(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 January 2025 (23.01.2025)

WIPO | PCT

(10) International Publication Number
WO 2025/019404 A2

(51) International Patent Classification:
G16H 50/20 (2018.01) *A61B 5/256* (2021.01)

(21) International Application Number:
PCT/US2024/037996

(22) International Filing Date:
15 July 2024 (15.07.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/527,258 17 July 2023 (17.07.2023) US

(71) Applicant: **UNIVERSITY OF PITTSBURGH - OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION** [US/US]; University of Pittsburgh Innovation Institute, 1st Floor Gardner Steel Conference Center (GSCC), 130 Thackeray Avenue, Pittsburgh, Pennsylvania 15260 (US).

(72) Inventors: **CHEN, Tao**; c/o UNIVERSITY OF PITTSBURGH - OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION, University of Pittsburgh Innovation Institute, 1st Floor Gardner Steel Conference Center (GSCC), 130 Thackeray Avenue, Pittsburgh, Pennsylvania 15260 (US). **SHANGGUAN, Longfel**; c/o UNIVERSITY OF PITTSBURGH - OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION, University of Pittsburgh Innovation Institute, 1st Floor Gardner Steel Conference Center (GSCC), 130 Thackeray Avenue, Pittsburgh, Pennsylvania 15260 (US). **YANG, Yongjie**; c/o UNIVERSITY OF PITTSBURGH - OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION, University of Pittsburgh Innovation Institute, 1st Floor Gardner Steel Conference Center (GSCC), 130 Thackeray Avenue, Pittsburgh, Pennsylvania 15260 (US).

(74) Agent: **SOPHIR, Eric** et al.; FOLEY & LARDNER LLP, 3000 K Street NW, Suite 600, Washington, District of Columbia 20007-5109 (US).

(54) **Title:** CARDIAC AUSCULTATION USING EARPHONES

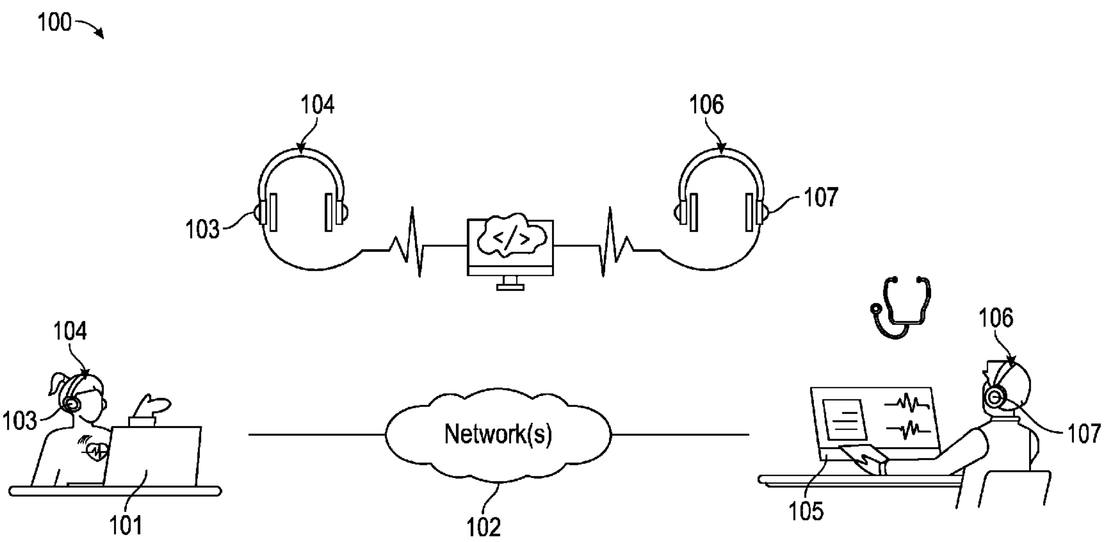

FIG. 1

(57) **Abstract:** Embodiments described herein provide for a hardware-and-software solution that effectively repurposes a patient's earphone into microphone or stethoscope, allowing a specialist to remotely perform auscultation during a remote visit. The heartbeat signal is captured and augmented from a voltage signal produced by an audio-output earphone containing phonocardiogram (PCG) signals. Hardware and software components accentuate and correct the PCG signal of the electrical signal. A patient's client device sends a clear and correct version of the PCG signal to a provider's client device to perform cardiac auscultation.

WO 2025/019404 A2

[Continued on next page]

(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, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:
- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

CARDIAC AUSCULTATION USING EARPHONES

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application No. 63/527,258, filed July 17, 2023, which is incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] This application generally relates to systems and methods for audio signal processing and, in particular, capturing and processing voltage signals from audio-output speakers of an earphone to capture and process cardiac signals of a human wearing the earphone into a format that can be reviewed by a care provider to perform cardiac auscultation.

BACKGROUND

[0003] Improvements to remote communications software programs, online meeting platforms, and related protocols and standards have facilitated technical improvements and increased popularity in purpose-driven products focusing on healthcare, including virtual care programs and online platforms (sometimes referred to herein as "remote care software" or "virtual care software"). The popularity of the remote care software increased drastically during the COVID pandemic. In response to the pandemic, the federal and state governments, healthcare providers and experts, and commercial insurers encouraged patients and healthcare providers to meet virtually (e.g., conduct remote video visits), allowing patients to access care providers from home, over the Internet, without the risk of COVID exposure.

[0004] A problem with remote care software is that the care provider relies primarily, or exclusively, on information volunteered by the patient or video observations. For instance, a specialist might ordinarily perform cardiac auscultations using a stethoscope to assess the patient's health. But in the context of a virtual meeting, the specialist cannot perform cardiac auscultation, which is a core metric for monitoring and evaluating a patient's health.

[0005] Cardiac auscultation is a crucial physical examination among different kinds of auscultations. Cardiac auscultation is performed to examine the circulatory system by listening to the heart sound, or Phonocardiogram (PCG) signals, emanating from the human heart. Recently, some pharmaceutical or medical device providers rolled out products for performing remote cardiac auscultation, such as at-home digital stethoscope that allows patients to measure the PCG signals at home and synchronize the patient's data with the specialists through Wi-Fi or Bluetooth connections. However, these products are often complex, expensive devices and are difficult to operate for unsophisticated users. Moreover, even with access to these new types of digital stethoscope products, patients typically lack professional training for properly placing or operating these new digital stethoscopes to accurately and consistently collect the heart sound.

[0006] As virtual meeting software grows more sophisticated and the advantages gain popularity, it is unlikely that the healthcare industry will abandon remote care software and digital solutions. As such, what is needed is a means for care specialists to perform auscultations remotely, over the Internet, using the virtual care software.

SUMMARY

[0007] Disclosed herein are systems and methods capable of addressing the above-described shortcomings and may provide any number of additional or alternative benefits and advantages. Embodiments described herein provide for a hardware and software solution that effectively repurposes a patient's earphones into a stethoscope, allowing a specialist to hear or watch a heartbeat signal of the patient during a remote visit using virtual care software, where the heartbeat signal is captured and augment from a voltage signal produced by an audio-output earphone of the patient.

[0008] The embodiments implementing the proposed solutions provide various advantages over other approaches, such as higher accuracy, easy to operate, and relatively inexpensive. Embodiments may accurately and consistently detect coarse-grained heartrate variation (HRV) and/or fine-grained cardiac features (e.g., S1 and S2 sound) that are essential to cardiac auscultation. The embodiments allow the care specialists to take remote cardiac auscultation with minimum human intervention. The embodiments need not require particularly expensive hardware components for the patients to purchase.

[0009] In an embodiment, a computer-implemented method for performing cardiac auscultation comprising obtaining, by a computer, a raw audio signal containing a heartbeat signal produced from an earphone, the earphone configured to output an analog audio waveform for an ear canal of an end-user; determining, by the computer, a cardiac cycle duration for the end-user;

segmenting, by the computer, the heartbeat signal into a plurality of heartbeat cycle segments according to the cardiac cycle duration of the end-user; generating, by the computer, a spectrogram of the heartbeat signal in a frequency-domain by applying a neural network architecture on the plurality of heartbeat cycle segments; applying, by the computer, a transform function on the spectrogram to transform the spectrogram to a time-domain, thereby generating a refined heartbeat signal; and transmitting, by the computer, the refined heartbeat signal in a machine-readable format to a provider computer.

[0010] In another embodiment, a computer-implemented method for processing signals from audio earphones for cardiac auscultation, the method comprising: instructing, by a computer, a signal-processing device to transmit a plurality of probing signals to a headset of a user according to a plurality of candidate impedance settings of the signal-processing device; for each candidate impedance setting of the plurality of candidate impedance settings: obtaining, by the computer, a reference signal received at the signal-processing device from an audio-output earphone of the headset, wherein the reference signal corresponds to a probing signal sent to the headset according to the candidate impedance setting; determining, by the computer, an amount of power of the reference signal occurring at the candidate impedance setting; and computing, by the computer, a signal-to-noise ratio using the amount of power of the reference signal occurring at the candidate impedance setting; identifying, by the computer, the candidate impedance setting having a comparative maximum signal-to-noise ratio as an optimal impedance match between a first impedance of the signal-processing device and a second impedance of the computer; and adjusting, by the computer, impedance of the signal-processing device according to the optimal impedance match.

[0011] In another embodiment, a device for processing signals to perform cardiac auscultation via earphones, where the device comprises a detection circuit and an impedance-matching circuit. The detection circuit includes an interface configured to obtain a raw voltage signal containing a heartbeat signal produced from an audio-output earphone of a headset; and an amplifier configured to amplify the heartbeat signal in the raw voltage signal, thereby generating an amplified voltage signal containing the heartbeat signal. The impedance-matching circuit includes a programmable controller configured to adjust a first impedance of the impedance-matching circuit to match a second impedance of a pairing device according to a tuning instruction received from the pairing device indicating an optimal impedance value for the first impedance.

[0012] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory and are intended to provide further explanation of the invention as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The present disclosure can be better understood by referring to the following figures. The components in the figures are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the disclosure. In the figures, reference numerals designate corresponding parts throughout the different views.

[0014] **FIG. 1** shows components of a system that provides remote cardiac auscultation, according to an embodiment.

[0015] **FIGS. 2A-2B** show arrangements of patient-side components, according to one or more embodiments, where **FIG. 2B** shows a comparatively closer-view of the patient-side components compared to the comparatively further view shown by **FIG. 2A**.

[0016] **FIG. 3** shows operational workflow amongst hardware and software components of a system providing remote auscultation, according to an embodiment.

[0017] **FIG. 4** is a schematic of a low-power voltage detection circuit of a signal processor coupled to other patient-side components, according to an embodiment.

[0018] **FIG. 5** is a schematic of an impedance circuit of a signal processor coupled to other patient-side components, according to an embodiment.

[0019] **FIG. 6** shows a schematic of devices for processing PCG signals before applying a software pipeline of a pairing device, according to an embodiment.

[0020] **FIG. 7** shows operational steps of method for processing signals to perform cardiac auscultation remotely, according to an embodiment.

[0021] **FIG. 8** shows operational component layers and data types of a machine-learning architecture executed for signal recovery and refinement, according to an embodiment.

DETAILED DESCRIPTION

[0022] Reference will now be made to the illustrative embodiments illustrated in the drawings, and specific language will be used here to describe the same. It will nevertheless be understood that no limitation of the scope of the invention is thereby intended. Alterations and further modifications of the inventive features illustrated here, and additional applications of the principles of the inventions as illustrated here, which would occur to a person skilled in the relevant art and having possession of this disclosure, are to be considered within the scope of the invention.

[0023] Embodiments described herein provide for a hardware and software-based solution that effectively repurposes a patient's earphones into microphone or stethoscope, allowing a specialist to hear or watch a heartbeat signal of the patient during a remote visit using virtual care software, where the heartbeat signal is captured and augment from a voltage signal produced by an audio-output earphone of the patient. Described herein are systems and methods gathering electrical signals containing heartbeat signals (sometimes referred to as "phonocardiogram (PCG) signals" or "cardiac signals") from an audio-output earphone worn by a patient (or type of end-user) and accentuating and correcting the heartbeat signal aspects of the electrical signal. A client device of the patient then sends a clear and corrected version or representation of the heartbeat signal to a client device of a care provider (or other type of end-user) who intends to perform cardiac auscultation or otherwise review the patient's cardiac health. The embodiments described herein effectively convert or repurpose the audio-output earphone of the patient into a microphone or stethoscope for the provider, even when the patient and provider meet virtually via a remote communications software program, executed or accessed by the client devices.

[0024] A hardware component includes an audio-processing device, which is plug-in peripheral to the patient device that effectively repurposes an audio speaker or transducer of an earphone into a microphone and uses the earphone to capture the patient's minute PCG signals from via the patient's ear canal.

[0025] Typically, the PCG signals suffer from strong attenuation and multi-path effects when propagating from the patient's heart to the patient's ear canals. A software component described herein includes software programming for executing various signal-processing operations and algorithms. The software component further includes a data-driven, machine-learning model pipeline, which may, for example, de-reverberate and further correct the amplitude and frequency distortion in raw PCG receptions. These software component functions enable the patient client device (sometimes referred to as a "pairing device" or "patient device") to recover or augment the PCG signals and prepare the PCG signals to send to the care provider's device over the Internet or other communications network.

[0026] Capturing PCG signals via an earphone presents a few technical and/or anatomical challenges, addressed by various embodiments described herein. As an example, a stethoscope captures near-ideal heart sounds (e.g., PCG signals), because the stethoscope's diaphragm is applied directly against the patient's body, just above the heart. By contrast, the PCG signals captured using an earphone (as a repurposed microphone) often suffer significant attenuation and frequency distortion as the signals propagate from the heart through the human bones, muscles, fat, and skins before arriving at the human ear canal and the earphone. Accordingly, these PCG signals tend to be very weak and thus are likely to be buried by ambient noises and human organ artifacts.

[0027] As another example, although using the audio speaker or transducer of the earphone as a microphone is feasible by leveraging the structural reciprocity of speakers and microphones, capturing the PCG signals with the earphone's audio speaker is still challenging because an audio-output speaker earphone is configured and optimally designed for signal output, rather than signal absorption. When a weak PCG signal from the patient's heart eventually arrives at the speaker diaphragm, for instance, the speaker diaphragm will transform only a small or minute portion of the PCG signal into a weak voltage signal. This weak voltage signal is unlikely to maintain sufficient structure to convey fine-grained PCG features such as S1 (the "lub" sound) and S2 (the "dub" sound) of the heart sound components. As such, simply capturing the voltage signal of an audio signal will be insufficient for cardiac auscultation.

[0028] As another example, an acoustic signal will get diffracted, reflected, and absorbed when propagating from amongst various audio cables and devices along an audio chain, to the pairing device. Generally, a proportion of a signal being absorbed by the pairing device is affected by a mismatch between a first impedance of the earphone and a second impedance of the pairing device. But because both the impedance of the earphone and the pairing device's impedance are unknown in circumstances described herein, the conventional, offline (without computing functionality) approaches to impedance-matching are generally insufficient. Moreover, conventional approaches to impedance matching also change dramatically with respect to the hardware type, form factor, and material of the circuits, which would inhibit broad interoperability amongst possible pairing devices, earphones, and circuit components of an audio-processing circuit.

[0029] Embodiments described herein include a hardware-plugin audio-processing device that couples to the earphone and the pairing device (e.g., patient client device) to improve the quality of the PCG signals received at the pairing device. The audio processing device effectively repurposes a speaker or transducer of an earphone in an agile microphone by capturing minute PCG signals at the ear canal in the voltage signals generated by the repurposed microphone. The audio-processing device comprising a printed circuit board (PCB) containing various hardware circuit devices that define a voltage detection circuit and an impedance-matching circuit. The voltage detection circuit may, for example, amplify the PCG signal and denoises the PCG signal by applying one or more filters that remove strong noises in an analog domain with a low-power analog circuit. The programmable impedance-matching circuit performs various operations to ensure that the PCG signals can be delivered to the pairing device with minimum signal reflections. To cope with the various dynamics that could affect impedance-matching, the impedance-matching circuit that performs an online (e.g., coupled to the pairing device or other client device) automatic impedance matching operation in conjunction with the pairing device. The pairing device and the impedance-matching circuit perform various operations for a feedback-loop-based control algorithm to automatically and dynamically balance an amount of impedance between the earphone and the pairing device, without any human intervention.

[0030] Embodiments may include a software component executed by the pairing device (or other computing device) that address certain challenges mentioned above and may provide additional benefits. Upon receiving the PCG signals from the audio-processing device, the software component includes programming that executes a signal-processing pipeline of functional engines running on the pairing device. The functional engines of the pipeline may, for example, de-reverberate the PCG signals received from the pairing device, segment the PCG signals into heartbeat cycles tailored to the patient. The pairing device may then apply various layers of a machine-learning architecture or other operations to, for example, correct frequency and/or phase distortion caused by the multi-path effect when the PCG signal propagated inside the patient's body. The pairing device then transmits the corrected PCG signal to the provider device in one or more machine-readable formats, such that care provider may review audio, video, and/or audiovisual format representing the PCG signal (including a heartbeat signal or refined heartbeat signal) and perform cardiac auscultation. In this way, the machine-readable format of the refined heartbeat signal includes at least one of an audio format, a visual format, or an audiovisual format.

[0031] **FIG. 1** shows components of a system **100** that provides remote cardiac auscultation, according to an embodiment. The system **100** includes a patient device **101** (sometimes referred to as a "pairing device **101**") coupled to one or more patient earphones **103**, a provider device **105** coupled to provider earphones **107** of a provider headset **106**, and one or more communications networks **102** through which the patient device **101** and the provider device **105** communicate.

[0032] The end-user devices **101**, **105** may include patient devices **101** (sometimes referred to as pairing devices **101**) and care provider devices **105**. The end-user devices **101**, **105** include any computing device comprising hardware and software components, capable of remote communications with other end-user devices **101**, **105**. The end-user devices **101**, **105** include an audio interface that connects the audio headsets **104**, **106** to the end-user devices **101**, **105**, such that a patient headset **104** connects to the patient device **101** and a provider headset **106** connects to a provider device **105**. The headsets **104**, **106** include one or more earphones **103** (e.g., left earphone, right earphone).

[0033] The end-user devices **101**, **105** execute software programming of virtual care software, allowing the patient and provider to conduct a remote health visit. The virtual care software establishes a communications session and communicate multimedia data according to any number of computer or device networking and media communications protocols. The end-user devices **101**, **105** may execute locally installed virtual care software or may access a host server that executes webserver software that hosts a website to conduct the remote care session. The end-user devices **101**, **105** may communicate the media data via one or more networks **102** to transmit and receive computer files and/or data streams containing media data, text data, or other types of data.

[0034] The patient earphone **103** may be connected to the patient device **101** via a hardware-based signal processing device (sometimes referred to as a “signal processor”), described further below (e.g., signal processing device **210** of **FIGS. 2A-2B**). The software of the patient device **101** and/or the components of the signal processor perform various functions for gathering and processing voltage signals containing PCG signals from an audio-output speaker or transducer of the patient earphones **103**. Notably, the functions for gathering and processing the PCG signals need not rely upon any microphone or audio-input device of the patient headset **104** or the patient device **101**. The patient device **101** may execute various software programming modules or engines for signal processing, which may define an audio-processing pipeline of functions. In some cases, the software modules executed by the patient device define one or more functional layers of a machine-learning architecture.

[0035] The headsets **104**, **106** and earphones **103**, **107** include any type of earphones product capable of outputting audio signals to the eardrums and ear canal of the end-user. The patient earphones **103** include a percussion diaphragm that agitates or vibrates to generate analog percussion waves to create the audible sound at the patient’s eardrums. The patient earphones **103** may have any number of form-factors, such as over-ear headphones, in-ear buds, or as a component of the patient headset **104** having a microphone (not shown), among others. In some cases, the patient earphones **103** are configured strictly for output; and, in some cases, the patient earphones **103** include or a situated proximate to a microphone allowing for the headphones to gather spoken audio signals from the patient. Typically, however, a microphone of a headset **104** cannot provide for cardiac auscultation, because the hardware of the microphone or headset **104** cannot capture the audio signals representing the PCG signals and/or because the software associated with the headset **104** filters or dilutes the audio signals containing the user’s pulse.

[0036] As mentioned, embodiments disclosed herein use the audio-output speaker portion of a patient earphone **103** to capture (as input) the audio signals for cardiac auscultation, and need not rely upon a microphone. Whenever the patient’s heart beats (or the patient causes other types of background noise), the diaphragm of the audio speaker of the patient earphones **103** will be agitated or vibrated, causing the patient earphone **103** to produce an electrical signal representing the heartbeat as a PCG signal and/or background noise. The patient earphones **103** generates and sends these electrical signals, via a copper wire or similar conductive wire, to the audio processing device and/or to the patient device **101**. The patient's body naturally produces the PCG signals that propagate from the patient's heart, through the body, to the ear canal. A speaker-diaphragm or transducer of the patient earphone **103** physically responds to these PCG signals (e.g., vibrates) and generates a weak voltage signal that is sent to the audio processing device and pairing device **101** through one or more wires.

[0037] The audio-processing device (sometimes referred to as an "signal processor") includes various components that process the electrical signals from the earphone **103**, including the electrical signals caused by the PCG signals. The functions of the signal processor effectively repurpose the speaker-transducer of the patient earphone **103** into an improvised microphone that captures and delivers the PCG signals in a useable form to the client device.

[0038] **FIGS. 2A-2B** show arrangements of patient-side components, according to one or more embodiments, where **FIG. 2B** shows a comparatively closer-view of the patient-side components compared to the comparatively further view shown by **FIG. 2A**. The patient-side components include a patient headset **204**, patient earphones **203a-203b** (generally referred to as patient earphones **203**), wires **211a-211b** (generally referred to as wires **211**), a signal-processing device (sometimes referred to as a signal processor **210**), and a client-side pairing device **201** (sometimes referred to as a client device **201**). The patient earphones **203** may be in the form of over-ear headphones **203a** (as shown in **FIG. 2A**) or in-ear buds **203b** (as shown in **FIG. 2B**).

[0039] The patient earphones **203** are coupled to the client device **201** via the signal processor **210**. A first wire **211a** couples the patient earphones **203** to the signal processor **210**, where the first wire **211a** connects at an interface of the patient earphones **203** and connects at a first interface of the signal processor **210**. A second wire **211b** couples the signal processor **210** to the client device **201**, where the second wire **211b** connects at a second interface of the signal processor **210** and connects at an interface of the client device **201**. With reference to **FIG. 2B**, the signal processor **210** includes a printed circuit board (PCB) containing various hardware components of a hardware-based signal processor. For ease of understanding, **FIG. 2B** shows an exposed PCB of the signal processor **210**, though embodiments of the signal processor **210** include a housing for the PCB that provides improved protection and aesthetic.

[0040] A patient's body naturally produces PCG signals that propagate to the patient's ear canal. A speaker-diaphragm and transducer of the patient earphones **203** physically respond to these PCG signals, generating a weak voltage signal that is typically offloaded to the pairing device **201** through the wires **211**.

[0041] The signal processor **210** includes various components that process the electrical signals from the earphones **203**, including the electrical signals caused by the PCG signals. The functions of the signal processor **210** effectively repurpose the speaker-transducer of the patient earphones **203** into a makeshift or improvised microphone that captures and delivers the PCG signals in a useable form to the client device **201** (e.g., a desktop or a tablet that the patient uses to talk to a specialist operating a specialist device).

[0042] The signal processor **210** includes hardware as a low-power plug-in peripheral that wires the patient earphone **203** to the patient's client device **201**. In some cases, the signal processor **210** includes 3.5 mm audio jacks as one or more interfaces. Additionally or alternatively, the signal processor **210** includes wireless interfaces (e.g., Wi-FI, Bluetooth) as the one or more interfaces. Correspondingly, the patient earphones **203** and the patient device **201** may include 3.5 mm audio jacks or wireless interfaces as the interfaces corresponding to the signal processor **210**. The signal processor **210** receives the electrical signals having the PCG signals (sometimes referred to as "cardiac signals") as input and feeds the electrical signals to an amplifier component that amplifies the features of the inputted electrical signals that define the cardiac signals. The signal processor **210** includes a programmable impedance circuit that adjusts or tunes an impedance of the earphone **203** to optimally deliver as much of the cardiac signals to the client device **201** as practicable. The programmable impedance circuit adjusts the impedance of the earphone **203** to match the impedance of the client device **201**.

[0043] The client device **201** includes software programming defining a data-driven, signal-processing pipeline, where the client device **201** execute a series of software modules for various signal-processing functions. The software modules executed on the pairing device **201** may, for example, recover fine-grained cardiac features of the cardiac signals within the raw electrical signals captured from the earphone **203**. The client device **201** then transmits data packets via one or more networks containing cardiac signal data to a provider's computing device, which may present the cardiac signal data as one or more user interface functions performed by a software application by the provider device.

[0044] **FIG. 3** shows operational workflow amongst hardware and software components of a system **300** providing remote auscultation, according to an embodiment. The system **300** includes a patient headset **304** having a left earphone **303a** and a right earphone **303b** (generally referred to as an earphone **303** or earphones **303**), an audio signal processor (signal processor **310**) as a hardware component, and a client pairing device **301** executing software components. As the patient's body naturally produces PCG signals that propagate to the patient's ear canal, a speaker diaphragm or transducer of the patient earphone **303** physically responds to the PCG signals and generates a weak voltage electrical signal. The electrical signal propagated through connecting wires to the client device **301** via the signal processor **310**.

[0045] The signal processor **310** is a hardware component that functionally converts or repurposes the speaker transducer of the patient earphones **303** into a makeshift microphone that captures and delivers the PCG signals to the client pairing device **301** (e.g., a desktop or a tablet that the patient uses to talk to a care specialist user operating a specialist device (not shown)), which executes software components for processing inputted signal data. As depicted in **FIG. 3**, the signal processor **310** includes a voltage detection circuit **312** and an impedance matching circuit **314**. The signal processor **310** is formed as a low-power plug-in peripheral device that wires the patient earphone **303** to the client pairing device **301**. The signal processor **310** takes the cardiac signals as input, amplifies the cardiac signal using a pre-amp device, and tunes the impedance of the earphone **303** with the programmable impedance circuit **314** for optimal delivery of as much of the cardiac signals as practicable to the client pairing device **301**. In some implementations, the signal processor **310** is configured to capture and process the voltage signals from the left earphone **303a** because the left earphone **303a** is closer to the patient's heart. As will be discussed, in some implementations, an impedance-matching circuit **314** transmits a probing signal or chirp to the right earphone **303b** of the headset **304**. The probing signal propagates across the patient's head to the left earphone **303a**. The left earphone **303a** responds to the probing signal, causing the left earphone **303a** to generate and return a reference signal for the impedance-matching circuit **314** to adjust impedance settings. As an example, in some implementations the processor **310** may receive, from the audio-output earphone (e.g., left earphone **303a**), the reference signal that is responsive to the probing signal, and transmits the reference signal corresponding to the probing signal to the pairing device **301**.

[0046] The client pairing device **301** includes software programming defining a data-driven, signal-processing pipeline, where the client pairing device **301** execute a series of software modules for various signal-processing functions. The software modules of the data-driven signal-processing pipeline, running on the client pairing device **301**, recover fine-grained cardiac features of the cardiac signals within the raw electrical signals captured from the earphones **303**. The signal processing pipeline contains four operational aspects, as performed by the software modules of the client pairing device **301**, include preprocessing operations **316**, segmentation operations **318**, spectrogram-recovery operations **320**, and waveform refinement operations **322**, among others.

[0047] **FIG. 4** is a schematic of a low-power voltage detection circuit **412** of a signal processor **410** coupled to other patient-side components, according to an embodiment. The patient-side components include a headset **404** comprising one or more patient earphones **403** connected to the signal processor **410**.

[0048] The voltage detection circuit **412** includes hardware devices that detect voltage signals obtained at a patient's ear canal and containing PCG signals produced from the patient's heart. The detection circuit **412** detects the voltage signal from an audio-output speaker or transducer of an earphone **403** of the patient headset **404**. In some implementations, the detection circuit **412** is coupled to and evaluates voltage signals from the speaker or transducer of the lefthand side of the patient earphone **403** because the human heart is relatively closer to the left ear. In some cases, the amplifier **414** connects to the left-ear speaker transducer of the left earphone **403** through a 3.5mm audio jack or similar interface. In some implementations, the righthand channel of the earphones **403** is reserved for sound playback.

[0049] The voltage detection circuit **412** includes a low-noise operational amplifier **414** and peripheral circuits, such as a set of one or more passive resistors and/or one or more capacitors. The amplifier **414** comprises any amplifier device(s) capable of predictable frequency responses on low frequencies (e.g., less than 1kHz) in order to avoid extra-frequency distortion on the PCG signals.

[0050] In some cases, the voltage detection circuit **412** includes one or more bypass capacitors **416** (1uF) situated before the amplifier **414**. The capacitors **416** are configured to filter away high-frequency noises from the PCG signals having a frequency above the frequency of the PCG signals. Optionally, an equivalent series of resistances (not shown) corresponding to the bypass capacitors **416** may improve a common-mode rejection ratio of the amplifier **414**, ensuring a high amplification gain. As mentioned, the inductive voltage signal (Ed) varies in proportion to the combination (Ze+ΔZ) of the impedance (Ze) at the earphones **403** and the impedance variation (ΔZ). A strong common-mode DC input may be caused by the impedance (Ze) at the earphone **403** to the amplifier **414**. Maintaining a high common-mode rejection coefficient would restrain the DC interference. The equivalent series of resistances of the bypass capacitors **416** address the issues of the common-mode rejection to the amplifier **414**.

[0051] In some embodiments, one or more circuits or devices include at least one of a resistor or a capacitor configured to filter noise of the heartbeat signal at a high-frequency above a heartbeat frequency of the heartbeat signal.

[0052] **FIG. 5** is a schematic of an impedance-matching circuit **514** of a signal processor **510** coupled to other patient-side components, according to an embodiment. The patient-side components include patient earphones **503** connected to the signal processor **510**, which is connected to a client pairing device (client device **501**). In some circumstances, the pairing device **501** receives faint or weak PCG signals when the pairing device **501** is plugged directly into a voltage detection circuit. The impedance matching circuit **514** resolves the issue to better deliver the PCG signals to the pairing device **501**.

[0053] The impedance circuit **514** performs impedance matching and impedance variation operations. The impedance circuit **514** generates an impedance variation measurement and automatically alters and matches impedance in response to variations occurring in the voltage of input electrical signals from the patient earphones **503**.

[0054] Oftentimes, when the earphone **503** connects to the pairing device **501**, a constant bias voltage electrical signal (Es) travels through an audio jack (or similar interface) of the earphone **503** and arrives at a diaphragm of the earphones **503**. The earphone **503** has an impedance value (Ze) and a voltage detection circuit of the processor **510** has an impedance value (Zp), which are serially connected to form the voltage detection circuit or voltage division circuit. An impedance variation (ΔZ) occurs at the patient earphones **503** due to the PCG signal, which is orders-of-magnitude smaller than the impedance values of the voltage division circuit (Ze, Zp). Typically, the impedance values (Ze and Zp) are constant values, so the voltage signal (Es) has a value that varies in proportion to the combined impedance values (Ze+ΔZ) of the impedance variation and the impedance (Ze) at the earphones **503**. Accordingly, it is feasible to detect the PCG signals by tracking the variance in the voltage signal (Ed). However, since the PCG signals are very weak after propagating along the human body, the variation of voltage signal (Ed) having the PCG signals would be very subtle. A voltage detection circuit described herein generates an amplified voltage signal.

[0055] Ordinarily, the amplified voltage signal (Ed) containing the PCG signals flows from the earphone **503** to the pairing device **501** through the wiring and devices of the audio chain. Unfortunately, since the impedance (Zs) of the pairing device **501** components (e.g., sound card of a laptop) differs from the impedance at the earphone **503** (e.g., Ze+Zp), only a small portion of the amplified voltage signal (Ed) will be absorbed by the pairing device **501**, which results in a very weak reception of the PCG signals at the pairing device **501**. The impedance matching circuit **514** of the signal processor **510** performs various operations for matching the impedance of the earphone **503** and the impedance of the pairing device **501**. But as explained above, impedance matching in is challenging because both the impedance of earphone **503** and the impedance of the pairing device **501** are unknown in advance, and conventional approaches to impedance matching are ineffective, inefficient, or insufficient.

[0056] To address provide impedance matching, the signal processor **510** includes the programmable impedance matching circuit **514**. The impedance matching circuit **514** includes uses a digital potentiometer circuit that determines an impedance measurement of the circuit **514** and adjusts the impedance (Zp)of the circuit **514** or other components in order to match the impedance of the earphone **503** with the impedance of the pairing device **501**. In some embodiments, the impedance of the impedance circuit **514** can be programmed with an SPI control signal from the pairing device or other administrative channel and adjusts the effective impedance of the earphone **503** to match various different pairing devices **501**.

[0057] Conventionally, approaches to impedance matching aim to match impedance at the earphones (Ze+Zp) to the impedance (Zs) of the pairing device **501**, such that most inductive voltage signal (Ed) can be delivered to the pairing device **501**. In these approaches, as the impedance (Zp) at the signal processor **510** increases to match the effective impedance (Ze+Zp) to the impedance (Zs) of the client device **501**, the voltage signal (Ed) will decline, further weaking the voltage signal any detectable PCG signals before arriving at the pairing device **501**.

[0058] To avoid the problems of conventional impedance matching, the impedance matching circuit **514** of the signal processor **510** implements an online, impedance matching function or tuning algorithm in conjunction with the pairing device **501**. The impedance matching function includes a feedback-loop-based impedance tuning algorithm that determines an optimal matching state. The impedance matching function iteratively adjusts or tunes the impedance of the circuit **514** until identifying a matching state that leads to a strongest, maximum received voltage signal (Erecv) having the highest or maximum signal-to-noise ratio (SNR). For example, a computing device or electronic circuit instructs a signal-processing device to iteratively send the probing signal to the headset according to one or more candidate impedance settings.

[0059] In many circumstances, simply using each heartbeat signal as a reference signal (Erecv) to tune the impedance of the circuit **514** would take a long time and cause delays. To expedite the impedance matching, the circuit **514** sends an active probing signal with a very short signal time period (e.g., 10ms) from the speaker of the righthand earphone (not shown). The probing signal propagates through the user's head and is captured by the speaker or transducer of the lefthand earphone **503**, which is captures and passed as a reference signal to the detection circuit (e.g., detection circuit **412** of **FIG. 4**). By taking this active probing signal as the reference signal, the impedance matching signal **514** may iterate through a searching domain within only a few seconds (e.g., 3 seconds) to identify an optimal impedance setting.

[0060] The probing signal includes consecutive chirps on the ultra-sound band (e.g., 17KHz – 22KHz), such that the chirps are imperceptible and not distracting to the patients, and the chirps do not interfere with the heart sound or motion noises. Noise-resilient chirp signals allows the circuit **514** to send the probing signals at a relatively lower power (e.g., 40dBA) and does not pose any risk of harm to the patient.

[0061] In operation, the impedance-matching circuit **514** or pairing device **501** invokes and executes an active matching function, which determines the optimal impedance value of the circuit **514**. The active matching function iterates through any number of candidate impedance value settings for the circuit **514**, which may be within the range of 0-10kΩ. The active matching function invokes a compute-energy function that measures an amount of power of a reference signal returned for a probing signal sent for a given candidate impedance setting.

[0062] In some cases, the compute-energy function generally includes four operations: removing noise of the received reference signal using a bandpass filter (BPF) with a cutoff threshold frequency at 17k and 22kHz; down-converting the waveform of the reference signal for a given candidate impedance setting to a baseband (e.g., 0–5kHz) and pass the reference signal through a lowpass filter (LPF); removing possible interference (e.g., modulated physiological signal on a chirp signal; hardware jitter noise) by applying a convolutional neural network or convolution function on the refence signal; and computing a power spectral density (PSD) indicating an amount of power. In some circumstances, down-converting a reference signal to the baseband will result in better signal quality, as LPF retains fewer residual noises at the 3dB cutoff frequency compared to a BPF.

[0063] **FIG. 6** shows a schematic of devices for processing PCG signals before applying a software pipeline of a pairing device **601**, according to an embodiment. The schematic shows components of a PCB of an signal processor **610** having a detection circuit **612** and an impedance matching circuit **614**. The detection circuit **612** includes a low-noise amplifier ("INA126") for signal detection, and the impedance matching circuit **614** includes a potentiometer chip ("MAX5402") for automatic impedance matching. Optionally, the signal processor **610** includes a switched capacitor voltage converter **616** ("LMC7660") for voltage transformation. In some embodiments, the patient can turn on/off the signal processor **610** by actuating an onboard switch button (ON/OFF switch). In some embodiments, the pairing device **601** provides a power supply to the PCB board of the signal processor **610** and sends control signals through, for example, a micro-USB interface with the pairing device **601**.

[0064] **FIG. 7** shows operational steps of method **700** for processing signals to perform cardiac auscultation remotely, according to an embodiment.

[0065] In step **701**, an audio-output speaker or transducer of a lefthand earphone of a patient captures PCG signals of the patient's heart. The diaphragm of the left hand earphone reacts to the PCG signals (e.g., vibrates) and generates a raw voltage signal, which passes to a hardware-based audio-processing device (e.g., signal processor **210**) coupled to a pairing device of the patient (e.g., patient client device **101**).

[0066] In step **703**, the signal processor receives, detects, and amplifies the PCG signals of the raw voltage signals received from the audio-output earphone worn by the patient. The signal processor includes various hardware devices (e.g., capacitors, resistors) that define a voltage detection circuit, which amplifies and detects each PCG signal in raw voltage signals that exceed a detection threshold power or frequency. The hardware of the signal processor may apply any number of filters and thresholds on the raw voltage signal in order to, for example, reduce noise or other forms of degradation affecting the PCG signals or reference signals, and/or to detect the weak PCG signals or reference signals carried with the raw voltage signals.

[0067] In step **705**, the signal processor performs an impedance matching or tuning operation to determine an optimal impedance value for a circuit of the signal processor. The optimal impedance value for the circuit occurs at an impedance matching state, when the impedance of earphone matches the impedance of the pairing device and when the reference signals produce a maximum or highest SNR at the circuit of the signal processor.

[0068] The signal processor then sends the amplified voltage signal having the PCG signal to the pairing device, which performs various software functions for correcting and improving the quality of the PCG signals.

[0069] In step **707**, the pairing device performs various pre-processing operations on a PCG signal. The pre-processing operations include, for example, filtering, spike removal, and normalization.

[0070] In some implementations, the pairing device executes pre-processing operations for filtering noise or degradations from the PCG signal. The pairing device may apply a low-pass filter (LPF) with a cutoff frequency threshold to eliminated out-band noise (e.g., ambient acoustic noise). In some cases, an administrator may preconfigure the cutoff frequency based upon a known range of heart sound. For instance, the cutoff frequency may be preconfigured by the administrator in view of the heart sound components (e.g., S1, S2, murmurs) known to be within the range of 0Hz-500Hz. As an example, in some embodiments the pairing device performs a first filter with a second-order Butterworth LPF with a cutoff frequency threshold at 500Hz to eliminate the out-band noises greater than the frequency of the heart sounds in the PCG signals.

[0071] In some implementations, the pairing device executes pre-processing operations for removing frequency energy spikes that interfere or degrade the PCG signal. These energy spikes are caused, for example, by friction and rubbing between an earphone and the patient's ear or head. These energy spikes may interfere with and degrade the PCG signal. In certain circumstances, after filtering, in-band energy spikes may persist after filtering the out-band frequencies. The pre-processing operations apply a spike removal function on the PCG signal to eliminate these energy spikes. In some embodiments, the spike removal function divides the PCG signal into time windows (e.g., 500ms time window with 250ms hop-length) and compute maximum absolute amplitudes (MAAs) over each time window. If the pairing device determines that the MAA of a time window exceeds a preconfigured energy threshold (e.g., three times a median value of all MAAs of the PCG signal), then the pairing device determines that the particular MAA is an outlier spike of the PCG signal. The spike-removal function then removes the MAA from the PCG signal.

[0072] In some implementations, the pairing device executes pre-processing operations for normalizing the PCG signal to normalized scale before feeding the PCG signal to a segmentation function of the pairing device pipeline of functions. The pairing device normalizes the PCG signal by scaling the PCG signal to a predetermine range (e.g., [-1, 1]), and then feeds the normalized PCG signal into a segmentation process (as in later step **709**). Beneficially, the normalization would not affect fine-grained cardiac characteristics hidden in the collected PCG signal because both a relative amplitude among different heart sound components and the frequencies are well preserved after normalization.

[0073] In step **709**, the pairing device performs various segmentation operations on the PCG signal, thereby segmenting the PCG signal into cardiac cycles having a cardiac cycle duration tailored for the patient. The segmentation operations include, for example, de-reverberation for reducing or removing reverberation degradation on the PCG signal, cycle boundary detection, and PCG signal refinement. The pairing device segments the pre-processed PCG signal into cardiac cycles, which later operations reference when performing frequency and energy distortion corrections. A cardiac cycle describes the sequence of electrical and mechanical events that occurs with every heartbeat. The heartbeat mechanics include a heart relaxation (diastole) and a heart contraction (systole). The duration of a cardiac cycle varies but normally lasts 0.6s – 1s. To ensure optimal performance of PCG recovery function (as in later step **711**), the segmentation operations should detect a precise boundary of each cardiac cycle of the patient.

[0074] Optionally, the pairing device may perform a de-reverberation function on the PCG signal. The PCG signals captured by the earphone must propagate over longer distances inside the human body (i.e., from the heart to the ear canal) compared to the PCG signals captured at a conventional stethoscope. The PCG signal suffers comparatively more degradation from the multi-path effect of propagating across the patient's body. The potential propagation paths followed by the PCG signals have various different lengths before reaching the audio-output speaker of the earphone, thereby creating different versions of the PCG signal that reaches the audio-output speaker of the earphone at different time intervals. Accordingly, severe reverberations are expected (i.e., inter-symbol interference) on the PCG signal, which renders the boundary of each heartbeat cardiac cycle less distinguishable. The pairing device may apply a de-reverberation filter (e.g., Wiener filter) to produce an uncorrupted or reverberation-free PCG signal by suppressing the reverberations during diastole intervals. Applying the de-reverberation filter on the PCG signal renders each cardiac cycle within the PCG signal readily distinguishable.

[0075] The pairing device detects the boundary of each cardiac cycle on the (optionally de-reverberated) PCG signal from the signal processor. For ease of explanation, the pairing device may be understood to apply the cardiac cycle segmentation operation on the PCG signal in two logical phases, boundary detection and refinement, though embodiments need not be coded or otherwise configured to act in two phases. A simplistic approach to detecting cardiac cycle boundaries within the PCG signal might include applying an amplitude threshold on the PCG signal to distinguish and filter noise from the cardiac signal (i.e., PCG signal). For instance, an electronic device or software function may apply an amplitude filter on the heartbeat cardiac signal configured for an energy threshold over a time window. However, such a design is susceptible to noise variations and could be less accurate. Rather, the pairing device applies a machine-learning architecture having execution layers that implement a hidden Markov model (HMM) for the segmentation or portion, fast boundary detection, and refinement.

[0076] For the boundary detection functions, the pairing device applies a homomorphic envelope detector, followed by a zero-phase low-pass filter, on the input PCG signal. A homomorphic envelope detector is a signal processing technique used to extract the envelope of a modulated signal, implementing functions of homomorphic signal processing and envelope detection to achieve accurate envelope extraction. The homomorphic signal processing operations include mathematical techniques that operate on a logarithm of a signal, enabling easier manipulation of the multiplication and addition operations in a frequency domain. Applying the homomorphic processing operations of the envelope detector on the PCG signal allows the pairing device to perform complex operations on the PCG signal, such as modulation and demodulation, using simple additions and multiplications. The envelope of a modulated signal refers to a slow-varying waveform that represents amplitude variations of an original signal. For example, in amplitude modulation (AM), the envelope corresponds to the original baseband signal. The homomorphic envelope detector function leverages the benefits of homomorphic signal processing to accurately extract the envelope of a modulated signal. As an example, the envelope detector executed by the pairing device performs a logarithmic conversion on the PCG signal, converting the PCG into a logarithmic representation. This conversion may include determining the natural logarithm of an absolute value of a portion of the PCG signal. The envelope detector then applies a LPF to the logarithmic representation of the PCG signal, thereby removing or filtering high-frequency portions of the PCG signal, leaving only the slower-varying portions corresponding to (revealing and detecting) the envelope. The envelope detector then performs an exponential conversion, in which the pairing device coverts the filtered signal back to the original form using an exponential operation. This process effectively recovers the envelope of the PCG signal.

[0077] The envelope detector applies the LPF (or other operations) to determine and store the profile or model of the PCG signal in a non-transitory storage medium of the pairing device contains various profiles of cardiac signals. In cases, the envelop detector removes any filtered high-frequency outliers identified by the LPF. In this way, the envelope detector applies the LPF to make the S1 and S2 peaks of the heart sound more prominent. The pairing device accesses and leverages the S1 and S2 peaks to detect a coarse-grained boundary of each cardiac cycle within the PCG signal.

[0078] In some cases, the pairing device applies an autocorrelation function on the PCG signal to detect a coarse-grained boundary of each cardiac cycle. The autocorrelation function includes various statistical techniques that measure a similarity or correlation between the PCG signal and a time-delayed version of the PCG signal. Generally, the autocorrelation function provides information about the similarity between the PCG signal and the delayed version at different time lags to, for example, identify patterns, periodicity, or statistical dependencies within the PCG signal. The autocorrelation coefficient measures or indicates the strength and direction of the correlation between the PCG signal and the delayed version of the PCG signal, obtained by dividing the autocorrelation value at lag k by the autocorrelation value at lag 0. The pairing device, for example, estimates the span of a cardiac cycle as the time from lag zero to a highest correlation coefficient for the PCG signal. As an example, the pairing device, envelope detector, or other electronic computing device applies the envelope detector having a cardiac signal profile (or types of models) on the heartbeat signal to identify one or more peaks of the heartbeat signal and detects a cardiac cycle boundary in the heartbeat signal using the one or more peaks correlated with one or more timestamps.

[0079] The output of the boundary detection function includes the cardiac cycle segments of the PCG signal based upon the coarse-grained boundaries detected for each cardiac cycle. The autocorrelation function detects an average length of multiple cardiac cycles. Oftentimes, however, the length of a cardiac cycle may change over time due to heart rate variability (HRV) of the human body, rendering the coarse-grained cycle boundaries insufficient for detailed review of the particular patient's cardiac signals for auscultation.

[0080] The pairing device applies a refinement operation that identifies and determines a precise, fine-grained boundary of each cardiac cycle in the vicinity of the timestamps of coarse-grained cardiac cycles obtained with the fast boundary detection. The pairing device feeds the portions of the PCG signals in the form of the segmented or truncated cardiac cycles into a segmentation model of the machine-learning architecture. In some embodiments, the segmentation model includes a HMM model trained to estimate a probability of an expected precise cardiac cycle boundary with logistic regression, under the supervision of PCG feature distributions, such as a S1 peaks and S2 peaks. In some implementations, the HMM of the segmentation model is trained on a public PCG feature distribution, where the pairing device (or other computing device) trained the PCG feature distribution on a large public or commercial cardiac database containing data records containing verified cardiac signal data. In some cases, the pairing device or administrative device may implement the verified cardiac signals as a ground truth of the public and/or particular patient for supervised training or fine-tuning the parameters of the executable layers of the machine-learning architecture, such as the HMM underlying the segmentation model for detecting fine-grained boundaries of cardiac cycles.

[0081] The output of the refinement function includes the cardiac cycle segments of the PCG signal based upon the fine-grained boundaries detected for each cardiac cycle.

[0082] With reference to **FIG. 8**, in some embodiments, the pairing device applies a machine-learning architecture **800** having execution layers, such as spectrogram recovery layers **802** for performing spectrogram recovery (as in step **711**) and waveform refinement layers **804** for refining the recovered PCG waveform (as in step **713**).

[0083] In step **711**, the pairing device performs a spectrogram recovery operation to recover the spectrogram of the PCG signal. The pairing device applies the machine-learning architecture **800** having the spectrogram recovery layers **802**.

[0084] In some embodiments, the spectrogram recovery layers **802** define a classic encoder-decoder model architecture (e.g., UNet) for performing the PCG spectrogram recovery. For example, UNet is a type of encoder-decoder model proven to effectively recover human vital signs and reconstruct signals, such as magnetic resonance (MR) signals. The illustrative recovery model layers **802** include, for example, six encoder layers and six decoder layers with skip connections. An encoder layer of the recovery model **802** includes a 2D convolution layer, a batch normalization (BN) layer, a ReLU function, and a dropout regularization function. In some configurations, the default stride is set to 2. A decoder layer of the recovery model **802** comprises, for example, a 2D transposed convolution layer, a BN layer, a ReLU function, and a dropout function. The S1 and S2 heart sound components normally last one-tenth of a second (0.1sec), so the kernel size of the first two convolution layers are set to 8×8, ensuring the reception field is sufficient to capture complete S1 and S2 components of the heart sound. The recovery layers **802** include an instance normalization (IN) layer, replacing the standard BN layer. The instance normalization (IN) layer may expedite training convergence. In some configurations, the frame length of each spectrogram input **801** is set to 2048, with a hop length of 1024. The machine-learning architecture **800** implements L1 loss layer that perform a loss function to measure a difference (shown as *Lspec*) between a reconstructed PCG spectrogram **805** and a ground-truth spectrogram, which may be the input spectrogram **801**, a pre-stored ground-truth spectrogram, or otherwise preconfigured ground-truth spectrogram.

[0085] The output of the spectrogram recovery layers **802** include the PCG spectrogram **805** having reconstructed frequency components. However, the phase values of the reconstructed PCG signals tend to be discontinuous, which could cause inconsistent group delay across frequencies, which could inject audible noises into the PCG signals.

[0086] In step **713**, the pairing device performs a waveform refinement operation that refines the spectrogram waveform of the PCG signal, using the reconstructed PCG spectrogram **805**, to generate an improved PCG waveform **809** for the PCG signal.

[0087] The machine-learning architecture **800** applies a transform function **806** (e.g., FFT, SFT, iSTFT) on the PCG spectrogram **805** to transform the reconstructed spectrogram **805** from a spectro-temporal energy domain representation to a time-domain waveform representation **807**. The machine-learning architecture **800** then applies the waveform refinement layers **804** on the time-domain waveform **807** of the PCG signal.

[0088] In some embodiments, the waveform refinement layers **804** include a 1D UNet encoder-decoder model for the PCG waveform refinement functions. Similar to the example spectrogram recovery model **802**, the 1D UNet of the waveform refinement model **804** includes, for example, six encoder layers and six decoder layers with skip connections. An encoder layer of the refinement model **804** comprises a 1D convolution, a BN layer, a PReLU activation function, and a dropout. The PReLU activation function allows the refinement model **804** to accept negative data sample inputs. In some configurations, the default stride is 2. A decoder layer of the refinement model **804** replaces the convolution with a 1D transposed convolution. In some configurations, because the audio wave is quasi-stationary within a very short time (e.g., 2 ms - 50 ms), the kernel size may be preconfigured to 128, thereby ensuring a 2ms reception field on the waveform **809** at 48kHz sampling rate.

[0089] The machine-learning architecture **800** applies another L1 loss layer on the output reconstructed waveform **809** of the refinement model **804**, similar to the L1 loss layer applied to reconstructed spectrogram **805** outputted by the spectrogram recovery model **802**. The second loss layer determines difference (shown as *Ltime*) between the reconstructed PCG waveform **809** and a ground-truth PCG waveform, which may be the input waveform **807**, a pre-stored ground-truth PCG waveform, or otherwise preconfigured ground-truth PCG waveform.

[0090] The machine-learning architecture may apply another loss layer that addresses certain problems caused during reconstruction and refinement operations (as in steps **711-713**). During the signal reconstruction of performed by the machine-learning architecture **800**, the changes to the signal sample may alter both the phase and frequency of the intended PCG signal, which may destroy the reconstructed spectrogram **805** and refined waveform **809**. To address this issue, the machine-learning architecture **800** includes another L1 loss function that measures the difference (*Lspec'*) between the reconstructed waveform **809** spectrum or spectrogram (not shown) after the waveform refinement model **804** compared against the transformed input waveform **807** or spectrogram **805** after the recovery model **802**. This additional loss layer helps tune the parameters of the refinement model **804** to be sensitive to phase refinement and preservation during overall PCG waveform reconstruction and refinement (in steps **711-713**).

[0091] In some embodiments, these reconstruction models **802**, **804** of the machine-learning architecture **800** are connected and executed in a series and applies an integrated loss layer that performs an integrated loss function using weighted combination of each of the three loss functions (*Lspec*, *Ltime*, *Lspec'*). The integrated loss function includes additional weighting coefficients (α, β). For example, α is a weight assigned to the first loss value (*Lspec*) and β is a weight assigned to the third loss value (*Lspec'*), where α is manually configured to be, for example, 10-times bigger than β to prioritize the spectrogram recovery performance during the training. Oftentimes, during the model training, the final output PCG waveform **809** contains some high-frequency artifacts above the PCG frequency band. The machine-learning architecture 800 apply an LPF with 500 Hz cutoff frequencies to the waveform **809** output to eliminate the out-band audio artifacts.

[0092] In step **715**, the pairing device outputs the PCG signal into one or more machine-readable format for transmitting and presenting the PCG signal to the care provider at the care provider's client device. The pairing device receives the final PCG waveforms **809** generated by the machine-learning architecture **800** and generates a machine-readable format for presenting the PCG signal using the final PCG waveforms **809**. The format of the PCG signal data may include any multimedia format, include audio, visual, or audiovisual data formats.

[0093] The pairing device may transmit the PCG signal data to a provider client device or to a server hosting a website accessed by the provider device. The PCG signal data may be transmitted to the provider device or the server as a media computer file or as media data stream. The pairing device and the provider device may execute software (e.g., browser, remote care software) for establishing a remote care session using any number of computing networking and remote communications protocols. The provider device executes the remote care software to obtain (receive, retrieve) the PCG signal data and present or playback the PCG signal to the care provider according to the particular data format, via a user interface of the provider device.

[0094] Embodiments may include electronic devices and computing processing hardware for performing various processes and methods for performing cardiac auscultation.

[0095] In embodiments, a system, device, or processor-implemented (e.g., at least one processor of a computer or electronic device) method is for performing cardiac auscultation. A processor may obtain a raw audio signal containing a heartbeat signal produced from an earphone. The earphone is configured to output an analog audio waveform for an ear canal of an end-user. The processor may determine a cardiac cycle duration for the end-user. The processor may segment the heartbeat signal into a plurality of heartbeat cycle segments according to the cardiac cycle duration of the end-user. The processor may generate a spectrogram of the heartbeat signal in a frequency-domain by applying a neural network architecture on the plurality of heartbeat cycle segments. The processor may apply a transform function on the spectrogram to transform the spectrogram to a time-domain, thereby generating a refined heartbeat signal. The processor may transmit the refined heartbeat signal in a machine-readable format to a provider computer.

[0096] When determining the cardiac cycle duration for segmenting, the processor may apply a de-reverberation filter on the heartbeat signal, thereby removing reverberation degradation on the heartbeat signal.

[0097] When determining the cardiac cycle duration for segmenting, processor may apply an envelope detector having a cardiac signal profile on the heartbeat signal to identify one or more peaks of the heartbeat signal. The processor may detect a cardiac cycle boundary in the heartbeat signal using the one or more peaks correlated with one or more timestamps.

[0098] For each cardiac cycle boundary of a plurality of cardiac cycle boundaries detected in the heartbeat signal, the processor may apply a hidden Markov model on a cardiac cycle portion of the heartbeat signal within the cardiac cycle boundary to determine a precise cardiac cycle boundary for the one or more timestamps correlated with the one or more peaks of the cardiac cycle boundary. The cardiac cycle boundary indicates the cardiac cycle duration of the end-user for segmenting the heartbeat signal.

[0099] Obtaining the heartbeat signal may include preprocessing the heartbeat signal, where the processor may apply a low-pass filter configured for a threshold low frequency on the heartbeat signal. Obtaining the heartbeat signal may include preprocessing the heartbeat signal, where the processor may apply an amplitude filter configured for an energy threshold over a time window on the heartbeat signal. Obtaining the heartbeat signal may preprocessing the heartbeat signal, where the processor may normalize the heartbeat signal.

[0100] The machine-readable format of the refined heartbeat signal may include at least one of an audio format, a visual format, or an audiovisual format.

[0101] In embodiments a system, device, or processor-implemented (e.g., processor of a computer or electronic device) method is for processing signals from audio earphones for cardiac auscultation. A processor may instruct a signal-processing device to transmit a plurality of probing signals to a headset of a user according to a plurality of candidate impedance settings of the signal-processing device. For each candidate impedance setting of the plurality of candidate impedance settings, the processor may obtain a reference signal received at the signal-processing device from an audio-output earphone of the headset. The reference signal corresponds to a probing signal sent to the headset according to the candidate impedance setting. The processor may determine an amount of power of the reference signal occurring at the candidate impedance setting. The processor may compute a signal-to-noise ratio using the amount of power of the reference signal occurring at the candidate impedance setting. The processor may identify the candidate impedance setting having a comparative maximum signal-to-noise ratio as an optimal impedance match between a first impedance of the signal-processing device and a second impedance of the computer. The processor may adjust impedance of the signal-processing device according to the optimal impedance match.

[0102] When determining the amount of power of the reference signal, the processor may compute a power spectral density for the reference signal indicating the amount of power of the reference signal.

[0103] When obtaining the reference signal for the candidate impedance setting, the processor may down-convert the reference signal to a preconfigured baseband. When obtaining the reference signal for the candidate impedance setting, the processor may apply a neural network architecture on the reference signal trained to remove one or more types of degradation. When obtaining the reference signal for the candidate impedance setting, the processor may apply one or more filters on the reference signal according to one or more threshold frequencies.

[0104] The processor may instruct the signal-processing device to iteratively send the probing signal to the headset according to one or more candidate impedance settings. The processor may iteratively adjust the first impedance of the signal-processing device according to one or more candidate impedance settings until identifying the optimal impedance match.

[0105] In embodiments, a device is for processing signals to perform cardiac auscultation via earphones. The device comprises a detection circuit and an impedance-matching circuit. The detection circuit includes: an interface configured to obtain a raw voltage signal containing a heartbeat signal produced from an audio-output earphone of a headset; and an amplifier configured to amplify the heartbeat signal in the raw voltage signal, thereby generating an amplified voltage signal containing the heartbeat signal. The impedance-matching circuit includes a programmable controller configured to: adjust a first impedance of the impedance-matching circuit to match a second impedance of a pairing device according to a tuning instruction received from the pairing device indicating an optimal impedance value for the first impedance.

[0106] The amplifier of the detection circuit may be coupled to a transducer of the audio-output earphone. The audio-output earphone may be a left-ear earphone of the headset.

[0107] The detection circuit may include one or more peripheral circuits, including at least one of a resistor or a capacitor configured to filter noise of the heartbeat signal at a high-frequency above a heartbeat frequency of the heartbeat signal.

[0108] The impedance-matching circuit may be further configured to: transmit one or more probing signals to at least one earphone of the headset according to one or more candidate impedance settings; and for each probing signal of the one or more probing signals transmitted for the one or more candidate impedance settings: receive from the audio-output earphone a reference signal responsive to the probing signal; and transmit the reference signal corresponding to the probing signal to the pairing device.

[0109] The impedance-matching circuit may be further configured to: obtain each reference signal from the audio-output earphone; determine an optimal impedance using each reference signal; and adjust impedance of the device according to the optimal impedance.

[0110] The various illustrative logical blocks, modules, circuits, and algorithm steps described in connection with the embodiments disclosed herein may be implemented as electronic hardware, computer software, or combinations of both. To clearly illustrate this interchangeability of hardware and software, various illustrative components, blocks, modules, circuits, and steps have been described above generally in terms of their functionality. Whether such functionality is implemented as hardware or software depends upon the particular application and design constraints imposed on the overall system. Skilled artisans may implement the described functionality in varying ways for each particular application, but such implementation decisions should not be interpreted as causing a departure from the scope of the present invention.

[0111] Embodiments implemented in computer software may be implemented in software, firmware, middleware, microcode, hardware description languages, or any combination thereof. A code segment or machine-executable instructions may represent a procedure, a function, a subprogram, a program, a routine, a subroutine, a module, a software package, a class, or any combination of instructions, data structures, or program statements. A code segment may be coupled to another code segment or a hardware circuit by passing and/or receiving information, data, arguments, attributes, or memory contents. Information, arguments, attributes, data, etc. may be passed, forwarded, or transmitted via any suitable means including memory sharing, message passing, token passing, network transmission, etc.

[0112] The actual software code or specialized control hardware used to implement these systems and methods is not limiting of the invention. Thus, the operation and behavior of the systems and methods were described without reference to the specific software code being understood that software and control hardware can be designed to implement the systems and methods based on the description herein.

[0113] When implemented in software, the functions may be stored as one or more instructions or code on a non-transitory computer-readable or processor-readable storage medium. The steps of a method or algorithm disclosed herein may be embodied in a processor-executable software module which may reside on a computer-readable or processor-readable storage medium. A non-transitory computer-readable or processor-readable media includes both computer storage media and tangible storage media that facilitate transfer of a computer program from one place to another. A non-transitory processor-readable storage media may be any available media that may be accessed by a computer. By way of example, and not limitation, such non-transitory processor-readable media may comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other tangible storage medium that may be used to store desired program code in the form of instructions or data structures and that may be accessed by a computer or processor. Disk and disc, as used herein, include compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk, and blu-ray disc where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Combinations of the above should also be included within the scope of computer-readable media. Additionally, the operations of a method or algorithm may reside as one or any combination or set of codes and/or instructions on a non-transitory processor-readable medium and/or computer-readable medium, which may be incorporated into a computer program product.

[0114] The preceding description of the disclosed embodiments is provided to enable any person skilled in the art to make or use the present invention. Various modifications to these embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without departing from the spirit or scope of the invention. Thus, the present invention is not intended to be limited to the embodiments shown herein but is to be accorded the widest scope consistent with the following claims and the principles and novel features disclosed herein.

[0115] While various aspects and embodiments have been disclosed, other aspects and embodiments are contemplated. The various aspects and embodiments disclosed are for purposes of illustration and are not intended to be limiting, with the true scope and spirit being indicated by the following claims.

CLAIMS

What is claimed is:

1. A computer-implemented method for performing cardiac auscultation, the method comprising:

obtaining, by a computer, a raw audio signal containing a heartbeat signal produced from an earphone, the earphone configured to output an analog audio waveform for an ear canal of an end-user;

determining, by the computer, a cardiac cycle duration for the end-user;

segmenting, by the computer, the heartbeat signal into a plurality of heartbeat cycle segments according to the cardiac cycle duration of the end-user;

generating, by the computer, a spectrogram of the heartbeat signal in a frequency-domain by applying a neural network architecture on the plurality of heartbeat cycle segments;

applying, by the computer, a transform function on the spectrogram to transform the spectrogram to a time-domain, thereby generating a refined heartbeat signal; and

transmitting, by the computer, the refined heartbeat signal in a machine-readable format to a provider computer.

2. The method according to claim 1, wherein determining the cardiac cycle duration for segmenting includes applying, by the computer, a de-reverberation filter on the heartbeat signal, thereby removing reverberation degradation on the heartbeat signal.

3. The method according to claim 1, wherein determining the cardiac cycle duration for segmenting includes:

applying, by the computer, an envelope detector having a cardiac signal profile on the heartbeat signal to identify one or more peaks of the heartbeat signal; and

detecting, by the computer, a cardiac cycle boundary in the heartbeat signal using the one or more peaks correlated with one or more timestamps.

4. The method according to claim 3, further comprising, for each cardiac cycle boundary of a plurality of cardiac cycle boundaries detected in the heartbeat signal:

applying, by the computer, a hidden Markov model on a cardiac cycle portion of the heartbeat signal within the cardiac cycle boundary to determine a precise cardiac cycle boundary for the one or more timestamps correlated with the one or more peaks of the cardiac cycle boundary,

wherein the cardiac cycle boundary indicates the cardiac cycle duration of the end-user for segmenting the heartbeat signal.

5. The method according to claim 1, wherein obtaining the heartbeat signal includes preprocessing the heartbeat signal by applying, by the computer, a low-pass filter configured for a threshold low frequency on the heartbeat signal.

6. The method according to claim 1, wherein obtaining the heartbeat signal includes preprocessing the heartbeat signal by applying, by the computer, an amplitude filter configured for an energy threshold over a time window on the heartbeat signal.

7. The method according to claim 1, wherein obtaining the heartbeat signal includes preprocessing the heartbeat signal by normalizing, by the computer, the heartbeat signal.

8. The method according to claim 1, wherein the machine-readable format of the refined heartbeat signal includes at least one of an audio format, a visual format, or an audiovisual format.

9. A computer-implemented method for processing signals from audio earphones for cardiac auscultation, the method comprising:

instructing, by a computer, a signal-processing device to transmit a plurality of probing signals to a headset of a user according to a plurality of candidate impedance settings of the signal-processing device;

for each candidate impedance setting of the plurality of candidate impedance settings:

obtaining, by the computer, a reference signal received at the signal-processing device from an audio-output earphone of the headset, wherein the reference signal corresponds to a probing signal sent to the headset according to the candidate impedance setting;

determining, by the computer, an amount of power of the reference signal occurring at the candidate impedance setting; and

computing, by the computer, a signal-to-noise ratio using the amount of power of the reference signal occurring at the candidate impedance setting;

identifying, by the computer, the candidate impedance setting having a comparative maximum signal-to-noise ratio as an optimal impedance match between a first impedance of the signal-processing device and a second impedance of the computer; and

adjusting, by the computer, impedance of the signal-processing device according to the optimal impedance match.

10. The method according to claim 9, wherein determining the amount of power of the reference signal includes computing, by the computer, a power spectral density for the reference signal indicating the amount of power of the reference signal.

11. The method according to claim 9, wherein obtaining the reference signal for the candidate impedance setting includes down-converting, by the computer, the reference signal to a preconfigured baseband.

12. The method according to claim 9, wherein obtaining the reference signal for the candidate impedance setting includes applying, by the computer, a neural network architecture on the reference signal trained to remove one or more types of degradation.

13. The method according to claim 9, wherein obtaining the reference signal for the candidate impedance setting includes applying, by the computer, one or more filters on the reference signal according to one or more threshold frequencies.

14. The method according to claim 9, wherein the computer instructs the signal-processing device to iteratively send the probing signal to the headset according to one or more candidate impedance settings.

15. The method according to claim 9, wherein the computer iteratively adjusts the first impedance of the signal-processing device according to one or more candidate impedance settings until identifying the optimal impedance match.

16. A device for processing signals to perform cardiac auscultation via earphones, the device comprising:

a detection circuit including:

an interface configured to obtain a raw voltage signal containing a heartbeat signal produced from an audio-output earphone of a headset; and

an amplifier configured to amplify the heartbeat signal in the raw voltage signal, thereby generating an amplified voltage signal containing the heartbeat signal; and

an impedance-matching circuit having a programmable controller configured to:

adjust a first impedance of the impedance-matching circuit to match a second impedance of a pairing device according to a tuning instruction received from the pairing device indicating an optimal impedance value for the first impedance.

17. The device of claim 16, wherein the amplifier of the detection circuit is coupled to a transducer of the audio-output earphone, and wherein the audio-output earphone is a left-ear earphone of the headset.

18. The device of claim 16, wherein the detection circuit further includes one or more peripheral circuits including at least one of a resistor or a capacitor configured to filter noise of the heartbeat signal at a high-frequency above a heartbeat frequency of the heartbeat signal.

19. The device of claim 16, wherein the impedance-matching circuit is further configured to:

transmit one or more probing signals to at least one earphone of the headset according to one or more candidate impedance settings; and

for each probing signal of the one or more probing signals transmitted for the one or more candidate impedance settings:

receive from the audio-output earphone a reference signal responsive to the probing signal; and

transmit the reference signal corresponding to the probing signal to the pairing device.

20. The device of claim 19, wherein the impedance-matching circuit is further configured to:

obtain each reference signal from the audio-output earphone;

determine an optimal impedance using each reference signal; and

adjust impedance of the device according to the optimal impedance.

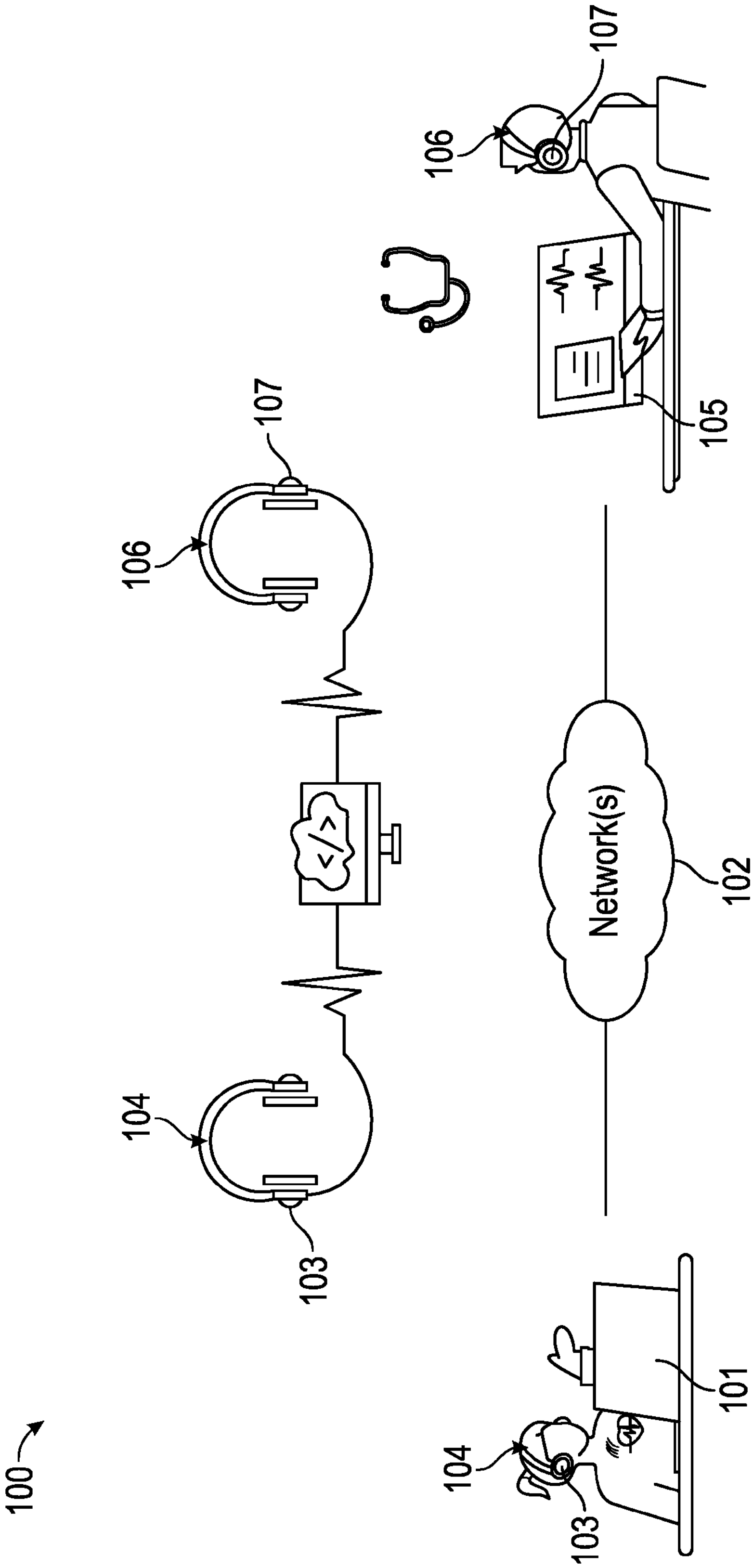

FIG. 1

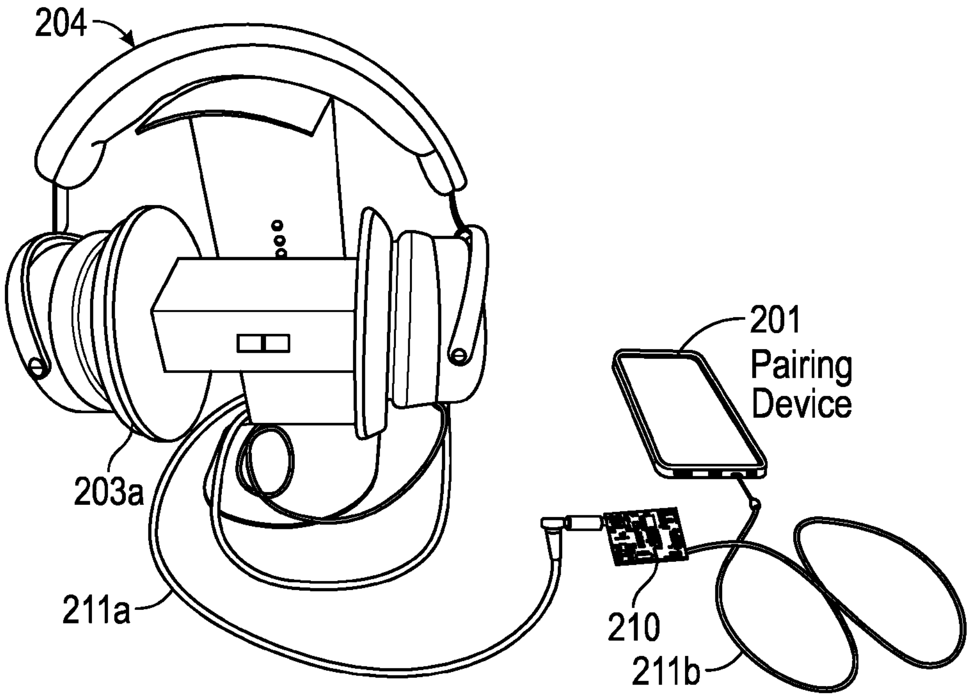
204
201
Pairing
Device
203a
211a
210
211b

FIG. 2A

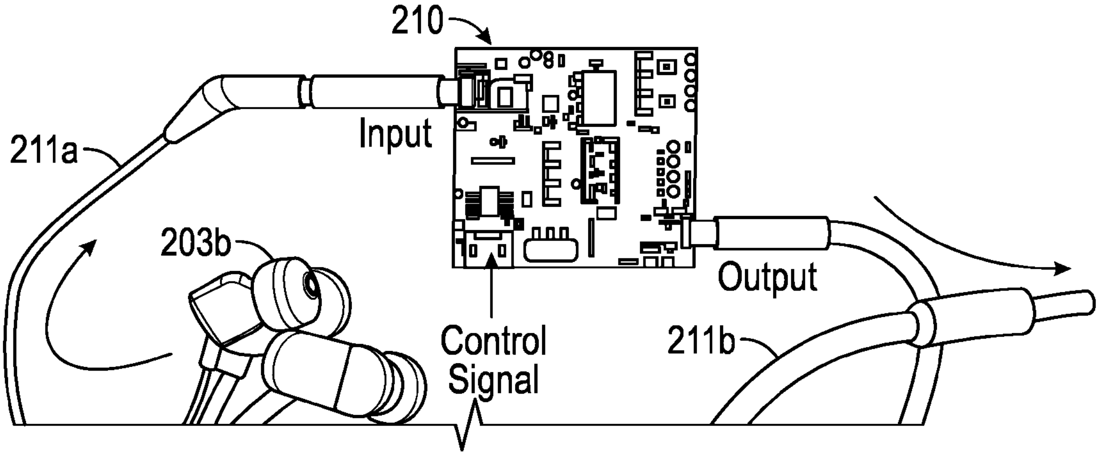
210
211a
Input
203b
Control
Signal
Output
211b

FIG. 2B

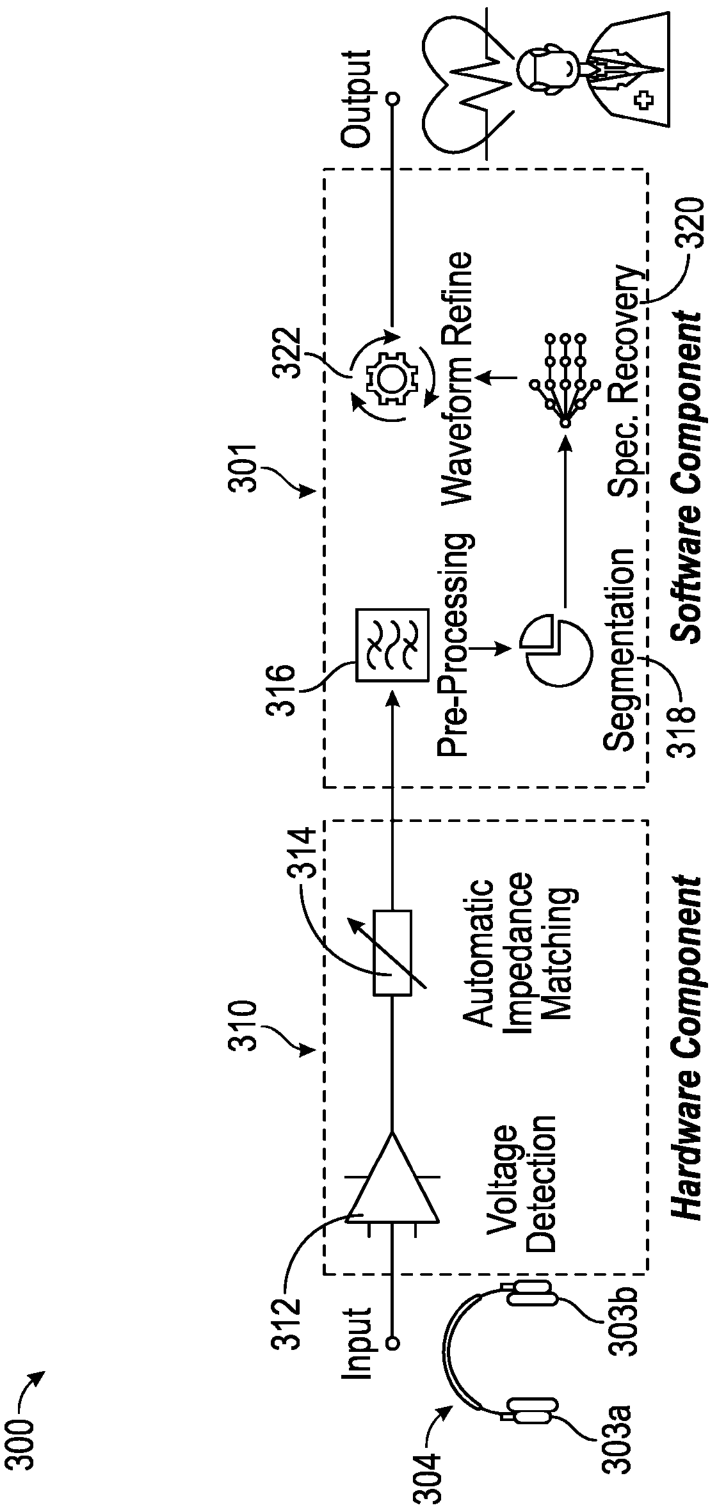
300
304
303a
303b
Input
312
Voltage
Detection
310
314
Automatic
Impedance
Matching
Hardware Component
316
Pre-Processing
318
Segmentation
301
Waveform Refine
322
Spec. Recovery
320
Software Component
Output

FIG. 3

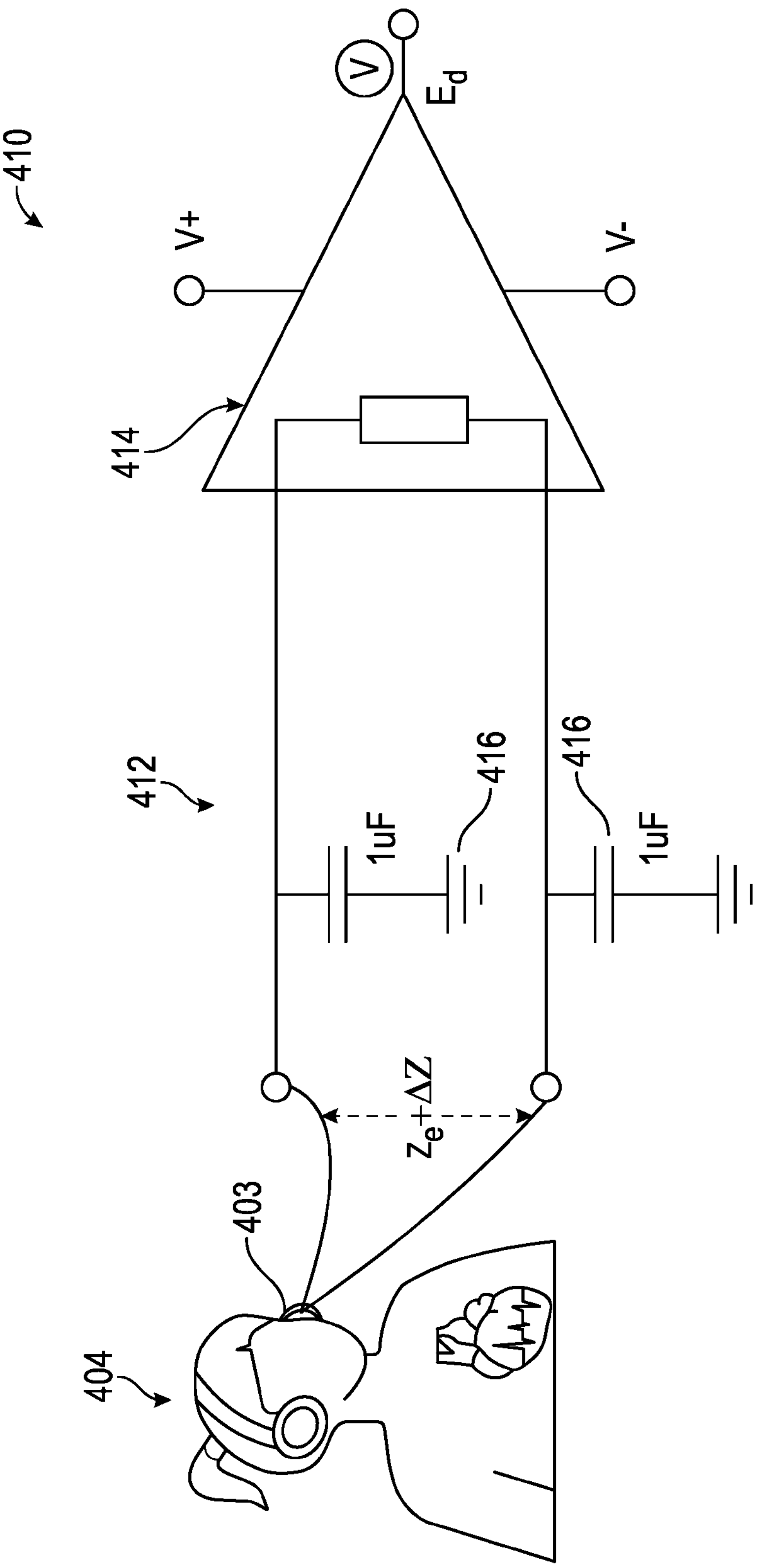

FIG. 4

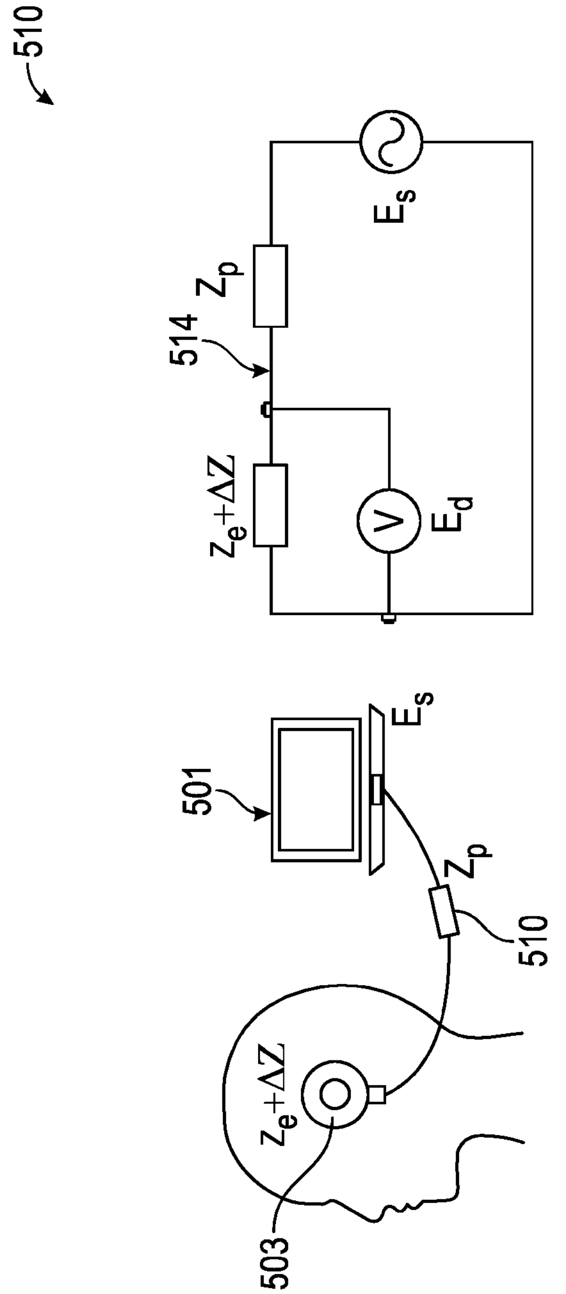

FIG. 5

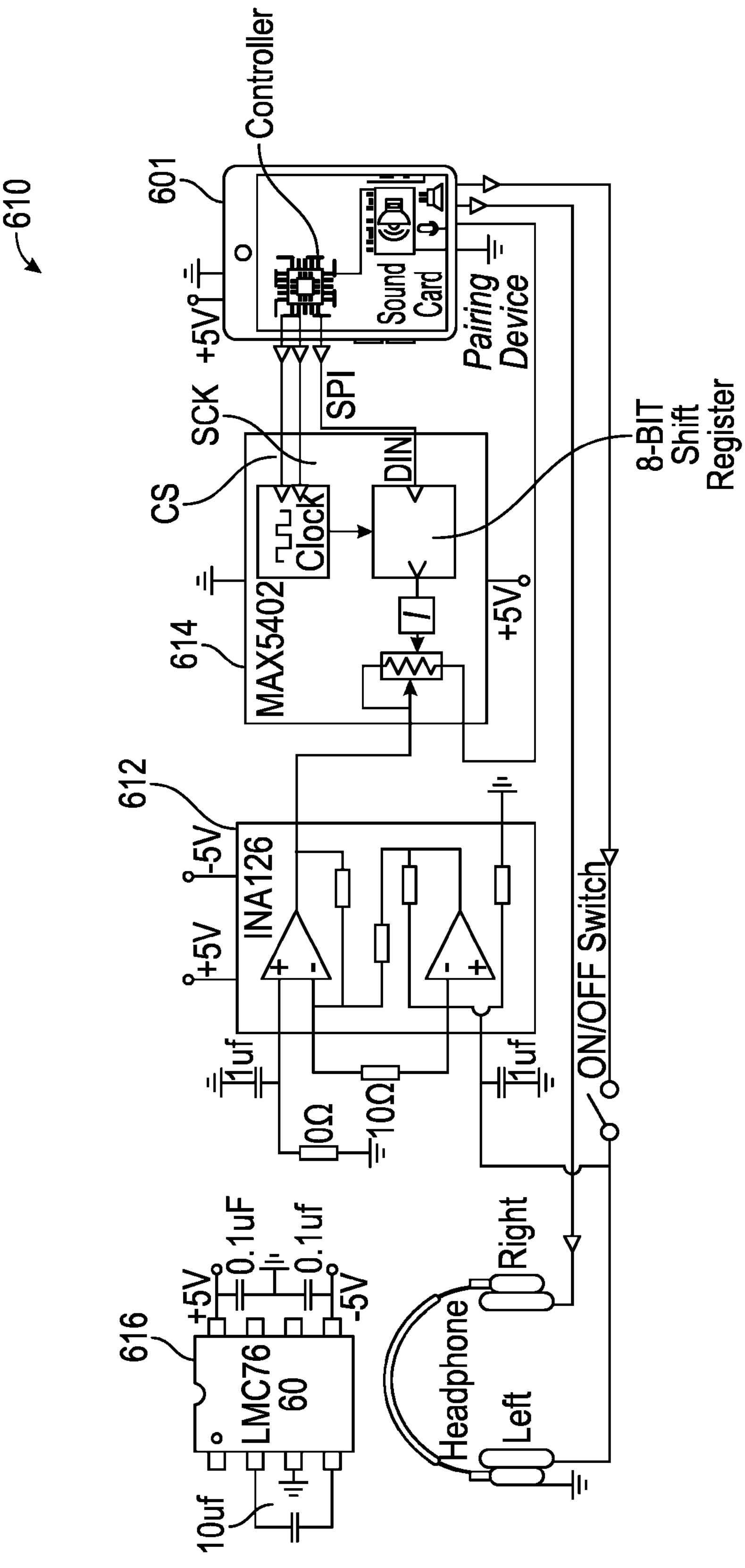
610
601
Controller
+5V
Sound
Card
Pairing
Device
SCK
SPI
CS
DIN
Clock
8-BIT
Shift
Register
MAX5402
614
+5V
612
INA126
-5V
+5V
1uf
1uf
0Ω
10Ω
ON/OFF Switch
616
+5V
0.1uF
0.1uf
-5V
LMC76
60
10uf
Right
Headphone
Left

FIG. 6

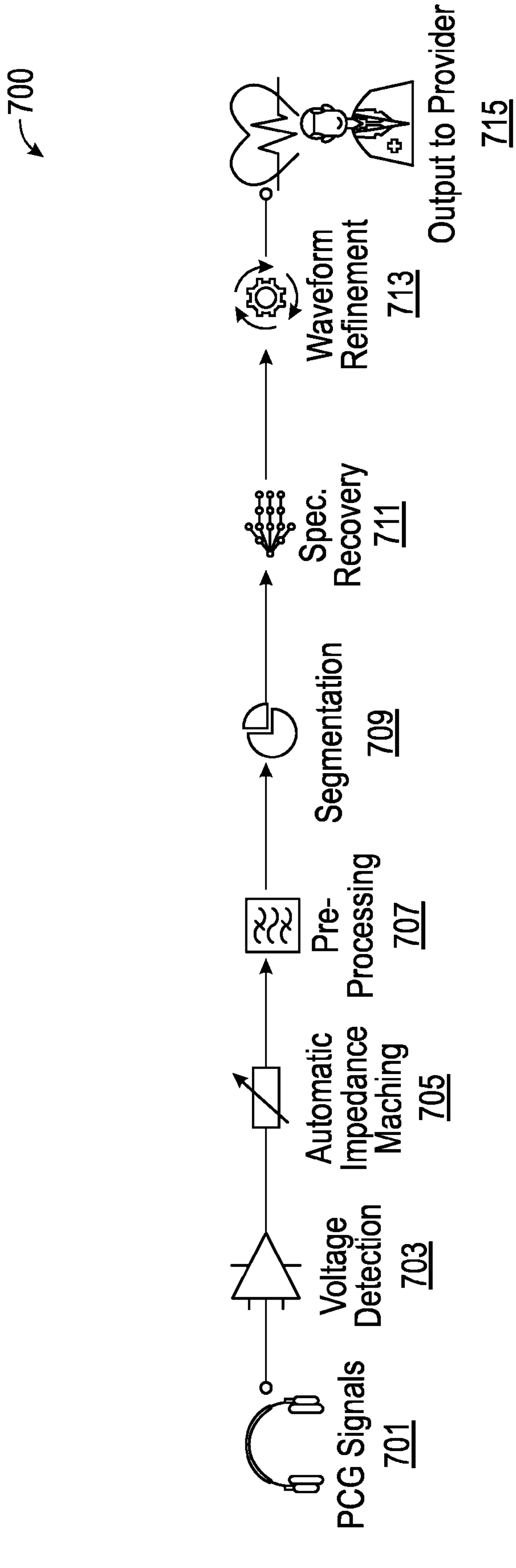
700
PCG Signals
701
Voltage Detection
703
Automatic Impedance Maching
705
Pre-Processing
707
Segmentation
709
Spec. Recovery
711
Waveform Refinement
713
Output to Provider
715

FIG. 7

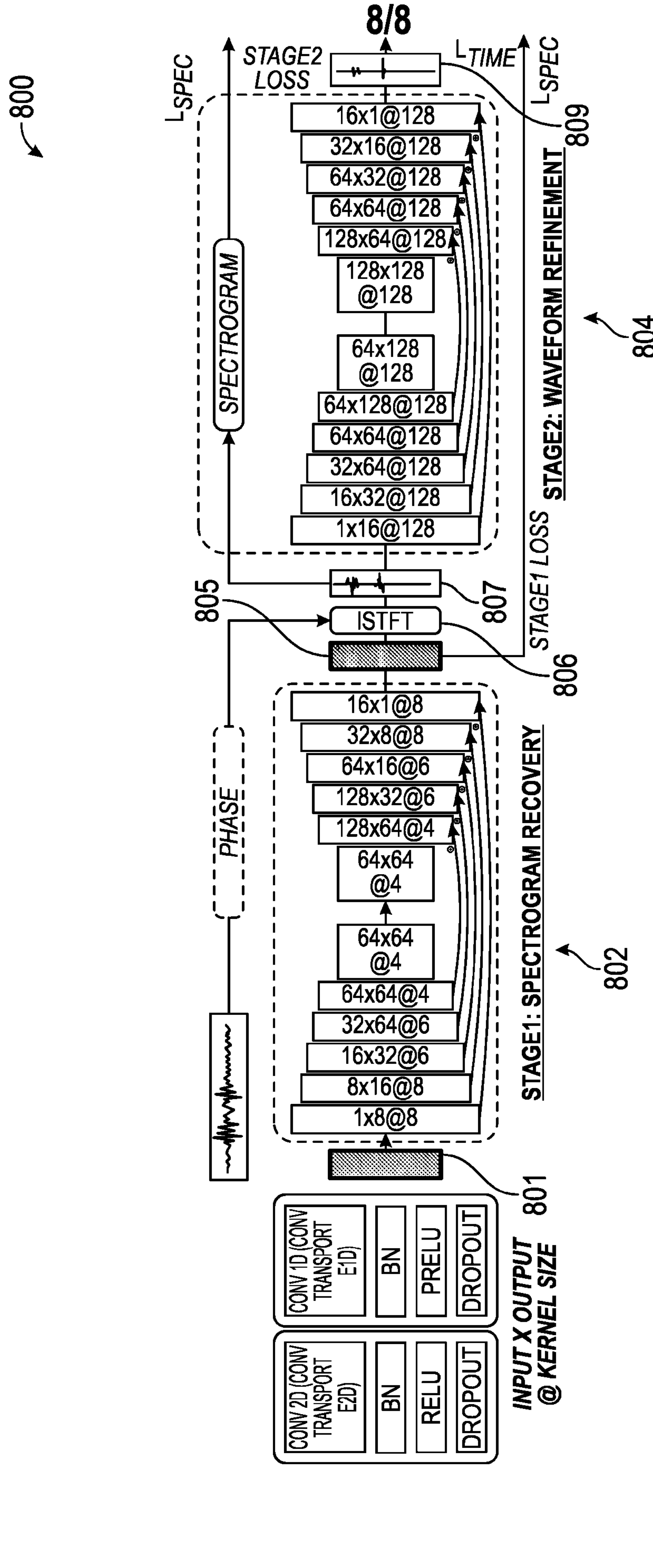
800
CONV 2D (CONV TRANSPORT E2D)
BN
RELU
DROPOUT
CONV 1D (CONV TRANSPORT E1D)
BN
PRELU
DROPOUT
INPUT X OUTPUT @ KERNEL SIZE
801
PHASE
1x8@8
8x16@8
16x32@6
32x64@6
64x64@4
64x64 @4
64x64 @4
128x64@4
128x32@6
64x16@6
32x8@8
16x1@8
STAGE1: SPECTROGRAM RECOVERY
802
805
ISTFT
806
807
STAGE1 LOSS
SPECTROGRAM
1x16@128
16x32@128
32x64@128
64x64@128
64x128@128
64x128 @128
128x128 @128
128x64@128
64x64@128
64x32@128
32x16@128
16x1@128
STAGE2: WAVEFORM REFINEMENT
804
L_{SPEC}
STAGE2 LOSS
L_{TIME}
L_{SPEC}
809

FIG. 8