control system 110 may, in some implementations, be configured for compensating for environmental noise during a teleconference.

[0029] In some examples, the control system 110 may be configured for estimating a current speech spectrum corresponding to speech of remote
15 teleconference participants. According to some examples, the control system 110 may be configured for estimating a current noise spectrum corresponding to environmental noise in a local environment in which a local teleconference participant is located. In some examples, the control system 110 may be configured for calculating a current speech intelligibility index (SII) based, at
20 least in part, on the current speech spectrum and the current noise spectrum. In some examples, the control system 110 may be configured for determining, based at least in part on the current SII, whether to make an adjustment of a local audio system used by the local teleconference participant. In some examples, the determining may involve evaluating the current SII according to one or more
25 target SII parameters. According to some examples, the control system 110 may be configured for updating at least one of the one or more target SII parameters responsive to user input corresponding to a playback volume change.

[0030] The interface system 105 may include one or more network interfaces and/or one or more external device interfaces (such as one or more universal
30 serial bus (USB) interfaces). According to some implementations, the interface system 105 may include one or more wireless interfaces. The interface system 105 may include one or more devices for implementing a user interface, such as one or more microphones, one or more speakers, a display system, a touch sensor

system and/or a gesture sensor system. In some examples, the interface system 105 may include one or more interfaces between the control system 110 and a memory system, such as the optional memory system 115 shown in Figure 1B. However, the control system 110 may include a memory system in some instances.

[0031] The control system 110 may, for example, include a general purpose single- or multi-chip processor, a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array (FPGA) or other programmable logic device, discrete gate or transistor logic, and/or discrete hardware components.

[0032] In some implementations, the control system 110 may reside in more than one device. For example, a portion of the control system 110 may reside in a device within an environment (such as a laptop computer, a tablet computer, a smart audio device, etc.) and another portion of the control system 110 may reside in a device that is outside the environment, such as a server. In other examples, a portion of the control system 110 may reside in a device within an environment and another portion of the control system 110 may reside in one or more other devices of the environment.

[0033] Some or all of the methods described herein may be performed by one or more devices according to instructions (e.g., software) stored on one or more non-transitory media. Such non-transitory media may include memory devices such as those described herein, including but not limited to random access memory (RAM) devices, read-only memory (ROM) devices, etc. The one or more non-transitory media may, for example, reside in the optional memory system 115 shown in Figure 1B and/or in the control system 110. Accordingly, various innovative aspects of the subject matter described in this disclosure can be implemented in one or more non-transitory media having software stored thereon. The software may, for example, include instructions for controlling at least one device to process audio data. The software may, for example, be executable by one or more components of a control system such as the control system 110 of Figure 1B.

[0034] In some examples, the apparatus 101 may include the optional microphone system 120 shown in Figure 1B. The optional microphone system

120 may include one or more microphones. In some implementations, one or more of the microphones may be part of, or associated with, another device, such as a loudspeaker, a smart audio device, etc.

5 [0035] According to some implementations, the apparatus 101 may include the optional loudspeaker system 125 shown in Figure 1B. The optional loudspeaker system 125 may include one or more loudspeakers. Loudspeakers may sometimes be referred to herein as “speakers.” In some examples, at least some loudspeakers of the optional loudspeaker system 125 may be arbitrarily located . For example, at least some speakers of the optional loudspeaker system 125 may
10 be placed in locations that do not correspond to any standard prescribed speaker layout, such as Dolby 5.1, Dolby 5.1.2, Dolby 7.1, Dolby 7.1.4, Dolby 9.1, Hamasaki 22.2, etc. In some such examples, at least some loudspeakers of the optional loudspeaker system 125 may be placed in locations that are convenient to the space (e.g., in locations where there is space to accommodate the
15 loudspeakers), but not in any standard prescribed loudspeaker layout.

[0036] In some implementations, the apparatus 101 may include the optional sensor system 130 shown in Figure 1B. The optional sensor system 130 may include a touch sensor system, a gesture sensor system, one or more cameras, etc.

20 [0037] In some implementations, the apparatus 101 may include the optional display system 135 shown in Figure 1B. The optional display system 135 may include one or more displays, such as one or more light-emitting diode (LED) displays. In some instances, the optional display system 135 may include one or more organic light-emitting diode (OLED) displays. In some examples wherein the apparatus 101 includes the display system 135, the sensor system 130 may
25 include a touch sensor system and/or a gesture sensor system proximate one or more displays of the display system 135. According to some such implementations, the control system 110 may be configured for controlling the display system 135 to present a graphical user interface (GUI), such as a GUI related to implementing one of the methods disclosed herein.

30 [0038] As noted above, reliable noise estimation and noise reduction techniques can help meeting participants to better understand what is being said by other teleconference participants. Recent developments that involve neural-network-based approaches have helped to resolve the noise reduction problem. Neural-

network-based methods are often capable of removing many types of noise, provided that the neural network training process noise has provided the neural network with sufficient exposure to each particular type of noise that is to be removed. However, it may be difficult or even impossible to train a neural
5 network to remove every possible type of noise.

[0039] Accordingly, noise compensation is still an important aspect of providing audio that is acceptable for teleconferencing. The basic principle of noise compensation is to adjust the playback volume up and down based on the ambient noise sensed by the microphone. In some examples, noise compensation
10 may be applied on a per-band basis, with different gains applied to different frequency bands.

[0040] In the context of noise compensation, a robust and stable stationary noise estimator is important. The quality of the stationary noise estimator generally determines the overall system performance. The basic functions of a stationary
15 noise estimator include:

- The stationary noise estimator should only estimate “stationary” noise to prevent the volume be adjusted up or down rapidly in the case of occasional (dynamic) noise, e.g., a dog’s barking, sounds caused by tapping on a table, sounds caused by keyboard strokes, etc. As used
20 herein, the term “stationary noise” does not refer to noise types having “strict-sense stationarity,” in which the statistical characteristics do not change over time, but instead refer to noise types having “wide-sense stationarity,” in which the first moment and autocovariance do not vary with respect to time and in which the second moment is finite for all
25 times. Stationary noise is one type of “stationary process,” which is a stochastic process whose unconditional joint probability distribution does not change when shifted in time.
- The stationary noise estimator should only estimate “true” ambient noise of the local environment, not the noise in the echo signal that is played
30 back. For example, referring to Figure 1A, The stationary noise estimator should only estimate the environmental noise 104, not the loudspeaker playback or “echo” sounds 106. If the noise estimated is actually caused primarily by echo (echo noise is dominant), then the environmental noise

compensation (ENC system) will form a positive feedback loop, which will cause any gain adjustment to result in the maximum or the minimum volume setting.

[0041] Figure 1C is a flow diagram that outlines an example of a method that may be performed by an apparatus or system such as those disclosed herein. The blocks of method 140, like other methods described herein, are not necessarily performed in the order indicated. In some implementations, one or more of the blocks of method 140 may be performed concurrently. Moreover, some implementations of method 140 may include more or fewer blocks than shown and/or described.

[0042] According to this example, method 140 is a method of compensating for environmental noise during a teleconference. The blocks of method 140 may be performed by one or more devices, which may be (or may include) a control system such as the control system 110 that is shown in Figure 1B and described above. According to this example, the blocks of method 140 are repeated for each new block of audio data.

[0043] In this example, method 140 involves calculating the current speech intelligibility index (SII) and making an incremental change in the system—such as a playback volume change—if necessary, such that the system stays within a speech intelligibility index target. The system may include a device that is configured to provide a teleconference. In some examples, the system may include a laptop, having a loudspeaker and a microphone, that is configured to provide a teleconference and is being used by a local participant during the teleconference. The output of the loudspeaker during the teleconference is one example of what is referred to as the “speech of remote teleconference participants” in this document. Referring again to Figure 1A, the audio captured by the local microphone will include the environmental noise 104 and the speech 102 of the local teleconference participant. For a noise detector, the speech 102 may be regarded as a type of interference.

[0044] According to the example shown in Figure 1C, block 145 involves obtaining a frame of audio data from a local microphone. In this example, block 150 involves estimating a speech spectrum corresponding to the speech 102 in the current frame of audio data from the microphone. According to this example,

block 155 involves calculating the current noise spectrum corresponding to the environmental noise 104 for the current frame of audio data from the microphone. In some examples, blocks 150 and 155 may be performed concurrently. According to some examples, the noise spectrum is based in part
5 on historical data. In some examples, block 155 will keep updating the noise spectrum as new data frames are received. In some examples, the current noise spectrum determined by block 155 is for the current frame of audio data from the microphone and is used to determine SII and other parameters. In this example, block 160 involves calculating the current SII, which is the SII corresponding to
10 the current frame of audio data from the microphone. According to this example, block 165 involves determining, based at least in part on the current SII, whether to make an adjustment of a local audio system used by the local teleconference participant.

[0045] In some examples, block 150 may involve estimating the speech spectrum
15 as described in *Methods for Calculation of the Speech Intelligibility Index*, which was originally published in 1969 and was revised in 1997 by the American National Standards Institute and the Acoustical Society of America (hereinafter ASA/ANSI S3.5-1997), which is hereby incorporated by reference. For example, block 150 may involve estimating the speech spectrum as described in the
20 “Methods for determining input variables for SII calculation” sections on pages 11–13 of ASA/ANSI S3.5-1997. However, in some alternative examples, block 150 may involve estimating the speech spectrum according to one or more other methods.

[0046] In some alternative examples, block 150 may involve estimating the
25 speech spectrum according to a modified version of what is described in ASA/ANSI S3.5-1997. Some such examples involve using the “insertion gain” somewhat differently from what is described in ASA/ANSI S3.5-1997. In ASA/ANSI S3.5-1997, for an amplification or attenuation device worn by a listener, at a specific frequency the insertion gain is the difference in decibels
30 between the pure-tone sound pressure level (SPL) at the eardrum with the amplification/attenuation device in place and the pure-tone SPL at the eardrum with the amplification/attenuation device removed. The insertion gain is used to calculate the equivalent speech spectrum level (see section 5.1.3, starting on page 12 of ASA/ANSI S3.5-1997) and is also used to calculate the equivalent noise

spectrum level (see section 5.1.4, on page 13 of ASA/ANSI S3.5-1997). For example, according to ASA/ANSI S3.5-1997, if the speech spectrum is measured at the center of a listener's head, the equivalent speech spectrum for a particular frequency band is calculated as the measured speech spectrum level for the particular frequency band plus the insertion gain for the particular frequency band.

[0047] Some disclosed examples involve mapping the equivalent SPL of the playback to this insertion gain. In some examples, the current hardware/software gain that is applied to the loudspeaker is converted to an equivalent SPL at the listener position. According to some examples, this conversion is aided by a tuning parameter during the tuning process for a particular device, such as a particular laptop model. In some examples, this equivalent SPL is subtracted by the nominal SPL associated with the "normal" speech spectrum. The final subtracted value is used as the insertion gain, with the same value used for all frequency bands.

[0048] In some examples, block 155 may involve calculating the current noise spectrum corresponding to the environmental noise 104 in the current frame of audio data from the microphone according to the methods described in ASA/ANSI S3.5-1997, for example on page 13. However, in some disclosed examples, block 155 may involve calculating the current noise spectrum according to alternative methods.

[0049] According to some such methods, block 155 involves calculating the current noise spectrum corresponding to a novel noise estimator (NE), examples of which are described in more detail with reference to Figures 2A–3A. Some implementations of the new noise estimator include at least a first or main noise estimator and a second, auxiliary noise estimator. The second noise estimator may be configured to respond to noise level changes relatively faster than the first or main noise estimator.

[0050] In some examples block 160 may involve determining, based on the current speech spectrum estimated by block 150 and the current noise spectrum estimated by block 155, the current SII as a value between 0 and 1, where 1 means very intelligible and 0 means not intelligible at all. In some examples, block 160 may involve determining the current SII based on the current speech

spectrum, the current noise spectrum and a measured or assumed hearing threshold level.

[0051] According to some such examples, the SII that is calculated in block 160 may be the SII metric described in ASA/ANSI S3.5-1997. In some such
 5 examples, the SII metric may be calculated as described in the “Methods for calculating Speech Intelligibility Index” section on pages 9–11 of ASA/ANSI S3.5-1997, which is specifically incorporated herein by reference.

[0052] According to some examples, the current SII may be determined as follows:

$$10 \quad Sii[current] = sii[current-1] * sii_alpha + instantaneous_sii * (1 - sii_alpha)$$

In the foregoing equation, $sii[current]$ represents the current smoothed SII value, $sii[current-1]$ represents the smoothed SII value calculated in the previous block, Sii_alpha represents a tuning parameter, and $Instantaneous_sii$ represents the
 immediate output of metric described in ASA/ANSI S3.5-1997. In one example,
 15 the tuning parameter Sii_alpha may be $2^{(-0.02/0.15)}$, where 0.02 represents the block length in seconds. Other examples may use larger or smaller tuning parameters.

[0053] Figure 2A shows example blocks of a novel noise estimator according to some examples. In this example, the noise estimator 200 includes a first noise
 20 estimator 210—also referred to herein as a main noise estimator 210—a second noise estimator 220—also referred to herein as an auxiliary noise estimator 220—and a fast tracking flag module 230. According to this example, the main noise estimator 210, the auxiliary noise estimator 220 and the fast tracking flag module 230 are implemented by an instance of the control system 110 that is described
 25 with reference to Figure 1B. As with other figures provided herein, the types and numbers of elements shown in Figure 2A are merely provided by way of example. Other implementations may include more elements, fewer elements, different types of elements, different arrangements of elements, or combinations thereof.

30 [0054] In the examples shown in Figure 2A, the main noise estimator 210 and the auxiliary noise estimator 220 are both configured to receive loudspeaker audio data 222 and microphone audio data 224, for example as described with reference

to block 145 of Figure 1C. In some examples, the loudspeaker audio data 222 may be, or may include, a reference signal corresponding to what is being played back by a local loudspeaker.

[0055] According to this example, the main noise estimator 210 is configured to
5 determine and output a noise estimate 208 and a confidence metric 207—also referred to herein as a “confidence value”—based at least in part on the loudspeaker audio data 222 and the microphone audio data 224. In this example, the noise estimate 208 and the confidence metric 207 correspond to a current noise spectrum, which in turn corresponds to a current input audio frame of the
10 microphone audio data 224. The confidence metric 207 indicates a likelihood of the current input audio frame corresponding mainly to ambient noise, such as the environmental noise 104 that is described with reference to Figure 1A. In some examples, the main noise estimator 210 may be configured to determine the noise estimate 208, the confidence metric 207, or both, based at least in part on fast
15 tracking flags 228 from the fast tracking flag module 230. Example blocks and functionalities of the main noise estimator 210 are described in more detail below with reference to Figures 2B and 3A.

[0056] In this example, the auxiliary noise estimator 220 is configured to determine and output a noise estimate 226 based at least in part on the
20 loudspeaker audio data 222 and the microphone audio data 224. Here, noise estimate 226 corresponds to the current input audio frame of the microphone audio data 224. According to some disclosed examples, the auxiliary noise estimator 220 is configured to respond to noise level changes—and to produce a responsive noise estimate 226—relatively faster than the main noise estimator
25 210, thereby increasing the response speed (decreasing the response time) of the noise estimator 200.

[0057] According to this example, the fast tracking flag module 230 is configured to determine and output fast tracking flags 228, based at least in part on the noise estimates 226 and 208. In some examples, the main noise estimator
30 210 may be configured to converge to a new noise level relatively faster if the main noise estimator 210 has received a fast tracking flag 228.

[0058] In some examples, if the noise estimate 226 from the auxiliary noise estimator 220 indicates that the power of the wideband noise is greater than the

power of the wideband noise indicated by the noise estimate 208 from the main noise estimator 210 by a threshold and for a determined time interval, the fast tracking flag module 230 is configured to set a value of the fast tracking flag 228 to a maximum value, for example to 1. In other words, the fast tracking flag module 230 may be configured to set a value of the fast tracking flag 228 to a maximum value, for example to 1, upon determining that the difference between (1) the noise power estimated by the auxiliary noise estimator 220 and (2) the noise power estimated by the main noise estimator 210 is greater than a difference threshold for at least a first time threshold. According to some examples, upon determining that the difference between (1) and (2) is less than the difference threshold for at least a second time threshold, the fast tracking flag module 230 may be configured to set a value of the fast tracking flag 228 to a minimum value, for example to 0. The first time threshold may or may not be equal to the second time threshold. In some examples, the power threshold may be in the range from 1dB to 5dB, e.g., 1dB, 2dB, 3dB, 4dB or 5dB. According to some examples, the time constant may be in the range from 0.5 seconds to 2 seconds, e.g., 0.5 seconds, 0.6 seconds, 0.7 seconds, 0.8 seconds, 0.9 seconds, 1.0 seconds, 1.1 seconds, 1.2 seconds, 1.3 seconds, 1.4 seconds, 1.5 seconds, 1.6 seconds, 1.7 seconds, 1.8 seconds, 1.9 seconds or 2 seconds.

[0059] According to some examples, the auxiliary noise estimator 220 may be configured with different tuning parameters than those of the main noise estimator 210. In some alternative examples, the auxiliary noise estimator 220 may be based on a deep neural network. Such noise estimators may have a very fast response.

[0060] Figure 2B shows blocks of the main noise estimator of Figure 2A according to some disclosed implementations. In this example, the main noise estimator 210 includes a voice activity detector (VAD) 201, an echo coupling gain estimator 202, a confidence calculation block 203, and a noise and confidence statistics calculation block 204. According to this example, the VAD 201, the echo coupling gain estimator 202, the confidence calculation block 203, and the noise and confidence statistics calculation block 204 are implemented by an instance of the control system 110 that is described with reference to Figure 1B. As with other figures provided herein, the types and numbers of elements shown in Figure 2B are merely provided by way of example. Other

implementations may include more elements, fewer elements, different types of elements, different arrangements of elements, or combinations thereof.

[0061] In this example, the VAD 201 is configured to detect local speech activity based on the microphone audio data 224 and to send speech activity signals 211 to the echo coupling gain estimator 202, the confidence calculation block 203 and the noise and confidence statistics calculation block 204. According to some examples, only when the speech activity signals 211 indicate that there is no local speech activity are the echo coupling gain estimator 202, the confidence calculation block 203 and the noise and confidence statistics calculation block 204 active.

[0062] According to this example, the echo coupling gain estimator 202 is configured to estimate, based on the microphone audio data 224 and the loudspeaker audio data 222, coupling gains from loudspeaker playback to microphone capturing for each of a plurality of frequency bands and to provide corresponding coupling gain estimations 205 to the confidence calculation block 203. In other words, with reference to scenario depicted in Figure 1A, the echo coupling gain estimator 202 is configured to estimate the contribution of the echo 106 to the sounds detected by the microphone 112 and that are present in the microphone audio data 224 and to provide corresponding coupling gain estimations 205. In the example shown in Figure 2B, the loudspeaker audio data 222 is, or includes, reference signals corresponding to what is being played back by a local loudspeaker. Figure 3A shows more details of the echo coupling gain estimator 202 according to some examples.

[0063] In this example, the confidence calculation block 203 is configured to calculate, for each frequency band, the likelihood that ambient noise is dominant in the current input audio frame of the microphone audio data 224. According to this example, the confidence calculation block 203 is configured to output confidence data 206 to the noise and confidence statistics calculation block 204. In this example, the confidence data 206 includes at least a confidence value corresponding to the current noise spectrum. Here, the confidence value indicates a likelihood of a current input audio frame corresponding mainly to ambient noise. According to this example, the confidence data 206 also includes a binary value (e.g., either 0 or 1) that indicates whether the current input audio

frame is likely to be primarily ambient noise and should be processed accordingly.

[0064] According to this example, the noise and confidence statistics calculation block 204 is configured to determine and output a noise estimate 208 and a confidence metric 207, based at least in part on the microphone audio data 224 and the confidence data 206. The confidence metric 207 indicates a likelihood of the current input audio frame corresponding mainly to ambient noise, such as the environmental noise 104 that is described with reference to Figure 1A. In some examples, the noise and confidence statistics calculation block 204 may be configured to determine the noise estimate 208, the confidence metric 207, or both, based at least in part on fast tracking flags 228 (not shown in Figure 2B) from the fast tracking flag module 230 of Figure 2A. Detailed examples of how the noise and confidence statistics calculation block 204 may function are provided below.

[0065] Figure 3A shows blocks of the echo coupling gain estimator of Figure 2B according to some disclosed implementations. In this example, the echo coupling gain estimator 202 includes a delay line module 301, a minimum follower 302, a threshold detector 303, a maximum follower 304, an update control block 305 and a subtraction node 312. According to this example, the delay line module 301, the minimum follower 302, the threshold detector 303, the maximum follower 304, the update control block 305 and the subtraction node 312 are implemented by an instance of the control system 110 that is described with reference to Figure 1B. According to this example, the echo coupling gain estimator 202 only operates when there is no local speech activity, as determined by the VAD 201 (see Figure 3A). As with other figures provided herein, the types and numbers of elements shown in Figure 3A are merely provided by way of example. Other implementations may include more elements, fewer elements, different types of elements, different arrangements of elements, or combinations thereof.

[0066] In the example shown in Figure 3A, the delay line module 301 is a length N delay line and is configured to receive the loudspeaker audio data 222—which is, or includes, loudspeaker reference signals corresponding to what is being played back by a local loudspeaker. In some examples, N may be in the range

from 8 to 16 frames, e.g., 8 frames, 9 frames, 10 frames, 11 frames, 12 frames, 13 frames, 14 frames, 15 frames, 16 frames, etc. According to some examples, each frame may be in the range from 10 milliseconds (ms) to 30 ms, e.g., 10 ms, 12 ms, 14 ms, 16 ms, 18 ms, 20 ms, 22 ms, 24 ms, 26 ms, 28 ms, 30 ms, etc.

- 5 According to this example, the input loudspeaker reference signals are in the form of frequency banded loudspeaker reference power. In this example, the delay line module 301 is configured to implement a maximum operation (MAX) across the entire delay line. The purpose of letting the loudspeaker reference signals go through a delay line is to compensate for the natural delay between the
- 10 loudspeaker reference signals and the corresponding microphone capture of the sounds played back by one or more local loudspeakers, in order to take into account the sound wave propagation delay, the electrical circuitry delay, the software buffering delay, etc., as well as the reverberant build-up that may be present in a local room. In this example, the delay line module 301 outputs
- 15 maximum power data 310, which indicates the maximum of the loudspeaker reference signal power for each frequency band for the current audio frame and for the previous N-1 audio frames.

- [0067] According to this example, the minimum follower 302 is configured to track the minimum power for each input frequency band for each frame of the
- 20 microphone audio data 224 and to output minimum power data 311, which indicates the minima of input power for each frequency band of each frame of the microphone audio data 224. In Figure 3A, the “B” indicates the number of frequency bands. In some examples, the time window size of the minimum follower 302 may be dynamically adjusted. For example, (referring again to
- 25 Figure 2A) in some implementations, when a fast-tracking flag 228 is input to the main noise estimator 210, the minimum follower 302 will shorten its window size in order to help the main noise estimator 210 to converge to a solution faster. In one such example, the minimum follower 302 may shorten its window size from 1.0 seconds to 250 ms. Other examples may involve different starting window
- 30 sizes, different shortened window sizes, or both.

[0068] In the example shown in Figure 3A, the threshold detector 303 is a simple threshold detector that is configured to determine when the power of the current frame of the input microphone audio data 224 is above a tracked minimum power level by at least a threshold amount and to provide corresponding threshold

detector output 313. The tracked minimum power level is highly dependent on the sensitivity of the particular microphone. Therefore, the tracked minimum power level range may vary widely from microphone to microphone. According to some examples, the threshold amount be in the range from 3dB to 10dB, e.g.,
5 3dB, 4dB, 5dB, 6dB, 7dB, 8dB, 9dB or 10dB. In some examples, the threshold detector output 313 of the threshold detector 303 has a value of 1 whenever the power of the current frame of the input microphone audio data 224 is above the tracked minimum by at least the threshold amount. In some such examples, the threshold detector output 313 has a value of 0 whenever the power of the current
10 frame of the input microphone audio data 224 is *not* above the tracked minimum by at least the threshold amount.

[0069] In this example, the subtraction node 312 is configured for subtracting (a) the input power of the current frame of input microphone audio data 224 from (b) the maximum power of the loudspeaker reference signals, as determined by the
15 maximum power data 310 output by the delay line module 301, to produce the subtraction node output 315. The subtraction node output 315 represents potential coupling gain estimations.

[0070] According to this example, the maximum follower 304 is configured to determine and output coupling gain estimations 205 based on the subtraction
20 node output 315 and update control signals 317 from the update control module 305.

[0071] In this example, the update control module 305 is configured to determine whether to disallow or allow the subtraction node output 315 to be output by the maximum follower 304 as the current coupling gain estimation 205. According
25 to some such examples, the update control signals 317 control whether the subtraction node output 315 will be received by the maximum follower 304 at all. Even though—according to this example—the echo coupling gain estimator 202 only operates when there is no local speech activity, it is nonetheless possible that local transient noise may still be included in the microphone signals 224.
30 Whenever there is some transient noise in a frequency band, the gain calculated by the subtraction node 312 will not be accurate and should be eliminated.

[0072] However, it can be challenging to determine whether a current input audio frame contains local transient noise or simply resulted from a powerful echo—in

other words, from loud loudspeaker playback—without prior knowledge of the coupling gain. And this coupling gain is what the echo coupling gain estimator 202 need to estimate here.

[0073] This problem can be overcome by using the fact that the coupling gain will generally not change abruptly and continuously for at least N frames, where N is the delay line length of the delay line module 301. Therefore, according to some examples, whenever there is a new maximum of gain, the control system 110 only allows the maximum follower 304 to track the new gain maximum if two conditions are satisfied:

1. The new gain maximum is within a certain range—such as within 1 dB, 2 dB, 3 dB, 4 dB, etc.—of a previously-tracked maximum gain. In some implementations, this condition may be ignored for the first update if there is no measured initial value (see below);
2. There is no new gain maximum within the previous N frames.

[0074] The above-mentioned problem also may be overcome by considering the fact that the coupling gain is primarily a characteristic of the device being used to participate in the teleconference, such as the phone, the laptop, the conferencing endpoint, etc. Although the coupling gain can be affected by the acoustic properties of a room in which the device resides, the coupling gain is mainly determined by the industrial design of the device itself. Once the product is manufactured, this coupling gain may be obtained and, in some examples, this coupling gain may be stored for future use as the initial value.

Examples of Confidence Calculation Block Functionality

[0075] This section includes examples of how the confidence calculation block 203 of Figure 2B may be implemented. Given the coupling gain and a current audio frame of the microphone audio data 224, the confidence calculation block 203 can calculate the likelihood of the current audio frame being predominantly ambient noise. Suppose the coupling gain (in dB) for frequency band b is g_b and the maximum reference power 310 is X_b (in dB). The estimated echo power $\widehat{E}_b(dB)$ in dB may be represented as follows:

$$\widehat{E}_b(dB) = X_b(dB) + g_b(dB)$$

[0076] The likelihood of the current audio frame of the microphone audio data 224, having power Y_b , being ambient noise may be represented as follows:

$$L_b = \frac{1.0}{1.0 + e^{-6.0(Y_b - \widehat{E}_b)/\sigma}}$$

[0077] In the foregoing equation, σ represents a threshold value, which may be 4 dB, 5 dB, 6 dB, 7 dB, 8 dB, etc.

[0078] According to some examples, the confidence calculation block 203 may determine and send an indication flag I_b —corresponding to frequency band b—to the noise and confidence statistics calculation block 204. In some such examples, the confidence calculation block 203 may determine the indication flag I_b by implementing the following set of conditions:

$$I_b = \begin{cases} 1, & \text{if } Y_b - \widehat{E}_b \geq \sigma \\ 1, & \text{if } n_b > Y_b \text{ and } \sum n_b \geq \sum Y_b \\ 0, & \text{otherwise} \end{cases}$$

[0079] In the foregoing equations, n_b represents the estimated noise mean for frequency band b. The second condition (if $n_b > Y_b$ and $\sum n_b \geq \sum Y_b$) allows noise statistics to be updated once the noise is gone. Even though the likelihood of current frame being ambient noise may be dubious, the level estimated may not be sensible under such conditions.

Noise and Confidence Statistics Calculations

[0080] This section includes examples of how the noise and confidence statistics calculation block 204 of Figure 2B may be implemented. In the example shown in Figure 3A, the input to the noise and confidence statistics calculation block 204 includes the minimum power data 311 from the minimum follower 302. The minimum power data 311 indicates the minima of input power for each frequency band of each frame of the microphone audio data 224. According to this example, the input to the noise and confidence statistics calculation block 204 also includes the microphone signal power, Y_b and the indication flag I_b —, both of which correspond to frequency band b.

[0081] To prevent transient noise being taken into control and to increase the estimation accuracy, in some examples the noise and confidence statistics calculation block 204 only takes Y_b into the accumulation if $I_b = 1$ and if the following condition is met:

$$Y_b \leq \underline{Y_b} + \beta$$

[0082] In the foregoing equation, $\underline{Y_b}$ represents the tracked minima of Y_b and β represents a threshold, such as 8 dB, 9 dB, 10 dB, 11 dB, 12 dB, etc. In some examples, only when noise accumulator is being updated, the confidence statistics are updated which takes L_b as input. According to some examples, the output of the the noise and confidence statistics calculation block 204 includes n_b , the estimated ambient noise power for each band, and p_b , the confidence value of corresponding to the current ambient noise power estimation.

Possible Uses of Output from the Noise and Confidence Statistics

10 *Calculation Block*

[0083] In some implementations, the confidence value p_b may be used to generate a wideband confidence flag P to control the behaviour of noise compensation control logic, for example as follows:

$$P = H(\sum p_b \cdot w_b)$$

15 In the foregoing equation, w_b represents a weighting factor wherein $\sum w_b = 1$, and H represents a hysteresis function output 0 or 1 with two hysteresis curve thresholds. According to some implementations, the compensation logic or mechanism(s) shall only be active when $P = 1$. In some examples, whenever $P = 0$ and the current volume (or/and other settings) are not at user's setting(s),
20 the compensation logic or mechanism(s) may cause the setting(s) to revert to the user's setting(s).

Determining and Applying System Adjustments

[0084] After the control system has estimated the current SII, in some implementations the control system will implement a multi-step process in order
25 to determine what the target SII is, whether change the system and, if so, how to change the system to achieve this target SII. In some examples, this process corresponds to block 165 of Figure 1C. Accordingly, this section describes various actions that may be performed according to various implementations of block 165.

30 [0085] In some examples, block 165 may calculate and apply the one or more changes to the system such that the SII calculated in block 160 of Figure 1C is

within a target SII range. Such changes may include, but are not limited to, the following:

- A change in the hardware/software gain of the speaker (e.g., volume control);
- 5 • An increase in the gain would map to an increase to SII. Similar logic for decrease;
- A change in the equalization (EQ);
- A change in the volume leveler effect of Dolby audio processing (DAP).

[0086] In some examples, such control(s) will only be implemented if the
10 wideband confidence indication is high (for example, 1), in which case there is a high confidence level that the estimated noise spectrum is indeed from background noise. If the confidence indicator is low (for example, 0) and any of the above controls are not at the user setting, in some examples the control system will cause the user setting(s) to be restored.

15 [0087] Following are various examples of how to ensure that the SII stays within a target SII range. Figure 3B shows a plot of estimated SII over time according to one example. The estimated SII may, for example correspond to the output of block 160 of Figure 1C. Figure 3A illustrates how SII can fluctuate over a period of time. During this period of time, in this example block 165 does not suggest
20 any system change(s), because the SII remains within the high and low targets, which are shown in Figure 3B as “High target sii” and Low target sii.” In this case, block 165 is in what is referred to herein as a “ref_gain_adj = NONE” state, indicating that no gain adjustments will be made to the audio being played back in a local system that is providing a teleconference to a local teleconference
25 participant.

[0088] Figure 4 shows a plot of estimated SII over time according to another example. Figure 4 show what happens when the SII increases beyond the high_target_sii. Between the start of time and t1, block 165 does not suggest any volume change (ref_gain_adj = NONE state). At t1, the SII value has breached
30 the range, going beyond the high_target_sii. In this case, Block 165 will suggest a system change, and will continue to do in subsequent frames (with optional pauses between suggestions as described below) until the SII is between

dec_aim_target_sii and *low_target_sii*. Until this is achieved, Block 165 is in *ref_gain_adj* = INC_REQ state.

[0089] The *dec_aim_target_sii* value may be the same as the *target_sii* or may differ from the *target_sii* based on the knowledge on how the system will react, and how the SII will tend to fluctuate. For example, in some systems it may be known that SII may be underestimated compared to the true SII due to a slow noise estimate adaptation. In these cases, a *dec_aim_target_sii* that is slightly lower than the *target_sii* (but higher than the *low_target_sii*) may be preferable. Once the SII is between *dec_aim_target_sii* and *low_target_sii* (at t2), in this example the system state will go back to the behavior described with reference to Figure 3B. In our implementation, the *dec_aim_target_sii* is the same as the *target_sii*.

[0090] One example of a change to the system that may be implemented is to decrease the volume. The decrease in volume could be specified such that it is proportional to the difference between the current SII and the SII that we are aiming for. In one implementation, the control system may determine a possible gain change in dB, for example a possible gain change such that:

$$\text{target_gain_change} = \text{gain_dec_delta} * (1 + \text{dec_speed} * (\text{sii} - \text{dec_aim_target_sii}))$$

[0091] In the foregoing equation, *Gain_dec_delta* represents a first tuning parameter, which in one example is 0.6 and *Dec_speed* represents a second tuning parameter, which in one example 1.0. Other examples may involve different tuning parameters. In some alternative examples, the first tuning parameter may be 0.5, 0.55, 0.65, 0.7, etc., and the second tuning parameter may be 0.9, 0.95, etc.

[0092] In some implementations, whenever the control system determines a possible change to the system, the control system may be configured to optionally wait for a time interval, which may be measured in input audio frames, to allow the system to stabilize before checking the current SII and suggesting a new change. The suggested change may or may not be implemented. For example, if a feature corresponding to one or more disclosed methods has been switched off or disabled—for example, according to user input—in some examples, the

suggested change will not be implemented. In some such implementations, the amount of time that the control system is configured to wait we wait between checking the current SII and suggesting a new change may be proportional to how close the SII is to the *dec_aim_target_sii*. In some implementations, the
 5 control system may first calculate:

$$\text{Diff_to_target} = \text{sii} - \text{dec_aim_target_sii}$$

$$\text{and Diff_range} = (\text{high_target_sii} - \text{low_target_sii}) / 2$$

[0093] Based on the value of Diff_to_target, the control system may calculate a
 10 wait_factor, for example as follows:

$$\text{Wait_factor} = 1, \text{ if } 0 \geq \text{diff or diff} \geq \text{Diff_range}$$

$$\text{Wait_factor} = 1, \text{ if diff} > \text{Diff_range}/3$$

$$\text{Wait_factor} = 2, \text{ if diff} > \text{Diff_range}/5$$

$$\text{Wait_factor} = 3, \text{ if diff} > \text{Diff_range}/7$$

15 Wait_factor = 4, in all other cases

[0094] In some such examples, the number of input audio frames corresponding to the waiting time interval is equal to:

$$\text{Wait_frames} = \text{gain_adj_holdon_frames} * \text{wait_factor}$$

[0095] In the foregoing equation, gain_adj_holdon_frames represents a tuning
 20 parameter. In one implementation the tuning factor is 96 for a 20ms block length. In other implementations, the tuning factor may be 90, 92, 94, 98 or 100 for a 20ms block length.

[0096] In some instances, the following may occur:

- At t0, the SII is within the high_target_sii and low_target_sii.
- 25 • At t1, the SII goes beyond high_target_sii. The control system is now trying to compensate by suggesting a change.
- The suggested change in the system occurs, and the control system waits.
- At t2, the SII is between high_target_sii and dec_aim_target_sii. The control system again suggests a new change.

- This suggested change occurs, and the control system waits.
- At t3, the SII is now below *low_target_sii*.

[0097] In the foregoing case, the last change in the system has made the system overshoot the target by time t3. To rectify this situation, in some examples block 165 will involve reacting and suggesting changes in the same way as if the SII had breached the *low_target_sii* from a NONE state. Therefore, in some such examples block 165 will involve transitioning to *ref_gain_adj* = DEC_REQ state.

[0098] The above example described a case in which the SII went from a NONE state to an INC_REQ state. A similar logic can be followed for cases in which the SII falls below the *low_target_sii*. In some such examples, block 165 will involve suggesting a system change. In some such examples, the control system will continue suggesting a system change in subsequent frames (with optional pauses between suggestions as described previously) until the SII is between *inc_aim_target_sii* and *high_target_sii*. Until this condition is achieved, block 165 may correspond to an *ref_gain_adj* = DEC_REQ state.

[0099] Figure 5 shows another plot of estimated SII over time. Figure 5 shows how the control system may implement block 165 when, and after, the SII falls below the low target SII according to some examples. The *inc_aim_target_sii* value shown in Figure 5 may be the same as the *target_sii* or may differ based on the knowledge on how the system will react, and how the SII will tend to fluctuate. For example, in some systems it may be known that SII may be overestimated compared to the true SII due to a slow noise estimate adaptation. In these cases, an *inc_aim_target_sii* that is slightly higher than the *target_sii* (but lower than the *high_target_sii*) may be preferable. Once the SII is between *inc_aim_target_sii* and *high_target_sii* (at t2), in some examples the system state will go back to the behavior described with reference to Figure 3B. In some implementations, the *inc_aim_target_sii* may be the same as the *target_sii*.

[0100] One example of a change to the system that the control system may suggest is to increase the volume. The increase in volume may be specified such that it is proportional to the difference between the current SII and a target SII. In one implementation, the control system may suggest a gain change in dB such that:

$$\text{target_gain_change} = \text{gain_inc_delta} * (1 + \text{inc_speed} * (\text{inc_aim_target_sii} - \text{sii}))$$

[0101] In the foregoing equation, Gain_inc_delta represents a tuning parameter.

In one implementation this tuning parameter is 0.8, but in alternative

- 5 implementations this tuning parameter may be 0.7, 0.75, 0.85, 0.9, etc. In the foregoing equation, Inc_speed represents another tuning parameter. In one implementation this tuning parameter is 1.0, but in alternative implementations this tuning parameter may be 0.9, 0.95, 1.05, 1.1, etc.

- [0102] Whenever the control system suggests a change to the system, the control
10 system may optionally wait a few frames to allow the system to stabilize before checking the current SII and suggesting a new change. The amount of time the control system waits between suggestions may be proportional to how close the SII is to the *inc_aim_target_sii*. In one implementation, the wait time may be determined as follows:

- 15 $\text{Wait_frames} = \text{gain_adj_holdon_frames}$

[0103] In the foregoing equation, gain_adj_holdon_frames represents a tuning parameter. In one implementation the tuning factor is 96 for a 20ms block length. In other implementations, the tuning factor may be 90, 92, 94, 98 or 100 for a 20ms block length.

- 20 [0104] In some cases, the control system may determine that the following occurs:

- At t0, the SII is within the high_target_sii and low_target_sii.
- At t1, the SII goes below the low_target_sii. The control system is now trying to compensate by suggesting a change.
- 25 • The suggested change in the system occurs, and the control system waits.
- At t2, the SII is between low_target_sii and inc_aim_target_sii . The control system again suggests a new change.
- This suggested change occurs, and the control system waits.
- At t3, the SII is now above high_target_sii.

- 30 [0105] In this case, the last change in the system has made the system overshoot the target SII. To rectify this, block 165 may involve reacting and suggesting changes in the same way as if the SII had breached the high_target_sii from an

NONE state. In some examples, block 165 will involve transitioning to a `ref_gain_adj = INC_REQ` state.

[0106] In the description so far of how block 165 may be implemented, we have not addressed how the `target_sii`, `high_target_sii`, `low_target_sii` may be determined. According to some examples, the control system may implement an initialization process during which these values are determined. In some instances the initialization process may be initiated by a user, whereas in other instances the initialization process may be initiated when a device used to provide a teleconference is powering up. In some examples, during the initialization step, the other operations described above are not performed, and only when initialization is complete, will the above operations be performed.

[0107] According to some examples, the control system may start the initialization process by determining the average SII over a time interval, which may correspond to a number of audio frames or blocks. (As used herein, the terms “audio frame” and “audio block” have the same meaning.) This period of blocks may be considered as a tuning parameter, referred to herein as “`sii_max_init_counter`.” In one example, this tuning parameter may be 50 blocks, whereas in other examples this tuning parameter may be 40 blocks, 45 blocks, 55 blocks, 60 blocks, etc. After determining the average SII over a time interval, in some implementations the control system will set this average SII to be the `target_sii`. In some such implementations, the high target SII and the low target SII may be determined as follows:

$$\text{high_target_sii} = \text{target_sii} + \text{target_sii_range}$$

$$\text{low_target_sii} = \text{target_sii} - \text{target_sii_range}$$

[0108] In the foregoing equations, `target_sii_range` represents a tuning parameter. In one example, this tuning parameter may be 0.15, whereas in other examples this tuning parameter may be 0.1, 0.2, etc.

[0109] During the initialization phase, there may be audio blocks in which the loudspeaker has insignificant energy. In these blocks, the SII calculation will not be indicative of the true SII of the system and therefore should not be taken into account during the averaging SII procedure. In some implementations, the

control system may determine whether the energy is insignificant by checking whether the following condition is met:

$$\text{mono_ref_level} > \text{ref_th_alpha} * \text{ref_th}$$

[0110] In the foregoing equation, mono_ref_level represents the energy of the speaker for the current block in dB, an input into the control system and ref_th_alpha represents a tuning parameter. In one example, this tuning parameter may be 4.0, whereas in other examples this tuning parameter may be 3.0, 3.5, 4.5, 5.0, etc. In the foregoing equation, ref_th represents a tuning parameter. In one example, this tuning parameter may be -30, whereas in other examples this tuning parameter may be -20, -25, -35, -40, etc.

[0111] When making the compensation, there may be a goal to ensure that the system does not go below the user's initial volume. To support this goal, during initialization, the control system may determine the current volume of the system and then set the current volume as our minimum volume. This current volume may, for example, be stored as "usr_ref_gain." When a decrease in the volume is suggested in block 165, in some implementations the control system may determine whether to implement this suggestion by ensuring that implement this suggestion would not cause the volume to go below the minimum volume.

[0112] During the initialization phase, in some systems the estimated SII may ramp up and may be invalid for the first few blocks of the audio session. This phenomenon may be caused by the instability of the speech level or the noise estimate during the first few blocks of the audio session. In some implementations, the control system may ensure that the SII averaging process does not take into account the first few blocks of the audio session. In some such implementations, the control system may ignore the first 20 blocks (assuming 20ms block length), the first 22 blocks, the first 24 blocks, the first 26 blocks, the first 28 blocks, the first 30 blocks, etc.

[0113] In some instances, during the implementation of block 165, the user may attempt to change the volume manually. If this occurs, the control system may optionally take the user's attempted volume change into account. In some such examples, the control system may interpret the user's selected volume as a new minimum volume (new usr_ref_gain) and a new target_sii. According to some

implementations, when the user interacts with the volume gain of the system, the control system may restart the initialization phase.

[0114] Figure 6 is a flow diagram that outlines another example of a method that may be performed by an apparatus or system such as those disclosed herein. The blocks of method 600, like other methods described herein, are not necessarily performed in the order indicated. In some implementation, one or more of the blocks of method 600 may be performed concurrently. Moreover, some implementations of method 600 may include more or fewer blocks than shown and/or described. The blocks of method 600 may be performed by one or more devices, which may be (or may include) a control system such as the control system 110 that is shown in Figure 1B and described above.

[0115] In this example, method 600 is a method of compensating for environmental noise during a teleconference. According to this example, block 605 involves estimating, by a control system, a current speech spectrum corresponding to speech of remote teleconference participants. In some examples, block 605 may correspond to block 150 of Figure 1C and may be performed according to the descriptions herein of block 150.

[0116] According to this example, block 610 involves estimating, by the control system, a current noise spectrum corresponding to environmental noise in a local environment in which a local teleconference participant is located. According to some examples, block 610 may correspond to block 155 of Figure 1C and may be performed according to the descriptions herein of block 155.

[0117] In this example, block 615 involves calculating, by the control system, a current speech intelligibility index (SII) based, at least in part, on the current speech spectrum and the current noise spectrum. In some examples, block 615 may correspond to block 160 of Figure 1C and may be performed according to the descriptions herein of block 160.

[0118] According to this example, block 620 involves determining, by the control system and based at least in part on the current SII, whether to make an adjustment of a local audio system used by the local teleconference participant. According to some examples, the determining may involve evaluating the current SII according to one or more target SII parameters. In some examples, the determining may involve determining whether the current SII is within a target SII range. In some examples, method 600 may involve, responsive to

determining that the adjustment should be made, adjusting, by the control system, at least a portion of the local audio system. According to some examples, block 620 may correspond to block 165 of Figure 1C and may be performed according to the descriptions herein of block 165. In some such examples, block 620 may involve one or more of the example implementations of block 165 that are described with reference to Figures 3B–5.

[0119] In this example, block 625 involves updating, by the control system, at least one of the one or more target SII parameters responsive to user input corresponding to a playback volume change.

[0120] In some examples, method 600 may involve determining a confidence value corresponding to the current noise spectrum. The confidence value may indicate the likelihood of a current input audio frame corresponding mainly to ambient noise. The confidence value may correspond with the confidence metric 207 that is described herein with reference to Figures 2A and 2B. In some examples, method 600 may involve determining whether to update one or more noise statistics based, at least in part, on the confidence value.

[0121] According to some examples, the confidence value may be a broadband confidence value. Determining whether to make the adjustment of the local audio system may be based, at least in part, on the broadband confidence value.

[0122] However, in some examples method 600 may involve determining a band-based confidence value for each frequency band of a plurality of frequency bands. Determining whether to make the adjustment of the local audio system may involve determining whether to update one or more noise statistics for each frequency band based, at least in part, on the band-based confidence value.

[0123] In some examples, estimating the current noise spectrum may involve estimating an echo coupling gain corresponding to local loudspeaker playback captured by a local microphone system. In some such examples, the echo coupling gain may be estimated by the echo coupling gain estimator 202 of Figures 2B and 3A. According to some examples, estimating the echo coupling gain may involve determining a maximum band loudspeaker reference power for each frequency band of a current audio frame and previous N-1 audio frames, where N is an integer corresponding to a number of audio frames in a delay line. In some examples, estimating the echo coupling gain may involve tracking a minimum power for each frequency band of an input microphone signal. In some examples, estimating the echo coupling gain may involve estimating an echo

coupling gain for each frequency band of the input microphone signal based, at least in part, on the maximum band loudspeaker reference power and the minimum power. Some disclosed examples involve determining whether to update an estimated echo coupling gain based, at least in part, on whether a current estimated echo coupling gain value has changed more than a threshold amount from a most recent estimated echo coupling gain value, on whether an estimated echo coupling gain has been updated within a threshold time interval, or on combinations thereof. In some such examples the threshold time interval may be measured in audio frames or audio blocks.

- 10 [0124] In some examples, method 600 may involve adjusting, by the control system and responsive to determining that the adjustment should be made, at least a portion of the local audio system to maintain the current SII within a target SII range. According to some examples, maintaining the current SII within a target SII range may involve increasing a loudspeaker playback volume of one or more audio frames until the current SII is greater than a first target SII and less than a high target SII. In some examples, the first target SII may be greater than a median target SII and less than the high target SII. According to some examples, maintaining the current SII within a target SII range involves decreasing a loudspeaker playback volume of one or more audio frames until the current SII may be less than a second target SII and greater than a low target SII. In some such examples, the second target SII may be less than a median target SII and greater than the low target SII.

[0125] Following are some enumerated example embodiments (EEEs):

- 25 [0126] EEE1. A method of compensating for environmental noise during a teleconference, the method comprising: estimating, by a control system, a current speech spectrum corresponding to speech of remote teleconference participants; estimating, by the control system, a current noise spectrum corresponding to environmental noise in a local environment in which a local teleconference participant is located; calculating, by the control system, a current speech intelligibility index (SII) based, at least in part, on the current speech spectrum and the current noise spectrum; determining a confidence value corresponding to the current noise spectrum, the confidence value indicating a likelihood of a current input audio frame corresponding mainly to ambient noise; and determining, by the control system and based at least in part on the current SII and the confidence value, whether to make an adjustment of a local audio system
- 30
35

used by the local teleconference participant, wherein the determining involves evaluating the current SII according to one or more target SII parameters.

[0127] EEE2. The method of EEE1, wherein the determining involves determining whether the current SII is within a target SII range.

5 [0128] EEE3. The method of EEE1 or EEE2, further comprising, responsive to determining that the adjustment should be made, adjusting, by the control system, at least a portion of the local audio system.

[0129] EEE4. The method of any one of EEEs 1–3, further comprising updating at least one of the one or more target SII parameters responsive to user input
10 corresponding to a playback volume change.

[0130] EEE5. The method of EEE1, wherein the confidence value is a broadband confidence value and wherein determining whether to make the adjustment of the local audio system is based, at least in part, on the broadband confidence value.

[0131] EEE6. The method of EEE4 or EEE5, further comprising determining
15 whether to update one or more noise statistics based, at least in part, on the confidence value.

[0132] EEE7. The method of EEE1, further comprising determining a band-based confidence value for each frequency band of a plurality of frequency bands and wherein determining whether to make the adjustment of the local audio
20 system involves determining whether to update one or more noise statistics for each frequency band based, at least in part, on the band-based confidence value.

[0133] EEE8. The method of any one of EEEs 1–7, wherein estimating the current noise spectrum involves estimating an echo coupling gain corresponding to local loudspeaker playback captured by a local microphone system.

25 [0134] EEE9. The method of EEE8, wherein estimating the echo coupling gain involves: determining a maximum band loudspeaker reference power for each frequency band of a current audio frame and previous N-1 audio frames, where N is an integer corresponding to a number of audio frames in a delay line; tracking a minimum power for each frequency band of an input microphone signal; and
30 estimating an echo coupling gain for each frequency band of the input microphone signal based, at least in part, on the maximum band loudspeaker reference power and the minimum power.

[0135] EEE10. The method of EEE9, further comprising determining whether to update an estimated echo coupling gain based, at least in part, on
35 whether a current estimated echo coupling gain value has changed more than a

threshold amount from a most recent estimated echo coupling gain value, on whether an estimated echo coupling gain has been updated within a threshold time interval, or on combinations thereof.

5 [0136] EEE11. The method of any one of EEEs1–10, further comprising, responsive to determining that the adjustment should be made, adjusting, by the control system, at least a portion of the local audio system to maintain the current SII within a target SII range.

10 [0137] EEE12. The method of EEE11, wherein maintaining the current SII within a target SII range involves increasing a loudspeaker playback volume of one or more audio frames until the current SII is greater than a first target SII and less than a high target SII.

[0138] EEE13. The method of EEE12, wherein the first target SII is greater than a median target SII and less than the high target SII.

15 [0139] EEE14. The method of EEE12 or EEE13, wherein maintaining the current SII within a target SII range involves decreasing a loudspeaker playback volume of one or more audio frames until the current SII is less than a second target SII and greater than a low target SII.

[0140] EEE15. The method of EEE14, wherein the second target SII is less than a median target SII and greater than the low target SII.

20 [0141] EEE16. An apparatus configured to perform the method of any one of EEEs1–15.

[0142] EEE17. A system configured to perform the method of any one of EEEs1–15.

25 [0143] EEE18. One or more non-transitory media having instructions stored thereon for controlling one or more devices to perform the method of any one of EEEs1–15.

30 [0144] Some aspects of the present disclosure include a system or device configured (e.g., programmed) to perform one or more examples of the disclosed methods, and a tangible computer readable medium (e.g., a disc) which stores code for implementing one or more examples of the disclosed methods or steps thereof. For example, some disclosed systems can be or include a programmable general purpose processor, digital signal processor, or microprocessor, programmed with software or firmware and/or otherwise configured to perform any of a variety of operations on data, including an embodiment of disclosed

methods or steps thereof. Such a general purpose processor may be or include a computer system including an input device, a memory, and a processing subsystem that is programmed (and/or otherwise configured) to perform one or more examples of the disclosed methods (or steps thereof) in response to data
5 asserted thereto.

[0145] Some embodiments may be implemented as a configurable (e.g., programmable) digital signal processor (DSP) that is configured (e.g., programmed and otherwise configured) to perform required processing on audio signal(s), including performance of one or more examples of the disclosed
10 methods. Alternatively, embodiments of the disclosed systems (or elements thereof) may be implemented as a general purpose processor (e.g., a personal computer (PC) or other computer system or microprocessor, which may include an input device and a memory) which is programmed with software or firmware and/or otherwise configured to perform any of a variety of operations including
15 one or more examples of the disclosed methods. Alternatively, elements of some embodiments of the inventive system are implemented as a general purpose processor or DSP configured (e.g., programmed) to perform one or more examples of the disclosed methods, and the system also includes other elements (e.g., one or more loudspeakers and/or one or more microphones). A general
20 purpose processor configured to perform one or more examples of the disclosed methods may be coupled to an input device (e.g., a mouse and/or a keyboard), a memory, and a display device.

[0146] Another aspect of the present disclosure is a computer readable medium (for example, a disc or other tangible storage medium) which stores code for
25 performing (e.g., coder executable to perform) one or more examples of the disclosed methods or steps thereof.

[0147] While specific embodiments and applications of the disclosure have been described herein, it will be apparent to those of ordinary skill in the art that many variations on the embodiments and applications described herein are possible
30 without departing from the scope of this disclosure.

CLAIMS

What Is Claimed Is:

1. A method of compensating for environmental noise during a teleconference, the method comprising:
 - 5 estimating, by a control system, a current speech spectrum corresponding to speech of remote teleconference participants;
estimating, by the control system, a current noise spectrum corresponding to environmental noise in a local environment in which a local teleconference participant is located;
 - 10 calculating, by the control system, a current speech intelligibility index (SII) based, at least in part, on the current speech spectrum and the current noise spectrum;
determining, by the control system and based at least in part on the current SII, whether to make an adjustment of a local audio system used by the
15 local teleconference participant, wherein the determining involves evaluating the current SII according to one or more target SII parameters; and
updating, by the control system, at least one of the one or more target SII parameters responsive to user input corresponding to a playback volume change.
2. The method of claim 1, wherein the determining involves determining
20 whether the current SII is within a target SII range.
3. The method of claim 1 or claim 2, further comprising, responsive to determining that the adjustment should be made, adjusting, by the control system, at least a portion of the local audio system.
4. The method of any one of claims 1–3, further comprising determining a
25 confidence value corresponding to the current noise spectrum, the confidence value indicating a likelihood of a current input audio frame corresponding mainly to ambient noise.
5. The method of claim 4, wherein the confidence value is a broadband confidence value and wherein determining whether to make the adjustment of the
30 local audio system is based, at least in part, on the broadband confidence value.
6. The method of claim 4 or claim 5, further comprising determining

whether to update one or more noise statistics based, at least in part, on the confidence value.

7. The method of claim 4, further comprising determining a band-based confidence value for each frequency band of a plurality of frequency bands and
5 wherein determining whether to make the adjustment of the local audio system involves determining whether to update one or more noise statistics for each frequency band based, at least in part, on the band-based confidence value.

8. The method of any one of claims 1–7, wherein estimating the current noise spectrum involves estimating an echo coupling gain corresponding to local
10 loudspeaker playback captured by a local microphone system.

9. The method of claim 8, wherein estimating the echo coupling gain involves:

determining a maximum band loudspeaker reference power for each frequency band of a current audio frame and previous N-1 audio frames, where N
15 s an integer corresponding to a number of audio frames in a delay line;

tracking a minimum power for each frequency band of an input microphone signal; and

estimating an echo coupling gain for each frequency band of the input microphone signal based, at least in part, on the maximum band loudspeaker
20 reference power and the minimum power.

10. The method of claim 9, further comprising determining whether to update an estimated echo coupling gain based, at least in part, on whether a current estimated echo coupling gain value has changed more than a threshold amount from a most recent estimated echo coupling gain value, on whether an estimated
25 echo coupling gain has been updated within a threshold time interval, or on combinations thereof.

11. The method of any one of claims 1–10, further comprising, responsive to determining that the adjustment should be made, adjusting, by the control system, at least a portion of the local audio system to maintain the current SII within a
30 target SII range.

12. The method of claim 11, wherein maintaining the current SII within a target SII range involves increasing a loudspeaker playback volume of one or

more audio frames until the current SII is greater than a first target SII and less than a high target SII.

13. The method of claim 12, wherein the first target SII is greater than a median target SII and less than the high target SII.

5 14. The method of claim 12 or claim 13, wherein maintaining the current SII within a target SII range involves decreasing a loudspeaker playback volume of one or more audio frames until the current SII is less than a second target SII and greater than a low target SII.

10 15. The method of claim 14, wherein the second target SII is less than a median target SII and greater than the low target SII.

16. An apparatus configured to perform the method of any one of claims 1–15.

17. A system configured to perform the method of any one of claims 1–15.

15 18. One or more non-transitory media having instructions stored thereon for controlling one or more devices to perform the method of any one of claims 1–15.

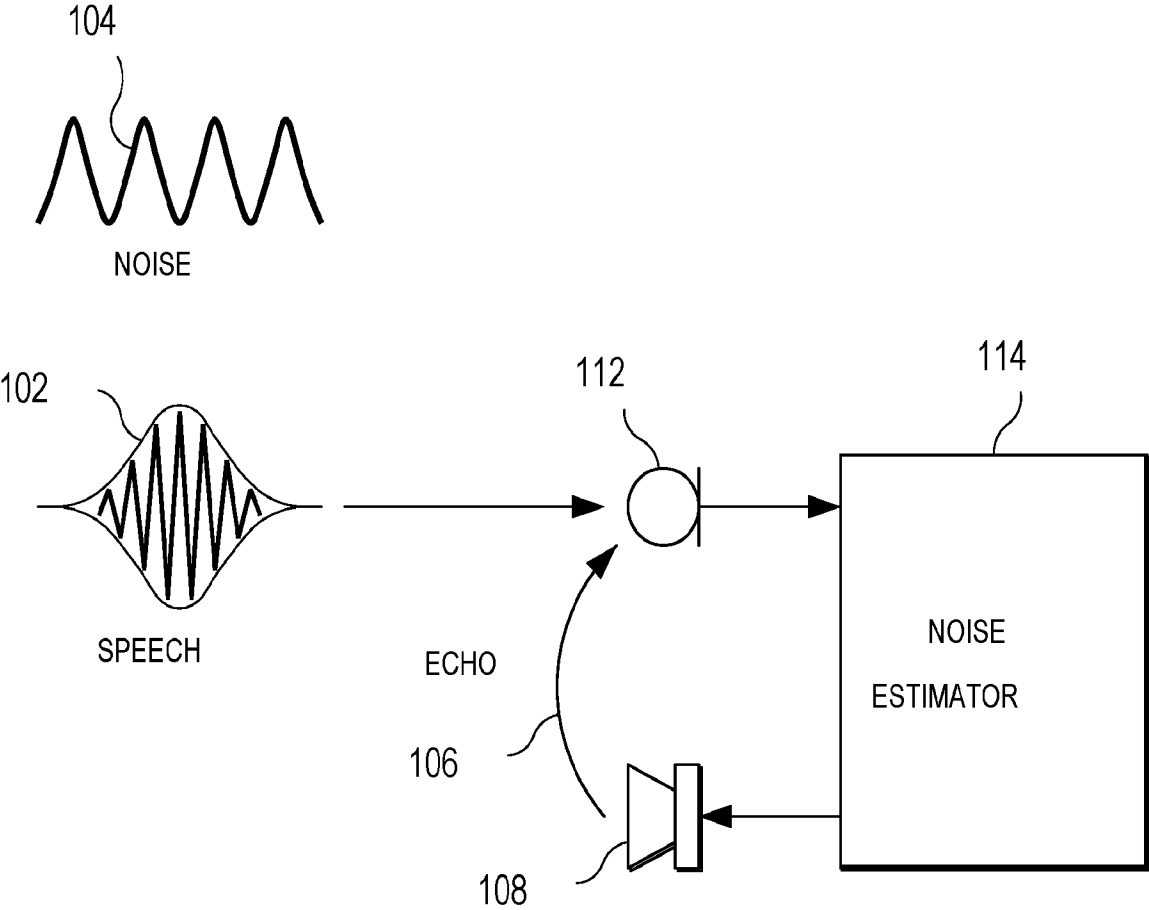


FIGURE 1A

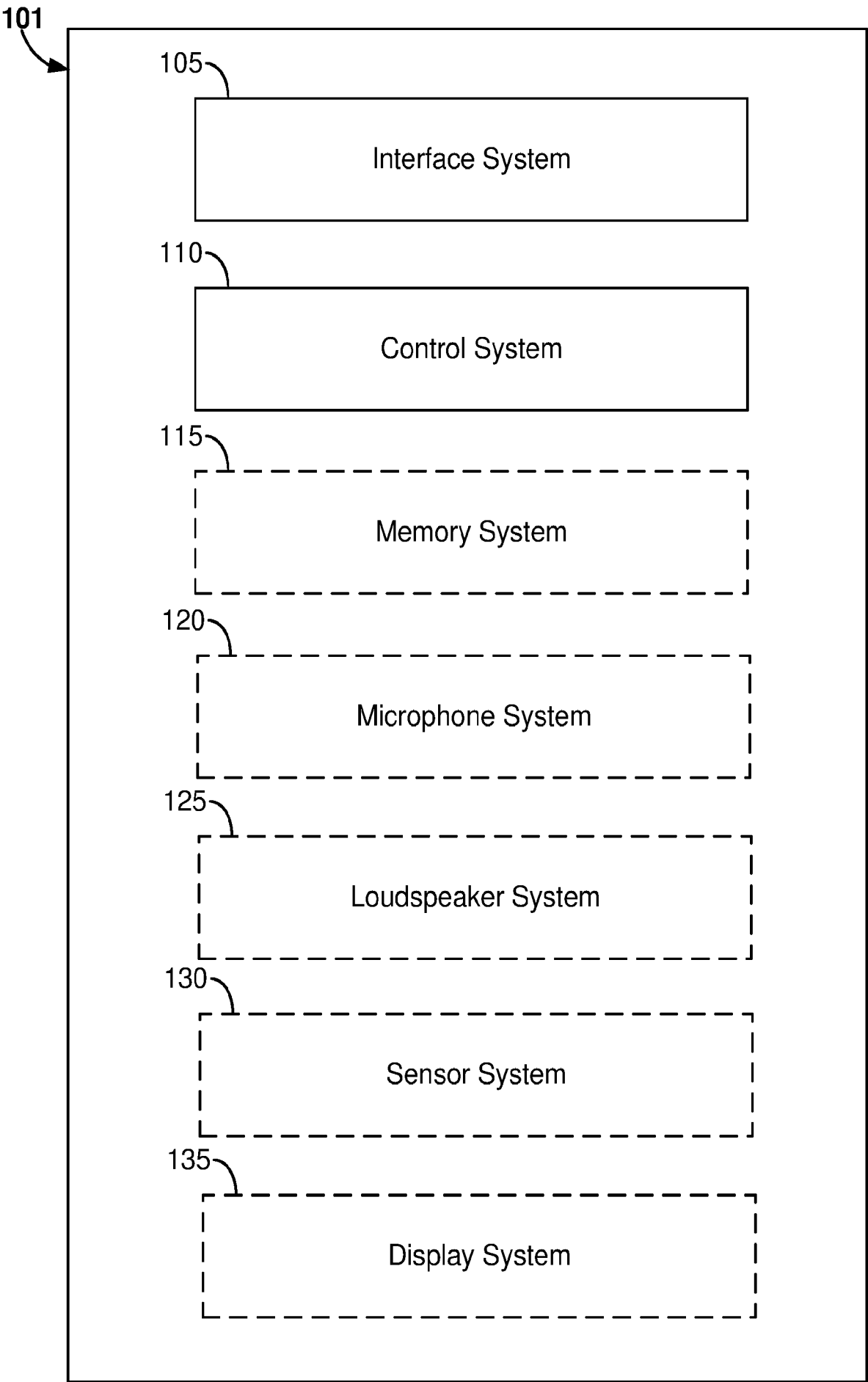
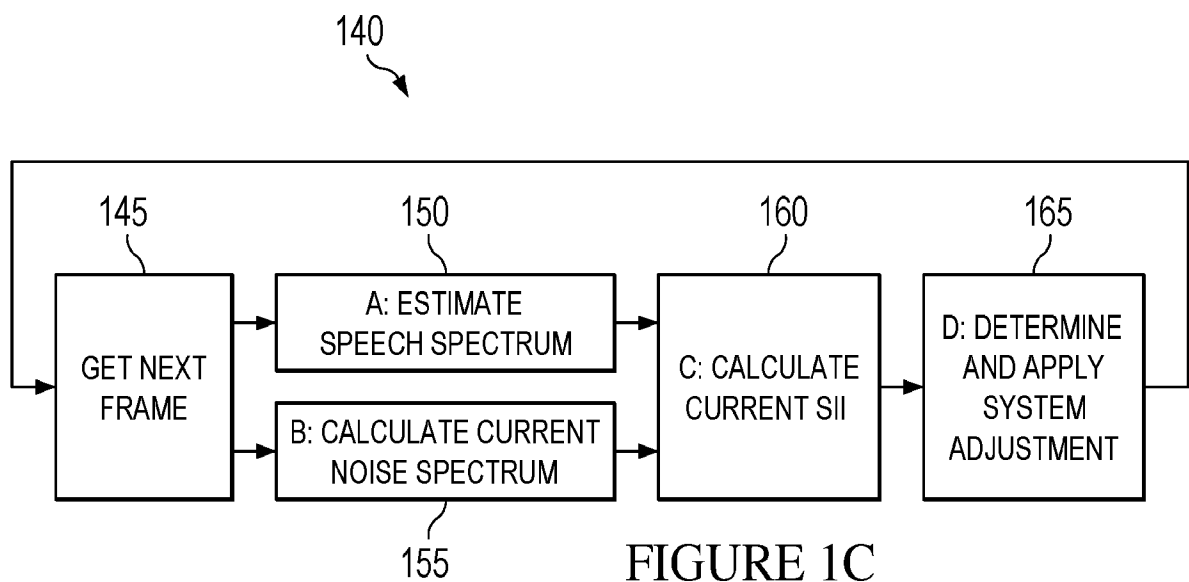


Figure 1B



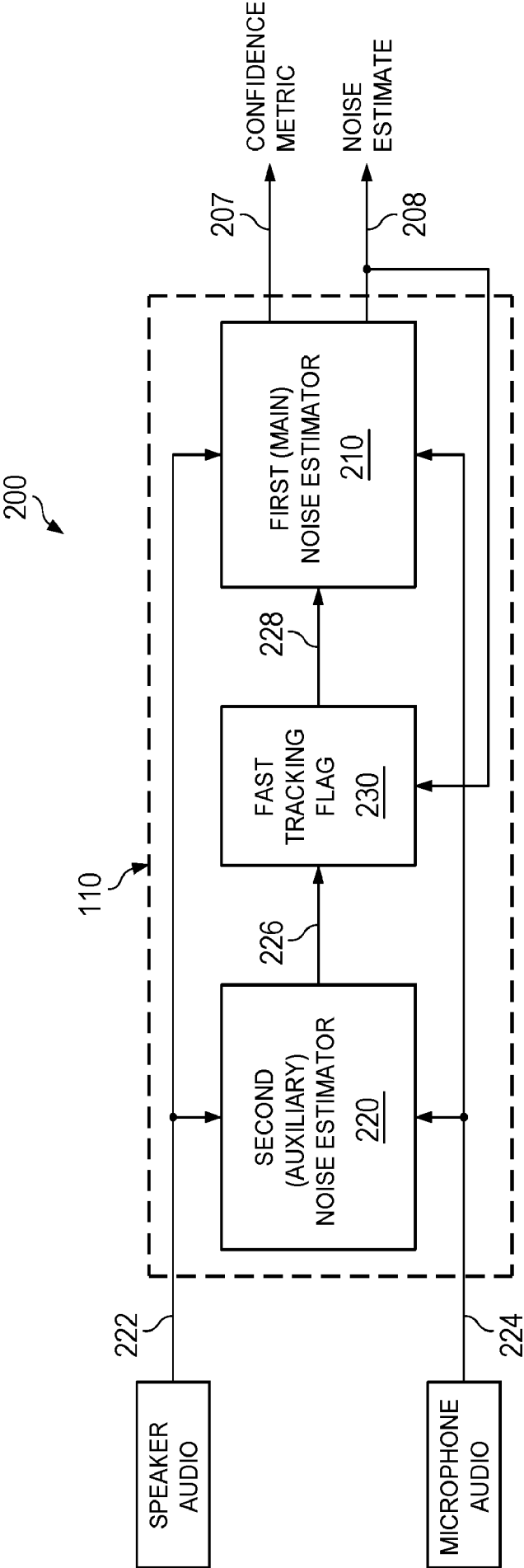


FIGURE 2A

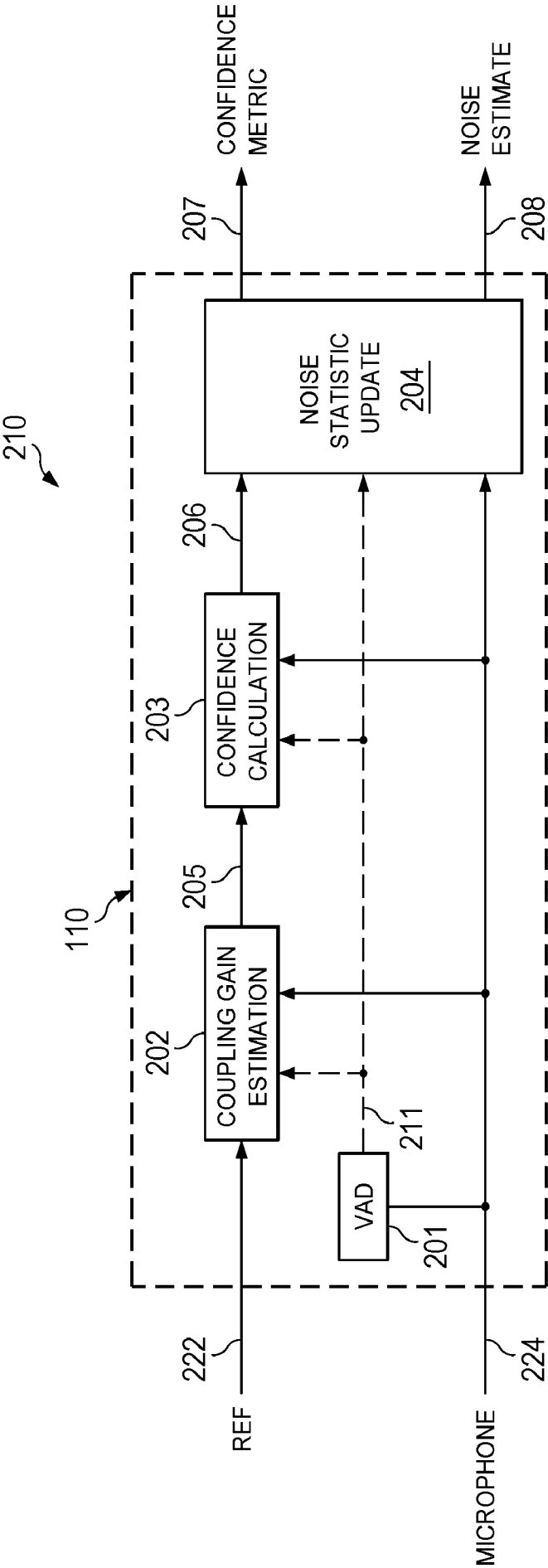


FIGURE 2B

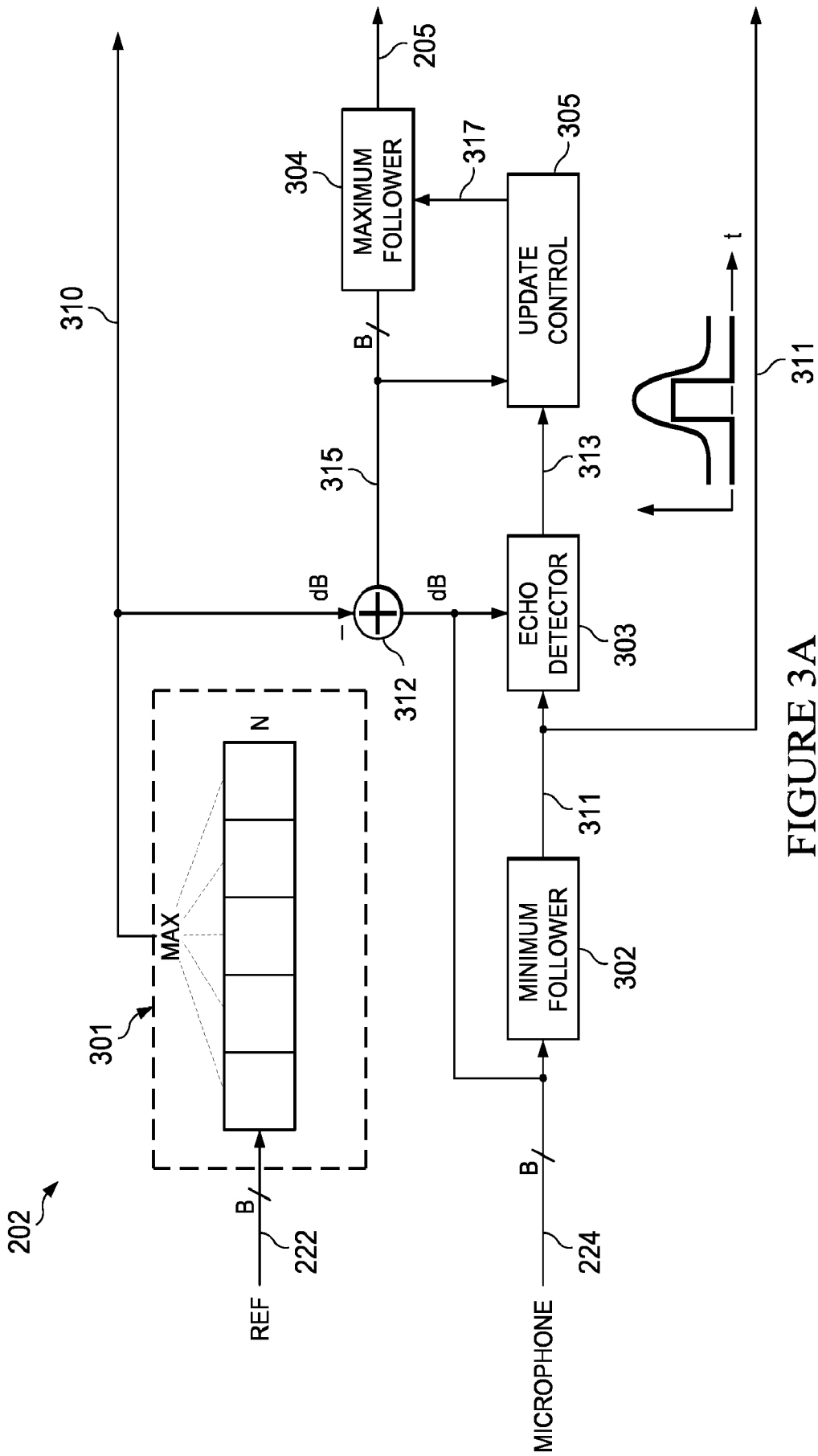


FIGURE 3A

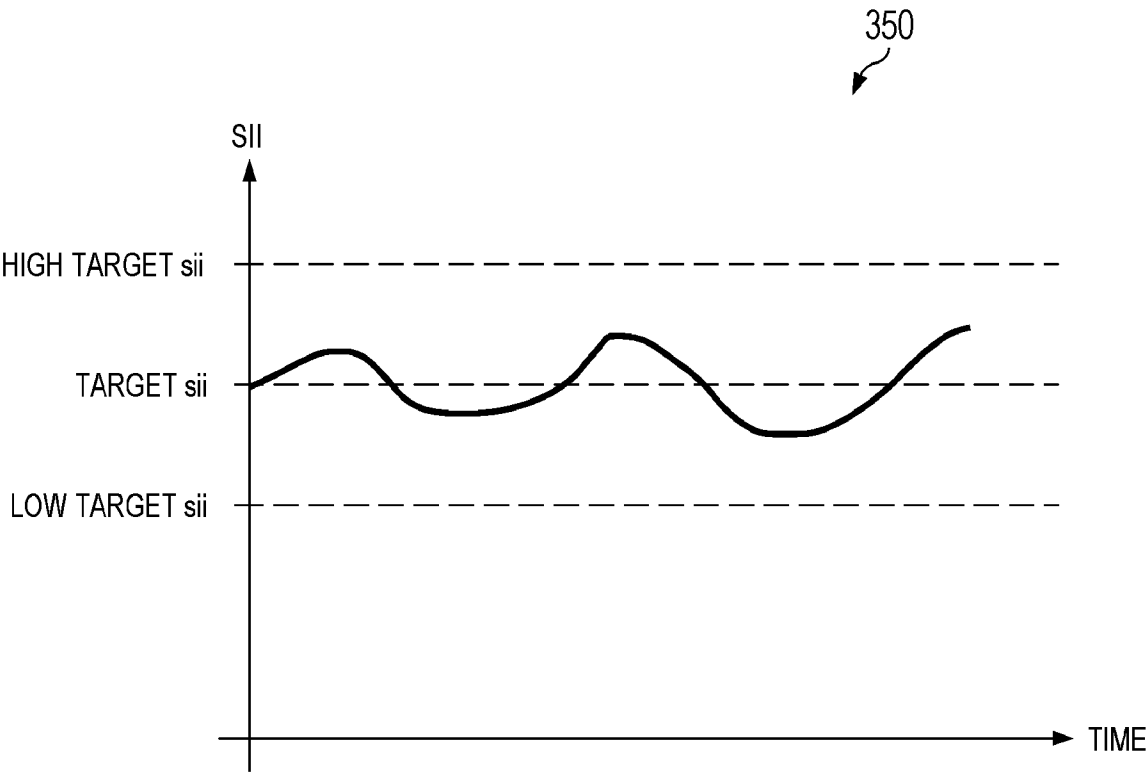


FIGURE 3B

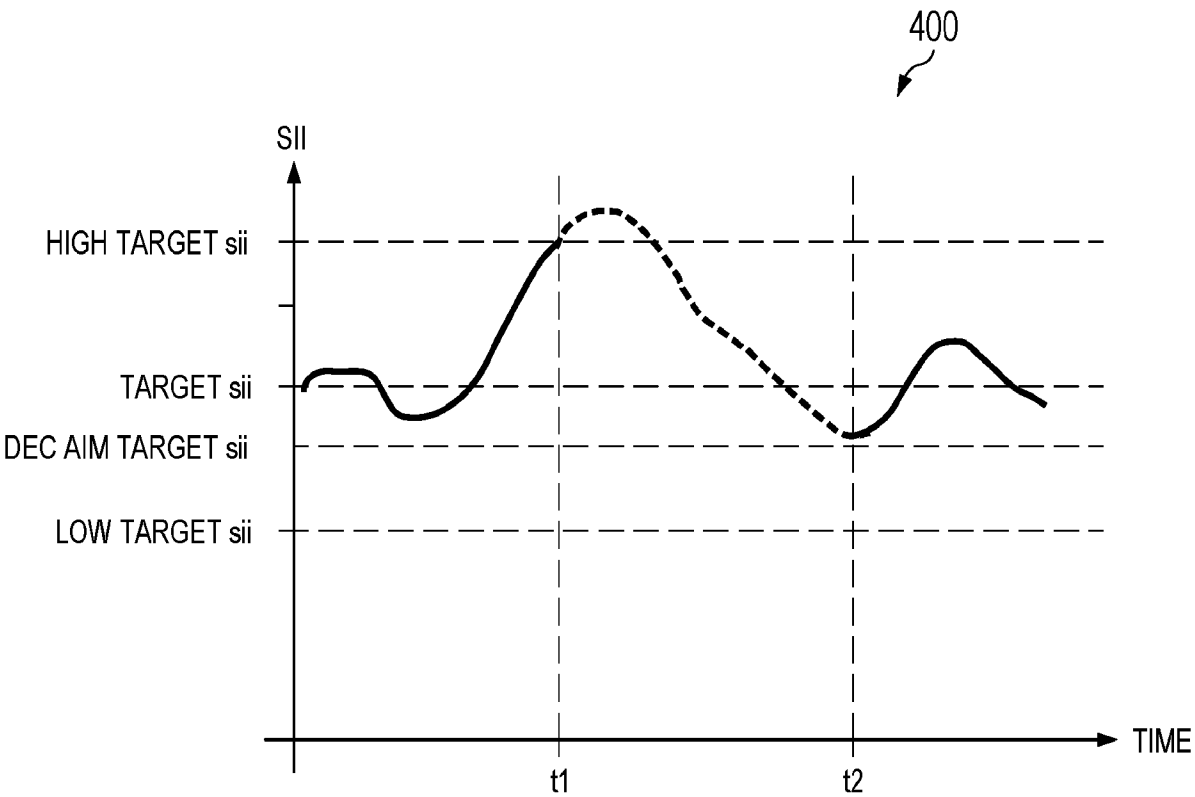


FIGURE 4

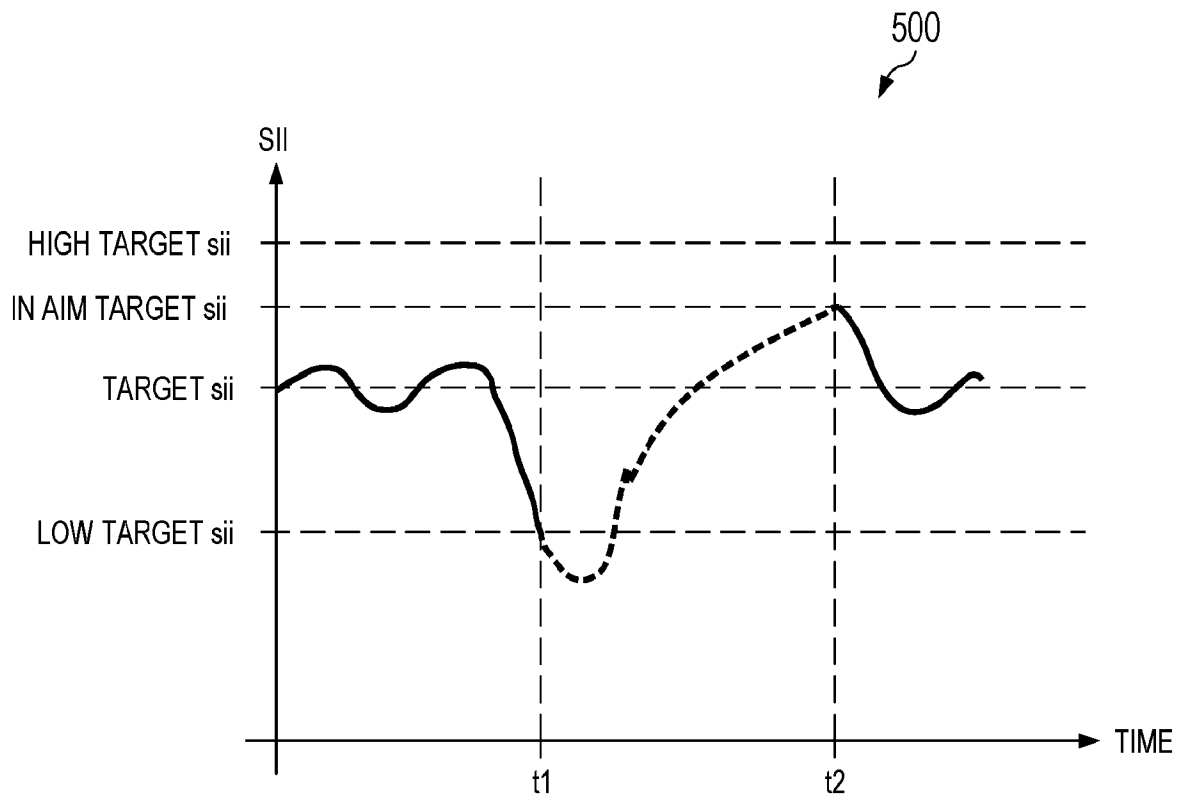
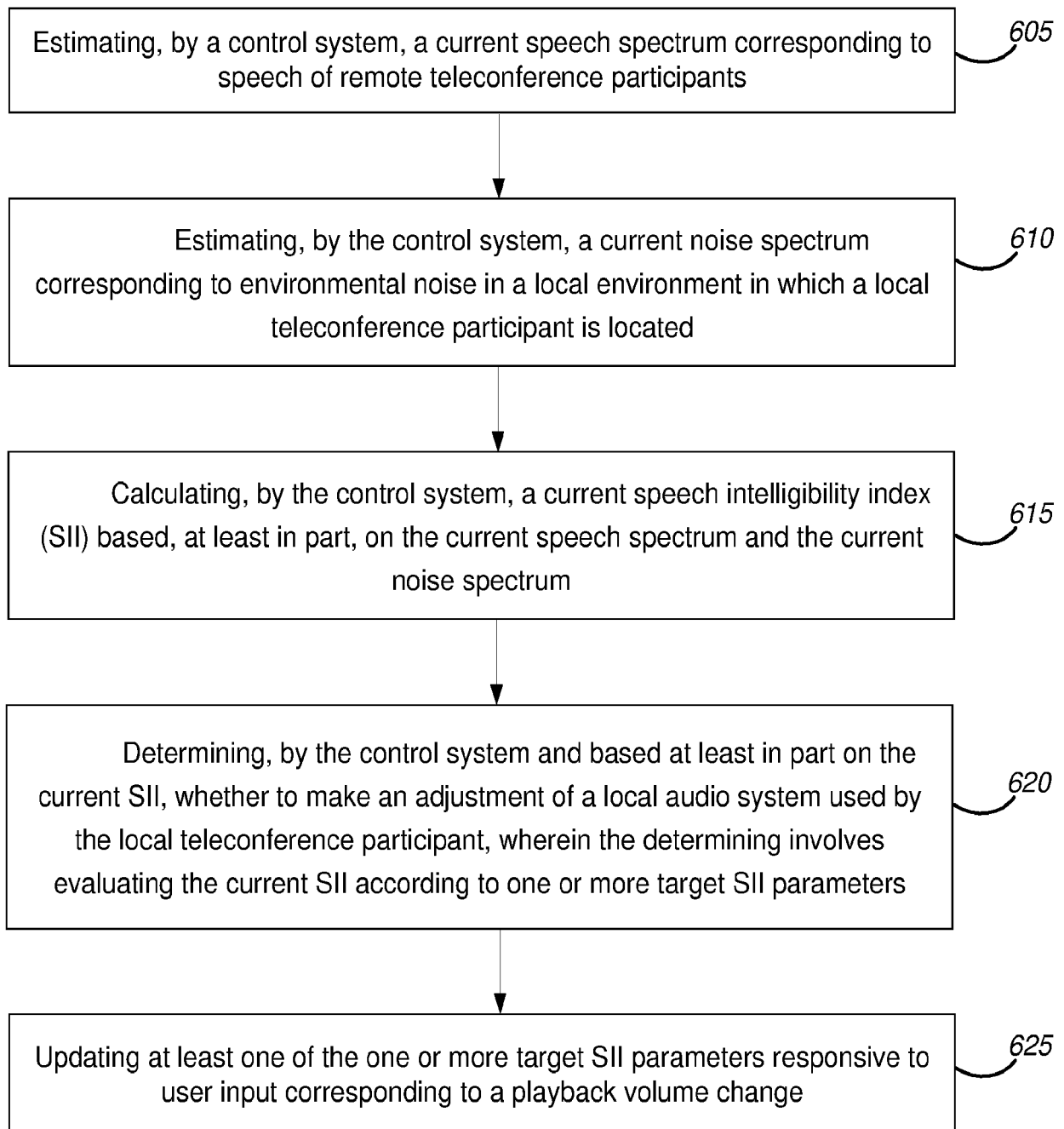


FIGURE 5



600 ↗

Figure 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/036469

A. CLASSIFICATION OF SUBJECT MATTER		
INV.	G10L21/0208	H04M3/00
		H04M9/08
		H03G3/32
ADD.	G10L25/60	G10L21/0364
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) G10L H04M H04Q H03G		
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) EPO-Internal, COMPENDEX, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2017/140772 A1 (PAK JUNHYEONG [KR] ET AL) 18 May 2017 (2017-05-18) figure 3 paragraphs [0005], [0008], [0022], [0035], [0037] -----	1 - 18
A	US 2021/174821 A1 (HWANG INWOO [KR] ET AL) 10 June 2021 (2021-06-10) figure 3 paragraphs [0012], [0015], [0024] -----	1 - 18
A	US 2012/121096 A1 (CHEN SHAOHAI [US] ET AL) 17 May 2012 (2012-05-17) figures 1,2 paragraphs [0004], [0005], [0031] -----	1 - 18
☐ 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	"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	
"E" earlier application or patent but published on or after the international filing date	"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	
"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)	"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	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 5 August 2024	Date of mailing of the international search report 13/08/2024	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Chétry, Nicolas	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2024/036469

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2017140772 A1	18-05-2017	KR 101715198 B1	10-03-2017
		US 2017140772 A1	18-05-2017

US 2021174821 A1	10-06-2021	CN 113038344 A	25-06-2021
		EP 3836140 A1	16-06-2021
		KR 20210072384 A	17-06-2021
		US 2021174821 A1	10-06-2021
		WO 2021118106 A1	17-06-2021

US 2012121096 A1	17-05-2012	EP 2453438 A1	16-05-2012
		HK 1170839 A1	08-03-2013
		JP 5525508 B2	18-06-2014
		JP 2012105287 A	31-05-2012
		KR 20120051595 A	22-05-2012
		NL 2007764 C2	31-07-2012
		TW 201227718 A	01-07-2012
		US 2012121096 A1	17-05-2012
		WO 2012064764 A1	18-05-2012
