



(11) **EP 3 485 655 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention  
of the grant of the patent:  
**03.01.2024 Bulletin 2024/01**

(51) International Patent Classification (IPC):  
**H04S 7/00** <sup>(2006.01)</sup> **H04R 27/00** <sup>(2006.01)</sup>  
**H04R 29/00** <sup>(2006.01)</sup>

(21) Application number: **17754501.9**

(52) Cooperative Patent Classification (CPC):  
**H04S 7/301; H04R 27/00; H04R 29/007;**  
**H04R 2227/005; H04R 2227/007**

(22) Date of filing: **14.07.2017**

(86) International application number:  
**PCT/US2017/042191**

(87) International publication number:  
**WO 2018/013959 (18.01.2018 Gazette 2018/03)**

(54) **SPECTRAL CORRECTION USING SPATIAL CALIBRATION**

**SPEKTRALE KORREKTUR MITTELS RÄUMLICHER KALIBRIERUNG**

**CORRECTION SPECTRALE À L'AIDE D'UN ÉTALONNAGE SPATIAL**

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB**  
**GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO**  
**PL PT RO RS SE SI SK SM TR**

(30) Priority: **15.07.2016 US 201615211835**  
**15.07.2016 US 201615211822**

(43) Date of publication of application:  
**22.05.2019 Bulletin 2019/21**

(60) Divisional application:  
**23212793.6**

(73) Proprietor: **Sonos Inc.**  
**Santa Barbara, CA 93101 (US)**

(72) Inventor: **SHEEN, Timothy**  
**Santa Barbara**  
**CA 93101 (US)**

(74) Representative: **EIP**  
**Fairfax House**  
**15 Fulwood Place**  
**London WC1V 6HU (GB)**

(56) References cited:  
**WO-A1-2011/139502 US-A1- 2014 003 635**  
**US-A1- 2016 011 850**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

## Description

[0001] The disclosure is related to consumer goods and, more particularly, to methods, systems, products, features, services, and other elements directed to media playback or some aspect thereof.

[0002] Options for accessing and listening to digital audio in an out-loud setting were limited until in 2003, when SONOS, Inc. filed for one of its first patent applications, entitled "Method for Synchronizing Audio Playback between Multiple Networked Devices," and began offering a media playback system for sale in 2005. The Sonos Wireless HiFi System enables people to experience music from many sources via one or more networked playback devices. Through a software control application installed on a smartphone, tablet, or computer, one can play what he or she wants in any room that has a networked playback device. Additionally, using the controller, for example, different songs can be streamed to each room with a playback device, rooms can be grouped together for synchronous playback, or the same song can be heard in all rooms synchronously.

[0003] Given the ever growing interest in digital media, there continues to be a need to develop consumer-accessible technologies to further enhance the listening experience.

[0004] US 2016/011850 A1 relates to a method of detecting a trigger condition that initiates calibration of a media playback system capable of multi-channel audio playback, emitting periodic calibration audio from playback devices, detecting the emitted calibration sound via a microphone, analyzing the calibration audio and calibrating the media playback system accordingly

[0005] WO 2011/139502 A1 relates to a calibration process for a surround sound system triggered by a user via a user interface as to determine filter settings for room equalisation.

[0006] US 2014/003635 A1 shows to the concept of microphone array beamforming in the context of speaker calibration of a surround sound system.

[0007] According to a first aspect of the present invention, there is provided a method as defined in appended claim 1. According to a second aspect of the present invention, there is provided a media playback system as defined in appended claim 15.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Features, aspects, and advantages of the presently disclosed technology may be better understood with regard to the following description, appended claims, and accompanying drawings where:

Figure 1 shows an example media playback system configuration in which certain embodiments may be practiced;

Figure 2 shows a functional block diagram of an example playback device;

Figure 3 shows a functional block diagram of an example control device;

Figure 4 shows an example controller interface;

Figure 5 shows an example control device;

Figure 6 shows a smartphone that is displaying an example control interface, according to an example implementation;

Figure 7 illustrates an example movement through an example environment in which an example media playback system is positioned;

Figure 8 illustrates an example chirp that increases in frequency over time;

Figure 9 shows an example brown noise spectrum;

Figures 10A and 10B illustrate transition frequency ranges of example hybrid calibration sounds,

Figure 11 shows a frame illustrating an iteration of an example periodic calibration sound;

Figure 12 shows a series of frames illustrating iterations of an example periodic calibration sound;

Figure 13 shows an example flow diagram to facilitate a spatial calibration;

Figure 14 shows example frames illustrating calibration audio that is divided into frames and slots;

Figure 15 shows a smartphone that is displaying an example control interface, according to an example implementation

Figure 16 shows a smartphone that is displaying an example control interface, according to an example implementation

Figure 17 shows a smartphone that is displaying an example control interface, according to an example implementation;

Figure 18 shows a smartphone that is displaying an example control interface, according to an example implementation;

Figure 19 shows an example flow diagram to facilitate a spatial calibration using an applied spatial calibration;

Figure 20 shows an example flow diagram to facilitate a spatial calibration using an applied spatial calibration; and

Figure 21 shows a smartphone that is displaying an example control interface, according to an example implementation.

[0009] The drawings are for the purpose of illustrating example embodiments, but it is understood that the inventions are not limited to the arrangements and instrumentality shown in the drawings.

## DETAILED DESCRIPTION

### I. Overview

[0010] Embodiments described herein involve, inter alia, techniques to facilitate calibration of a media playback system. Calibration procedures contemplated herein involve a recording device (e.g., a networked micro-

phone device (NMD)) detecting sound waves (e.g., one or more calibration sounds) that were emitted by one or more playback devices of a media playback system. A processing device, such as a recording device, a playback device or another device that is communicatively coupled to the media playback system, analyzes the detected sound waves to determine one or more calibrations for the one or more playback devices of the media playback system. When applied, such calibrations configure the one or more playback devices to a given listening area (i.e., the environment in which the playback device(s) were positioned while emitting the sound waves).

**[0011]** The processing device determines a spatial calibration that configures the one or more playback devices to a given listening area spatially. Such a calibration configures the one or more playback devices to one or more particular locations within the environment (e.g., one or more preferred listening positions, such as favorite seating location), perhaps by adjusting time-delay and/or loudness for those particular locations. A spatial calibration includes one or more filters that may include delay and/or phase adjustment, gain adjustment, and/or any other adjustment to correct for the spatial placement of the playback device(s) relative to the one or more particular locations within the environment.

**[0012]** As noted above, during a calibration procedure, one or more playback devices of a media playback system output calibration sounds. The media playback systems according to the invention include multiple audio drivers, which are divided among the playback device(s) of a media playback system in various arrangements. For instance, an example media playback system may include a soundbar-type playback device with multiple audio drivers (e.g., nine audio drivers). Another playback device might include multiple audio drivers of different types (e.g., tweeters and woofers, perhaps of varying size). Other example playback devices may include a single audio driver (e.g., a single full-range woofer in a playback device, or a large low-frequency woofer in a subwoofer-type device).

**[0013]** In operation, multiple audio drivers of a media playback system form one or more "sound axes." Each such "sound axis" corresponds to a respective input channel of audio content. In some implementations, two or more audio drivers are arrayed to form a sound axis. For instance, a sound-bar type device might include nine audio drivers which form multiple sound axes (e.g., front, left, and right surround sound channels). Any audio driver contributes to any number of sound axes. For example, a left axis of a surround sound system may be formed by contributions from all nine audio drivers in the example sound-bar type device. Alternatively, an axis may be formed by a single audio driver.

**[0014]** Example media playback systems described herein adopt various playback configurations representing respective sets of sound axes. Example playback configurations may include respective configurations

based on the number of input channels (e.g., mono, stereo, surround, or any of the above in combination with a subwoofer). Other example playback configurations may be based on the content type. For instance, a first set of axes may be formed by audio drivers of a media playback system when playing music and a second set of axes formed by the audio drivers when playing audio that is paired with video (e.g., television audio). Other playback configurations may be invoked by various groupings of playback devices within the media playback system. Many examples are possible.

**[0015]** During calibration procedures, the multiple audio drivers of the media playback system form the one or more sound axes, such that each sound axis outputs sound during the calibration procedure. For instance, calibration audio emitted by multiple audio drivers is divided into constituent frames. Each frame is in turn divided into slots. During each slot of a given frame, a respective sound axis is formed by outputting audio. In this manner, an NMD that is recording the audio output of the audio drivers can obtain samples from each sound axis. The frames may repeat, so as to produce multiple samples for each sound axis when recorded by the NMD.

**[0016]** Another type of calibration that is produced by calibration procedures described herein is a spectral calibration. A spectral calibration may configure the playback device(s) of a media playback system across a given listening area spectrally. Such a calibration may help offset acoustic characteristics of the environment generally instead of being relatively more directed to particular listening locations like the spatial calibrations. A spectral calibration may include one or more filters that adjust the frequency response of the playback devices. In operation, one of the two or more calibrations may be applied to playback by the one or more playback devices, perhaps for different use cases. Example uses cases might include music playback or surround sound (i.e., home theater), among others.

**[0017]** In some example calibration procedures contemplated herein, a media playback system may perform a first calibration to determine a spatial calibration for playback device(s) of the media playback system. The media playback system then applies the spatial calibration while the playback devices are emitting audio during a second calibration to determine a spectral calibration. Such a calibration procedure may yield a calibration that includes both spatial and spectral correction.

**[0018]** Example techniques may involve performing aspects of a spatial calibration. A first implementation may include detecting a trigger condition that initiates calibration of a media playback system including multiple audio drivers that form multiple sound axes, each sound axis corresponding to a respective channel of multi-channel audio content. The first implementation may also include causing the multiple audio drivers to emit calibration audio that is divided into constituent frames, the multiple sound axes emitting calibration audio during respective slots of each constituent frame. The first implemen-

tation may further include recording, via a microphone, the emitted calibration audio. The first implementation may include causing delays for each sound axis of the multiple sound axes to be determined, the determined delay for each sound axis based on the slots of recorded calibration audio corresponding to the sound axes and causing the multiple sound axes to be calibrated. Calibrating the multiple sound axes may involve causing audio output of the multiple sound axes to be delayed according to the respective determined delays.

**[0019]** The implementation may include receiving data representing one or more spatial filters that correspond to respective playback configurations. Each playback configuration represents a particular set of sound axes formed via one or more audio drivers and each sound axis corresponds to a respective input channel of audio content. The implementation also involves causing the one or more audio drivers to output calibration audio that is divided into a repeating set of frames, the set of frames including a respective frame for each playback configuration. Causing the one or more audio drivers to output the calibration audio involves causing an audio stage to apply, during each frame, the spatial filter corresponding to the respective playback configuration. The second implementation may also include receiving data representing one or more spectral filters that correspond to respective playback configurations, the one or more spectral filters based on the calibration audio output by the one or more audio drivers. When playing back audio content in a given playback configuration, the audio stage may apply a particular spectral filter corresponding to the given playback configuration.

**[0020]** The implementation includes detecting a trigger condition that initiates calibration of a media playback system for multiple playback configurations. Each playback configuration represents a particular set of sound axes formed via multiple audio drivers of the media playback system and each sound axis corresponds to a respective channel of audio content. The implementation also involves causing the multiple audio drivers to output calibration audio that is divided into a repeating set of frames, the set of frames including a respective frame for each playback configuration. Causing the multiple audio drivers to output the calibration audio involves causing, during each frames of the set of frames, a respective set of spatial filters to be applied to the multiple audio drivers, each set of spatial filters including a respective spatial filter for each sound axis. The implementation further involves recording, via the microphone, the calibration audio output by the multiple audio drivers and causing a processing device to determine respective sets of spectral filters for the multiple playback configurations based on the recorded calibration audio, each set of spectral filters including a respective spectral filter for each sound axis.

**[0021]** The above implementation may be embodied as a method, a device configured to carry out the implementation, a system of devices configured to carry out

the implementation, or a non-transitory computer-readable medium containing instructions that are executable by one or more processors to carry out the implementation, among other examples. It will be understood by one of ordinary skill in the art that this disclosure includes numerous other embodiments, including combinations of the example features described herein. Further, any example operation described as being performed by a given device to illustrate a technique may be performed by any suitable devices, including the devices described herein. Yet further, any device may cause another device to perform any of the operations described herein.

**[0022]** While some examples described herein may refer to functions performed by given actors such as "users" and/or other entities, it should be understood that this description is for purposes of explanation only. The claims should not be interpreted to require action by any such example actor unless explicitly required by the language of the claims themselves.

## II. Example Operating Environment

**[0023]** Figure 1 illustrates an example configuration of a media playback system 100 in which one or more embodiments disclosed herein may be practiced or implemented. The media playback system 100 as shown is associated with an example home environment having several rooms and spaces, such as for example, a master bedroom, an office, a dining room, and a living room. As shown in the example of Figure 1, the media playback system 100 includes playback devices 102-124, control devices 126 and 128, and a wired or wireless network router 130.

**[0024]** Further discussions relating to the different components of the example media playback system 100 and how the different components may interact to provide a user with a media experience may be found in the following sections. While discussions herein may generally refer to the example media playback system 100, technologies described herein are not limited to applications within, among other things, the home environment as shown in Figure 1. For instance, the technologies described herein may be useful in environments where multi-zone audio may be desired, such as, for example, a commercial setting like a restaurant, mall or airport, a vehicle like a sports utility vehicle (SUV), bus or car, a ship or boat, an airplane, and so on.

### a. Example Playback Devices

**[0025]** Figure 2 shows a functional block diagram of an example playback device 200 that may be configured to be one or more of the playback devices 102-124 of the media playback system 100 of Figure 1. The playback device 200 may include a processor 202, software components 204, memory 206, audio processing components 208, audio amplifier(s) 210, speaker(s) 212, and a network interface 214 including wireless interface(s) 216

and wired interface(s) 218. In one case, the playback device 200 may not include the speaker(s) 212, but rather a speaker interface for connecting the playback device 200 to external speakers. In another case, the playback device 200 may include neither the speaker(s) 212 nor the audio amplifier(s) 210, but rather an audio interface for connecting the playback device 200 to an external audio amplifier or audio-visual receiver.

**[0026]** In one example, the processor 202 may be a clock-driven computing component configured to process input data according to instructions stored in the memory 206. The memory 206 may be a tangible computer-readable medium configured to store instructions executable by the processor 202. For instance, the memory 206 may be data storage that can be loaded with one or more of the software components 204 executable by the processor 202 to achieve certain functions. In one example, the functions may involve the playback device 200 retrieving audio data from an audio source or another playback device. In another example, the functions may involve the playback device 200 sending audio data to another device or playback device on a network. In yet another example, the functions may involve pairing of the playback device 200 with one or more playback devices to create a multi-channel audio environment.

**[0027]** Certain functions may involve the playback device 200 synchronizing playback of audio content with one or more other playback devices. During synchronous playback, a listener will preferably not be able to perceive time-delay differences between playback of the audio content by the playback device 200 and the one or more other playback devices. U.S. Patent No. 8,234,395 entitled, "System and method for synchronizing operations among a plurality of independently clocked digital data processing devices," provides in more detail some examples for audio playback synchronization among playback devices.

**[0028]** The memory 206 may further be configured to store data associated with the playback device 200, such as one or more zones and/or zone groups the playback device 200 is a part of, audio sources accessible by the playback device 200, or a playback queue that the playback device 200 (or some other playback device) may be associated with. The data maybe stored as one or more state variables that are periodically updated and used to describe the state of the playback device 200. The memory 206 may also include the data associated with the state of the other devices of the media system, and shared from time to time among the devices so that one or more of the devices have the most recent data associated with the system. Other embodiments are also possible.

**[0029]** The audio processing components 208 may include one or more digital-to-analog converters (DAC), an audio preprocessing component, an audio enhancement component or a digital signal processor (DSP), and so on. In one embodiment, one or more of the audio processing components 208 may be a subcomponent of the proc-

essor 202. In one example, audio content may be processed and/or intentionally altered by the audio processing components 208 to produce audio signals. The produced audio signals may then be provided to the audio amplifier(s) 210 for amplification and playback through speaker(s) 212. Particularly, the audio amplifier(s) 210 may include devices configured to amplify audio signals to a level for driving one or more of the speakers 212. The speaker(s) 212 may include an individual transducer (e.g., a "driver") or a complete speaker system involving an enclosure with one or more drivers. A particular driver of the speaker(s) 212 may include, for example, a sub-woofer (e.g., for low frequencies), a mid-range driver (e.g., for middle frequencies), and/or a tweeter (e.g., for high frequencies). In some cases, each transducer in the one or more speakers 212 may be driven by an individual corresponding audio amplifier of the audio amplifiers) 210. In addition to producing analog signals for playback by the playback device 200, the audio processing components 208 may be configured to process audio content to be sent to one or more other playback devices for playback.

**[0030]** Audio content to be processed and/or played back by the playback device 200 may be received from an external source, such as via an audio line-in input connection (e.g., an auto-detecting 3.5mm audio line-in connection) or the network interface 214.

**[0031]** The network interface 214 may be configured to facilitate a data flow between the playback device 200 and one or more other devices on a data network. As such, the playback device 200 may be configured to receive audio content over the data network from one or more other playback devices in communication with the playback device 200, network devices within a local area network, or audio content sources over a wide area network such as the Internet. In one example, the audio content and other signals transmitted and received by the playback device 200 may be transmitted in the form of digital packet data containing an Internet Protocol (IP)-based source address and IP-based destination addresses. In such a case, the network interface 214 may be configured to parse the digital packet data such that the data destined for the playback device 200 is properly received and processed by the playback device 200.

**[0032]** As shown, the network interface 214 may include wireless interface(s) 216 and wired interface(s) 218. The wireless interface(s) 216 may provide network interface functions for the playback device 200 to wirelessly communicate with other devices (e.g., other playback device(s), speaker(s), receiver(s), network device(s), control device(s) within a data network the playback device 200 is associated with) in accordance with a communication protocol (e.g., any wireless standard including IEEE 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, 802.15, 4G mobile communication standard, and so on). The wired interface(s) 218 may provide network interface functions for the playback device 200 to communicate over a wired connection with other devices

in accordance with a communication protocol (e.g., IEEE 802.3). While the network interface 214 shown in Figure 2 includes both wireless interface(s) 216 and wired interface(s) 218, the network interface 214 may in some embodiments include only wireless interface(s) or only wired interface(s).

**[0033]** In one example, the playback device 200 and one other playback device may be paired to play two separate audio components of audio content. For instance, playback device 200 may be configured to play a left channel audio component, while the other playback device may be configured to play a right channel audio component, thereby producing or enhancing a stereo effect of the audio content. The paired playback devices (also referred to as "bonded playback devices") may further play audio content in synchrony with other playback devices.

**[0034]** In another example, the playback device 200 may be sonically consolidated with one or more other playback devices to form a single, consolidated playback device. A consolidated playback device may be configured to process and reproduce sound differently than an unconsolidated playback device or playback devices that are paired, because a consolidated playback device may have additional speaker drivers through which audio content may be rendered. For instance, if the playback device 200 is a playback device designed to render low frequency range audio content (i.e. a subwoofer), the playback device 200 may be consolidated with a playback device designed to render full frequency range audio content. In such a case, the full frequency range playback device, when consolidated with the low frequency playback device 200, may be configured to render only the mid and high frequency components of audio content, while the low frequency range playback device 200 renders the low frequency component of the audio content. The consolidated playback device may further be paired with a single playback device or yet another consolidated playback device.

**[0035]** By way of illustration, SONOS, Inc. presently offers (or has offered) for sale certain playback devices including a "PLAY:1," "PLAY:3," "PLAY:5," "PLAYBAR," "CONNECT:AMP," "CONNECT," and "SUB." Any other past, present, and/or future playback devices may additionally or alternatively be used to implement the playback devices of example embodiments disclosed herein. Additionally, it is understood that a playback device is not limited to the example illustrated in Figure 2 or to the SONOS product offerings. For example, a playback device may include a wired or wireless headphone. In another example, a playback device may include or interact with a docking station for personal mobile media playback devices. In yet another example, a playback device may be integral to another device or component such as a television, a lighting fixture, or some other device for indoor or outdoor use.

## b. Example Playback Zone Configurations

**[0036]** Referring back to the media playback system 100 of Figure 1, the environment may have one or more playback zones, each with one or more playback devices. The media playback system 100 may be established with one or more playback zones, after which one or more zones may be added, or removed to arrive at the example configuration shown in Figure 1. Each zone may be given a name according to a different room or space such as an office, bathroom, master bedroom, bedroom, kitchen, dining room, living room, and/or balcony. In one case, a single playback zone may include multiple rooms or spaces. In another case, a single room or space may include multiple playback zones.

**[0037]** As shown in Figure 1, the balcony, dining room, kitchen, bathroom, office, and bedroom zones each have one playback device, while the living room and master bedroom zones each have multiple playback devices. In the living room zone, playback devices 104, 106, 108, and 110 may be configured to play audio content in synchrony as individual playback devices, as one or more bonded playback devices, as one or more consolidated playback devices, or any combination thereof. Similarly, in the case of the master bedroom, playback devices 122 and 124 may be configured to play audio content in synchrony as individual playback devices, as a bonded playback device, or as a consolidated playback device.

**[0038]** In one example, one or more playback zones in the environment of Figure 1 may each be playing different audio content. For instance, the user may be grilling in the balcony zone and listening to hip hop music being played by the playback device 102 while another user may be preparing food in the kitchen zone and listening to classical music being played by the playback device 114. In another example, a playback zone may play the same audio content in synchrony with another playback zone. For instance, the user may be in the office zone where the playback device 118 is playing the same rock music that is being playing by playback device 102 in the balcony zone. In such a case, playback devices 102 and 118 may be playing the rock music in synchrony such that the user may seamlessly (or at least substantially seamlessly) enjoy the audio content that is being played out-loud while moving between different playback zones. Synchronization among playback zones may be achieved in a manner similar to that of synchronization among playback devices, as described in previously referenced U.S. Patent No. 8,234,395.

**[0039]** As suggested above, the zone configurations of the media playback system 100 may be dynamically modified, and in some embodiments, the media playback system 100 supports numerous configurations. For instance, if a user physically moves one or more playback devices to or from a zone, the media playback system 100 may be reconfigured to accommodate the change(s). For instance, if the user physically moves the playback device 102 from the balcony zone to the office zone, the

office zone may now include both the playback device 118 and the playback device 102. The playback device 102 may be paired or grouped with the office zone and/or renamed if so desired via a control device such as the control devices 126 and 128. On the other hand, if the one or more playback devices are moved to a particular area in the home environment that is not already a playback zone, a new playback zone may be created for the particular area.

**[0040]** Further, different playback zones of the media playback system 100 may be dynamically combined into zone groups or split up into individual playback zones. For instance, the dining room zone and the kitchen zone 114 may be combined into a zone group for a dinner party such that playback devices 112 and 114 may render audio content in synchrony. On the other hand, the living room zone may be split into a television zone including playback device 104, and a listening zone including playback devices 106, 108, and 110, if the user wishes to listen to music in the living room space while another user wishes to watch television.

#### c. Example Control Devices

**[0041]** Figure 3 shows a functional block diagram of an example control device 300 that may be configured to be one or both of the control devices 126 and 128 of the media playback system 100. Control device 300 may also be referred to as a controller 300. As shown, the control device 300 may include a processor 302, memory 304, a network interface 306, and a user interface 308. In one example, the control device 300 may be a dedicated controller for the media playback system 100. In another example, the control device 300 may be a network device on which media playback system controller application software may be installed, such as for example, an iPhone™, iPad™ or any other smart phone, tablet or network device (e.g., a networked computer such as a PC or Mac™).

**[0042]** The processor 302 may be configured to perform functions relevant to facilitating user access, control, and configuration of the media playback system 100. The memory 304 may be configured to store instructions executable by the processor 302 to perform those functions. The memory 304 may also be configured to store the media playback system controller application software and other data associated with the media playback system 100 and the user.

**[0043]** In one example, the network interface 306 may be based on an industry standard (e.g., infrared, radio, wired standards including IEEE 802.3, wireless standards including IEEE 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, 802.15, 4G mobile communication standard, and so on). The network interface 306 may provide a means for the control device 300 to communicate with other devices in the media playback system 100. In one example, data and information (e.g., such as a state variable) may be communicated between control

device 300 and other devices via the network interface 306. For instance, playback zone and zone group configurations in the media playback system 100 may be received by the control device 300 from a playback device or another network device, or transmitted by the control device 300 to another playback device or network device via the network interface 306. In some cases, the other network device may be another control device.

**[0044]** Playback device control commands such as volume control and audio playback control may also be communicated from the control device 300 to a playback device via the network interface 306. As suggested above, changes to configurations of the media playback system 100 may also be performed by a user using the control device 300. The configuration changes may include adding/removing one or more playback devices to/from a zone, adding/removing one or more zones to/from a zone group, forming a bonded or consolidated player, separating one or more playback devices from a bonded or consolidated player, among others. Accordingly, the control device 300 may sometimes be referred to as a controller, whether the control device 300 is a dedicated controller or a network device on which media playback system controller application software is installed.

**[0045]** The user interface 308 of the control device 300 may be configured to facilitate user access and control of the media playback system 100, by providing a controller interface such as the controller interface 400 shown in Figure 4. The controller interface 400 includes a playback control region 410, a playback zone region 420, a playback status region 430, a playback queue region 440, and an audio content sources region 450. The user interface 400 as shown is just one example of a user interface that may be provided on a network device such as the control device 300 of Figure 3 (and/or the control devices 126 and 128 of Figure 1) and accessed by users to control a media playback system such as the media playback system 100. Other user interfaces of varying formats, styles, and interactive sequences may alternatively be implemented on one or more network devices to provide comparable control access to a media playback system.

**[0046]** The playback control region 410 may include selectable (e.g., by way of touch or by using a cursor) icons to cause playback devices in a selected playback zone or zone group to play or pause, fast forward, rewind, skip to next, skip to previous, enter/exit shuffle mode, enter/exit repeat mode, enter/exit cross fade mode. The playback control region 410 may also include selectable icons to modify equalization settings, and playback volume, among other possibilities.

**[0047]** The playback zone region 420 may include representations of playback zones within the media playback system 100. In some embodiments, the graphical representations of playback zones may be selectable to bring up additional selectable icons to manage or configure the playback zones in the media playback system,

such as a creation of bonded zones, creation of zone groups, separation of zone groups, and renaming of zone groups, among other possibilities.

**[0048]** For example, as shown, a "group" icon may be provided within each of the graphical representations of playback zones. The "group" icon provided within a graphical representation of a particular zone may be selectable to bring up options to select one or more other zones in the media playback system to be grouped with the particular zone. Once grouped, playback devices in the zones that have been grouped with the particular zone will be configured to play audio content in synchrony with the playback device(s) in the particular zone. Analogously, a "group" icon may be provided within a graphical representation of a zone group. In this case, the "group" icon may be selectable to bring up options to deselect one or more zones in the zone group to be removed from the zone group. Other interactions and implementations for grouping and ungrouping zones via a user interface such as the user interface 400 are also possible. The representations of playback zones in the playback zone region 420 may be dynamically updated as playback zone or zone group configurations are modified.

**[0049]** The playback status region 430 may include graphical representations of audio content that is presently being played, previously played, or scheduled to play next in the selected playback zone or zone group. The selected playback zone or zone group may be visually distinguished on the user interface, such as within the playback zone region 420 and/or the playback status region 430. The graphical representations may include track title, artist name, album name, album year, track length, and other relevant information that may be useful for the user to know when controlling the media playback system via the user interface 400.

**[0050]** The playback queue region 440 may include graphical representations of audio content in a playback queue associated with the selected playback zone or zone group. In some embodiments, each playback zone or zone group may be associated with a playback queue containing information corresponding to zero or more audio items for playback by the playback zone or zone group. For instance, each audio item in the playback queue may comprise a uniform resource identifier (URI), a uniform resource locator (URL) or some other identifier that may be used by a playback device in the playback zone or zone group to find and/or retrieve the audio item from a local audio content source or a networked audio content source, possibly for playback by the playback device.

**[0051]** In one example, a playlist may be added to a playback queue, in which case information corresponding to each audio item in the playlist may be added to the playback queue. In another example, audio items in a playback queue may be saved as a playlist. In a further example, a playback queue may be empty, or populated but "not in use" when the playback zone or zone group

is playing continuously streaming audio content, such as Internet radio that may continue to play until otherwise stopped, rather than discrete audio items that have playback durations. In an alternative embodiment, a playback queue can include Internet radio and/or other streaming audio content items and be "in use" when the playback zone or zone group is playing those items. Other examples are also possible.

**[0052]** When playback zones or zone groups are "grouped" or "ungrouped," playback queues associated with the affected playback zones or zone groups may be cleared or re-associated. For example, if a first playback zone including a first playback queue is grouped with a second playback zone including a second playback queue, the established zone group may have an associated playback queue that is initially empty, that contains audio items from the first playback queue (such as if the second playback zone was added to the first playback zone), that contains audio items from the second playback queue (such as if the first playback zone was added to the second playback zone), or a combination of audio items from both the first and second playback queues. Subsequently, if the established zone group is ungrouped, the resulting first playback zone may be re-associated with the previous first playback queue, or be associated with a new playback queue that is empty or contains audio items from the playback queue associated with the established zone group before the established zone group was ungrouped. Similarly, the resulting second playback zone may be re-associated with the previous second playback queue, or be associated with a new playback queue that is empty, or contains audio items from the playback queue associated with the established zone group before the established zone group was ungrouped. Other examples are also possible.

**[0053]** Referring back to the user interface 400 of Figure 4, the graphical representations of audio content in the playback queue region 440 may include track titles, artist names, track lengths, and other relevant information associated with the audio content in the playback queue. In one example, graphical representations of audio content may be selectable to bring up additional selectable icons to manage and/or manipulate the playback queue and/or audio content represented in the playback queue. For instance, a represented audio content may be removed from the playback queue, moved to a different position within the playback queue, or selected to be played immediately, or after any currently playing audio content, among other possibilities. A playback queue associated with a playback zone or zone group may be stored in a memory on one or more playback devices in the playback zone or zone group, on a playback device that is not in the playback zone or zone group, and/or some other designated device. Playback of such a playback queue may involve one or more playback devices playing back media items of the queue, perhaps in sequential or random order.

**[0054]** The audio content sources region 450 may in-



clude graphical representations of selectable audio content sources from which audio content may be retrieved and played by the selected playback zone or zone group. Discussions pertaining to audio content sources may be found in the following section.

**[0055]** Figure 5 depicts a smartphone 500 that includes one or more processors, a tangible computer-readable memory, a network interface, and a display. Smartphone 500 might be an example implementation of control device 126 or 128 of Figure 1, or control device 300 of Figure 3, or other control devices described herein. By way of example, reference will be made to smartphone 500 and certain control interfaces, prompts, and other graphical elements that smartphone 500 may display when operating as a control device of a media playback system (e.g., of media playback system 100). Within examples, such interfaces and elements may be displayed by any suitable control device, such as a smartphone, tablet computer, laptop or desktop computer, personal media player, or a remote control device.

**[0056]** While operating as a control device of a media playback system, smartphone 500 may display one or more controller interface, such as controller interface 400. Similar to playback control region 410, playback zone region 420, playback status region 430, playback queue region 440, and/or audio content sources region 450 of Figure 4, smartphone 500 might display one or more respective interfaces, such as a playback control interface, a playback zone interface, a playback status interface, a playback queue interface, and/or an audio content sources interface. Example control devices might display separate interfaces (rather than regions) where screen size is relatively limited, such as with smartphones or other handheld devices.

#### d. Example Audio Content Sources

**[0057]** As indicated previously, one or more playback devices in a zone or zone group may be configured to retrieve for playback audio content (e.g., according to a corresponding URI or URL for the audio content) from a variety of available audio content sources. In one example, audio content may be retrieved by a playback device directly from a corresponding audio content source (e.g., a line-in connection). In another example, audio content may be provided to a playback device over a network via one or more other playback devices or network devices.

**[0058]** Example audio content sources may include a memory of one or more playback devices in a media playback system such as the media playback system 100 of Figure 1, local music libraries on one or more network devices (such as a control device, a network-enabled personal computer, or a networked-attached storage (NAS), for example), streaming audio services providing audio content via the Internet (e.g., the cloud), or audio sources connected to the media playback system via a line-in input connection on a playback device or network device, among other possibilities.

**[0059]** In some embodiments, audio content sources may be regularly added or removed from a media playback system such as the media playback system 100 of Figure 1. In one example, an indexing of audio items may be performed whenever one or more audio content sources are added, removed or updated. Indexing of audio items may involve scanning for identifiable audio items in all folders/directory shared over a network accessible by playback devices in the media playback system, and generating or updating an audio content database containing metadata (e.g., title, artist, album, track length, among others) and other associated information, such as a URI or URL for each identifiable audio item found. Other examples for managing and maintaining audio content sources may also be possible.

#### e. Example Calibration Sequence

**[0060]** As indicated above, example calibration procedures involve one or more playback devices emitting a calibration sound, which may be detected by a recording device (or multiple recording devices).

**[0061]** In some embodiments, the detected calibration sounds may be analyzed across a range of frequencies over which the playback device is to be calibrated (i.e., a calibration range). Accordingly, the particular calibration sound that is emitted by a playback device covers the calibration frequency range. The calibration frequency range may include a range of frequencies that the playback device is capable of emitting (e.g., 15 - 30,000 Hz) and may be inclusive of frequencies that are considered to be in the range of human hearing (e.g., 20 - 20,000 Hz). By emitting and subsequently detecting a calibration sound covering such a range of frequencies, a frequency response that is inclusive of that range may be determined for the playback device. Such a frequency response may be representative of the environment in which the playback device emitted the calibration sound.

**[0062]** In some embodiments, a playback device may repeatedly emit the calibration sound during the calibration procedure such that the calibration sound covers the calibration frequency range during each repetition. With a moving microphone, repetitions of the calibration sound are continuously detected at different physical locations within the environment. For instance, the playback device might emit a periodic calibration sound. Each period of the calibration sound may be detected by the recording device at a different physical location within the environment thereby providing a sample (i.e., a frame representing a repetition) at that location. Such a calibration sound may therefore facilitate a space-averaged calibration of the environment. When multiple microphones are utilized, each microphone may cover a respective portion of the environment (perhaps with some overlap).

**[0063]** Yet further, the recording devices may measure both moving and stationary samples. For instance, while the one or more playback devices output a calibration sound, a recording device may move within the environ-

ment. During such movement, the recording device may pause at one or more locations to measure stationary samples. Such locations may correspond to preferred listening locations. In another example, a first recording device and a second recording device may include a first microphone and a second microphone respectively. While the playback device emits a calibration sound, the first microphone may move and the second microphone may remain stationary, perhaps at a particular listening location within the environment (e.g., a favorite chair).

**[0064]** In some cases, the one or more playback devices may be joining into a grouping, such as a bonded zone or zone group. In such cases, the calibration procedure may calibrate the one or more playback devices as a group. Example groupings include zone groups or bonded pairs, among other example configurations.

**[0065]** The playback device(s) under calibration initiates the calibration procedure based on a trigger condition. For instance, a recording device, such as control device 126 of media playback system 100, may detect a trigger condition that causes the recording device to initiate calibration of one or more playback devices (e.g., one or more of playback devices 102-124). Alternatively, a playback device of a media playback system may detect such a trigger condition (and then perhaps relay an indication of that trigger condition to the recording device).

**[0066]** In some embodiments, detecting the trigger condition may involve detecting input data indicating a selection of a selectable control. For instance, a recording device, such as control device 126, may display an interface (e.g., control interface 400 of Figure 4), which includes one or more controls that, when selected, initiate calibration of a playback device, or a group of playback devices (e.g., a zone).

**[0067]** To illustrate such a control, Figure 6 shows smartphone 500 which is displaying an example control interface 600. Control interface 600 includes a graphical region 602 that prompts to tap selectable control 604 (Start) when ready. When selected, selectable control 604 may initiate the calibration procedure. As shown, selectable control 604 is a button control. While a button control is shown by way of example, other types of controls are contemplated as well.

**[0068]** Control interface 600 further includes a graphical region 606 that includes a video depicting how to assist in the calibration procedure. Some calibration procedures may involve moving a microphone through an environment in order to obtain samples of the calibration sound at multiple physical locations. In order to prompt a user to move the microphone, the control device may display a video or animation depicting the step or steps to be performed during the calibration.

**[0069]** To illustrate movement of the control device during calibration, Figure 7 shows media playback system 100 of Figure 1. Figure 7 shows a path 700 along which a recording device (e.g., control device 126) might be moved during calibration. As noted above, the recording

device may indicate how to perform such a movement in various ways, such as by way of a video or animation, among other examples. A recording device might detect iterations of a calibration sound emitted by one or more playback devices of media playback system 100 at different points along the path 700, which may facilitate a space-averaged calibration of those playback devices.

**[0070]** In other examples, detecting the trigger condition may involve a playback device detecting that the playback device has become uncalibrated, which might be caused by moving the playback device to a different position. For example, the playback device may detect physical movement via one or more sensors that are sensitive to movement (e.g., an accelerometer). As another example, the playback device may detect that it has been moved to a different zone (e.g., from a "Kitchen" zone to a "Living Room" zone), perhaps by receiving an instruction from a control device that causes the playback device to leave a first zone and join a second zone.

**[0071]** In further examples, detecting the trigger condition may involve a recording device (e.g., a control device or playback device) detecting a new playback device in the system. Such a playback device may have not yet been calibrated for the environment. For instance, a recording device may detect a new playback device as part of a set-up procedure for a media playback system (e.g., a procedure to configure one or more playback devices into a media playback system). In other cases, the recording device may detect a new playback device by detecting input data indicating a request to configure the media playback system (e.g., a request to configure a media playback system with an additional playback device).

**[0072]** In some cases, the first recording device (or another device) instructs the one or more playback devices to emit the calibration sound. For instance, a recording device, such as control device 126 of media playback system 100, may send a command that causes a playback device (e.g., one of playback devices 102-124) to emit a calibration sound. The control device may send the command via a network interface (e.g., a wired or wireless network interface). A playback device may receive such a command, perhaps via a network interface, and responsively emit the calibration sound.

**[0073]** Acoustics of an environment may vary from location to location within the environment. Because of this variation, some calibration procedures may be improved by positioning the playback device to be calibrated within the environment in the same way that the playback device will later be operated. In that position, the environment may affect the calibration sound emitted by a playback device in a similar manner as playback will be affected by the environment during operation.

**[0074]** Further, some example calibration procedures may involve one or more recording devices detecting the calibration sound at multiple physical locations within the environment, which may further assist in capturing acoustic variability within the environment. To facilitate

detecting the calibration sound at multiple points within an environment, some calibration procedures involve a moving microphone. For example, a microphone that is detecting the calibration sound may be moved through the environment while the calibration sound is emitted. Such movement may facilitate detecting the calibration sounds at multiple physical locations within the environment, which may provide a better understanding of the environment as a whole.

**[0075]** In some embodiments, the one or more playback devices may repeatedly emit the calibration sound during the calibration procedure such that the calibration sound covers the calibration frequency range during each repetition. With a moving microphone, repetitions of the calibration sound are detected at different physical locations within the environment, thereby providing samples that are spaced throughout the environment. In some cases, the calibration sound may be periodic calibration signal in which each period covers the calibration frequency range.

**[0076]** To facilitate determining a frequency response, the calibration sound should be emitted with sufficient energy at each frequency to overcome background noise. To increase the energy at a given frequency, a tone at that frequency may be emitted for a longer duration. However, by lengthening the period of the calibration sound, the spatial resolution of the calibration procedure is decreased, as the moving microphone moves further during each period (assuming a relatively constant velocity). As another technique to increase the energy at a given frequency, a playback device may increase the intensity of the tone. However, in some cases, attempting to emit sufficient energy in a short amount of time may damage speaker drivers of the playback device.

**[0077]** Some implementations may balance these considerations by instructing the playback device to emit a calibration sound having a period that is approximately 3/8th of a second in duration (e.g., in the range of 1/4 to 1 second in duration). In other words, the calibration sound may repeat at a frequency of 2-4 Hz. Such a duration may be long enough to provide a tone of sufficient energy at each frequency to overcome background noise in a typical environment (e.g., a quiet room) but also be short enough that spatial resolution is kept in an acceptable range (e.g., less than a few feet assuming normal walking speed).

**[0078]** In some embodiments, the one or more playback devices may emit a hybrid calibration sound that combines a first component and a second component having respective waveforms. For instance, an example hybrid calibration sound might include a first component that includes noises at certain frequencies and a second component that sweeps through other frequencies (e.g., a swept-sine). A noise component may cover relatively low frequencies of the calibration frequency range (e.g., 10-50 Hz) while the swept signal component covers higher frequencies of that range (e.g., above 50 Hz). Such a hybrid calibration sound may combine the advantages

of its component signals.

**[0079]** A swept signal (e.g., a chirp or swept sine) is a waveform in which the frequency increases or decreases with time. Including such a waveform as a component of a hybrid calibration sound may facilitate covering a calibration frequency range, as a swept signal can be chosen that increases or decreases through the calibration frequency range (or a portion thereof). For example, a chirp emits each frequency within the chirp for a relatively short time period such that a chirp can more efficiently cover a calibration range relative to some other waveforms. Figure 8 shows a graph 800 that illustrates an example chirp. As shown in Figure 8, the frequency of the waveform increases over time (plotted on the X-axis) and a tone is emitted at each frequency for a relatively short period of time.

**[0080]** However, because each frequency within the chirp is emitted for a relatively short duration of time, the amplitude (or sound intensity) of the chirp must be relatively high at low frequencies to overcome typical background noise. Some speakers might not be capable of outputting such high intensity tones without risking damage. Further, such high intensity tones might be unpleasant to humans within audible range of the playback device, as might be expected during a calibration procedure that involves a moving microphone. Accordingly, some embodiments of the calibration sound might not include a chirp that extends to relatively low frequencies (e.g., below 50 Hz). Instead, the chirp or swept signal may cover frequencies between a relatively low threshold frequency (e.g., a frequency around 50-100 Hz) and a maximum of the calibration frequency range. The maximum of the calibration range may correspond to the physical capabilities of the channel(s) emitting the calibration sound, which might be 20,000 Hz or above.

**[0081]** A swept signal might also facilitate the reversal of phase distortion caused by the moving microphone. As noted above, a moving microphone causes phase distortion, which may interfere with determining a frequency response from a detected calibration sound. However, with a swept signal, the phase of each frequency is predictable (as Doppler shift). This predictability facilitates reversing the phase distortion so that a detected calibration sound can be correlated to an emitted calibration sound during analysis. Such a correlation can be used to determine the effect of the environment on the calibration sound.

**[0082]** As noted above, a swept signal may increase or decrease frequency over time. In some embodiments, the recording device may instruct the one or more playback devices to emit a chirp that descends from the maximum of the calibration range (or above) to the threshold frequency (or below). A descending chirp may be more pleasant to hear to some listeners than an ascending chirp, due to the physical shape of the human ear canal. While some implementations may use a descending swept signal, an ascending swept signal may also be effective for calibration.

**[0083]** As noted above, example calibration sounds may include a noise component in addition to a swept signal component. Noise refers to a random signal, which is in some cases filtered to have equal energy per octave. In embodiments where the noise component is periodic, the noise component of a hybrid calibration sound might be considered to be pseudorandom. The noise component of the calibration sound may be emitted for substantially the entire period or repetition of the calibration sound. This causes each frequency covered by the noise component to be emitted for a longer duration, which decreases the signal intensity typically required to overcome background noise.

**[0084]** Moreover, the noise component may cover a smaller frequency range than the chirp component, which may increase the sound energy at each frequency within the range. As noted above, a noise component might cover frequencies between a minimum of the frequency range and a threshold frequency, which might be, for example around a frequency around 50-100 Hz. As with the maximum of the calibration range, the minimum of the calibration range may correspond to the physical capabilities of the channel(s) emitting the calibration sound, which might be 20 Hz or below.

**[0085]** Figure 9 shows a graph 900 that illustrates an example brown noise. Brown noise is a type of noise that is based on Brownian motion. In some cases, the playback device may emit a calibration sound that includes a brown noise in its noise component. Brown noise has a "soft" quality, similar to a waterfall or heavy rainfall, which may be considered pleasant to some listeners. While some embodiments may implement a noise component using brown noise, other embodiments may implement the noise component using other types of noise, such as pink noise or white noise. As shown in Figure 9, the intensity of the example brown noise decreases by 6 dB per octave (20 dB per decade).

**[0086]** Some implementations of a hybrid calibration sound may include a transition frequency range in which the noise component and the swept component overlap. As indicated above, in some examples, the control device may instruct the playback device to emit a calibration sound that includes a first component (e.g., a noise component) and a second component (e.g., a sweep signal component). The first component may include noise at frequencies between a minimum of the calibration frequency range and a first threshold frequency, and the second component may sweep through frequencies between a second threshold frequency and a maximum of the calibration frequency range.

**[0087]** To overlap these signals, the second threshold frequency may be a lower frequency than the first threshold frequency. In such a configuration, the transition frequency range includes frequencies between the second threshold frequency and the first threshold frequency, which might be, for example, 50-100 Hz. By overlapping these components, the playback device may avoid emitting a possibly unpleasant sound associated with a harsh

transition between the two types of sounds.

**[0088]** Figures 10A and 10B illustrate components of example hybrid calibration signals that cover a calibration frequency range 1000. Figure 10A illustrates a first component 1002A (i.e., a noise component) and a second component 1004A of an example calibration sound. Component 1002A covers frequencies from a minimum 1008A of the calibration range 1000 to a first threshold frequency 1008A. Component 1004A covers frequencies from a second threshold 1010A to a maximum of the calibration frequency range 1000. As shown, the threshold frequency 1008A and the threshold frequency 1010A are the same frequency.

**[0089]** Figure 10B illustrates a first component 1002B (i.e., a noise component) and a second component 1004B of another example calibration sound. Component 1002B covers frequencies from a minimum 1008B of the calibration range 1000 to a first threshold frequency 1008A. Component 1004A covers frequencies from a second threshold 1010B to a maximum 1012B of the calibration frequency range 1000. As shown, the threshold frequency 1010B is a lower frequency than threshold frequency 1008B such that component 1002B and component 1004B overlap in a transition frequency range that extends from threshold frequency 1010B to threshold frequency 1008B.

**[0090]** Figure 11 illustrates one example iteration (e.g., a period or cycle) of an example hybrid calibration sound that is represented as a frame 1100. The frame 1100 includes a swept signal component 1102 and noise component 1104. The swept signal component 1102 is shown as a downward sloping line to illustrate a swept signal that descends through frequencies of the calibration range. The noise component 1104 is shown as a region to illustrate low-frequency noise throughout the frame 1100. As shown, the swept signal component 1102 and the noise component overlap in a transition frequency range. The period 1106 of the calibration sound is approximately 3/8ths of a second (e.g., in a range of 1/4 to 1/2 second), which in some implementation is sufficient time to cover the calibration frequency range of a single channel.

**[0091]** Figure 12 illustrates an example periodic calibration sound 1200. Five iterations (e.g., periods) of hybrid calibration sound 1100 are represented as frames 1202, 1204, 1206, 1208, and 1210. In each iteration, or frame, the periodic calibration sound 1200 covers a calibration frequency range using two components (e.g., a noise component and a swept signal component).

**[0092]** In some embodiments, a spectral adjustment may be applied to the calibration sound to give the calibration sound a desired shape, or roll off, which may avoid overloading speaker drivers. For instance, the calibration sound may be filtered to roll off at 3 dB per octave, or  $1/f$ . Such a spectral adjustment might not be applied to vary low frequencies to prevent overloading the speaker drivers.

**[0093]** In some embodiments, the calibration sound

may be pre-generated. Such a pre-generated calibration sound might be stored on the control device, the playback device, or on a server (e.g., a server that provides a cloud service to the media playback system). In some cases, the control device or server may send the pre-generated calibration sound to the playback device via a network interface, which the playback device may retrieve via a network interface of its own. Alternatively, a control device may send the playback device an indication of a source of the calibration sound (e.g., a URI), which the playback device may use to obtain the calibration sound.

**[0094]** Alternatively, the control device or the playback device may generate the calibration sound. For instance, for a given calibration range, the control device may generate noise that covers at least frequencies between a minimum of the calibration frequency range and a first threshold frequency and a swept sine that covers at least frequencies between a second threshold frequency and a maximum of the calibration frequency range. The control device may combine the swept sine and the noise into the periodic calibration sound by applying a crossover filter function. The cross-over filter function may combine a portion of the generated noise that includes frequencies below the first threshold frequency and a portion of the generated swept sine that includes frequencies above the second threshold frequency to obtain the desired calibration sound. The device generating the calibration sound may have an analog circuit and/or digital signal processor to generate and/or combine the components of the hybrid calibration sound.

**[0095]** Further example calibration procedures are described in U.S. Patent Application No. 14/805,140 filed July 21, 2015, entitled "Hybrid Test Tone For Space-Averaged Room Audio Calibration Using A Moving Microphone," U.S. Patent Application No. 14/805,340 filed July 21, 2015, entitled "Concurrent Multi-Loudspeaker Calibration with a Single Measurement," and U.S. Patent Application No. 14/864,393 filed September 24, 2015, entitled "Facilitating Calibration of an Audio Playback Device."

**[0096]** Calibration may be facilitated via one or more control interfaces, as displayed by one or more devices. Example interfaces are described in U.S. Patent Application No. 14/696,014 filed April 24, 2015, entitled "Speaker Calibration," and U.S. Patent Application No. 14/826,873 filed August 14, 2015, entitled "Speaker Calibration User Interface."

**[0097]** Moving now to several example implementations, implementations 1300, 1900, and 2000 shown in Figures 13, 19 and 20, respectively present example embodiments of techniques described herein. These example embodiments that can be implemented within an operating environment including, for example, the media playback system 100 of Figure 1, one or more of the playback device 200 of Figure 2, or one or more of the control device 300 of Figure 3, as well as other devices described herein and/or other suitable devices. Further, operations illustrated by way of example as being per-

formed by a media playback system can be performed by any suitable device, such as a playback device or a control device of a media playback system, implementations 1300, 1900, and 2000 may include one or more operations, functions, or actions as illustrated by one or more of blocks shown in Figures 13, 19, and 20. Although the blocks are illustrated in sequential order, these blocks may also be performed in parallel, and/or in a different order than those described herein. Also, the various blocks may be combined into fewer blocks, divided into additional blocks, and/or removed based upon the desired implementation.

**[0098]** In addition, for the implementations disclosed herein, the flowcharts show functionality and operation of one possible implementation of present embodiments. In this regard, each block may represent a module, a segment, or a portion of program code, which includes one or more instructions executable by a processor for implementing specific logical functions or steps in the process. The program code may be stored on any type of computer readable medium, for example, such as a storage device including a disk or hard drive. The computer readable medium may include non-transitory computer readable medium, for example, such as computer-readable media that stores data for short periods of time like register memory, processor cache, and Random Access Memory (RAM). The computer readable medium may also include non-transitory media, such as secondary or persistent long term storage, like read only memory (ROM), optical or magnetic disks, compact-disc read only memory (CD-ROM), for example. The computer readable media may also be any other volatile or non-volatile storage systems. The computer readable medium may be considered a computer readable storage medium, for example, or a tangible storage device. In addition, for the implementations disclosed herein, each block may represent circuitry that is wired to perform the specific logical functions in the process.

### III. Example Techniques To Facilitate Spatial Calibration

**[0099]** As discussed above, embodiments described herein may facilitate the calibration of one or more playback devices by determining a spatial calibration. Figure 13 illustrates an example implementation 1300 by which a media playback system facilitates such a calibration.

#### a. Detect Trigger Condition

**[0100]** At block 1302, implementation 1300 involves detecting a trigger condition. For instance, a networked microphone device detects a trigger condition that initiates calibration of a media playback system (or perhaps a set of playback devices in a media playback system). Example networked microphone devices include any suitable device that includes a network interface and a microphone. For instance, playback devices (e.g., play-

back device 200) and control devices (e.g., control device 300) may each operate as a networked microphone device. Other example networked microphone devices include control devices 126 and 128 of Figure 1.

**[0101]** The trigger condition initiates calibration of multiple audio drivers. In some cases, the multiple audio drivers may be housed in a single playback device. For instance, a soundbar-type playback device may include multiple audio drivers (e.g., nine audio drivers). In other cases, the multiple audio drivers may be divided among two or more playback devices. For example, a soundbar with multiple audio drivers may be calibrated with one or more other playback devices each with one or more respective audio drivers. Some example playback devices include multiple audio drivers of different types (e.g., tweeters and woofers, perhaps of varying size).

**[0102]** The particular playback devices (and audio drivers) under calibration may correspond to zones of a media playback system. For instance, an example trigger condition may initiate calibration of a given zone of a media playback system (e.g., the Living Room zone of media playback system 100 shown in Figure 1). According to this example, the Living Room zone includes playback devices 104, 106, 108, and 110 that together include multiple audio drivers, and the example trigger condition may therefore initiate calibration of multiple audio drivers.

**[0103]** As noted above in connection with the example calibration sequence, various trigger conditions are contemplated herein. Some example triggers conditions include input data instructing the media playback system to initiate calibration. Such input data may be received via a user interface (e.g., control interface 600 of Figure 6) of a networked microphone device, as illustrated in Figure 6 or perhaps via another device that relays the instruction to the networked microphone device and/or the playback devices under calibration.

**[0104]** Other example trigger conditions might be based on sensor data. For instance, sensor data from an accelerometer or other suitable sensor may indicate that a given playback device has moved, which may prompt calibration of that playback device (and perhaps other playback devices associated with the given playback device, such as those in a bonded zone or zone group with the playback device).

**[0105]** Some trigger conditions may involve a combination of input data and sensor data. For instance, sensor data may indicate a change in the operating environment of a media playback system, which may cause a prompt to initiate calibration to be displayed on a networked microphone device. The media playback system might proceed with calibration after receiving input data at the prompt indicating confirmation to initiate calibration.

**[0106]** Further example trigger conditions may be based on changes in configuration of a media playback system. For instance, example trigger conditions include addition or removal of a playback device from a media playback system (or grouping thereof). Other example trigger conditions include receiving new types of input

content (e.g., receiving multi-channel audio content).

**[0107]** In operation, multiple audio drivers form one or more sound axes. For instance, two playback devices each with a respective audio driver may form respective sound axes. In some cases, two or more audio drivers may be arrayed to form a sound axis. For example, a playback device with multiple audio drivers (e.g., a soundbar with nine audio drivers) may form multiple sound axes (e.g., three sound axes). Any audio driver may contribute to any number of sound axes. For example, a given sound axis may be formed by contributions from all nine audio drivers of a soundbar.

**[0108]** Each sound axis corresponds to a respective input channel of audio content. For instance, audio drivers of a media playback system may form two sound axes corresponding, respectively, to left and right channels of stereo content. As another example, the audio drivers may form sound axes corresponding to respective channels of surround sound content (e.g., center, front left, front right, rear left, and rear right channels).

**[0109]** Arraying two or more audio drivers to form a given sound axis enables the two or more audio drivers to "direct" the sound output for the given sound axis in a certain direction. For instance, where nine audio drivers of a soundbar are each contributing a portion of a sound axis corresponding to a left channel of surround sound content, the nine audio drivers may be arrayed (i.e., acoustically summed, perhaps using a DSP) in such a way that the net polar response of the nine audio drivers directs sound to the left. Concurrently with the sound axis corresponding to the left channel, the nine audio drivers also form sound axes corresponding to center and right channels of the surround sound content to direct sound to the center and to the right, respectively.

**[0110]** A particular set of sound axes formed by playback devices of a media playback system may be referred to as a playback configuration. In operation, playback devices of a media playback system may be configured into a given one of multiple possible playback configurations. While in a given playback configuration, the audio drivers of the playback devices may form a particular set of sound axes. In some cases, configuration of playback devices into a new playback configuration may act as a trigger condition to initiate calibration of the playback devices.

**[0111]** To illustrate, referring back to Figure 1, playback devices 104, 106, 108, and 110 of the Living Room zone may be configurable into multiple playback configurations. In a first playback configuration, perhaps associated with surround sound audio content, playback device 104 may form one or more sound axes (e.g., front, left, and right channels) while playback devices 108 and 110 form respective sound axes (e.g., left and right surround channels). Playback device 110, being a subwoofer-type device, may contribute a separate low-frequency sound axis or a low-frequency portion of the sound axes formed by playback devices 104, 106, and/or 108. In another playback configuration, the audio drivers of playback de-

vices 104, 106, 108, and 110 may combine to form sound axes corresponding to left and right channels of stereo audio content. Another playback configuration may involve the audio drivers forming a single sound axis corresponding to mono audio content.

**[0112]** In operation, playback devices may utilize a given playback configuration according to various factors. Such factors may include the zone configuration (e.g., whether the playback devices are in a 5.1, 5.0, or other surround sound configuration, a stereo pair configuration, a play bar-only configuration, among others). The factors may also include the specific types and capabilities of the playback devices. The factors may further include the specific type of content provided to the playback devices (or expected to be provided). For instance, playback devices may adopt a first playback configuration when playing surround sound content and another when playing stereo content. As another example, playback devices may use a given playback configuration when playing music and another when playing audio that is paired with video (e.g., television content). Further example playback configurations include any of the above example configurations with (or without) a subwoofer-type playback device, as addition (or subtraction) of such a device from the playback configuration may change the acoustic characteristics and/or allocation of playback responsibilities in the playback configuration.

**[0113]** Calibration sequences involve calibrating playback devices for multiple playback configurations. Such calibration sequences may yield multiple calibration profiles that are applied the playback devices are in a given playback configuration. For instance, a given calibration procedure may calibrate the Living Room zone of media playback system 100 for a surround sound playback configuration and a music playback configuration. While in the surround sound playback configuration, the playback devices of the Living Room zone may apply a first calibration profile (e.g., one or more filters that adjusts one or more of magnitude response, frequency response, phase, etc.) corresponding to the surround sound playback configuration. Likewise, while in the music playback configuration, the playback devices of the Living Room zone may apply a second calibration profile corresponding to the music playback configuration.

#### b. Cause Multiple Audio Drivers To Emit Calibration Audio

**[0114]** In Figure 13, at block 1304, implementation 1300 involves causing the multiple audio drivers to emit calibration audio. For instance, the NMD instructs the playback device(s) that include the multiple audio drivers to emit calibration audio via the multiple audio drivers. For instance, control device 126 of media playback system 100 may send a command that causes a playback device (e.g., one of playback devices 102-124) to emit calibration audio. The NMD sends the command via a network interface (e.g., a wired or wireless network in-

terface). A playback device may receive such a command, perhaps via a network interface, and responsively emit the calibration audio.

**[0115]** The calibration audio may include one or more calibration sounds, such as a frequency sweep ("chirp"), brown noise or other types of noise, or a song, among other example sounds. Additional details on example calibration sounds are noted above in connection with the example calibration sequence described in section II. e, as well as generally throughout the disclosure.

**[0116]** The calibration audio is divided into frames. As illustrated in Figures 11 and 12 and described herein, frames may represent iterations (e.g., a period or cycle) of an example calibration sound. When recorded, frames may produce respective samples of the calibration sound as emitted by one or more audio drivers.

**[0117]** As noted above, a calibration sequence involves calibration of multiple sound axes. Example calibration audio to calibrate multiple sound axes may be divided into constituent frames, wherein each frame includes calibration audio for every sound axis under calibration. Accordingly, when recorded, each frame may include samples of the calibration audio produced by each sound axis. The frames may repeat to produce multiple samples for each sound axis.

**[0118]** To include the calibration audio for every sound axis under calibration, each frame is further divided into slots. Each slot includes the calibration audio for a respective sound axis under calibration. For instance, an example frame for a playbar-type playback device (e.g., playback device 104 shown in Figure 1) that forms three sound axes (such as left, right, and center channels) may include three slots. If, for example, that device was to be calibrated with a subwoofer-type device, each frame might include four slots, one for each sound axis formed by the playbar-type playback device and one for the sound axis produced by the subwoofer. As a further example, where the playbar-type playback device is calibrated with two additional playback devices that produce respective sound axes (e.g., rear left and right channels), each frame may include five slots (or six slots if calibrated with a subwoofer).

**[0119]** As noted above, each slot includes the calibration audio for a respective sound axis under calibration. The calibration audio in each slot may include a frequency sweep ("chirp"), brown noise or other types of noise, among other examples. For instance, referring back to Figures 11 and 12, the calibration audio in each sound may include a hybrid calibration sound. Slots may occur sequentially in a known order, so as to facilitate matching slots within recorded calibration audio to respective sound axes. Each slot may have a known duration, which may also facilitate matching slots within recorded calibration audio to respective sound axis. In further examples, each slot and/or frame may include a watermark (e.g., a particular pattern of sound) to identify the slot or frame, which may be used to match slots within recorded calibration audio to respective sound axes.

**[0120]** To illustrate, Figure 14 shows an example calibration audio 1400. Calibration sound 1400 includes frames 1402, 1404, and 1406. Frames 1402, 1404, and 1406 are divided into respective three respective slots. In particular, frame 1402 includes slots 1402A, 1402B and 1402C. Likewise, frames 1404 and 1406 include slots 1404A, 1404B, & 1404C and 1406A, 1406B, & 1406C, respectively. Each slot includes an iteration of hybrid calibration sound 1 100 of Figure 1 1. During a calibration procedure, the calibration sound in each slot may be emitted by a respective sound axis (perhaps formed via multiple audio drivers). For instance, slots 1402A, 1404A, and 1406A may correspond to a first sound axis (e.g., a left channel) while slots 1402B, 1404B, and 1406B correspond to a second sound axis (and slots 1402C, 1404C, and 1406C to a third sound axis). In such manner, when recorded, calibration audio 1400 may produce three samples of each sound axis, provided that a sufficient portion of frames 1402, 1404 and 1406 are recorded.

**[0121]** As noted above, the playback devices of a media playback system are calibrated for multiple playback configurations. Calibration audio to calibrate multiple playback configurations may include a repeating series of frames. Each frame in a series corresponds to a respective playback configuration. For instance, example calibration audio to calibrate three playback configurations may include a series of three frames (e.g., frames 1402, 1404, and 1406 of Figure 14).

**[0122]** As shown in Figure 14, each frame in the series is divided into slots corresponding to the sound axes of the playback configuration corresponding to that frame. Since different playback configurations might form different sets of sound axes perhaps with different numbers of total axes, frames in a series may have different numbers of slots. The series of frames may repeat so as to produce multiple samples for each sound axis of each playback configuration,

#### c. Record Calibration Audio

**[0123]** In Figure 13, at block 1306, implementation 1300 involves recording the emitted calibration audio. For instance, an NMD records calibration audio as emitted by playback devices of a media playback system (e.g., media playback system 100) via a microphone. As noted above, example NMDs include control devices (e.g., control device 126 or 128 of Figure 1), playback devices, or any suitable device with a microphone or other sensor to record calibration audio. In some cases, multiple NMDs may record the calibration audio via respective microphones.

**[0124]** In practice, some of the calibration sound may be attenuated or drowned out by the environment or by other conditions, which may interfere with the recording device recording all of the calibration sound. As such, the NMD may measure a portion of the calibration sounds as emitted by playback devices of a media playback sys-

tem. The calibration audio may be any of the example calibration sounds described above with respect to the example calibration procedure, as well as any suitable calibration sound.

**[0125]** In some cases, the NMD(s) may remain more or less stationary while recording the calibration audio. For instance, the NMDs may be positioned at one or more particular locations (e.g., a preferred listening location). Such positioning may facilitate recording the calibration audio as would be perceived by a listener at that particular location.

**[0126]** Certain playback configurations may suggest particular preferred listening locations. For example, playback configurations corresponding to surround sound audio or audio that is coupled with video may suggest the location at which users will watch television while listening to the playback devices (e.g., on a couch or chair). In some examples, an NMD may prompt to move to a particular location (e.g., a preferred listening location) to begin the calibration. When calibration multiple playback configurations, the NMD may prompt to move to certain listening locations corresponding to each playback configurations.

**[0127]** To illustrate such prompts, in Figure 15, smartphone 500 is displaying control interface 1500 which includes graphical region 1502. Graphical region 1502 prompts to move to a particular location (i.e., where the user will usually watch TV in the room). Such a prompt may be displayed to guide a user to begin the calibration sequence in a preferred location. Control interface 1500 also includes selectable controls 1504 and 1506, which respectively advance and step backward in the calibration sequence.

**[0128]** Figure 16 depicts smartphone 500 displaying control interface 1600 which includes graphical region 1602. Graphical region 1602 prompts the user to raise the recording device to eye level. Such a prompt may be displayed to guide a user to position the phone in a position that facilitates measurement of the calibration audio. Control interface 1600 also includes selectable controls 1604 and 1606, which respectively advance and step backward in the calibration sequence.

**[0129]** Next, Figure 17 depicts smartphone 500 displaying control interface 1700 which includes graphical region 1702. Graphical region 1702 prompts the user to "set the sweet spot." (i.e., a preferred location within the environment). After smartphone 500 detects selection of selectable control 1704, smartphone 500 may begin measuring the calibration sound at its current location (and perhaps also instruct one or more playback devices to output the calibration audio). As shown, control interface 1700 also includes selectable control 1706, which advances the calibration sequence (e.g., by causing smartphone to begin measuring the calibration sound at its current location, as with selectable control 1704).

**[0130]** In Figure 18, smartphone 500 is displaying control interface 1800 which includes graphical region 1802. Graphical region 1802 indicates that smartphone 500 is



recording the calibration audio. Control interface 1800 also includes selectable control 1804, which steps backwards in the calibration sequence.

d. Cause Processing of Recorded Calibration Audio

**[0131]** In Figure 13, at block 1308, implementation 1300 involves causing the recorded calibration audio to be processed. For instance, the NMD causes a processing device to process the recorded calibration audio. In some cases, the NMD may include the processing device. Alternatively, the NMD may transmit the recorded audio to one or more other processing devices for processing. Example processing devices include playback devices, control devices, a computing device connected to the media playback system via a local area network, a remote computing device such as a cloud server, or any combination of the above.

**[0132]** Processing of the calibration audio involves determining one or more calibrations for each of the one or more sound axes. Each calibration of the multiple sound axes may involve modifying one or more of magnitude response, frequency response, phase adjustment, or any other acoustic characteristic. Such modifications may spatially calibrate the multiple sound axes to one or more locations (e.g., one or more preferred listening locations).

**[0133]** Such modifications may be applied using one or more filters implemented in a DSP or as analog filters. The calibration data may include the parameters to implement the filters (e.g., as the coefficients of a bi-quad filter). Filters may be applied per audio driver or per set of two or more drivers (e.g., two or more drivers that form a sound axis or two or more of the same type of audio driver, among other examples). In some cases, respective calibrations may be determined for the multiple playback configurations under calibration.

**[0134]** The recorded calibration audio may be processed as it is recorded or after recording is complete. For instance, where the calibration audio is divided into frames, the frames may be transmitted to the processing device as they are recorded, possibly in groups of frames. Alternatively, the recorded frames may be transmitted to the processing device after the playback devices finish emitting the calibration audio.

**[0135]** Processing may involve determining respective delays for each sound axis of the multiple sound axes. Ultimately, such delays may be used to align time-of-arrival of respective sound from each sound axis at a particular location (e.g., a preferred listening location). For instance, a calibration profile for a given playback configuration may include filters that delay certain sound axes of the playback configuration to align time-of-arrival of the sound axes of the playback configuration at a preferred listening location. Sound axes may have different times-of-arrival at a particular location because they are formed by audio drivers at different distances from the particular location. Further, some sound axes may be directed away from the particular location (e.g., left and

right channels of a soundbar-type playback device), and as such, reflect off of the environment before arriving at the particular location. Such a sound path may increase the effective distance between the audio drivers forming a sound axis and the particular location, which may cause a later time-of-arrival as compared to sound axes that have a more direct path. As noted above, such a preferred listening location might be a couch or chair for a surround sound playback configuration.

**[0136]** Within examples, the processing device may separate the recorded audio into parts corresponding to the different sound axes and/or playback configurations that emitted each part. For instance, where the calibration sound emitted by the playback devices was divided into frames, the processing device may divide the recorded audio back into the constituent frames. Where the calibration sound included a series of frames, the processing device may attribute the frames from each series to the respective playback configuration corresponding to those frames. Further, the processing device may divide each frame into respective slots corresponding to each sound axis. As noted above, the playback devices may emit frames and slots in a known sequence and each slot may have a known duration to facilitate dividing the recorded audio into its constituent parts. In some examples, each slot and/or frame may include a watermark to identify the slot or frame, which may be used to match frames within recorded calibration audio to respective playback configurations and/or slots to respective sound axes.

**[0137]** The processing device may determine an impulse response for each sound axis. Each impulse response may be further processed by generating frequency filtered responses so as to divide the impulse responses into frequency bands. Audio drivers of different types may array better at different frequency bands. For instance, mid-range woofers may array well to form a sound axis in a range from 300 Hz to 2.5 kHz. As another example, tweeters may array well in a range from 8 kHz to 14 kHz. Where an example sound axis is configured to form a center channel of a surround sound configuration, the sound axis should be maximum on-axis and attenuated to the right and left. Conversely, for sound axes forming left and right channels of the surround sound configuration, each array should be attenuated (e.g., NULL) on-axis and maximum to the left or right, respectively. Outside of the certain ranges such as those provided above, the audio drivers might not form the sound axis as well in the intended direction. Such frequency ranges are provided by way of example and may vary according to capabilities and characteristics of different audio drivers.

**[0138]** As a further example, in a playback device with multiple audio drivers of different types (e.g., tweeters and woofers), a processing device may determine three band-limited responses. Such responses might include a full-range response, a response covering a mid-range for woofers (e.g., 300 Hz to 2.5 kHz), and a response

covering high frequencies for the tweeters (e.g., 3 kHz to 14 kHz). Such frequency-filtered responses may facilitate further processing by more clearly representing each sound axis.

**[0139]** Processing the recorded audio may involve comparisons between the responses for each sound axis. To facilitate such comparisons, the impulse responses for each slot may be time-aligned with one another (as they were emitted during different periods of time). For instance, the impulse responses may be aligned to a first reference point, such as the beginning of each slot. Such time-alignment of the impulse responses facilitates identification of particular reference points in each response.

**[0140]** In an example implementation, identification of particular reference points in each response involves identifying a given second reference point in an impulse response of a reference sound axis. By way of example, the reference sound axis may be a sound axis corresponding to a center channel of a surround sound system (e.g., a 3.0, 3.1, 5.0, 5.1 or other multi-channel playback configuration). This sound axis may be used as the reference sound axis because sound from this axis travels more directly to typical preferred listening locations than other sound axes (e.g., sound axis that form left and right channels). The given second reference point in this impulse response may be the first peak value. The first peak can be assumed to correspond to the direct signal from the audio driver(s) to the NMD (rather than a reflection). This given second reference point (i.e., the first peak) is used as a reference for subsequent times-of-arrival of other sound axes at the NMD.

**[0141]** To compare times-of-arrival other sound axes at the NMD to the reference sound axis, the processing device may identify second reference points in the other impulse responses. These other second reference points correspond to the same second reference point as in the reference sound axis. For instance, if the first peak in the impulse response of the reference sound axis was used as the given second reference point, then the first peaks in the other impulse responses are identified as the second reference points.

**[0142]** Where the approximate physical configuration of the multiple audio drivers is known, a time window may be applied to limit the portion of each impulse response where the second reference points are to be identified. For instance, where the sound axes form left, right, and center channels, the impulse responses for the sound axes forming the left and right channels can be limited to a time window subsequent to the peak value in the impulse response for the sound axis forming the center channels. Sound from the sound axes forming the left and right channels travels outward to the left and right (rather than on-axis) and thus the peak value of interest will be a reflection of the sound from these axes off the environment. However, sound axes forming left and/or right surround channels and/or a subwoofer channel may have been physically closer to the NMD than the audio driver(s) forming the center channel. As such, a window

for impulse responses corresponding to those axes may encompass time before and after the given reference point in the reference sound axis so as to account for the possibility of either positive or negative delay relative to that reference sound axis.

**[0143]** Once respective second reference points in the impulse responses have been identified, the respective times-of-arrival of sound from each sound axis at the NMD (i.e., the microphone of the NMD) can be determined. In particular, the processing device may determine the respective times-of-arrival at the microphone by comparing respective differences from the first reference point to the second reference points in each impulse response.

**[0144]** With respective times-of-arrival of sound from each sound axis at the NMD determined, the processing device may determine respective delays to be applied for each sound axis. The processing device may determine the delays relative to a delay target. This delay target may be the sound axis that has the latest time-of-arrival. The sound axis acting as the delay target might not receive any delay. Other sound axes may be assigned a delay to match the time-of-arrival of the sound axis acting as the delay target. A sound axis that forms a center channel may not be used as the delay target in some instances because sound axes with later times-of-arrival cannot be assigned "negative" delay to match the time-of-arrival of the sound axis forming the center channel.

**[0145]** In some cases, the delay for any given sound axis may be capped at a maximum delay threshold. Such capping may prevent issues with large amounts of delay causing apparent mismatch between audio content output by the sound axes and video content that is coupled to that audio content (e.g., lip-sync issues). Such capping may be applied only to playback configurations that include audio paired with video, as large delays may not impact user experience when the audio is not paired with video. Alternatively, if the video display is synchronized with the playback device(s), then the video might be delayed to avoid apparent mismatch between audio content output by the sound axes and video content that is coupled to that audio content, which may eliminate the need for a maximum delay threshold.

**[0146]** As noted above, the NMD that recorded the calibration audio might not perform certain portions of the processing (or might not process the calibration audio at all). Rather, the NMD may transmit data representing the recorded calibration audio to a processing device, perhaps with one or more instructions on how to process the recorded calibration audio. In other cases, the processing device may be programmed to process recorded calibration audio using certain techniques. In such embodiments, transmitting data representing the recorded calibration audio (e.g., data representing raw samples of calibration audio and/or data representing partially processed calibration audio) may cause the processing device to determine calibration profiles (e.g.,

filter parameters).

#### e. Cause Calibration Of Multiple Sound Axes

**[0147]** In Figure 13, at block 1310, implementation 1300 involves causing calibration of the multiple sound axes. For instance, the NMD may send calibration data to the playback device(s) that form the multiple sound axes. Alternatively, the NMD may instruct another processing device to transmit calibration data to the playback device. Such calibration data may cause the playback device(s) to calibrate the multiple sound axes to a certain response.

**[0148]** As noted above, calibration of the multiple sound axes may involve modifying one or more of magnitude response, frequency response, phase adjustment, or any other acoustic characteristic. Such modifications may be applied using one or more filters implemented in a DSP or as analog filters. The calibration data may include the parameters to implement the filters (e.g., as the coefficients of a bi-quad filter). Filters may be applied per audio driver or per set of two or more drivers (e.g., two or more drivers that form a sound axis or two or more of the same type of audio driver, among other examples).

**[0149]** Calibrating the multiple sound axes may include causing audio output of the multiple sound axes to be delayed according to the respective determined delays for the sound axes. Such delays may be formed by causing respective filters to delay audio output of the multiple audio drivers according to the respective determined delays for the multiple sound axes. Such filters may implement a circular buffer delay line, among other examples.

**[0150]** In some cases, the delays are dynamic. For instance, the response of one axis may overlap with the response of another in a given range, yet the sound axes may have different times-of-arrival (thus suggesting different delays). In such situations, the delays of each sound axis may be smoothed across the overlapping range. For instance, a delay curve may be implemented across the range to smooth the delay. Such smoothing may improve user experience by avoiding possibly sharp differences in delay between sound axes in overlapping ranges.

**[0151]** As noted above, in some cases, sound produced by certain sound axes may reflect off of the environment before

#### IV. Example Techniques To Facilitate Spectral Calibration Using Applied Spatial Calibration

**[0152]** As noted above, embodiments described herein facilitate the calibration of one or more playback devices. Figure 19 illustrates an example implementation 1900 by which a playback device facilitates spectral calibration using applied spatial calibration.

#### a. Receive Data Representing Spatial Calibration(s)

**[0153]** At block 1902, implementation 1900 involves receiving data, representing one or spatial calibrations. For instance, a playback device (e.g., any playback device of media playback system 100 in Figure 1 or playback device 300 in Figure 3) may receive data representing one or more spatial calibrations (e.g., any of the multiple calibrations described above in connection with implementation 1300 of Figure 13) via a network interface from a device such as a processing device or a NMD, among other possible sources. Each calibration may have been previously determined by way of a calibration sequence, such as the example calibration sequences described above.

**[0154]** A calibration includes one or more filters. Such filters may modify one or more of magnitude response, frequency response, phase adjustment, or any other acoustic characteristic. Further, such filters may calibrate the playback device(s) under calibration to one or more particular listening locations within a listening area. As noted above, the filters may be implemented in a DSP (e.g., as the coefficients of a bi-quad filter) or as an analog filter, or a combination thereof. The received calibration data may include a filter for each audio channel, axis or device under calibration. Alternatively, a filter may be applied to more than one audio channel, axis or device.

**[0155]** In some cases, multiple calibrations may correspond to respective playback configurations. As noted above, a playback configuration refers to a specific set of sound axes formed by multiple audio drivers. Further, a spatial calibration includes calibration of audio drivers in multiple playback configurations. As such, there may be more than one filter (or set of filters) for each audio channel, axis or device. Each filter (or set of filters) corresponds to a different playback configuration.

**[0156]** As noted above, playback configurations may involve variance in the assignment of audio drivers to form sound axes. Each sound axis in a playback configuration may correspond to a respective input channel of audio content. Example playback configurations may correspond to difference numbers of input channels, such as mono, stereo, surround (e.g., 3.0, 5.0, 7.0) or any of the above in combination with a subwoofer (e.g., 3.1, 5.1, 7.1). Other playback configurations may be based on input content type. For instance, example playback configurations may correspond to input audio content including music, home theater (i.e., audio paired with video), surround sound audio content, spoken word, etc. These example playback configurations should not be considered limiting. The received calibrations may include filter(s) corresponding to any individual playback configuration or any combination of playback configurations.

**[0157]** The playback device may maintain these calibrations in data storage. Alternatively, such calibrations may be maintained on a device or system that is communicatively coupled to the playback device via a net-

work. The playback device may receive the calibrations from this device or system, perhaps upon request from the playback device.

#### b. Cause Audio Driver(s) To Output Calibration Audio

**[0158]** In Figure 19, at block 1904, implementation 1900 involves causing the audio driver(s) to output calibration audio. For instance the playback device may cause an audio stage to drive the audio drivers to output calibration audio. An example audio stage may include amplifier(s), signal processing (e.g., a DSP), as well as other possible components. In some cases, the playback device may instruct other playback devices under calibration to output calibration audio, perhaps when acting as a group coordinator for the playback devices under calibration.

**[0159]** The calibration audio may include one or more calibration sounds, such as a frequency sweep ("chirp"), brown noise or other types of noise, or a song, among other examples. Additional details on example calibration sounds are noted above in connection with the example calibration sequences described above.

**[0160]** The calibration audio is divided into frames. As illustrated in Figures 11 and 12 and described herein, frames may represent iterations of an example calibration sound. When recorded, frames may produce respective samples of the calibration sound as emitted by one or more audio drivers. The frames may repeat to produce multiple samples.

**[0161]** As noted above, a calibration sequence involves calibration of multiple sound axes. In such cases, the calibration audio output is divided into constituent frames, wherein each frame includes calibration audio for every sound axis under calibration. Accordingly, when recorded, each frame may include samples of the calibration audio produced by each sound axis. The frames may repeat to produce multiple samples for each sound axis.

**[0162]** As noted above, in calibration procedures, the playback devices of a media playback system is calibrated for multiple playback configurations. Alternatively, different playback configurations for a set of audio drivers may be calibrated in respective calibration sequences. Example calibration audio to calibrate multiple playback configurations may include a repeating set of frames. Each frame in a set corresponds to a respective playback configuration. For instance, example calibration audio to calibrate three playback configurations may include a series of three frames (e.g., frames 1402, 1404, and 1406 of Figure 14).

**[0163]** During each frame, the playback device may apply a spatial calibration corresponding to a respective playback configuration. Applying a spatial calibration involves causing an audio stage (or multiple audio stages) to apply respective filter(s) corresponding to each playback configuration. As the input signal passes through the filter(s), the calibration is applied to modify one or

more of magnitude response, frequency response, phase adjustment, or any other acoustic characteristic of the audio driver(s) as the calibration audio is emitted. As noted above, such filters may modify the emitted calibration audio to suit a particular listening location. For instance, example spatial filters may at least partially balance time-of-arrival of sound from multiple sound axes at the particular listening location.

**[0164]** In other embodiments, the spatial calibration may be applied to calibration audio by a device other than the playback device. A spatial calibration may be applied by any device that stores and/or generates the calibration audio for output by the audio drivers using a processor or DSP of that device. Yet further, a spatial calibration may be applied by any intermediary device between the device that stores the calibration audio and the playback device(s) under calibration.

**[0165]** To include the calibration audio for every sound axis under calibration, each frame is further divided into slots. Each slot includes the calibration audio for a respective sound axis under calibration. For instance, an example frame for a playbar-type playback device (e.g., playback device 104 shown in Figure 1) that forms three sound axes (such as left, right, and center channels) may include three slots. If, for example, that device was to be calibrated with a subwoofer-type device, each frame might include four slots, one for each sound axis formed by the playbar-type playback device and one for the sound axis produced by the subwoofer. As a further example, where the playbar-type playback device is calibrated with two additional playback devices that produce respective sound axes (e.g., surround left and right channels), each frame may include five slots (or six slots if calibrated with a subwoofer). Figure 14 illustrates example calibration audio with constituent frames that are divided into slots.

**[0166]** As noted above, each slot includes the calibration audio for a respective sound axis under calibration. The calibration audio in each slot may include a frequency sweep ("chirp"), brown noise or other types of noise, among other examples. For instance, as shown in Figures 11 and 12, the calibration audio in each sound may include a hybrid calibration sound. Slots may occur sequentially in a known order, so as to facilitate matching slots within recorded calibration audio to respective sound axes. Each slot may have a known duration, which may also facilitate matching slots within recorded calibration audio to respective sound axis. In further examples, each slot and/or frame may include a watermark (e.g., a particular pattern of sound) to identify the slot or frame, which may be used to match slots within recorded calibration audio to respective sound axes.

#### c. Receive Data Representing Spectral Calibration(s)

**[0167]** In Figure 19, at block 1906, implementation 1900 involves receiving data representing one or spectral calibrations. For instance, the playback device may re-

ceive data representing one or more spectral calibrations from a processing device. These spectral calibrations are based on the calibration audio output by the audio driver(s). In particular, the calibration audio output from the audio driver(s) is recorded by one or more recording devices (e.g., an NMD). Before being recorded, the calibration audio may be interact (e.g., be reflected or absorbed) by the surrounding environment and thereby represent characteristics of the environment.

**[0168]** Example spectral calibrations may offset acoustics characteristics of the environment to achieve a given response (e.g., a flat response, a response that is considered desirable, or a set equalization). For instance, if a given environment attenuates frequencies around 500 Hz and amplifies frequencies around 14000 Hz, a calibration might boost frequencies around 500 Hz and cut frequencies around 14000 Hz so as to offset these environmental effects.

**[0169]** Some example techniques for determining a calibration are described in U.S. Patent Application No. 13/536,493 filed June 28, 2012, entitled "System and Method for Device Playback Calibration" and published as US 2014/0003625 A1. Example techniques are described in paragraphs [0019]-[0025] and [0068]-[01 18] as well as generally throughout the specification.

**[0170]** Further example techniques for determining a calibration are described in U.S. Patent Application No. 14/216,306 filed March 17, 2014, entitled "Audio Settings Based On Environment" and published as US 2015/0263692 A1. Example techniques are described in paragraphs [0014]-[0025] and [0063]-[01 14] as well as generally throughout the specification.

**[0171]** Additional example techniques for determining a calibration are described in U.S. Patent Application No. 14/481,511 filed September 9, 2014, entitled "Playback Device Calibration" and published as US 2016/0014534 A1. Example techniques are described in paragraphs [0017]-[0043] and [0082]-[0184] as well as generally throughout the specification.

**[0172]** Example processing devices include NMDs, other playback devices, control devices, a computing device connected to the media playback system via a local area network, a remote computing device such as a cloud server, or any combination of the above. In some cases, the processing device(s) may transmit the spatial calibrations to one or more intermediary devices which may transmit the spatial calibrations to the playback device. Such intermediary devices may store the data representing one or spatial calibrations.

#### d. Apply Particular Spectral Filter

**[0173]** At block 1908, implementation 1900 involves applying a particular spectral calibration. For instance, the playback device may apply a particular filter corresponding to a given playback configuration when playing back audio content in that playback configuration. The playback device may maintain or have access to respec-

tive spectral calibrations corresponding to multiple playback configurations.

**[0174]** In some examples, the playback device may be instructed to enter a particular playback configuration and accordingly apply a particular calibration corresponding to that playback configuration. For instance, a control device may transmit a command to form a specific set of sound axes corresponding to a given playback configuration.

**[0175]** Alternatively, the playback device may detect the proper spectral calibration to apply based on its current configuration. As noted above, playback devices may be joined into various groupings, such as a zone group or bonded zone. Each grouping may represent a playback configuration. In some implementations, upon being joined into a grouping with another playback device, the playback device may apply a particular calibration associated with the playback configuration of that grouping. For instance, based on detecting that the playback device has joined a particular zone group, the playback device may apply a certain calibration associated with zone groups (or with the particular zone group).

**[0176]** The playback device may detect the spectral calibration to apply based the audio content being provided to the playback device (or that it has been instructed to play back). For instance, the playback device may detect that it is playing back media content that consists of only audio (e.g., music). In such cases, the playback device may apply a particular calibration associated with a playback configuration that corresponds to music playback. As another example, the playback device may receive media content that is associated with both audio and video (e.g., a television show or movie). When playing back such content, the playback device may apply a particular calibration corresponding to audio that is paired with video, or perhaps a calibration corresponding to home theater (e.g., surround sound).

**[0177]** The playback device may apply a certain calibration based on the source of the audio content. Receiving content via a particular one of these sources may trigger a particular playback configuration. For instance, receiving content via a network interface may indicate music playback. As such, while receiving content via the network interface, the playback device may apply a particular calibration associated with a particular playback configuration corresponding to music playback. As another example, receiving content via a particular physical input may indicate home theater use (i.e., playback of audio from a television show or movie). While playing back content from that input, the playback device may apply a different calibration associated with a playback configuration corresponding to home theater playback.

**[0178]** A given zone scene may be associated with a particular playback configuration. Upon entering a particular zone scene and thus the particular playback configuration, the playback device may apply a particular calibration associated with that playback configuration. Alternatively, the content or configuration associated with

a zone scene may cause the playback device to apply a particular calibration. For example, a zone scene may involve playback of a particular media content or content source, which causes the playback device to apply a particular calibration.

[0179] In yet further examples, the playback configuration may be indicated to the playback device by way of one or more messages from a control device or another playback device. For instance, after receiving input that selects a particular playback configuration, a device may indicate to the playback device that a particular playback configuration is selected. The playback device may apply a calibration associated with that playback configuration. As another example, the playback device may be a member of a group, such as a bonded zone group. Another playback device, such as a group coordinator device of that group, may detect a playback configuration of the group and send a message indicating that playback configuration (or the calibration for that configuration) to the playback device.

[0180] In some cases, the playback device may also apply the calibration to one or more additional playback devices. For instance, the playback device may be a member (e.g., the group coordinator) of a group (e.g., a zone group). The playback device may send messages instructing other playback devices in the group to apply the calibration. Upon receiving such a message, these playback devices may apply the calibration.

[0181] In some examples, the calibration or calibration state may be shared among devices of a media playback system using one or more state variables. Some examples techniques involving calibration state variables are described in U.S. Patent Application No. 14/793,190 filed July 7, 2015, entitled "Calibration State Variable," and U.S. Patent Application No. 14/793,205 filed July 7, 2015, entitled "Calibration Indicator."

## V. Example Techniques To Facilitate Spectral Calibration Using Applied Spatial Calibration

[0182] As noted above, embodiments described herein facilitate the calibration of one or more playback devices. Figure 20 illustrates an example implementation 2,00 by which an NMD facilitates spectral calibration of a media playback system using applied spatial calibration.

### a. Detect Trigger Condition

[0183] At block 2002, implementation 2000 involves detecting a trigger condition that initiates calibration. For instance, a NMD detects a trigger condition that initiates calibration of a media playback system. The trigger condition initiates calibration of the playback device(s) in the media playback system for multiple playback configurations, either explicitly or perhaps because the audio driver(s) of the playback device(s) have been set up with multiple playback configurations. Example trigger condi-

tions to initiate a calibration are described above in section III. a, as well as generally throughout the disclosure.

### b. Cause Audio Driver(s) To Output Calibration Audio

[0184] In Figure 20, at block 2004, implementation 2000 involves causing audio driver(s) to output calibration audio. For instance the NMD causes multiple audio drivers to output calibration audio. The NMD transmits an instruction the playback device(s) under calibration, perhaps via a network interface. Example calibration audio is described above in connection with the example calibration techniques.

### c. Record Calibration Audio

[0185] In Figure 20, at block 2006, implementation 2000 involves recording the calibration audio. For instance, the NMD records the calibration audio as output by the audio driver(s) of the playback device(s) under calibration via a microphone. In some cases, multiple NMDs may record the calibration audio via respective microphones.

[0186] The NMD may be moving through the environment while recording the calibration audio so as to measure the calibration sounds at different locations. With a moving microphone, repetitions of the calibration sound are detected at different physical locations within the environment. Samples of the calibration sound at different locations may provide a better representation of the surrounding environment as compared to samples in one location. For example, referring back to Figure 7, control device 126 of media playback system 100 may detect calibration audio emitted by one or more playback devices (e.g., playback devices 104, 106, 108, and/or 110 of the Living Room Zone) at various points along the path 700 (e.g., at point 702 and/or point 704). Alternatively, the control device may record the calibration signal along the path.

[0187] As such, an NMD may display one or more prompts to move the NMD while the calibration audio is being emitted. Such prompts may guide a user in moving the recording device during the calibration. To illustrate, in Figure 21, smartphone 500 is displaying control interface 2100 which includes graphical regions 2102 and 2104. Graphical region 2102 prompts to watch an animation in graphical region 2104. Such an animation may depict an example of how to move the smartphone within the environment during calibration to measure the calibration audio at different locations. While an animation is shown in graphical region 2104 by way of example, the control device may alternatively show a video or other indication that illustrates how to move the control device within the environment during calibration. Control interface 2100 also includes selectable controls 2106 and 2108, which respectively advance and step backward in the calibration sequence.

[0188] Further examples for recording calibration au-

dio are described above in section III a. as well as generally throughout the disclosure.

#### d. Determine Spectral Calibration(s)

**[0189]** At block 2008, implementation 2000 involves determining spectral calibrations. For instance, the NMD causes a processing device to determine respective sets of spectral filters for the multiple playback configurations under calibrations. These spectral calibrations may be based on the recorded calibration audio output by the audio driver(s). In some cases, the NMD may include the processing device. Alternatively, the NMD may transmit the recorded audio to one or more other processing devices. Example processing devices and processing techniques are described above.

**[0190]** The NMD may causing a particular calibration (e.g., a particular set of spectral filters) corresponding to a given playback configuration to be applied to the sound axes formed by the multiple audio drivers when the media playback system is playing back audio content in the given playback configuration. Additional examples of applying calibrations are described above.

#### VI. Conclusion

**[0191]** The description above discloses, among other things, various example systems, methods, apparatus, and articles of manufacture including, among other components, firmware and/or software executed on hardware. It is understood that such examples are merely illustrative and should not be considered as limiting. For example, it is contemplated that any or all of the firmware, hardware, and/or software aspects or components can be embodied exclusively in hardware, exclusively in software, exclusively in firmware, or in any combination of hardware, software, and/or firmware. Accordingly, the examples provided are not the only way(s) to implement such systems, methods, apparatus, and/or articles of manufacture.

**[0192]** The specification is presented largely in terms of illustrative environments, systems, procedures, steps, logic blocks, processing, and other symbolic representations that directly or indirectly resemble the operations of data processing devices coupled to networks. These process descriptions and representations are typically used by those skilled in the art to most effectively convey the substance of their work to others skilled in the art. Numerous specific details are set forth to provide a thorough understanding of the present disclosure. However, it is understood to those skilled in the art that certain embodiments of the present disclosure can be practiced without certain, specific details. In other instances, well known methods, procedures, components, and circuitry have not been described in detail to avoid unnecessarily obscuring aspects of the embodiments. Accordingly, the scope of the present disclosure is defined by the appended claims rather than the foregoing description of embod-

iments.

**[0193]** When any of the appended claims are read to cover a purely software and/or firmware implementation, at least one of the elements in at least one example is hereby expressly defined to include a tangible, non-transitory medium such as a memory, DVD, CD, Blu-ray, and so on, storing the software and/or firmware.

#### 10 Claims

1. A method for a networked microphone device, NMD, the method comprising:

detecting (1302,2002) a trigger condition that initiates calibration of a media playback system comprising multiple audio drivers, causing (1304,2006), via a network interface of the networked microphone device, each of the multiple audio drivers of the media playback system to output calibration audio that is divided into a repeating set of frames (1402,1404,1406), recording (1306), via a microphone of the NMD, the calibration audio output by the multiple audio drivers; and causing (1308,2008) a processing device to determine respective sets of spectral filters for the media playback system, the method **characterised in that:**

the calibration of the media playback system is for multiple playback configurations, each playback configuration representing a respective set of one or more sound axes formed via the multiple audio drivers of the media playback system, wherein each sound axis corresponds to a respective input channel of audio content; and each frame of the output calibration audio corresponds to a respective playback configuration; and the trigger condition causes each of the multiple audio drivers of the media playback system to output the calibration audio via the one or more sound axes corresponding to the given playback configuration during respective slots (1402A-C,1404A-C,1406A-C) of each frame corresponding to the respective playback configuration such that, during each frame of the set of frames, a respective set of spatial filters comprising a respective spatial filter for each of the one or more sound axes corresponding to the respective playback configuration is applied to the multiple audio drivers, wherein the spatial filter calibrates the media playback system to a given listening area spatially by directing sound output for

- a particular sound axis of the set of sound axes in a certain direction by arraying multiple audio drivers to form the particular sound axis; and  
the processing device determines the respective sets of spectral filters for the multiple playback configurations based on the recorded calibration audio, each set of spectral filters comprising a respective spectral filter for each sound axis.
2. The method of claim 1, further comprising, when the media playback system is playing back audio content in a given playback configuration, causing (1310) the determined set of spectral filters corresponding to the given playback configuration to be applied to the sound axes formed by the multiple audio drivers.
3. The method of any preceding claim, wherein causing the respective set of spatial filters to be applied to the multiple audio drivers comprises causing the processing device to apply the spatial filters to the calibration audio and transmit the calibration audio with the applied spatial filters to one or more playback devices (200) comprising the multiple audio drivers (212).
4. The method of any preceding claim, wherein the media playback system comprises multiple playback devices each comprising a subset of the multiple audio drivers.
5. The method of any preceding claim, wherein:  
in a surround sound playback configuration:  
each sound axis corresponds to a respective channel of surround sound audio content; and a first spatial filter corresponds to the surround sound playback configuration; in a stereo playback configuration:  
each sound axis corresponds to a respective channel of stereo audio content; and a second spatial filter corresponds to the stereo playback configuration; and in a mono playback configuration:  
the multiple audio drivers form a single sound axis; and  
a third spatial filter corresponds to the mono playback configuration.
6. The method of claim 5, wherein:  
the mono playback configuration is a first mono playback configuration,  
the stereo playback configuration is a first stereo playback configuration,
- the surround sound playback configuration is a first surround sound configuration; and  
the multiple playback configurations comprise at least one of:  
a second mono playback configuration, the multiple audio drivers configured to form one or more full frequency range sound axes and a subwoofer sound axis to synchronously output mono audio content when playing back audio content in the mono playback configuration, wherein a fourth spatial filter corresponds to the second mono playback configuration;  
a second stereo playback configuration, the multiple audio drivers configured to form one or more sound axes to synchronously output channels of stereo content audio content with a subwoofer sound axis when playing back audio content in the second stereo playback configuration, wherein a third spatial filter corresponds to the second stereo playback configuration; and  
a second surround sound playback configuration, the multiple audio drivers configured to form one or more full frequency range sound axes to synchronously output respective channels of surround sound content audio content with a subwoofer sound axis when playing back audio content in the second surround sound playback configuration, wherein a fourth spatial filter corresponds to the second surround sound playback configuration.
7. The method of any preceding claim, wherein the multiple playback configurations comprise two or more of:  
a music playback configuration, the multiple audio drivers configured to form sound axes to output music content when playing back audio content in the music playback configuration, wherein a music playback spatial filter corresponds to the music playback configuration, and  
a home theater playback configuration, the multiple audio drivers configured to form sound axes to output audio content that is paired to video content when playing back audio content in the home theater playback configuration, wherein a home theater playback spatial filter corresponds to the home theater playback configuration.
8. The method of any preceding claim, wherein the calibration audio is second calibration audio, the method further comprising:  
before causing the multiple audio drivers to output the second calibration audio, causing the multiple



audio drivers to output first calibration audio that is divided into a repeating set of frames comprising a respective frame for each playback configuration of the multiple playback configurations:

recording, via the microphone, the first calibration audio output by the multiple audio drivers; and  
causing the processing device to determine the respective sets of spatial filters for the multiple playback configurations based on the recorded first calibration audio, each set of spatial filters comprising a respective spatial filter for each sound axis.

**9.** The method of claim 8, wherein:

the determined sets of spatial filters calibrate the playback device to a particular listening location within a listening area of the playback device, and  
the determined spectral filters offset acoustic characteristics of the listening area.

**10.** The method of claim 8 or 9, wherein:

causing the multiple audio drivers to output the first calibration audio comprises causing multiple audio drivers to emit calibration audio via multiple sound axes at respective slots in each frame, each sound axis corresponding to a respective channel of multi-channel audio content; and

causing the processing device to determine the respective sets of spatial filters comprises:

causing respective spatial delays for each sound axis of the multiple sound axes to be determined based on the slots of recorded calibration audio corresponding to the sound axes according to the respective determined delays, wherein causing respective delays for each sound axis of the multiple sound axes to be determined comprises:

causing a processing device to determine respective times-of-arrival at the microphone for each sound axis of the multiple sound axes from the slots of recorded calibration audio corresponding to each sound axis; and  
causing the delays for each sound axis of the multiple sound axes to be determined, each determined delay corresponding to the determined time-of-arrival of a respective sound axis.

**11.** The method of claim 10, wherein causing the processing device to determine respective times-of-arrival at the microphone for each sound axis of the multiple sound axes comprises:

dividing the recorded calibration audio into the constituent frames and each constituent frame into the respective slots for each sound axis; determining respective impulse responses for the sound axes from the respective slots corresponding to each sound axis; aligning the respective impulse responses to a first reference point; identifying respective second reference points in each impulse response; and determining the respective times-of-arrival at the microphone based on respective differences from the first reference point to the second reference points in each impulse response, wherein the sound axes consist of a reference sound axis and one or more other sound axes, and wherein identifying respective second reference points in each impulse response comprises:

identifying, as a given second reference point, a peak value in the impulse response of the reference sound axis; and  
identifying, as the other second reference points, respective peak values of the impulse responses of the one or more other sound axes in a time window subsequent to the given second reference point.

**12.** The method of claim 10 or 11, wherein causing the processing device to determine respective times-of-arrival at the microphone for each sound axis of the multiple sound axes comprises:

sending, via the network interface to the processing device:

the recorded calibration audio, and  
an instruction to determine respective times-of-arrival at the microphone for each sound axis of the multiple sound axes; and  
receiving, via the network interface, the determined respective times-of-arrival.

**13.** The method of claim 10, wherein causing delays for each sound axis of the multiple sound axes to be determined comprises:

determining that time-of-arrival of a given sound axis exceeds a maximum delay threshold; and  
causing delay of the given sound axis to be set at the maximum delay threshold when the media playback system is playing back audio content that is paired to video content.

14. The method of any preceding claim, wherein detecting the trigger condition that initiates calibration of a media playback system comprises one of:

detecting, via a user interface, input data indicating a command to initiate calibration of the media playback system; and  
detecting configuration of the media playback system into a particular axis configuration, wherein the multiple audio drivers form a particular set of sound axes.

15. A media playback system comprising:

multiple audio drivers;  
a networked microphone device comprising a network interface and a microphone;  
a playback device (200) comprising at least one audio driver from the multiple audio drivers and a network interface; and  
a processing device,  
wherein the networked microphone device is configured to perform the method according to any of claims 1 to 14, and  
the playback device is configured to, when playing back audio content in a given playback configuration, cause an audio stage to apply a particular spectral filter corresponding to the given playback configuration.

## Patentansprüche

1. Verfahren für eine vernetzte Mikrofonvorrichtung, NMD, wobei das Verfahren umfasst:

Erkennen (1302, 2002) einer Auslösebedingung, die eine Kalibrierung eines Medienwiedergabesystems initiiert, das mehrere Audiotreiber umfasst,  
Bewirken (1304, 2006), über eine Netzwerkschnittstelle der vernetzten Mikrofonvorrichtung, dass jeder der mehreren Audiotreiber des Medienwiedergabesystems ein Kalibrierungsaudio ausgibt, das in einen sich wiederholenden Satz von Frames (1402, 1404, 1406) unterteilt ist,  
Aufzeichnen (1306), über ein Mikrofon der NMD, des von den mehreren Audiotreibern ausgegebenen Kalibrierungsaudios; und  
Bewirken (1308, 2008), dass eine Verarbeitungsvorrichtung jeweilige Sätze von Spektralfiltern für das Medienwiedergabesystem bestimmt, wobei das Verfahren **dadurch gekennzeichnet ist, dass:**

die Kalibrierung des Medienwiedergabesystems für mehrere Wiedergabekonfigurationen gilt, wobei jede Wiedergabekonfiguration einen jeweiligen Satz einer oder mehrerer Klangachsen darstellt, die über die mehreren Audiotreiber des Medienwiedergabesystems gebildet werden, wobei jede Klangachse einem jeweiligen Eingangskanal von Audioinhalten entspricht; und  
jeder Frame des ausgegebenen Kalibrierungsaudios einer jeweiligen Wiedergabekonfiguration entspricht; und  
die Auslösebedingung bewirkt, dass jeder der mehreren Audiotreiber des Medienwiedergabesystems das Kalibrierungsaudio über die eine oder die mehreren Klangachsen, die der gegebenen Wiedergabekonfiguration entsprechen, während jeweiliger Slots (1402A-C, 1404A-C, 1406A-C) jedes Frames, der der jeweiligen Wiedergabekonfiguration entspricht, ausgibt, sodass während jedes Frames des Satzes von Frames ein jeweiliger Satz von räumlichen Filtern, umfassend einen jeweiligen räumlichen Filter für jede der einen oder mehreren Klangachsen, die der jeweiligen Wiedergabekonfiguration entsprechen, auf die mehreren Audiotreiber angewendet wird, wobei der räumliche Filter das Medienwiedergabesystem räumlich auf einen gegebenen Hörbereich kalibriert, indem eine Klangausgabe für eine bestimmte Klangachse des Satzes von Klangachsen in eine bestimmte Richtung gelenkt wird, indem mehrere Audiotreiber so angeordnet werden, dass sie die bestimmte Klangachse bilden; und  
die Verarbeitungsvorrichtung die jeweiligen Sätze von Spektralfiltern für die mehreren Wiedergabekonfigurationen basierend auf dem aufgezeichneten Kalibrierungsaudio bestimmt, wobei jeder Satz von Spektralfiltern einen jeweiligen Spektralfilter für jede Klangachse umfasst.

2. Verfahren nach Anspruch 1, ferner umfassend, wenn das Medienwiedergabesystem Audioinhalte in einer gegebenen Wiedergabekonfiguration wiedergibt, das Bewirken (1310), dass der bestimmte Satz von Spektralfiltern, die der gegebenen Wiedergabekonfiguration entsprechen, auf die Klangachsen angewendet wird, die durch die mehreren Audiotreiber gebildet werden.
3. Verfahren nach einem der vorstehenden Ansprüche, wobei das Bewirken, dass der jeweilige Satz von räumlichen Filtern auf die mehreren Audiotreiber angewendet wird, das Bewirken umfasst, dass die Verarbeitungsvorrichtung die räumlichen Filter auf das Kalibrierungsaudio anwendet und das Kalibrierungsaudio mit den angewandten räumlichen Filtern

an eine oder mehrere Wiedergabevorrichtungen (200), umfassend die mehreren Audiotreiber (212), übertragen wird.

4. Verfahren nach einem der vorstehenden Ansprüche, wobei das Medienwiedergabesystem mehrere Wiedergabevorrichtungen umfasst, die jeweils eine Teilmenge der mehreren Audiotreiber umfassen. 5

5. Verfahren nach einem der vorstehenden Ansprüche, wobei:  
in einer Surround-Sound-Wiedergabekonfiguration: 10

jede Klangachse einem jeweiligen Kanal von Surround-Sound-Audioinhalten entspricht; und ein erster räumlicher Filter der Surround-Sound-Wiedergabekonfiguration entspricht; in einer Stereo-Wiedergabekonfiguration: 15

jede Klangachse einem jeweiligen Kanal von Stereo-Audioinhalten entspricht; und ein zweiter räumlicher Filter der Stereo-Wiedergabekonfiguration entspricht; und in einer Mono-Wiedergabekonfiguration: 20

die mehreren Audiotreiber eine einzige Klangachse bilden; und ein dritter räumlicher Filter der Mono-Wiedergabekonfiguration entspricht. 25

6. Verfahren nach Anspruch 5, wobei: 30

die Mono-Wiedergabekonfiguration eine erste Mono-Wiedergabekonfiguration ist, die Stereo-Wiedergabekonfiguration eine erste Stereo-Wiedergabekonfiguration ist, die Surround-Sound-Wiedergabekonfiguration eine erste Surround-Sound-Konfiguration ist; und die mehreren Wiedergabekonfigurationen mindestens eines der Folgenden umfassen: 35 40

eine zweite Mono-Wiedergabekonfiguration, wobei die mehreren Audiotreiber dazu konfiguriert sind, eine oder mehrere Vollfrequenzbereichsklangachsen und eine Subwoofer-Klangachse zu bilden, um Mono-Audioinhalte synchron auszugeben, wenn Audioinhalte in der Mono-Wiedergabekonfiguration wiedergegeben werden, wobei ein vierter räumlicher Filter der zweiten Mono-Wiedergabekonfiguration entspricht; 45 50

eine zweite Stereo-Wiedergabekonfiguration, wobei die mehreren Audiotreiber dazu konfiguriert sind, eine oder mehrere Klangachsen zu bilden, um Kanäle von Stereo-Audioinhalten mit einer Subwoofer-Klang- 55

achse zu auszugeben, wenn Audioinhalte in der zweiten Stereo-Wiedergabekonfiguration wiedergegeben werden, wobei ein dritter räumlicher Filter der zweiten Stereo-Wiedergabekonfiguration entspricht; und eine zweite Surround-Sound-Wiedergabekonfiguration, wobei die mehreren Audiotreiber dazu konfiguriert sind, eine oder mehrere Vollfrequenzbereichsklangachsen zu bilden, um jeweilige Kanäle von Surround-Sound-Audioinhalten mit einer Subwoofer-Klangachse synchron auszugeben, wenn Audioinhalte in der zweiten Surround-Sound-Wiedergabekonfiguration wiedergegeben werden, wobei ein vierter räumlicher Filter der zweiten Surround-Sound-Wiedergabekonfiguration entspricht.

7. Verfahren nach einem der vorstehenden Ansprüche, wobei die mehreren Wiedergabekonfigurationen eines oder mehrere der Folgenden umfassen:

eine Musikwiedergabekonfiguration, wobei die mehreren Audiotreiber dazu konfiguriert sind, Klangachsen zu bilden, um Musikinhalte auszugeben, wenn Audioinhalte in der Musikwiedergabekonfiguration wiedergegeben werden, wobei ein räumlicher Filter für Musikwiedergabe der Musikwiedergabekonfiguration entspricht, und

eine Heimkino-Wiedergabekonfiguration, wobei die mehreren Audiotreiber dazu konfiguriert sind, Klangachsen zu bilden, um Audioinhalte auszugeben, die mit Videoinhalten kombiniert werden, wenn Audioinhalte in der Heimkino-Wiedergabekonfiguration wiedergegeben werden, wobei ein räumlicher Filter für Heimkino-Wiedergabe der Heimkino-Wiedergabekonfiguration entspricht.

8. Verfahren nach einem der vorstehenden Ansprüche, wobei das Kalibrierungsaudio ein zweites Kalibrierungsaudio ist, wobei das Verfahren ferner umfasst: vor dem Bewirken, dass die mehreren Audiotreiber das zweite Kalibrierungsaudio ausgeben, Bewirken, dass die mehreren Audiotreiber ein erstes Kalibrierungsaudio ausgeben, das in einen sich wiederholenden Satz von Frames unterteilt ist, umfassend einen jeweiligen Frame für jede Wiedergabekonfiguration der mehreren Wiedergabekonfigurationen:

Aufzeichnen, über das Mikrofon, des von den mehreren Audiotreibern ausgegebenen Kalibrierungsaudios; und

Bewirken, dass die Verarbeitungsvorrichtung die jeweiligen Sätze von räumlichen Filtern für die mehreren Wiedergabekonfigurationen basierend auf dem aufgezeichneten ersten Kalib-

rierungsaudio bestimmt, wobei jeder Satz von räumlichen Filtern einen jeweiligen räumlichen Filter für jede Klangachse umfasst.

**9. Verfahren nach Anspruch 8, wobei:**

5

die bestimmten Sätze von räumlichen Filtern die Wiedergabevorrichtung auf einen bestimmten Hörplatz innerhalb eines Hörbereichs der Wiedergabevorrichtung kalibrieren, und die bestimmten Spektralfilter akustische Eigenschaften des Hörbereichs ausgleichen.

10

**10. Verfahren nach Anspruch 8 oder 9, wobei:**

15

das Bewirken, dass die mehreren Audiotreiber das erste Kalibrierungsaudio ausgeben, das Bewirken umfasst, dass mehrere Audiotreiber Kalibrierungsaudios über mehrere Klangachsen bei jeweiligen Slots in jedem Frame emittieren, wobei jede Klangachse einem jeweiligen Kanal von Mehrkanal-Audioinhalten entspricht; und das Bewirken, dass die Verarbeitungsvorrichtung die jeweiligen Sätze von räumlichen Filtern bestimmt, umfasst:

20

Bewirken, dass jeweilige räumliche Verzögerungen für jede Klangachse der mehreren Klangachsen basierend auf den Slots von aufgezeichneten Kalibrierungsaudios, die den Klangachsen entsprechen, gemäß den jeweiligen bestimmten Verzögerungen bestimmt werden, wobei das Veranlassen der Bestimmung der jeweiligen Verzögerungen für jede Schallachse der mehreren Schallachsen umfasst:

25

30

35

Bewirken, dass eine Verarbeitungsvorrichtung jeweilige Ankunftszeiten an dem Mikrofon für jede Klangachse der mehreren Klangachsen aus den Slots von aufgezeichneten Kalibrierungsaudios entsprechend jeder Klangachse bestimmt; und

40

45

Bewirken, dass die Verzögerungen für jede Klangachse der mehreren Klangachsen bestimmt werden, wobei jede bestimmte Verzögerung der bestimmten Ankunftszeit einer jeweiligen Klangachse entspricht.

50

**11. Verfahren nach Anspruch 10, wobei das Bewirken, dass die Verarbeitungsvorrichtung jeweilige Ankunftszeiten an dem Mikrofon für jede Klangachse der mehreren Klangachsen bestimmt, umfasst:**

55

Unterteilen des aufgezeichneten Kalibrierungs-

audios in die einzelnen Frames und jedes einzelnen Frames in die jeweiligen Slots für jede Klangachse;

Bestimmen jeweiliger Impulsantworten für die Klangachsen aus den jeweiligen Slots, die jeder Klangachse entsprechen;

Ausrichten der jeweiligen Impulsantworten auf einen ersten Referenzpunkt;

Identifizieren jeweiliger zweiter Referenzpunkte in jeder Impulsantwort; und Bestimmen der jeweiligen Ankunftszeit an dem Mikrofon basierend auf jeweiligen Differenzen von dem ersten Referenzpunkt zu den zweiten Referenzpunkten in jeder Impulsantwort, wobei die Klangachsen aus einer Referenzklangachse und einer oder mehreren anderen Klangachsen bestehen, und wobei das Identifizieren jeweiliger zweiter Referenzpunkte in jeder Impulsantwort umfasst:

Identifizieren, als einen gegebenen zweiten Referenzpunkt, eines Spitzenwerts in der Impulsantwort der Referenzklangachse; und

Identifizieren, als die anderen zweiten Referenzpunkte, von jeweiligen Spitzenwerten der Impulsantworten der einen oder der mehreren anderen Klangachsen in einem auf den gegebenen zweiten Referenzpunkt folgenden Zeitfenster.

**12. Verfahren nach Anspruch 10 oder 11, wobei das Bewirken, dass die Verarbeitungsvorrichtung jeweilige Ankunftszeiten an dem Mikrofon für jede Klangachse der mehreren Klangachsen bestimmt, umfasst: Senden, über die Netzwerkschnittstelle an die Verarbeitungsvorrichtung:**

des aufgezeichneten Kalibrierungsaudios; und einer Anweisung zum Bestimmen jeweiliger Ankunftszeiten an dem Mikrofon für jede Klangachse der mehreren Klangachsen; und Empfangen, über die Netzwerkschnittstelle, der bestimmten jeweiligen Ankunftszeiten.

**13. Verfahren nach Anspruch 10, wobei das Bewirken, dass Verzögerungen für jede Klangachse der mehreren Klangachsen bestimmt werden, umfasst:**

Bestimmen, dass die Ankunftszeit einer gegebenen Klangachse einen maximalen Verzögerungsschwellenwert überschreitet; und Bewirken, dass die Verzögerung der gegebenen Klangachse auf den maximalen Verzögerungsschwellenwert eingestellt wird, wenn das Medienwiedergabesystem Audioinhalte wiedergibt, die mit Videoinhalten kombiniert werden.

14. Verfahren nach einem der vorstehenden Ansprüche, wobei das Erkennen der Auslösebedingung, die die Kalibrierung eines Medienwiedergabesystems initiiert, eines von Folgendem umfasst:

Erkennen, über eine Benutzerschnittstelle, von Eingabedaten, die einen Befehl zum Initiieren der Kalibrierung des Medienwiedergabesystems angeben; und  
Erkennen der Konfiguration des Medienwiedergabesystems in einer bestimmten Achsenkonfiguration, wobei die mehreren Audiotreiber einen bestimmten Satz von Klangachsen bilden.

15. Medienwiedergabesystem, umfassend:

mehrere Audiotreiber;  
eine vernetzte Mikrofonvorrichtung, umfassend eine Netzwerkschnittstelle und ein Mikrofon;  
eine Wiedergabevorrichtung (200), umfassend mindestens einen Audiotreiber von den mehreren Audiotreibern und eine Netzwerkschnittstelle; und  
eine Verarbeitungsvorrichtung, wobei die vernetzte Mikrofonvorrichtung dazu konfiguriert ist, das Verfahren nach einem der Ansprüche 1 bis 14 durchzuführen, und die Wiedergabevorrichtung dazu konfiguriert ist, wenn Audioinhalte in einer gegebenen Wiedergabekonfiguration wiedergegeben werden, zu bewirken, dass eine Audiostufe einen bestimmten Spektralfilter anwendet, der der gegebenen Wiedergabekonfiguration entspricht.

## Revendications

1. Procédé destiné à un dispositif de microphone en réseau, NMD, le procédé comprenant :

la détection (1302, 2002) d'une condition de déclenchement qui initie l'étalonnage d'un système de lecture multimédia comprenant plusieurs pilotes audio,  
l'incitation (1304, 2006), par l'intermédiaire d'une interface réseau du dispositif de microphone en réseau, de chacun des multiples pilotes audio du système de lecture multimédia afin de produire un audio d'étalonnage divisé en un ensemble répétitif de trames (1402, 1404, 1406),  
l'enregistrement (1306), par l'intermédiaire d'un microphone du NMD, de l'audio d'étalonnage produit par les multiples pilotes audio ; et  
l'incitation (1308, 2008) d'un dispositif de traitement afin de déterminer des ensembles respectifs de filtres spectraux destinés au système de lecture multimédia, le procédé étant **caracté-**

sé en ce que :

l'étalonnage du système de lecture multimédia concerne plusieurs configurations de lecture, chaque configuration de lecture représentant un ensemble respectif d'un ou plusieurs axes sonores formés par l'intermédiaire des multiples pilotes audio du système de lecture multimédia, dans lequel chaque axe sonore correspond à un canal d'entrée respectif du contenu audio ; et chaque trame de l'audio d'étalonnage de sortie correspond à une configuration de lecture respective ; et

la condition de déclenchement incite chacun des multiples pilotes audio du système de lecture multimédia afin de produire l'audio d'étalonnage par l'intermédiaire d'un ou plusieurs axes sonores correspondant à la configuration de lecture donnée pendant les créneaux respectifs (1402A-C, 1404A-C, 1406A-C) de chaque trame correspondant à la configuration de lecture respective, de sorte que, pendant chaque trame de l'ensemble de trames, un ensemble respectif de filtres spatiaux comprenant un filtre spatial respectif pour chacun des un ou plusieurs axes sonores correspondant à la configuration de lecture respective est appliqué aux multiples pilotes audio, dans lequel le filtre spatial étalonne le système de lecture multimédia à une zone d'écoute donnée dans l'espace en dirigeant la sortie sonore d'un axe sonore particulier de l'ensemble d'axes sonores dans une certaine direction en mettant en réseau plusieurs pilotes audio afin de former l'axe sonore particulier ; et

le dispositif de traitement détermine les ensembles respectifs de filtres spectraux destinés aux multiples configurations de lecture en fonction de l'audio d'étalonnage enregistré, chaque ensemble de filtres spectraux comprenant un filtre spectral respectif destiné à chaque axe sonore.

2. Procédé selon la revendication 1, comprenant en outre, lorsque le système de lecture multimédia lit un contenu audio dans une configuration de lecture donnée, l'application (1310) de l'ensemble déterminé de filtres spectraux correspondant à la configuration de lecture donnée aux axes sonores formés par les multiples pilotes audio.
3. Procédé selon l'une quelconque revendication précédente, dans lequel l'application de l'ensemble respectif de filtres spatiaux aux multiples pilotes audio comprend l'application par le dispositif de traitement

des filtres spatiaux à l'audio d'étalonnage et la transmission de l'audio d'étalonnage aux filtres spatiaux appliqués à un ou plusieurs dispositifs de lecture (200) comprenant les multiples pilotes audio (212).

4. Procédé selon l'une quelconque revendication précédente, dans lequel le système de lecture multimédia comprend plusieurs dispositifs de lecture, chacun comprenant un sous-ensemble de multiples pilotes audio.

5. Procédé selon l'une quelconque revendication précédente, dans lequel :  
dans une configuration de lecture de son ambiophonique :

chaque axe sonore correspond à un canal respectif de contenu audio de son ambiophonique ;  
et

un premier filtre spatial correspond à la configuration de lecture de son ambiophonique ;  
dans une configuration de lecture stéréo :

chaque axe sonore correspond à un canal respectif de contenu audio stéréo ; et  
un deuxième filtre spatial correspond à la configuration de lecture stéréo ; et  
dans une configuration de lecture mono :

les multiples pilotes audio forment un axe sonore unique ; et  
un troisième filtre spatial correspond à la configuration de lecture mono.

6. Procédé selon la revendication 5, dans lequel :  
la configuration de lecture mono est une première configuration de lecture mono, la configuration de lecture stéréo est une première configuration de lecture stéréo, la configuration de lecture de son ambiophonique est une première configuration de son ambiophonique ; et les configurations de lecture multiples comprennent au moins l'une des configurations suivantes :

une seconde configuration de lecture mono, les multiples pilotes audio étant configurés pour former un ou plusieurs axes sonores de gamme complète de fréquences et un axe sonore de caisson de basse afin de produire de manière synchronisée un contenu audio mono lors de la lecture d'un contenu audio dans la configuration de lecture mono, dans lequel un quatrième filtre spatial correspond à la seconde configuration de lecture mono ;

une seconde configuration de lecture stéréo, les multiples pilotes audio étant configurés pour former un ou plusieurs axes sonores afin de produire de manière synchronisée des canaux de

contenu audio stéréo avec un axe sonore de caisson de basse lors de la lecture de contenu audio dans la seconde configuration de lecture stéréo, dans lequel un troisième filtre spatial correspond à la seconde configuration de lecture stéréo ; et

une seconde configuration de lecture de son ambiophonique, les multiples pilotes audio étant configurés pour former un ou plusieurs axes sonores de gamme complète de fréquences afin de produire de manière synchronisée des canaux respectifs de contenu audio de son ambiophonique avec un axe sonore de caisson de basse lors de la lecture de contenu audio dans la seconde configuration de lecture de son ambiophonique, dans lequel un quatrième filtre spatial correspond à la seconde configuration de lecture de son ambiophonique.

7. Procédé selon l'une quelconque revendication précédente, dans lequel les configurations de lecture multiples comprennent au moins deux des éléments suivants

une configuration de lecture musicale, les multiples pilotes audio étant configurés pour former des axes sonores afin de produire un contenu musical lors de la lecture d'un contenu audio dans la configuration de lecture musicale, dans lequel un filtre spatial de lecture musicale correspond à la configuration de lecture musicale, et

une configuration de lecture home cinéma, les multiples pilotes audio étant configurés pour former des axes sonores afin de produire un contenu audio associé à un contenu vidéo lors de la lecture d'un contenu audio dans la configuration de lecture home cinéma, dans lequel un filtre spatial de lecture home cinéma correspond à la configuration de lecture home cinéma.

8. Procédé selon l'une quelconque revendication précédente, dans lequel l'audio d'étalonnage est un second audio d'étalonnage, le procédé comprenant en outre :

avant l'incitation des multiples pilotes audio afin de produire le second audio d'étalonnage, l'incitation des multiples pilotes audio afin de produire le premier audio d'étalonnage divisé en un ensemble répétitif de trames comprenant une trame respective de chaque configuration de lecture parmi les configurations de lecture multiples :

l'enregistrement, par l'intermédiaire du microphone, du premier audio d'étalonnage produit par les multiples pilotes audio ; et  
l'incitation du dispositif de traitement afin de déterminer les ensembles respectifs de filtres spa-

tiaux destinés aux configurations de lecture multiples en fonction du premier audio d'étalonnage enregistré, chaque ensemble de filtres spatiaux comprenant un filtre spatial respectif destiné à chaque axe sonore.

5

**9.** Procédé selon la revendication 8, dans lequel :

les ensembles déterminés de filtres spatiaux étalonnent le dispositif de lecture à un emplacement d'écoute particulier dans une zone d'écoute du dispositif de lecture, et les filtres spectraux déterminés compensent les caractéristiques acoustiques de la zone d'écoute.

10

15

**10.** Procédé selon la revendication 8 ou 9, dans lequel :

l'incitation des multiples pilotes audio afin de produire le premier audio d'étalonnage comprend l'incitation des multiples pilotes audio afin d'émettre un audio d'étalonnage par l'intermédiaire de multiples axes sonores à des emplacements respectifs dans chaque trame, chaque axe sonore correspondant à un canal respectif d'un contenu audio multicanal ; et l'incitation du dispositif de traitement afin de déterminer les ensembles respectifs de filtres spatiaux comprend :

20

25

30

la détermination des retards spatiaux respectifs pour chaque axe sonore des multiples axes sonores en fonction des créneaux d'audio d'étalonnage enregistrés correspondant aux axes sonores en fonction des retards respectifs déterminés, dans lequel la détermination des retards respectifs pour chaque axe sonore des multiples axes sonores comprend :

35

40

l'incitation d'un dispositif de traitement afin de déterminer les temps d'arrivée respectifs au microphone pour chaque axe sonore des axes sonores multiples à partir des créneaux d'audio d'étalonnage enregistrés correspondant à chaque axe sonore ; et la détermination des retards pour chaque axe sonore des multiples axes sonores, chaque retard déterminé correspondant au temps d'arrivée déterminé d'un axe sonore respectif.

45

50

**11.** Procédé selon la revendication 10, dans lequel l'incitation du dispositif de traitement permettant de déterminer des moments d'arrivée respectifs au microphone pour chaque axe sonore des multiples axes sonores comprend :

55

la division de l'audio d'étalonnage enregistré en trames constitutives et de chaque trame constitutive en créneaux respectifs pour chaque axe sonore ;

la détermination des réponses impulsionnelles respectives pour les axes sonores à partir des fentes respectives correspondant à chaque axe sonore ;

l'alignement des réponses impulsionnelles respectives sur un premier point de référence ;

l'identification de seconds points de référence respectifs dans chaque réponse impulsionnelle ; et la détermination des heures d'arrivée respectives au microphone en fonction des différences respectives entre le premier point de référence et les seconds points de référence dans chaque réponse impulsionnelle, dans lequel les axes sonores sont constitués d'un axe sonore de référence et d'un ou plusieurs autres axes sonores, et dans lequel l'identification des seconds points de référence respectifs dans chaque réponse impulsionnelle comprend :

l'identification, en tant que second point de référence donné, d'une valeur de crête dans la réponse impulsionnelle de l'axe sonore de référence ; et

l'identification, en tant qu'autres seconds points de référence, des valeurs de crête respectives des réponses impulsionnelles d'un ou de plusieurs autres axes sonores dans une fenêtre temporelle postérieure au second point de référence donné.

**12.** Procédé selon la revendication 10 ou 11, dans lequel la détermination par le dispositif de traitement des heures d'arrivée respectives au microphone pour chaque axe sonore des multiples axes sonores comprend :

l'envoi, par l'intermédiaire de l'interface réseau au dispositif de traitement :

l'audio d'étalonnage enregistré, et une instruction afin de déterminer les temps d'arrivée respectifs au microphone pour chaque axe sonore des axes sonores multiples ; et la réception, par l'intermédiaire de l'interface réseau, des temps d'arrivée respectifs déterminés.

**13.** Procédé selon la revendication 10, dans lequel la détermination des retards pour chaque axe sonore des multiples axes sonores comprend :

la détermination que le temps d'arrivée d'un axe sonore donné dépasse un seuil de retard maximal ; et

le retard de l'axe sonore donné est réglé sur le seuil de retard maximal lorsque le système de lecture multimédia lit un contenu audio associé à un contenu vidéo.

5

14. Procédé selon l'une quelconque revendication précédente, dans lequel la détection de la condition de déclenchement qui initie l'étalonnage d'un système de lecture multimédia comprend l'un des éléments suivants :

10

la détection, par l'intermédiaire d'une interface utilisateur, de données d'entrée indiquant une commande pour initier l'étalonnage du système de lecture multimédia ; et

15

la détection de la configuration du système de lecture multimédia dans une configuration d'axe particulière, dans lequel les multiples pilotes audio forment un ensemble particulier d'axes sonores.

20

15. Système de lecture multimédia comprenant :

de multiples pilotes audio ;

un dispositif de microphone en réseau comprenant une interface réseau et un microphone ;

25

un dispositif de lecture (200) comprenant au moins un pilote audio à partir des multiples pilotes audio et une interface réseau ; et

un dispositif de traitement,

30

dans lequel le dispositif de microphone en réseau est configuré pour effectuer le procédé selon l'une quelconque des revendications 1 à 14, et

le dispositif de lecture est configuré pour, lors de la lecture d'un contenu audio dans une configuration de lecture donnée, activer un étage audio afin d'appliquer un filtre spectral particulier correspondant à la configuration de lecture donnée.

35

40

45

50

55



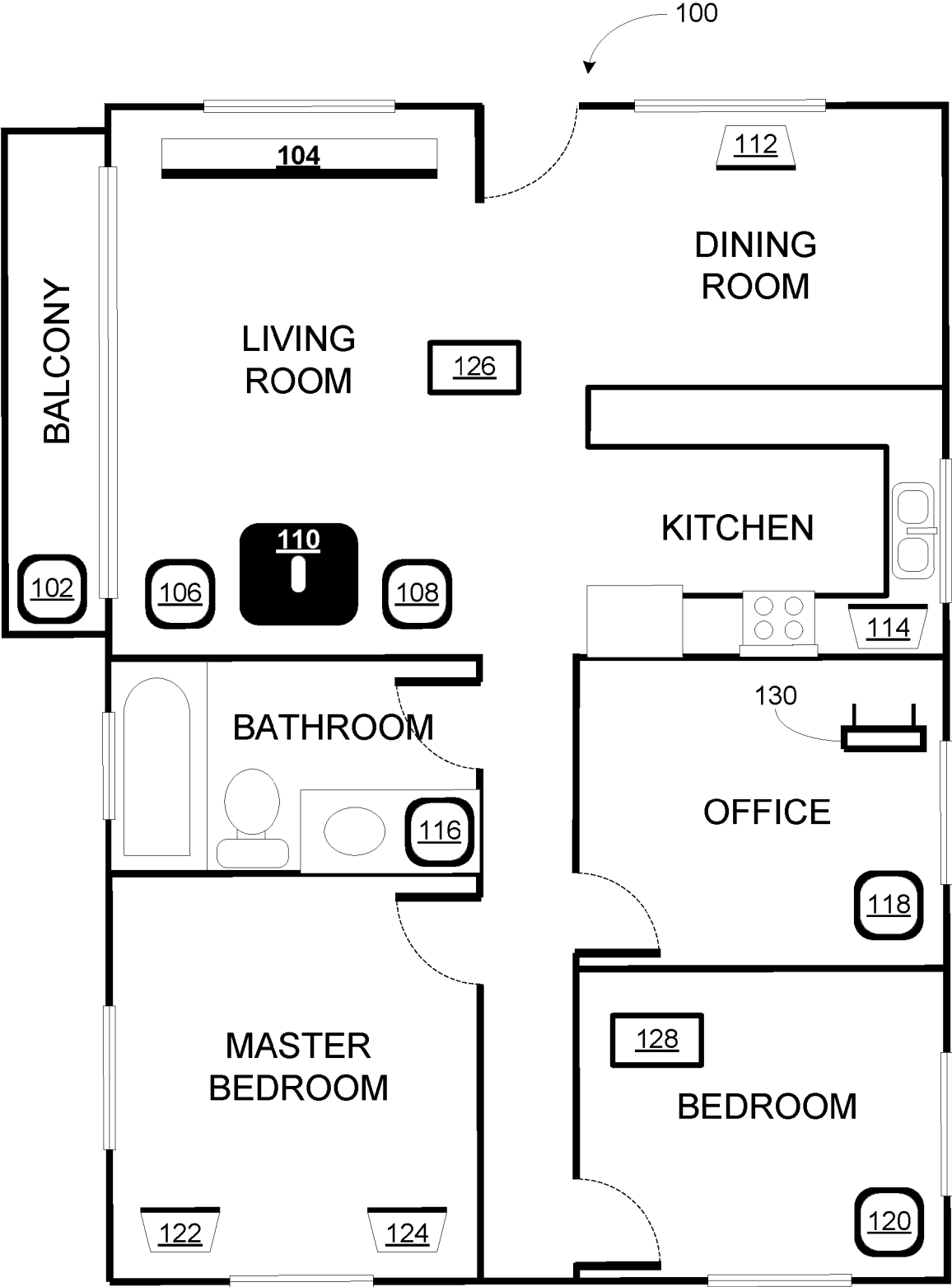


FIGURE 1

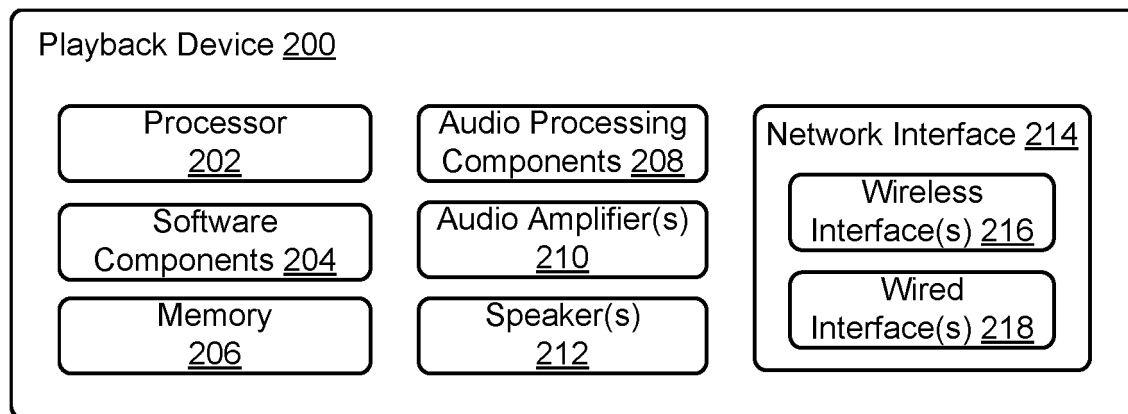


FIGURE 2

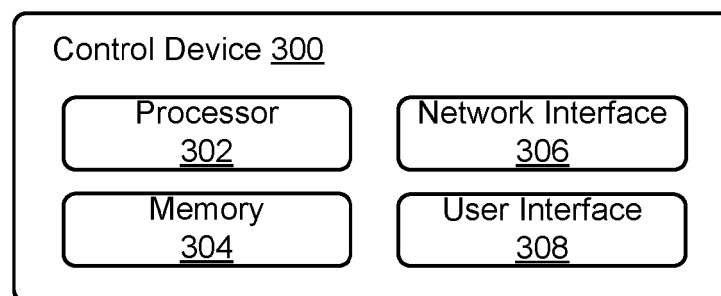


FIGURE 3

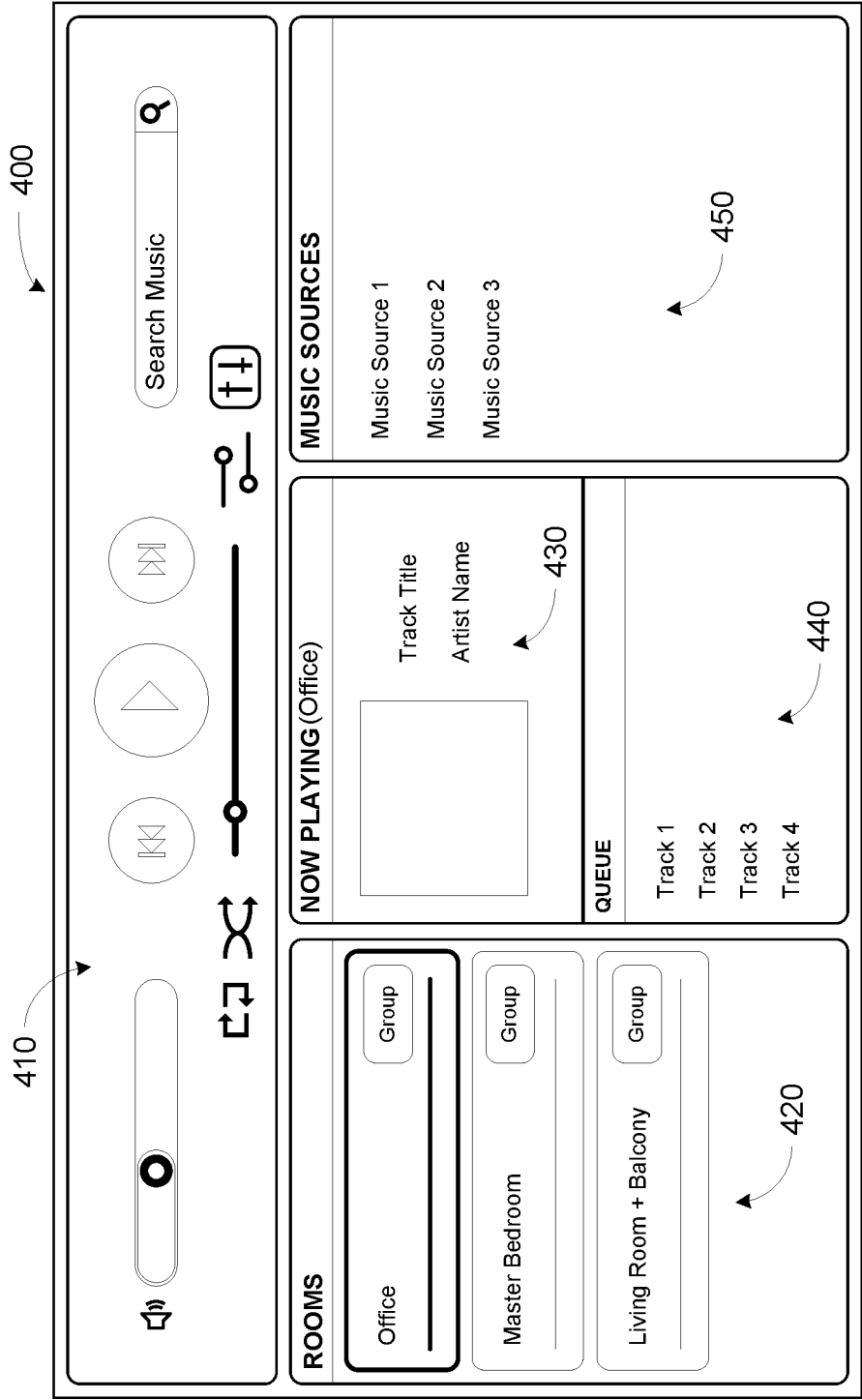


FIGURE 4

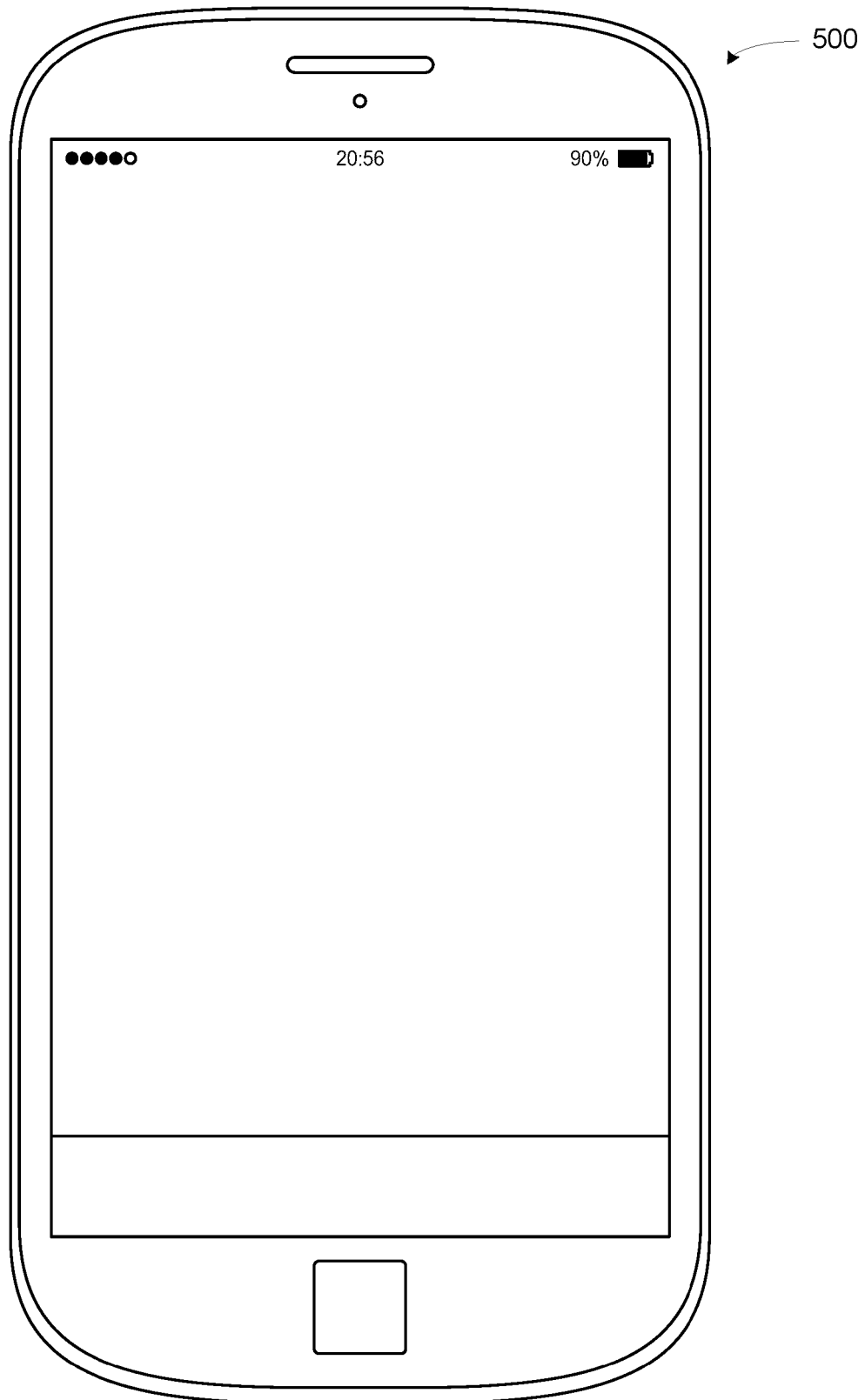


FIGURE 5

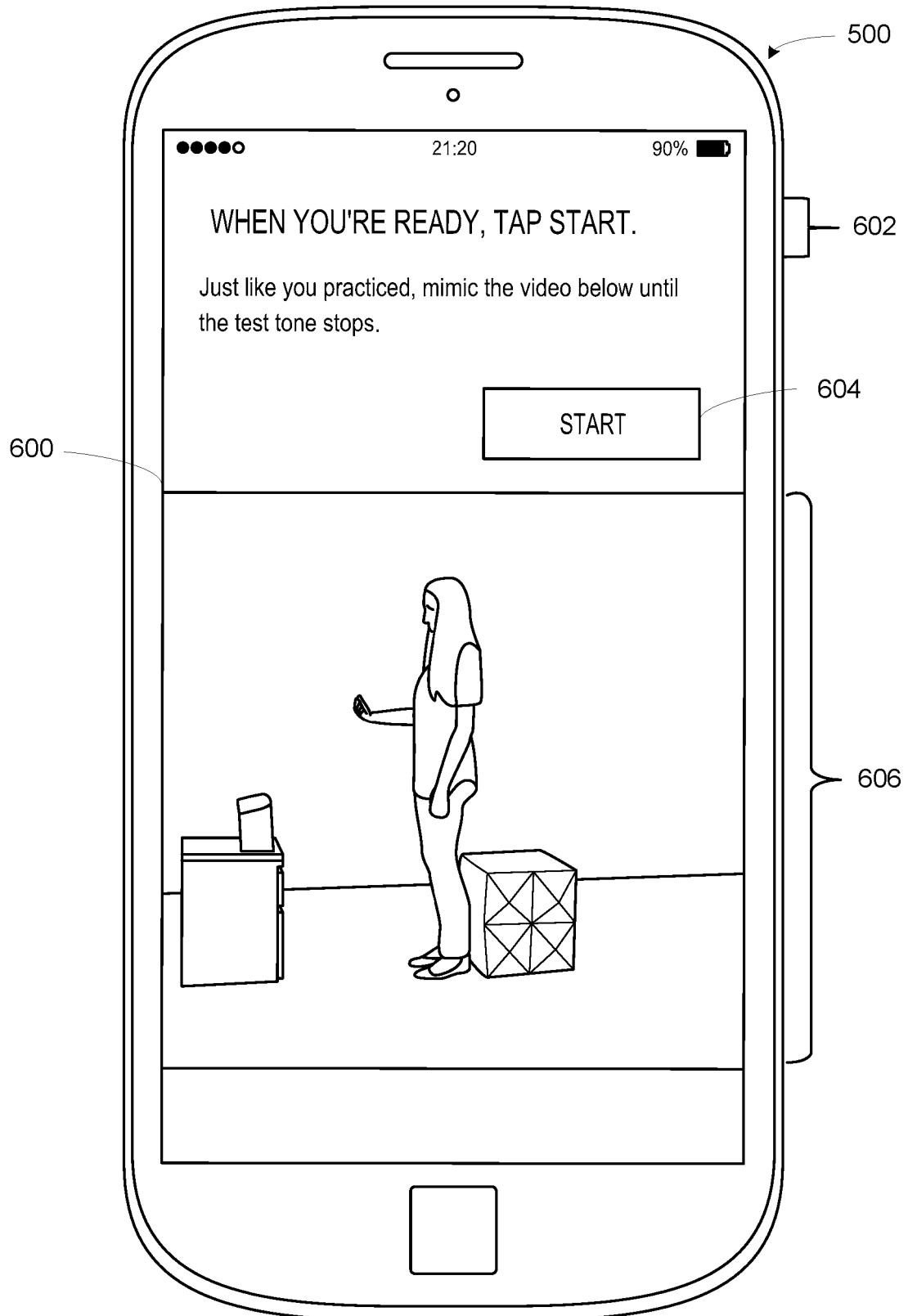


FIGURE 6

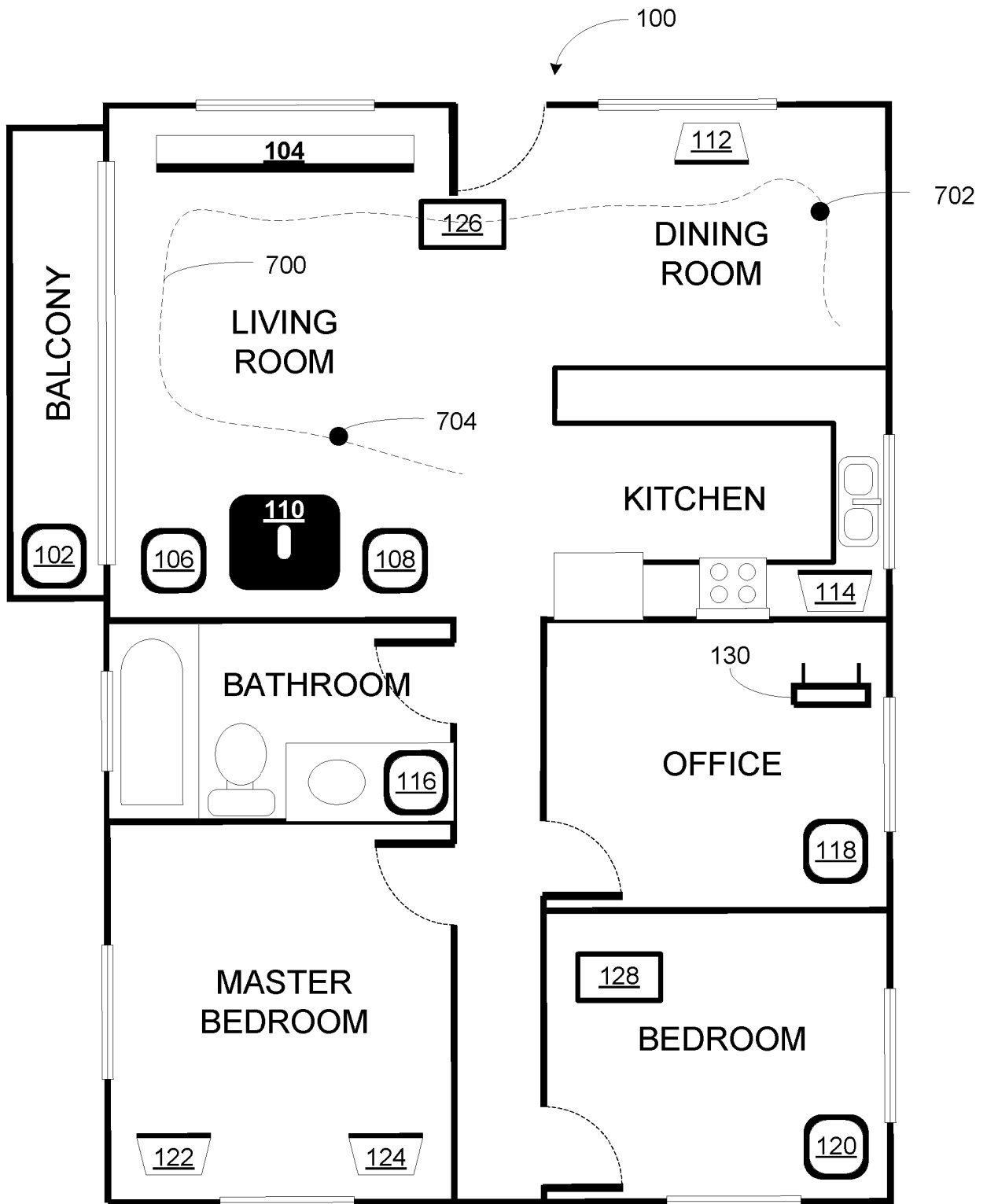


FIGURE 7

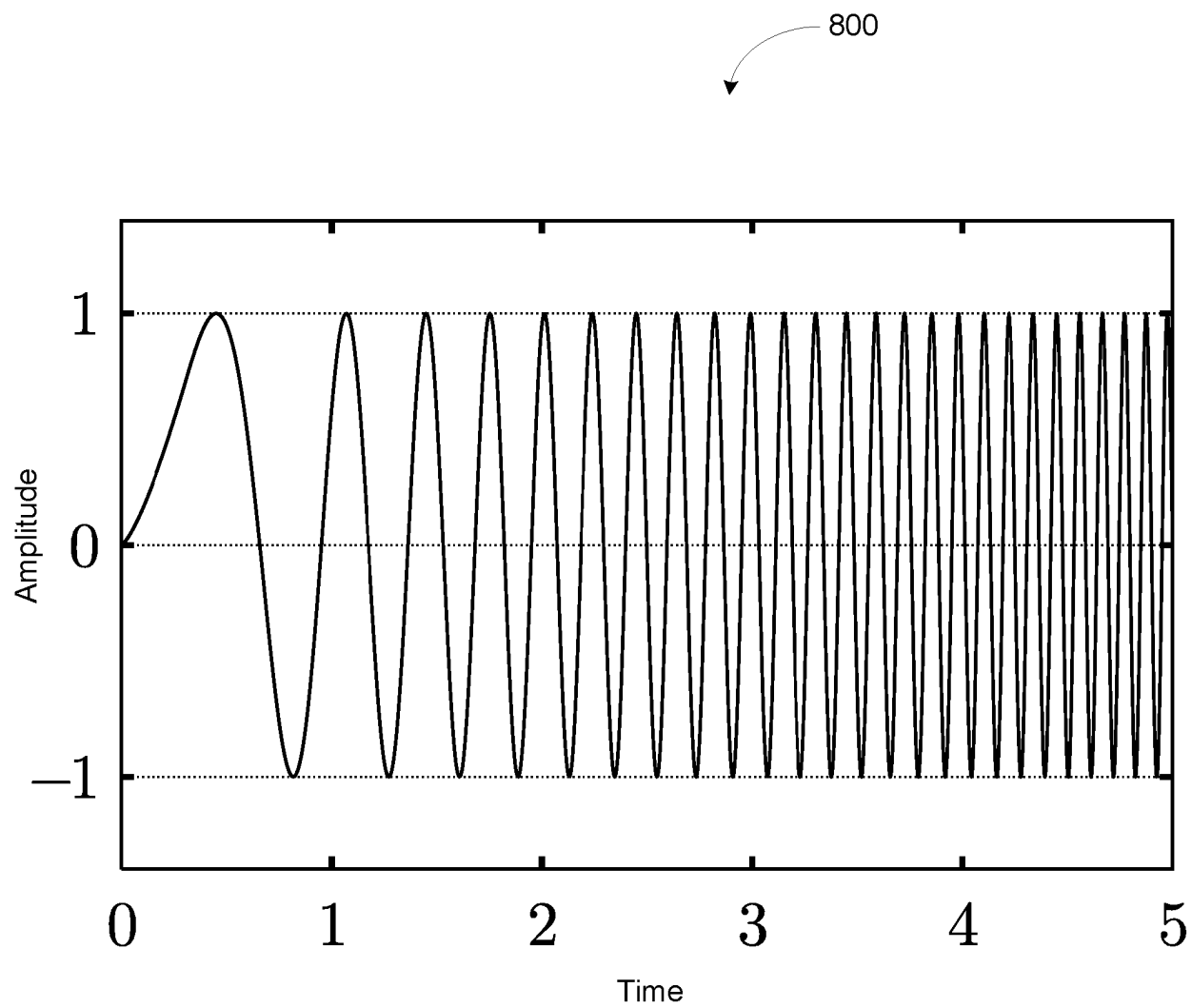


FIGURE 8

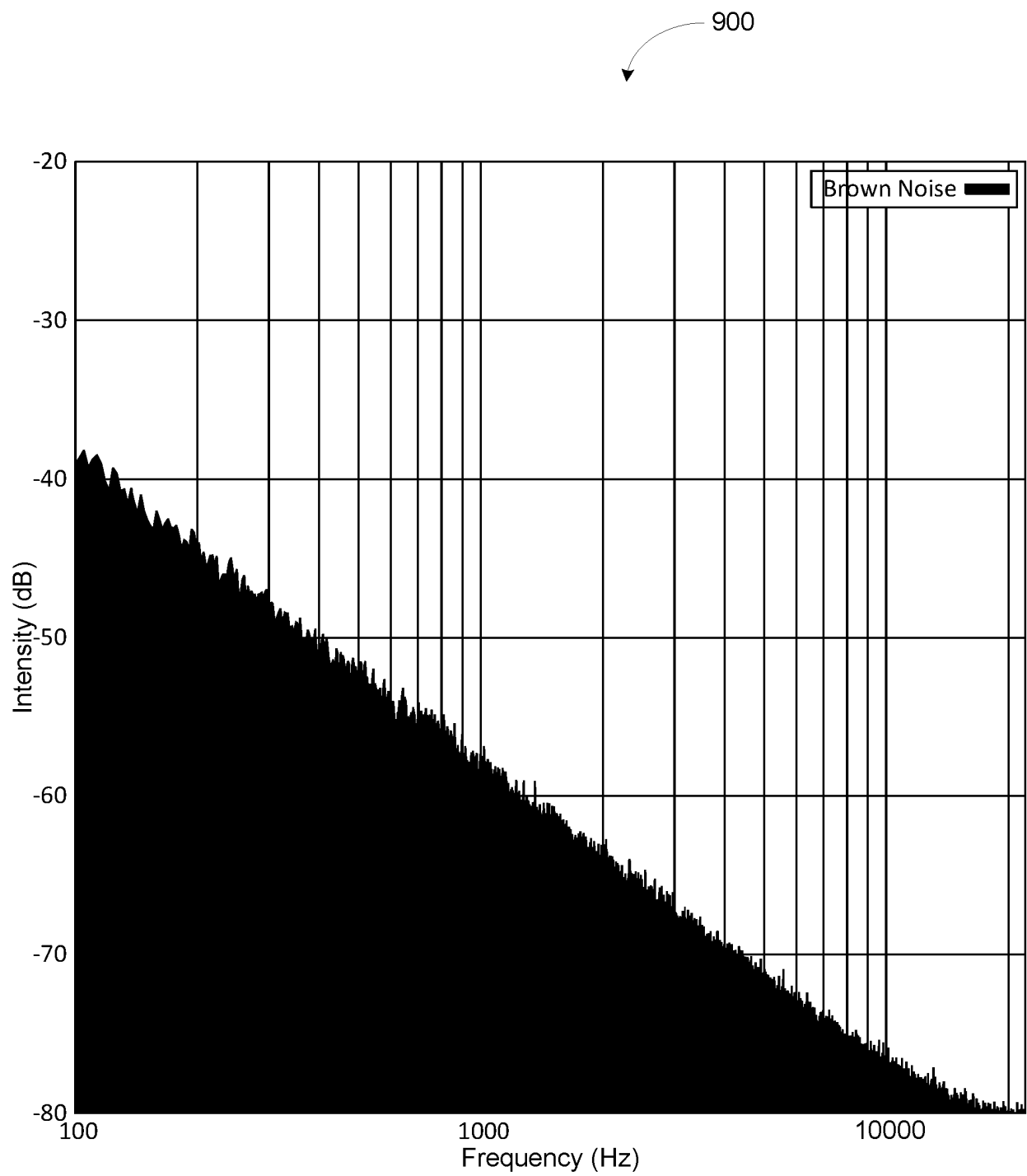


FIGURE 9



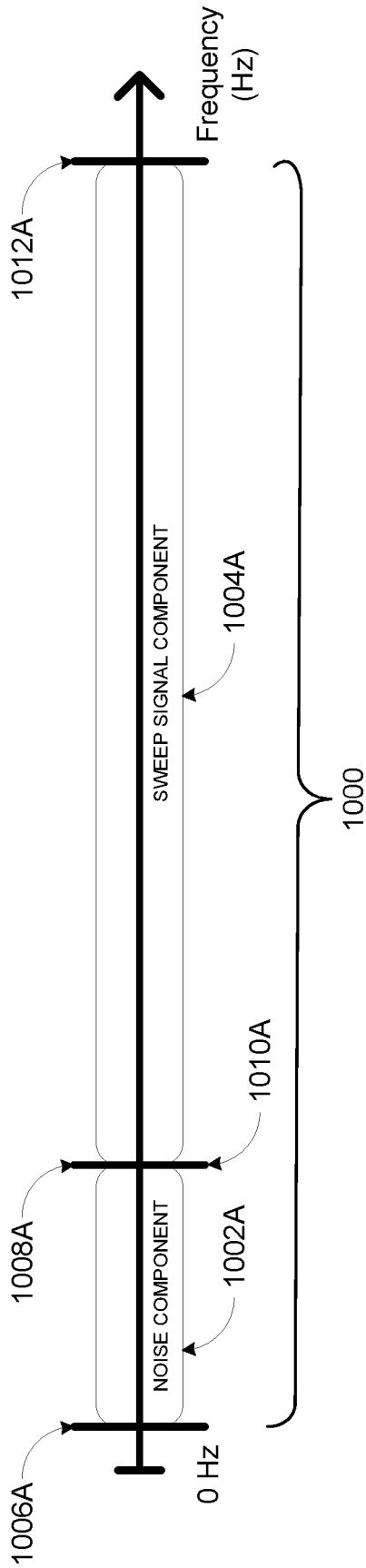


FIGURE 10A

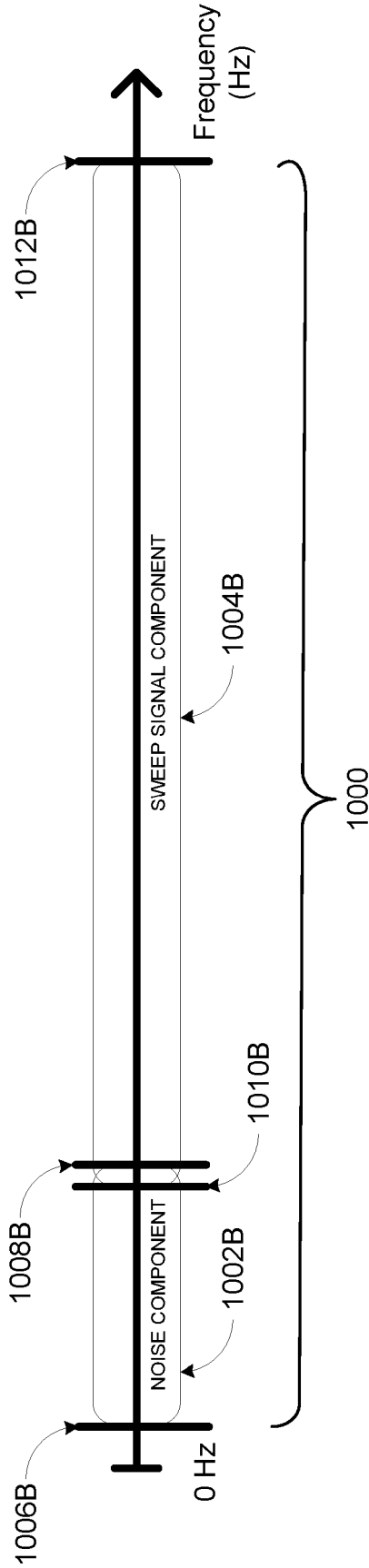


FIGURE 10B

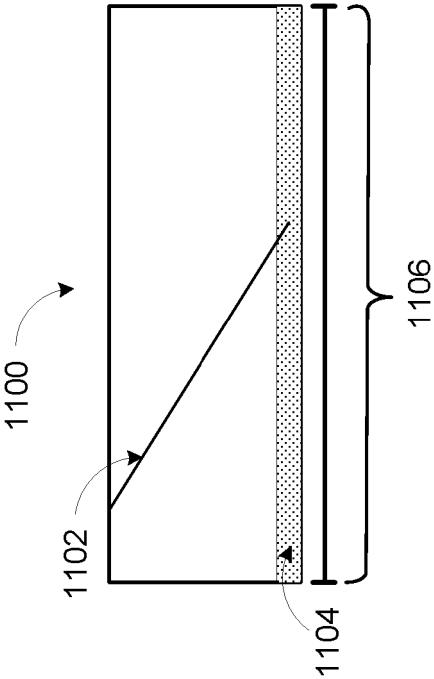


FIGURE 11

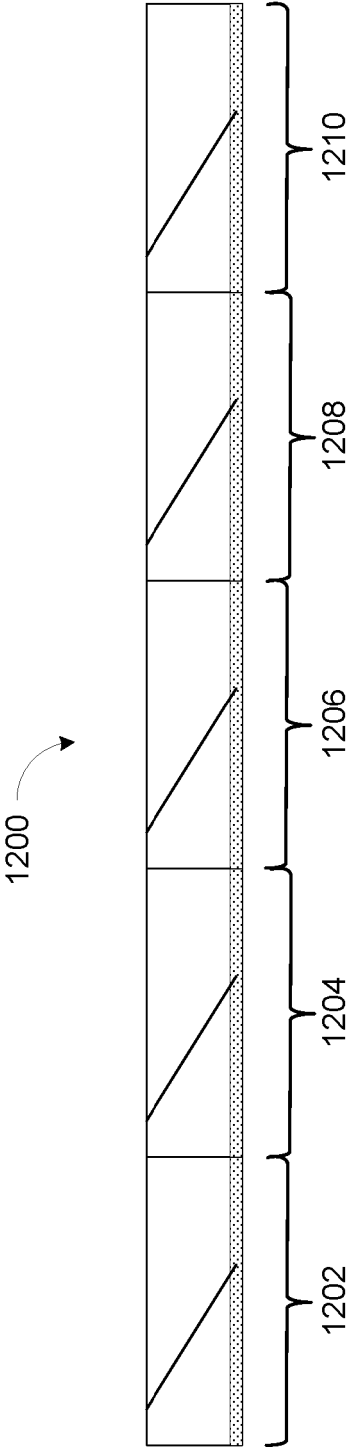


FIGURE 12

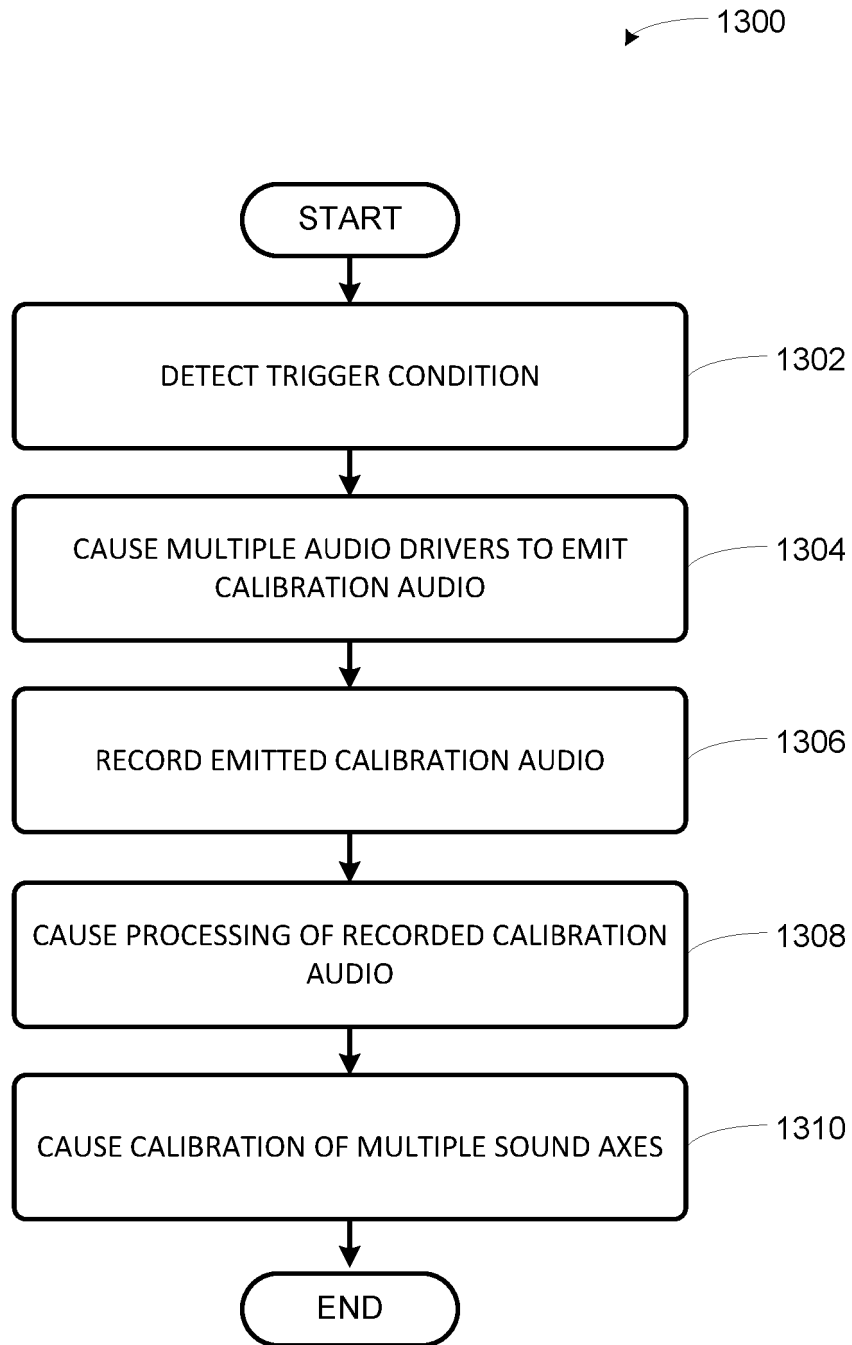


FIGURE 13

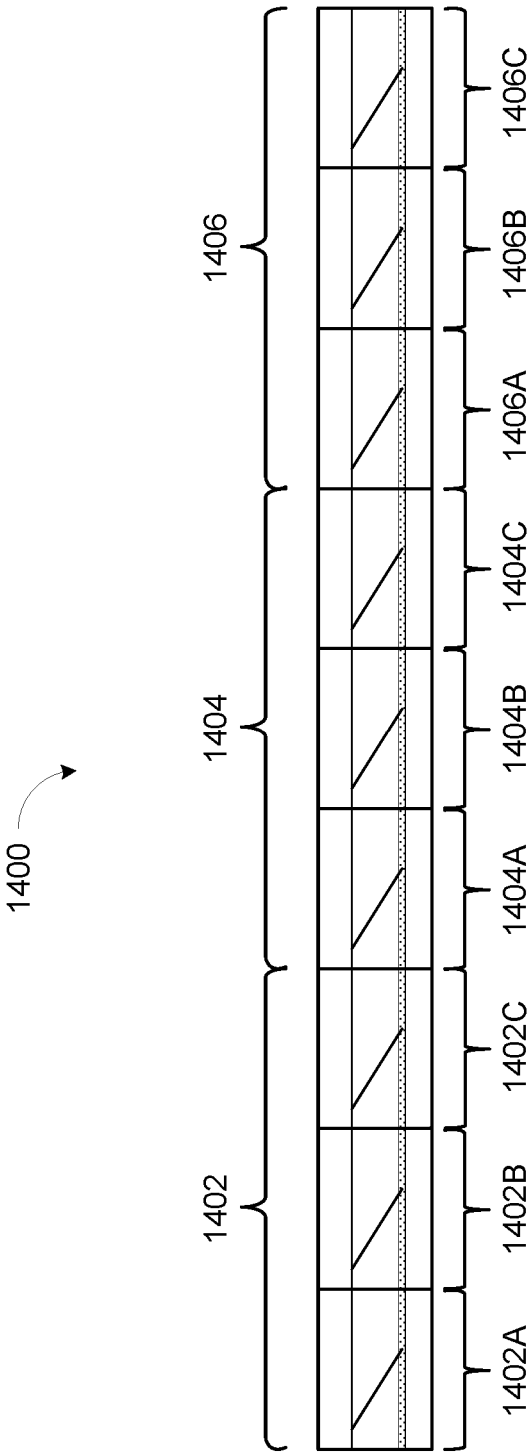


FIGURE 14

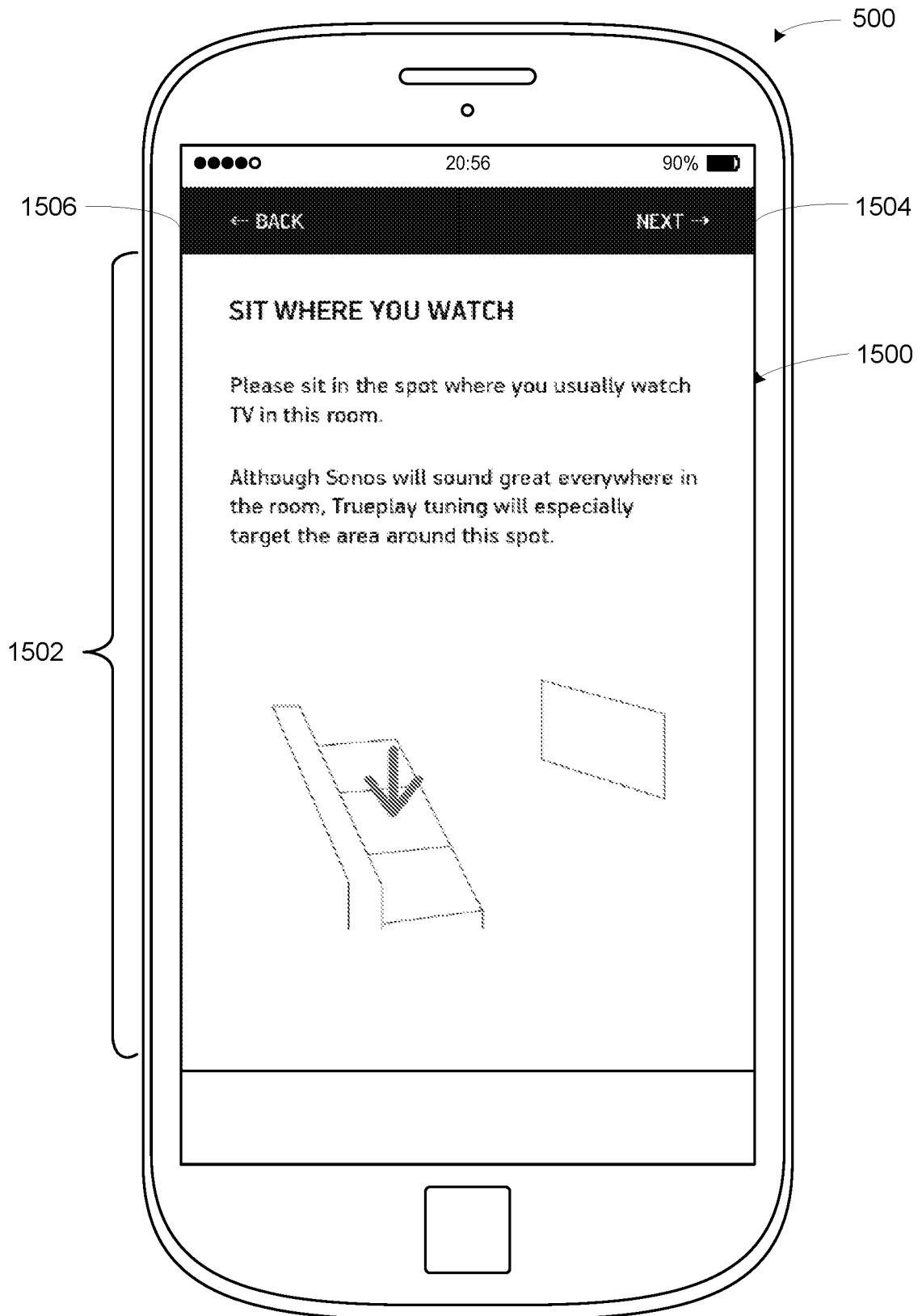


FIGURE 15

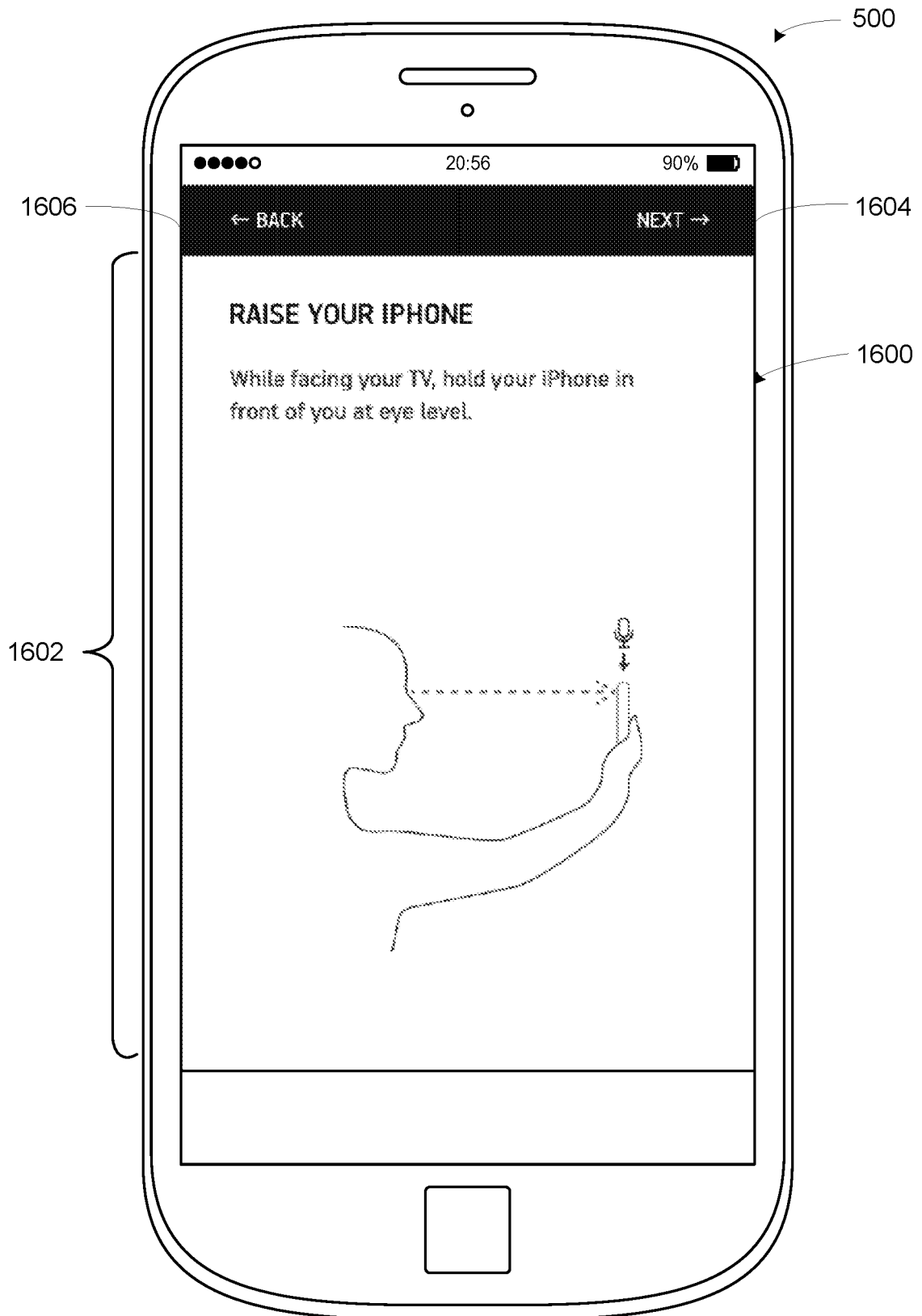


FIGURE 16

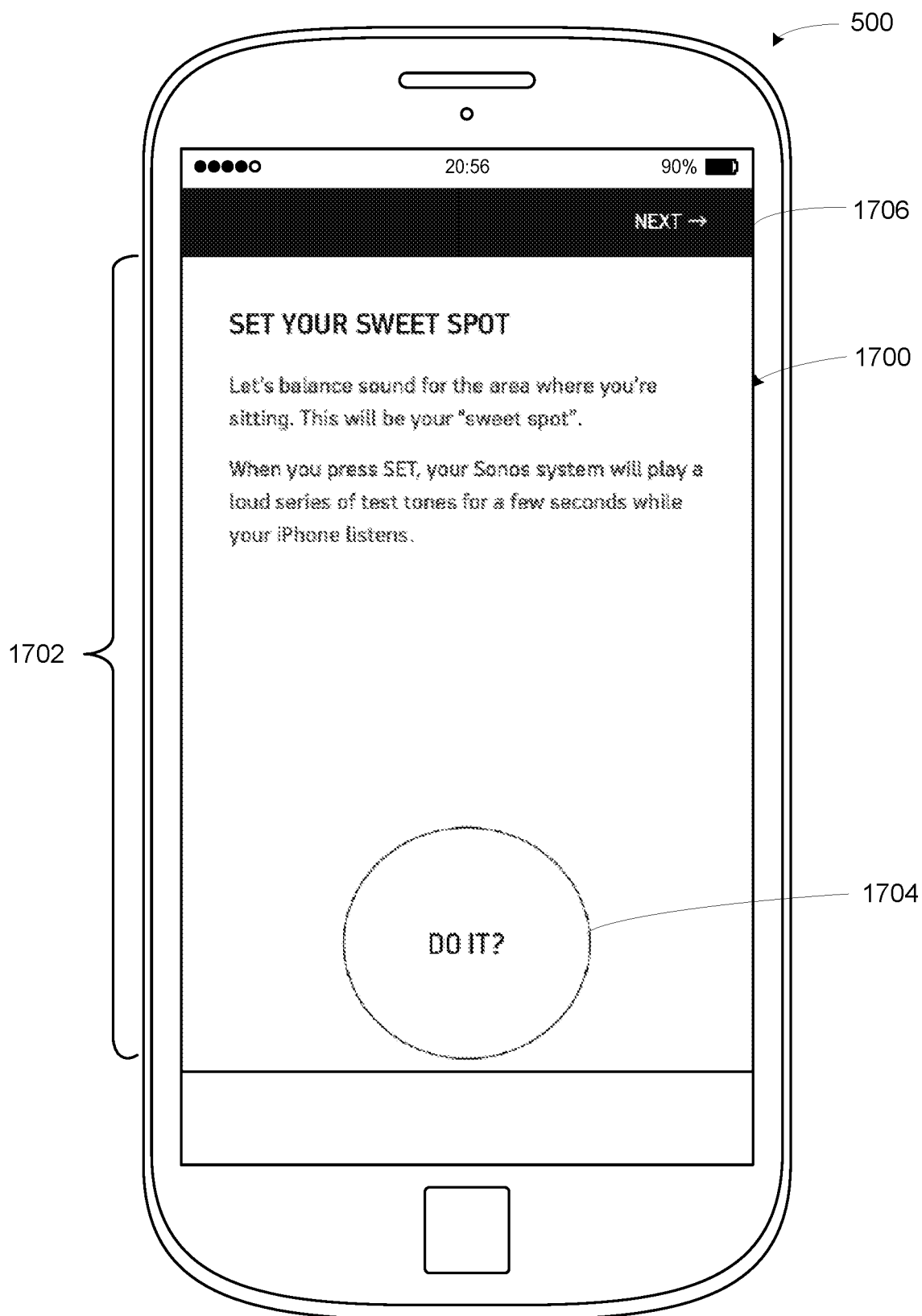


FIGURE 17

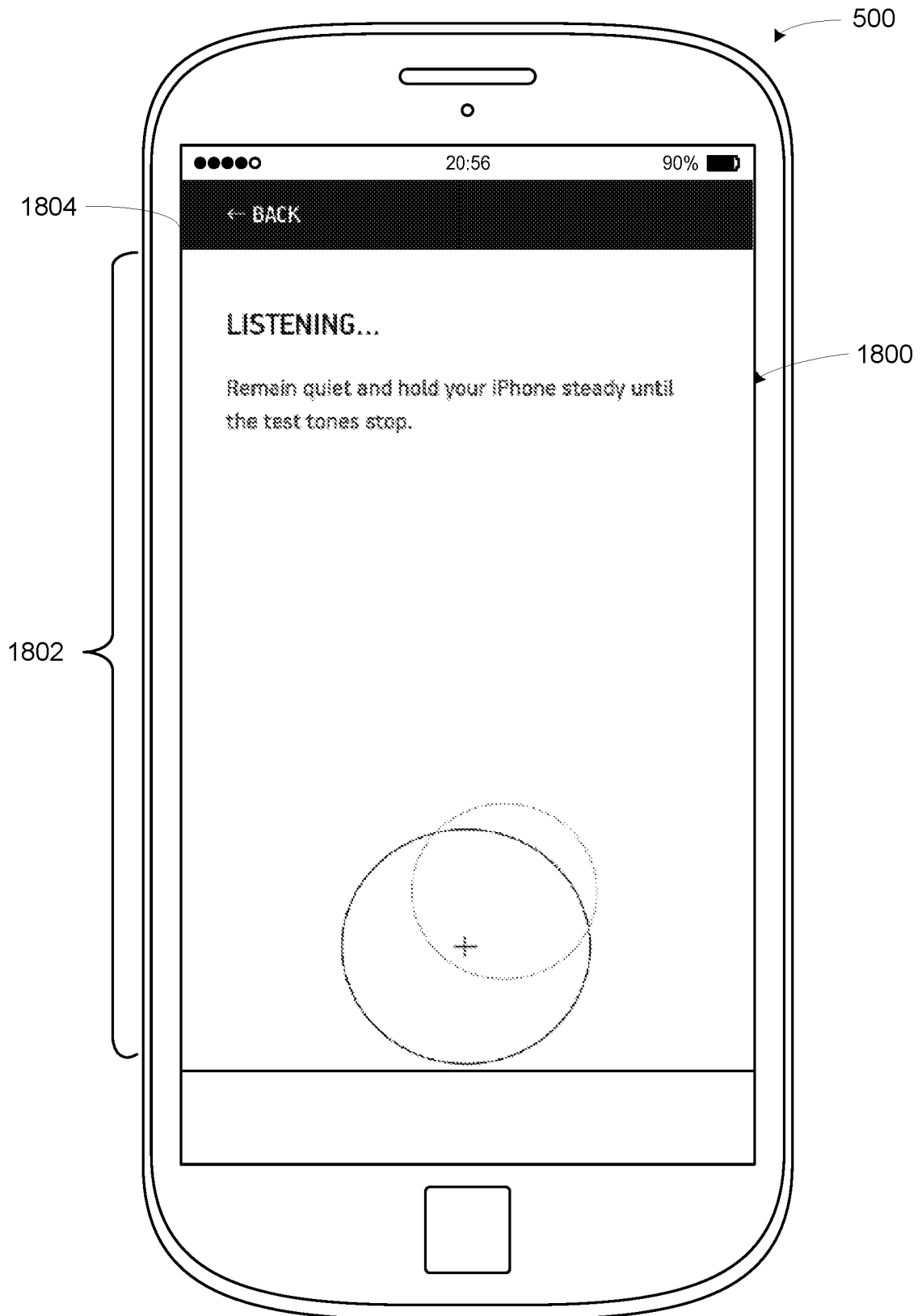


FIGURE 18



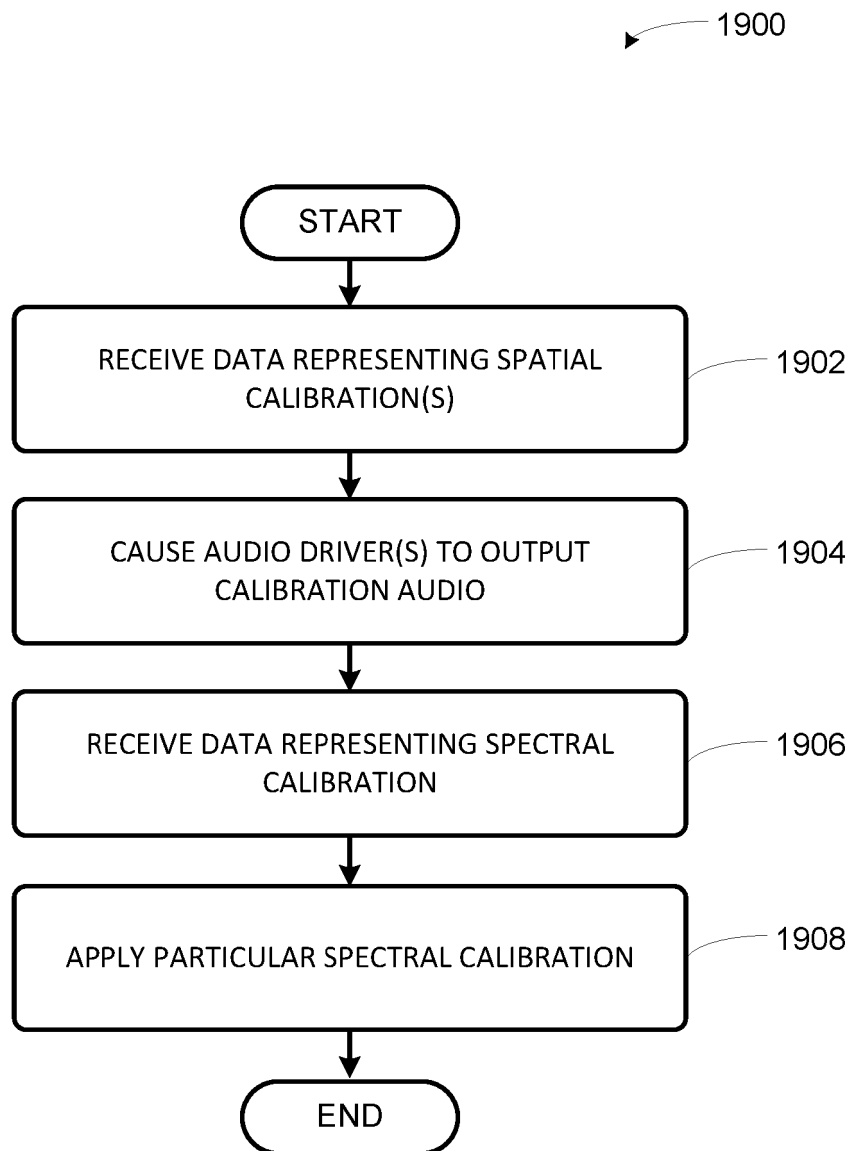


FIGURE 19

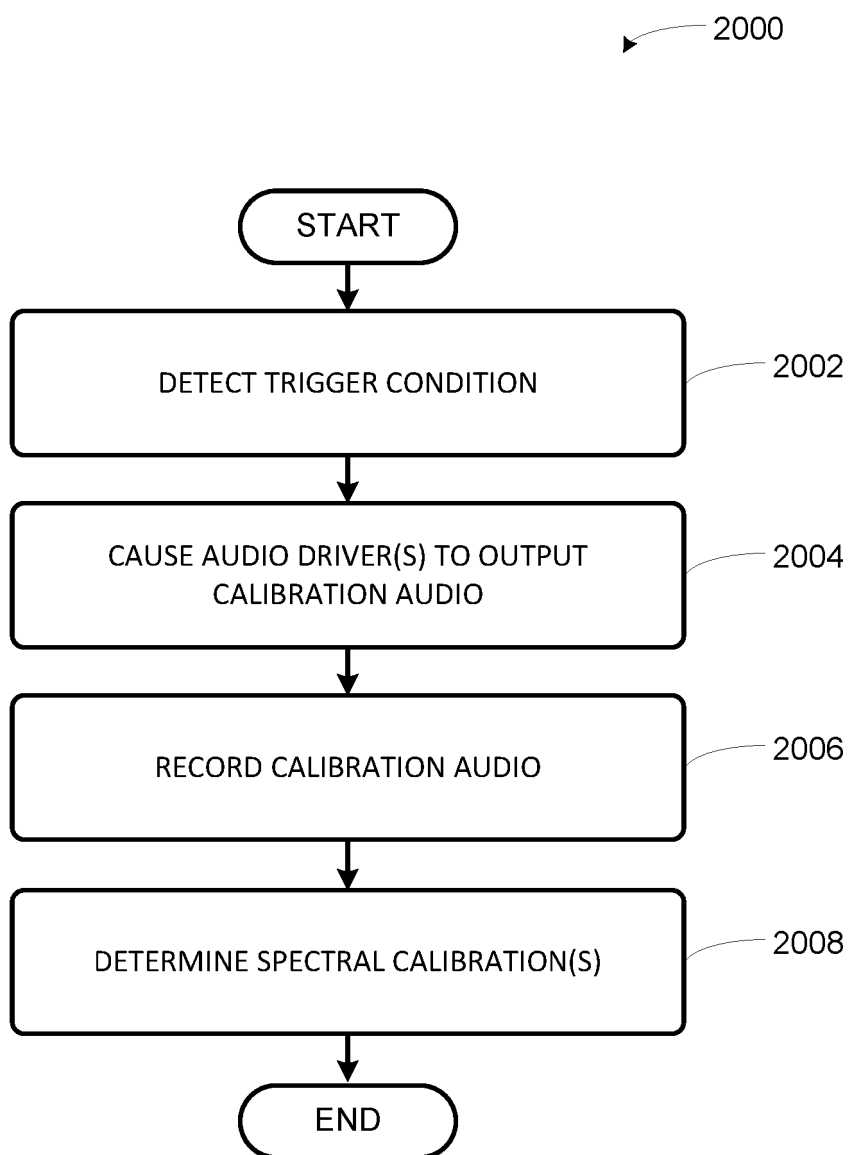


FIGURE 20

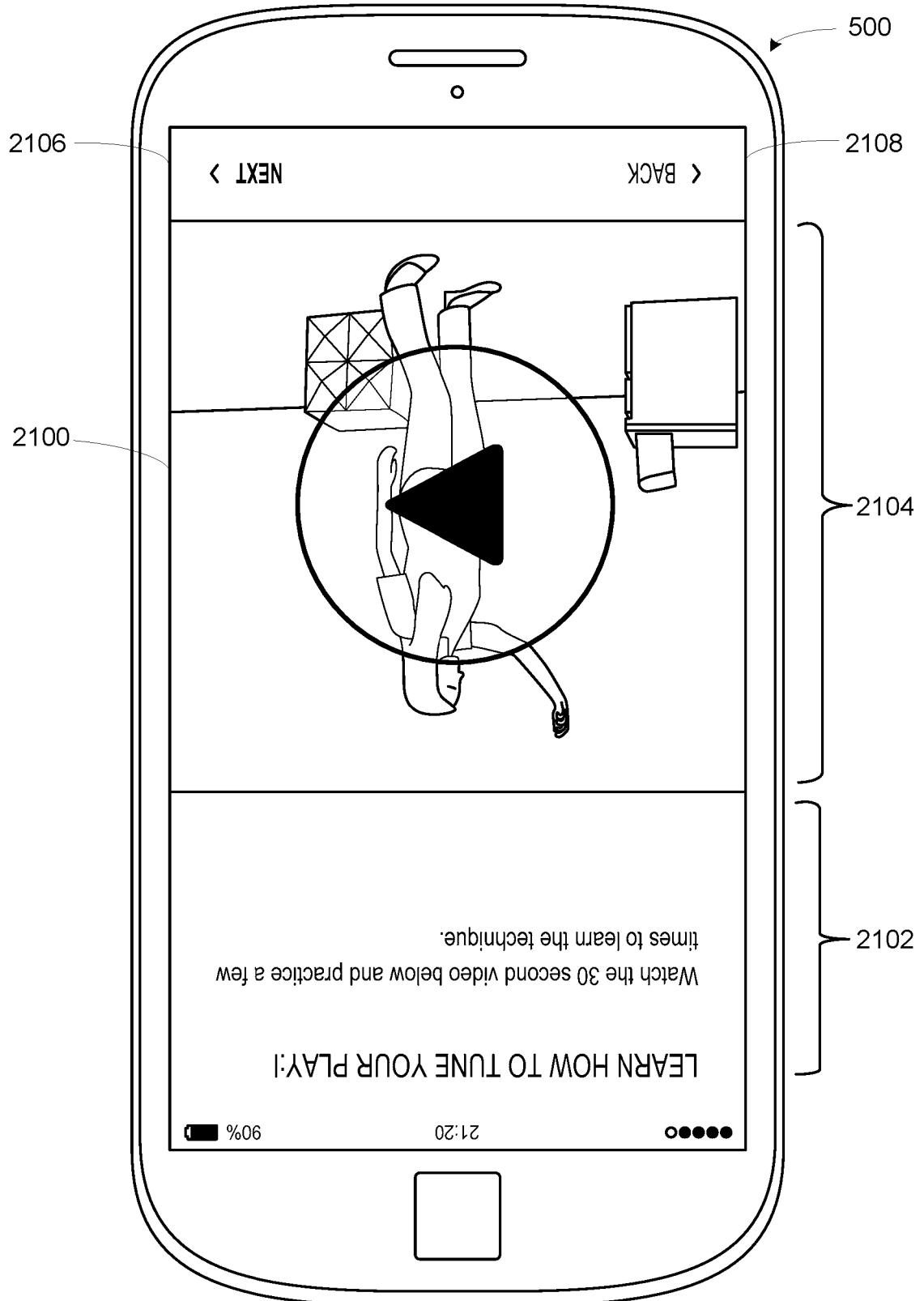


FIGURE 21

## REFERENCES CITED IN THE DESCRIPTION

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

### Patent documents cited in the description

- US 2016011850 A1 [0004]
- WO 2011139502 A1 [0005]
- US 2014003635 A1 [0006]
- US 8234395 B [0027] [0038]
- US 80514015 [0095]
- US 80534015 [0095]
- US 86439315 [0095]
- US 69601415 [0096]
- US 82687315 [0096]
- US 53649312 [0169]
- US 20140003625 A1 [0169]
- US 21630614 [0170]
- US 20150263692 A1 [0170]
- US 48151114 [0171]
- US 20160014534 A1 [0171]
- US 79319015 A [0181]
- US 79320515 [0181]