

(19) United States

(12) Patent Application Publication

Pan et al.

(10) Pub. No.: US 2023/0289420 A1

(43) Pub. Date: Sep. 14, 2023

(54) METHOD FOR MULTIFACTOR AUTHENTICATION USING BONE CONDUCTION AND AUDIO SIGNALS

(71) Applicant: University of Houston System, Houston, TX (US)

(72) Inventors: Miao Pan, Houston, TX (US); Chenpei Huang, Houston, TX (US); Dian Shi, Houston, TX (US)

(21) Appl. No.: 18/180,277

(22) Filed: Mar. 8, 2023

Related U.S. Application Data

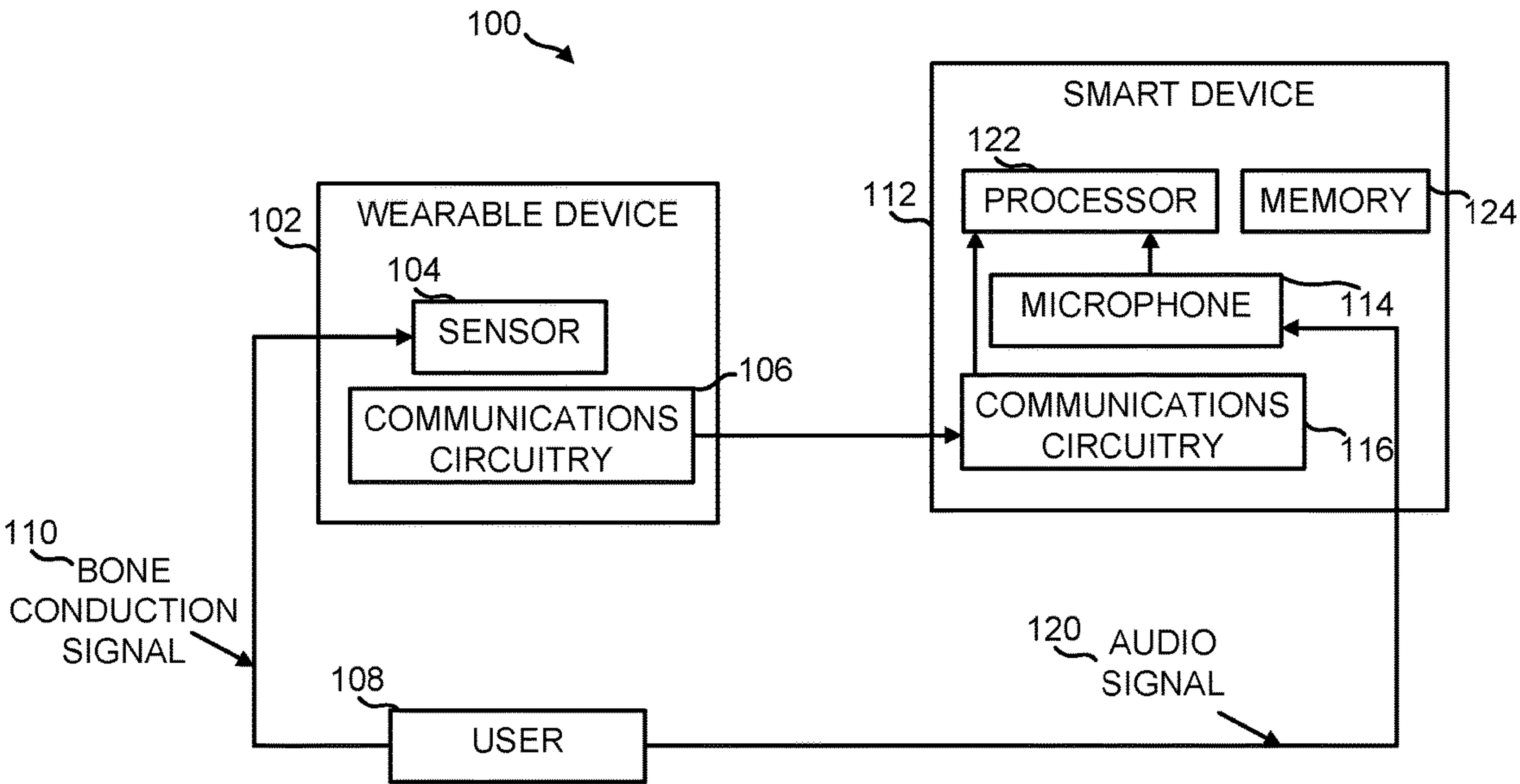
(60) Provisional application No. 63/268,999, filed on Mar. 8, 2022, provisional application No. 63/269,001, filed on Mar. 8, 2022, provisional application No. 63/380,229, filed on Oct. 19, 2022.

Publication Classification

(51) Int. Cl.
G06F 21/32 (2006.01)
G10L 17/26 (2006.01)

G10L 17/02 (2006.01)
G10L 21/0208 (2006.01)
(52) U.S. Cl.
CPC G06F 21/32 (2013.01); G10L 17/26 (2013.01); G10L 17/02 (2013.01); G10L 21/0208 (2013.01)

(57) ABSTRACT
Provided here are systems and method for two-way authentication of a user. In an embodiment, the method may include, in response to reception of an audio signal from a user, determining, via a smart device and/or authentication circuitry, whether a corresponding bone conduction signal is received from one of one or more separate wearable devices. The method may include, in response to determination that the corresponding bone conduction signal is received: determining, via the smart device and/or authentication circuitry, whether the audio signal and corresponding conduction signal are consistent. The method may include, in response to a determination that the audio signal and corresponding conduction signal are consistent: verifying, via the smart device, the audio signal; verifying, via the smart device, the corresponding bone conduction signal; and, in response to verification of the audio signal and the corresponding bone conduction signal, authenticating a user.



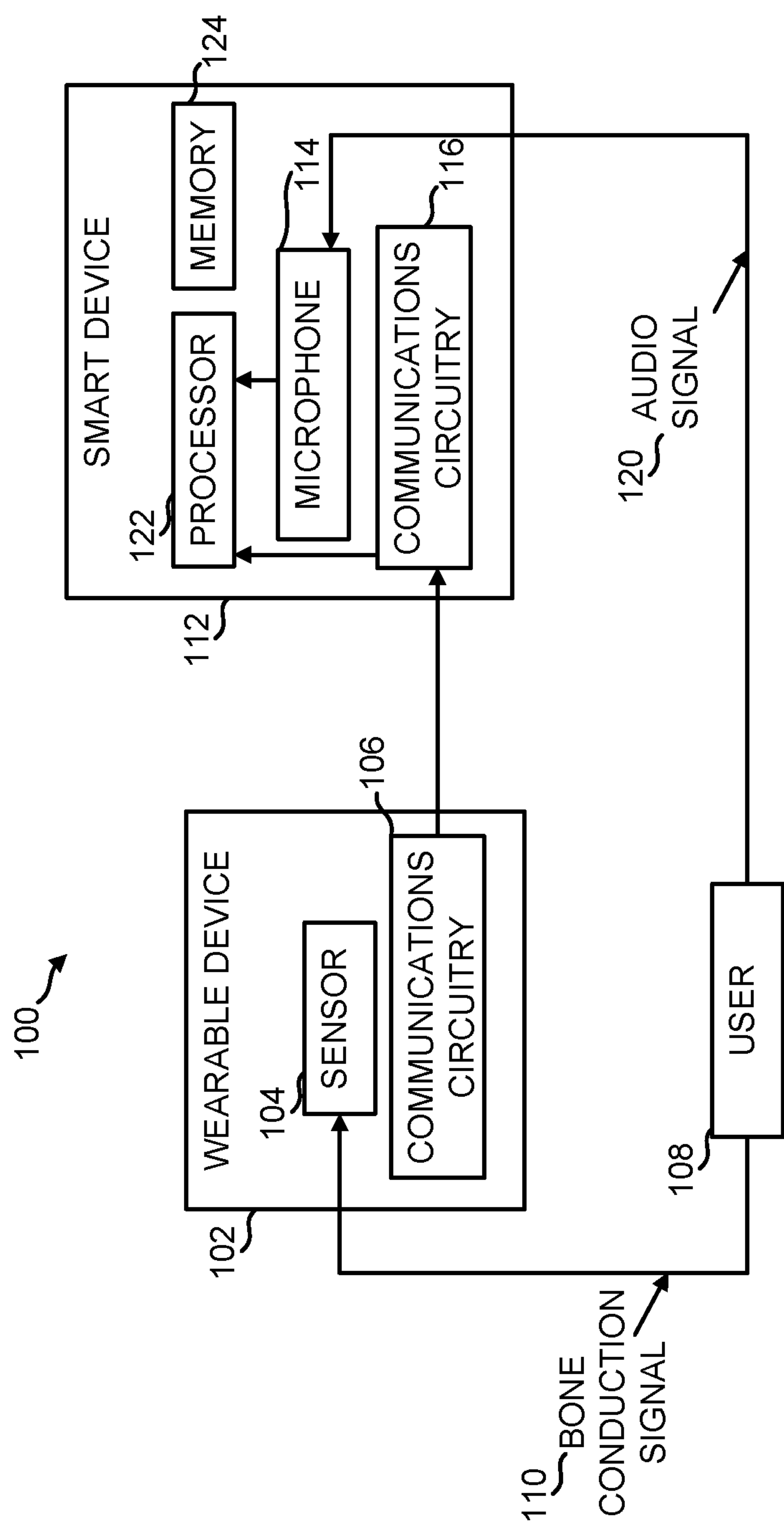


FIG. 1A

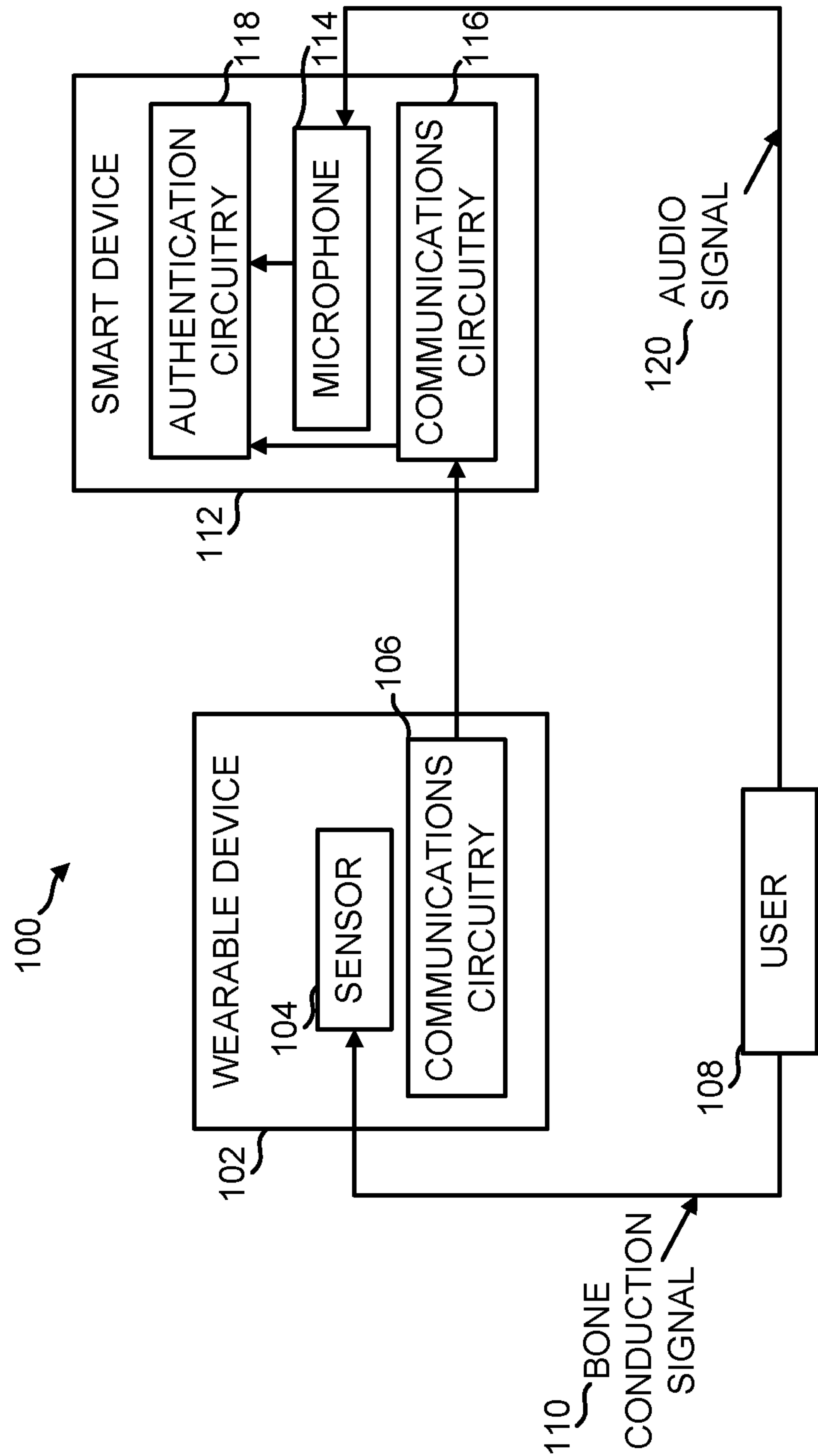


FIG. 1B

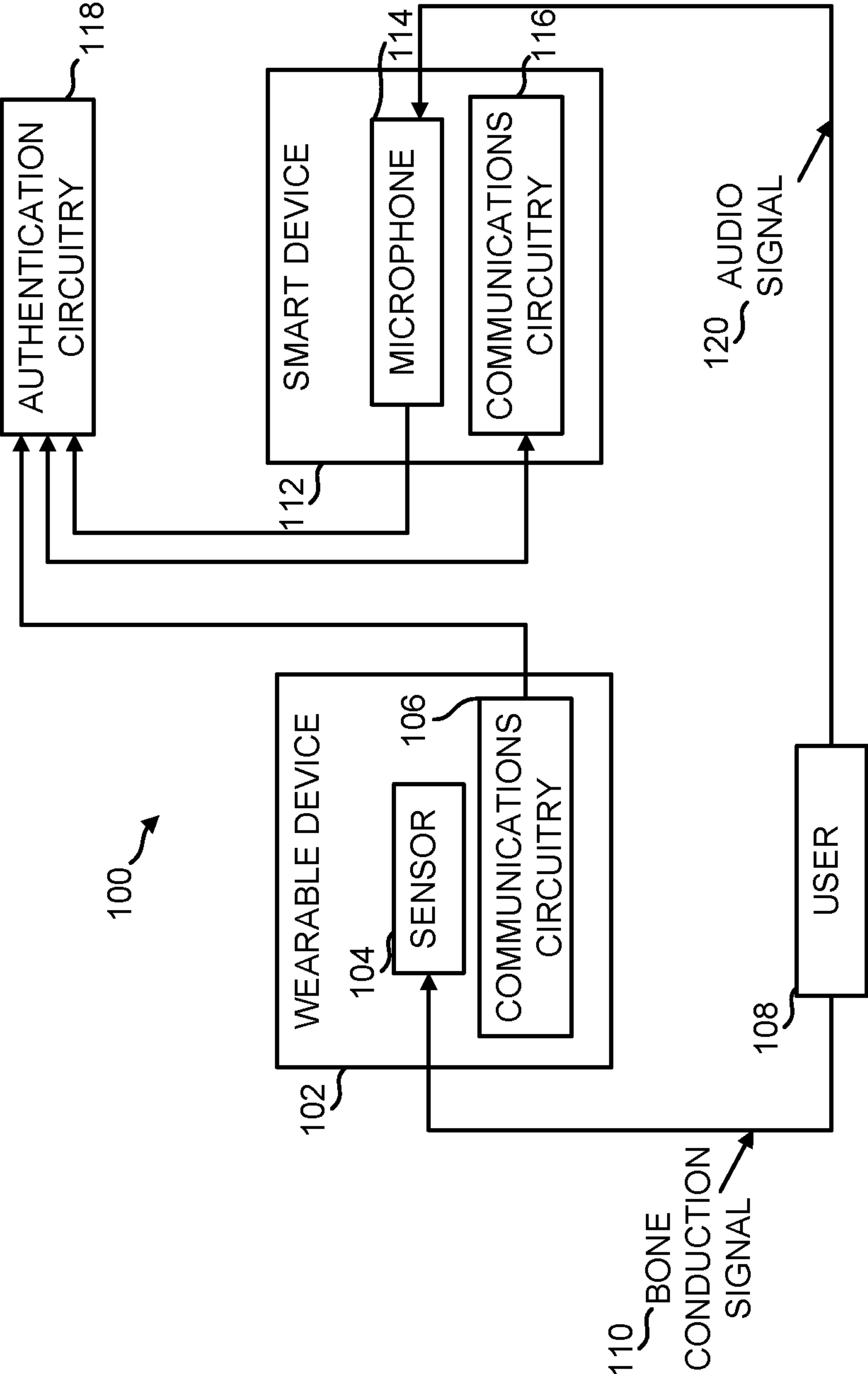


FIG. 1C

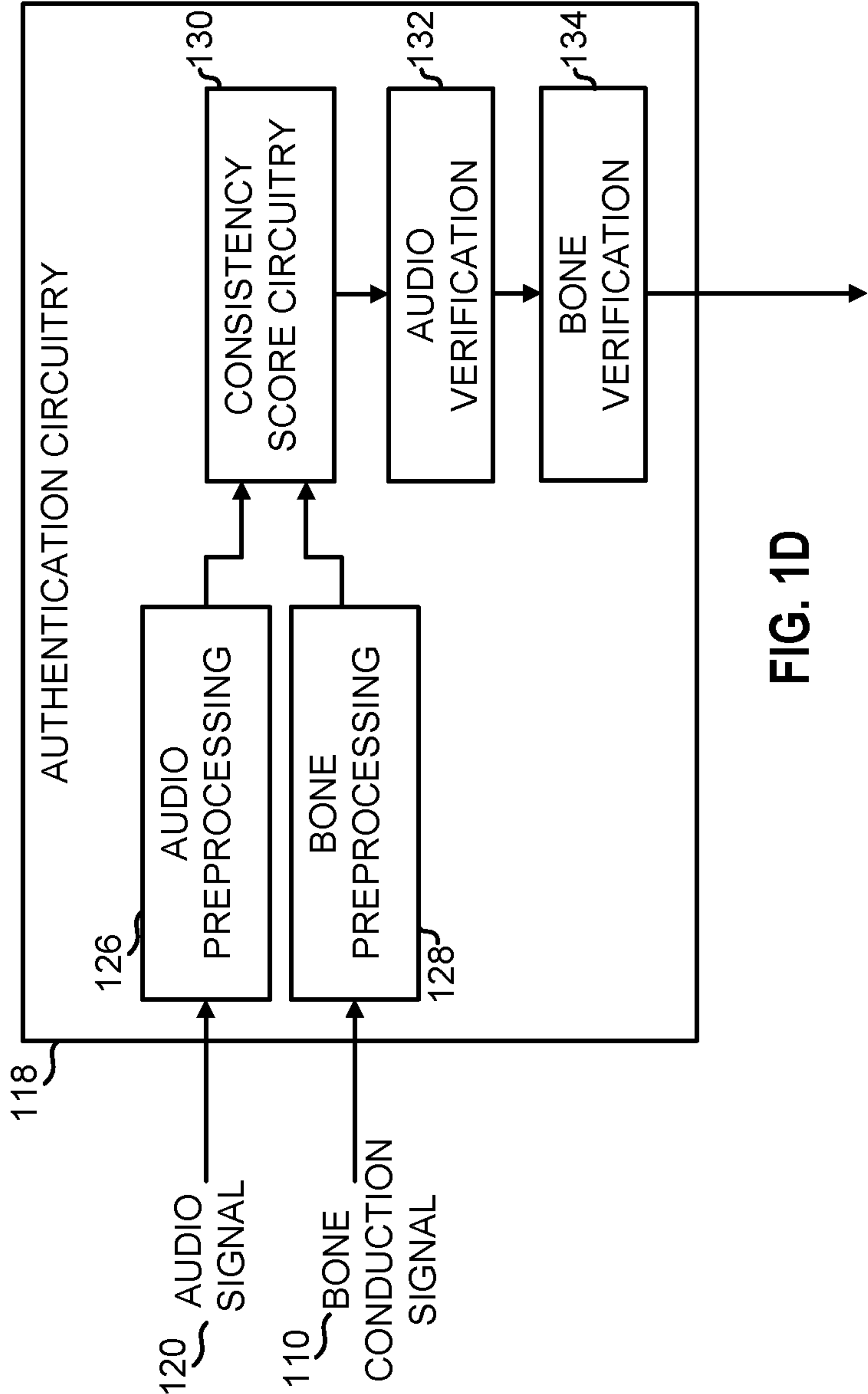


FIG. 1D

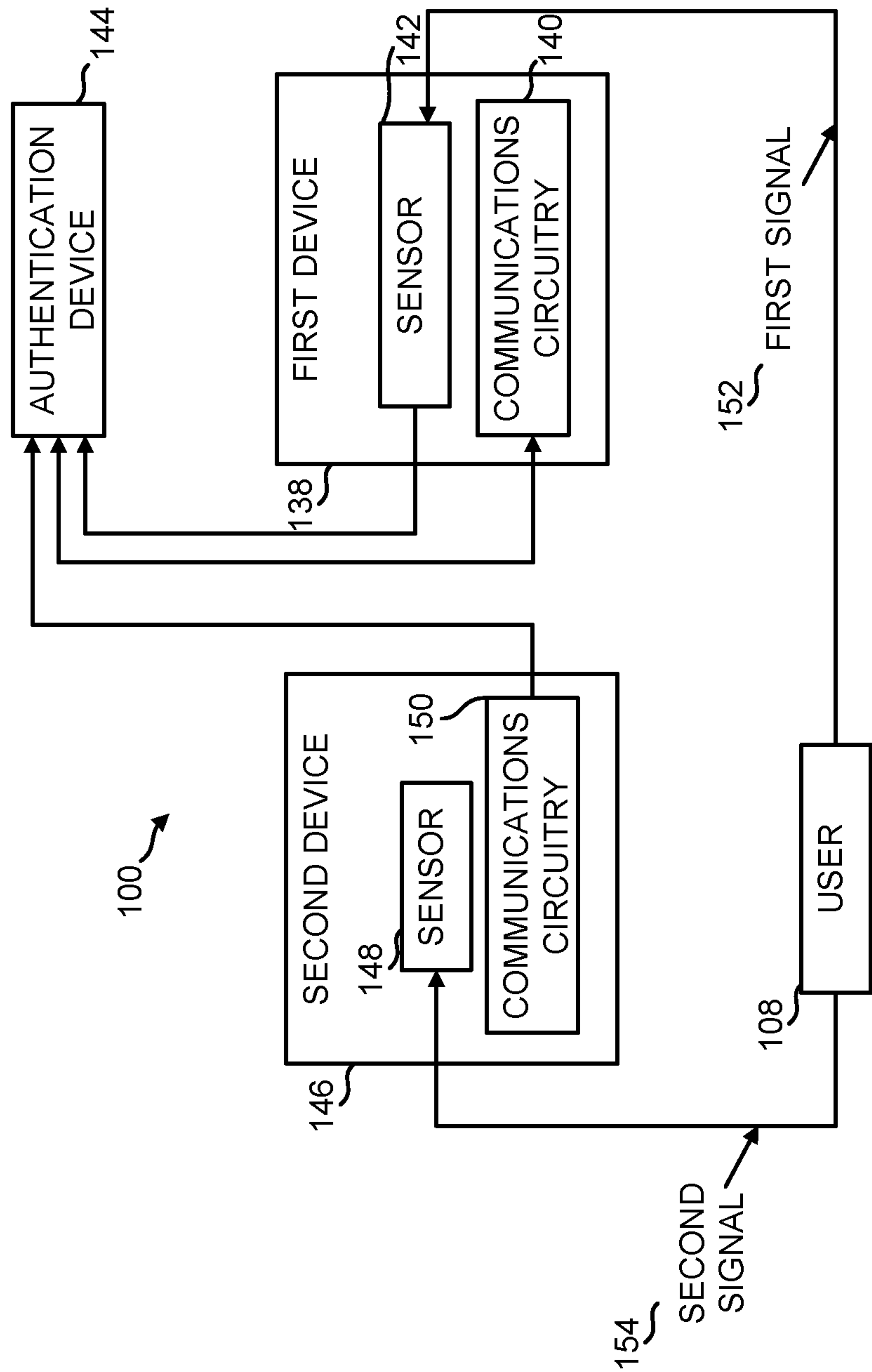


FIG. 1E

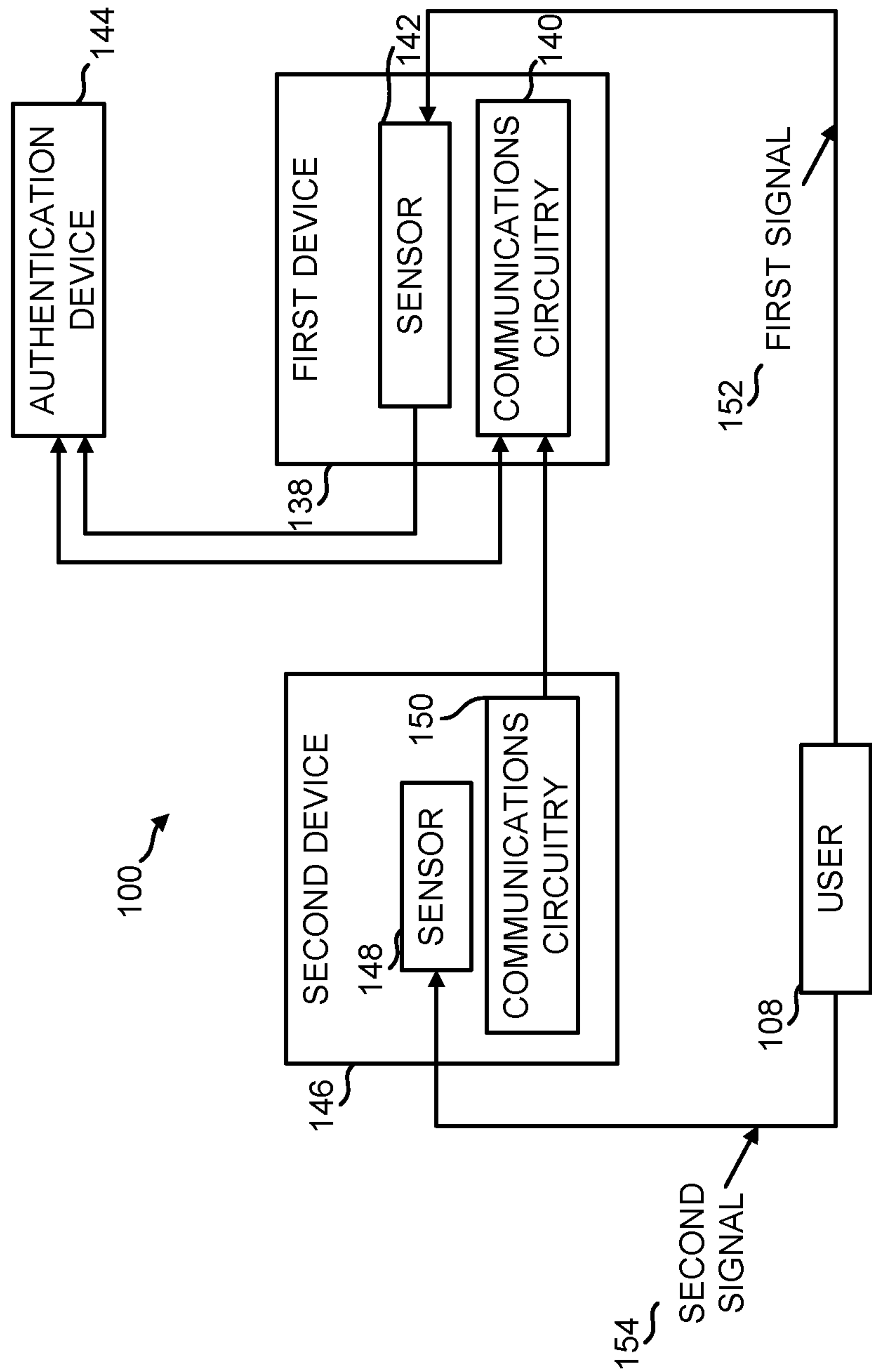


FIG. 1F

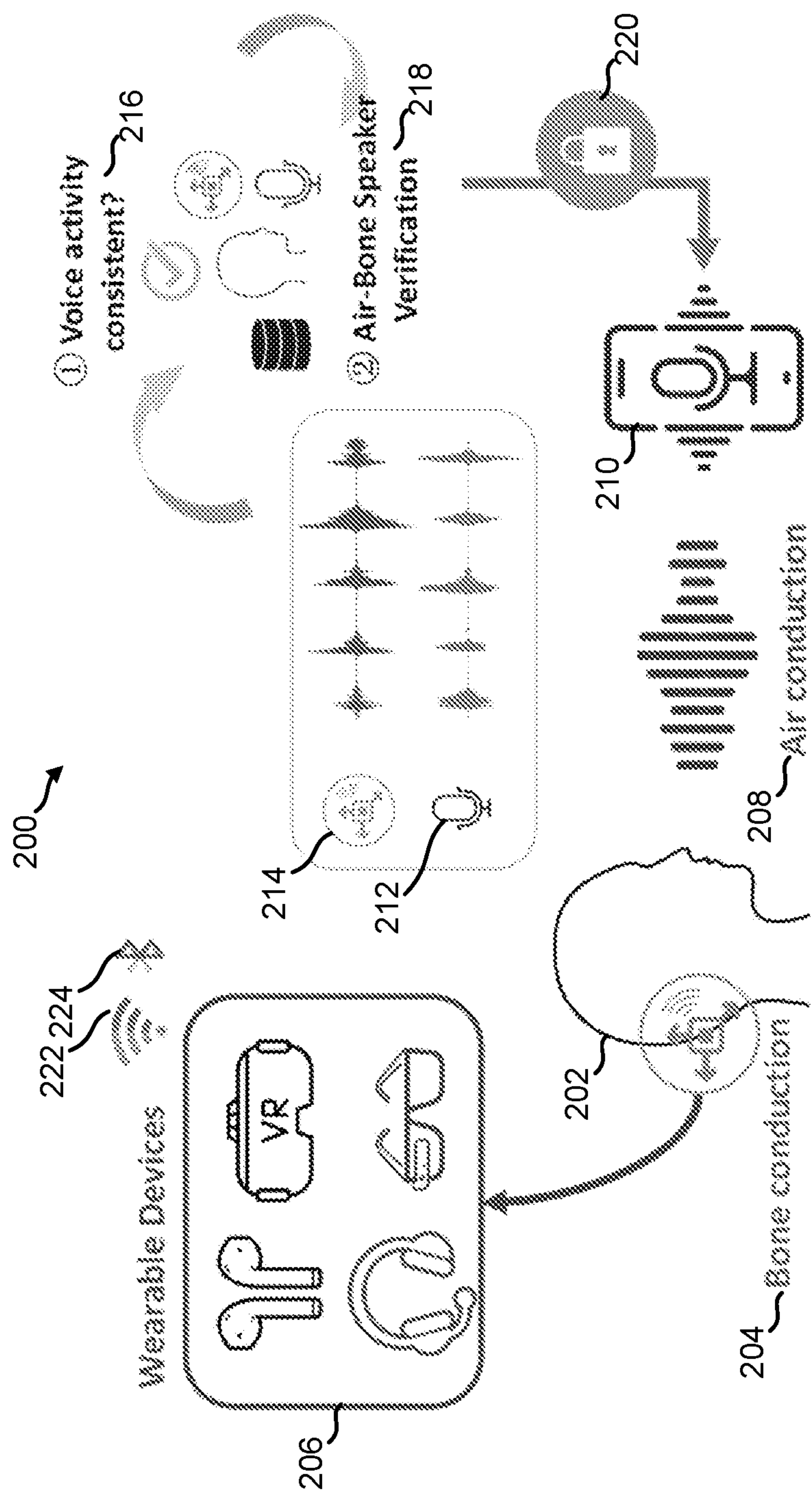


FIG. 2

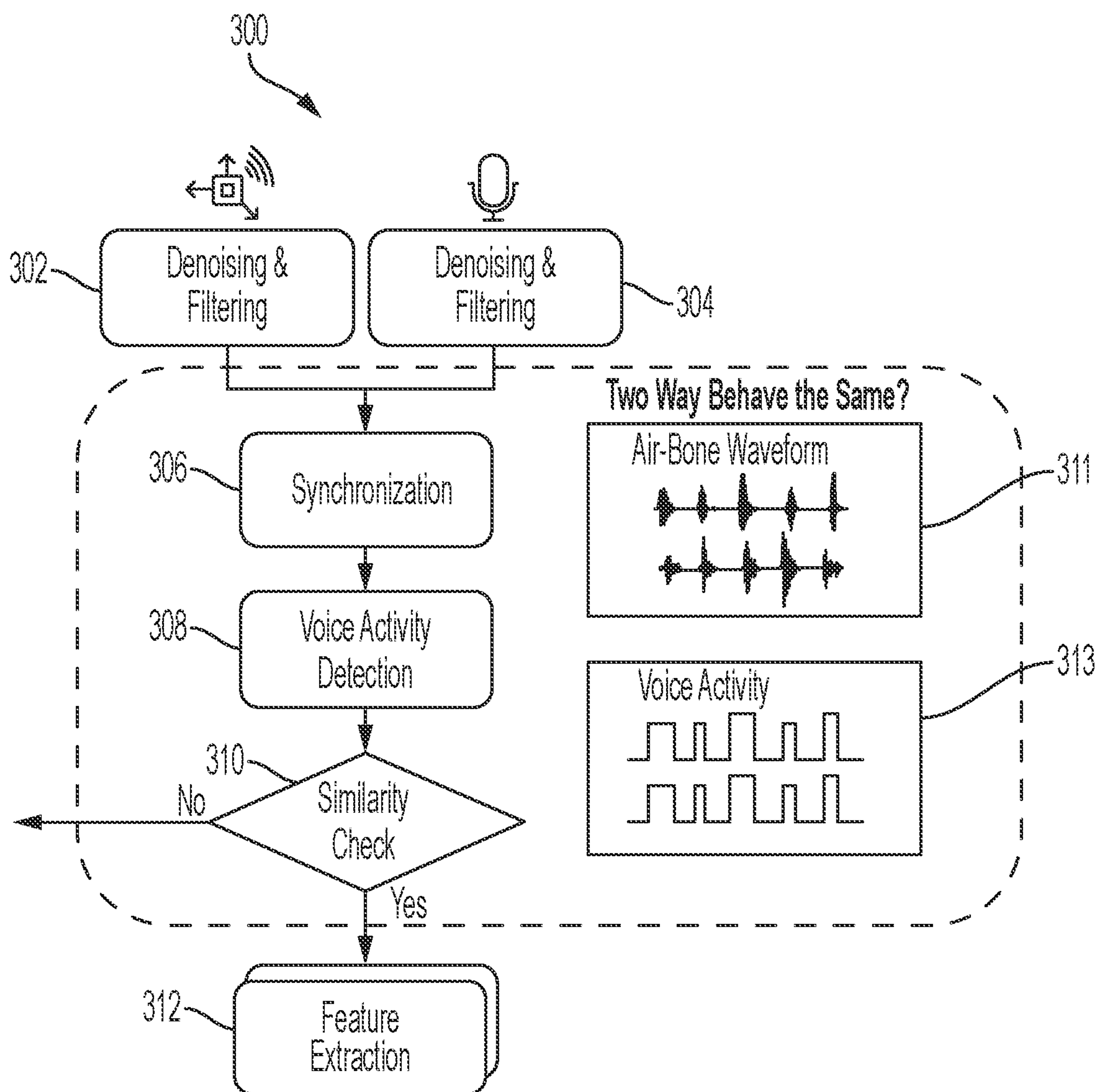


FIG. 3A

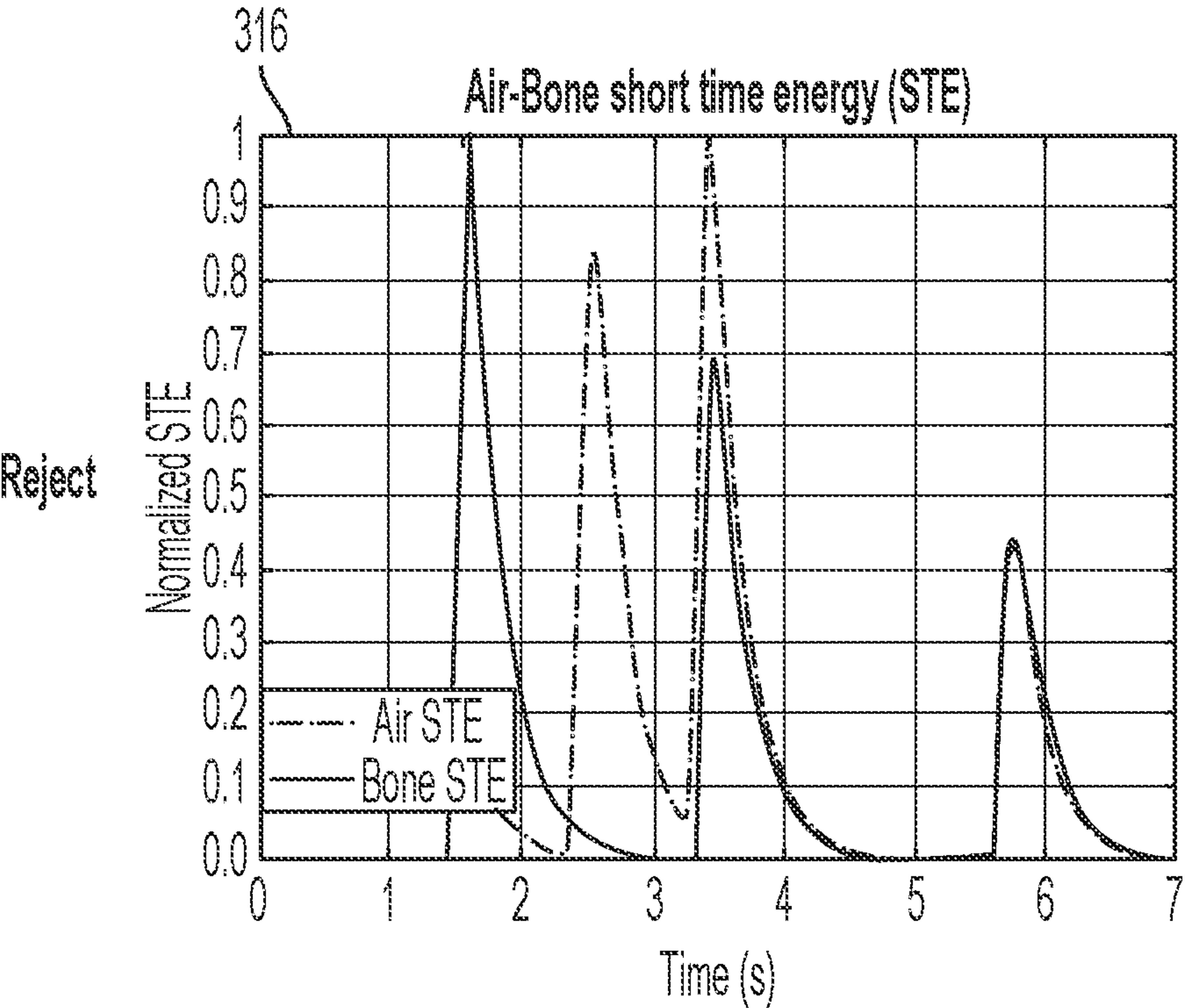
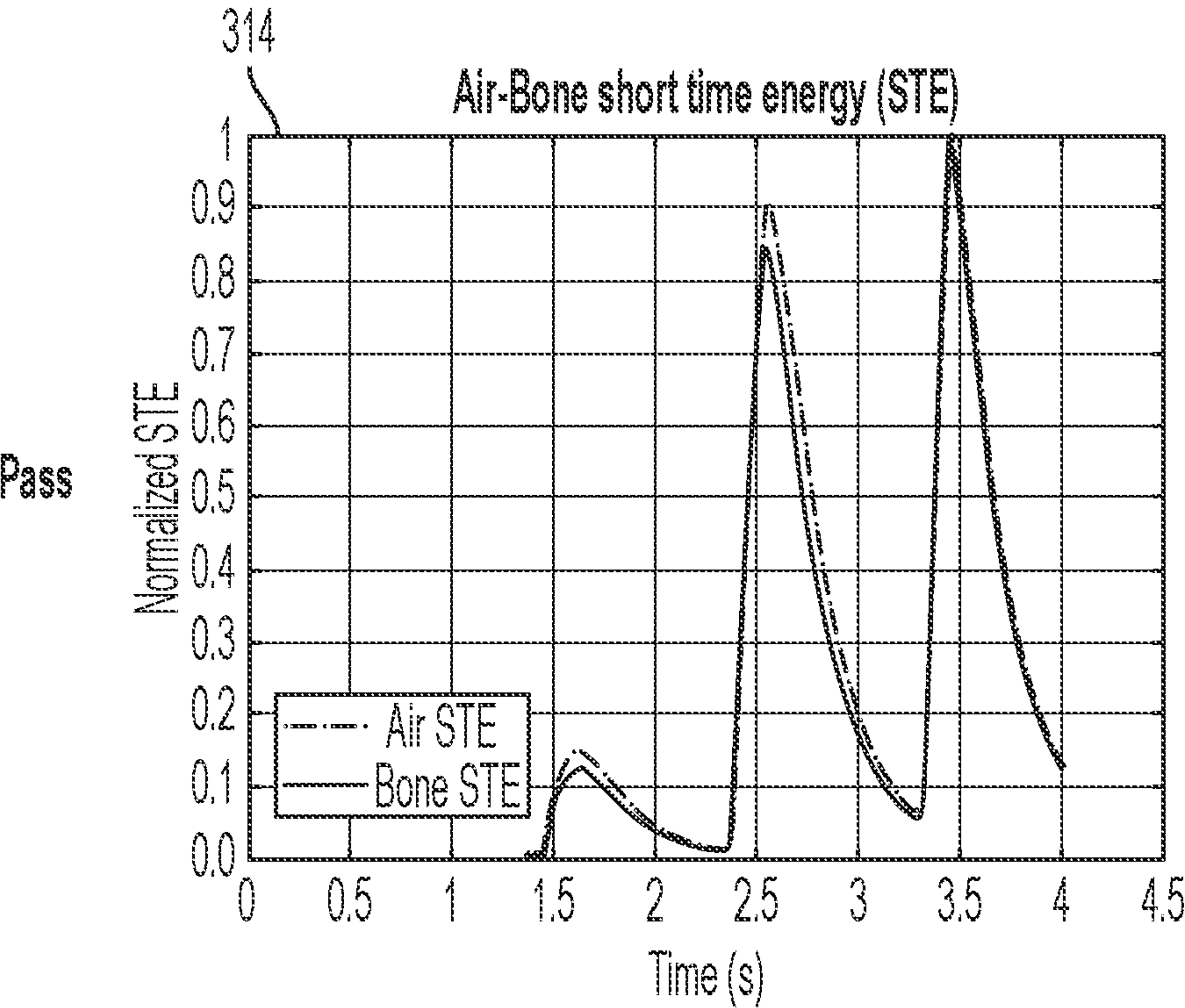


FIG. 3B

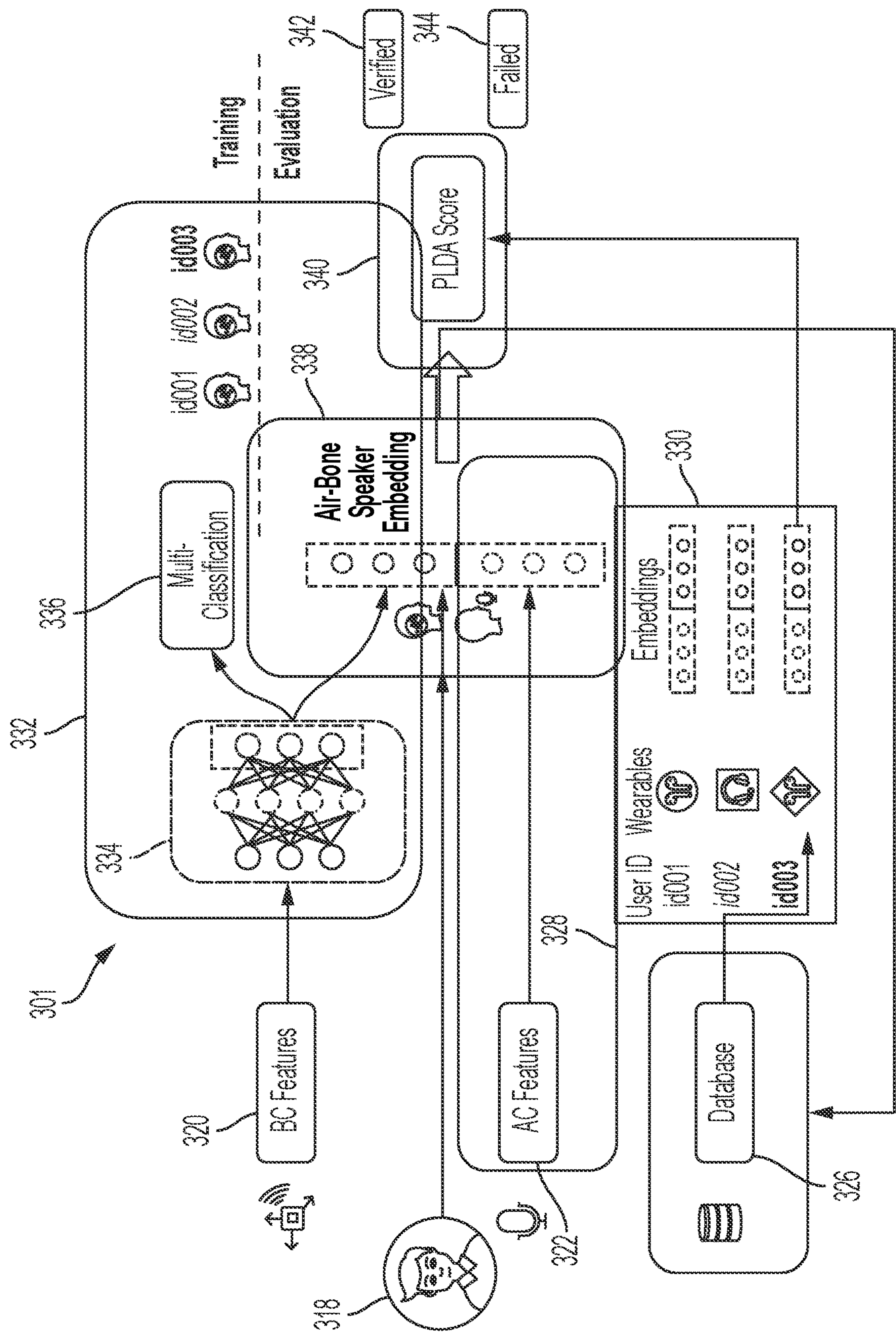


FIG. 3C

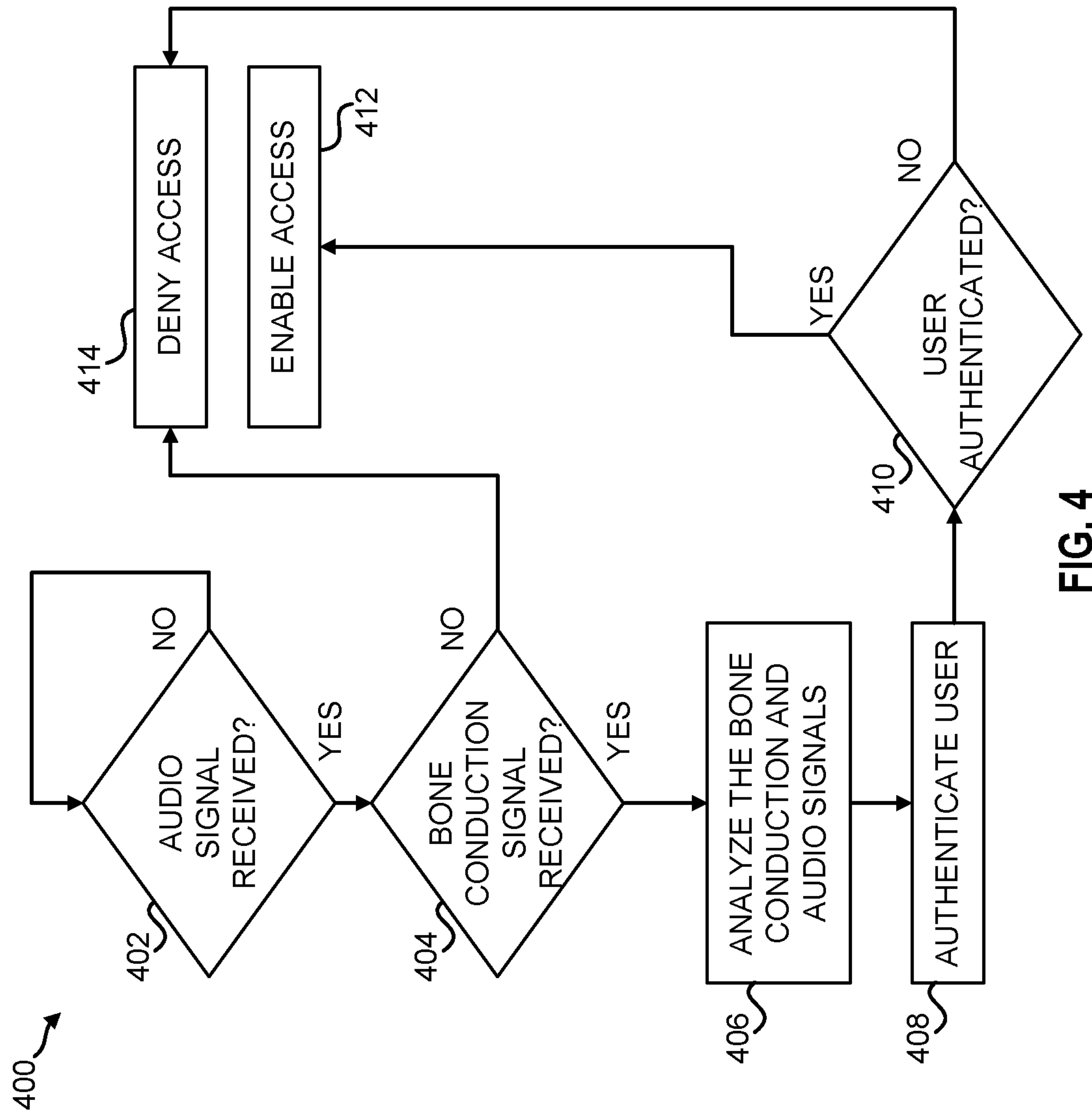


FIG. 4

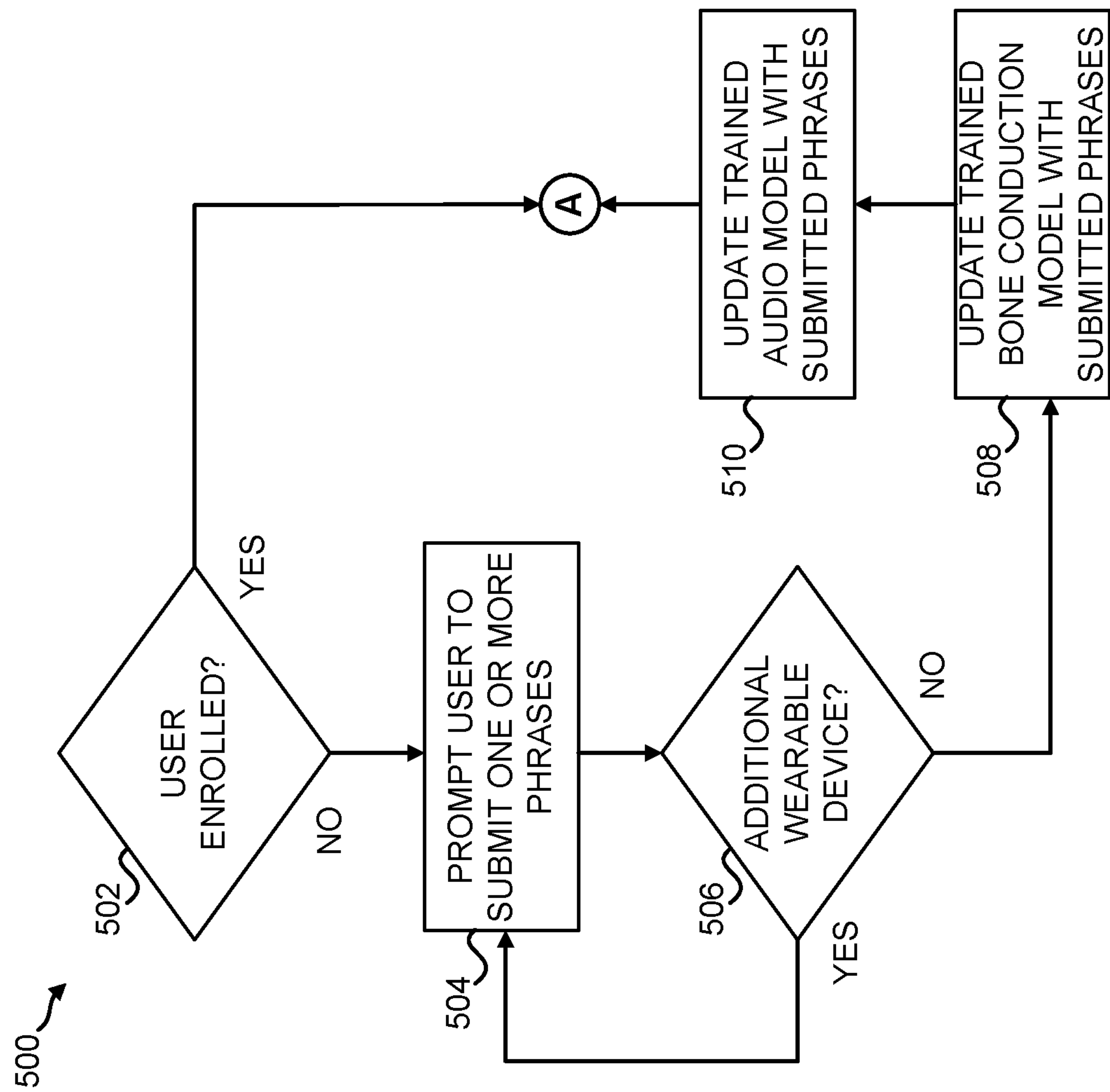


FIG. 5A

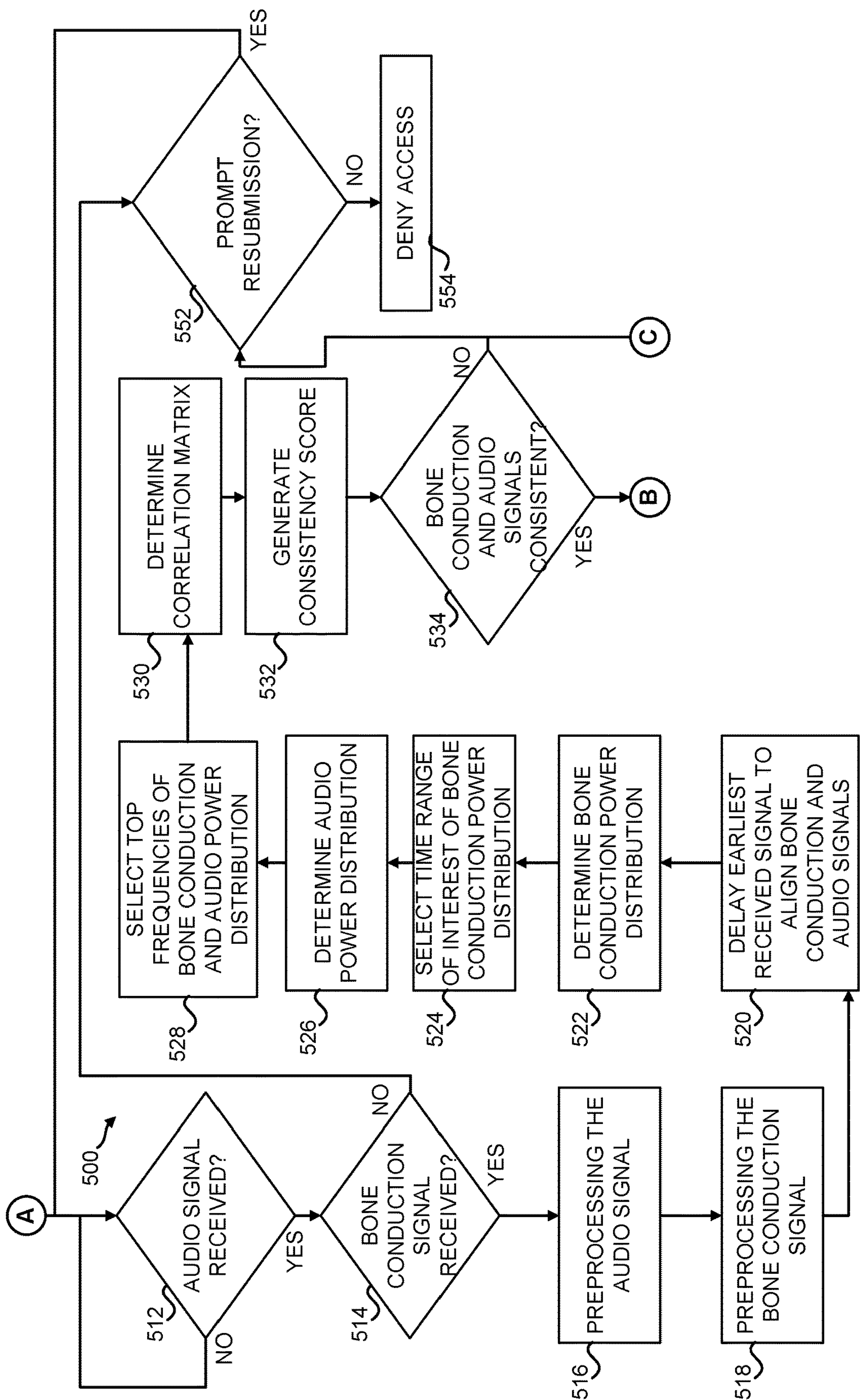


FIG. 5B

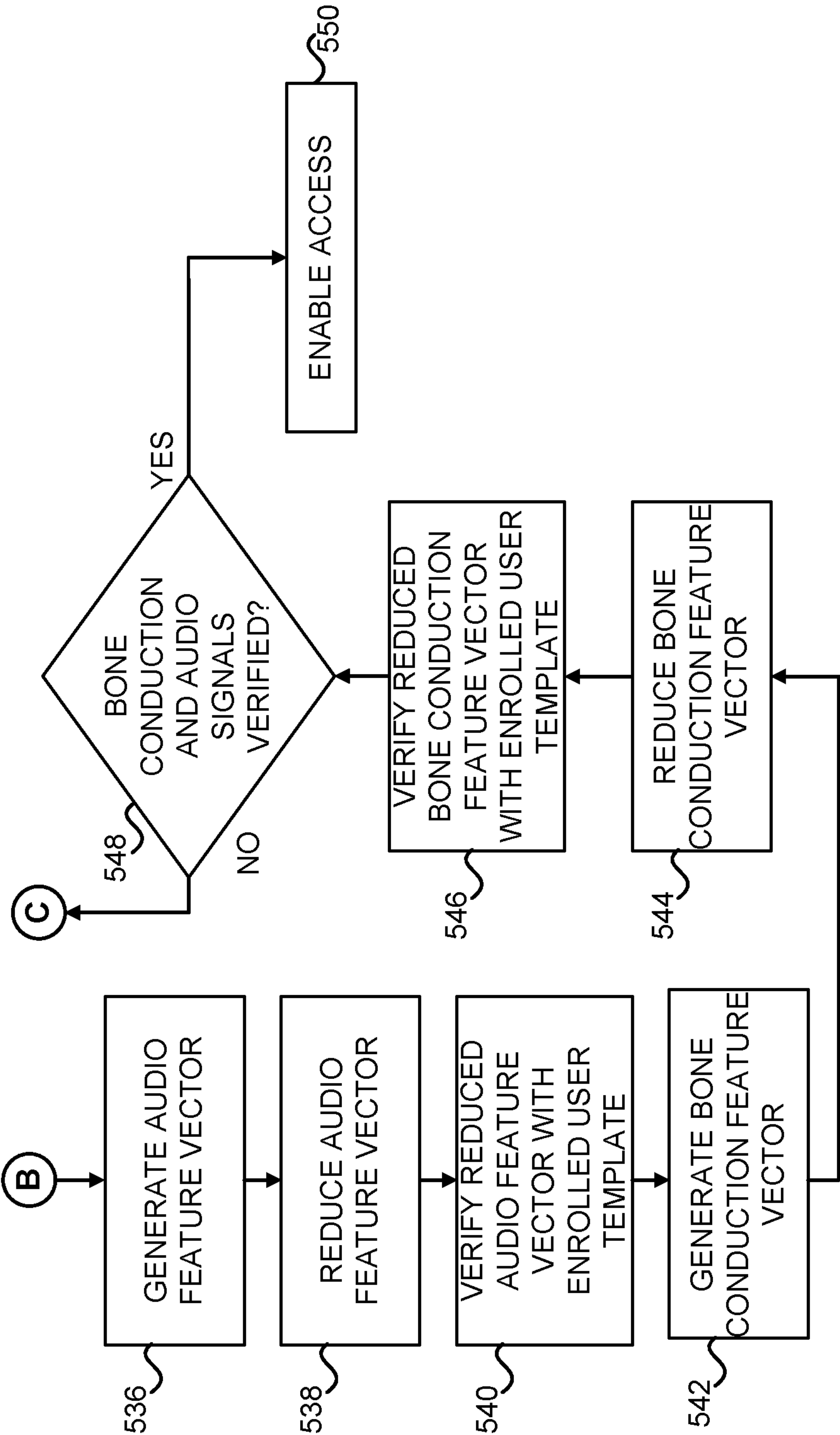


FIG. 5C

METHOD FOR MULTIFACTOR AUTHENTICATION USING BONE CONDUCTION AND AUDIO SIGNALS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is related to U.S. Provisional Application No. 63/268,999, filed Mar. 8, 2022, titled “SYSTEMS AND APPARATUS FOR MULTIFACTOR AUTHENTICATION USING BONE CONDUCTION AND AUDIO SIGNALS,” U.S. Provisional Application No. 63/269,001, filed Mar. 8, 2022, titled “METHOD FOR MULTIFACTOR AUTHENTICATION USING BONE CONDUCTION AND AUDIO SIGNALS,” and U.S. Provisional Application No. 63/380,229, filed Oct. 19, 2022, titled “SYSTEMS AND METHODS FOR CONTINUOUS, ACTIVE, AND NON-INTRUSIVE USER AUTHENTICATION,” the disclosures of which are incorporated herein by reference in their entirety.

TECHNICAL FIELD

[0002] The disclosure relates to methods and systems for using a bone conduction signal and an audio signal for two-way or multifactor authentication of user. More specifically, the methods and systems use a bone conduction signal from a wearable device and an audio signal from a smart device to authenticate a user of the smart device.

BACKGROUND

[0003] Smart voice assistants or smart devices typically include a vocal or audio based authentication process, if any authentication at all. For example, a user may speak a specific phrase or statement and the smart device may recognize the user based on the unique audio characteristics of the user’s vocals when speaking the specific phrase. Such audio characteristics are usually determined during an enrollment or initialization period. However, a user’s submission during enrollment or initialization may not match a current attempt to authenticate due to differences in tone and volume, thus resulting in a mismatch or denial of access to the smart device. Further, a third party may spoof or copy a user’s audio signal and utilize such a spoof or copy of the audio signal to imitate the user, thus gaining access to, potentially, sensitive and/or private user data or information.

SUMMARY

[0004] Accordingly, Applicants have recognized a need for systems and methods to utilize two-factor or multifactor authentication to enable user access to a smart device, the two-factor or multifactor authentication including verifying a user via a bone conduction signal from a wearable device and an audio signal obtained by the smart device. The present disclosure is directed to embodiments of such systems and methods.

[0005] The present disclosure is generally directed to systems and methods for utilizing two-factor or multifactor authentication to enable user access to a smart device, the two-factor or multifactor authentication including verifying a user via a bone conduction signal from a wearable device and an audio signal obtained by the smart device. In such embodiments, a smart device may request an input from a user prior to access to a portion of or all of the functionality and/or data of the smart device. Prior to a request for such

an input, the smart device may prompt the user to initialize or enroll with the smart device. Such an enrollment or initialization may include prompting a user to speak a particular phrase or phrases using each of one or more wearable devices associated with the user. In other words, the user may speak or submit a phrase for each wearable device the user desires to utilize for subsequent verifications and/or authentications. In such embodiments, the smart device and/or an authentication circuitry may store the audio signals, obtained from a microphone or other sensor of the smart device, and bone conduction signals, obtained from each of the one or more wearable devices. Further, an identifier or tag may be stored alongside each submission or entry, such an identifier or tag associated with one of the one or more wearable devices. The identifier or tag may be a device name obtained via data in a connection (for example, Bluetooth, WiFi, or other signal communication standard) between the wearable device and the smart device and/or authentication circuitry. Further still, more than one user may access a smart device. In such embodiments, different users may be able to access different functionality and/or data of the smart device. In such examples, when initializing or enrolling a user, a user identifier may be generated and may be stored with associated wearable devices and level of access. In an embodiment, functionality may include the smart device, after determining consistency and verifying the bone conduction signal and audio signal, utilizing natural language processing to interpret voice commands from the user and, based on the interpretation, performing an action or operation. The smart device may be configured to provide the additional functionality of granting access to virtual or physical locations and/or data, enabling a user to purchase goods and/or utilize services (for example, play music, view video of cameras connected to and/or proximate the smart device, among other services and/or functions), allowing a user to access a computing device or user profile, among other functionality.

[0006] The smart device and/or authentication circuitry may include one or more trained models. The smart device and/or authentication circuitry may include a bone conduction verification model and an audio conduction verification model. The bone conduction verification model may be trained based on previous users’ bone conduction signals, previous users’ enrollment or initialization bone conduction signals, and the outcome of previous verifications. Further, when a user submits a bone conduction signal enrollment or initialization, the bone conduction signal enrollment or initialization may be utilized to refine, re-train, or further train the bone conduction verification model. The audio conduction verification model may be trained based on previous users’ audio signals, previous users’ enrollment or initialization audio signals, and the outcome of previous verifications. Further, when a user submits an audio signal enrollment or initialization, the audio signal enrollment or initialization may be utilized to refine, re-train, or further train the audio conduction verification model.

[0007] As noted, prior to providing access to some or all of the functionality of the smart device, the smart device and/or authentication circuitry may request authentication and/or verification. In response to such a prompt, the smart device and/or authentication circuitry may wait until an audio signal and/or bone conduction signal is received. Once one signal (for example, the audio signal and/or bone conduction signal) is received by the smart device and/or

authentication circuitry, the smart device and/or authentication circuitry may wait for the other signal to be received (for example, the bone conduction signal and/or the audio signal, respectively). If the other signal is not received within a pre-selected or specified time interval, the smart device and/or authentication circuitry may deny the user access to some or all functionality (for example, a user may play a song, but not access private or sensitive data or information relating to the user). Further, the smart device and/or authentication circuitry may prompt the user to resubmit a phrase or phrases for authentication or verification for an amount of pre-determined or pre-selected times.

[0008] Once the smart device and/or authentication circuitry receives both signals (for example, the bone conduction signal and the audio signal), the smart device and/or authentication circuitry may determine whether the bone conduction signals and audio signals are consistent. In other words, whether the bone conduction signals and audio signals are based on the same phrase and/or from the same user. Such a determination may include pre-processing each signal (for example, delaying or aligning one of the signals to match the other and/or reducing noise of the signals by passing the signals through a Wiener filter or other audio or signal filter, among other pre-processing steps), determining a consistency score, and/or determining whether the consistency score is greater than or equal to a consistency threshold. Such a consistency score may be based on similarities between the audio signal and/or bone conduction signal. In such embodiments, since the audio signal and bone conduction signal are from the same user, some aspects of each signal should be consistent. After consistency is determined, the audio signal may be verified and the bone conduction signal may be verified. Such verifications may occur simultaneously, substantially simultaneously, and/or in sequence. Once verified, the smart device and/or authentication circuitry may grant access to some or all functionality of the smart device to the user.

[0009] In other embodiments, rather than a smart device obtaining an audio signal to be utilized for two-factor or multi-factor authentication of a user, a device may obtain an identification signal via a sensor of the device. Such an identification signal may include a signal generated by one or more of a badge scan, an identification card scan, a retinal scan, a fingerprint scan, or other scan configured to obtain a signal to be utilized for authentication. In such embodiments, authentication of a user may include authenticating a user based on the bone conduction signal and the identification signal.

[0010] Accordingly, an embodiment of the disclosure is directed to a method for two-way authentication of a user. The method may include receiving a bone conduction signal from a user via one or more wearable devices. The method may include receiving an audio signal from the user via a microphone separate from the one or more wearable devices, the audio signal corresponding to the bone conduction signal. The method may include determining a consistency score for the audio signal and corresponding bone conduction signal. The method may include, in response to the consistency score being greater than or equal to a consistency threshold, verifying, using an audio conduction model (AC model), that the audio signal is associated with a particular user. The method may include verifying, using a bone conduction model (BC model), that the bone conduction signal is associated with the particular user. The method

may include, in response to verification of the audio signal and bone conduction signal, enabling, for the particular user, access to a smart device.

[0011] In an embodiment, the determination of the consistency score may include, (a) determining a marginal bone conduction power distribution with respect to time; (b) selecting a time range of interest based on an average of the margin marginal bone conduction power distribution with respect to time; (c) determining a marginal audio conduction and bone conduction power distribution with respect to frequency; (d) selecting the top frequency index M from the marginal bone conduction power distribution with respect to frequency and the top frequency index N from the marginal audio conduction power distribution with respect to frequency; (e) determining a correlation matrix between the top frequency index M and the top frequency index N; and (f) generating, based on the correlation matrix, the consistency score.

[0012] In an embodiment, the method may include, prior to determining the consistency score, pre-processing the bone conduction signal. The pre-processing of the bone conduction signal may include passing the bone conduction signal through (a) a low-pass filter to remove noise generated by human motion and (b) a Wiener filter to remove noise. The method may include, prior to determining the consistency score, passing the audio signal through a Wiener filter to remove noise. The method may include, prior to determination of a consistency score and verification, prompting a user to submit an initial or enrollment (a) bone conduction signal and (b) audio signal. The AC model may be trained using a plurality of responses prior to submission of an initial or enrollment audio signal. The AC model may further be trained using a submitted initial or enrollment audio signal. The BC model may be a convolutional neural network (CNN). The CNN may be trained using the initial or enrollment bone conduction signal.

[0013] In an embodiment, the method may include, if one or more of the audio signal or bone conduction signal are not verified, preventing the particular user from accessing the smart device. The method may include, prior to determining the consistency score, delaying the earliest received of the bone conduction signal and the audio signal to align the bone conduction signal and the audio signal.

[0014] Another embodiment of the disclosure is directed to a non-transitory machine-readable storage medium storing processor-executable instructions that, when executed by at least one processor. Such execution may cause the at least one processor to receive a bone conduction signal from a user via one or more wearable devices. The execution may cause the at least one processor to receive an audio signal from the user via a microphone separate from the one or more wearable devices, the audio signal corresponding to the bone conduction signal. The execution may cause the at least one processor to determine a consistency score for the audio signal and corresponding bone conduction signal. The execution may cause the at least one processor to, in response to the consistency score being greater than or equal to a consistency threshold: (a) verify, using an audio conduction model (AC model), that the audio signal is associated with a particular user; (b) verify, using a bone conduction model (BC model), that the bone conduction signal is associated with the particular user; and (c) in response to verification of the audio signal and bone conduction signal, enable, for the particular user, access to a smart device.

[0015] In an embodiment, the smart device may include the microphone. The bone conduction signal may be received via wireless communication. The consistency threshold may be based on similarities between bone conduction signals and audio signals that indicate the bone conduction signal and the audio signal are from the particular user.

[0016] Another embodiment of the disclosure is directed to a method for two-way authentication of a user. The method may include prompting a particular user to submit initial or enrollment (a) audio signals and (b) bone conduction signals. The method may include updating an audio conduction model (AC model) and a bone conduction model (BC model) based on the received initial or enrollment (a) audio signal and (b) bone conduction signals. The method may include, after reception of the initial or enrollment (a) audio signal and (b) bone conduction signals, receiving a bone conduction signal from a user via one or more wearable devices. The method may further include receiving an audio signal from the user via a microphone separate from the one or more wearable devices, the audio signal corresponding to the bone conduction signal. The method may include determining a consistency score for the audio signal and corresponding bone conduction signal. The method may include, in response to the consistency score being greater than or equal to a consistency threshold, verifying, using the AC model, that the audio signal is associated with a particular user. The method may include verifying, using the BC model, that the bone conduction signal is associated with the particular user. The method may include, in response to verification of the audio signal and bone conduction signal, enabling, for the particular user, access to a smart device.

[0017] In an embodiment the initial or enrollment audio signal may be received via the microphone. The microphone may be included in the smart device. The initial or enrollment bone conduction signals may be received from one of one or more wearable devices. The method may include prompting the particular user to submit initial or enrollment (a) audio signals and (b) bone conduction signals for each of the one or more wearable devices. The initial or enrollment (a) audio signals and (b) bone conduction signals may include one or more specific phrases. The BC model may be a convolutional neural network (CNN). The CNN may be trained using the initial or enrollment bone conduction signal to generate corresponding initial embedded bone conduction vectors. The method may include, prior to verification via the BC model, generating embedded bone conduction vectors using the bone conduction signals and the initial embedded bone conduction vectors.

[0018] Still other aspects and advantages of these embodiments and other embodiments, are discussed in detail herein. Moreover, it is to be understood that both the foregoing information and the following detailed description provide merely illustrative examples of various aspects and embodiments, and are intended to provide an overview or framework for understanding the nature and character of the claimed aspects and embodiments. Accordingly, these and other objects, along with advantages and features herein disclosed, will become apparent through reference to the following description and the accompanying drawings. Furthermore, it is to be understood that the features of the various embodiments described herein are not mutually exclusive and may exist in various combinations and permutations.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] These and other features, aspects, and advantages of the disclosure will become better understood with regard to the following descriptions, claims, and accompanying drawings. It is to be noted, however, that the drawings illustrate only several embodiments of the disclosure and, therefore, are not to be considered limiting of the scope of the disclosure.

[0020] FIG. 1A, FIG. 1B, FIG. 1C, FIG. 1D, FIG. 1E, and FIG. 1F are block diagrams of systems to authenticate a user, according to an embodiment of the present disclosure.

[0021] FIG. 2 is a block diagram of a system to authenticate a user, according to an embodiment of the present disclosure.

[0022] FIG. 3A, FIG. 3B, and FIG. 3C are block diagrams of systems to authenticate a user, according to an embodiment of the present disclosure.

[0023] FIG. 4 is a flow diagram of two-way user authentication, according to an embodiment of the present disclosure.

[0024] FIG. 5A, FIG. 5B, and FIG. 5C are flow diagrams of two-way user authentication, according to an embodiment of the present disclosure.

DETAILED DESCRIPTION

[0025] So that the manner in which the features and advantages of the embodiments of the systems and methods disclosed herein, as well as others that will become apparent, may be understood in more detail, a more particular description of embodiments of systems and methods briefly summarized above may be had by reference to the following detailed description of embodiments thereof, in which one or more are further illustrated in the appended drawings, which form a part of this specification. It is to be noted, however, that the drawings illustrate only various embodiments of the systems and methods disclosed herein and are therefore not to be considered limiting of the scope of the systems and methods disclosed herein as it may include other effective embodiments as well.

[0026] The present disclosure is directed to systems and methods for two-way authentication or multifactor authentication for a user. Such authentication may occur in response to a user's attempt to access a device or smart device. Typically, a smart device or other device may utilize audio or actual passwords or pins to authenticate a user. However, a user's speech may be spoofed or copied or a user's password determined or stolen, thus allowing a third party to take control of a user's account. To prevent such issues, the systems and methods described herein may utilize a two-factor or multifactor authentication. For example, a user attempting to access functionality or data associated with a particular device or smart device may be prompted to submit a phrase or one or more phrases. The data may include personal data, data associated with the user, data stored on the smart device, data associated with the smart device, and/or data stored in a specified location accessible via the smart device. Functionality of the smart device may include (for example, the smart device may be configured to) the smart device accepting voice commands, the smart device performing an action based on the received voice commands, granting access to virtual or physical locations and/or data, and/or enabling a user to purchase goods and/or utilize services (for example, play music, view

video of cameras connected to and/or proximate the smart device, among other services and/or functions), among other functionality. In an example, the smart device may, after determining consistency and verifying the bone conduction signal and audio signal, utilize natural language processing to interpret voice commands from the user and, based on the interpretation, perform an action or operation. Further, to ensure proper authentication, the user may submit such a phrase or one or more phrases while wearing a wearable device. While the smart device may include a microphone or other sensor to obtain an audio signal or other identification signal from the user, the wearable device may be configured to obtain a bone conduction signal from the user and transmit the bone conduction signal to the smart device and/or an authentication circuitry. Thus, at least two different factors may be gathered to authenticate a user.

[0027] Further, such systems and methods may authenticate the user when the bone conduction signal and audio signal or other signal are received. In such examples, the smart device and/or authentication circuitry may first determine whether the bone conduction signal and audio signal or other signal are consistent. If the signals (for example, the bone conduction signal and audio signal or other signal) are consistent, then each signal (for example, the bone conduction signal and audio signal or other signal) may be verified. If each signal (for example, the bone conduction signal and audio signal or other signal) is verified, then the smart device and/or authentication circuitry may authenticate the user.

[0028] The systems and methods may utilize various machine learning and/or signal processing techniques to analyze the bone conduction and the audio signal or other signal. Additionally, machine learning models or classifiers used may be trained based on a user's enrollment or initial submission of one or more phrases. In other words, the smart device and/or authentication circuitry may prompt the user to submit the enrollment or initialization. Such an enrollment or initialization may be prompted for each wearable device the user utilizes (for example, a headphone, headset, earbud, smart glasses, smart phone, virtual reality (VR) headset, and/or augmented reality (AR) headset, among other devices), as the interaction between such wearable devices and the user's skull may produce different bone conduction signals. Once enrolled or initialized, the smart device and/or authentication circuitry may update, refine, re-train, or further train any machine learning models, probabilistic models, and/or statistical models utilized in such systems and methods. Thus, when the smart device and/or authentication device verifies a bone conduction signal and/or audio signal or other signal, the smart device and/or authentication device may grant the user access to all or some portion of the functionality and/or data for the smart device or other devices.

[0029] A solution to such issues, as noted, include the use of a wearable device in conjunction with a microphone or other sensor of a smart device or other device. The user can provide an audio signal and bone conduction signal and, based on such signals, be authenticated. Other and/or different signals may be utilized in the systems and methods. Further, a user may submit an enrollment or initialization for each of one or more wearable devices the user owns or that the user wishes to utilize for authentication.

[0030] FIG. 1A, FIG. 1B, FIG. 1C, FIG. 1D, FIG. 1E, and FIG. 1F are block diagrams of systems to authenticate a user, according to an embodiment of the present disclosure.

Turning first to FIG. 1A, the system 100 may include a wearable device 102 and a smart device 112. In such embodiments, a user 108 may be wearing the wearable device 102. The wearable device 102 may be a number of different devices, for example, such as audio headphones (wired or wireless), a headset, smart glasses, earbuds, a VR headset, an AR headset, a smart watch, and/or another wearable device including a bone conduction sensor. In another embodiment, rather than a wearable device 102, another device may be utilized to gather, obtain, or receive a bone conduction signal 110 via a bone conduction sensor, for example, such as a smart phone with an embedded bone conduction sensor. Such wearable devices 102, describe herein, may include a sensor 104, such as a sensor 104 configured to detect or sense bone conduction signals from a user 108, and communications circuitry 106. The sensor 104, in an example, may be an accelerometer to measure the vibrations created when the user 108 speaks and that travel through the user's 108 skull. The communications circuitry 106 may be configured to allow communication with the smart device 112 and/or an authentication circuitry (for example, authentication circuitry in FIG. 1C). Such a configuration may be based on a Bluetooth, WiFi, and/or other communication standard. The wearable device may detect or sense bone conduction signal 110 from a user when a user speaks. The bone conduction signal 110 may then be transmitted to the smart device and/or authentication circuitry. In an embodiment, the user 108 may utilize one or more different wearable devices at different times.

[0031] A user may include a person attempting to utilize the smart device 112. A user may also include a person, third person, and/or designee who may be directed, permitted, and/or appointed by a person who is verified as a user or accessor able to authorize other users.

[0032] The system 100 in FIG. 1A may illustrate, at a high level, the interaction between the wearable device 102, the smart device 112, and the user 108, when the user 108 seeks authentication from and/or access to the smart device 112 (for example, to perform one or more functions, including, but not limited to, operating a drone, playing music or videos, accessing user data or information, and/or performing smart home functions, among other operations or functions). In such embodiments, the smart device 112 may include a voice enabled smart assistant (for example, Siri, Alexa, Google, among others), a voice controlled device, a smart voice controlled device, a connected or internet of things (IoT) enabled device, smart appliances, a remote control, a television, a video game entertainment system, a household internet enabled device, a household smart device, a voice controlled intelligent personal assistant device, a voice controlled intelligent personal vehicle management device, computing devices (for example, cell phones, smart phones, laptops, desktops, tablets, among other computing devices), among other smart devices configured to receive an input from a user and, based on the input, perform a function or operation.

[0033] In an embodiment, and as illustrated in FIG. 1A, the smart device 112 may include a processor 122, a memory 124 (for example, a non-transitory machine readable medium), a microphone 114 or other sensor, and/or a communications circuitry 116. In some examples, the smart device 112 may be a computing device. The term "computing device" is used herein to refer to any one or all of servers, virtual computing device or environment, desktop comput-

ers, personal data assistants (PDAs), laptop computers, tablet computers, smart books, palm-top computers, personal computers, smartphones, wearable devices (such as headsets, smartwatches, or the like), and similar electronic devices equipped with at least a processor and any other physical components necessarily to perform the various operations described herein. Devices such as smartphones, laptop computers, tablet computers, and wearable devices are generally collectively referred to as mobile devices.

[0034] The term “server” or “server device” is used to refer to any computing device capable of functioning as a server, such as a master exchange server, web server, mail server, document server, or any other type of server. A server may be a dedicated computing device or a server module (for example, an application) hosted by a computing device that causes the computing device to operate as a server. A server module (for example, server application) may be a full function server module, or a light or secondary server module (for example, light or secondary server application) that is configured to provide synchronization services among the dynamic databases on computing devices. A light server or secondary server may be a slimmed-down version of server type functionality that can be implemented on a computing device, such as a smart phone, thereby enabling it to function as an Internet server (for example, an enterprise e-mail server) only to the extent necessary to provide the functionality described herein.

[0035] As used herein, a “non-transitory machine-readable storage medium” may be any electronic, magnetic, optical, or other physical storage apparatus to contain or store information such as executable instructions, data, and the like. For example, any machine-readable storage medium described herein may be any of random access memory (RAM), volatile memory, non-volatile memory, flash memory, a storage drive (for example, a hard drive), a solid state drive, any type of storage disc, and the like, or a combination thereof. The memory may store or include instructions executable by the processor.

[0036] As used herein, a “processor” or “processing circuitry” may include, for example one processor or multiple processors included in a single device or distributed across multiple computing devices. The processor (for example, processor 122 shown in FIG. 1A) may be at least one of a central processing unit (CPU), a semiconductor-based microprocessor, a graphics processing unit (GPU), a field-programmable gate array (FPGA) to retrieve and execute instructions, a real time processor (RTP), other electronic circuitry suitable for the retrieval and execution instructions stored on a machine-readable storage medium, or a combination thereof.

[0037] The smart device 112, as noted, may include a memory 124 (for example, non-transitory machine readable storage medium). The memory 124 may include instructions or one or more sets of instructions. The instructions may be executed by the processor 122 of the smart device 112. Such instructions may include instructions to authenticate the user 108. The instructions to authenticate a user 108 may comprise one or more sets of instructions, sub-instructions, modules, and/or other routines or sub-routines. For example, such instructions may include a first set of instructions. The first set of instructions may include instructions to, when executed, determine whether a received bone conduction signal 110 is consistent with an audio signal 120. Prior to or in conjunction with execution of the first set of instructions

to determine consistency, instructions (for example, a second set of instructions) to pre-process the bone conduction signal 110 and audio signal 120 may be executed. Such pre-processing may include aligning the bone conduction signal 110 with the audio signal 120 (for example, by delaying an earliest received signal to align similar features or aspects of the signals) and passing each of the bone conduction signal 110 and the audio signal 120 through a noise reduction program, algorithm, and/or filter (for example, a Wiener filter). In another embodiment, pre-processing may include converting the bone conduction signal 110 and audio signal 120 to a binary waveform, for determining consistency.

[0038] After pre-processing, consistency of the signals may be determined via execution of the instructions noted above. For example, filtered and/or noise reduced signal waveforms (for example, waveforms for the bone conduction signal 110 and audio signal 120) may be compared to determine whether the waveforms or features of the waveforms are similar, substantially similar, or consistent. The instructions, when executed, may determine a probability or score based on an amount of similar or consistent features or an amount of differences between the waveforms. The probability or score may indicate whether or not the signals (for example, the bone conduction signal 110 and audio signal 120) are consistent. In such examples, the smart device 112 or instructions may utilize a threshold or consistency threshold to determine whether the generated probability or score indicates consistency or not. For example, if the generated probability or score is greater than or equal to the consistency threshold, then the instructions, when executed, may determine that the signals (for example, the bone conduction signal 110 and audio signal 120) are consistent, while if the generated probability or score is less than the consistency threshold, then the instructions, when executed, may determine that the signals (for example, the bone conduction signal 110 and audio signal 120) are not consistent. If the signals (for example, the bone conduction signal 110 and audio signal 120) are determined to be not consistent, instructions, when executed, may notify the user 108 of inconsistent signals (for example, such a notification may be generated and sent to the user 108). Further, if the bone conduction signal 110 and audio signal 120 are inconsistent, then instructions to prompt the user 108 to resubmit a phrase or phrases for authentication may be executed. In an embodiment, a prompt for the user 108 to resubmit a phrase or phrases may be submitted a predetermined or preselected number of times. After the prompt to resubmit has been given the predetermined or preselected number of times and the user 108 has submitted corresponding responses resulting in additional inconsistent bone conduction signals 110 and audio signals 120, the user 108 may be locked out of the smart device 112 until the user’s 108 identity can be verified (for example, via an email or phone call) or until after a predetermined amount of time has passed. In such examples, locking a user may prevent the user from accessing data on or associated with the smart device 112 and/or utilizing functionality of the smart device 112. Further, after one or more attempts or a pre-determined or pre-selected number of attempts, the smart device 112 may generate a notification indicating a number of unsuccessful attempts at access. The smart device 112 may then transmit such a notification to the user 108. Further, the

notification may be sent via an alternative form of communication, such as email, text message, phone call, or other form of communication.

[0039] If the bone conduction signal 110 and audio signal 120 are determined to be consistent, instructions to verify each signal may be executed. The steps or operation for verification may include passing the bone conduction signal 110 and audio signal 120 through a corresponding trained model or classifier. Each corresponding trained model or classifier may be trained using a number of previous or pre-collected and stored signals and outcomes (for example, previously submitted bone conduction signals 110 or audio signals 120, previously submitted enrollment bone conduction signals 110 or audio signals 120, and an indicator to indicate whether the bone conduction signals 110 or audio signals 120 were verified). The trained model or classifier may also be re-trained or further trained using an initial or enrollment submission from the user 108 (for example, an initial or enrollment bone conduction signal or audio signal). In an embodiment, the trained model or classifier may be a supervised or unsupervised learning model. In an embodiment, the trained model or classifier may be based on one or more of decision trees, random forest models, random forests utilizing bagging or boosting (as in, gradient boosting), neural network methods, support vector machines (SVM), other supervised learning models, other semi-supervised learning models, other unsupervised learning models, or some combination thereof, as will be readily understood by one having ordinary skill in the art. Other types of models may be utilized to verify a bone conduction signal 110 or audio signal 120, such as meta-analytical model or another statistical or probabilistic model. Upon verification or not of the bone conduction signal 110 and audio signals 120, instructions may be executed to grant or deny access of the user 108 to the smart device 112.

[0040] Turning to FIG. 1B, the smart device 112 may include an authentication circuitry 118. The authentication circuitry 118 may include or may perform the same or similar steps or operations as described herein for the instructions stored in the memory 124 and executed by the processor 122. The authentication circuitry 118 may be included in or may be a part of the processor 122 and/or memory 124. The authentication circuitry 118 may be a separate circuit in signal communication with the processor 122 and memory 124. The authentication circuitry 118, as illustrated in FIG. 1C for example, may be separate from the smart device 112. The authentication circuitry 118, in such examples, may be a cloud computing service, may be stored on a server, may be accessible via a network (for example, local area network, wide area network, wireless network, and/or other network capable of allowing communication between the smart device 112, authentication circuitry 118, and/or the wearable device 102), or may be stored or be a part of circuitry in signal communication with many smart devices and wearable devices. The authentication circuitry 118, for example, may reside or may be a part of a server in signal communication with the smart device 112 and/or wearable device 102. The authentication circuitry 118 may connect to the smart device 112 and wearable device 102 via the communications circuitry of each (for example, communications circuitry 116 and communications circuitry 106, respectively). In another embodiment, the authentication circuitry 118 may be included in another device. The other device may utilize the authentication circuitry to

authenticate users 108. Other devices may include a security checkpoint (for example, a location where a user 108 swipes a badge or enters a code to access).

[0041] The wearable device 102, as shown in FIG. 1C, may connect to the authentication circuitry 118 to provide the bone conduction signal 110 to the authentication circuitry 118. In another embodiment, the wearable device 102 may connect to the smart device 112 and provide the bone conduction signal 110 to the authentication circuitry 118 via the smart device 112. In other words, the wearable device 102 may provide a bone conduction signal 110 to the smart device 112 and the smart device 112 may then provide the bone conduction signal 110 to the authentication circuitry 118.

[0042] In an embodiment and referring to FIG. 1C, once the authentication circuitry 118 authenticates the user 108, the authentication circuitry 118 may transmit a signal to the smart device 112 indicating that the user 108 has been authenticated. Once such a signal is received by the smart device 112, the smart device 112 may allow or enable the user 108 to access the smart device 112 (for example, access features and/or functionality of the smart device 112 and/or data associated with or stored in the smart device 112).

[0043] Turning to FIG. 1D, the authentication circuitry 118 may include one or more circuits, chips, sub-circuits, sets of instructions or code, routines or sub-routines, or components. The components included in the authentication circuitry 118 may include audio preprocessing circuitry 126. The audio preprocessing circuitry 128 may preprocess the audio signal 120 by reducing noise. For example, the audio preprocessing circuitry 126 may include a low band pass filter, such as a Wiener filter, to remove background noise and other unwanted audio. A bone preprocessing circuitry 128 may preprocess the bone conduction signal 110 by aligning or synchronizing the bone conduction signal 110 with the audio signal 120. The bone preprocessing circuitry 128 may additionally reduce noise from the bone conduction signal 110, such as noise created by movement of the user. After pre-processing the bone conduction signal 110 and the audio signal 120, the preprocessed bone conduction signal 110 and the preprocessed audio signal 120 may be transmitted to a consistency score circuitry 130. The consistency score circuitry 130 may determine whether the bone conduction signal 110 and the audio signal 120 are consistent or, in other words, whether the bone conduction signal 110 and the audio signal 120 are or were produced by the user 108. Following a determination of consistency, the audio signal may be verified via an audio verification circuitry 132. The audio verification circuitry 132 may include a trained model or classifier to determine whether the audio signal 120 is from or originates from the user 108. Following audio verification, the bone conduction signal 110 may be verified via the bone verification circuitry 134. The bone verification circuitry 134 may include a trained model or classifier to determine whether the bone conduction signal 110 is from or originates from the user 108. In another embodiment, the bone conduction signal 110 may be verified prior to verification of the audio signal 120. In yet another embodiment, the audio signal 120 and bone conduction signal 110 may be verified in parallel or at a substantially similar time.

[0044] Turning to FIG. 1E and FIG. 1F, the system 100 may include an authentication device 144. The authentication device 144 may include the authentication circuitry 118 described herein or may include similar or substantially

similar functionality as the authentication circuitry **118**. The authentication device **144**, in such examples, may be included in or may be a cloud computing service. The authentication device **144** may be stored on or may be a server. In yet another embodiment, the authentication device **144** may be accessible via a network (for example, local area network, wide area network, wireless network, and/or other network capable of allowing communication between a first device **138**, authentication device **144**, and/or the second device **146**) or may be stored or be a part of circuitry or a computing device in signal communication with many different devices. The authentication device **144** may include a communications interface configured to receive signals from one or more devices.

[0045] The system **100** may include a first device **138** and a second device **146**. The first device **138** may include a computing device, a security device, or a device configured to receive a signal, communicate with the second device **146**, offer functionality or services, offer access to data, and/or offer physical or virtual access to a location, among other functionality. The first device **138** may include a sensor **142** and communications circuitry **140**. The sensor **142** may receive or sense a first signal **152** from the user **108** (for example, an identification signal). The first signal **152** may be an audio signal, a bone conduction signal, and/or signals associated with a badge scan, identification scan, retinal scan, fingerprint scan, facial scan (for example, facial recognition scan), gesture scan (for example, limb or body gesture recognition scan), and/or a scan of some other aspect of a user **108**. The first device **138** may include a communications circuitry **140**. The communications circuitry **140** may provide or transmit the first signal **152** to communications circuitry or a communications interface of the authentication device **144**.

[0046] As noted, the system **100** may include a second device **146**. The second device **146** may include a sensor **148** and a communications circuitry **150**. The sensor **148** may sense a second signal **154** from the user **108**. In an embodiment, the second signal **154** may be a different type of signal than the first signal **152** (for example, the first signal **152** may be an audio signal, while the second signal **154** may be a bone conduction signal). The second signal **154** may be an audio signal, a bone conduction signal, and/or signals associated with a badge scan, identification scan, retinal scan, fingerprint scan, facial scan (for example, facial recognition scan), gesture scan (for example, limb or body gesture recognition scan), and/or a scan of some other aspect of a user **108**. The second device **146** may include a communications circuitry **150**. The communications circuitry **150** may provide or transmit the second signal **154** to the communications circuitry **140** of the first device **138** (as illustrated in FIG. 1E) and/or the authentication device **144** (as illustrated in FIG. 1F). The second device **146**, in an embodiment, may be a device associated with and/or related to the user **108** or a personal device owned or utilized by the user **108**.

[0047] For example, the user **108** may generate a first signal **152** and a second signal **154**. The first signal **152** may be received or sensed by the sensor **142** of the first device **138**, while the second signal **154** may be received or sensed by the sensor **148** of the second device **146**. In another embodiment, additional signals may be utilized and/or sensed by the devices and/or other additional devices. The second signal **154** may be transmitted to the first device **138**

or the authentication device **144** via communications circuitry **150**. The first device **138** may then transmit the first signal **152**, in addition to, in some embodiments, the second signal **154**, to the authentication device **144** via communications circuitry **140**. The authentication device **144** may determine whether the signals are consistent (for example, from the same user **108**). The authentication device **144** may determine a consistency score based on the two signals. For example, a user **108** may swipe or move a badge against a badge scanner (for example, the first device **138**) and speak into an audio device (for example, the second device **146**), thus generating a first signal (for example, based on the badge scan) and a second signal (for example, based on the audio signal). The authentication device **144** may determine the consistency score based on whether the badge associated with the user **108** matches an audio signal from the same user **108**. As noted, different and/or additional signals may be utilized in such operations, such as a bone conduction signal and/or signals associated with an identification scan, retinal scan, fingerprint scan, facial scan (for example, facial recognition scan), gesture scan (for example, limb or body gesture recognition scan), and/or a scan of some other aspect of a user **108**. After the consistency is determined, the authentication device **144** may verify each signal from the user **108** (for example, whether each of the signals are authentic or, in other words, actually from the user **108**). Verification, as described, may utilize models trained using submitted or previously submitted enrollment or submissions from the user **108**. After verification, the authentication device **144** may grant the user **108** access to utilize functionality of one or more devices (for example, the first device **138**, the second device **146**, and/or additional devices), access to data, and/or physical or virtual access to a location.

[0048] FIG. 2 is a block diagram of a system to authenticate a user, according to an embodiment of the present disclosure. As illustrated a user **202** may wear a wearable device **206** (for example, a headset, headphones, earbuds, smart glasses, and/or VR or AR headset, among other devices). As the user **202** speaks, vibrations from the user's **202** vocal chords may travel through the user's **202** jaw or skull to the wearable device **206** (for example, bone conduction **204**). The wearable device **206** may sense the vibrations and convert the vibrations to a bone conduction signal **214**. The bone conduction signal **214** may be transmitted from the wearable device **206** to a smart device and/or authentication circuitry for user **202** authentication via a WiFi **222** or Bluetooth **224** interface. In addition, the as the user **202** speaks, a user's **202** vocal chords may generate sound or speech (for example, an audio signal **212** generated via air conduction **208**). The sound or speech may be recorded or received by a microphone **210** or other sensor of the smart device. The sound or speech, as an audio signal **212**, may be remain with the smart device (for example, where authentication occurs at or in the smart device) or be transmitted to an authentication circuitry (for example, where authentication occurs separate from the smart device) for authentication.

[0049] Before granting the user **202** access, the smart device and/or authentication circuitry may determine whether the bone conduction signal **214** is consistent with the audio signal **212** (see voice activity consistent **216**). In other words, the smart device and/or authentication circuitry may determine whether the bone conduction signal **214** is

from or originates from the same user **202** that provided the audio signal **212**. During the consistency check, the smart device and/or authentication circuitry may compare and/or analyze the bone conduction signal **214** and audio signal **212** to determine to determine consistency. If the bone conduction signal **214** and audio signal **212** are determined to be inconsistent, then the smart device may remain locked (see **220**).

[0050] If the bone conduction signal **214** and audio signal **212** are determined to be consistent, then the smart device and/or authentication circuitry may perform bone conduction signal **214** and audio signal **212** verification. In other words, the bone conduction signal **214** may be passed through a trained model or classifier and/or compared to previous submissions or enrollments to determine whether the bone conduction signal **214** is from or originates from a particular user **202**. Similarly, the audio signal **212** may be passed through a trained model or classifier and/or compared to previous submissions or enrollments to determine whether the audio signal **214** is from or originates from a particular user **202**. Such operations may occur simultaneously or in sequence. If at any point the smart device and/or authentication circuitry determines that either the bone conduction signal **214** or audio signal **212** are not from the user **202**, then the smart device may lock the user out or disable further access to the smart device by the user. If both the bone conduction signal **214** and audio signal **212** are verified, then the smart device may unlock or allow the user to access features of, functionality of, or data stored in the smart device.

[0051] FIG. 3A, FIG. 3B, and FIG. 3C are block diagrams of systems to authenticate a user, according to an embodiment of the present disclosure. Turning first to FIG. 3A, signals generated by a user may initially be transmitted to a denoising and filtering module. For example, a bone conduction signal may be received at the denoising and filtering module **302**, while an audio signal may be received at the denoising and filtering module **304**. The denoising and filtering module **302** may remove different frequencies than that of denoising and filtering module **304**, due to the nature of the signals captured. For example, denoising and filtering module **302** may remove signals at about 20 Hz and under. Such signals at about 20 Hz and under may represent noise generated by gravity and human motion. Further a frequency of interest for bone conduction signals may be about 1.5 kHz. The denoising and filtering module **302** may remove frequencies outside of a range defined by the frequency of interest. For example, a band-pass filter, included in the denoising and filtering module **302**, may remove frequencies under about 20 Hz and over about 2 kHz. The denoising and filtering module **304** may include a low-pass band filter.

[0052] After denoising, the bone conduction signal and audio signal may be passed or transmitted to a synchronization module **306**. The synchronization module **306** may utilize cross-correlation to align the bone conduction signal with the audio signal. The synchronization module **306** may delay the earlier of the bone conduction signal and audio signal so that the bone conduction signal and audio signal reach a maximum cross-correlation, such that the signals may be further processed and analyzed to determine if they are from the same user.

[0053] After synchronization, a consistency score may be determined. The score may be generated using a voice activity detection module **308** and a similarity check module

310. The voice activity module **308** may convert the bone conduction signal and the audio signal to binary waveforms. The binary waveforms may be compared to, at the least, ensure that voice and bone conduction activity is occurring at the same time. Such analysis ensures that an in depth consistency check is not performed if the bone conduction signal and audio signal do not at least include matching activity time frames.

[0054] The bone conduction signal and audio signal may then be analyzed and compared in the similarity check module **310**. The similarity check module **310** may normalize, chunk the signals into frames, and transform the signals into a time-frequency grid by a short-time Fourier Transform. The processed signals may then be compared and a score generated. Depending on the score, it may be determined whether the bone conduction signal and the audio signal are from the same user (for example, a high score indicating the bone conduction signal and the audio signal is consistent and a low score indicating bone conduction signal and the audio signal are inconsistent). Additionally, the consistency score may be compared to a threshold (for example, a consistency threshold) and if the score exceeds the threshold, then the bone conduction signal and the audio signal may be determined to be consistent.

[0055] In another embodiment, the bone conduction signal and the audio signal may be converted to a waveform and the waveform may be compared to determine whether certain features are similar. For example, the waveforms may simply be compared to determine each time of activity (for example, as illustrated in the air-bone waveform **311**). In another example, specific frequencies may be compared. In another example, or in addition to the waveform, the signals may be converted to a binary waveform (as noted above) which simply denotes activity using a square waveform (for example, as illustrated in voice activity **313**, where a 1 would indicate activity and 0 would indicate non-activity).

[0056] In an embodiment, if the bone conduction signal and the audio signal are determined to be consistent, then features may be extracted from each of the bone conduction signal and the audio signal. The bone conduction features may be extracted using a convolutional neural network (CNN) or other neural network or machine learning model. The CNN may leverage an image-classification method to extract image-like feature maps from using a time-frequency analysis. The audio features extracted may include various spectral characteristics of the audio signal. In an embodiment, extracting features may include converting the audio signal to an audio signal feature vector and converting the bone conduction signal to a bone conduction signal vector.

[0057] FIG. 3B illustrates two examples of a pass and a fail in regards to consistency between the audio signal and bone conduction signal. After preprocessing, the audio signal and bone conduction signal may align or may be substantially similar, in the case of an audio signal and bone conduction signal being from the same user (see **314**). In another example, after preprocessing some audio signals and bone conduction signals may not match (see **316**). In such an example, the audio signal and bone conduction signal may be so different as to indicate a potential attack or spoof, thus access to the smart device may be denied. Further, in some examples, no bone conduction signal may be detected, in which case access to the smart device may be denied.

[0058] Turning to FIG. 3C, after consistency is determined and the features extracted, the bone conduction features 320 may be transmitted to a trained model or classifier 332. The trained model or classifier may be trained using a user's 318 enrollment or submission bone conduction signals. The trained model or classifier may be based on a CNN or other neural network 334. The output of the CNN or classifier may be a bone conduction vector. The system may include a multi-classification module 336 or a multi-class image classifier. An output of the multi-classification module 336 may include a probability indicating whether the original bone conduction signal is from or originates from a particular user. In another embodiment, the bone conduction vector may be combined with the audio vector at 338. A score may then be generated at a probability linear discriminant analysis (PLDA) module 340 using the air-bone speaker embedding (for example, the bone conduction vector combined with the audio vector). In another embodiment, a linear discriminant analysis (LDA), a cosine similarity scorer (CSS), or another statistical or probabilistic model or classifier may be utilized, rather than the PLDA. Based on the score, the user may be verified 342 or filed 344.

[0059] As noted, the audio features may be extracted after the consistency check. After the features are extracted, the audio features or vectors may be applied to a model or classifier to further reduce the amount of vectors. The reduced audio vectors, as noted, may be added to the air-bone speaker embedding 338.

[0060] In another embodiment, prior to authentication of the user 318, the smart device and/or authentication circuitry may prompt the user to submit an audio signal and bone conduction signal. Such a signal may be submitted for one or more wearable devices of the user. The results may be stored in a database 326 for future use. Further the model or classifier used in relation to the bone conducting signal and the audio signal may be trained using the submission.

[0061] FIG. 4 is a flow diagram of two-way user authentication, according to an embodiment of the present disclosure. The method is detailed with reference to the smart device, wearable device, and system 100 of FIG. 1A through FIG. 1D. Unless otherwise specified, the actions of method 400 may be completed within the smart device, wearable device, and/or authentication circuitry. Specifically, method 400 may be included in one or more programs, protocols, or instructions loaded into the memory of the smart device or wearable device and executed on the processor or one or more processors of the smart device or wearable device. The order in which the operations are described is not intended to be construed as a limitation, and any number of the described blocks may be combined in any order and/or in parallel to implement the methods.

[0062] At block 402, the smart device may determine whether an audio signal has been received. The smart device may wait until such a signal is received. In an embodiment, the smart device may prompt the user to submit the response (for example, the audio signal and the bone conduction signal).

[0063] At block 404, the wearable device and/or smart device may determine whether the bone conduction signal has been received. In an embodiment, the smart device and/or authentication circuitry may wait for a specified period of time for the bone conduction signal. For example, the smart device and/or authentication circuitry may wait about 30 seconds, about 1 minute, or up to about 5 minutes.

Longer periods of time between reception of signals may indicate a potential attack or spoof via a third party. After the period of time has lapsed, the smart device and/or authentication circuitry may deny access for the user to utilize the smart device. In an embodiment, a portion of functionality of the smart device may be available for use by the user.

[0064] At block 406, the smart device and/or authentication circuitry may analyze the audio and bone conduction signals. Such analysis may include, at a high level, a consistency check and verification. The consistency check may include preprocessing the bone conduction signal and the audio signal. If the signals are consistent, then smart device and/or authentication circuitry may verify the audio signal and the bone conduction signal. The verification may include applying the audio signal or audio features or vectors and the bone conduction signal or bone conduction features or vectors to a trained model or classifier corresponding to the signal (for example, a model or classifier specific to either the bone conduction signal or the audio signal). The output of such a model may be transmitted to a PLDA or other statistical or probabilistic model to determine a score.

[0065] At block 408, the smart device and/or authentication circuitry may authenticate the user. The smart device and/or authentication circuitry may determine, based on scores generated from during analysis of the bone conduction signal and audio signal, that the user is authentic or is verified.

[0066] At block 410, if the user is authenticated, the smart device and/or authentication circuitry may grant access, at block 412, to the smart device. If the user is not authenticated, the smart device and/or authentication circuitry, at block 414, may deny access to the smart device.

[0067] FIG. 5A and FIG. 5B are flow diagrams of two-way user authentication, according to an embodiment of the present disclosure. The method is detailed with reference to the smart device, wearable device, and system 100 of FIG. 1A through FIG. 1D. Unless otherwise specified, the actions of method 400 may be completed within the smart device, wearable device, and/or authentication circuitry. Specifically, method 400 may be included in one or more programs, protocols, or instructions loaded into the memory of the smart device or wearable device and executed on the processor or one or more processors of the smart device or wearable device. The order in which the operations are described is not intended to be construed as a limitation, and any number of the described blocks may be combined in any order and/or in parallel to implement the methods.

[0068] At block 502, the smart device and/or authentication circuitry may determine whether a user is enrolled or has been initialized. In other words, the smart device and/or authentication circuitry may determine whether the user has submitted an initial phrase or one or more phrases.

[0069] At block 504, if the user has not been enrolled or initialized, then the smart device and/or the authentication circuitry may prompt a user to submit one or more phrases for initialization or enrollment. The one or more phrases may be random phrases and/or specified phrases (for example, specified by the smart device). In such examples, the smart device may issue or transmit a prompt as a text based message or as an audio prompt. In an embodiment, the smart device and/or authentication circuitry may issue or transmit the prompt for each of the user's wearable devices.

[0070] At block 506, the smart device and/or authentication circuitry may issue or transmit a prompt as to whether an additional wearable device is to be enrolled or initialized. If an additional wearable device is to be initialized or enrolled, then the smart device and/or authentication circuitry may prompt the user to wear the wearable device and submit one or more phrases.

[0071] At block 508, the smart device and/or authentication circuitry may update a trained bone conduction model with the submitted one or more phrases. The trained bone conduction model may be utilized to verify a user's bone conduction signal. Utilizing the user's submitted one or more phrases may ensure that an accurate verification may occur. At block 510, the smart device and/or authentication circuitry may update a trained audio model with the submitted one or more phrases. The trained bone conduction model may be utilized to verify a user's bone conduction signal. Utilizing the user's submitted one or more phrases ensure that an accurate verification may occur.

[0072] At block 512, the smart device and/or authentication circuitry may determine whether an audio signal has been received. The smart device and/or authentication circuitry may wait for such a signal prior to proceeding to the next operation. In another embodiment, the smart device and/or authentication circuitry may receive the bone conduction signal prior to the audio signal.

[0073] At block 514, the smart device and/or the authentication circuitry may determine whether the bone conduction signal has been received. In an embodiment, the smart device and/or authentication circuitry may wait for a specified or preselected period of time prior to denying access to the user. The smart device and/or the authentication circuitry may wait for about 30 seconds, about 1 minute, or up to about 5 minutes. Longer periods of time between reception of signals may indicate a potential attack or spoof via a third party. In an embodiment, the bone conduction signal is received by the wearable device at substantially the same time as the audio signal is received by the microphone or other sensor.

[0074] If the bone conduction signal is received, at block 516, the smart device and/or authentication circuitry may preprocess the audio signal. Preprocessing the audio signal may include passing the audio signal through a low-band pass filter, through a Wiener filter, and/or utilizing other noise reduction techniques, as will be understood by those skilled in the art.

[0075] At block 518, the smart device and/or the authentication circuitry may preprocess the bone conduction signal. Preprocessing the bone conduction signal may include passing the bone conduction signal through a band pass filter, through a Wiener filter, and/or utilizing other noise reduction techniques, as will be understood by those skilled in the art. The band pass filter may remove frequencies below about 20 Hz and frequencies above about 2 kHz.

[0076] At block 520, the smart device and/or the authentication circuitry may delay the earliest received signal to align the preprocessed bone conduction signal and the preprocessed audio signal. The time that either signal is received may not match the other. As such, to ensure an accurate consistency check, the smart device and/or the authentication circuitry may shift the earliest received signal in relation to time. Thus, at least from the perspective of

alignment according to or in relation to time, the preprocessed bone conduction signal and the preprocessed audio signal may align.

[0077] At block 522, the smart device and/or the authentication circuitry may determine a bone conduction power distribution with respect to time. Higher spectral power for each signal may generate higher correlations considering the amplitude. Therefore, selecting the frequencies with the highest or largest power ensures accurate consistency checks. Such a determination may be represented by $y_b(t) \leftarrow \sum_f y_b(f, t)$.

[0078] At block 524, the smart device and/or the authentication circuitry may select the time range of interest for the bone conduction power distribution. As indicated, the higher the power for a particular frequency the likelier correlation between a bone conduction signal and audio signal is, if the signals are from the same user. As such, the smart device and/or the authentication circuitry may select the time range with the highest power or spectral power. Such an operation may be represented by $t' \leftarrow \arg_t [y_b(t) \geq \theta]$.

[0079] At block 526, the smart device and/or the authentication circuitry may determine an audio power distribution and/or bone conduction power distribution with respect to frequency. Such operations may be represented by, for audio, $y_a(f) \leftarrow \sum_{t'} y_a(f, t')$ and, for bone conduction $y_b(f) \leftarrow \sum_{t'} y_b(f, t')$.

[0080] At block 528, the smart device and/or the authentication circuitry may select the top frequencies of the bone conduction signal and audio signal power distribution. Such an operation may be represented by $m \leftarrow \arg_l \text{Sort}(y_a(f))_{1:M}$, and $n \leftarrow \arg_l \text{Sort}(y_b(f))_{1:N}$. At block 530, the smart device and/or authentication circuitry may determine a correlation matrix, as represented by $C \leftarrow \text{Corr}^{M \times N} [y_a(f, t'), y_b(f, t')]$. Using the correlation matrix, at block 532, the smart device and/or the authentication circuitry may generate a consistency score, as represented by

$$S \leftarrow \max_{m \leq M, n \leq N} C(m, n).$$

[0081] At block 534, the smart device and/or the authentication circuitry may determine whether the consistency score indicates that the bone conduction signal and audio signal are consistent or from the same user. The smart device and/or the authentication circuitry may utilize a consistency threshold to determine whether the consistency score indicates that the bone conduction signal and audio signal are consistent or from the same user (for example, the consistency score is greater than or equal to the consistency threshold).

[0082] At block 536, the smart device and/or the authentication circuitry may generate an audio feature vector. Using a neural network, a max pooling layer, or a fully connected layer, at block 538, the smart device and/or the authentication circuitry may reduce the audio feature vectors. At block 540, the smart device and/or the authentication circuitry may verify the reduced audio feature vector with the enrolled user template and/or a classifier.

[0083] At block 542, the smart device and/or the authentication circuitry may generate a bone conduction feature vector. Using a neural network, a max pooling layer, or a fully connected layer, at block 544, the smart device and/or the authentication circuitry may reduce the bone conduction

feature vector. At block 546, the smart device and/or the authentication circuitry may verify the reduced bone conduction feature vector with the enrolled user template and/or a classifier.

[0084] At block 544, the smart device and/or the authentication circuitry may utilize the results of block 546 and block 540 to determine whether the bone conduction signal and the audio signals have been verified. If the signals have been verified, at block 550, the smart device and/or the authentication circuitry may enable access for the user.

[0085] If the signals have not been verified, at block 552, the smart device and/or the authentication circuitry may prompt the user to resubmit an audio and bone conduction signal for resubmission (for example, submit an additional attempt). The smart device and/or the authentication circuitry may allow for a specified number of resubmissions or at least one resubmission. Once that number has been met, the smart device and/or the authentication circuitry may deny access to the user at block 554.

[0086] FIG. 4 and FIG. 5A through FIG. 5C illustrate flowcharts describing sets of operations performed by apparatuses, methods, and computer program products according to various example embodiments. It will be understood that each block of the flowcharts, and combinations of blocks in the flowcharts, may be implemented by various means, embodied as hardware, firmware, circuitry, and/or other devices associated with execution of software including one or more software instructions. For example, one or more of the operations described above may be embodied by software instructions. In this regard, the software instructions which embody the procedures described above may be stored by a memory of an apparatus employing an embodiment and executed by a processor of that apparatus. As will be appreciated, any such software instructions may be loaded onto a computing device or other programmable apparatus (for example, hardware) to produce a machine, such that the resulting computing device or other programmable apparatus implements the functions specified in the flowchart blocks. These software instructions may also be stored in a computer-readable memory that may direct a computing device or other programmable apparatus to function in a particular manner, such that the software instructions stored in the computer-readable memory produce an article of manufacture, the execution of which implements the functions specified in the flowchart blocks. The software instructions may also be loaded onto a computing device or other programmable apparatus to cause a series of operations to be performed on the computing device or other programmable apparatus to produce a computer-implemented process such that the software instructions executed on the computing device or other programmable apparatus provide operations for implementing the functions specified in the flowchart blocks.

[0087] The flowchart blocks support combinations of means for performing the specified functions and combinations of operations for performing the specified functions. It will be understood that one or more blocks of the flowcharts, and combinations of blocks in the flowcharts, can be implemented by special purpose hardware-based computing devices which perform the specified functions, or combinations of special purpose hardware and software instructions.

[0088] In some embodiments, some of the operations above may be modified or further amplified. Furthermore, in some embodiments, additional optional operations may be

included. Modifications, amplifications, or additions to the operations above may be performed in any order and in any combination.

[0089] This application is related to U.S. Provisional Application No. 63/268,999, filed Mar. 8, 2022, titled “SYSTEMS AND APPARATUS FOR MULTIFACTOR AUTHENTICATION USING BONE CONDUCTION AND AUDIO SIGNALS,” U.S. Provisional Application No. 63/269,001, filed Mar. 8, 2022, titled “METHOD FOR MULTIFACTOR AUTHENTICATION USING BONE CONDUCTION AND AUDIO SIGNALS,” and U.S. Provisional Application No. 63/380,229, filed Oct. 19, 2022, titled “SYSTEMS AND METHODS FOR CONTINUOUS, ACTIVE, AND NON-INTRUSIVE USER AUTHENTICATION,” the disclosures of which are incorporated herein by reference in their entirety.

[0090] In the drawings and specification, several embodiments of systems and methods to provide two-way authentication for a user via a smart device or device and a wearable device have been disclosed, and although specific terms are employed, the terms are used in a descriptive sense only and not for purposes of limitation. Embodiments of systems and methods have been described in considerable detail with specific reference to the illustrated embodiments. However, it will be apparent that various modifications and changes can be made within the spirit and scope of the embodiments of systems and methods as described in the foregoing specification, and such modifications and changes are to be considered equivalents and part of this disclosure.

What is claimed is:

1. A method for two-way authentication of a user, the method comprising:
 - receiving a bone conduction signal from the user via one or more wearable devices;
 - receiving an audio signal from the user via a microphone separate from the one or more wearable devices, the audio signal to correspond with the bone conduction signal;
 - determining a consistency score for the audio signal and corresponding bone conduction signal;
 - in response to the consistency score being greater than or equal to a consistency threshold:
 - verifying, using an audio conduction model (AC model), that the audio signal is associated with the user;
 - verifying, using a bone conduction model (BC model), that the bone conduction signal is associated with the user; and
 - in response to verification of the audio signal and the bone conduction signal, enabling, for the user, access to a smart device.
2. The method of claim 1, wherein determination of the consistency score includes:
 - determining a marginal bone conduction power distribution with respect to time;
 - selecting a time range of interest based on an average of the margin marginal bone conduction power distribution with respect to time;
 - determining a marginal audio conduction and bone conduction power distribution with respect to frequency;
 - selecting a top frequency index M from the marginal bone conduction power distribution with respect to fre-

quency and a top frequency index N from the marginal audio conduction power distribution with respect to frequency;

determining a correlation matrix between the top frequency index M and the top frequency index N; and generating, based on the correlation matrix, the consistency score.

3. The method of claim 1, further comprising, prior to determining the consistency score, pre-processing the bone conduction signal.

4. The method of claim 3, wherein pre-processing the bone conduction signal includes passing the bone conduction signal through (1) a low-pass filter to remove noise generated by human motion and (2) a Wiener filter to remove the noise.

5. The method of claim 1, further comprising, prior to determining the consistency score, passing the audio signal through a Wiener filter to remove noise.

6. The method of claim 1, further comprising, prior to determination of a consistency score and verification, prompting the user to submit an initial or enrollment (1) bone conduction signal and (2) audio signal.

7. The method of claim 6, further comprising:
 training the AC model with a plurality of responses prior to submission of an initial or enrollment audio signal, and
 training the AC model with a submitted initial or enrollment audio signal.

8. The method of claim 6, wherein the BC model comprises a convolutional neural network (CNN).

9. The method of claim 8, further comprising training the CNN with a stored/pre-collected bone conduction dataset.

10. The method of claim 1, further comprising, if one or more of the audio signal or bone conduction signal are not verified, preventing the user from accessing the smart device.

11. The method of claim 1, further comprising, prior to determining the consistency score, delaying an earliest received of the bone conduction signal and the audio signal to align the bone conduction signal and the audio signal.

12. A method for two-way authentication of a user, the method comprising:
 prompting the user to submit initial or enrollment (a) audio signals and (b) bone conduction signals;
 updating an audio conduction model (AC model) and a bone conduction model (BC model) based on the received initial or enrollment (a) audio signals and (b) bone conduction signals;
 after reception of the initial or enrollment (a) audio signals and (b) bone conduction signals:
 receiving a bone conduction signal from the user via one or more wearable devices;
 receiving an audio signal from the user via a microphone separate from the one or more wearable devices, the audio signal corresponding to the bone conduction signal;
 determining a consistency score for the audio signal and corresponding bone conduction signal;
 in response to the consistency score being greater than or equal to a consistency threshold:
 verifying, using the AC model, that the audio signal is associated with the user;

verifying, using the BC model, that the bone conduction signal is associated with the user; and
 in response to verification of the audio signal and bone conduction signal, enabling, for the user, access to a smart device.

13. The method of claim 12, wherein the initial or enrollment audio signal is received via the microphone, and wherein the smart device includes the microphone.

14. The method of claim 13, wherein the initial or enrollment bone conduction signals are received from one of one or more wearable devices.

15. The method of claim 12, further comprising prompting the user to submit initial or enrollment (a) audio signals and (b) bone conduction signals for each of the one or more wearable devices.

16. The method of claim 15, wherein the initial or enrollment (a) audio signals and (b) bone conduction signals include one or more specific phrases.

17. The method of claim 12, wherein the BC model comprises a convolutional neural network (CNN), and wherein the CNN is trained using a stored/pre-collected bone conduction dataset to generate corresponding initial embedded bone conduction vectors.

18. The method of claim 17, further comprising, prior to verification via the BC model, generating embedded bone conduction vectors using the bone conduction signals and the initial embedded bone conduction vectors.

19. A non-transitory machine-readable storage medium storing processor-executable instructions that, when executed by at least one processor, cause the at least one processor to:

receive a bone conduction signal from a user via one or more wearable devices;
 receive an audio signal from the user via a microphone separate from the one or more wearable devices, the audio signal corresponding to the bone conduction signal;
 determine a consistency score for the audio signal and the bone conduction signal;
 in response to the consistency score being greater than or equal to a consistency threshold:
 verify, using an audio conduction model (AC model), that the audio signal is associated with the user;
 verify, using a bone conduction model (BC model), that the bone conduction signal is associated with the user; and
 in response to verification of the audio signal and bone conduction signal, enable, for the user, access to a smart device.

20. The non-transitory machine-readable storage medium of claim 19, wherein the smart device includes the microphone.

21. The non-transitory machine-readable storage medium of claim 20, wherein the bone conduction signal is received via wireless communication.

22. The non-transitory machine-readable storage medium of claim 19, wherein the consistency threshold is based on similarities between bone conduction signals and audio signals that indicate the bone conduction signal and the audio signal are from the user.