



(51) International Patent Classification:

H04R 27/00 (2006.01)

(21) International Application Number:

PCT/US2020/020511

(22) International Filing Date:

28 February 2020 (28.02.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/811,962 28 February 2019 (28.02.2019) US

(71) Applicant: SONOS, INC. [US/US]; 614 Chapala Street, Santa Barbara, CA 93101 (US).

(72) Inventors: WILBERDING, Dayn; 614 Chapala Street, Santa Barbara, CA 93101 (US). NAM, Youjin; 614 Chapala Street, Santa Barbara, CA 93101 (US). TOULIS, Tadeo, T.; 614 Chapala Street, Santa Barbara, CA 93101 (US). HARRIS, Cole; 614 Chapala Street, Santa Barbara, CA 93101 (US). DEVINE, Patrick; 614 Chapala Street, Santa Barbara, CA 93101 (US). MILLINGTON, Nicholas, A.J.; 614 Chapala Street, Santa Barbara, CA 93101 (US). RAMOS, Maria-Jose; 614 Chapala Street, Santa Barbara, CA 93101 (US). BANNON, Richard; 614 Chapala Street, Santa Barbara, CA 93101 (US). SHELDON, Michael; 614

Chapala Street, Santa Barbara, CA 93101 (US). LIN, Ted; 614 Chapala Street, Santa Barbara, CA 93101 (US). YORE, Jason; 614 Chapala Street, Santa Barbara, CA 93101 (US). KUPER, Ron; 614 Chapala Street, Santa Barbara, CA 93101 (US). VAUTRIN, Jodi; 614 Chapala Street, Santa Barbara, CA 93101 (US).

(74) Agent: URBAN, Benjamin, M.; McDonnell Boehnen Hulbert & Berghoff LLP, 300 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

(54) Title: PLAYBACK TRANSITIONS BETWEEN AUDIO DEVICES

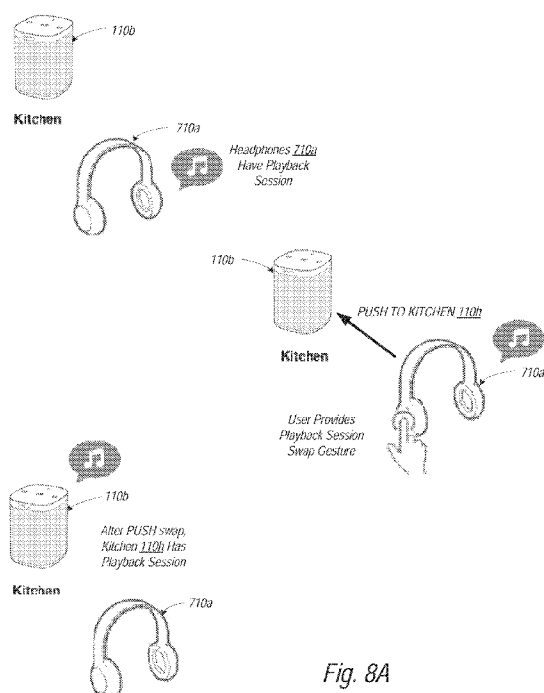


Fig. 8A

(57) Abstract: Examples described herein relate to transitioning a playback session between portable playback devices such as "smart" headphones, earbuds, and handheld speakers with playback devices of a zone-based media playback system. Exemplary techniques facilitate continuity of playback when transitioning between locations (e.g., from at home to on-the-go or vice versa) or between listening paradigms (e.g., personal or out-loud playback of audio content). An example implementation includes detecting a swap trigger, determining the source playback device(s) and target playback device(s), and performing a playback session swap between the source playback device(s) and target playback device(s).



EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

PLAYBACK TRANSITIONS BETWEEN AUDIO DEVICES

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit under 35 U.S.C. § 119 of U.S. provisional App. No. 62/811,962 filed on February 28, 2019, entitled "Playback Transitions," which is incorporated herein by reference in its entirety.

FIELD OF THE DISCLOSURE

[0002] The present 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.

BACKGROUND

[0003] Options for accessing and listening to digital audio in an out-loud setting were limited until in 2002, when SONOS, Inc. began development of a new type of playback system. Sonos then filed one of its first patent applications in 2003, entitled "Method for Synchronizing Audio Playback between Multiple Networked Devices," and began offering its first media playback systems for sale in 2005. The Sonos Wireless Home Sound System enables people to experience music from many sources via one or more networked playback devices. Through a software control application installed on a controller (e.g., smartphone, tablet, computer, voice input device), one can play what she wants in any room having a networked playback device. Media content (e.g., songs, podcasts, video sound) can be streamed to playback devices such that each room with a playback device can play back corresponding different media content. In addition, rooms can be grouped together for synchronous playback of the same media content, and/or the same media content can be heard in all rooms synchronously.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] Features, aspects, and advantages of the presently disclosed technology may be better understood with regard to the following description, appended claims, and accompanying drawings, as listed below. A person skilled in the relevant art will understand that the features shown in the drawings are for purposes of illustrations, and variations, including different and/or additional features and arrangements thereof, are possible.

[0005] Figure 1A is a partial cutaway view of an environment having a media playback system configured in accordance with aspects of the disclosed technology.

[0006] Figure 1B is a schematic diagram of the media playback system of Figure 1A and one or more networks.

[0007] Figure 1C is a block diagram of a playback device.

[0008] Figure 1D is a block diagram of a playback device.

[0009] Figure 1E is a block diagram of a network microphone device.

[0010] Figure 1F is a block diagram of a network microphone device.

[0011] Figure 1G is a block diagram of a playback device.

[0012] Figure 1H is a partially schematic diagram of a control device.

[0013] Figures 1-I, II, IK, and 1L are schematic diagrams of corresponding media playback system zones.

[0014] Figure 1M is a schematic diagram of media playback system areas.

[0015] Figure 2A is a front isometric view of a playback device configured in accordance with aspects of the disclosed technology.

[0016] Figure 2B is a front isometric view of the playback device of Figure 3A without a grille.

[0017] Figure 2C is an exploded view of the playback device of Figure 2A.

[0018] Figure 3A is a front view of a network microphone device configured in accordance with aspects of the disclosed technology.

[0019] Figure 3B is a side isometric view of the network microphone device of Figure 3A.

[0020] Figure 3C is an exploded view of the network microphone device of Figures 3A and 3B.

[0021] Figure 3D is an enlarged view of a portion of Figure 3B.

[0022] Figures 4A, 4B, 4C, and 4D are schematic diagrams of a control device in various stages of operation in accordance with aspects of the disclosed technology.

[0023] Figure 5 is front view of a control device.

[0024] Figure 6 is a message flow diagram of a media playback system.

[0025] Figure 7A is a partial cutaway view of an environment having a media playback system configured in accordance with aspects of the disclosed technology.

[0026] Figure 7B is a block diagram of a portable playback device configured in accordance with aspects of the disclosed technology.

[0027] Figure 7C is a front isometric view of a portable playback device implemented as headphones configured in accordance with aspects of the disclosed technology.

[0028] Figure 7D is a front isometric view of a portable playback device implemented as earbuds configured in accordance with aspects of the disclosed technology.

[0029] Figure 7E is a front isometric view of a portable playback device configured in accordance with aspects of the disclosed technology.

[0030] Figure 7F is a front isometric view of a portable playback device with a device base configured in accordance with aspects of the disclosed technology.

[0031] Figure 7G is a schematic diagram illustrating an example pairing configuration between a portable playback device and a control device.

[0032] Figure 8A is a schematic diagram illustrating an example push swap in accordance with aspects of the disclosed technology.

[0033] Figure 8B is a schematic diagram illustrating an example pull swap in accordance with aspects of the disclosed technology.

[0034] Figure 8C is a schematic diagram illustrating an example push swap in accordance with aspects of the disclosed technology.

[0035] Figure 9 is a schematic diagram illustrating an example audio-based identification technique in accordance with aspects of the disclosed technology.

[0036] Figure 10 is a schematic diagram illustrating example control schemes in accordance with aspects of the disclosed technology.

[0037] Figure 11 is a schematic diagram illustrating example feedback techniques in accordance with aspects of the disclosed technology.

[0038] Figures 12A and 12B are example messaging diagrams to illustrating example playback session swap techniques.

[0039] Figure 13A is a method flow diagram illustrating an example swap pull technique in accordance with aspects of the disclosed technology.

[0040] Figure 13B is a method flow diagram illustrating an example swap push technique in accordance with aspects of the disclosed technology.

[0041] Figure 14 is a method flow diagram illustrating an example home theater swap technique in accordance with aspects of the disclosed technology.

[0042] Figure 15 is a method flow diagram illustrating a technique to facilitate a playback session swap in accordance with aspects of the disclosed technology.

[0043] Figure 16A is a schematic diagram illustrating an example pairing configuration between a portable playback device and a bridge device.

[0044] Figure 16B is a block diagram of a bridge device configured in accordance with aspects of the disclosed technology.

[0045] Figure 16C is a front isometric view of a bridge device configured in accordance with aspects of the disclosed technology.

[0046] Figure 16D is a view of a touch-sensitive region implemented in a bridge device configured in accordance with aspects of the disclosed technology.

[0047] Figure 16E is a front view of a bridge device configured in accordance with aspects of the disclosed technology.

[0048] Figures 17A, 17B, 17C, 17D, 17E, and 17F are schematic diagrams of a bridge device user interface in various stages of operation in accordance with aspects of the disclosed technology.

[0049] Figure 18A is a view of an example arrangement between a bridge device and a device base.

[0050] Figure 18B is a view of an example arrangement between a portable playback device, a bridge device and a device base.

[0051] Figure 18C is a view of an example arrangement between a first portable playback device, a bridge device, a second portable playback device, and a device base.

[0052] Figures 19A, 19B, and 19C are schematic diagrams of a control device user interface in various stages of operation in accordance with aspects of the disclosed technology.

[0053] Figures 20A, 20B, and 20C are example messaging diagrams to illustrating example playback session swap techniques.

[0054] Figures 21A, 21B, and 21C are schematic diagrams of a control device user interface in various stages of operation in accordance with aspects of the disclosed technology.

[0055] Figures 22A, 22B, 22C, and 22D are diagrams illustrating example proximity-based playback session swaps.

[0056] Figure 23A is a front isometric view of earbuds configured in accordance with aspects of the disclosed technology.

[0057] Figure 23B is a bottom view of a charging case configured in accordance with aspects of the disclosed technology.

[0058] Figure 23C is a top view of the charging case.

[0059] Figure 23D is a first side view of the charging case.

[0060] Figure 23E is a second side view of the charging case.

[0061] Figure 23F is a front isometric view of earbuds illustrating exemplary arrangement with the charging case.

[0062] Figure 23G is an isometric view of the earbud.

[0063] Figure 23H is a first side view of the earbud.

- [0064] Figure 23I is a second side view of the earbud.
- [0065] Figure 23J is a third side view of the earbud.
- [0066] Figure 23K is a fourth side view of the earbud.
- [0067] Figure 23L is a fifth side view of the earbud.
- [0068] Figure 23M is a sixth side view of the earbud.
- [0069] Figure 24A is a front isometric view of a portable playback device implemented as a handheld speaker configured in accordance with aspects of the disclosed technology.
- [0070] Figure 24B is a side view of the portable playback device.
- [0071] Figure 24C is a top view of the portable playback device.
- [0072] Figure 24D is a bottom view of the portable playback device.
- [0073] Figure 24E is a front isometric view of the portable playback device illustrating exemplary arrangement with a device base.
- [0074] Figure 24F is a front isometric view of the portable playback device illustrating exemplary user inputs to the portable playback device.
- [0075] Figure 25A is a front view of headphones configured in accordance with aspects of the disclosed technology.
- [0076] Figure 25B is a first side view of the headphones.
- [0077] Figure 25C is a second side view of the headphones.
- [0078] Figure 26A is a front view of headphones configured in accordance with aspects of the disclosed technology.
- [0079] Figure 26B is a first side view of the headphones.
- [0080] Figure 26C is a second side view of the headphones.
- [0081] The drawings are for the purpose of illustrating example embodiments, but those of ordinary skill in the art will understand that the technology disclosed herein is not limited to the arrangements and/or instrumentality shown in the drawings.

DETAILED DESCRIPTION

I. Overview

[0082] Example techniques described herein relate to transitioning a playback session between wearable playback devices, such as “smart” headphones and earbuds, and playback devices of a zone-based media playback system. Further example techniques relate to transitioning a playback session between portable (e.g., battery-operated, carryable) playback devices and playback devices of a zone-based media playback system. Such transitions are referred to herein as “swaps” or “playback session swaps.” Such exemplary swap techniques facilitate continuity of playback when transitioning between locations (e.g., from at home to on-the-go or vice versa) or between listening paradigms (e.g., personal or out-loud). Further, some exemplary techniques may reduce the extent of user input (or other user involvement) involved with transitioning playback as compared with some other techniques.

[0083] In an illustrative example, a user begins a playback session on exemplary headphones while on-the-go. For instance, a user begins listening to KEXP Seattle using earbuds paired with a mobile device (e.g., a smartphone) over a wireless connection such as 802.15 (Bluetooth®) or 802.11, among other examples. In this example, KEXP radio is streamed via the Internet to the mobile device.

[0084] Upon arriving home, the user may desire to continue listening to KEXP radio out-loud. To initiate a playback session swap from the earbuds to a playback device within the kitchen, the user may provide an input to the earbuds. Since the earbuds have a playback session on-going, this input designates the earbuds as the source of the playback session swap. The target of the swap (i.e., the kitchen zone) may have been previously designated in a pre-defined swap pair with the earbuds or may be determined after the input using a proximity detection technique, such as audio chirps, as described in further detail herein. The earbuds and/or mobile device perform the playback session swap with the kitchen zone, and playback of KEXP radio continues uninterrupted out-loud on the playback device within the kitchen.

[0085] In another illustrative example, the user may begin a playback session on an exemplary portable speaker. For instance, a user begins listening to WBEZ Chicago using the handheld speaker in a dining room. In this example, WBEZ Chicago is streamed via the Internet to the handheld speaker over a home local area network. Desiring to meditate, the user carries the handheld speaker to the living room and asks a voice assistant service to play meditation music. The handheld speaker plays an acknowledgement from the voice assistant service and begins playing back a curated meditation playlist from a streaming audio service.

[0086] While playing the curated meditation playlist, the user's friend enters the living room and suggests that the user check out the new Childish Gambino track, which is playing via a control application on her smartphone. To initiate a playback session swap from the smartphone to the handheld speaker, the friend holds the smartphone close to the handheld speaker to initiate a Near-Field Communication (NFC) exchange between the smartphone and handheld speaker. This exchange designates the smartphone as the source of the playback session swap and the handheld speaker as the target. The smartphone performs the playback session swap with the handheld speaker, and playback of the Childish Gambino track continues uninterrupted out-loud on the handheld speaker.

[0087] To enjoy the Childish Gambino track with more powerful amplifier(s) and/or larger transducer(s), the user initiates a playback session swap from the handheld speaker to a playback device in the living room by providing an input to the handheld speaker. This input designates the handheld speaker as the source of the playback session swap. The handheld speaker automatically designates the living room zone as the swap target based on detected proximity of the handheld speaker to the living room zone. The handheld speaker performs the playback session swap with the living room zone, and playback of the Childish Gambino track continues uninterrupted out-loud on the playback device within the living room.

[0088] In a third illustrative example, in the evening, the user may begin a playback session in the bedroom on a soundbar device that plays back audio content from a television. Desiring to keep the volume down so as not to disturb her partner trying to put their baby to bed in an adjacent room, the user initiates a playback session swap from the soundbar device to the handheld speaker, which she has placed on her nightstand. Since the handheld speaker is physically closer to the user, she can comfortably hear the audio from the television at a lower volume level.

[0089] After putting the baby to sleep, the partner comes into the bedroom to find the user asleep. To initiate a playback session swap from the handheld speaker to a pair of headphones, the partner may provide an input to the headphones. This input designates the headphones as the target of the playback session swap. The source of the swap (i.e., the handheld speaker) is determined based on context (i.e., based on an active playback session). The handheld speaker performs the playback session swap with the headphones, and playback of the television audio continues uninterrupted out-loud on the headphones.

[0090] As noted above, example techniques described herein involve playback session swaps. An example implementation includes detecting a swap trigger, determining the source

playback device(s) and target playback device(s), and performing a playback session swap between the source playback device(s) and target playback device(s).

[0091] While some examples described herein may refer to functions performed by given actors such as “users,” “listeners,” and/or other entities, it should be understood that this 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.

[0092] Moreover, some functions are described herein as being performed “based on” or “in response to” another element or function. “Based on” should be understood that one element or function is related to another function or element. “In response to” should be understood that one element or function is a necessary result of another function or element. For the sake of brevity, functions are generally described as being based on another function when a functional link exists; however, such disclosure should be understood as disclosing either type of functional relationship.

[0093] In the Figures, identical reference numbers identify generally similar, and/or identical, elements. To facilitate the discussion of any particular element, the most significant digit or digits of a reference number refers to the Figure in which that element is first introduced. For example, element 110a is first introduced and discussed with reference to Figure 1A. Many of the details, dimensions, angles and other features shown in the Figures are merely illustrative of particular embodiments of the disclosed technology. Accordingly, other embodiments can have other details, dimensions, angles and features without departing from the spirit or scope of the disclosure. In addition, those of ordinary skill in the art will appreciate that further embodiments of the various disclosed technologies can be practiced without several of the details described below.

II. Suitable Operating Environment

[0094] Figure 1A is a partial cutaway view of a media playback system 100 distributed in an environment 101 (e.g., a house). The media playback system 100 includes one or more playback devices 110 (identified individually as playback devices 110a-n), one or more network microphone devices (“NMDs”), 120 (identified individually as NMDs 120a-c), and one or more control devices 130 (identified individually as control devices 130a and 130b).

[0095] As used herein the term “playback device” can generally refer to a network device configured to receive, process, and output data of a media playback system. For example, a playback device can be a network device that receives and processes audio content. In some embodiments, a playback device includes one or more transducers or speakers powered by one or more amplifiers. In other embodiments, however, a playback device includes one of (or

neither of) the speaker and the amplifier. For instance, a playback device can comprise one or more amplifiers configured to drive one or more speakers external to the playback device via a corresponding wire or cable.

[0096] Moreover, as used herein the term NMD (i.e., a “network microphone device”) can generally refer to a network device that is configured for audio detection. In some embodiments, an NMD is a stand-alone device configured primarily for audio detection. In other embodiments, an NMD is incorporated into a playback device (or vice versa).

[0097] The term “control device” can generally refer to a network device configured to perform functions relevant to facilitating user access, control, and/or configuration of the media playback system 100.

[0098] Each of the playback devices 110 is configured to receive audio signals or data from one or more media sources (e.g., one or more remote servers, one or more local devices) and play back the received audio signals or data as sound. The one or more NMDs 120 are configured to receive spoken word commands, and the one or more control devices 130 are configured to receive user input. In response to the received spoken word commands and/or user input, the media playback system 100 can play back audio via one or more of the playback devices 110. In certain embodiments, the playback devices 110 are configured to commence playback of media content in response to a trigger. For instance, one or more of the playback devices 110 can be configured to play back a morning playlist upon detection of an associated trigger condition (e.g., presence of a user in a kitchen, detection of a coffee machine operation). In some embodiments, for example, the media playback system 100 is configured to play back audio from a first playback device (e.g., the playback device 100a) in synchrony with a second playback device (e.g., the playback device 100b). Interactions between the playback devices 110, NMDs 120, and/or control devices 130 of the media playback system 100 configured in accordance with the various embodiments of the disclosure are described in greater detail below with respect to Figures 1B-6.

[0099] In the illustrated embodiment of Figure 1A, the environment 101 includes a household having several rooms, spaces, and/or playback zones, including (clockwise from upper left) a master bathroom 101a, a master bedroom 101b, a second bedroom 101c, a family room or den 101d, an office 101e, a living room 101f, a dining room 101g, a kitchen 101h, and an outdoor patio 101i. While certain embodiments and examples are described below in the context of a home environment, the technologies described herein may be implemented in other types of environments. In some embodiments, for example, the media playback system 100 can be implemented in one or more commercial settings (e.g., a restaurant, mall, airport, hotel,

a retail or other store), one or more vehicles (e.g., a sports utility vehicle, bus, car, a ship, a boat, an airplane), multiple environments (e.g., a combination of home and vehicle environments), and/or another suitable environment where multi-zone audio may be desirable.

[0100] The media playback system 100 can comprise one or more playback zones, some of which may correspond to the rooms in the environment 101. The media playback system 100 can be established with one or more playback zones, after which additional zones may be added, or removed to form, for example, the configuration shown in Figure 1A. Each zone may be given a name according to a different room or space such as the office 101e, master bathroom 101a, master bedroom 101b, the second bedroom 101c, kitchen 101h, dining room 101g, living room 101f, and/or the balcony 101i. In some aspects, a single playback zone may include multiple rooms or spaces. In certain aspects, a single room or space may include multiple playback zones.

[0101] In the illustrated embodiment of Figure 1A, the master bathroom 101a, the second bedroom 101c, the office 101e, the living room 101f, the dining room 101g, the kitchen 101h, and the outdoor patio 101i each include one playback device 110, and the master bedroom 101b and the den 101d include a plurality of playback devices 110. In the master bedroom 101b, the playback devices 110l and 110m may be configured, for example, to play back audio content in synchrony as individual ones of playback devices 110, as a bonded playback zone, as a consolidated playback device, and/or any combination thereof. Similarly, in the den 101d, the playback devices 110h-j can be configured, for instance, to play back audio content in synchrony as individual ones of playback devices 110, as one or more bonded playback devices, and/or as one or more consolidated playback devices. Additional details regarding bonded and consolidated playback devices are described below with respect to Figures 1B and 1E, as well as Figures 1-I-1M.

[0102] In some aspects, one or more of the playback zones in the environment 101 may each be playing different audio content. For instance, a user may be grilling on the patio 101i and listening to hip hop music being played by the playback device 110c while another user is preparing food in the kitchen 101h and listening to classical music played by the playback device 110b. 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 101e listening to the playback device 110f playing back the same hip hop music being played back by playback device 110c on the patio 101i. In some aspects, the playback devices 110c and 110f play back the hip hop music in synchrony such that the user perceives that the audio

content is being played seamlessly (or at least substantially seamlessly) while moving between different playback zones.

a. Suitable Media Playback System

[0103] Figure 1B is a schematic diagram of the media playback system 100 and a cloud network 102. For ease of illustration, certain devices of the media playback system 100 and the cloud network 102 are omitted from Figure 1B. One or more communication links 103 (referred to hereinafter as “the links 103”) communicatively couple the media playback system 100 and the cloud network 102.

[0104] The links 103 can comprise, for example, one or more wired networks, one or more wireless networks, one or more wide area networks (WAN), one or more local area networks (LAN), one or more personal area networks (PAN), one or more telecommunication networks (e.g., one or more Global System for Mobiles (GSM) networks, Code Division Multiple Access (CDMA) networks, Long-Term Evolution (LTE) networks, 5G communication network networks, and/or other suitable data transmission protocol networks), etc. The cloud network 102 is configured to deliver media content (e.g., audio content, video content, photographs, social media content) to the media playback system 100 in response to a request transmitted from the media playback system 100 via the links 103. In some embodiments, the cloud network 102 is further configured to receive data (e.g. voice input data) from the media playback system 100 and correspondingly transmit commands and/or media content to the media playback system 100.

[0105] The cloud network 102 includes computing devices 106 (identified separately as a first computing device 106a, a second computing device 106b, and a third computing device 106c). The computing devices 106 can comprise individual computers or servers, such as, for example, a media streaming service server storing audio and/or other media content, a voice service server, a social media server, a media playback system control server, etc. In some embodiments, one or more of the computing devices 106 comprise modules of a single computer or server. In certain embodiments, one or more of the computing devices 106 comprise one or more modules, computers, and/or servers. Moreover, while the cloud network 102 is described above in the context of a single cloud network, in some embodiments the cloud network 102 includes a plurality of cloud networks comprising communicatively coupled computing devices. Furthermore, while the cloud network 102 is shown in Figure 1B as having three of the computing devices 106, in some embodiments, the cloud network 102 includes fewer (or more than) three computing devices 106.

[0106] The media playback system 100 is configured to receive media content from the networks 102 via the links 103. The received media content can comprise, for example, a Uniform Resource Identifier (URI) and/or a Uniform Resource Locator (URL). For instance, in some examples, the media playback system 100 can stream, download, or otherwise obtain data from a URI or a URL corresponding to the received media content. A network 104 communicatively couples the links 103 and at least a portion of the devices (e.g., one or more of the playback devices 110, NMDs 120, and/or control devices 130) of the media playback system 100. The network 104 can include, for example, a wireless network (e.g., a WiFi network, a Bluetooth, a Z-Wave network, a ZigBee, and/or other suitable wireless communication protocol network) and/or a wired network (e.g., a network comprising Ethernet, Universal Serial Bus (USB), and/or another suitable wired communication). As those of ordinary skill in the art will appreciate, as used herein, “WiFi” can refer to several different communication protocols including, for example, Institute of Electrical and Electronics Engineers (IEEE) 802.11a, 802.11b, 802.11g, 802.11n, 802.11ac, 802.11ac, 802.11ad, 802.11af, 802.11ah, 802.11ai, 802.11aj, 802.11aq, 802.11ax, 802.11ay, 802.15, etc. transmitted at 2.4 Gigahertz (GHz), 5 GHz, and/or another suitable frequency.

[0107] In some embodiments, the network 104 includes a dedicated communication network that the media playback system 100 uses to transmit messages between individual devices and/or to transmit media content to and from media content sources (e.g., one or more of the computing devices 106). In certain embodiments, the network 104 is configured to be accessible only to devices in the media playback system 100, thereby reducing interference and competition with other household devices. In some examples, the dedicated communication network is implemented as a mesh network where the devices in the media playback system form the nodes in the mesh network. One or more root nodes of the mesh network then connect the mesh network to a household WiFi network, which functions in parallel to the mesh network.

[0108] In other embodiments, however, the network 104 includes an existing household communication network (e.g., a household WiFi network). In some embodiments, the links 103 and the network 104 comprise one or more of the same networks. In some aspects, for example, the links 103 and the network 104 comprise a telecommunication network (e.g., an LTE network, a 5G network). Moreover, in some embodiments, the media playback system 100 is implemented without the network 104, and devices comprising the media playback system 100 can communicate with each other, for example, via one or more direct connections, PANs, telecommunication networks, and/or other suitable communication links.

[0109] In some embodiments, audio content sources may be regularly added or removed from the media playback system 100. In some embodiments, for example, the media playback system 100 performs an indexing of media items when one or more media content sources are updated, added to, and/or removed from the media playback system 100. The media playback system 100 can scan identifiable media items in some or all folders and/or directories accessible to the playback devices 110, and generate or update a media content database comprising metadata (e.g., title, artist, album, track length) and other associated information (e.g., URIs, URLs) for each identifiable media item found. In some embodiments, for example, the media content database is stored on one or more of the playback devices 110, network microphone devices 120, and/or control devices 130.

[0110] In the illustrated embodiment of Figure 1B, the playback devices 110l and 110m comprise a group 107a. The playback devices 110l and 110m can be positioned in different rooms in a household and be grouped together in the group 107a on a temporary or permanent basis based on user input received at the control device 130a and/or another control device 130 in the media playback system 100. When arranged in the group 107a, the playback devices 110l and 110m can be configured to play back the same or similar audio content in synchrony from one or more audio content sources. In certain embodiments, for example, the group 107a includes a bonded zone in which the playback devices 110l and 110m comprise left audio and right audio channels, respectively, of multi-channel audio content, thereby producing or enhancing a stereo effect of the audio content. In some embodiments, the group 107a includes additional playback devices 110. In other embodiments, however, the media playback system 100 omits the group 107a and/or other grouped arrangements of the playback devices 110. Additional details regarding groups and other arrangements of playback devices are described in further detail below with respect to Figures 1-I through 1M.

[0111] The media playback system 100 includes the NMDs 120a and 120d, each comprising one or more microphones configured to receive voice utterances from a user. In the illustrated embodiment of Figure 1B, the NMD 120a is a standalone device and the NMD 120d is integrated into the playback device 110n. The NMD 120a, for example, is configured to receive voice input 121 from a user 123. In some embodiments, the NMD 120a transmits data associated with the received voice input 121 to a voice assistant service (VAS) configured to (i) process the received voice input data and (ii) transmit a corresponding command to the media playback system 100. In some aspects, for example, the computing device 106c includes one or more modules and/or servers of a VAS (e.g., a VAS operated by one or more of SONOS®, AMAZON®, GOOGLE®, APPLE®, MICROSOFT®). The computing device 106c

can receive the voice input data from the NMD 120a via the network 104 and the links 103. In response to receiving the voice input data, the computing device 106c processes the voice input data (i.e., “Play Hey Jude by The Beatles”), and determines that the processed voice input includes a command to play a song (e.g., “Hey Jude”). The computing device 106c accordingly transmits commands to the media playback system 100 to play back “Hey Jude” by the Beatles from a suitable media service (e.g., via one or more of the computing devices 106) on one or more of the playback devices 110.

b. Suitable Playback Devices

[0112] Figure 1C is a block diagram of the playback device 110a comprising an input/output 111. The input/output 111 can include an analog I/O 111a (e.g., one or more wires, cables, and/or other suitable communication links configured to carry analog signals) and/or a digital I/O 111b (e.g., one or more wires, cables, or other suitable communication links configured to carry digital signals). In some embodiments, the analog I/O 111a is an audio line-in input connection comprising, for example, an auto-detecting 3.5mm audio line-in connection. In some embodiments, the digital I/O 111b includes a Sony/Philips Digital Interface Format (S/PDIF) communication interface and/or cable and/or a Toshiba Link (TOSLINK) cable. In some embodiments, the digital I/O 111b includes an High-Definition Multimedia Interface (HDMI) interface and/or cable. In some embodiments, the digital I/O 111b includes one or more wireless communication links comprising, for example, a radio frequency (RF), infrared, WiFi, Bluetooth, or another suitable communication protocol. In certain embodiments, the analog I/O 111a and the digital I/O 111b comprise interfaces (e.g., ports, plugs, jacks) configured to receive connectors of cables transmitting analog and digital signals, respectively, without necessarily including cables.

[0113] The playback device 110a, for example, can receive media content (e.g., audio content comprising music and/or other sounds) from a local audio source 105 via the input/output 111 (e.g., a cable, a wire, a PAN, a Bluetooth connection, an ad hoc wired or wireless communication network, and/or another suitable communication link). The local audio source 105 can comprise, for example, a mobile device (e.g., a smartphone, a tablet, a laptop computer) or another suitable audio component (e.g., a television, a desktop computer, an amplifier, a phonograph, a Blu-ray player, a memory storing digital media files). In some aspects, the local audio source 105 includes local music libraries on a smartphone, a computer, a networked-attached storage (NAS), and/or another suitable device configured to store media files. In certain embodiments, one or more of the playback devices 110, NMDs 120, and/or control devices 130 comprise the local audio source 105. In other embodiments, however, the

media playback system omits the local audio source 105 altogether. In some embodiments, the playback device 110a does not include an input/output 111 and receives all audio content via the network 104.

[0114] The playback device 110a further includes electronics 112, a user interface 113 (e.g., one or more buttons, knobs, dials, touch-sensitive surfaces, displays, touchscreens), and one or more transducers 114 (referred to hereinafter as “the transducers 114”). The electronics 112 is configured to receive audio from an audio source (e.g., the local audio source 105) via the input/output 111, one or more of the computing devices 106a-c via the network 104 (Figure 1B)), amplify the received audio, and output the amplified audio for playback via one or more of the transducers 114. In some embodiments, the playback device 110a optionally includes one or more microphones 115 (e.g., a single microphone, a plurality of microphones, a microphone array) (hereinafter referred to as “the microphones 115”). In certain embodiments, for example, the playback device 110a having one or more of the optional microphones 115 can operate as an NMD configured to receive voice input from a user and correspondingly perform one or more operations based on the received voice input.

[0115] In the illustrated embodiment of Figure 1C, the electronics 112 comprise one or more processors 112a (referred to hereinafter as “the processors 112a”), memory 112b, software components 112c, a network interface 112d, one or more audio processing components 112g (referred to hereinafter as “the audio components 112g”), one or more audio amplifiers 112h (referred to hereinafter as “the amplifiers 112h”), and power 112i (e.g., one or more power supplies, power cables, power receptacles, batteries, induction coils, Power-over Ethernet (POE) interfaces, and/or other suitable sources of electric power). In some embodiments, the electronics 112 optionally include one or more other components 112j (e.g., one or more sensors, video displays, touchscreens).

[0116] The processors 112a can comprise clock-driven computing component(s) configured to process data, and the memory 112b can comprise a computer-readable medium (e.g., a tangible, non-transitory computer-readable medium, data storage loaded with one or more of the software components 112c) configured to store instructions for performing various operations and/or functions. The processors 112a are configured to execute the instructions stored on the memory 112b to perform one or more of the operations. The operations can include, for example, causing the playback device 110a to retrieve audio data from an audio source (e.g., one or more of the computing devices 106a-c (Figure 1B)), and/or another one of the playback devices 110. In some embodiments, the operations further include causing the playback device 110a to send audio data to another one of the playback devices 110a and/or

another device (e.g., one of the NMDs 120). Certain embodiments include operations causing the playback device 110a to pair with another of the one or more playback devices 110 to enable a multi-channel audio environment (e.g., a stereo pair, a bonded zone).

[0117] The processors 112a can be further configured to perform operations causing the playback device 110a to synchronize playback of audio content with another of the one or more playback devices 110. As those of ordinary skill in the art will appreciate, during synchronous playback of audio content on a plurality of playback devices, a listener will preferably be unable to perceive time-delay differences between playback of the audio content by the playback device 110a and the other one or more other playback devices 110. Additional details regarding audio playback synchronization among playback devices can be found, for example, in U.S. Patent No. 8,234,395, which was incorporated by reference above.

[0118] In some embodiments, the memory 112b is further configured to store data associated with the playback device 110a, such as one or more zones and/or zone groups of which the playback device 110a is a member, audio sources accessible to the playback device 110a, and/or a playback queue that the playback device 110a (and/or another of the one or more playback devices) can be associated with. The stored data can comprise one or more state variables that are periodically updated and used to describe a state of the playback device 110a. The memory 112b can also include data associated with a state of one or more of the other devices (e.g., the playback devices 110, NMDs 120, control devices 130) of the media playback system 100. In some aspects, for example, the state data is shared during predetermined intervals of time (e.g., every 5 seconds, every 10 seconds, every 60 seconds) among at least a portion of the devices of the media playback system 100, so that one or more of the devices have the most recent data associated with the media playback system 100.

[0119] The network interface 112d is configured to facilitate a transmission of data between the playback device 110a and one or more other devices on a data network such as, for example, the links 103 and/or the network 104 (Figure 1B). The network interface 112d is configured to transmit and receive data corresponding to media content (e.g., audio content, video content, text, photographs) and other signals (e.g., non-transitory signals) comprising digital packet data including an Internet Protocol (IP)-based source address and/or an IP-based destination address. The network interface 112d can parse the digital packet data such that the electronics 112 properly receives and processes the data destined for the playback device 110a.

[0120] In the illustrated embodiment of Figure 1C, the network interface 112d includes one or more wireless interfaces 112e (referred to hereinafter as “the wireless interface 112e”). The wireless interface 112e (e.g., a suitable interface comprising one or more antennae) can be

configured to wirelessly communicate with one or more other devices (e.g., one or more of the other playback devices 110, NMDs 120, and/or control devices 130) that are communicatively coupled to the network 104 (Figure 1B) in accordance with a suitable wireless communication protocol (e.g., WiFi, Bluetooth, LTE). In some embodiments, the network interface 112d optionally includes a wired interface 112f (e.g., an interface or receptacle configured to receive a network cable such as an Ethernet, a USB-A, USB-C, and/or Thunderbolt cable) configured to communicate over a wired connection with other devices in accordance with a suitable wired communication protocol. In certain embodiments, the network interface 112d includes the wired interface 112f and excludes the wireless interface 112e. In some embodiments, the electronics 112 excludes the network interface 112d altogether and transmits and receives media content and/or other data via another communication path (e.g., the input/output 111).

[0121] The audio components 112g are configured to process and/or filter data comprising media content received by the electronics 112 (e.g., via the input/output 111 and/or the network interface 112d) to produce output audio signals. In some embodiments, the audio processing components 112g comprise, for example, one or more digital-to-analog converters (DAC), audio preprocessing components, audio enhancement components, a digital signal processors (DSPs), and/or other suitable audio processing components, modules, circuits, etc. In certain embodiments, one or more of the audio processing components 112g can comprise one or more subcomponents of the processors 112a. In some embodiments, the electronics 112 omits the audio processing components 112g. In some aspects, for example, the processors 112a execute instructions stored on the memory 112b to perform audio processing operations to produce the output audio signals.

[0122] The amplifiers 112h are configured to receive and amplify the audio output signals produced by the audio processing components 112g and/or the processors 112a. The amplifiers 112h can comprise electronic devices and/or components configured to amplify audio signals to levels sufficient for driving one or more of the transducers 114. In some embodiments, for example, the amplifiers 112h include one or more switching or class-D power amplifiers. In other embodiments, however, the amplifiers include one or more other types of power amplifiers (e.g., linear gain power amplifiers, class-A amplifiers, class-B amplifiers, class-AB amplifiers, class-C amplifiers, class-D amplifiers, class-E amplifiers, class-F amplifiers, class-G and/or class H amplifiers, and/or another suitable type of power amplifier). In certain embodiments, the amplifiers 112h comprise a suitable combination of two or more of the foregoing types of power amplifiers. Moreover, in some embodiments, individual ones of the amplifiers 112h correspond to individual ones of the transducers 114. In other embodiments,

however, the electronics 112 includes a single one of the amplifiers 112h configured to output amplified audio signals to a plurality of the transducers 114. In some other embodiments, the electronics 112 omits the amplifiers 112h.

[0123] The transducers 114 (e.g., one or more speakers and/or speaker drivers) receive the amplified audio signals from the amplifier 112h and render or output the amplified audio signals as sound (e.g., audible sound waves having a frequency between about 20 Hertz (Hz) and 20 kilohertz (kHz)). In some embodiments, the transducers 114 can comprise a single transducer. In other embodiments, however, the transducers 114 comprise a plurality of audio transducers. In some embodiments, the transducers 114 comprise more than one type of transducer. For example, the transducers 114 can include one or more low frequency transducers (e.g., subwoofers, woofers), mid-range frequency transducers (e.g., mid-range transducers, mid-woofers), and one or more high frequency transducers (e.g., one or more tweeters). As used herein, “low frequency” can generally refer to audible frequencies below about 500 Hz, “mid-range frequency” can generally refer to audible frequencies between about 500 Hz and about 2 kHz, and “high frequency” can generally refer to audible frequencies above 2 kHz. In certain embodiments, however, one or more of the transducers 114 comprise transducers that do not adhere to the foregoing frequency ranges. For example, one of the transducers 114 may comprise a mid-woofer transducer configured to output sound at frequencies between about 200 Hz and about 5 kHz.

[0124] By way of illustration, SONOS, Inc. presently offers (or has offered) for sale certain playback devices including, for example, a “SONOS ONE,” “PLAY:1,” “PLAY:3,” “PLAY:5,” “PLAYBAR,” “PLAYBASE,” “CONNECT:AMP,” “CONNECT,” and “SUB.” Other suitable playback devices may additionally or alternatively be used to implement the playback devices of example embodiments disclosed herein. Additionally, one of ordinary skill in the art will appreciate that a playback device is not limited to the examples described herein or to SONOS product offerings. In some embodiments, for example, one or more playback devices 110 includes wired or wireless headphones (e.g., over-the-ear headphones, on-ear headphones, in-ear earphones). In other embodiments, one or more of the playback devices 110 comprise a docking station and/or an interface configured to interact with a docking station for personal mobile media playback devices. In certain embodiments, 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. In some embodiments, a playback device omits a user interface and/or one or more transducers. For example, FIG. 1D is a block

diagram of a playback device 110p comprising the input/output 111 and electronics 112 without the user interface 113 or transducers 114.

[0125] Figure 1E is a block diagram of a bonded playback device 110q comprising the playback device 110a (Figure 1C) sonically bonded with the playback device 110i (e.g., a subwoofer) (Figure 1A). In the illustrated embodiment, the playback devices 110a and 110i are separate ones of the playback devices 110 housed in separate enclosures. In some embodiments, however, the bonded playback device 110q includes a single enclosure housing both the playback devices 110a and 110i. The bonded playback device 110q can be configured to process and reproduce sound differently than an unbonded playback device (e.g., the playback device 110a of Figure 1C) and/or paired or bonded playback devices (e.g., the playback devices 110l and 110m of Figure 1B). In some embodiments, for example, the playback device 110a is full-range playback device configured to render low frequency, mid-range frequency, and high frequency audio content, and the playback device 110i is a subwoofer configured to render low frequency audio content. In some aspects, the playback device 110a, when bonded with the first playback device, is configured to render only the mid-range and high frequency components of a particular audio content, while the playback device 110i renders the low frequency component of the particular audio content. In some embodiments, the bonded playback device 110q includes additional playback devices and/or another bonded playback device. Additional playback device embodiments are described in further detail below with respect to Figures 2A-3D.

c. Suitable Network Microphone Devices (NMDs)

[0126] Figure 1F is a block diagram of the NMD 120a (Figures 1A and 1B). The NMD 120a includes one or more voice processing components 124 (hereinafter “the voice components 124”) and several components described with respect to the playback device 110a (Figure 1C) including the processors 112a, the memory 112b, and the microphones 115. The NMD 120a optionally includes other components also included in the playback device 110a (Figure 1C), such as the user interface 113 and/or the transducers 114. In some embodiments, the NMD 120a is configured as a media playback device (e.g., one or more of the playback devices 110), and further includes, for example, one or more of the audio components 112g (Figure 1C), the amplifiers 114, and/or other playback device components. In certain embodiments, the NMD 120a includes an Internet of Things (IoT) device such as, for example, a thermostat, alarm panel, fire and/or smoke detector, etc. In some embodiments, the NMD 120a includes the microphones 115, the voice processing 124, and only a portion of the components of the electronics 112 described above with respect to Figure 1B. In some aspects, for example, the

NMD 120a includes the processor 112a and the memory 112b (Figure 1B), while omitting one or more other components of the electronics 112. In some embodiments, the NMD 120a includes additional components (e.g., one or more sensors, cameras, thermometers, barometers, hygrometers).

[0127] In some embodiments, an NMD can be integrated into a playback device. Figure 1G is a block diagram of a playback device 110r comprising an NMD 120d. The playback device 110r can comprise many or all of the components of the playback device 110a and further include the microphones 115 and voice processing 124 (Figure 1F). The playback device 110r optionally includes an integrated control device 130c. The control device 130c can comprise, for example, a user interface (e.g., the user interface 113 of Figure 1B) configured to receive user input (e.g., touch input, voice input) without a separate control device. In other embodiments, however, the playback device 110r receives commands from another control device (e.g., the control device 130a of Figure 1B).

[0128] Referring again to Figure 1F, the microphones 115 are configured to acquire, capture, and/or receive sound from an environment (e.g., the environment 101 of Figure 1A) and/or a room in which the NMD 120a is positioned. The received sound can include, for example, vocal utterances, audio played back by the NMD 120a and/or another playback device, background voices, ambient sounds, etc. The microphones 115 convert the received sound into electrical signals to produce microphone data. The voice processing 124 receives and analyzes the microphone data to determine whether a voice input is present in the microphone data. The voice input can comprise, for example, an activation word followed by an utterance including a user request. As those of ordinary skill in the art will appreciate, an activation word is a word or other audio cue that signifying a user voice input. For instance, in querying the AMAZON® VAS, a user might speak the activation word "Alexa." Other examples include "Ok, Google" for invoking the GOOGLE® VAS and "Hey, Siri" for invoking the APPLE® VAS.

[0129] After detecting the activation word, voice processing 124 monitors the microphone data for an accompanying user request in the voice input. The user request may include, for example, a command to control a third-party device, such as a thermostat (e.g., NEST® thermostat), an illumination device (e.g., a PHILIPS HUE ® lighting device), or a media playback device (e.g., a Sonos® playback device). For example, a user might speak the activation word "Alexa" followed by the utterance "set the thermostat to 68 degrees" to set a temperature in a home (e.g., the environment 101 of Figure 1A). The user might speak the same activation word followed by the utterance "turn on the living room" to turn on illumination devices in a living room area of the home. The user may similarly speak an activation word

followed by a request to play a particular song, an album, or a playlist of music on a playback device in the home. Additional description regarding receiving and processing voice input data can be found in further detail below with respect to Figures 3A-3F.

d. Suitable Control Devices

[0130] Figure 1H is a partially schematic diagram of the control device 130a (Figures 1A and 1B). As used herein, the term “control device” can be used interchangeably with “controller” or “control system.” Among other features, the control device 130a is configured to receive user input related to the media playback system 100 and, in response, cause one or more devices in the media playback system 100 to perform an action(s) or operation(s) corresponding to the user input. In the illustrated embodiment, the control device 130a includes a smartphone (e.g., an iPhone™, an Android phone) on which media playback system controller application software is installed. In some embodiments, the control device 130a includes, for example, a tablet (e.g., an iPad™), a computer (e.g., a laptop computer, a desktop computer), and/or another suitable device (e.g., a television, an automobile audio head unit, an IoT device). In certain embodiments, the control device 130a includes a dedicated controller for the media playback system 100. In other embodiments, as described above with respect to Figure 1G, the control device 130a is integrated into another device in the media playback system 100 (e.g., one more of the playback devices 110, NMDs 120, and/or other suitable devices configured to communicate over a network).

[0131] The control device 130a includes electronics 132, a user interface 133, one or more speakers 134, and one or more microphones 135. The electronics 132 comprise one or more processors 132a (referred to hereinafter as “the processors 132a”), a memory 132b, software components 132c, and a network interface 132d. The processor 132a can be configured to perform functions relevant to facilitating user access, control, and configuration of the media playback system 100. The memory 132b can comprise data storage that can be loaded with one or more of the software components executable by the processor 302 to perform those functions. The software components 132c can comprise applications and/or other executable software configured to facilitate control of the media playback system 100. The memory 112b can be configured to store, for example, the software components 132c, media playback system controller application software, and/or other data associated with the media playback system 100 and the user.

[0132] The network interface 132d is configured to facilitate network communications between the control device 130a and one or more other devices in the media playback system 100, and/or one or more remote devices. In some embodiments, the network interface 132 is

configured to operate according to one or more suitable communication industry standards (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, LTE). The network interface 132d can be configured, for example, to transmit data to and/or receive data from the playback devices 110, the NMDs 120, other ones of the control devices 130, one of the computing devices 106 of Figure 1B, devices comprising one or more other media playback systems, etc. The transmitted and/or received data can include, for example, playback device control commands, state variables, playback zone and/or zone group configurations. For instance, based on user input received at the user interface 133, the network interface 132d can transmit a playback device control command (e.g., volume control, audio playback control, audio content selection) from the control device 304 to one or more of the playback devices 100. The network interface 132d can also transmit and/or receive configuration changes such as, for example, adding/removing one or more playback devices 100 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. Additional description of zones and groups can be found below with respect to Figures 1-I through 1M.

[0133] The user interface 133 is configured to receive user input and can facilitate control of the media playback system 100. The user interface 133 includes media content art 133a (e.g., album art, lyrics, videos), a playback status indicator 133b (e.g., an elapsed and/or remaining time indicator), media content information region 133c, a playback control region 133d, and a zone indicator 133e. The media content information region 133c can include a display of relevant information (e.g., title, artist, album, genre, release year) about media content currently playing and/or media content in a queue or playlist. The playback control region 133d can include selectable (e.g., via touch input and/or via a cursor or another suitable selector) icons to cause one or more playback devices in a selected playback zone or zone group to perform playback actions such as, for example, 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, etc. The playback control region 133d may also include selectable icons to modify equalization settings, playback volume, and/or other suitable playback actions. In the illustrated embodiment, the user interface 133 includes a display presented on a touch screen interface of a smartphone (e.g., an iPhone™, an Android phone). In some embodiments, however, 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.

[0134] The one or more speakers 134 (e.g., one or more transducers) can be configured to output sound to the user of the control device 130a. In some embodiments, the one or more speakers comprise individual transducers configured to correspondingly output low frequencies, mid-range frequencies, and/or high frequencies. In some aspects, for example, the control device 130a is configured as a playback device (e.g., one of the playback devices 110). Similarly, in some embodiments the control device 130a is configured as an NMD (e.g., one of the NMDs 120), receiving voice commands and other sounds via the one or more microphones 135.

[0135] The one or more microphones 135 can comprise, for example, one or more condenser microphones, electret condenser microphones, dynamic microphones, and/or other suitable types of microphones or transducers. In some embodiments, two or more of the microphones 135 are arranged to capture location information of an audio source (e.g., voice, audible sound) and/or configured to facilitate filtering of background noise. Moreover, in certain embodiments, the control device 130a is configured to operate as playback device and an NMD. In other embodiments, however, the control device 130a omits the one or more speakers 134 and/or the one or more microphones 135. For instance, the control device 130a may comprise a device (e.g., a thermostat, an IoT device, a network device) comprising a portion of the electronics 132 and the user interface 133 (e.g., a touch screen) without any speakers or microphones. Additional control device embodiments are described in further detail below with respect to Figures 4A-4D and 5.

e. Suitable Playback Device Configurations

[0136] Figures 1-1 through 1M show example configurations of playback devices in zones and zone groups. Referring first to Figure 1M, in one example, a single playback device may belong to a zone. For example, the playback device 110g in the second bedroom 101c (FIG. 1A) may belong to Zone C. In some implementations described below, multiple playback devices may be “bonded” to form a “bonded pair” which together form a single zone. For example, the playback device 110l (e.g., a left playback device) can be bonded to the playback device 110i (e.g., a left playback device) to form Zone A. Bonded playback devices may have different playback responsibilities (e.g., channel responsibilities). In another implementation described below, multiple playback devices may be merged to form a single zone. For example, the playback device 110h (e.g., a front playback device) may be merged with the playback device 110i (e.g., a subwoofer), and the playback devices 110j and 110k (e.g., left and right surround speakers, respectively) to form a single Zone D. In another example, the playback devices 110g and 110h can be merged to form a merged group or a zone group 108b. The

merged playback devices 110g and 110h may not be specifically assigned different playback responsibilities. That is, the merged playback devices 110h and 110i may, aside from playing audio content in synchrony, each play audio content as they would if they were not merged.

[0137] Each zone in the media playback system 100 may be provided for control as a single user interface (UI) entity. For example, Zone A may be provided as a single entity named Master Bathroom. Zone B may be provided as a single entity named Master Bedroom. Zone C may be provided as a single entity named Second Bedroom.

[0138] Playback devices that are bonded may have different playback responsibilities, such as responsibilities for certain audio channels. For example, as shown in Figure 1-I, the playback devices 110l and 110m may be bonded so as to produce or enhance a stereo effect of audio content. In this example, the playback device 110l may be configured to play a left channel audio component, while the playback device 110k may be configured to play a right channel audio component. In some implementations, such stereo bonding may be referred to as “pairing.”

[0139] Additionally, bonded playback devices may have additional and/or different respective speaker drivers. As shown in Figure 1J, the playback device 110h named Front may be bonded with the playback device 110i named SUB. The Front device 110h can be configured to render a range of mid to high frequencies and the SUB device 110i can be configured render low frequencies. When unbonded, however, the Front device 110h can be configured render a full range of frequencies. As another example, Figure 1K shows the Front and SUB devices 110h and 110i further bonded with Left and Right playback devices 110j and 110k, respectively. In some implementations, the Right and Left devices 110j and 102k can be configured to form surround or “satellite” channels of a home theater system. The bonded playback devices 110h, 110i, 110j, and 110k may form a single Zone D (FIG. 1M).

[0140] Playback devices that are merged may not have assigned playback responsibilities, and may each render the full range of audio content the respective playback device is capable of. Nevertheless, merged devices may be represented as a single UI entity (i.e., a zone, as discussed above). For instance, the playback devices 110a and 110n the master bathroom have the single UI entity of Zone A. In one embodiment, the playback devices 110a and 110n may each output the full range of audio content each respective playback devices 110a and 110n are capable of, in synchrony.

[0141] In some embodiments, an NMD is bonded or merged with another device so as to form a zone. For example, the NMD 120b may be bonded with the playback device 110e, which together form Zone F, named Living Room. In other embodiments, a stand-alone

network microphone device may be in a zone by itself. In other embodiments, however, a stand-alone network microphone device may not be associated with a zone. Additional details regarding associating network microphone devices and playback devices as designated or default devices may be found, for example, in previously referenced U.S. Patent Application No. 15/438,749.

[0142] Zones of individual, bonded, and/or merged devices may be grouped to form a zone group. For example, referring to Figure 1M, Zone A may be grouped with Zone B to form a zone group 108a that includes the two zones. Similarly, Zone G may be grouped with Zone H to form the zone group 108b. As another example, Zone A may be grouped with one or more other Zones C-I. The Zones A-I may be grouped and ungrouped in numerous ways. For example, three, four, five, or more (e.g., all) of the Zones A-I may be grouped. When grouped, the zones of individual and/or bonded playback devices may play back audio in synchrony with one another, as described in previously referenced U.S. Patent No. 8,234,395. Playback devices may be dynamically grouped and ungrouped to form new or different groups that synchronously play back audio content.

[0143] In various implementations, the zones in an environment may be the default name of a zone within the group or a combination of the names of the zones within a zone group. For example, Zone Group 108b can have be assigned a name such as “Dining + Kitchen”, as shown in Figure 1M. In some embodiments, a zone group may be given a unique name selected by a user.

[0144] Certain data may be stored in a memory of a playback device (e.g., the memory 112c of Figure 1C) as one or more state variables that are periodically updated and used to describe the state of a playback zone, the playback device(s), and/or a zone group associated therewith. The memory 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.

[0145] In some embodiments, the memory may store instances of various variable types associated with the states. Variables instances may be stored with identifiers (e.g., tags) corresponding to type. For example, certain identifiers may be a first type “a1” to identify playback device(s) of a zone, a second type “b1” to identify playback device(s) that may be bonded in the zone, and a third type “c1” to identify a zone group to which the zone may belong. As a related example, identifiers associated with the second bedroom 101c may indicate that the playback device is the only playback device of the Zone C and not in a zone group. Identifiers associated with the Den may indicate that the Den is not grouped with other

zones but includes bonded playback devices 110h-110k. Identifiers associated with the Dining Room may indicate that the Dining Room is part of the Dining + Kitchen zone group 108b and that devices 110b and 110d are grouped (FIG. 1L). Identifiers associated with the Kitchen may indicate the same or similar information by virtue of the Kitchen being part of the Dining + Kitchen zone group 108b. Other example zone variables and identifiers are described below.

[0146] In yet another example, the media playback system 100 may store variables or identifiers representing other associations of zones and zone groups, such as identifiers associated with Areas, as shown in Figure 1M. An area may involve a cluster of zone groups and/or zones not within a zone group. For instance, Figure 1M shows an Upper Area 109a including Zones A-D, and a Lower Area 109b including Zones E-I. In one aspect, an Area may be used to invoke a cluster of zone groups and/or zones that share one or more zones and/or zone groups of another cluster. In another aspect, this differs from a zone group, which does not share a zone with another zone group. Further examples of techniques for implementing Areas may be found, for example, in U.S. Application No. 15/682,506 filed August 21, 2017 and titled "Room Association Based on Name," and U.S. Patent No. 8,483,853 filed September 11, 2007, and titled "Controlling and manipulating groupings in a multi-zone media system." Each of these applications is incorporated herein by reference in its entirety. In some embodiments, the media playback system 100 might not implement Areas, in which case the system may not store variables associated with Areas.

[0147] In further examples, the playback devices 110 of the media playback system 100 are named and arranged according to a control hierarchy referred to as home graph. Under the home graph hierarchy, the base unit of the home graph hierarchy is a "Set." A "Set" refers to an individual device or multiple devices that operate together in performing a given function, such as an individual playback device 110 or a bonded zone of playback devices. After Sets, the next level of the hierarchy is a "Room." Under the home graph hierarchy, a "Room" can be considered a container for Sets in a given room of a home. For example, an example Room might correspond to the kitchen of a home, and be assigned the name "Kitchen" and include one or more Sets (e.g., "Kitchen Island"). The next level of the example home graph hierarchy is "Area," which includes two or more Rooms (e.g., "Upstairs" or "Downstairs"). The highest level of the home graph hierarchy is "Home." A Home refers to the entire home, and all of the Sets within. Each level of the home graph hierarchy is assigned a human-readable name, which facilitates control via GUI and VUI. Additional details regarding the home graph control hierarchy can be found, for example, in U.S. Patent App. No. 16/216,357 entitled, "Home Graph," which is incorporated herein by reference in its entirety.

III. Example Systems and Devices

[0148] Figure 2A is a front isometric view of a playback device 210 configured in accordance with aspects of the disclosed technology. Figure 2B is a front isometric view of the playback device 210 without a grille 216e. Figure 2C is an exploded view of the playback device 210. Referring to Figures 2A-2C together, the playback device 210 includes a housing 216 that includes an upper portion 216a, a right or first side portion 216b, a lower portion 216c, a left or second side portion 216d, the grille 216e, and a rear portion 216f. A plurality of fasteners 216g (e.g., one or more screws, rivets, clips) attaches a frame 216h to the housing 216. A cavity 216j (Figure 2C) in the housing 216 is configured to receive the frame 216h and electronics 212. The frame 216h is configured to carry a plurality of transducers 214 (identified individually in Figure 2B as transducers 214a-f). The electronics 212 (e.g., the electronics 112 of Figure 1C) is configured to receive audio content from an audio source and send electrical signals corresponding to the audio content to the transducers 214 for playback.

[0149] The transducers 214 are configured to receive the electrical signals from the electronics 112, and further configured to convert the received electrical signals into audible sound during playback. For instance, the transducers 214a-c (e.g., tweeters) can be configured to output high frequency sound (e.g., sound waves having a frequency greater than about 2 kHz). The transducers 214d-f (e.g., mid-woofers, woofers, midrange speakers) can be configured output sound at frequencies lower than the transducers 214a-c (e.g., sound waves having a frequency lower than about 2 kHz). In some embodiments, the playback device 210 includes a number of transducers different than those illustrated in Figures 2A-2C. For example, the playback device 210 can include fewer than six transducers (e.g., one, two, three). In other embodiments, however, the playback device 210 includes more than six transducers (e.g., nine, ten). Moreover, in some embodiments, all or a portion of the transducers 214 are configured to operate as a phased array to desirably adjust (e.g., narrow or widen) a radiation pattern of the transducers 214, thereby altering a user's perception of the sound emitted from the playback device 210.

[0150] In the illustrated embodiment of Figures 2A-2C, a filter 216i is axially aligned with the transducer 214b. The filter 216i can be configured to desirably attenuate a predetermined range of frequencies that the transducer 214b outputs to improve sound quality and a perceived sound stage output collectively by the transducers 214. In some embodiments, however, the playback device 210 omits the filter 216i. In other embodiments, the playback device 210

includes one or more additional filters aligned with the transducers 214b and/or at least another of the transducers 214.

[0151] Figures 3A and 3B are front and right isometric side views, respectively, of an NMD 320 configured in accordance with embodiments of the disclosed technology. Figure 3C is an exploded view of the NMD 320. Figure 3D is an enlarged view of a portion of Figure 3B including a user interface 313 of the NMD 320. Referring first to Figures 3A-3C, the NMD 320 includes a housing 316 comprising an upper portion 316a, a lower portion 316b and an intermediate portion 316c (e.g., a grille). A plurality of ports, holes or apertures 316d in the upper portion 316a allow sound to pass through to one or more microphones 315 (Figure 3C) positioned within the housing 316. The one or more microphones 316 are configured to received sound via the apertures 316d and produce electrical signals based on the received sound. In the illustrated embodiment, a frame 316e (Figure 3C) of the housing 316 surrounds cavities 316f and 316g configured to house, respectively, a first transducer 314a (e.g., a tweeter) and a second transducer 314b (e.g., a mid-woofer, a midrange speaker, a woofer). In other embodiments, however, the NMD 320 includes a single transducer, or more than two (e.g., two, five, six) transducers. In certain embodiments, the NMD 320 omits the transducers 314a and 314b altogether.

[0152] Electronics 312 (Figure 3C) includes components configured to drive the transducers 314a and 314b, and further configured to analyze audio data corresponding to the electrical signals produced by the one or more microphones 315. In some embodiments, for example, the electronics 312 comprises many or all of the components of the electronics 112 described above with respect to Figure 1C. In certain embodiments, the electronics 312 includes components described above with respect to Figure 1F such as, for example, the one or more processors 112a, the memory 112b, the software components 112c, the network interface 112d, etc. In some embodiments, the electronics 312 includes additional suitable components (e.g., proximity or other sensors).

[0153] Referring to Figure 3D, the user interface 313 includes a plurality of control surfaces (e.g., buttons, knobs, capacitive surfaces) including a first control surface 313a (e.g., a previous control), a second control surface 313b (e.g., a next control), and a third control surface 313c (e.g., a play and/or pause control). A fourth control surface 313d is configured to receive touch input corresponding to activation and deactivation of the one or microphones 315. A first indicator 313e (e.g., one or more light emitting diodes (LEDs) or another suitable illuminator) can be configured to illuminate only when the one or more microphones 315 are activated. A second indicator 313f (e.g., one or more LEDs) can be configured to remain solid during normal

operation and to blink or otherwise change from solid to indicate a detection of voice activity. In some embodiments, the user interface 313 includes additional or fewer control surfaces and illuminators. In one embodiment, for example, the user interface 313 includes the first indicator 313e, omitting the second indicator 313f. Moreover, in certain embodiments, the NMD 320 comprises a playback device and a control device, and the user interface 313 comprises the user interface of the control device.

[0154] Referring to Figures 3A-3D together, the NMD 320 is configured to receive voice commands from one or more adjacent users via the one or more microphones 315. As described above with respect to Figure 1B, the one or more microphones 315 can acquire, capture, or record sound in a vicinity (e.g., a region within 10m or less of the NMD 320) and transmit electrical signals corresponding to the recorded sound to the electronics 312. The electronics 312 can process the electrical signals and can analyze the resulting audio data to determine a presence of one or more voice commands (e.g., one or more activation words). In some embodiments, for example, after detection of one or more suitable voice commands, the NMD 320 is configured to transmit a portion of the recorded audio data to another device and/or a remote server (e.g., one or more of the computing devices 106 of Figure 1B) for further analysis. The remote server can analyze the audio data, determine an appropriate action based on the voice command, and transmit a message to the NMD 320 to perform the appropriate action. For instance, a user may speak “Sonos, play Michael Jackson.” The NMD 320 can, via the one or more microphones 315, record the user’s voice utterance, determine the presence of a voice command, and transmit the audio data having the voice command to a remote server (e.g., one or more of the remote computing devices 106 of Figure 1B, one or more servers of a VAS and/or another suitable service). The remote server can analyze the audio data and determine an action corresponding to the command. The remote server can then transmit a command to the NMD 320 to perform the determined action (e.g., play back audio content related to Michael Jackson). The NMD 320 can receive the command and play back the audio content related to Michael Jackson from a media content source. As described above with respect to Figure 1B, suitable content sources can include a device or storage communicatively coupled to the NMD 320 via a LAN (e.g., the network 104 of Figure 1B), a remote server (e.g., one or more of the remote computing devices 106 of Figure 1B), etc. In certain embodiments, however, the NMD 320 determines and/or performs one or more actions corresponding to the one or more voice commands without intervention or involvement of an external device, computer, or server.

[0155] Figures 4A-4D are schematic diagrams of a control device 430 (e.g., the control device 130a of Figure 1H, a smartphone, a tablet, a dedicated control device, an IoT device, and/or another suitable device) showing corresponding user interface displays in various states of operation. A first user interface display 431a (Figure 4A) includes a display name 433a (i.e., “Rooms”). A selected group region 433b displays audio content information (e.g., artist name, track name, album art) of audio content played back in the selected group and/or zone. Group regions 433c and 433d display corresponding group and/or zone name, and audio content information audio content played back or next in a playback queue of the respective group or zone. An audio content region 433e includes information related to audio content in the selected group and/or zone (i.e., the group and/or zone indicated in the selected group region 433b). A lower display region 433f is configured to receive touch input to display one or more other user interface displays. For example, if a user selects “Browse” in the lower display region 433f, the control device 430 can be configured to output a second user interface display 431b (Figure 4B) comprising a plurality of music services 433g (e.g., Spotify, Radio by Tuncin, Apple Music, Pandora, Amazon, TV, local music, line-in) through which the user can browse and from which the user can select media content for play back via one or more playback devices (e.g., one of the playback devices 110 of Figure 1A). Alternatively, if the user selects “My Sonos” in the lower display region 433f, the control device 430 can be configured to output a third user interface display 431c (Figure 4C). A first media content region 433h can include graphical representations (e.g., album art) corresponding to individual albums, stations, or playlists. A second media content region 433i can include graphical representations (e.g., album art) corresponding to individual songs, tracks, or other media content. If the user selections a graphical representation 433j (Figure 4C), the control device 430 can be configured to begin play back of audio content corresponding to the graphical representation 433j and output a fourth user interface display 431d fourth user interface display 431d includes an enlarged version of the graphical representation 433j, media content information 433k (e.g., track name, artist, album), transport controls 433m (e.g., play, previous, next, pause, volume), and indication 433n of the currently selected group and/or zone name.

[0156] Figure 5 is a schematic diagram of a control device 530 (e.g., a laptop computer, a desktop computer). The control device 530 includes transducers 534, a microphone 535, and a camera 536. A user interface 531 includes a transport control region 533a, a playback status region 533b, a playback zone region 533c, a playback queue region 533d, and a media content source region 533e. The transport control region includes one or more controls for controlling media playback including, for example, volume, previous, play/pause, next, repeat, shuffle,

track position, crossfade, equalization, etc. The audio content source region 533e includes a listing of one or more media content sources from which a user can select media items for playback and/or adding to a playback queue.

[0157] The playback zone region 533b can include representations of playback zones within the media playback system 100 (Figures 1A and 1B). 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, renaming of zone groups, etc. In the illustrated embodiment, a “group” icon is 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 can 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 the illustrated embodiment, 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. In some embodiments, the control device 530 includes other interactions and implementations for grouping and ungrouping zones via the user interface 531. In certain embodiments, the representations of playback zones in the playback zone region 533b can be dynamically updated as playback zone or zone group configurations are modified.

[0158] The playback status region 533c includes 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 533b and/or the playback queue region 533d. 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 100 via the user interface 531.

[0159] The playback queue region 533d includes 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. In some embodiments, for example, a playlist can be added to a playback queue, in which information corresponding to each audio item in the playlist may be added to the playback queue. In some embodiments, audio items in a playback queue may be saved as a playlist. In certain embodiments, 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 some embodiments, 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.

[0160] 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.

[0161] Figure 6 is a message flow diagram illustrating data exchanges between devices of the media playback system 100 (Figures 1A-1M).

[0162] At step 650a, the media playback system 100 receives an indication of selected media content (e.g., one or more songs, albums, playlists, podcasts, videos, stations) via the control device 130a. The selected media content can comprise, for example, media items stored locally on or more devices (e.g., the audio source 105 of Figure 1C) connected to the media playback system and/or media items stored on one or more media service servers (one or more of the

remote computing devices 106 of Figure 1B). In response to receiving the indication of the selected media content, the control device 130a transmits a message 651a to the playback device 110a (Figures 1A-1C) to add the selected media content to a playback queue on the playback device 110a.

[0163] At step 650b, the playback device 110a receives the message 651a and adds the selected media content to the playback queue for play back.

[0164] At step 650c, the control device 130a receives input corresponding to a command to play back the selected media content. In response to receiving the input corresponding to the command to play back the selected media content, the control device 130a transmits a message 651b to the playback device 110a causing the playback device 110a to play back the selected media content. In response to receiving the message 651b, the playback device 110a transmits a message 651c to the computing device 106a requesting the selected media content. The computing device 106a, in response to receiving the message 651c, transmits a message 651d comprising data (e.g., audio data, video data, a URL, a URI) corresponding to the requested media content.

[0165] At step 650d, the playback device 110a receives the message 651d with the data corresponding to the requested media content and plays back the associated media content.

[0166] At step 650e, the playback device 110a optionally causes one or more other devices to play back the selected media content. In one example, the playback device 110a is one of a bonded zone of two or more players (Figure 1M). The playback device 110a can receive the selected media content and transmit all or a portion of the media content to other devices in the bonded zone. In another example, the playback device 110a is a coordinator of a group and is configured to transmit and receive timing information from one or more other devices in the group. The other one or more devices in the group can receive the selected media content from the computing device 106a, and begin playback of the selected media content in response to a message from the playback device 110a such that all of the devices in the group play back the selected media content in synchrony.

IV. Example Synchrony Grouping Techniques

[0167] Example synchrony techniques involve a group coordinator providing audio content and timing information to one or more group members to facilitate synchronous playback among the group coordinator and the group members. In some embodiments, at least some aspects of the technical solutions derive from the technical structure and organization of the audio information, playback timing, and clock timing information that the playback devices

use to play audio content from audio sources in synchrony with each other, including how different playback devices generate playback timing based on clock timing (local clock timing or remote clock timing) and play audio content based on playback timing (generated locally or remotely) and clock timing (generated locally or remotely). Therefore, to aid in understanding certain aspects of the disclosed technical solutions, certain technical details of the audio information, playback timing, and clock timing information, as well as how playback devices generate and/or use playback timing and clock timing for playing audio content in different configurations, are described below.

a. Audio Content

[0168] Audio content may be any type of audio content now known or later developed. For example, in some embodiments, the audio content includes any one or more of: (i) streaming music or other audio obtained from a streaming media service, such as Spotify, Pandora, or other streaming media services; (ii) streaming music or other audio from a local music library, such as a music library stored on a user's laptop computer, desktop computer, smartphone, tablet, home server, or other computing device now known or later developed; (iii) audio content associated with video content, such as audio associated with a television program or movie received from any of a television, set-top box, Digital Video Recorder, Digital Video Disc player, streaming video service, or any other source of audio-visual media content now known or later developed; (iv) text-to-speech or other audible content from a voice assistant service (VAS), such as Amazon Alexa or other VAS services now known or later developed; (v) audio content from a doorbell or intercom system such as Nest, Ring, or other doorbells or intercom systems now known or later developed; and/or (vi) audio content from a telephone, video phone, video/teleconferencing system or other application configured to allow users to communicate with each other via audio and/or video.

[0169] In operation, a "sourcing" playback device obtains any of the aforementioned types of audio content from an audio source via an interface on the playback device, e.g., one of the sourcing playback device's network interfaces, a "line-in" analog interface, a digital audio interface, or any other interface suitable for receiving audio content in digital or analog format now known or later developed.

[0170] An audio source is any system, device, or application that generates, provides, or otherwise makes available any of the aforementioned audio content to a playback device. For example, in some embodiments, an audio source includes any one or more of a streaming media (audio, video) service, digital media server or other computing system, VAS service, television,

cable set-top-box, streaming media player (e.g., AppleTV, Roku, gaming console), CD/DVD player, doorbell, intercom, telephone, tablet, or any other source of digital audio content.

[0171] A playback device that receives or otherwise obtains audio content from an audio source for playback and/or distribution to other playback devices is sometimes referred to herein as the “sourcing” playback device, “master” playback device, or “group coordinator.” One function of the “sourcing” playback device is to process received audio content for playback and/or distribution to other playback devices. In some embodiments, the sourcing playback device transmits the processed audio content to all the playback devices that are configured to play the audio content. In some embodiments, the sourcing playback device transmits the processed audio content to a multicast network address, and all the other playback devices configured to play the audio content receive the audio content via that multicast address. In some embodiments, the sourcing playback device alternatively transmits the processed audio content to each unicast network address of each other playback device configured to play the audio content, and each of the other playback devices configured to play the audio content receive the audio content via its unicast address.

[0172] In some embodiments, the “sourcing” playback device receives audio content from an audio source in digital form, e.g., as a stream of packets. In some embodiments, individual packets in the stream of packets have a sequence number or other identifier that specifies an ordering of the packets. Packets transmitted over a data packet network (e.g., Ethernet, WiFi, or other packet networks) may arrive out of order, so the sourcing playback device uses the sequence number or other identifier to reassemble the stream of packets in the correct order before performing further packet processing. In some embodiments, the sequence number or other identifier that specifies the ordering of the packets is or at least comprises a timestamp indicating a time when the packet was created. The packet creation time can be used as a sequence number based on an assumption that packets are created in the order in which they should be subsequently played out.

[0173] In some embodiments, the sourcing playback device does not change the sequence number or identifier of the received packets during packet processing. In some embodiments, the sourcing playback device reorders at least a first set packets in a packet stream based on each packet’s sequence identifier, extracts audio content from the received packets, reassembles a bitstream of audio content from the received packets, and then repacketizes the reassembled bitstream into a second set of packets, where packets in the second set of packets have sequence numbers that differ from the sequence numbers of the packets in the first set of packets. In some embodiments, individual packets in the second set of packets are a different

length (i.e., shorter or longer) than individual packets in the first set of packets. In some embodiments, reassembling a bitstream from the incoming packets and then subsequently repacketizing the reassembled bitstream into a different set of packets facilitates uniform processing and/or transmission of audio content by the sourcing playback device and by other playback devices that receive the audio content from the sourcing playback device. However, for some delay-sensitive audio content, reassembly and repacketization may be undesirable, and therefore, in some embodiments, the sourcing playback device may not perform reassembly and repacketization for some (or all) audio content that it receives before playing the audio content and/or transmitting the audio content to other playback devices.

[0174] In some embodiments an audio source provides audio content in digital form to a sourcing playback device, e.g., via a digital line-in interface. In such embodiments, the sourcing playback device packetizes the digital audio into packets of audio content before transmitting the audio content to other playback devices. In some embodiments, individual packets of the audio content comprise a sequence number or other identifier so that, when other playback devices receive the audio content, those other playback devices will be able to reliably arrange the received packets in the correct order before performing further packet processing.

[0175] In some embodiments an audio source provides audio content in analog form to a sourcing playback device, e.g., via an analog line-in interface. In such embodiments, the sourcing playback device converts the received analog audio into digital audio and packetizes the digital audio into packets of audio content before transmitting the audio content to other playback devices. In some embodiments, individual packets of the audio content comprise a sequence number or other identifier so that, when other playback devices receive the audio content, those other playback devices will be able to reliably arrange the received packets in the correct order before performing further packet processing.

[0176] After obtaining audio content from an audio source or from another playback device, a playback device in some embodiments one or more of (i) plays the audio content individually, (ii) plays the content in synchrony with one or more additional playback devices, and/or (iii) transmits the audio content to one or more other playback devices.

b. Playback Timing

[0177] The playback devices disclosed and described herein use playback timing to play audio content in synchrony with each other. An individual playback device can generate playback timing and/or playback audio content according to playback timing, based on the playback device's configuration in the media playback network. The sourcing playback device

that generates the playback timing for audio content also transmits that generated playback timing to all the playback devices that are configured to play the audio content. In some embodiments, the sourcing playback device transmits the playback timing to a multicast network address, and all the other playback devices configured to play the audio content receive the playback timing via that multicast address. In some embodiments, the sourcing playback device alternatively transmits the playback timing to each unicast network address of each other playback device configured to play the audio content, and each of the other playback devices configured to play the audio content receive the playback timing via its unicast address.

[0178] In operation, a playback device (or a computing device associated with the playback device) generates the playback timing for audio content based on clock timing (described below), which can be “local” clock timing (i.e., clock timing generated by the sourcing playback device) or “remote” clock timing received from a different playback device (or different computing device).

[0179] In some embodiments, the playback timing is generated for individual frames (or packets) of audio content. As described above, in some embodiments, the audio content is packaged in a series of frames (or packets) where individual frames (or packets) comprise a portion of the audio content. In some embodiments, the playback timing for the audio content includes a playback time for each frame (or packet) of audio content. In some embodiments, the playback timing for an individual frame (or packet) is included within the frame (or packet), e.g., in the header of the frame (or packet), in an extended header of the frame (or packet), and/or in the payload portion of the frame (or packet).

[0180] In some embodiments, the playback time for an individual frame (or packet) is identified within a timestamp or other indication. In such embodiments, the timestamp (or other indication) represents a time to play audio content within that individual frame (or packet). In operation, when the playback timing for an individual frame (or packet) is generated, the playback timing for that individual frame (or packet) is a future time relative to a current clock time of a reference clock at the time that the playback timing for that individual frame (or packet) is generated. The reference clock can be a “local” clock at the playback device or a “remote” clock at a separate network device, e.g., another playback device, a computing device, or another network device configured to provide clock timing for use by playback devices to generate playback timing and/or playback audio content.

[0181] In operation, a playback device tasked with playing particular audio content will play the portion(s) of the particular audio content within an individual frame (or packet) at the playback time specified by the playback timing for that individual frame (or packet), as

adjusted to accommodate for clocking differences between the sourcing playback device, the clock timing, and the playback device that is tasked with playing the audio content, as describe in more detail below.

c. Clock Timing

[0182] The playback devices disclosed and described herein use clock timing to generate playback timing for audio content and to play the audio content based on the generated playback timing. In some embodiments, the sourcing playback device uses clock timing from a reference clock (e.g., a device clock, a digital-to-audio converter clock, a playback time reference clock, or any other clock) to generate playback timing for audio content that the sourcing playback device receives from an audio source. For an individual playback device, the reference clock can be a “local” clock at the playback device or a “remote” clock at a separate network device, e.g., another playback device, a computing device, or another network device configured to provide clock timing for use by playback devices to generate playback timing and/or playback audio content.

[0183] In some embodiments, all of the playback devices tasked with playing particular audio content in synchrony use the same clock timing from the reference clock to play back that particular audio content. In some embodiments, playback devices use the same clock timing to play audio content that was used to generate the playback timing for the audio content.

[0184] In operation, the network device that generates the clock timing also transmits the clock timing to all the playback devices in the network that need to use the clock timing for generating playback timing and/or playing back audio content. In some embodiments, the network device that generates the clock timing transmits the clock timing to a multicast network address, and all the other playback devices configured to generate playback timing and/or play audio content receive the clock timing via that multicast address. In some embodiments, the network device alternatively transmits the clock timing to each unicast network address of each other playback device configured to play the audio content, and each of the other playback devices configured to play the audio content receive the clock timing via its unicast address.

d. Generating Playback Timing with Clock Timing from a Local Clock

[0185] In some embodiments, a sourcing playback device (i) generates playback timing for audio content based on clock timing from a local clock at the sourcing playback device, and (ii) transmits the generated playback timing to all other playback devices configured to play the audio content. In operation, when generating playback timing for an individual frame (or packet), the “sourcing” playback device adds a “timing advance” to the current clock time of

the local clock of the sourcing playback device that the sourcing playback device is using for generating the playback timing.

[0186] In some embodiments, the “timing advance” is based on an amount of time that is greater than or equal to the sum of (i) the network transit time required for frames and/or packets comprising audio content transmitted from the sourcing playback device to arrive at all other playback devices that are configured to use the playback timing for playing the audio content in synchrony and (ii) the amount of time required for all the other playback devices configured to use that playback timing for synchronous playback to process received frames/packets from the sourcing playback device for playback.

[0187] In some embodiments, the sourcing playback device determines a timing advance by sending one or more test packets to one or more (or perhaps all) of the other playback devices configured to play the audio content that the sourcing device is transmitting, and then receiving test response packets back from those one or more of the other playback devices. In some embodiments, the sourcing playback device and the one or more other playback devices negotiate a timing advance via multiple test and response messages. In some embodiments with more than two additional playback devices, the sourcing playback device determines a timing advance by exchanging test and response messages with all of the playback devices, and then setting a timing advance that is sufficient for the playback device having the longest total of network transmit time and packet processing time.

[0188] In some embodiments, the timing advance is less than about 50 milliseconds. In some embodiments, the timing advance is less than about 20-30 milliseconds. And in still further embodiments, the timing advance is less than about 10 milliseconds. In some embodiments, the timing advance remains constant after being determined. In other embodiments, the playback device that generates the playback timing can change the timing advance in response to a request from a receiving device indicating that a greater timing advance is required (e.g., because the receiving device is not receiving packets comprising portions of audio content until after other devices have already played the portions of audio content) or a shorter timing advance would be sufficient (e.g., because the receiving device is buffering more packets comprising portions of audio content than necessary to provide consistent, reliable playback).

[0189] As described in more detail below, all the playback devices configured to play the audio content in synchrony will use the playback timing and the clock timing to play the audio content in synchrony with each other.

e. Playing Audio Content using Local Playback Timing and Local Clock Timing

[0190] In some embodiments, the sourcing playback device is configured to play audio content in synchrony with one or more other playback devices. And if the sourcing playback device is using clock timing from a local clock at the sourcing playback device to generate the playback timing, then the sourcing playback device will play the audio content using locally-generated playback timing and the locally-generated clock timing. In operation, the sourcing playback device plays an individual frame (or packet) comprising portions of the audio content when the local clock that the sourcing playback device used to generate the playback timing reaches the time specified in the playback timing for that individual frame (or packet).

[0191] For example, recall that when generating playback timing for an individual frame (or packet), the sourcing playback device adds a “timing advance” to the current clock time of the reference clock used for generating the playback timing. In this instance, the reference clock used for generating the playback timing is a local clock at the sourcing playback device. So, if the timing advance for an individual frame is, for example, 30 milliseconds, then the sourcing playback device plays the portion (e.g., a sample or set of samples) of audio content in an individual frame (or packet) 30 milliseconds after creating the playback timing for that individual frame (or packet).

[0192] In this manner, the sourcing playback device plays the audio content using locally-generated playback timing and clock timing from a local reference clock. As described further below, by playing the portion(s) of the audio content of an individual frame and/or packet when the clock time of the local reference clock reaches the playback timing for that individual frame or packet, the sourcing playback device plays that portion(s) of the audio content in that individual frame and/or packet in synchrony with other playback devices.

f. Playing Audio Content using Remote Playback Timing and Remote Clock Timing

[0193] Recall that, in some embodiments, the sourcing playback device transmits the audio content and the playback timing for the audio content to one or more other playback devices. And further recall that, in some embodiments, the network device providing the clock timing can be a different device than the sourcing playback device. Playback devices that receive the audio content, the playback timing, and the clock timing from another playback device are configured to playback the audio content using the playback timing from the sourcing playback device (i.e., remote playback timing) and clock timing from a clock at another playback device (i.e., remote clock timing). In this manner, the receiving playback device in this instance plays audio content using remote playback timing and remote clock timing.

[0194] To play an individual frame (or packet) of the audio content in synchrony with every other playback device tasked with playing the audio content, the receiving playback device (i) receives the frames (or packets) comprising the portions of the audio content from the sourcing playback device, (ii) receives the playback timing for the audio content from the sourcing playback device (e.g., in the frame and/or packet headers of the frames and/or packets comprising the portions of the audio content or perhaps separately from the frames and/or packets comprising the portions of the audio content), (iii) receives the clock timing from another network device, e.g., another playback device, a computing device, or another network device configured to provide clock timing for use by playback devices to generate playback timing and/or playback audio content, and (iv) plays the portion(s) of the audio content in the individual frame (or packet) when the local clock that the receiving playback device uses for audio content playback reaches the playback time specified in the playback timing for that individual frame (or packet) received from the sourcing playback device, as adjusted by a “timing offset.”

[0195] In operation, after the receiving playback device receives clock timing from the other network device, the receiving device determines a “timing offset” for the receiving playback device. This “timing offset” comprises (or at least corresponds to) a difference between the “reference” clock at the network device that the network device used to generate the clock timing and a “local” clock at the receiving playback device that the receiving playback device uses to play the audio content. In operation, each playback device that receives the clock timing from another network device calculates its own “timing offset” based on the difference between its local clock and the clock timing, and thus, the “timing offset” that each playback determines is specific to that particular playback device.

[0196] In some embodiments, when playing back the audio content, the receiving playback device generates new playback timing (specific to the receiving playback device) for individual frames (or packets) of audio content by adding the previously determined “timing offset” to the playback timing for each frame (or packet) received from the sourcing playback device. With this approach, the receiving playback device converts the playback timing for the audio content received from the sourcing playback device into “local” playback timing for the receiving playback device. Because each receiving playback device calculates its own “timing offset,” each receiving playback device’s determined “local” playback timing for an individual frame is specific to that particular playback device.

[0197] And when the “local” clock that the receiving playback device is using for playing back the audio content reaches the “local” playback time for an individual frame (or packet),

the receiving playback device plays the audio content (or portions thereof) associated with that individual frame (or packet). As described above, in some embodiments, the playback timing for a particular frame (or packet) is in the header of the frame (or packet). In other embodiments, the playback timing for individual frames (or packets) is transmitted separately from the frames (or packets) comprising the audio content.

[0198] Because the receiving playback device plays frames (or packets) comprising portions of the audio content according to the playback timing as adjusted by the “timing offset” relative to the clock timing, and because the sourcing playback generated the playback timing for those frames (or packets) relative to the clock timing and plays the same frames (or packets) comprising portions of the audio content according to the playback timing and its determined “timing offset,” the receiving playback device and the sourcing playback device play the same frames (or packets) comprising the same portions of the audio content in synchrony, i.e., at the same time or at substantially the same time.

[0199] Additional details regarding audio playback synchronization among playback devices and/or zones can be found, for example, in U.S. Patent No. 8,234,395 entitled, “System and method for synchronizing operations among a plurality of independently clocked digital data processing devices,” which is incorporated herein by reference in its entirety.

V. Example Portable Playback Devices

[0200] As noted above, certain playback device implementations may be configured for portable use. These portable implementations include wearable playback devices, such as headphones and earbuds, generally designed for personal listening by one user at a time, and portable devices designed for out loud playback. Figure 7A is a partial cutaway view of the media playback system 100 with the inclusion of one or more portable playback devices 710 (identified individually as portable playback devices 710a, 710b, and 710c). The portable playback devices 710 are similar to the playback devices 110, but are configured for portable use. While they are shown in the home in Figure 7A, the portable playback devices 710 are configured to play back audio content while in the home and while “on the go.”

[0201] As shown in the block diagram of Figure 7B, a portable playback device 710a includes the same or similar components as the playback device 110a. However, to facilitate portable use, the playback device 710a may be implemented in a certain form factor (e.g., headphones or earbuds) and includes one or more batteries in power 712i to provide portable power.

[0202] Referring to Figure 7B, the portable playback device 710a includes an input/output 711, which can include an analog I/O 711a and/or a digital I/O 711b similar to the components

of the playback device 110. To facilitate portable usage, the input/output 711 of the portable playback device 710a may include an interface (such as a Bluetooth interface) to facilitate connection to a bridge device (e.g., a mobile device), which the portable playback device 710a may use to stream audio content and otherwise communicate with the bridge device.

[0203] The playback device 710a further includes electronics 712, a user interface 713 (e.g., one or more buttons, knobs, dials, touch-sensitive surfaces, displays, touchscreens), and one or more transducers 714 (referred to hereinafter as “the transducers 714”). The electronics 712 is configured to receive audio from an audio source via the input/output 711, one or more of the computing devices 106a-c via the network 104 (Figure 1B), amplify the received audio, and output the amplified audio for playback via one or more of the transducers 714.

[0204] In some embodiments, the playback device 710a optionally includes one or more microphones 715 (e.g., a single microphone, a plurality of microphones, a microphone array) (hereinafter referred to as “the microphones 715”). In some examples, the microphones 715 may include one or more voice microphones to facilitate voice input for telephone calls and the like. In certain embodiments, for example, playback device 710a can operate as an NMD (similar to the NMD 120 of Figure 1F) configured to receive voice input from a user using the voice microphones and correspondingly perform one or more operations based on the received voice input. In further examples, the microphones 715 may include one or more acoustic noise cancelling (ANC) microphones which, in operation, capture ambient noise in the environment to facilitate cancelling of this ambient noise by the playback device 710a.

[0205] In the illustrated embodiment of Figure 7B, the electronics 712 include one or more processors 712a (referred to hereinafter as “the processors 112a”), memory 712b, software components 712c, a network interface 712d, one or more audio processing components 712g (referred to hereinafter as “the audio components 712g”), one or more audio amplifiers 712h (referred to hereinafter as “the amplifiers 712h”), and power 712i (e.g., one or more power supplies, power cables, power receptacles, batteries, induction coils, Power-over Ethernet (POE) interfaces, and/or other suitable sources of electric power). In some embodiments, the electronics 712 optionally include one or more other components 712j (e.g., one or more sensors, video displays, touchscreens).

[0206] The network interface 712d is configured to facilitate a transmission of data between the playback device 710a and one or more other devices on a data network such as, for example, the links 103 and/or the network 104 (Figure 1B). The network interface 712d is configured to transmit and receive data corresponding to media content (e.g., audio content, video content, text, photographs) and other signals (e.g., non-transitory signals) comprising digital packet data

including an Internet Protocol (IP)-based source address and/or an IP-based destination address. The network interface 712d can parse the digital packet data such that the electronics 712 properly receives and processes the data destined for the playback device 110a.

[0207] In the illustrated embodiment of Figure 7B, the network interface 712d includes one or more wireless interfaces 712e (referred to hereinafter as “the wireless interface 712e”). The wireless interface 712e (e.g., a suitable interface comprising one or more antennae) can be configured to wirelessly communicate with one or more other devices (e.g., one or more of the playback devices 110, NMDs 120, control devices 130, other portable playback devices 710, as well as other devices disclosed herein, such as bridge devices) that are communicatively coupled to the network 104 (Figure 1B) in accordance with a suitable wireless communication protocol (e.g., WiFi, Bluetooth, LTE). In some embodiments, the network interface 712d optionally includes a wired interface 712f (e.g., an interface or receptacle configured to receive a network cable such as an Ethernet, a USB-A, USB-C, and/or Thunderbolt cable) configured to communicate over a wired connection with other devices in accordance with a suitable wired communication protocol. In some embodiments, the electronics 712 excludes the network interface 712d altogether and transmits and receives media content and/or other data via another communication path (e.g., the input/output 711).

[0208] The audio components 712g are configured to process and/or filter data comprising media content received by the electronics 712 (e.g., via the input/output 711 and/or the network interface 712d) to produce output audio signals. In some embodiments, the audio processing components 712g comprise, for example, one or more digital-to-analog converters (DAC), audio preprocessing components, audio enhancement components, a digital signal processors (DSPs), and/or other suitable audio processing components, modules, circuits, etc. In certain embodiments, one or more of the audio processing components 712g can comprise one or more subcomponents of the processors 712a. In some embodiments, the electronics 712 omits the audio processing components 712g. In some aspects, for example, the processors 712a execute instructions stored on the memory 712b to perform audio processing operations to produce the output audio signals.

[0209] The amplifiers 712h are configured to receive and amplify the audio output signals produced by the audio processing components 712g and/or the processors 712a. The amplifiers 712h can comprise electronic devices and/or components configured to amplify audio signals to levels sufficient for driving one or more of the transducers 714. In some embodiments, for example, the amplifiers 712h include one or more switching or class-D power amplifiers. In other embodiments, however, the amplifiers include one or more other types of power

amplifiers (e.g., linear gain power amplifiers, class-A amplifiers, class-B amplifiers, class-AB amplifiers, class-C amplifiers, class-D amplifiers, class-E amplifiers, class-F amplifiers, class-G and/or class H amplifiers, and/or another suitable type of power amplifier). In certain embodiments, the amplifiers 712h comprise a suitable combination of two or more of the foregoing types of power amplifiers. Moreover, in some embodiments, individual ones of the amplifiers 712h correspond to individual ones of the transducers 714. In other embodiments, however, the electronics 712 includes a single one of the amplifiers 712h configured to output amplified audio signals to a plurality of the transducers 714.

[0210] The transducers 714 (e.g., one or more speakers and/or speaker drivers) receive the amplified audio signals from the amplifier 712h and render or output the amplified audio signals as sound (e.g., audible sound waves having a frequency between about 20 Hertz (Hz) and 20 kilohertz (kHz)). In some embodiments, the transducers 714 can comprise a single transducer. In other embodiments, however, the transducers 714 comprise a plurality of audio transducers. In some embodiments, the transducers 714 comprise more than one type of transducer. For example, the transducers 714 can include one or more low frequency transducers (e.g., subwoofers, woofers), mid-range frequency transducers (e.g., mid-range transducers, mid-woofers), and one or more high frequency transducers (e.g., one or more tweeters).

[0211] Figure 7C is a front isometric view of the portable playback device 710a configured in accordance with aspects of the disclosed technology. As shown in Figure 7C, the portable playback device 710a is implemented as headphones to facilitate more private playback as compared with the out loud playback of the playback device(s) 110. As shown, the portable playback device 710a (also referred to as headphones 710a) includes a housing 716a to support a pair of transducers 714a on or around user's head over the user's ears.

[0212] The headphones 710a also include a user interface 713a with a touch-sensitive region to facilitate playback controls such as transport and/or volume controls. The touch-sensitive region of the user interface 713a may support gesture controls. For instance, a swipe forward or backward across the touch-sensitive region may skip forward or backwards, respectively. Other gestures include a touch-and-hold, as well as a touch-and-continued hold, which may correspond to various swapping and grouping functions, as described in further detail below. In some implementations, the user interface 713a may include respective touch-sensitive regions on the exterior of each earcup.

[0213] Figure 7D is a front isometric view of the portable playback device 710b configured in accordance with aspects of the disclosed technology. As shown in Figure 7D, the portable

playback device 710b is implemented as earbuds to facilitate more private playback as compared with the out loud playback of the playback device(s) 110, similar to the headphones 710a. As shown, the portable playback device 710b (also referred to as earbuds 710b) includes a housing 716b to support a pair of transducers 714b within a user's ears. The earbuds 710b also include a user interface 713b with a touch-sensitive region to facilitate playback controls such as transport and/or volume controls. The earbuds 9710b can be in the form of wired, wireless, or true wireless earbuds.

[0214] Figure 7E is a front isometric view of the portable playback device 710c. As compared with the headphones 710a and the earbuds 710b, the portable playback device 710c includes one or more larger transducers to facilitate out loud audio content playback. A speaker grill 716a covers the transducers. Relative to the playback device(s) 110, the portable playback device 710c may include less powerful amplifier(s) and/or smaller transducer(s) to balance battery life, sound output capability, and form factor (i.e., size, shape, and weight) of the portable playback device 710c. The portable playback device 710c includes a user interface 713c with a touch-sensitive region to facilitate playback controls such as transport and/or volume controls.

[0215] Some portable playback devices 710 are configured to be placed upon a device base 718. To illustrate, Figure 7F is a front isometric view of the portable playback device 710d, which is configured to be placed upon a device base 718a. Like the portable playback device 710c, as compared with the headphones 710a and the earbuds 710b, the portable playback device 710d includes one or more larger transducers to facilitate out loud audio content playback. A speaker grill 716b covers the transducers. The portable playback device 710c includes a user interface 713d with a touch-sensitive region to facilitate playback controls such as transport and/or volume controls.

[0216] The device base 718a includes protrusions 719a and 719b, which align with recesses 717a and 717b on the portable playback device 710c. Such protrusions and recesses may facilitate placing the portable playback device 710c on the device base 718a and may improve stability of the playback device while it is positioned on the device base 718a.

[0217] In example implementations, the portable playback device 710c is rotatable about the device base 718a to control volume of the portable playback device 710c. For instance, the portable playback device 710c may rotate with respect to the device base 718a, which may generate a volume control signal in a sensor of the portable playback device 710c and/or device base 718a. In another example, a first portion of the device base 718a is rotatable with respect to a second portion of the device base 718a. Rotation of these two portions generates a volume

control signal in a sensor of the device base 718a that controls volume of the portable playback device 710c when the portable playback device 710c is placed upon the device base 718a.

[0218] The device base 718a includes a device charging system. When the playback device 710c is placed on device base 718a, the playback device 710c may draw current from the charging system to charge one or more of its batteries. In some examples, the charging system of the device base 718a includes an inductive charging circuit (*e.g.*, a coil that induces a current in a corresponding coil in the playback device 710c that wirelessly charges one or more batteries of the playback device 710c). Alternatively, the charging system of the device base 718a includes conductive terminals by which the playback device 710c may draw current from the device base 718.

[0219] Within examples, the device base 718a carries an identifier that distinguishes that device base 718a from at least some other device bases (*e.g.*, other device bases of the media playback system 100, or perhaps other devices bases more broadly). In some implementations, the device base 718a may passively communicate this identifier to the playback device 710c when it is placed upon the device base 718a. For instance, a charging circuit of the device base 718 may include a current or voltage signature (*i.e.*, a pattern) that is unique as compared to other device bases. The playback device 710c may use this unique signature to identify the device base 718. Alternatively, a charging circuit may superimpose a signal onto the current delivered from the device base 718a (*e.g.*, current from the device base 718a may include a higher frequency signal carrying the identifier of the device base 718a). In further examples, the device base 718a includes an RFID tag, QR code, or other identifying component that is read by the playback device 710c when the playback device 710c is placed upon the device base 718a.

[0220] In some implementations, the device base(s) 718 of the media playback system 100 are associated with respective zones. Placing a portable playback device 710 on a device base causes the device base to join the associated zone. Additional details regarding device bases can be found, for example, in U.S. Patent No. 9,544,701 entitled, "Base Properties in a Media Playback System," which is incorporated herein by reference in its entirety.

[0221] In some implementations, the device base 718a includes a control system. Example control systems of the device base 718a include one or more processors and memory. The processor(s) may be clock-driven computing components that process input data according to instructions stored in the memory. Example operations include communicating via a communications interface (*e.g.*, a BLUETOOTH® interface) with playback device 710c (*e.g.*,

to cause the playback device 710c to join an associated zone via one or more instructions) and causing the charging system to supply current to playback device 710c, among other examples.

[0222] Within example implementations, the playback device 710 may operate in one of a first mode and a second mode. Generally, the playback device 710 operates in the first mode while in the physical proximity of the media playback system 100 (e.g., while in the home and connected to the network 104) to facilitate interoperability with the playback devices 110a-n of the media playback system 100 and operates in the second mode while “on the go,” but the playback device 710 may also be operable in the second mode while in the physical proximity of the media playback system 100. The portable playback device 710 may switch between modes manually (e.g., via user input to a user interface 713) or automatically (e.g., based on proximity to one or more playback devices 110a-n, connection to the network 104, and/or based on the location of the mobile device).

[0223] The playback device 710 may operate in the first mode while connected to a wireless local area network (e.g., the network 104). Through the connection to the wireless local area network, the playback device 710 may stream audio content from one or more audio sources, including local and remote (e.g., cloud) network locations. Further, in the first mode, the portable playback device 710 may interface with other devices of the media playback system 100. For instance, the portable playback device 710 may form synchrony groupings or other arrangements with the playback devices 110a-n and/or other portable playback devices 710 in the first mode. Further, in the first mode, the portable playback device 710 may be controlled by the control device(s) 130 in the same or similar manner as the playback device(s) 110.

[0224] The playback device 710 may operate in a second mode when connected via BLUETOOTH (802.15) to a mobile device. In some aspects, in the second mode, the portable device operates similarly to conventional BLUETOOTH speakers or wearable devices. That is, the playback device 710 may pair with a mobile device, such as a smartphone or tablet, and the user may play back the audio output of the mobile device. Similarly, the microphones 715a of the portable playback device 710 may provide audio input to the mobile device. As noted above, this mode can be utilized “on the go” to facilitate playback away from the media playback system 100, for example, outside the range of a home network. Further, this mode can be used in proximity to the media playback system 100, which may facilitate more private use of the portable playback device 710a or provide convenient access to content on the mobile device for playback.

[0225] Figure 7G illustrates an example pairing arrangement between the headphones 710a and a mobile device configured as a control device 130a. As noted above, a mobile device may

become a control device 130 via the installation of control application software, which may further provide bridging features to facilitate the control device 130a operating as an interface between the headphones 710a and the media playback system 100.

[0226] The control device 130a may include communications interface(s), processing capabilities, and/or other features that are not necessarily implemented in the portable playback device 710a. By “pairing” the portable playback device 710a to the control device 130a, the portable playback device 710 is able to utilize some of these features. This arrangement may permit the portable playback device 710a to be smaller and more portable, to draw less power, and/or to be less expensive, among other possible benefits.

[0227] For instance, in various implementations, the portable playback device 710a may be implemented with or without a communications interface to connect to the Internet while “on the go” (e.g., a cellular data connection). By pairing the portable playback device 710a to the control device 130a via a personal area connection such as Bluetooth® (IEEE 802.15) or a wireless local area network connection (IEEE 802.11), the portable playback device 710a may stream music via an Internet connection of the control device 130a and the pairing connection. In embodiments that include a wireless local area network interface, the portable playback device 710a may connect directly to a wireless local area network (e.g., the network 104 (Figure 1B)) if available.

[0228] Similarly, in various implementations, the portable playback device 710a may be implemented with or without a wireless local area network interface. By pairing the portable playback device 710 to the control device 130a via a personal area connection such as Bluetooth® (IEEE 802.15), the portable playback device 710a may stream music via an Internet connection of the control device 130a and the pairing connection. In this example, the Internet connection of the bridge device 860 may be a wireless local area network having a gateway to the Internet or via a cellular data connection.

[0229] In an example implementation, the control device 130a is bonded or defaults to a particular playback device (e.g., playback device 110c), bonded zone of playback device (e.g., playback devices 110l and 110m) or group of playback devices e.g., a “Kitchen + Dining Room” Group). Alternatively, if the home graph hierarchy is utilized, the control device 130a may be bonded to or defaults to joining a particular Set, Room, or Area. Then, in this configuration, control of the bonded playback device(s) 110 via a NMD 120 or a control device 130 also controls the paired portable playback device 710a.

[0230] Alternatively, the control device 130a may itself form a zone or Set. For instance, in one example, control device 130a may be configured as a “Anne’s Portable” zone or an

“Anne’s Headphones” Set. Configuring the control device 130a as a zone or Set facilitates control of the paired headphones 710a with the NMD(s) 120 and/or the control device(s) 130 of the media playback system 100.

[0231] In an alternative implementation, the portable playback device 710a may independently interface with the media playback system 100 as its own zone or Set. Such an implementation of the portable playback device 710a may include a cellular data connection to facilitate portable streaming (i.e., streaming away from the media playback system 100 and/or the network 104). In this example, the portable playback device 710a may join the media playback system 100 as a zone or Set when connected to the network 104 or otherwise in proximity to the playback device(s) 110.

VI. Example Swap Techniques

[0232] As noted above, example techniques described herein relate to transitions (or “swaps”) of a playback session between a portable playback device 710 and one or more playback device(s) 110. During a playback session swap, playback of audio content stops at a “source” playback device and starts at a “target” playback device at the same or substantially the same offset within the audio content. For instance, the media playback system 100 may swap playback between a “source” portable playback device 710 and the one or more “target” playback device(s) 110. In further examples, the media playback system 100 may swap playback between one or more “source” playback device(s) 110 and a “target” portable playback device 710.

[0233] To illustrate, a user may begin listening to audio content via the headphones 710a or earbuds 710b while “on the go” and then swap playback of the audio content to one or more playback devices 110a-n to continue listening to the audio content out loud at home. In another example, a user may begin listening to audio content via the headphones 710a or earbuds 710b at home (perhaps so as not to disturb another person in the home) and then swap to one or more playback devices 110a-n to continue listening to the audio content out loud. In a third example, a user may begin listening to audio content out loud via the portable playback device 710c and then swap to one or more one or more playback devices 110a-n, as the target playback device(s) may have greater sound output capabilities (e.g., due to more powerful amplifier(s) and/or larger transducer(s)), be located in a different room, be configured in a synchrony group, or for any other reason.

[0234] Similarly, a user may be listening to audio content via one or more playback devices 110a-n and the swap playback of the audio content to the portable playback device 710. For example, a user may be listening to television audio in the den 101d (including playback

devices 110h, 110i, 110j, and 110k) and then swap playback to earbuds 710b for more personal listening. As another example, a user may be listening to an Internet radio station in the kitchen 101h (including playback device 110b) and then swap playback to headphones 710a to continue listening on the go. As a third example, a user may be listening to music in the bedroom 101c (including playback device 110g) and swap playback to portable playback device 710c to take the music out to the yard.

[0235] A playback device that has an on-going playback session may maintain or have access to playback session data that defines and/or identifies the playback session. The playback session data may include data representing a source of the audio content (e.g., a URI or URL indicating the location of the audio content), as well as an offset indicating a position within the audio content to start playback. The offset may be defined as a time (e.g., in milliseconds) from the beginning of the audio track or as a number of samples, among other examples. In example implementations, the offset may be set to a playback position in the audio content of the current playback position to allow time for the target device to start buffering the audio content. Then, the source playback device stops playback of the audio content at the offset and the target playback device starts playback of the audio content at the offset.

[0236] The playback session data may include data representing a source of the audio content (e.g., a URI or URL indicating the location of the audio content), as well as an offset indicating a position within the audio content to start playback. The offset may be defined as a time (e.g., in milliseconds) from the beginning of the audio track or as a number of samples, among other examples.

[0237] The playback session data may further include data representing playback state. Playback state may include a playback state of the session (e.g., playing, paused, or stopped). If the playback session implements a playback queue, the playback session data may include the playback queue state, such as the current playback position within the queue.

[0238] The playback queue state may also include a queue version. For example, in a cloud queue embodiment, the cloud queue server and the media playback system 100 may use the queue version to maintain consistency. The queue version may be incremented each time the queue is modified and then shared between the media playback system 100 and cloud queue server to indicate the most recent version of the queue.

[0239] Further, the playback session data may also include authorization data, such as one or more keys and/or tokens. Such authorization data may include a token associated with the user account. During a playback session swap, the media playback system 100 may verify that the token is authorized on both the source and target playback devices. The authorization data may

further include a token associated with the streaming audio service, which may enable the target playback device to access the audio content at the source. Yet further, the authorization data may include a token associate with the playback session, which enables the target playback device to access the session. Other example authorization data is contemplated as well

[0240] In some implementations, an input to a playback device triggers a swap. This input may be referred to as a “playback session swap input.” Within examples, a playback session swap input may be provided to a user interface on the playback device, such as the user interface 313 on the playback device 320 (Figure 7D) or the user interface 713a of the headphones 710 (Figure 7B). Alternatively, the playback session swap input may be provided to a user interface on a control device 130, such as the user interface 430 (Figures 4A-4D) when the user interface 430 is controlling a particular playback device or device(s) (e.g., a zone or zone group).

[0241] The playback device receiving a playback session swap input may be referred to as the “initiating playback device.” Within examples, the initiating playback device is either the source or the target for the swap. When a portable playback device 710 has a playback session on-going (e.g., the portable playback device is actively playing back audio content, or has an active, but paused, playback session) and receives a playback session swap input, the portable playback device 710 may assume that the user wants to “push” the playback session to a nearby playback device 110. Accordingly, the portable playback device 710 is identified as the source of the swap and the nearby playback device 110 is identified as the target.

[0242] To illustrate, Figure 8A is a schematic diagram illustrating an example push swap between the headphones 710a (Figure 7C) and the playback device 110b in the Kitchen 110h zone (Figure 7A). As shown in Figure 8A, initially, the headphones 710a have a playback session on-going. Then, the user provides a playback session swap gesture to the headphones 710a. The playback session on the headphones 710a is pushed to the playback device 110b. After the push, the Kitchen 110h zone receives information about the playback session and continues playback of the playback session, which is on-going on the playback device 110b.

[0243] Conversely, if the portable playback device 710 does not have a playback session on-going and receives a playback session swap input, the portable playback device 710 will assume that the user wants to “pull” a playback session from a nearby playback device 110. Here, the portable playback device 710 is identified as the target of the swap and the nearby playback device 110 is identified as the source. To illustrate, Figure 8B is a schematic diagram illustrating an example pull swap between the headphones 710a (Figure 7C) and the playback device 110b in the Kitchen 110h zone (Figure 7A). As shown in Figure 8B, initially, the

playback device 110b has a playback session on-going. Then, the user provides a playback session swap gesture to the headphones 710a. The playback session on the playback device 110b is pushed to the headphones 710a. After the push, the headphones 710a have the playback session.

[0244] If both the portable playback device 710 and the nearby playback device 110 have on-going playback sessions, it may be unclear whether the user wants to push the playback session on the portable playback device 710 to the nearby playback device 110 or pull the playback session on the nearby playback device 110 to the portable playback device 710. In some implementations the portable playback device 710 may assume that the user would like to “push” the playback session to a nearby playback device 110. To illustrate, Figure 8C is a schematic diagram illustrating an example push swap between the headphones 710a (Figure 7C) and the playback device 110b in the Kitchen 110h zone (Figure 7A). As shown in Figure 8C, initially, both the headphones 710a and the playback device 110b have a playback session on-going. Then, the user provides a playback session swap gesture to the headphones 710a. The playback session on the headphones 710a is pushed to the playback device 110b. After the push, the Kitchen 110h zone has the playback session from the headphones 710a, which is on-going on the playback device 110b. If the user would instead like to “pull” the playback session on the nearby playback device 110 instead, the user may stop the playback session on the portable playback device 710 first, then provide a playback session swap input to the portable playback device 710. In alternate implementations, the portable playback device 710 may be configured to make the opposite assumption.

[0245] In the Figure 8A-8C examples, the initiating device is the headphones 710a. In other examples, the user may provide a playback session swap input to one of the playback devices 110, such as the playback device 110b. In this scenario, similar assumptions may apply in designating the source and target of the swap. In particular, when the playback device 110b has a playback session on-going and receives a playback session swap input, the playback device 110b may assume that the user wants to “push” the playback session to a nearby portable playback device 710. Conversely, if the playback device 110b does not have a playback session on-going and receives a playback session swap input, the playback device 110b will assume that the user wants to “pull” a playback session from a nearby portable playback device 710.

[0246] Within examples, the initiating playback device of a swap may identify the other playback devices in the swap based on proximity to the initiating playback device. That is, the initiating playback device may identify one or more nearby playback devices as targets for a push swap from the initiating playback device or as a source for a pull swap to the initiating

playback device. After or based on receiving a playback session swap input, the initiating playback device may identify such nearby playback devices automatically (i.e., without necessarily receiving further user input from the playback session swap input).

[0247] Some example techniques for identifying nearby playback devices involve audio-based identification. In exemplary audio-based identification techniques, the initiating playback device requests that swap-eligible playback devices emit an identifiable sound (such as an audio chirp), which can be detected by one or more microphones of the initiating playback device. The initiating playback device may then identify nearby playback devices based on the characteristics of the detected sounds.

[0248] To illustrate, Figure 9 is a schematic diagram illustrating an audio-based identification technique using an audio chirp. An audio chirp includes acoustic characteristics (e.g., one or more tones) that enable identification of the playback device transmitting the audio chirp. In Figure 9, a user has initiated a swap on a portable playback device 710 (here, the headphones 710a). As noted above, if a playback session is on-going on the headphones 710a, the headphones 710a will assume that the user desires to push the playback session to one or more nearby playback devices. Otherwise, the headphones 710a will assume that the user desires to pull a playback session on one or more nearby playback devices to the headphones 710a.

[0249] After or based on receiving a playback session swap input, the headphones 710a may identify swap-eligible playback devices in the media playback system 100. For a push swap, the set of swap-eligible playback devices may include playback devices of a certain type or that are assigned a certain role in the media playback system 100. For instance, other portable playback devices may be configured as ineligible for swap. As another example, only the master in a bonded zone (e.g., a stereo pair or surround sound configuration) may be considered swap-eligible. For a pull swap, the set of swap-eligible playback devices include playback devices that have on-going playback sessions. This set may be further narrowed based on other factors, such as those noted above.

[0250] As noted above in section II, the playback devices 110 in the media playback system 100 may maintain or have access to state variables representing the state of the playback devices and other configuration information. This state information is updated periodically or on an event-based basis (e.g., when state changes), such as via subscriptions to certain types of events or statuses (e.g., playback event, grouping event, topology change event, player volume event, group volume event, playback metadata event) and notifications of specific events. The protocol used for the subscriptions may be uPnP based or a proprietary controller protocol or API. Portable playback devices 710, including the headphones 710a and earbuds 719b, may

similarly maintain or have access to these state variables and determine the set of swap-eligible playback devices based on the information in the state variables. The state variables may be received from another playback device in the media playback system and/or received from state information stored in a remote computing system in the cloud. In the Figure 9 example, the headphones 710a have identified the playback device 110b, the playback device 110g, and the playback device 100g as swap-eligible playback devices.

[0251] After identifying the swap-eligible playback devices, the headphones 710a, as the initiating playback device in the swap, causes the swap-eligible playback devices to emit respective audio chirps. For example, the headphones 710a may send instructions to the playback device 110b, the playback device 110g, and the playback device 100g to cause these swap-eligible playback devices to emit unique audio chirps. In some examples, the audio chirps may be ultrasonic (e.g., greater than 20 kHz) or near-ultrasonic (e.g., 19-20 kHz), to avoid propagation of the audio chirp outside the proximity of the emitting playback device and/or to avoid user distraction.

[0252] Each audio chirp may include data in the form of an encoded identifier. Each encoded identifier may be different and encoded as a set of tones, which is known to the initiating playback device. The audio chirps can be transmitted by the swap-eligible playback devices simultaneously, concurrently, or sequentially or when the particular playback device receives the instruction to transmit the audio chirp. In some examples, a device in the media playback system may provide timing information as to when to transmit the audio chirp for each of the playback devices.

[0253] After instructing the swap-eligible playback devices to emit the audio chirps, the initiating playback device in the swap attempts to detect the emitted audio chirps via one or more microphones (e.g., the microphones 715). For instance, the headphones 710a may attempt to detect the emitted audio chirps via one or more voice microphones in the housing of the headphones 710a. Alternatively, the headphones 710a may attempt to detect the emitted audio chirps via one or more ANC microphones in the housing of the headphones 710a. In some instances, particular microphones (ANC or voice) may be selected or tuned to be sensitive to the ultrasonic or near-ultrasonic range such that these microphones are particularly suited to receiving the audio chirp. Other examples are possible as well.

[0254] To identify the “nearby” playback device, the initiating playback device may compare the detected audio chirps. For instance, the headphones 710a may compare various metrics such as sound pressure and signal-to-noise ratio of the detected audio chirps to identify the “loudest” audio chirp, which may be assumed to have been emitted by the playback device that

is physically nearest to the initiating playback device. In example implementations, the initiating playback device may list or otherwise rank the swap-eligible playback devices by relative signal strength (e.g., SNR) and then select the highest ranked swap-eligible playback device as the source or target for the swap.

[0255] As shown in Figure 9, the headphones 710a detect audio chirps emitted by the playback devices 110n and the playback device 110g, which are in the Bathroom 101a and the Bedroom 101c, respectively. However, the headphones 710a did not detect the audio chirp emitted by the playback device 110b, perhaps because the audio chirp could not propagate to the headphones 710a from the Kitchen 101h, which is on a different floor of the house relative to the other zones. In this example, the playback device 110n is determined to be the nearest playback device, as a comparison of the metrics of the audio chirp emitted by the playback device 110n and the audio chirp emitted by the playback device 110g indicated that the audio chirp emitted by the playback device 110n was “loudest.”

[0256] To facilitate comparison between the detect audio chirps, the swap-eligible playback devices may emit the audio chirps at the same or substantially the same volume level. In some cases, the instructions to emit the audio chirps include instructions to change to a certain volume level. Since different playback devices have different types of transducers and/or amplifiers, the volume level for each playback device emitting the chirp may vary based on the type of device. Alternatively, the playback devices may be pre-configured to emit audio chirps at the certain volume level.

[0257] The playback session swap input may take various forms. For instance, a particular input to user interface 713a of the headphones 710 (Figure 7B) such as a tap or gesture to the touch-sensitive region (or a portion thereof) may trigger the swap. In further examples, the portable playback device 710 may include a physical button to trigger the swap. Yet further, a pattern of touch-inputs (e.g. short, long, short) or a tracing pattern (e.g., a shape such as a zig-zag or triangle) may trigger a swap. Other types of inputs are contemplated as well.

[0258] In some particular examples, a touch-and-hold or a continued touch-and-hold to a particular area on the touch-sensitive region (e.g., a play/pause area) triggers a swap. To illustrate, Figure 10 is a chart illustrating an exemplary control scheme for the portable playback device 710c, the playback device(s) 110, and the headphones 710a. As shown in Figure 10, a user may provide a press input (a.k.a. touch) to a touch-sensitive region to perform a primary action (i.e., play or pause). If a physical button is available for swap, the user can push and hold the physical button to invoke a swap.

[0259] If the user continues to hold the press input (touch-and-hold), a secondary action is performed. For the portable playback device 710c and the playback device(s) 110, the secondary action is to group with nearby playback device(s). That is, the initiating playback device (i.e., the portable playback device 710c or the playback device(s) 110) forms a synchrony group with nearby playback devices. In contrast, for the headphones 710a, the secondary action is to perform a push swap or a pull swap, as described in connection with Figures 8A-8C. This configuration allows a user to more quickly access the swap function when using the headphones 710a. Since wearable playback devices are designed for relatively private listening compared with the portable playback device 710c and the playback device(s) 110, the user is less likely to desire to group the headphones for synchronous playback with these types of devices. Other example implementations may vary this control scheme.

[0260] If the user continues to further hold the press input (touch-and-continued hold), a tertiary action is performed. For the portable playback device 710c and the playback device(s) 110, the tertiary action is to perform a push swap or a pull swap with nearby playback device(s). For the headphones 710a, no tertiary action is configured. Other example control schemes may configure a tertiary action for the headphones 710a. In some examples, continuing to hold the touch after the last action in the chain may cancel the input.

[0261] For a user, this control scheme provides audible feedback for the swap action. When a user provides a playback session swap input to a first playback device, the user will be confident that the first playback device will either be source or target of the swap (depending on whether the first playback device has an on-going playback session) as they have provided the triggering input to the first playback device. However, the user may be less confident that the initiating playback device has correctly identified the user's desired target (for a push swap) or source (for a pull swap). In particular, when using the example audio-based identification techniques described above, the initiating playback device has the potential to identify a different playback device as the source or target than the user intended, perhaps due to unique acoustic characteristics of the environment causing an audio chirp emitted from a further away playback device to appear to be nearest.

[0262] By using this control scheme, when providing a press-and-hold input to the portable playback device 710c or the playback device(s) 110, a grouping between the sources and targets of the potential swap (if the user continues to hold) occurs, causing synchronous out-loud playback on the sources and targets of the potential swap. In particular, when the initiating playback device has an on-going playback session, a push-group is performed, which causes the initiating playback device and the nearby playback devices to play back the on-going

playback session in synchrony. Conversely, when the initiating playback device does not have an on-going playback session, a pull-group is performed, which causes the initiating playback device and the nearby playback devices to play back an on-going playback session of the nearby playback devices in synchrony. This out-loud synchronous playback provides a preview to the user on the sources and targets of the swap that will occur if the user continues to hold the input. If the “previewed” playback devices in the group are different from the user’s desired swap source or target, the user can provide input to cancel the group and/or swap actions.

[0263] Further, in some implementations, the control scheme may facilitate a user selecting the desired source or target of a swap by providing additional input. In particular, in some examples, the user may cycle through swap eligible playback devices by providing one or more additional press-and-hold inputs within a threshold period of providing the preceding input. As noted above, the initiating playback device may list the swap eligible playback devices by signal strength. In an example, a second press-and-hold input following the initial press-and-hold input will select the second swap eligible playback devices in the list. Similarly, a third press-and-hold input following the initial press-and-hold input will select the third swap eligible playback devices in the list. Subsequent inputs will continue to cycle down the list (if further swap eligible playback devices are listed).

[0264] In some examples, the portable playback device 710c may similarly traverse a list of group eligible playback devices via consecutive touch-and-hold inputs. The group eligible playback devices may be the same as the swap eligible playback devices, and identified using the same or similar audio-based identification techniques. For example, to push/pull group with the nearest playback device 110, a user may provide a first touch-and-hold input to the playback device 710c. To push/pull group with the next nearest playback device 110, a user may provide a second touch-and-hold input to the playback device 710c within a threshold period of time from the first input. Subsequent touch-and-hold inputs may further traverse the ranked list of swap and/or group eligible playback devices in the ranked order of nearest to furthest. After the threshold period of time, the user would need to start the input sequence again to perform the grouping.

[0265] If a push-swap gesture is performed on the portable playback device 710c while the portable playback device 710c is already grouped with the nearest playback device 110, instead of performing a push-swap, the portable playback device 710c may be configured to push-ungroup instead.

[0266] To aid a user in understanding the control scheme, the source and/or target playback device(s) may provide feedback, including audio and/or visual feedback. To illustrate, Figure

11 Figure 10 is a chart showing an exemplary feedback scheme for the portable playback device 710c and the playback device(s) 110. As shown in Figure 11, at each stage in the control scheme, in conjunction with the respective action, the initiating playback device (the portable playback device 710c), which is the source in this example, provides audio and/or visual feedback. Further, when performing the group and swap actions, the target playback device also provides audio and/or visual feedback. For instance, when grouping the portable playback device 710c and the playback device(s) 110, each playback device provides respective tonal feedback (denoted in Figure 11 as two different tones “Marco” and “Polo”) and the source playback device provides visual feedback. When swapping the portable playback device 710c and the playback device(s) 110, each playback device provides tonal and visual feedback or only the portable playback device 710c provides tonal and visual feedback.

[0267] After identifying the source or target playback devices of a pull or push swap, respectively, the initiating playback device causes the playback session to transition from the source playback device(s) to the target playback devices. In example implementations, a swap involves forming a synchrony group including the source playback device(s) and the target playback devices. Exemplary synchrony grouping is described in greater detail in sections III and IV above. In forming the synchrony group, the target playback device(s) begin to play back audio content in synchrony with the source playback device(s). The source playback device(s) may then be removed from the synchrony group, which completes the swap. The source playback device may be removed or ungrouped from the synchrony group by sending a command to the target device to ungroup from the source playback device.

[0268] In another example implementation, in a pull swap, the target device may send a request to the source playback device for playback session information. The playback session information includes playback state information such as current playlist, track, and offset. In yet another example implementation, in a push swap, the initiating device may send a command to start playback and include the playback state information. The target playback device may use the playback state information to continue playing back the playback session without grouping and ungrouping with the source playback device.

[0269] Referring back to the example of Figure 8A, to push the playback session from the headphones 710a to the playback device 110b, the headphones 710a form a synchrony group including the headphones 710a and the playback device 110b, which causes the headphones 710a and the playback device 110b to synchronously play back the playback session. To finish the push swap, the headphones 710a leave the synchrony group.

[0270] In the Figure 8A example, since the playback session begins on the headphones 710a, the headphones 710a may initially operate as sourcing device of the synchrony group. As noted above, the sourcing device, or group coordinator, obtains the audio for the synchrony group. After the headphones 710a leave the synchrony group, the playback device 110b may assume the role of sourcing device.

[0271] In some examples, during the swap, to avoid the user being disrupted by audio playback during the swap, playback may be manipulated by the source or target playback device. For instance, the playback session may be paused concurrently with creation of the synchrony group and then resumed after the headphones 710a leave the synchrony group. In other examples, either the headphones 710a or the playback device 110b, or both, may be muted until the swap is completed. In yet another example, the headphones 710a may continue playback for x seconds (e.g., 1, 2, 3 seconds or the like) before pausing to allow for any delays in transitioning the playback session to the target playback device. Other examples are possible as well.

[0272] Referring back to the example of Figure 8B, to pull the playback session from the playback device 110b to the headphones 710a, the headphones 710a form a synchrony group including the headphones 710a and the playback device 110b, which causes the headphones 710a and the playback device 110b to synchronously play back the playback session. To finish the push swap, the playback device 110b leaves the synchrony group.

[0273] In the Figure 8B example, since the playback session begins on the playback device 110b, the playback device 110b may initially operate as sourcing device of the synchrony group. As noted above, the sourcing device, or group coordinator, obtains the audio for the synchrony group. After the playback device 110b leave the synchrony group, the headphones 710a may assume the role of sourcing device.

[0274] In alternate embodiments, instead of leaving the synchrony group to complete the swap, the source playback device remains in the synchrony group as the sourcing device. While this would normally cause the source playback device and the target playback devices to play content in synchrony, in these examples, the source playback device is placed into a mute state. Since the source playback device is muted, the playback session appears to have been swapped from the perspective of the user. This may be a true mute that disables or places certain components, such as the audio amplifiers, in a low-power state, which reduces power consumption relative to out loud playback when not in the mute state.

VII. Example Home Theater Swap Techniques

[0275] In some examples, a user may desire to transition a playback session from a soundbar-type playback device to a wearable playback device, to enable more private listening of audio from a television or other home theater source. Example soundbar-type playback devices include the playback device 110h (Figures 1K and 1J). A soundbar-type playback device is capable of receiving audio via an audio input interface from a television, media player (e.g., set-top box, streaming media playback device, computer), or other home theater source. Further, a soundbar-type playback device may operate as a sourcing device for a bonded zone that includes one or more satellites, which may play back certain channels (e.g., the playback devices 110j and 110k) and/or certain frequency ranges (e.g., the playback device 110i), as shown in Figures 1K and 1J illustrating the Den 101d. While some sound-bar playback devices implement a bar-shaped housing to enable multiple audio drivers to be carried linearly along a front surface, a soundbar-type playback device need not necessarily have a bar-shaped housing.

[0276] An exemplary sound-bar playback device may be considered to operate in one of two modes for receiving audio content, referred to herein as a home theater mode and a music mode. In the home theater mode, the soundbar-type playback device receives audio from a physically-connected source (e.g., a television) via the audio input interface. When streaming audio via a network interface, the soundbar-type playback device may be considered to be in the music mode. Notably, the streamed audio is not necessarily music, but may be other types of streamed audio content, such as podcasts or news programs. When streaming audio content in the music mode, the soundbar-type playback device may perform a swap in the same or similar manner as described in section VI.

[0277] While in the home theater mode, to perform a swap, the soundbar-type playback device may enter another mode, referred to herein as a “home theater swap mode,” or simply “swap mode.” When performing a swap action with a wearable playback device using the swap mode, the wearable playback device effectively becomes a satellite of the soundbar-type playback device. In the swap mode, the soundbar-type playback device functions as the source device if playing back audio from the audio input interface in the home theater mode. The wearable playback device then functions as the target playback device to receive and play back the audio from the audio input interface. Conversely, if the wearable playback device is already playing back the audio from the audio input interface in the swap mode, the soundbar-type playback device functions as the target playback device.

[0278] In some cases, the wearable playback device initiates the swap mode. Figure 12A is an example message flow diagram showing instructions exchanged between the headphones

710a, a soundbar-type playback device 110h, one or more satellites in a bonded zone (the Den 101d) with the soundbar-type playback device 110h (illustrated as the playback devices 110j, 110k, and 110i, as shown in Figure 1K), and one or more group members (if the bonded zone is in a zone group with any additional zones) in an example swap mode initiated by the headphones 710a.

[0279] Before entering the swap mode, at 1281a, the soundbar-type playback device 110h is playing back audio from an audio input interface in a home theater mode. As the sourcing device of the bonded zone that includes the satellites, in the home theater mode, the soundbar-type playback device 110h distributes audio to the satellites according to their roles in the bonded zone. Further, if the Den 101d is in a zone group with one or more other zones, the soundbar-type playback device 110h distributes full-range audio content to the group members of the zone group as the sourcing device of the zone group.

[0280] At 1282b, the headphones 710a receive a playback session swap input, which may be a touch-and-hold input, among other examples as described in connection with section VI. In this example, the headphones then identify the soundbar-type playback device 110h as the source for the swap (e.g., based on determining that the soundbar-type playback device 110h is the physically nearest playback device using an audio-based identification technique).

[0281] Then, at 1283a, the headphones 710a send data to the soundbar-type playback device 110h representing instructions to transition to the swap mode, which are received by the soundbar-type playback device 110h. The headphones 710a and the soundbar-type playback device 110h may send and receive the data representing the instructions via respective 802.11-compatible network interfaces. The headphones 710a may send this data based on receiving the playback session swap input.

[0282] Based on receiving the data representing the instructions to enter the swap mode, the soundbar-type playback device 110h transitions from the home theater mode to the swap mode. More particularly, at 1284a, the soundbar-type playback device 110h adds the headphones 710a to a bonded zone, which may be the same bonded zone as the Den 101d or a new bonded zone.

[0283] In some examples, in the home theater mode, the soundbar-type playback device 110h and the satellites operate as nodes in a mesh network. As described above in connection with Figure 1B, in some implementations, the network(s) 104 may include a dedicated communication network implemented as a mesh network. In the home theater mode, the soundbar-type playback device 110h distributes playback timing information and audio to the satellites using the mesh network.

[0284] To facilitate adding headphones 710a to the bonded zone, the soundbar-type playback device 110h transitions its 802.11-compatible network interface from operating as a node in a mesh network to operating as an access point. The access point forms a first wireless local area network (LAN) in a first wireless frequency band (e.g., the 5 GHz band). The soundbar-type playback device 110h then sends, via the 802.11-compatible network interface to the first wearable playback device, data representing a service set identifier (SSID) of the first wireless LAN and credentials for the first wireless LAN, which allows the headphones 710a to connect to the first wireless LAN. After the first wearable playback device connects to the first wireless LAN formed by the soundbar-type playback device, the soundbar-type playback device 110h forms a bonded zone that includes the soundbar-type playback device 110h and the headphones 710a. This may be considered to be the same bonded zone as the Den 101d or a new bonded zone. At 1285a, after connecting to the first wireless LAN, the headphones 710a send a message to the soundbar-type playback device 110h to start streaming the HT audio stream.

[0285] Further, in some examples, while in the swap mode, the headphones 710b effectively becomes a satellite of the soundbar-type playback device 110h. As such, the soundbar-type playback device 110h “parks” the satellite playback devices 110j, 110k, and 110i on a second wireless LAN in a second wireless frequency band (e.g., the 2.4 GHz band) because the satellite playback devices 110j, 110k, and 110i will not be playing back audio. Parking the satellites on the second LAN allows the satellites to remain contactable (e.g., to eventually re-form the bonded zone when transitioning back to the home theater mode) and to receive updates on the state of the media playback system 100 (e.g., state variable events). The soundbar-type playback device 110h may form this second wireless LAN using its 802.11-compatible network interface.

[0286] At 1286a, the soundbar-type playback device 110h stops streaming the HT audio stream to the satellites (e.g., 110j, 110k, and 110i). This may be performed as part of or in connection with parking the satellite playback devices 110j, 110k, and 110i on a second wireless LAN. Similarly, at 1287a, the soundbar-type playback device 110h may stop streaming the HT audio stream to the group members, if any. Forming a new bonded zone at 1284a may remove the soundbar-type playback device 110h from any existing zone group, which causes the group members to stop receiving the HT audio stream.

[0287] At 1288a, the soundbar-type playback device 110h streams the HT audio stream to the headphones 710a for playback. In connection with the headphones 710a receiving the stream and playing back the audio, the soundbar-type playback device 110h mutes to complete the swap. When muted, the soundbar-type playback device 110h may continue to process audio

data for playback synchronously with the headphones 710a. The HT audio stream may include data representing playback timing information for the bonded zone and the audio. In some examples, the audio is multi-channel audio such as a surround sound track. In such examples, the soundbar-type playback device 110h may down-mix the surround sound audio track to an audio track with fewer channels, such as stereo audio track. The surround sound audio track can be down-mixed to contain the same number of channels supported by the wearable device or the portable playback device.

[0288] While in the swap mode, the soundbar-type playback device 110h may detect an event representing a trigger to transition from operating in the swap mode to operating in the home theater mode. Such an event may include receiving, from the headphones 710a, data representing instructions to transition to the home theater mode (e.g., to end the swap mode), which the headphones 710a may send after receiving a playback session swap input while in the swap mode. As another example, the soundbar-type playback device 110h may detect that the headphones 710a have disconnected from the first wireless LAN (and are such no longer operating as a satellite) or been paused for x amount of time. Based on detecting such as event, the soundbar-type playback device 110h may transition to the home theater mode.

[0289] Transition to the home theater mode from the swap mode may involve the soundbar-type playback device 110h transitioning its 802.11-compatible network interface from operating as the access point to operating as the node in the mesh network. Further, the soundbar-type playback device 110h may cause the satellite playback devices to connect to the mesh network. Yet further, the soundbar-type playback device 110h may re-form the bonded zone that includes the soundbar-type playback device 110h and the satellite playback devices 110j, 110k, and 110i.

[0290] Yet further, while in the swap mode, additional wearable playback devices may connect to the soundbar-type playback device 110h as satellites. This may allow, for example, two partners to listen to television audio using individual wearable devices in the Den 101h without waking up sleeping children in the adjacent Bedroom 101c. A user may cause a second wearable playback device (e.g., the earbuds 710b) to join the swap mode by providing a playback session swap input (e.g., a touch-and-hold) to the second wearable device, which causes the second wearable playback device to send data representing instructions to transition to the swap mode to the soundbar-type playback device 110h. The soundbar-type playback device 110h then joins the second wearable device using the techniques shown in Figure 12A.

[0291] In some cases, a control device initiates the swap mode. Figure 12B is an example message flow diagram showing instructions exchanged between the control device 130a, the

headphones 710a, a soundbar-type playback device 110h, and one or more satellites in a bonded zone (the Den 101d) with the soundbar-type playback device 110h in an example swap mode initiated by the control device 130a.

[0292] Before entering the swap mode, at 1281b, the soundbar-type playback device 110h is playing back audio from an audio input interface in a home theater mode. As the sourcing device of the bonded zone that includes the satellites, in the home theater mode, the soundbar-type playback device 110h is a master device that distributes audio to the satellites according to their roles in the bonded zone. Further, if the Den 101d is in a zone group with one or more other zones, the soundbar-type playback device 110h distributes full-range audio content to the group members of the zone group as the sourcing device of the zone group.

[0293] At 1282b, the control device 130a, receive a playback session swap input. The control device 130 may receive a playback session swap input via a user interface, such as the user interface 430. More particularly, a particular user interface 430 may control the headphones 430a, and may include one or more controls that when selected, correspond to the playback session swap input.

[0294] Then, at 1283b, the control device 130a sends a swap command to the headphones, and the headphones send instructions to the soundbar-type playback device to transition to the swap mode. Alternatively, the control device 130a send data to the soundbar-type playback device 110h representing instructions to transition to the swap mode, which are received by the soundbar-type playback device 110h. The control device 130a and the soundbar-type playback device 110h may send and receive the data representing the instructions via respective 802.11-compatible network interfaces. The control device 130a may send this data based on receiving the playback session swap input.

[0295] Based on receiving the data representing the instructions to enter the swap mode, the soundbar-type playback device 110h transitions from the home theater mode to the swap mode. More particularly, at 1284b, the soundbar-type playback device 110h adds the headphones 710a to a bonded zone, which may be the same bonded zone as the Den 101d (e.g., identified as “Den” or a new bonded zone (e.g., identified as “Den + Ben’s Headphones”).

[0296] Similar to the Figure 12A example, in some examples, in the home theater mode, the soundbar-type playback device 110h and the satellites operate as nodes in a mesh network. To facilitate adding headphones 710a to the bonded zone, the soundbar-type playback device 110h transitions its 802.11-compatible network interface from operating as a node in a mesh network to operating as an access point. The access point forms a first wireless local area network (LAN) in a first wireless frequency band (e.g., the 5 GHz band). The soundbar-type playback

device 110h then sends, via the 802.11-compatible network interface to the first wearable playback device, data representing a service set identifier (SSID) of the first wireless LAN and (ii) credentials for the first wireless LAN, which allows the headphones 710a to connect to the first wireless LAN.

[0297] After the first wearable playback device connects to the first wireless LAN formed by the soundbar-type playback device, the soundbar-type playback device 110h forms a bonded zone that includes the soundbar-type playback device 110h and the headphones 710a. This may be considered to be the same bonded zone as the Den 101d or a new bonded zone. At 1285b, after connecting to the first wireless LAN, the headphones 710a send a message to the soundbar-type playback device 110h to start streaming the HT audio stream. At 1286c, the control device receives data indicating that the headphones 710a are ready to receive audio from the soundbar-type playback device 110h.

[0298] Further, in some examples, while in the swap mode, the headphones 710b effectively become a satellite of the soundbar-type playback device 110h. As such, since the headphones 710b are using the first wireless LAN on the first wireless band, the soundbar-type playback device 110h “parks” the satellite playback devices 110j, 110k, and 110i on a second wireless LAN in a second wireless frequency band (e.g., the 2.4 Ghz band). Parking the satellites on the second LAN allows the satellites to remain contactable (e.g., to eventually re-form the bonded zone when transitioning back to the home theater mode) and receive updates on the state of the media playback system 100 (e.g., state variable events). The soundbar-type playback device 110h may form this second wireless LAN using its 802.11-compatible network interface.

[0299] At 1287b, the soundbar-type playback device 110h stops streaming the HT audio stream to the satellites. This may be performed as part of or in connection with parking the satellite playback devices 110j, 110k, and 110i on a second wireless LAN.

[0300] At 1288b, the soundbar-type playback device 110h streams the HT audio stream to the headphones 710a for playback. In connection with the headphones 710a receiving the stream and playing back the audio, the soundbar-type playback device 110h mutes to complete the swap. The HT audio stream may include data representing playback timing information for the bonded zone and the audio. In some examples, the audio is multi-channel audio such as a surround sound track. In such examples, the soundbar-type playback device 110h may down-mix the surround sound audio track to an audio track with fewer channels, such as stereo audio track.

VIII. Example Swap Methods

[0301] Methods 1300A, 1300B, 1400, and 1500 shown in Figures 13A, 13B, 14, and 15 present an example swap techniques, according to example embodiments described herein. These example technique can be implemented within an operating environment including, for example, the media playback system 100 of Figure 7A, one or more of the playback device 110a-n, one or more of the NMDs 130, one or more of the control devices 130, one or one or more of the portable playback devices 710, as well as other devices described herein and/or other suitable devices. Further, operations illustrated by way of example as being performed 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. Methods 1300A, 1300B, 1400, and 1500 may include one or more operations, functions, or actions as illustrated by one or more of blocks shown in Figures 13A, 13B, 14, and 15. 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.

[0302] 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.

a. Example Method To Pull Swap

[0303] Method 1300A illustrates an example pull swap technique. A portable playback device, such as a the headphones 710a, the earbuds 710b, or the portable playback device 710c,

may perform the pull swap technique to pull audio content in a playback session on a playback device 110 to the portable playback device.

[0304] At block 1302A, the method 1300A includes receiving a playback session swap input. For instance, the portable playback device 710 may receive data representing a first playback session swap input. As described in connection with section VI, when the portable playback device 710 is not presently playing audio content, a playback session swap input may initiate a pull swap between the portable playback device 710 and one or more source playback device(s). In some examples, the portable playback device 710 receives the playback session swap input via a user interface. For instance, as discussed in connection with Figure 10, the headphones 710a may receive a touch-and-hold input. Alternatively, the portable playback device 710c may receive a continued touch-and-hold input. In further examples, a control device may receive the playback session swap input and instruct a particular wearable or portable playback device to initiate the playback session swap.

[0305] At block 1304A, the method 1300A includes identifying one or more source playback devices within a media playback system. For instance, the portable playback device 710 may identify one or more eligible playback devices 110 as source playback devices. Eligible source playback devices for a pull swap include playback devices 110 that are connected to a first wireless LAN (e.g., the network 104 in Figure 1B) and also playing back audio content in a playback session. As discussed in section VI, the set of eligible source playback devices may be filtered using various other factors, such as playback device type or role.

[0306] In some examples, the portable playback device 710 identifies the one or more source playback devices via an audio-based identification technique, as described in sections VI. In such examples, identifying the one or more source playback devices may include identifying a set of swap-eligible playback devices in the media playback system and then causing the set of swap-eligible playback devices to emit respective audio chirps that identify the emitting swap-eligible playback devices. The portable playback device 710 may then detect, via one or more microphones, the audio chirps emitted by one or more swap-eligible playback devices and select the one or more source playback devices from among the one or more swap-eligible playback devices based on the audio chirp from the one or more source playback devices indicating that the one or more source playback devices are physically nearest to the portable playback device 710 among the one or more swap-eligible playback devices. Selecting the one or more source playback devices may include comparing one or more respective metrics of the detected audio chirps emitted by one or more swap-eligible playback devices to determine that

the one or more source playback devices are physically nearest to the portable playback device 710 among the one or more swap-eligible playback devices.

[0307] At block 1306A, the method 1300A includes swapping a playback session from the source playback devices to the portable playback device. For example, the portable playback device 710 may transition the playback session from the determined one or more source playback devices to the portable playback device 710. Transitioning the playback session may include forming a first synchrony group including the portable playback device 710 and the one or more source playback devices. Forming the first synchrony group causes the portable playback device 710 to start playing the particular audio content of the playback session.

[0308] Transitioning the playback session may further include causing playback of the particular audio content on the one or more source playback devices to stop. In some examples, playback of the particular audio content on the one or more source playback devices is stopped by the one or more source playback devices leaving the first synchrony group. Alternatively, playback of the particular audio content on the one or more source playback devices is stopped by muting the one or more source playback devices. Other examples are possible as well.

b. Example Method To Push Swap

[0309] Method 1300B illustrates an example push swap technique. A portable playback device, such as the headphones 710a, the earbuds 710b, or the portable playback device 710c, may perform the push swap technique to push audio content in a playback session on the portable playback device to a nearby playback device 110.

[0310] At block 1302B, the method 1300B includes receiving a playback session swap input. For instance, the portable playback device 710 may receive data representing a first playback session swap input. As described in connection with section VI, when the portable playback device 710 is presently playing audio content, a playback session swap input may initiate a push swap between the portable playback device 710 and one or more target playback device(s). In some examples, the portable playback device 710 receives the playback session swap input via a user interface. For instance, as discussed in connection with Figure 10, the headphones 710a may receive a touch-and-hold input. Alternatively, the portable playback device 710c may receive a continued touch-and-hold input. In further examples, a control device may receive the playback session swap input and instruct a particular wearable or portable playback device to initiate the playback session swap.

[0311] At block 1304B, the method 1300B includes identifying one or more source playback devices within a media playback system. For instance, the portable playback device 710 may identify one or more eligible playback devices 110 as target playback devices. Eligible target

playback devices for a pull swap include playback devices 110 that are connected to a first wireless LAN (e.g., the network 104 in Figure 1B) and also playing back audio content in a playback session. As discussed in section VI, the set of eligible target playback devices may be filtered using various other factors, such as playback device type or role.

[0312] In some examples, the portable playback device 710 identifies the one or more target playback devices via an audio-based identification technique, as described in sections VI. In such examples, identifying the one or more target playback devices may include identifying a set of swap-eligible playback devices in the media playback system and then causing the set of swap-eligible playback devices to emit respective audio chirps that identify the emitting swap-eligible playback devices. The portable playback device 710 may then detect, via one or more microphones, the audio chirps emitted by one or more swap-eligible playback devices and select the one or more target playback devices from among the one or more swap-eligible playback devices based on the audio chirp from the one or more source playback devices indicating that the one or more target playback devices are physically nearest to the portable playback device 710 among the one or more swap-eligible playback devices. Selecting the one or more target playback devices may include comparing one or more respective metrics of the detected audio chirps emitted by one or more swap-eligible playback devices to determine that the one or more target playback devices are physically nearest to the portable playback device 710 among the one or more swap-eligible playback devices. The comparison can be performed by any of the devices in the media playback system and/or a remote computing system.

[0313] At block 1306B, the method 1300B includes swapping a playback session from the portable playback device to the one or more target playback devices. For example, the portable playback device 710 may transition its playback session to the one or more target playback devices. Transitioning the playback session may include forming a first synchrony group including the portable playback device 710 and the one or more target playback devices. Forming the first synchrony group causes the one or more target playback devices to start playing the particular audio content of the playback session.

[0314] Transitioning the playback session may further include causing playback of the particular audio content on the portable playback device 710 to stop. In some examples, playback of the particular audio content on the one or more source playback devices is stopped by removing the portable playback device 710 from the first synchrony group. Other examples are possible as well.

c. Example Home Theater Swap Method

[0315] Method 1400 illustrates an example home theater swap technique. A soundbar-type playback device may perform the home theater swap technique to cause a wearable playback device or portable playback device to play back audio received by the soundbar-type playback device and transmitted to the swap target device.

[0316] At block 1402, the method 1400 includes playing back audio while in a home theater mode. For instance, a soundbar-type playback device may play back audio while in a home theater mode. In some examples, the soundbar-type playback device is a master device of a first synchrony group. For instance, an example soundbar-type is the playback device 110h, which may operate as the sourcing device of the Den 101d bonded zone. This bonded zone includes the playback devices 110j and 110k and/or the playback device 110i, as illustrated in Figures 1K and 1J.

[0317] At block 1404, the method 1400 includes receiving instructions to transition to a swap mode. For instance, as illustrated in Figure 12A, the playback device 110h may receive data representing instructions to transition to a swap mode from a wearable playback device, such as the headphones 710a. As another example, as illustrated in Figure 12B, the playback device 110h may receive data representing instructions to transition to a swap mode from a control device 130.

[0318] At block 1406, the method 1400 includes transitioning from the home theater mode to the swap mode. The soundbar-type playback device may transition from the home theater mode to the swap mode based on receiving the data representing the instructions to enter the swap mode.

[0319] As described in connection with Figures 12A and 12B, transitioning from the home theater mode to the swap mode may include various steps. For instance, to facilitate the wearable playback device connecting to the playback device 110h as a satellite, the playback device 110h may transition its 802.11-compatible network interface from operating as a node in a mesh network to operating as an access point that forms a first wireless local area network (LAN) in a first wireless frequency band. Further, the playback device 110h may send, via the 802.11-compatible network interface to the wearable playback device, data representing a service set identifier (SSID) of the first wireless LAN and credentials for the first wireless LAN, which the wearable playback device may use to connect to the first wireless LAN.

[0320] Transitioning from the home theater mode to the swap mode may further include forming a second synchrony group that includes the soundbar-type playback device and the wearable playback device. For instance, the playback device 110h and the headphones 710a may form a second bonded zone after the headphones 710a connect to the first wireless LAN.

After forming the second bonded zone, the playback device 110h may operate as a sourcing device for the second bonded zone. In this role, the playback device 110h sends, to the headphones 710a, data representing playback timing information for the second synchrony group and the audio. The headphones 710a play back the audio according to the timing information, as described in section IV. After forming the second synchrony group, the playback device 110h mutes playback of the audio while the headphones 710a play back the audio.

[0321] Transitioning from the home theater mode to the swap mode may further include parking one or more satellite playback device in a second wireless LAN. For instance, the playback device 110h may cause the playback devices 110j and 110k and/or the playback device 110i to connect to a second wireless LAN in a second wireless frequency band and leave the first synchrony group.

[0322] In further examples, the soundbar-type playback device may add one or more additional wearable playback devices to the swap mode concurrently with a first wearable playback device. For example, while in the swap mode, the playback device 110h may receive, from a second wearable playback device such as the earbuds 710b or another instance of the headphones 710a, data representing instructions to transition to the swap mode. Based on receiving the data representing the instructions to enter the swap mode, the playback device 110h causes the second wearable playback device to join the second synchrony group.

[0323] Causing the second wearable playback device to join the second synchrony group may include sending, to the second wearable playback device, data representing the SSID of the first wireless LAN and credentials for the first wireless LAN. For instance, after the second wearable playback device connects to the first wireless LAN formed by the playback device 110h, the playback device 110h receives an indication from the second playback device that the second playback device is ready for playback and adds the second wearable playback device to the second synchrony group that includes the playback device 110h and the headphones 710b. Then, the playback device 110h sends, to the second wearable playback device, data representing the playback timing information for the second synchrony group and the audio. The second wearable playback device plays back the audio in synchrony with the first wearable playback device based on the playback timing information, as described in connection with section VI.

d. Example Swap Method

[0324] Method 1500 illustrates another example swap method.

[0325] At block 1502, the method 1500 includes detecting a swap trigger. A swap trigger may initiate a playback session swap between one or more source playback device(s) and one or more target playback device(s). In various implementations, a source playback device or a target playback device detects the swap trigger and initiates the playback session swap. Alternatively, another associated device, such as the control device 130 or the bridge device 860, detects the trigger and initiates the playback session swap.

[0326] As described herein, some example swap triggers involve detecting user actions, such as user inputs. For instance, a source playback device (e.g., the portable playback device 710) may detect a particular input representing a swap command and initiate a playback session swap based on detecting the particular input. As another example, the control device 130 may detect a particular input representing a swap command and initiate a playback session swap based on detecting the particular input. Other examples are contemplated as well.

[0327] Other example swap triggers are based on proximity. For instance, some example swap triggers involve detecting proximity between a source playback device (or a paired device, such as a control device 130a) and a target playback device. Further example swap triggers include detecting proximity of a source playback device (or a paired device, such as a control device 130a) to a particular location, such as the home location of the media playback system 100. Other example swap triggers are described throughout, and other suitable swap triggers are contemplated as well.

[0328] At block 1504, the method 1500 includes determining one or more source playback device(s) and one or more target playback device(s). As described above, example implementations involve swapping playback between one or more portable playback devices 710 and one or more playback devices 110. The portable playback device(s) 710 may operate as the source playback device(s) or target playback device(s), depending on context. The playback device(s) 110 may likewise participate in playback session swap as the source playback device(s) or the target playback device(s).

[0329] Within examples, a source playback device is determined based on context. For instance, if the playback device 710 detects a particular input representing a swap command, the playback device 710 may initiate a playback session swap as the source playback device based on detecting this particular input. In another example, if the control device 130 detects a particular input representing a command to swap playback from the playback device 110, the control device 130 may initiate a playback session swap with the playback device 110 as the source playback device or may send data indicating the command to the playback device 110

to cause the playback device 110 to initiate a playback session swap as the source playback device.

[0330] In further examples, the context is based on proximity. For instance, if the portable playback device 710 detects proximity of one or more potential target playback devices 110, the portable playback device 710 may initiate a playback session swap with the portable playback device 710 as the source playback device. As another example, if a paired control device 130 or bridge device 860 detect proximity of one or more potential target playback devices 110 and the paired portable playback device 710 is playing back audio content, the paired control device 130 or bridge device 860 may initiate a playback session swap with the paired portable playback device 710 as the source playback device or may send data indicating the proximate playback device(s) 110 to the paired portable playback device 710 to cause the paired portable playback device 710 to initiate a playback session swap as the source playback device.

[0331] As described above in section V, in some examples, the one or more target devices are determined based on a pre-defined swap pair with the source playback devices. For instance, as illustrated in Figure 11A, the Kitchen 101h is designated as a pre-defined swap pair with the headphones 710a. As described above, swap pairs can be configured and/or re-configured via the control device 130, or other suitable devices.

[0332] Alternatively, as described above in section V, the one or more target devices are determined based on proximity with the source playback devices. Proximity between the source playback device(s) and one or more target devices may be determined using any suitable proximity detection technique including the proximity detection techniques described above in section V. Further, as described above, “proximity” may be defined at one or more ranges, such as a location (e.g., home), zone, area, or individual device.

[0333] Yet further, in other examples, the one or more target devices are determined based on context. For instance, one or more playback devices may detect a particular input indicating a command designating the one or more playback devices as target playback devices. In further examples, the one or more target playback devices are determined based on an association between the target playback devices and a device base. For instance, if the device base 718a is associated with the kitchen 101h, placement of the portable playback device 710c on the device base 718a may trigger a playback session swap between the portable playback device 710c and the playback device 110b.

[0334] When a first playback device 110 is determined as a source or target based on context, one or more additional playback devices 110 may be determined based on a synchrony

grouping between the first playback device 110 and the one or more additional playback devices 110. For instance, if the playback device 110l in the master bedroom 101b is determined as a target device, the playback device 110m is also determined as a source playback device based on the bonded pair configuration of the playback device 110m and the playback device 110l. In another example, if a Kitchen + Dining Room Zone group is configured and the playback device 110d in the Dining Room receives a swap input, the playback device 110b is determined as a source playback device as well. This facilitates a session swap from all playback devices 110 participating in the playback session.

[0335] At block 1506, the method 1500 includes swapping the playback session from the one or more source playback device(s) to the one or more target playback device(s). Within examples, the method 1500 may implement any suitable technique to swap the playback session, such as the example messaging, cloud queue, and grouping techniques described in section V. Other examples are contemplated as well.

IX. Example Bridge Device

[0336] In some example implementations, portable playback devices, such as the headphones 710a, the earbuds 710b, or the portable playback device 710c, may interface with the media playback system 100 via a bridge device 860. Figure 16A illustrates an example pairing arrangement between the headphones 710a and a bridge device 860a. In contrast to a general purpose smartphone or tablet which includes bridging features when configured as the control device 130, the bridge device 860a is configured with hardware and software to interface a portable playback device 710a with the media playback system 100. The bridge device 860a may also include other features to support or enhance the media playback system 100.

[0337] Like the control device 130a, the bridge device 860a may include communications interface(s), processing capabilities, and/or other features that are not necessarily implemented in the portable playback device 710a. When “the portable playback device 710a is “paired” with the bridge device 860a, the portable playback device 710a is able to utilize some of these features. This arrangement may permit the portable playback device 710a to be smaller and more portable, to draw less power, and/or to be less expensive, among other possible benefits. For instance, similar to the control device 130a, the bridge device 860a may include additional communications interfaces as compared with the portable playback device 710a. For instance, the headphones 710a may utilize a cellular data connection of the bridge device 860a to connect to the Internet. As another example, the headphones 710a may utilize a wireless network

interface of the bridge device 860a to connect to the playback device(s) 110 via network 104 or to connect to the Internet.

[0338] In a further example, the portable playback device 710 may be paired with both a mobile device (e.g., a smartphone or tablet, possibly implementing a control device 130 via installation of controller application software) and a bridge device 860. In such an implementation, the portable playback device 710a may stream audio content from the mobile device via a first network interface (e.g., a Bluetooth® network interface) and connect to the bridging device 860 via a second network interface (e.g., a wireless local area network interface). In this arrangement, the mobile device provides a connection to the Internet to facilitate audio streaming and the bridging device 860 functions as an interface to the media playback system 100.

[0339] In an example implementation, the bridging device 860a is bonded to a particular playback device (e.g., playback device 110c), bonded zone of playback device (e.g., playback devices 110l and 110m) or group of playback devices e.g., a “Kitchen + Dining Room” Group). Alternatively, if the home graph hierarchy is utilized, the bridging device 860a may be bonded to a particular Set, Room, or Area. Then, control of the playback device(s) 110 that are bonded to the bridging device 860a via a NMD 120 or a control device 130 also controls the paired portable playback device 710a.

[0340] Alternatively, the bridge device 860a may itself form a zone or Set. For instance, in one example, the bridge device 860a may be configured as a “Ben’s Headphones” zone or “Ben’s Headphones” Set. Configuring the bridge device 860a facilitates control of the paired headphones 710a with the NMD(s) 120 and/or the control device(s) 130 of the media playback system 100.

[0341] Figure 16B is a block diagram of a bridge device 860a comprising an input/output 811. The input/output 811 can include an analog I/O 811a (e.g., one or more wires, cables, and/or other suitable communication links configured to carry analog signals) and/or a digital I/O 811b. The bridge device 860a further includes electronics 812 and a user interface 813 (e.g., one or more buttons, knobs, dials, touch-sensitive surfaces, displays, touchscreens). The bridge device 860a may optionally implement a NMD 820 and include one or more microphones 815 (e.g., a single microphone, a plurality of microphones, a microphone array) (hereinafter referred to as “the microphones 815”) to facilitate voice input.

[0342] In the illustrated embodiment of Figure 16B, the electronics 812 includes one or more processors 812a (referred to hereinafter as “the processors 812a”), memory 812b, software components 812c, a network interface 812d, and power 812i. In some embodiments, the

electronics 112 optionally include one or more other components 812j (e.g., one or more sensors, video displays, touchscreens).

[0343] In some examples, the electronics 812 includes one or more audio processing components 812g (referred to hereinafter as “the audio components 812g”), one or more audio amplifiers 812h (referred to hereinafter as “the amplifiers 812h”), and one or more transducers 814 to facilitate voice responses from the NMD 820. However, audio playback is not the intended purpose of the bridging device and so the audio playback capabilities are generally very limited compared with the playback devices 110 and the portable playback devices 710.

[0344] The processors 812a can comprise clock-driven computing component(s) configured to process data, and the memory 812b can comprise a computer-readable medium (e.g., a tangible, non-transitory computer-readable medium, data storage loaded with one or more of the software components 812c) configured to store instructions for performing various operations and/or functions. The processors 812a are configured to execute the instructions stored on the memory 112b to perform one or more of the operations. The operations can include, for example, pairing with a particular portable playback device 710 and related functions.

[0345] The network interface 812d is configured to facilitate a transmission of data between the bridging device 860a and one or more other devices on a data network such as, for example, the links 103 and/or the network 104 (Figure 1B). In the illustrated embodiment of Figure 16B, the network interface 812d includes one or more wireless interfaces 812e (referred to hereinafter as “the wireless interface 812e”). The wireless interface 812e (e.g., a suitable interface comprising one or more antennae) can be configured to wirelessly communicate with one or more other devices (e.g., one or more of the playback devices 110, NMDs 120, control devices 130, and/or portable playback devices 710) that are communicatively coupled to the network 104 (Figure 1B) in accordance with a suitable wireless communication protocol (e.g., WiFi, Bluetooth, LTE). In some examples, the wireless interface 812e forms an ad-hoc network with a paired portable playback device 710. In some embodiments, the network interface 812d optionally includes a wired interface 812f (e.g., an interface or receptacle configured to receive a network cable such as an Ethernet, a USB-A, USB-C, and/or Thunderbolt cable) configured to communicate over a wired connection with other devices in accordance with a suitable wired communication protocol.

[0346] Figure 16C is a front isometric view of the bridge device 860a configured in accordance with aspects of the disclosed technology as a command device 862a of the media playback system 100. To configure the bridge device 860a as the command device 862a, the

user interface 813a of the bridge device 860a includes playback controls. Example playback controls include transport (e.g., play/pause, skip forward/backward) and volume controls, among other examples. Similar to the control device 130, input to these playback controls are translated into playback commands via the software components 812c and transmitted to one or more playback devices 110 and/or 710 to control playback via the network interface 812d.

[0347] In example implementations, the command device, as compared with the control device(s) 130, is configured to control only paired and/or bonded playback devices, rather than the playback devices 110a-n of the media playback system 100 generally. For instance, in the Figure 16A example in which the bridge device 860a is paired with the portable playback device 710a, playback commands issued on the command device 862a are carried out on the portable playback device 710a. Further, when the bridge device 860a is bonded with one or more playback devices 110, playback commands issued on the command device 862a are also carried out on the bonded playback devices 110.

[0348] The user interface 813a of the bridge device 860a includes a dial 863a to facilitate volume control of paired playback device(s) 710 and/or bonded playback device(s) 110. In this example, the dial 863a is formed by a first portion of the housing 816a rotating about a base of the housing 816a, as illustrated in Figure 16C. Clockwise and counter-clockwise rotations of the dial 863a correspond to upward and downward volume adjustments.

[0349] The user interface 813a of the bridge device 860a also includes a touch-sensitive region 864a to facilitate transport control of paired playback device(s) 710 and/or bonded playback device(s) 110, as shown in Figure 16D. The touch-sensitive region 864a is formed on the top surface of the housing 816a, as shown in Figure 16C. The touch-sensitive region 864a may be implemented as a capacitive or resistive touch-sensitive region, among other examples. In this example, a touch-input to the center of the touch-sensitive region 864a is interpreted as a play/pause toggle. The touch-sensitive region 864a may also interpret certain inputs as skip forward and backward. For instance, touch-inputs to the right and left sides of the touch-sensitive region 864a may be interpreted as a skip forward and skip backward, respectively. Alternatively, a left-to-right swipe gesture on the touch-sensitive region 864a may be interpreted as a skip forward while a right-to-left swipe gesture is interpreted as a skip backward.

[0350] In certain implementations, the user interface 813a of the bridge device 813a is intentionally limited to a certain subset of playback commands as compared with the “full-featured” control supported by the control device(s) 130. As illustrated in Figures 16C and 16D, such a subset may include volume control and transport control (and perhaps only certain transport controls). Such a simplified, minimalist user interface may enhance the user

experience of the paired playback device(s) 710 or bonded playback device(s) 110 by reducing distraction, among other possible benefits.

[0351] In embodiments of the command device 862a that exclude library and/or search controls to select audio content for playback, initiating playback via the command device 862a may start a particular audio container. The particular audio container may be pre-configured by the user via the control device 130 or selected automatically by the media playback system. Example audio containers include playlist, Internet radio stations, albums, and podcasts.

[0352] Figure 16E is a front view of an example bridge device 860b. In contrast to the circular housing 816a of the bridge device 860a, the housing 816b of the bridge device 860b is more rectangular. A user interface 813b of the bridge device 860b includes a dial 863b on the front surface of the housing 816b, a touch-sensitive region 864b on the top surface of the housing 816b, as well as buttons 865a-d on the front surface of the housing 816b. Like the dial 863a, the dial 863b facilitates volume control of portable playback devices 710 paired to the bridge device 860b and/or playback devices 110 bonded to the bridge device 860b. Furthermore, the touch-sensitive region 864b facilitates transport control of paired portable playback devices 710 and/or bonded playback devices 110 in a similar manner to the touch-sensitive region 864a.

[0353] The buttons 865a-d correspond to respective audio containers. The particular audio containers may be pre-configured by the user via the control device 130 or selected automatically by the media playback system (e.g., based on user designated favorites or listening frequency). Selection of a particular button 865 causes the paired playback device 710 and/or bonded playback devices 110 to initiate playback of the corresponding container, similar to how a radio preset tunes a radio to a corresponding radio station.

[0354] For instance, selection of the button 865a causes the bridge device 860b to send one or more instructions to the paired playback device 710 to play back the audio container corresponding to the button 865a. The one or more instructions may include a URI indicating the location of the audio container at a computing device 106 (e.g., a content server of a streaming audio service). The paired playback device 710 then streams the audio container from the computing device 106 and plays back that audio container.

[0355] In some implementations, the bridge device 860 may include a graphical display. In such examples, the user interface 813 of the bridge device 860 may include a graphical user interface displayed on the touch-sensitive graphical display. In some examples, the graphical display is touch-sensitive to facilitate touch input to the graphical user interface. Yet, the graphical user interface may have limited playback controls as compared with the control

devices 430 and 530, which may reduce distraction caused by the presence of the graphical display, among other possible benefits.

[0356] To illustrate, Figure 17A presents a first user interface display 1770a configured for display on a bridge device having a circular touch-sensitive graphical display. For instance, an example implementation of the bridge device 860a may implement the touch-sensitive region 864a as a circular touch-sensitive graphical display. Other shapes and arrangements of touch-sensitive graphical displays are contemplated as well.

[0357] The first user interface display 1770a includes a plurality of regions 1771a-f, which are similar to the buttons 865a-d (Figure 8F). In the first user interface display 1770a, the regions 1770a-g are selectable via touch input to the respective region. Each region 1771 corresponds to a respective audio container. The particular audio containers may be pre-configured by the user via the control device 130 or selected automatically by the media playback system (e.g., based on user designated favorites or listening frequency). Example audio containers include Internet radio stations, playlists, albums, podcasts, as well as other streaming audio content. Selection of a particular button 865 causes the paired playback device 710 and/or bonded playback devices 110 to initiate playback of the corresponding container.

[0358] Presently, by way of illustration, region 1770a is shown in a central position in the first user interface display 1770a. Regions 1771b and 1771f are partially shown at bottom and top positions, respectively, in the first user interface display 1770a. By scrolling the first user interface display 1770a upwards or downwards using upwards or downwards swipe gestures, respectively, regions 1771b or 1771f can be fully displayed and the regions 1771c-e can be displayed as well in round robin fashion. To illustrate, Figure 17B illustrates an upwards swipe causing the region 1771b to move towards the central position. Figure 17C illustrates the region 1771b in the central position after the upwards swipe of Figure 17B. As shown in Figure 17C, regions 1771a and 1771c are partially displayed when region 1771b is in the central position.

[0359] When a particular region 1771 (e.g., region 1771a) is selected, the bridge device 860 causes the paired playback device(s) 710 and/or bonded playback device(s) 110 to initiate playback of the corresponding container. When the region 1771a is selected again while the container is playing back, the bridge device 860 causes the paired playback device(s) 710 and/or bonded playback device(s) 110 to stop playback of the corresponding container. In this manner, the regions 1771 function as a play/pause button.

[0360] Other transport controls may be implemented by the graphical user interface. For instance, as illustrated in Figure 17C, swipe gestures in the first user interface display 1770a

correlate to skip forward and skip backward. In particular, a swipe left may cause a skip forward while a swipe right causes a skip backward.

[0361] Figure 17D presents a second user interface display 1770b, which may be displayed based on selection of region 1771a. The second user interface display 1770b includes a region 1772 that includes a graphical representation of audio content playing back on the paired playback devices 710 and/or bonded playback devices 110, as well as media content information corresponding to the audio content. For instance, if the selected audio container is playing an audio track, metadata corresponding to the audio track is displayed in region 1772.

[0362] The second user interface display 1770b may also include one or more transport controls. To illustrate, the second user interface display 1770b includes a jump forward control 1773a and a jump backward control 1773b. In various implementations, the second user interface display 1770b may include other transport controls as well. For instance, swipe gestures in the second user interface display 1770b may correlate to skip forward and skip backward, similar to the first user interface display 1770a.

[0363] The second user interface display 1770b may further include navigational controls. By way of example, the second user interface display 1770b includes navigational controls 1774a and 1774b. Navigational control 1774a causes the bridge device 860 to display the first user interface display 1770a. Navigational control 1774b causes the bridge device 860 to display a third user interface display 1770c including a queue.

[0364] To illustrate, Figure 17E presents a third user interface display 1770c, which may be displayed based on selection of navigational control 1774b. As shown, the third user interface display 1770c includes an interface to browse within the audio container. Selecting an individual audio track or other media item within the audio container causes playback of that media item. For instance, if the selected audio container is a playlist, the third user interface display 1770c lists the audio tracks of the playlist. As another example, if the selected audio container is a podcast, the third user interface display 1770c may display other audio content available within that container (e.g., episodes of a podcast).

[0365] In some implementations, the graphical user interface facilitates selecting portable playback device(s) 710 to pair with the bridge device 860 and/or playback device(s) 110 to bond with the command device 862. To illustrate, Figure 17F presents a fourth user interface display 1770d with a plurality of toggle controls 1775 corresponding to respective portable playback devices 710 and zones. Toggling a toggle control 1775 pairs or bonds the corresponding portable playback device(s) 710 or playback device(s) 110 with the bridge

device. As shown, the toggle control 1775a corresponding to the headphones 710a is toggled on such that the headphones 710a are paired with the bridge device 860.

[0366] Selection of multiple zone names causes a zone group to be formed between the zones (if not already formed) and the bridge device 860 to be paired with the zone group (and thereby control all playback devices 110 in the zone group). Selection of the "Everywhere" toggle places the media playback system 100 into party mode (in which all playback devices 110 play music synchronously) and pairs the bridge device 860 with all playback devices 110 in the media playback system 100.

[0367] Within example embodiments, the bridge device 860 charges one or more batteries via placement on the device base 718. Figure 18A illustrates placement of the bridge device 860a on a device base 718b. The bridge device 860a may interact with the device base 718b in the same or similar manner as the portable playback device 710c. For instance, if the device base 718b is associated with a zone of the media playback system 100, placement of the bridge device 860a on the device base 718b causes the bridge device 860a (and paired portable device(s) 710) to join the associated zone.

[0368] In example implementations, the bridge device 860a is rotatable about the device base 718b to control volume of the portable playback device 710 paired to the bridge device 860a. In some implementations, rotation of the bridge device 860a about the device base 718b also controls volume of playback device(s) 110 bonded to the bridge device 860a. Similar to the bridge device 718a, the bridge device 860a may rotate with respect to the device base 718b, which may generate a volume control signal in a sensor of the bridge device 860a and/or device base 718b. In another example, a first portion of the device base 718b is rotatable with respect to a second portion of the device base 718b. Rotation of these two portions generates a volume control signal in a sensor of the device base 718b that controls volume of the paired playback device 710 when the bridge device 860a is placed upon the device base 718b.

[0369] The bridge device(s) 860 of the media playback system 100 may also have other features that support the portable device(s) 710 of the media playback system. For instance, the bridge device(s) 860 may support charging the portable device 170. To illustrate, Figure 18B illustrates an example stacking arrangement that includes the device base 718b charging the bridge device 860a, and the bridge device 860a charging the earbuds 710b via a charging case 1080. Similar to the device base 718, the bridge device 860a may charge the earbuds 710b via inductive charging or via conductive terminals. In some implementations, the device base 718b may directly charge the earbuds 710b via placement of the charging case 1080 on the device

base 718b. Other form factors of the charging case 1080 may be used to charge other form factors of the portable playback devices 710 (e.g., the headphones 710a).

[0370] Figure 18C illustrates another example stacking arrangement to facilitate device charging. In this example, the device base 718a charges the portable playback device 710c. The portable playback device 710c charges the bridge device 860a. The bridge device 860a charges the earbuds 710b via a charging case 1080. In this arrangement, only the device base 718a requires external power to charge the various stacked devices.

X. Additional Swap Examples

[0371] In some examples, the source and target of a swap are pre-defined. In a pre-defined swap pair, the source is the playback device 710 or the one or more playback devices 110 that are playing audio content and the target is the other playback device that is not playing back audio content. Playback swapping between the source and targeted playback devices is performed when a swap trigger action such as a button press or other user input is detected.

[0372] In some implementations, an input to the source device of a swap pair triggers a swap. For instance, a particular input to user interface 713a of the headphones 710 (Figure 7B) such as a tap or gesture to the touch-sensitive region (or a portion thereof) may trigger the swap. In further examples, the portable playback device 710 may include a physical button to trigger the swap. Yet further, a pattern of touch-inputs (e.g. short, long, short) or a tracing pattern (e.g., a shape such as a zig-zag or triangle) may trigger a swap. Other types of inputs are contemplated as well.

[0373] Additionally or alternatively, an input to the target device triggers a swap. For instance, a particular input to user interface 113 of the playback device 110a (Figure 1C) may trigger the swap. In further examples, the playback device(s) 110 may include a physical button to trigger the swap. Manipulating the button (e.g., by selecting, touching, sliding, etc.) triggers the swap. Other types of inputs are contemplated as well.

[0374] Within examples, a user interface, such as the user interface 133 of the control device 130a, or the user interface 813 of the bridge device 860a, may facilitate defining a pre-defined swap pair. To illustrate, Figure 19A presents a first user interface display 1931a to facilitate defining a swap pair for the headphones 710a ("Ben's Headphones"). The first user interface display 1931a is configured for display on the control device 430 by way of example but may be adapted to be displayed on other example devices disclosed herein. The control device 430 may display the first user interface display 1931a during a set-up procedure for the headphones 710a. Further, the user may display the first user interface 1931a via a settings user interface display, among other examples.

[0375] As shown, the first user interface display 1931a includes graphical indications of the zones (i.e., zone names) within the media playback system 100 and a toggle control corresponding to each zone. Toggling a toggle control configures the corresponding zone as a swap pair with the headphones 710a. In this example, the kitchen 101h is defined as a swap pair with the headphones 710c. While toggle controls are shown by way of example, other types of controls may be used in alternative implementations. Example user interfaces may include functionally similar user interface displays to define swap pairs for other portable playback devices 710 of the media playback system 100 (e.g., the earbuds 710b and/or the portable playback device 710c). Pre-defined swap pairs may be stored in data storage of the control device(s) 130, the playback device(s) and/or the portable playback devices 710, perhaps as one or more state variables shared among these devices.

[0376] Alternatively, if the home graph hierarchy is implemented, a similar user interface display may include graphical indications of the Sets, Rooms and/or Areas of the Home Graph configured in the media playback system 100. This user interface display may include toggle controls or other similar controls corresponding to each Set, Room and/or Area. In this example, toggling a toggle control configures the corresponding Set, Room and/or Area as a swap pair with the headphones 710a.

[0377] In some implementations, the media playback system 100 may define two or more swap pairs for a portable playback device. To illustrate, Figure 19B presents a second user interface display 1931b to facilitate defining multiple swap pairs for the earbuds 710b. As shown, each pre-defined swap pair corresponds to a different input (e.g., a different gesture). Providing an input corresponding to a particular pre-defined swap pair triggers a swap for that swap pair.

[0378] A user may define custom inputs corresponding to the pre-defined swap pairs. To illustrate, Figure 19C presents a third user interface display 1931c to facilitate defining a custom gesture. As shown, the third user interface display 1931c includes a prompt to provide a custom gesture. After pressing start, the earbuds 710b and the playback devices 110l and 110m in the swap pair monitor their respective user interfaces 713b and 113 to detect the custom input and then store the custom input in data storage.

[0379] In further examples, placement of a portable playback device 710 upon a charging base triggers a swap. For instance, placement of the portable playback device 710c upon the device base 718a (Figure 7F) may trigger a swap. In some implementations, the swap target is pre-defined for the portable playback device 710c.

[0380] Alternatively, the device base 718a may be bonded with one or more particular zones. Then, placement of the portable playback device 710c upon the device base 718a triggers a swap to the one or more particular zones. Additional details regarding bonding zones to a device base can be found, for example, in U.S. Patent No. 9,544,701 entitled, “Base Properties in a Media Playback System,” which, as noted above, is incorporated herein by reference in its entirety.

[0381] In further examples, an input to a user interface of the device base 718a may trigger the swap. Example inputs include a button press (or other manipulation) or a touch-input to a touch-sensitive region, similar to the example inputs described above. For instance, a particular gesture may be interpreted by the device base 718a as a swap trigger.

[0382] In further examples, an input to the user interface 113 of the NMD 120a triggers a swap. For instance, a user may speak a voice input as “swap to kitchen.” As described above in connection with Figures 3A-3D, a user may activate a voice assistance service to process voice input with an activation word or a button press (e.g., push-to-talk). This voice input includes a first command indicating the action (“swap”), and a second command indicating the target playback device for the action (“kitchen”). Here, as described above in connection with Figures 3A-3D, the voice input is transmitted to the voice assistance service and processed. In some cases, the instructions corresponding to the processed voice commands are transmitted back to the source playback device or target playback device to cause the playback session swap to be performed. Alternatively, instructions corresponding to the processed voice commands are transferred to a server to cause the playback session swap to be performed, as described in further detail below in connection with Figures 12B and 12C. Following the swap, the NMD 120a may audibly confirm the swap with a voice response, such as “<audio content name> is now playing in the kitchen.”

[0383] In some cases, both the source playback device and the target playback device are playing audio content when a swap trigger is detected. In such examples, respective playback sessions of the source playback device and the target playback device may be swapped, such that the source playback device begins playing back audio content previously playing on the target and the target begins playing back audio content previously playing on the source. Alternatively, the playback session of the source playback device is swapped to the target playback device and playback is stopped on the target.

[0384] Within example implementations, the source playback device may facilitate a swap by sending playback session data to the target device. The playback session data may include data representing a source of the audio content (e.g., a URI or URL indicating the location of the

audio content), as well as an offset indicating a position within the audio content to start playback. The offset may be defined as a time (e.g., in milliseconds) from the beginning of the audio track or as a number of samples, among other examples. In example implementations, the offset may be set to a playback position in the audio content of the current playback position to allow time for the target device to start buffering the audio content. Then, the source playback device stops playback of the audio content at the offset and the target playback device starts playback of the audio content at the offset.

[0385] The playback session data may further include one or more identifiers corresponding to the playback session. For instance, the playback session data may include a session identifier that distinguishes the playback sessions from other playback sessions. The playback session data may also include an application identifier identifying the media playback system controller application software controlling the playback session. In addition, the playback session data may include a streaming audio service identifier identifying the streaming audio service hosting the audio content at the source, as well as an audio item identifier (e.g., a unique identifier used by the streaming audio service to identify the audio content). As another example, a household identifier may be included in the playback session data to distinguish the media playback system 100 from other media playback systems. As a further example, a group identifier may identify the device(s) in a zone, bonded zone, or zone group.

[0386] The playback session data may further include data representing playback state. Playback state may include a playback state of the session (e.g., playing, paused, or stopped). If the playback session implements a playback queue, the playback session data may include the playback queue state, such as the current playback position within the queue.

[0387] The playback queue state may also include a queue version. For example, in a cloud queue embodiment, the cloud queue server and the media playback system 100 may use the queue version to maintain consistency. The queue version may be incremented each time the queue is modified and then shared between the media playback system 100 and cloud queue server to indicate the most recent version of the queue.

[0388] Further, the playback session data may also include authorization data, such as one or more keys and/or tokens. Such authorization data may include a token associated with the user account. During a playback session swap, the media playback system 100 may verify that the token is authorized on both the source and target playback devices. The authorization data may further include a token associated with the streaming audio service, which may enable the target playback device to access the audio content at the source. Yet further, the authorization data

may include a token associate with the playback session, which enables the target playback device to access the session. Other example authorization data is contemplated as well.

[0389] To illustrate, Figure 20A is an example message flow diagram showing instructions exchanged between the source playback device, the target playback device, and a content server during an example swap of a playback session. Such messages are representative and may include additional or fewer messages. In some implementations, rather than the messages being sent from a portable playback device 710 (as a source or target playback device), the messages are sent from a paired control device 130a (Figure 7G) or paired bridge device 860a (Figure 16A).

[0390] At 2081a, the source playback device starts a playback session. The playback session may be initiated on the source playback device, the control device 130, or the bridge device 860, among other examples. In some cases, the playback session may include one or more additional playback devices playing back in synchrony with the source playback device as part of a group.

[0391] At 2082a, the source playback device detects the swap trigger, such as any of the example swap triggers described above, among other examples. In some cases, another device (e.g., the target playback device, the control device 130, the device base 718, or the bridge device 860) detects the swap trigger and transmits data indicating that a swap trigger was detected to the source playback device.

[0392] At 2083a, the source playback device sends playback session data to the target playback device. As shown by way of example, the playback session data includes data representing URI indicating a source of the currently playing audio content in the session (e.g., the currently playing audio track). The playback session data also includes data representing an offset in the audio content indicating a position in the audio content to start playback. In addition, if the source playback device is playing back audio content from a queue, the playback session data may further include data representing the queue, which may include URIs corresponding to respective media items in the queue, as well as an order of the queued media items. Further, the playback session data includes one or more identifiers, as described above.

[0393] At 2084a, the target playback device sends a fetch message to the content server to request a stream of the audio content from the content server. The fetch message may include a URI indicating the source of the audio content at the content server. The fetch message may further include an offset. The fetch message may include other data as well, such as one or more identifiers and/or authorization data.

[0394] Based on this fetch message, at 2085a, the content server streams the audio content to the target playback device for playback. The content service may begin the stream at the offset in the audio content. The target playback device then begins playing back the audio content at the offset in the audio content.

[0395] At 2086a, the target playback device sends an acknowledge message to the source playback device after receiving the playback session data. In example implementations, the source playback device might not stop the playback session until receiving the acknowledge message from the target playback device. The acknowledgement message may indicate that the swap was successful.

[0396] Other example implementations utilize a cloud queue to facilitate the playback session swap. In contrast to a queue in data storage of the playback devices 110 (i.e., a local queue), a cloud queue for a playback session is maintained in the cloud on a computing device 106. In this implementation, rather than controlling the playback devices 110a-n locally via the networks 104, the control device 130a controls the playback devices 110a-n via the computing device 106 by manipulating the cloud queue on the computing device 106. The computing device 106 synchronizes the cloud queue (or a portion thereof) with the playback device(s) 110 participating in the playback session.

[0397] To illustrate, Figure 20B is an example message flow diagram showing instructions exchanged between the source playback device, a cloud queue server, the target playback device, and a content server during an example swap of a playback session. Such messages are representative and may include additional or fewer messages. In some implementations, rather than the messages being sent from a portable playback device 710 (as a source or target playback device), the messages are sent from a paired control device 130a (Figure 7G) or paired bridge device 860a (Figure 16A).

[0398] At 2081b, the source playback device starts a playback session. The playback session may be initiated on the source playback device, the control device 130, or the bridge device 860, among other examples. In some cases, the playback session may include one or more additional playback devices playing back in synchrony with the source playback device as part of a group.

[0399] At 2082b, the source playback device detects the swap trigger, such as any of the example swap triggers described above, among other examples. In some cases, another device (e.g., the target playback device, the control device 130, the device base 718, or the bridge device 860) detects the swap trigger and transmits data indicating that a swap trigger was detected to the source playback device.

[0400] At 2087, the source playback device sends a swap session message to the cloud queue server that includes playback session data. The swap session message may indicate the target playback device via one or more identifiers. In some examples, such as with a pre-defined swap pair, the cloud queue server may maintain the pre-defined swap pairs for the media playback system 100. The swap session message may also include data representing an offset in the audio content indicating a position in the audio content to start playback. Within examples, the cloud queue server may also track play position in the playback session and may use the position in the swap session message to verify play position. Yet further, the swap session message may include a household identifier identifying the media playback system 100 (so as to distinguish from other media playback systems in other households) as well as one or more player identifiers to identify the source and/or target playback devices.

[0401] Based on receiving the swap session message, the cloud queue server re-targets the session from the source device to the target device. For instance, the cloud queue server may use a household identifier in the playback session data to identify cloud queues of the media playback system 100 and then use the group identifier (or a queue identifier) to identify the cloud queue utilized in the playback session. The cloud queue server may swap this session to the target playback devices altering the cloud queue data to associate the cloud queue with the target playback device. Alternatively, the cloud queue server may mirror the cloud queue of the source device with a cloud queue of the target playback device, then set the playback state of this cloud queue to match the playback state indicated in the playback session data.

[0402] For instance, at 2088, the cloud queue server sends playback session data to the target playback device. The playback session data includes data representing URI indicating a source of the currently playing audio content in the session (e.g., the currently playing audio track). The playback session data also includes data representing an offset in the audio content indicating a position in the audio content to start playback. In addition, if the source playback device is playing back audio content from a cloud queue with multiple audio, the playback session data may further include data representing a window from the cloud queue. The window may indicate media items subsequent to the currently playing audio content as well as possibly media items before the currently playing audio content. The target playback device may queue this window in a local queue to facilitate further playback of the cloud queue in the transferred session.

[0403] At 2084b, the target playback device sends a fetch message to the content server to request a stream of the audio content from the content server. Based on this fetch message, at 2085b, the content server streams the audio content to the target playback device for playback.

The target playback device then begins playing back the audio content at the offset in the audio content.

[0404] Figure 20C is an example message flow diagram showing instructions exchanged between the source playback device, the target playback device, and one or more servers (e.g., the cloud queue server and/or the content server, which may be implemented by one or more cloud servers) during another example swap of a playback session. Such messages are representative and may include additional or fewer messages. In some implementations, rather than the messages being sent from a portable playback device 710 (as a source or target playback device), the messages are sent from a paired control device 130a (Figure 7G) or paired bridge device 860a (Figure 16A).

[0405] At 2081c, the source playback device starts a playback session. The playback session may be initiated on the source playback device, the control device 130, or the bridge device 860, among other examples. In some cases, the playback session may include one or more additional playback devices playing back in synchrony with the source playback device as part of a group.

[0406] At 2082c, the source playback device detects the swap trigger, such as any of the example swap triggers described above, among other examples. In some cases, another device (e.g., the target playback device, the control device 130, the device base 718, or the bridge device 860) detects the swap trigger and transmits data indicating that a swap trigger was detected to the source playback device.

[0407] At 2083b, the source playback device sends playback session data to the target playback device. The playback session data includes one or more identifier, such as a playback session identifier and a queue identifier. The playback session data may also include a URI indicating the source of the audio content, as well as an offset within that content.

[0408] At 2089, the target playback device sends a swap session request to the one or more servers. In a cloud queue implementation, the swap session request may be in the form of a load queue request indicating instructions to load the current cloud queue state of the cloud queue being played back by the source playback device onto the target playback device. To facilitate such requests, the swap session request includes one or more identifiers corresponding to the playback session (e.g., a household identifier, a playback device identifier of the target device, a queue identifier, a playback session identifier).

[0409] Upon receiving the swap session request, the one or more servers facilitate streaming the audio content to the target playback device. For instance, the one or more servers (the content server) may create a new session on the target playback device, such as by instructing

a cloud queue server to create a new session on the target playback device. This request may include the household identifier, application identifier, and user account, as well as other identifiers. The playback session data may be used to mirror the playback session on the source playback device in the new session on the target playback device.

[0410] At 2085c, the content server streams the audio content to the target playback device for playback. The content service may begin the stream at the offset in the audio content. The target playback device then begins playing back the audio content at the offset in the audio content.

[0411] At 2086b, the target playback device sends an acknowledge message to the source playback device after receiving the playback session data. In example implementations, the source playback device might not stop the playback session until receiving the acknowledge message from the target playback device. The acknowledgement message may indicate that the swap was successful.

[0412] In further examples, the source and target playback devices perform a swap by forming a synchrony group. As noted above, example playback device(s) 110 and/or playback device(s) 710 may dynamically form and de-form synchrony groups. As noted above, additional details regarding audio playback synchronization among playback devices and/or zones can be found, for example, in U.S. Patent No. 8,234,395 entitled, "System and method for synchronizing operations among a plurality of independently clocked digital data processing devices," which has been incorporated herein by reference in its entirety.

[0413] In some implementations, the source playback device forms a synchrony group with the target playback device and then mutes its output. When the synchrony group is formed, the target playback device begins playing back the audio content of a given session in synchrony with the source device. To complete the "swap," the source device is muted. From the user's perspective, the playback session appears swapped even though both the source and target playback devices are participating in the session. This mute can be a hidden (e.g., system) mute that is different from a mute command via a user interface. A hidden mute may be performed by lowering the volume or setting the volume to zero on the source device while displaying on the user interface that the source device is unmuted and playback is paused.

[0414] To swap playback back to the source playback device, the target playback device is removed from the synchrony group. A possible advantage of this implementation is that the session can be swapped back to the source device with relatively no delay, as the audio content does not need to re-buffer. Another possible advantage of this implementation is that the source playback device maintains control of the audio stream.

[0415] In further examples, detecting proximity between the source playback device and the target playback device triggers a swap. For instance, detecting that a source playback device and a target playback device of a pre-defined swap pair are in proximity may initiate a swap of a playback session between the source playback device and the target playback device. In some implementations, the source and target playback devices of a swap are defined by proximity of the source playback device to the target playback device. Example proximity detection may be implemented at one or more ranges, such as proximity to the media playback system 100 (i.e., the home or some other known location), proximity to a zone, or proximity to a playback device.

[0416] For instance, in some implementations, proximity of the portable playback device 710 to the media playback system 100 initiates a playback session swap with one or more target playback devices 110 within the home. In an example, upon a user returning home with the portable playback device 710, the paired control device 130a (Figure 7G), or the paired bridge device 860 (Figure 16A), proximity of the portable playback device 710 to the media playback system 100 is detected via a sensor or wireless communications interface of the portable playback device 710, the paired control device 130a (Figure 7G), or the paired bridge device 860 (Figure 16A). This proximity detection initiates a playback session swap between the portable playback device 710 and one or more target playback devices 110 within the home.

[0417] To illustrate, in example implementations, the paired control device 130a (Figure 7G) detects a wireless signal that indicates proximity of the portable playback device 710 to the playback devices 110. For instance, the paired control device 130a (Figure 7G) may detect (e.g., connect to) an 802.11 network (e.g., the network 104) within the home via the network interfaces 132d. Since the playback devices 110a-n are connected to the network 104, detection of this network indicates that the paired control device 130a (and by proxy, the paired portable playback device 710) is in proximity to the home. Other example wireless signals include near-field communication (NFC) and 802.15 (Bluetooth®, Bluetooth® Low Energy) signals which may be transmitted by the playback devices 110a-n within the home. In other examples, the paired bridge device 860 (Figure 16A) may detect such signals or the portable playback device 710 may detect the signals directly via their respective network interface(s).

[0418] Alternatively, the paired control device 130a (Figure 7G) detects proximity to the playback devices 110a-n via one or more sensors. For instance, the paired control device 130a may include a GPS sensor and compare current GPS coordinates to stored GPS coordinates of the home (or other known location of the playback devices 110a-n) to determine whether the paired control device 130a is in proximity to this stored location. In further examples, the paired

control device 130a may detect proximity by using a microphone to detect ultrasonic tones (or other signals) emitted by one or more of the playback devices 110a-n. Alternatively, the paired control device 130a may utilize a camera to detect known objects or signals within the home. Other examples are possible as well.

[0419] In some examples, prior to performing a playback session swap based on proximity, verification is required from the user. In some examples, verification is implemented via an input to a user interface on the source portable playback device 710 (or a paired control device 130 or bridge device 860a). For instance, verification may be implemented via a push notification (or other prompt, such as a widget) displayed on the paired control device 130a. To illustrate, Figure 21A presents a first user interface display 2131a that includes an example push notification 2191a. The paired control device 130a may display the first user interface display 2131a based on detecting proximity to the playback device(s) 110a-n.

[0420] As shown in Figure 21A, the push notification 2191a of the first user interface display 2131a includes multiple selectable controls. A first selectable control (“Swap”) causes the paired control device 130a to perform a playback session swap between the headphones 710a (“Ben’s Headphones”) with the kitchen 101h, which may be a pre-defined swap pair (Figure 11A) or the nearest playback device 110, among other examples. A second selectable control (“Cancel”) cancels the proximity-based swap.

[0421] Also shown in Figure 21A are selectable controls 2192a and 2193b. Rather than transferring an in-progress playback session, selectable control 2192a causes the kitchen 101h to continue a stopped playback session (e.g., playback of a podcast). In example implementations, selectable control 2192a may represent the last-stopped playback session on the portable playback device 710a, the last-stopped playback session in the kitchen 101h, or the last stopped playback session in the media playbacks system 100, among other examples. Alternatively, the push notification 2191a may include multiple selectable controls 2192 to select different last-stopped playback sessions.

[0422] Selectable control 2193a causes the kitchen 101h to start a new playback session that includes playback of a given playlist. In various implementations, example push notifications 2191 may include selectable control 2193 to start new playback sessions with various types of audio containers corresponding to the user. For instance, respective selectable controls 2193 may start new playback sessions with favorite playlists, radio stations, podcasts, albums, or artists, among other examples, similar to the buttons 865 (Figure 16E) and/or the regions 1771 (Figure 17A).

[0423] As further shown in Figure 21A, a third selectable control of the push notification 2191a causes display of a user interface display to select a different swap target. To illustrate, Figure 21B presents a second user interface display 2131b to facilitate selecting a swap target. The user interface display 2131b includes a plurality of toggle controls corresponding to respective zones of the media playback system 100 to facilitate selecting one or more target playback devices 110n.

[0424] In some implementations, proximity of the portable playback device 710 to a zone initiates a playback session swap with the playback device(s) 110 within that zone. Detecting that the portable playback device 710 is in proximity to a given zone may involve detecting a signal (e.g., wireless, ultrasonic) emitted by the playback devices within that zone. In some implementations, detecting signals emitted by other smart devices within the zone may indicate proximity.

[0425] For example, the paired control device 130a may determine a profile corresponding to one or more zones. For instance, while in the kitchen 101h, the paired control device 130a may detect a signal emitted by the playback device 110b, as well as other smart devices (e.g., a smart oven, a smart fridge, a smart power outlet) and save these signals as markers for the kitchen 101h in a profile corresponding to the kitchen 101h. Further, the paired control device 130a may combine this signal data with other sensor data (such as altitude) captured while in the kitchen 101h. The markers in a given profile may also be weighted (e.g., signals of a playback device in a given zone may be weighted more heavily than other smart devices within that zone).

[0426] Given stored profiles for multiple zones in the media playback system 100, to detect whether the portable playback device 710 is in proximity to a given zone, the paired control device 130a may compare current signals and/or sensor data to the stored profile(s) corresponding to the zones. For instance, the paired control device 130a may determine the closest match to the current signals and/or sensor data by comparing how many markers in each profile are present in the current signals and/or sensor data. The paired control device 130a may also threshold the markers by determining proximity to a particular zone when a pre-determined number (or percentage) of markers in the stored profile of the particular zone are also present in the current signals and/or sensor data. While these operations are described by example as being performed by the paired control device 130a, other devices such as the portable playback device 710 and/or the bridge device 860 may also determine the profile and/or detect proximity using the stored profiles.

[0427] Additional techniques to facilitate determining zone proximity can be found, for example, in U.S. Patent App. Pub. No. 2016/0062606 A1 entitled, “Zone Recognition,” which is incorporated herein by reference in its entirety.

[0428] Similar to proximity to the home, prior to performing a playback session swap to a zone based on proximity, the media playback system 100 may request verification that the user intends a swap to be performed. To illustrate, Figure 21C presents a third user interface display 2131c that includes an example push notification 2191b. The paired control device 130a may display the third user interface display 2131c based on detecting proximity to the Den 101d.

[0429] In further examples, proximity to a given zone is determined via a user input to a playback device of that zone. For instance, a particular user input to the playback device 710 (or paired control device 130a or bridge device 860a) may initiate a playback session swap with the playback device 710 as the source playback device. Then, a user input to a given playback device 110 selects that playback device (or associated zone) as the target playback device(s). The source and target playback devices may be configured to perform a swap on the condition that the second input is detected within a pre-determined period of time (e.g., 5 seconds) following the first input, so as to indicate proximity between the source and target playback devices.

[0430] In further examples, another trigger, such as a button press, initiates playback session swap to target playback devices with proximity to a source playback device. To illustrate, Figure 22A illustrates an example playback session swap between the portable playback device 710c and the playback device 110e, which is in proximity to the portable playback device 710c. As shown, a particular swap input (e.g., a long press to the user interface 713c) triggers a playback session swap. In this example, the source playback device (i.e., the portable playback device 710c) is identified via the particular swap input. The target playback device (i.e., the playback device 110e) is identified via proximity detection by the portable playback device.

[0431] As another example, Figure 14B illustrates an example playback session swap between the headphones 710a and the playback device 110e, which is in proximity to the portable playback device 710a. As shown, a hold-close action triggers a playback session swap. In this example, the source playback device (i.e., the headphones 710a) and the target playback device (i.e., the playback device 110e) are both identified by the hold-close action, which causes a near-field communication exchange between the headphones 710a and the playback device 110e. Since near-field communication has a limited range (e.g., 4 cm), the near-field communication exchange indicates proximity between the headphones 710a and the playback device 110e.

[0432] As a further example, Figure 22C illustrates an example playback session swap between the earbuds 710b and the playback device 110e. In this example, a hold close action by the control device 130a (paired to the earbuds 710b) triggers a playback session swap. In this example, the source playback device (i.e., the earbuds 710b) and the target playback device (i.e., the playback device 110e) are both identified by the hold-close action, which causes a near-field communication exchange between the paired control device 130a and the playback device 110e.

[0433] In another example, Figure 22D illustrates another example playback session swap between the earbuds 710b and the playback device 110e. In this example, a hold close action by the bridge device 860a (paired to the earbuds 710b) triggers a playback session swap. In this example, the source playback device (i.e., the earbuds 710b) and the target playback device (i.e., the playback device 110e) are both identified by the hold-close action, which causes a near-field communication exchange between the paired bridge device 860a and the playback device 110e.

[0434] In some instances, the target playback device(s) are members of a synchrony group such as a bonded zone (e.g., a stereo pair such as the master bedroom 101b or a surround sound configuration such as the den 101d) or a zone group (e.g., a “Kitchen + Dining Room” zone group). As noted above, example synchrony techniques involve a group coordinator providing audio content and timing information to one or more group members to facilitate synchronous playback among the group coordinator and the group members. In such examples, a target playback device might be a group coordinator (that provides audio content and timing information to the group members) or a group member (that receives audio content and timing information from the group coordinator).

[0435] In example implementations, when a group coordinator is designated as the target playback device, the group coordinator may, as a result of the synchrony group arrangement, automatically “take along” group members during a playback session swap by providing the group members audio content and timing information corresponding to the swapped playback session. That is, since the group members receive the audio content and timing information from the group coordinator, when the group coordinator starts playing back the swapped playback session, the group members start playing back the swapped playback session as well.

[0436] Generally, when initiating a playback session swap via a GUI of the control device 130a or VUI of the NMD 120a, the bonded zone or zone group is targeted as a whole by reference to the name of the bonded zone, the zone group, or a member zone. In a local implementation, the control device 130a or the NMD 120a may transmit one or more messages

to the group coordinator indicating the playback session swap, which then performs the swap. In a cloud implementation, the control device 130a or the NMD 120a may transmit one or more messages to the cloud queue server indicating the playback session swap to cause the cloud queue server to perform the swap or to relay instructions to the group coordinator to perform the playback session swap.

[0437] In other cases, a group member is targeted for a swap (e.g., by providing input indicating a swap command to the user interface of the group member). In a local implementation, the group coordinator can transmit one or more messages indicating the swap command to the group coordinator, which then performs the playback session swap. In a cloud implementation, the group member may transmit one or more messages to the cloud queue server indicating the playback session swap to cause the cloud queue server to perform the swap or to relay instructions to the group coordinator to perform the playback session swap. Alternatively, the group member may transmit one or more messages to the group coordinator indicating the playback session swap, which cause the group coordinator to transmit a playback session swap request to the cloud server.

XI. Additional Portable Playback Device Examples

[0438] Figure 23A is a front isometric view of earbuds 2310 including an earbud 2310a and an earbud 2310b configured in accordance with aspects of the disclosed technology. As shown, the earbuds 2300 are carried in a charging case 2380.

[0439] Figure 23B is a bottom view of the charging case 2380.

[0440] Figure 23C is a top view of the charging case 2380.

[0441] Figure 23D is a first side view of the charging case 2380.

[0442] Figure 23E is a second side view of the charging case 2380.

[0443] Figure 23F is a front isometric view of the earbud 2310a and the earbud 2310b illustrating exemplary arrangement with the charging case 2380.

[0444] Figure 23F is an isometric view of the earbud 2310a.

[0445] Figure 23H is a first side view of the earbud 2310a.

[0446] Figure 23I is a second side view of the earbud 2310a.

[0447] Figure 23J is a third side view of the earbud 2310a.

[0448] Figure 23K is a fourth side view of the earbud 2310a.

[0449] Figure 23L is a fifth side view of the earbud 2310a.

[0450] Figure 23M is a sixth side view of the earbud 2310a.

[0451] Figure 24A is a front isometric view of a portable playback device 2410 implemented as a handheld speaker configured in accordance with aspects of the disclosed technology.

- [0452] Figure 24B is a side view of the portable playback device 2410.
- [0453] Figure 24C is a top view of the portable playback device 2410.
- [0454] Figure 24D is a bottom view of the portable playback device 2410.
- [0455] Figure 24E is a front isometric view of the portable playback device 2410 illustrating exemplary arrangement with a device base 2418.
- [0456] Figure 24F is a front isometric view of the portable playback device 2410 illustrating exemplary user inputs to the portable playback device 2410.
- [0457] Figure 25A is a front view of headphones 2510 configured in accordance with aspects of the disclosed technology.
- [0458] Figure 25B is a first side view of the headphones 2510.
- [0459] Figure 25C is a second side view of the headphones 2510.
- [0460] Figure 26A is a front view of headphones 2610 configured in accordance with aspects of the disclosed technology.
- [0461] Figure 26B is a first side view of the headphones 2610.
- [0462] Figure 26C is a second side view of the headphones 2610.

XII. Conclusion

[0463] The above discussions relating to portable playback devices, playback devices, control devices, playback zone configurations, and media content sources provide only some examples of operating environments within which functions and methods described below may be implemented. Other operating environments and configurations of media playback systems, playback devices, and network devices not explicitly described herein may also be applicable and suitable for implementation of the functions and methods.

[0464] 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 ways) to implement such systems, methods, apparatus, and/or articles of manufacture.

[0465] Additionally, references herein to “embodiment” means that a particular feature, structure, or characteristic described in connection with the embodiment can be included in at least one example embodiment of an invention. The appearances of this phrase in various

places in the specification are not necessarily all referring to the same embodiment, nor are separate or alternative embodiments mutually exclusive of other embodiments. As such, the embodiments described herein, explicitly and implicitly understood by one skilled in the art, can be combined with other embodiments.

[0466] 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 embodiments.

[0467] 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.

[0468] Example 1: A method comprising: while a first playback device is playing back audio content during a playback session, detecting a playback session swap trigger corresponding to the playback session; determining (a) one or more source playback devices, the one or more source playback devices comprising the first playback device and (b) one or more target playback devices comprising a second playback device; and based on the playback session swap trigger, transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices.

[0469] Example 2: The method of example 1, wherein transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices comprises: forming a synchrony group that includes the first playback device and the second playback device such that the first playback device and the second playback device play back the audio content in synchrony; and muting the first playback device.

[0470] Example 3: The method of example 1 or 2, wherein transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices comprises: transmitting, to a cloud queue server, instructions to transfer the playback

session from the first playback device to the second playback device, wherein the cloud queue server transfers the playback session to the second playback device based on the instructions.

[0471] Example 4: The method of any preceding example, wherein transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices comprises: transmitting, to the second playback device, data representing (i) an uniform resource identifier (URI) indicating a source of the audio content and (ii) an offset within the audio content, wherein the second playback device streams the audio content from the source of the audio content and plays back the audio content beginning at the offset, and wherein the first playback device stops playing back the audio content at the offset.

[0472] Example 5: The method of any preceding example, wherein the first playback device comprises: at least one processor; data storage; one or more amplifiers; one or more transducers; one or more batteries configured to drive the one or more amplifiers and the one or more transducers; and one or more housings carrying the at least one processor, the data storage, the one or more amplifiers, the one or more transducers, and the one or more batteries, wherein the one or more housings are formed into at least one of (a) headphones or (b) earbuds.

[0473] Example 6: The method of any preceding example, wherein the first playback device is paired, via a first type of wireless connection, to a control device, and wherein the first playback device is connected to the second playback device via the first type of wireless connection and a second type of wireless connection between the control device and the second playback device.

[0474] Example 7: The method of any preceding example, wherein detecting the playback session swap trigger comprises: detecting, via a user interface of the control device, input representing a command to swap the playback session.

[0475] Example 8: The method of any preceding example, wherein detecting the input representing the command to swap the playback session comprises: detecting a touch-and-hold input to a touch-sensitive region on the first playback device, wherein a touch input performs a first action that is not a swap.

[0476] Example 9: The method of any preceding example, wherein detecting the input representing the command to swap the playback session comprises: detecting a touch-and-continued hold input to a touch-sensitive region on the first playback device, wherein a touch input performs a first action and a touch-and-hold performs a group action, and wherein the first action is not a swap.

[0477] Example 10: The method of one of examples 1-5, wherein the first playback device is paired, via a first type of wireless connection, to a bridge device, and wherein the first playback device is connected to the second playback device via the first type of wireless connection and a second type of wireless connection between the bridge device and the second playback device.

[0478] Example 11: The method of example 10, wherein detecting the playback session swap trigger comprises: detecting, via a user interface of the bridge device, input representing a command to swap the playback session.

[0479] Example 12: The method of example 10 or 11, wherein the bridge device comprises a circular housing, and wherein the method further comprises: detecting a rotation of the circular housing; and adjusting a playback volume of the first playback device in proportion to the rotation.

[0480] Example 13: The method of one of examples 1-12, wherein the first playback device comprises: at least one processor; data storage; one or more amplifiers; one or more transducers; one or more batteries configured to drive the one or more amplifiers and the one or more transducers; and a housing carrying the at least one processor, the data storage, the one or more amplifiers, the one or more transducers, and the one or more batteries, wherein the housing is formed into a handheld speaker.

[0481] Example 14: The method of example 13, wherein detecting the playback session swap trigger comprises: detecting that the housing is placed into a device base.

[0482] Example 15: The method of any preceding example, wherein the second playback device excludes a battery and draws current from wall power.

[0483] Example 16: The method of any preceding example, wherein the detecting the playback session swap trigger comprises: detecting proximity of the second playback device to the first playback device.

[0484] Example 17: The method of any preceding example, wherein determining the one or more target playback devices comprises: detecting proximity of the second playback device to the first playback device.

[0485] Example 18: The method of any preceding example, wherein the one or more target playback devices further comprise a third playback device, and wherein determining the one or more target playback devices comprises: determining that the third playback device is configured into a synchrony group with the second playback device.

[0486] Example 19: A system configured to perform the method of any of examples 1-18.

[0487] Example 20: A device configured to perform the method of any of examples 1-18.

[0488] Example 21: A tangible, non-transitory computer-readable media having stored therein instructions executable by one or more processors to perform the method of any of examples 1-18.

[0489] Example 22: A portable playback device comprising: at least one processor; a network interface; one or more amplifiers; one or more transducers; one or more batteries configured to drive the one or more amplifiers and the one or more transducers; and one or more housings formed into (a) earbuds or (b) headphones, the one or more housings carrying the at least one processor, the network interface, the one or more amplifiers, the one or more transducers, and the one or more batteries, and data storage having stored therein instructions executable by one or more processors to perform the method of any of examples 1-18.

[0490] Example 23: A method involving a wearable device, the method comprising: receiving data representing a first playback session swap input; based on receiving the data representing the first playback session swap input, identifying one or more source playback devices within a media playback system that are (a) connected to a first wireless local area network (LAN) and (b) playing back particular audio content in a playback session, wherein the wearable playback device is connected to the first wireless LAN via a 802.11-compatible network interface; and transitioning the playback session from the determined one or more source playback devices to the wearable playback device, wherein transitioning the playback session comprises (i) forming a first synchrony group including the wearable playback device and the one or more source playback devices, wherein forming the first synchrony group causes the wearable playback device to start playing the particular audio content of the playback session, and (ii) causing playback of the particular audio content on the one or more source playback devices to stop.

[0491] Example 24: The method of example 23, wherein identifying the one or more source playback devices comprises: identifying a set of swap-eligible playback devices in the media playback system; causing the set of swap-eligible playback devices to emit respective audio chirps that identify the emitting swap-eligible playback devices; detecting, via one or more microphones, the audio chirps emitted by one or more swap-eligible playback devices; and selecting the one or more source playback devices from among the one or more swap-eligible playback devices based on the audio chirp from the one or more source playback devices indicating that the one or more source playback devices are physically nearest to the wearable playback device among the one or more swap-eligible playback devices.

[0492] Example 25: The method of example 24, wherein the one or more microphones comprise one or more acoustic noise cancelling microphones carried on one or more exterior

surfaces of the one or more wearable housings, and wherein detecting the audio chirps emitted by one or more swap-eligible playback devices comprises: detecting the audio chirps emitted by one or more swap-eligible playback devices via the one or more acoustic noise cancelling microphones.

[0493] Example 26: The method of any preceding example 23-25, wherein selecting the one or more source playback devices from among the one or more swap-eligible playback devices comprises: comparing one or more respective metrics of the detected audio chirps emitted by one or more swap-eligible playback devices to determine that the one or more source playback devices are physically nearest to the wearable playback device among the one or more swap-eligible playback devices.

[0494] Example 27: The method of any preceding example 23-26, further comprising: while playing back audio content in the transitioned playback session, receiving data representing a second playback session swap input; based on the second playback session swap input, identifying one or more target playback devices within the media playback system that are connected to the first wireless LAN; and transitioning the playback session from the determined one or more target playback devices to the wearable playback device, wherein transitioning the playback session comprises (i) forming a second synchrony group including the wearable playback device and the one or more target playback devices, wherein forming the second synchrony group causes the one or more target playback devices to start playing the particular audio content of the playback session, and (ii) removing the wearable playback device from the second synchrony group.

[0495] Example 28: The method of any preceding example 23-27, wherein one or more wearable housings of the wearable playback device comprise a touch-sensitive region, and wherein receiving the data representing the playback session swap input comprises receiving input data representing a touch-and-hold input on the touch-sensitive region.

[0496] Example 29: The method of any preceding example 23-28, wherein receiving the data representing the playback session swap input comprises receiving, via the 802.11-compatible network interface from a controller application on a mobile device, data representing instructions to perform a playback session swap.

[0497] Example 30: The method of any preceding example 23-29, wherein causing playback of the particular audio content on the one or more source playback devices to stop comprises: after forming the synchrony group including the wearable playback device and one or more source devices, causing the one or more source devices to be removed from the synchrony group.

[0498] Example 31: The method of any preceding example 23-30, wherein the one or more source devices comprises a master playback device configured to play back multi-channel audio, and wherein transitioning the playback session comprises: sending, via the 802.11-compatible network interface to the master playback device, data representing instructions to enter a swap mode; sending, via the 802.11-compatible network interface to the master playback device, data representing instructions to enter a swap mode; disconnecting from the first wireless LAN and connecting to the second wireless LAN via the 802.11-compatible network interface; and while connected to the second wireless LAN, receiving, via the 802.11-compatible network interface, data representing (i) playback timing information for the first synchrony group and (ii) the multi-channel audio.

[0499] Example 32: The method of any preceding example 23-31, wherein the wearable playback device comprises: one or more network interfaces, wherein the one or more network interfaces comprises an 802.11-compatible network interface; one or more transducers; one or more amplifiers configured to drive the one or more transducers; one or more batteries; one or more processors; one or more wearable housings, the one or more wearable housings carrying the one or more network interfaces, the one or more transducers, the one or more amplifiers, the one or more batteries, the one or more processors, and data storage having instructions stored thereon that are executable by the one or more processors to cause the wearable playback device to perform the method of any preceding example 23-31.

[0500] Example 33: The method of example 32, wherein one or more wearable housings of the wearable playback device are formed into one of (a) headphones or (b) one or more earbuds.

[0501] Example 34: A system configured to perform the method of any of examples 23-32.

[0502] Example 35: A device configured to perform the method of any of example 23-32.

[0503] Example 36: A tangible, non-transitory computer-readable media having stored therein instructions executable by one or more processors to perform the method of any of examples 23-32.

[0504] Example 37: A method involving a wearable device, the method comprising: receiving data representing a first playback session swap input; based on receiving the data representing the first playback session swap input, identifying one or more source playback devices within a media playback system that are (a) connected to a first wireless local area network (LAN) and (b) playing back particular audio content in a playback session, wherein the wearable playback device is connected to the first wireless LAN via the 802.11-compatible network interface; and transitioning the playback session from the determined one or more source playback devices to the wearable playback device, wherein transitioning the playback

session comprises (i) forming a first synchrony group including the wearable playback device and the one or more source playback devices, wherein forming the first synchrony group causes the wearable playback device to start playing the particular audio content of the playback session, and (ii) causing playback of the particular audio content on the one or more source playback devices to stop.

[0505] Example 38: The method of example 37, wherein identifying the one or more source playback devices comprises: identifying a set of swap-eligible playback devices in the media playback system; causing the set of swap-eligible playback devices to emit respective audio chirps that identify the emitting swap-eligible playback devices; detecting, via the one or more microphones, the audio chirps emitted by one or more swap-eligible playback devices; and selecting the one or more source playback devices from among the one or more swap-eligible playback devices based on the audio chirp from the one or more source playback devices indicating that the one or more source playback devices are physically nearest to the wearable playback device among the one or more swap-eligible playback devices.

[0506] Example 39: The method of example 38, wherein the one or more microphones comprise one or more acoustic noise cancelling microphones carried on one or more exterior surfaces of the one or more wearable housings, and wherein detecting the audio chirps emitted by one or more swap-eligible playback devices comprises: detecting the audio chirps emitted by one or more swap-eligible playback devices via the one or more acoustic noise cancelling microphones.

[0507] Example 40: The method of any preceding example 37-39, wherein selecting the one or more source playback devices from among the one or more swap-eligible playback devices comprises: comparing one or more respective metrics of the detected audio chirps emitted by one or more swap-eligible playback devices to determine that the one or more source playback devices are physically nearest to the wearable playback device among the one or more swap-eligible playback devices.

[0508] Example 41: The method of any preceding example 37-40, further comprising: while playing back audio content in the transitioned playback session, receiving data representing a second playback session swap input; based on the second playback session swap input, identifying one or more target playback devices within the media playback system that are connected to the first wireless LAN; and transitioning the playback session from the determined one or more target playback devices to the wearable playback device, wherein transitioning the playback session comprises (i) forming a second synchrony group including the wearable playback device and the one or more target playback devices, wherein forming the second

synchrony group causes the one or more target playback devices to start playing the particular audio content of the playback session, and (ii) removing the wearable playback device from the second synchrony group.

[0509] Example 42: The method of any preceding example 37-41, wherein one or more wearable housings of the wearable playback device comprise a touch-sensitive region, and wherein receiving the data representing the playback session swap input comprises receiving input data representing a touch-and-hold input on the touch-sensitive region.

[0510] Example 43: The method of any preceding example 37-42, wherein receiving the data representing the playback session swap input comprises receiving, via the 802.11-compatible network interface from a controller application on a mobile device, data representing instructions to perform a playback session swap.

[0511] Example 44: The method of any preceding example 37-43, wherein causing playback of the particular audio content on the one or more source playback devices to stop comprises: after forming the synchrony group including the wearable playback device and one or more source devices, causing the one or more source devices to be removed from the synchrony group.

[0512] Example 45: The method of any preceding example 37-44, wherein the one or more source devices comprises a master playback device configured to play back multi-channel audio, and wherein transitioning the playback session comprises: sending, via the 802.11-compatible network interface to the master playback device, data representing instructions to enter a swap mode; receiving, via the 802.11-compatible network interface to the master playback device, data representing (i) a service set identifier (SSID) of a second wireless LAN, the second wireless LAN formed by the master playback device and (ii) credentials for the second wireless LAN; disconnecting from the first wireless LAN and connecting to the second wireless LAN via the 802.11-compatible network interface; and while connected to the second wireless LAN, receiving, via the 802.11-compatible network interface, data representing (i) playback timing information for the first synchrony group and (ii) the multi-channel audio.

[0513] Example 45: The method of any preceding example 37-44, wherein the wearable playback device comprises: one or more network interfaces, wherein the one or more network interfaces comprises an 802.11-compatible network interface; one or more transducers; one or more amplifiers configured to drive the one or more transducers; one or more batteries; one or more processors; one or more wearable housings, the one or more wearable housings carrying the one or more network interfaces, the one or more transducers, the one or more amplifiers, the one or more batteries, the one or more processors, and data storage having instructions

stored thereon that are executable by the one or more processors to cause the wearable playback device to perform the method of any preceding example 37-44.

[0514] Example 46: The method of example 45, wherein one or more wearable housings of the wearable playback device are formed into one of (a) headphones or (b) one or more earbuds.

[0515] Example 47: A system configured to perform the method of any of examples 37-46.

[0516] Example 48: A device configured to perform the method of any of examples 37-46.

[0517] Example 49: A tangible, non-transitory computer-readable media having stored therein instructions executable by one or more processors to perform the method of any of examples 37-46.

[0518] Example 50: A method involving a first playback device and a second playback device, the method comprising: while in a home theater mode, playing back audio received via the audio input interface, wherein the first playback device is a master device of a first synchrony group; while in the home theater mode, receiving, via an 802.11-compatible network interface from the second playback device, data representing instructions to transition to a swap mode; based on receiving the data representing the instructions to enter the swap mode with the second playback device, transitioning from the home theater mode to the swap mode with the second playback device, wherein transitioning from the home theater mode to the swap mode comprises: transitioning the 802.11-compatible network interface from operating as a node in a mesh network to operating as an access point that forms a first wireless local area network (LAN) in a first wireless frequency band; sending, via the 802.11-compatible network interface to the second playback device, data representing (i) a service set identifier (SSID) of the first wireless LAN and (ii) credentials for the first wireless LAN; after connecting to the first wireless LAN formed by the first playback device, forming a second synchrony group that includes the first playback device and the second playback device; receiving, via an 802.11-compatible network interface to the second playback device, data representing (i) playback timing information for the second synchrony group and (ii) the audio, wherein the second playback device plays back the audio; and after joining the second synchrony group, playing back the audio in synchrony with the first playback device, wherein the first playback device mutes playback of the audio while the second playback device plays back the audio.

[0519] Example 51: The method of example 50, wherein the first synchrony group includes the first playback device and one or more satellite playback devices, wherein the audio comprises multi-channel audio, and wherein playing back the multi-channel audio comprises sending, via the 802.11-compatible network interface to the one or more satellite playback devices, data representing (i) playback timing information for the first synchrony group and

(ii) respective channels of the multi-channel audio, and wherein transitioning from the home theater mode to the swap mode further comprises: causing the one or more satellite playback devices to (i) connect to a second wireless LAN in a second wireless frequency band and (ii) leave the first synchrony group.

[0520] Example 52: The method of example 51, further comprising: detecting an event representing a trigger to transition from operating in the swap mode to operating in the home theater mode; after detecting the event, transitioning from the swap mode to the home theater mode, wherein transitioning from the swap mode to the home theater mode comprises: causing the one or more satellite playback devices to connect to the mesh network; transitioning the 802.11-compatible network interface from operating as the access point to operating as the node in the mesh network; and re-forming the first synchrony group that includes the first playback device and the one or more satellite playback devices; operating in the home theater mode, sending, via the 802.11-compatible network interface to the one or more satellite playback devices, data representing (i) playback timing information for the first synchrony group and (ii) respective channels of the multi-channel audio; and playing back the one or more channels of the multi-channel audio in synchrony with the one or more satellite playback devices playing back the respective channels of the multi-channel audio.

[0521] Example 53: The method of example 52, wherein detecting the event comprises detecting that the first wireless playback device has disconnected from the first wireless LAN.

[0522] Example 54: The method of any preceding example 50-53, wherein the audio received via the audio input interface comprises a surround sound audio track, and wherein the functions further comprise down-mixing the surround sound audio track to a stereo audio track, and wherein sending the data representing the audio comprises sending data representing the stereo audio track to the second playback device.

[0523] Example 55: The method of any preceding example 50-54, further comprising: while in the swap mode, receiving, via the 802.11-compatible network interface from a third playback device, data representing instructions to transition to the swap mode; based on receiving, from the third playback device, the data representing the instructions to enter the swap mode, causing the third playback device to join the second synchrony group, wherein causing the third playback device to join the second synchrony group comprises: sending, via the 802.11-compatible network interface to the third playback device, data representing (i) the SSID of the first wireless LAN and (ii) credentials for the first wireless LAN; after the third playback device connects to the first wireless LAN formed by the first playback device, adding the third playback device to the second synchrony group that includes the first playback device and the

second playback device; and sending, via the 802.11-compatible network interface to the third playback device, data representing (i) the playback timing information for the second synchrony group and (ii) the audio, wherein the third playback device plays back the audio in synchrony with the second playback device.

[0524] Example 56: The method of any preceding example 50-55, further comprising: while in a music mode, playing back audio content received via the one or more network interfaces; while playing back the audio content in the music mode, receiving, via the 802.11-compatible network interface from the second playback device, data representing instructions to form a third synchrony group with the second playback device; forming the third synchrony group with the second playback device, wherein forming the third synchrony group with the second playback device configures the first playback device to play back the audio content in synchrony with the second playback device; and after forming the third synchrony group with the second playback device, leaving the third synchrony group, wherein the second playback device is the master device of the third synchrony group.

[0525] Example 57: The method of example 56, further comprising: receiving data representing a playback session swap trigger; based on receiving the data representing the playback session swap trigger, identifying one or more source playback devices within a media playback system that playing back particular audio content in a playback session, wherein identifying the one or more source devices comprises: identifying a set of swap-eligible playback devices in the media playback system, the set including the first playback device; causing the set of swap-eligible playback devices to emit respective audio chirps that identify the emitting swap-eligible playback devices; detecting, via the one or more microphones, the audio chirps emitted by one or more swap-eligible playback devices from the set of swap-eligible playback devices, the one or more swap-eligible playback devices including the first playback device; and selecting the first playback device as the one or more source playback devices from among the one or more swap-eligible playback devices based on the audio chirp from the first playback device indicating that the first playback device is physically nearest to the second playback device among the one or more swap-eligible playback devices.

[0526] Example 58: The method of example 56, further comprising: while in the music mode and before receiving the data representing instructions to form the third synchrony group with the second playback device, receiving, via the 802.11-compatible network interface from the second playback device, data representing instructions to emit a particular audio chirp; and based on receiving the data representing instructions to emit a particular audio chirp, emitting the particular audio chirp via the one or more transducers.

[0527] Example 59: The method of any preceding example 50-58, wherein the second playback device comprises one or more housings, and wherein the one or more housings are formed into one of (a) headphones or (b) a set of earbuds.

[0528] Example 60: The method of any preceding example 50-59, wherein the first playback device comprises: an audio input interface; one or more network interfaces, wherein the one or more network interfaces comprises an 802.11-compatible network interface; one or more transducers; one or more amplifiers configured to drive the one or more transducers; and a housing carrying the an audio input interface, the one or more network interfaces, the one or more transducers, the one or more amplifiers, the one or more processors, and data storage having instructions stored thereon that are executable by the one or more processors to cause the soundbar-type playback device to perform the method of any preceding example 50-59.

[0529] Example 61: The method of any preceding example 50-60, wherein the second playback device comprises: one or more network interfaces, wherein the one or more network interfaces comprises an 802.11-compatible network interface; one or more transducers; one or more amplifiers configured to drive the one or more transducers; one or more batteries; one or more processors; a housing carrying the one or more network interfaces, the one or more transducers, the one or more amplifiers, the one or more batteries, the one or more processors, and data storage having instructions stored thereon that are executable by the one or more processors to cause the wearable playback device to perform the method of any preceding example 50-60.

[0530] Example 62: A system configured to perform the method of any of examples 50-61.

[0531] Example 64: A device configured to perform the method of any of examples 50-61.

[0532] Example 65: A tangible, non-transitory computer-readable media having stored therein instructions executable by one or more processors to perform the method of any of examples 50-61.

CLAIMS

1. A method comprising:
 - while a first playback device is playing back audio content during a playback session,
 - detecting a playback session swap trigger corresponding to the playback session;
 - determining (a) one or more source playback devices, the one or more source playback devices comprising the first playback device and (b) one or more target playback devices comprising a second playback device; and
 - based on the playback session swap trigger, transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices.
2. The method of claim 1, wherein transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices comprises:
 - forming a synchrony group that includes the first playback device and the second playback device such that the first playback device and the second playback device play back the audio content in synchrony; and
 - muting the first playback device.
3. The method of claim 1 or 2, wherein transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices comprises:
 - transmitting, to a cloud queue server, instructions to transfer the playback session from the first playback device to the second playback device, wherein the cloud queue server transfers the playback session to the second playback device based on the instructions.
4. The method of any preceding claim, wherein transitioning the playback session from the determined one or more source playback devices to the one or more target playback devices comprises:
 - transmitting, to the second playback device, data representing (i) an uniform resource identifier (URI) indicating a source of the audio content and (ii) an offset within the audio content, wherein the second playback device streams the audio content from the source of the audio content and plays back the audio content beginning at the offset, and wherein the first playback device stops playing back the audio content at the offset.

5. The method of any preceding claim, wherein the first playback device comprises:

- at least one processor;
- data storage;
- one or more amplifiers;
- one or more transducers;
- one or more batteries configured to drive the one or more amplifiers and the one or more transducers; and

one or more housings carrying the at least one processor, the data storage, the one or more amplifiers, the one or more transducers, and the one or more batteries, wherein the one or more housings are formed into at least one of (a) headphones or (b) earbuds.

6. The method of any preceding claim, wherein the first playback device is paired, via a first type of wireless connection, to a control device, and wherein the first playback device is connected to the second playback device via the first type of wireless connection and a second type of wireless connection between the control device and the second playback device.

7. The method of any preceding claim, wherein detecting the playback session swap trigger comprises:

- detecting, via a user interface of the control device, input representing a command to swap the playback session.

8. The method of any preceding claim, wherein detecting the input representing the command to swap the playback session comprises:

- detecting a touch-and-hold input to a touch-sensitive region on the first playback device, wherein a touch input performs a first action that is not a swap.

9. The method of any preceding claim, wherein detecting the input representing the command to swap the playback session comprises:

- detecting a touch-and-continued hold input to a touch-sensitive region on the first playback device, wherein a touch input performs a first action and a touch-and-hold performs a group action, and wherein the first action is not a swap.

10. The method of one of claims 1-5, wherein the first playback device is paired, via a first type of wireless connection, to a bridge device, and wherein the first playback device is connected to the second playback device via the first type of wireless connection and a second type of wireless connection between the bridge device and the second playback device.

11. The method of claim 10, wherein detecting the playback session swap trigger comprises:

detecting, via a user interface of the bridge device, input representing a command to swap the playback session.

12. The method of claim 10 or 11, wherein the bridge device comprises a circular housing, and wherein the method further comprises:

detecting a rotation of the circular housing; and

adjusting a playback volume of the first playback device in proportion to the rotation.

13. The method of one of claims 1-12, wherein the first playback device comprises:

at least one processor;

data storage;

one or more amplifiers;

one or more transducers;

one or more batteries configured to drive the one or more amplifiers and the one or more transducers; and

a housing carrying the at least one processor, the data storage, the one or more amplifiers, the one or more transducers, and the one or more batteries, wherein the housing is formed into a handheld speaker.

14. The method of claim 13, wherein detecting the playback session swap trigger comprises:

detecting that the housing is placed into a device base.

15. The method of any preceding claim, wherein the second playback device excludes a battery and draws current from wall power.

16. The method of any preceding claim, wherein the detecting the playback session swap trigger comprises:

detecting proximity of the second playback device to the first playback device.

17. The method of any preceding claim, wherein determining the one or more target playback devices comprises:

detecting proximity of the second playback device to the first playback device.

18. The method of any preceding claim, wherein the one or more target playback devices further comprise a third playback device, and wherein determining the one or more target playback devices comprises:

determining that the third playback device is configured into a synchrony group with the second playback device.

19. A system configured to perform the method of any of claims 1-18.

20. A device configured to perform the method of any of claims 1-18.

21. A tangible, non-transitory computer-readable media having stored therein instructions executable by one or more processors to perform the method of any of claims 1-18.

22. A portable playback device comprising:

at least one processor;

a network interface;

one or more amplifiers;

one or more transducers;

one or more batteries configured to drive the one or more amplifiers and the one or more transducers; and

one or more housings formed into (a) earbuds or (b) headphones, the one or more housings carrying the at least one processor, the network interface, the one or more amplifiers, the one or more transducers, and the one or more batteries, and data storage having stored therein instructions executable by one or more processors to perform the method of any of claims 1-18.

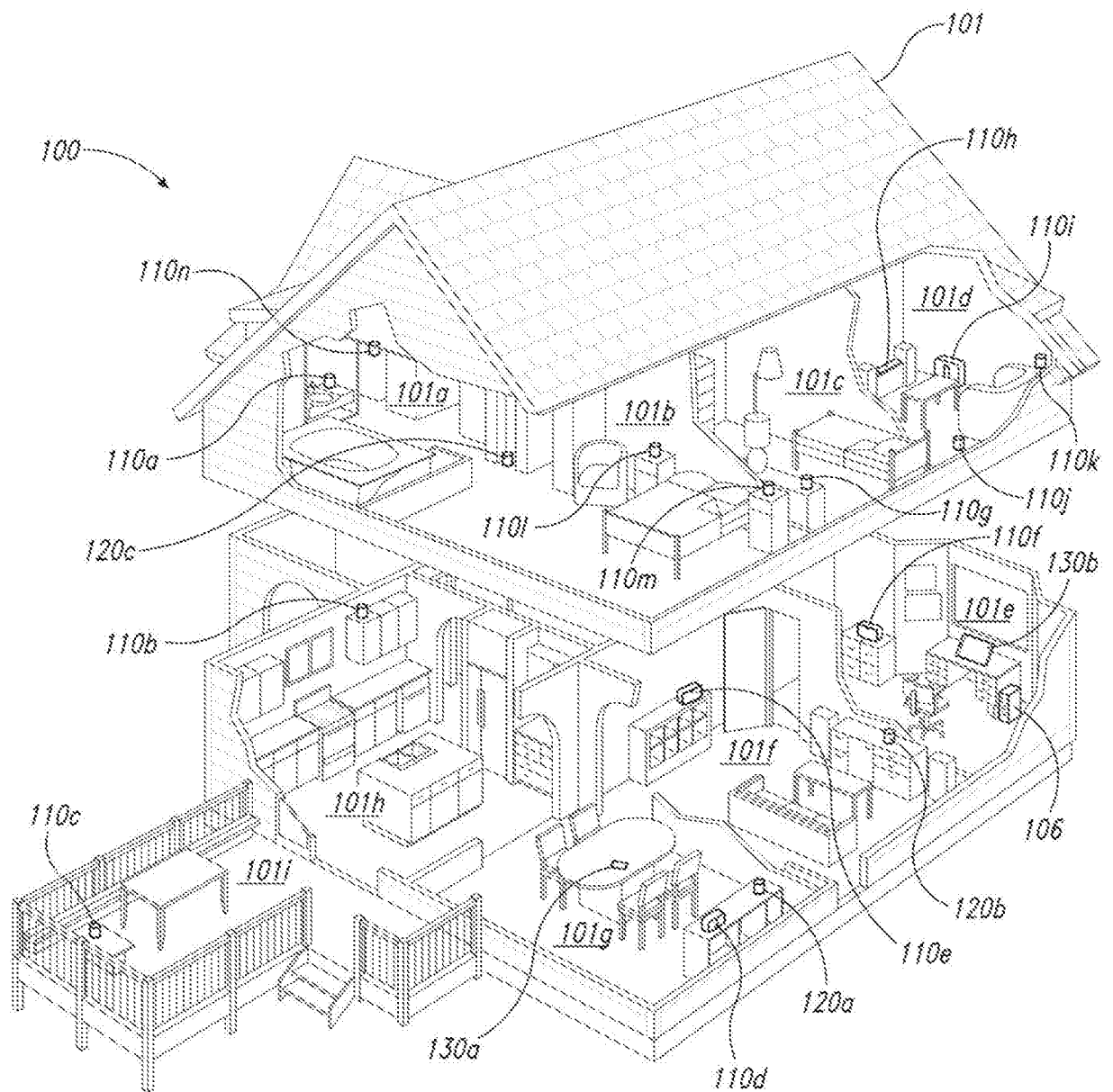


Fig. 1A

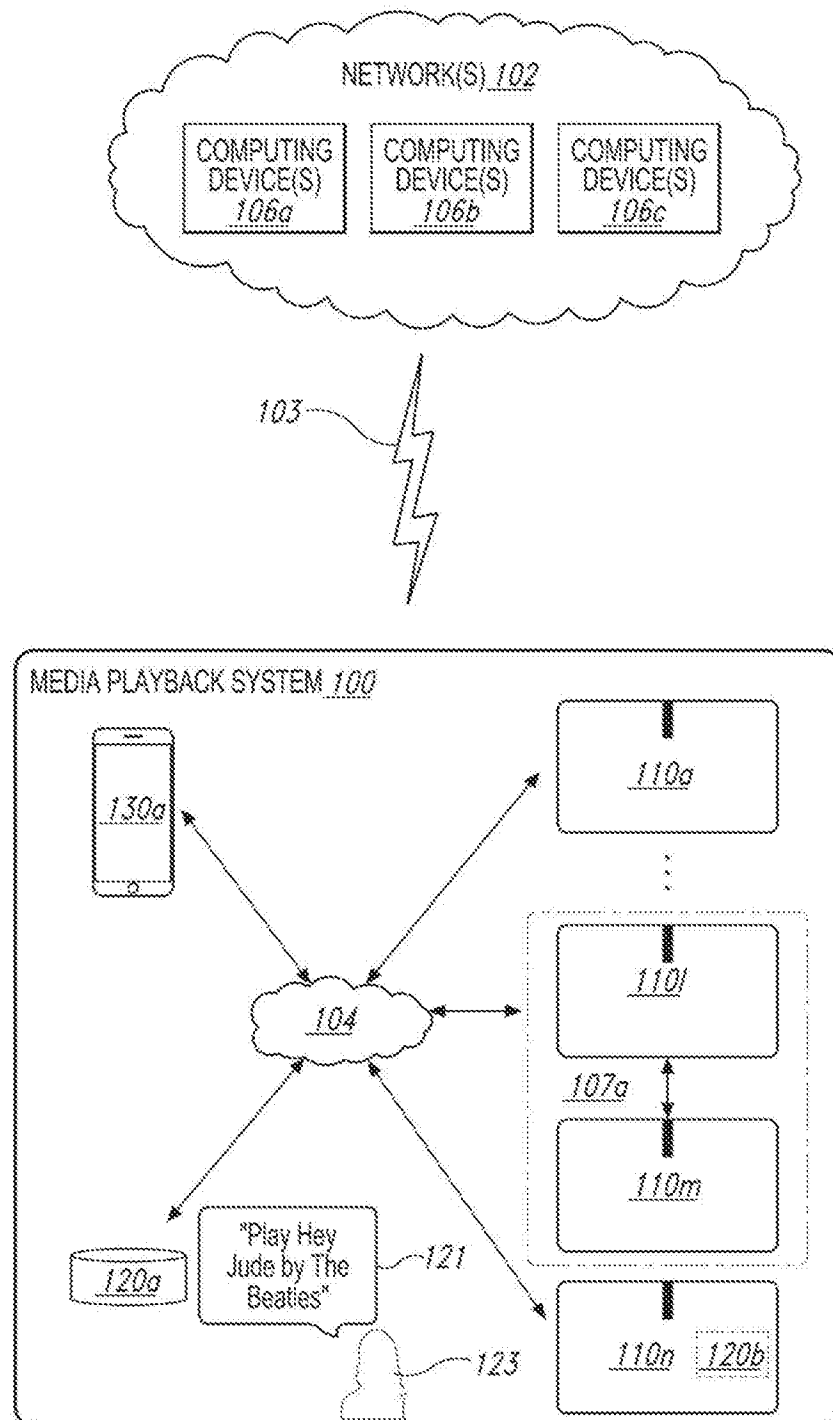


Fig. 1B

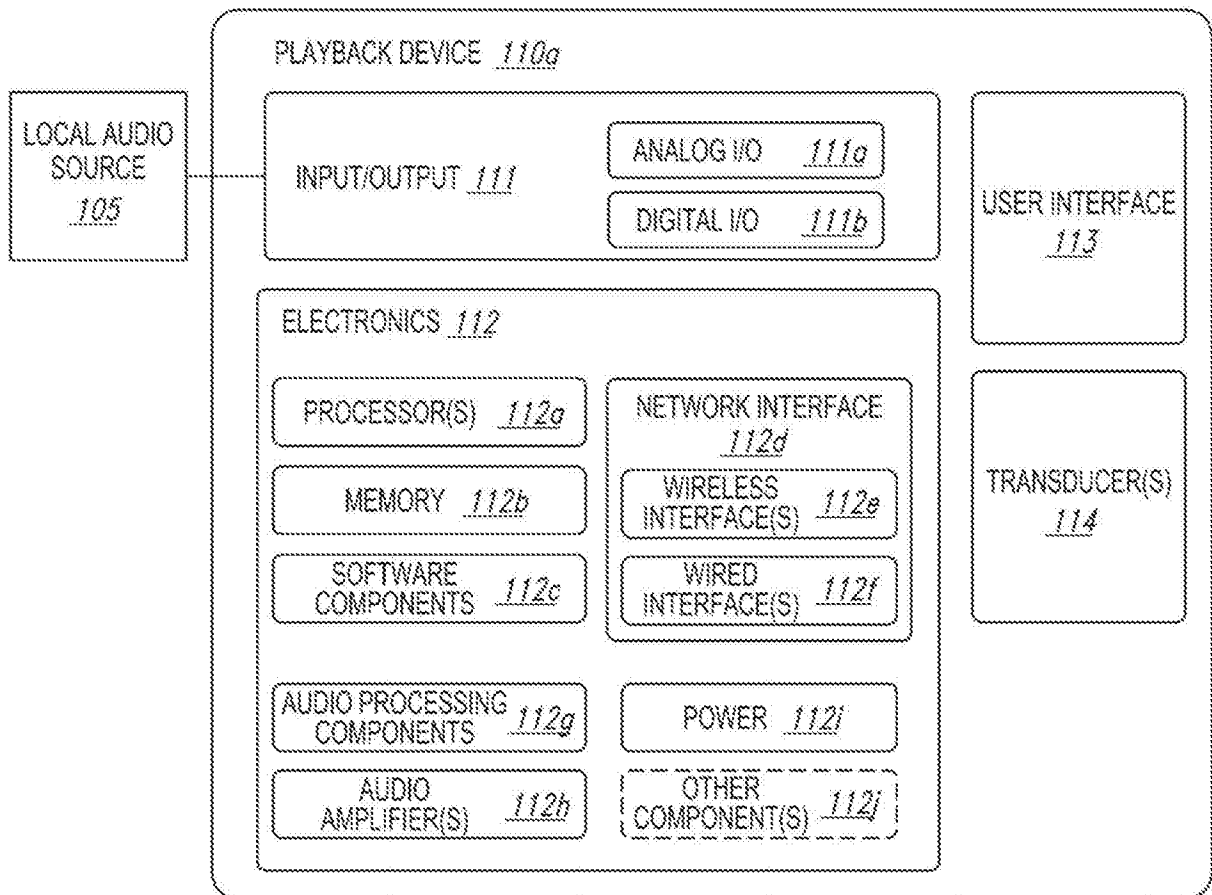


Fig. 1C

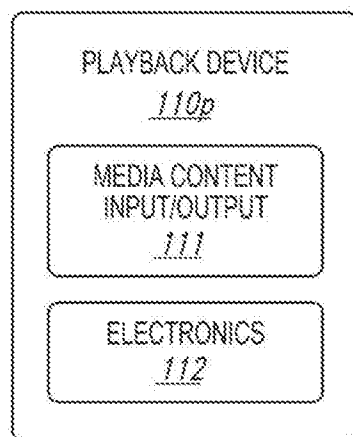


Fig. 1D

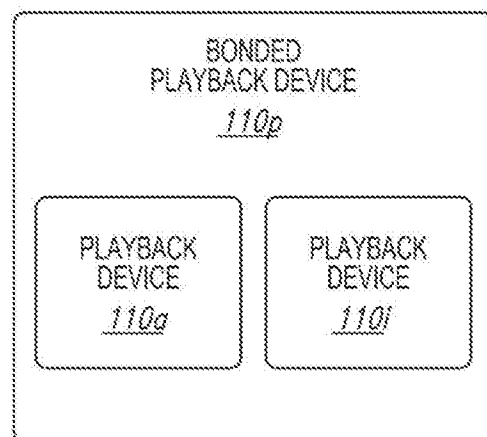


Fig. 1E

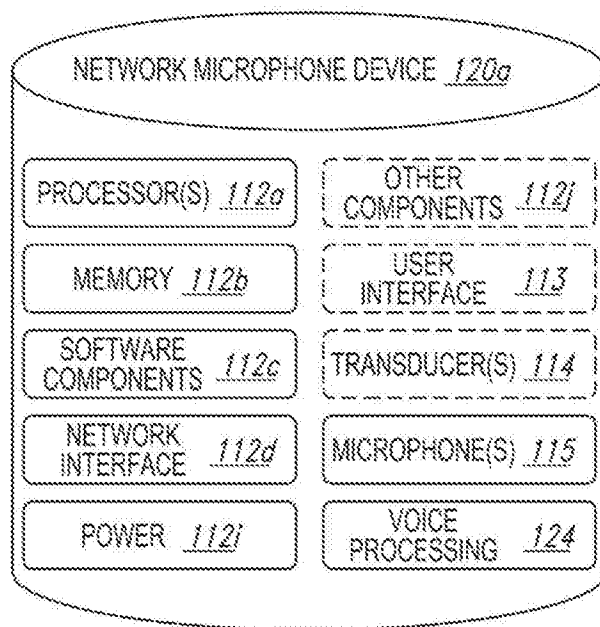


Fig. 1F

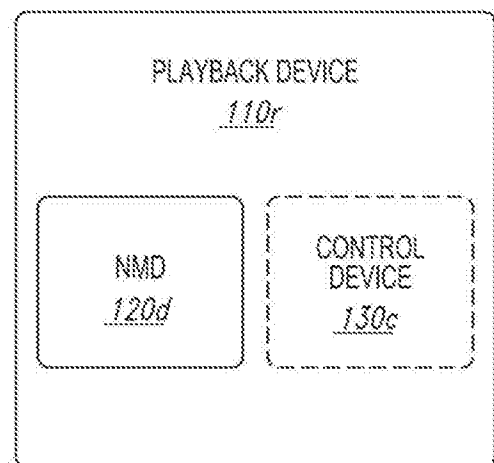


Fig. 1G

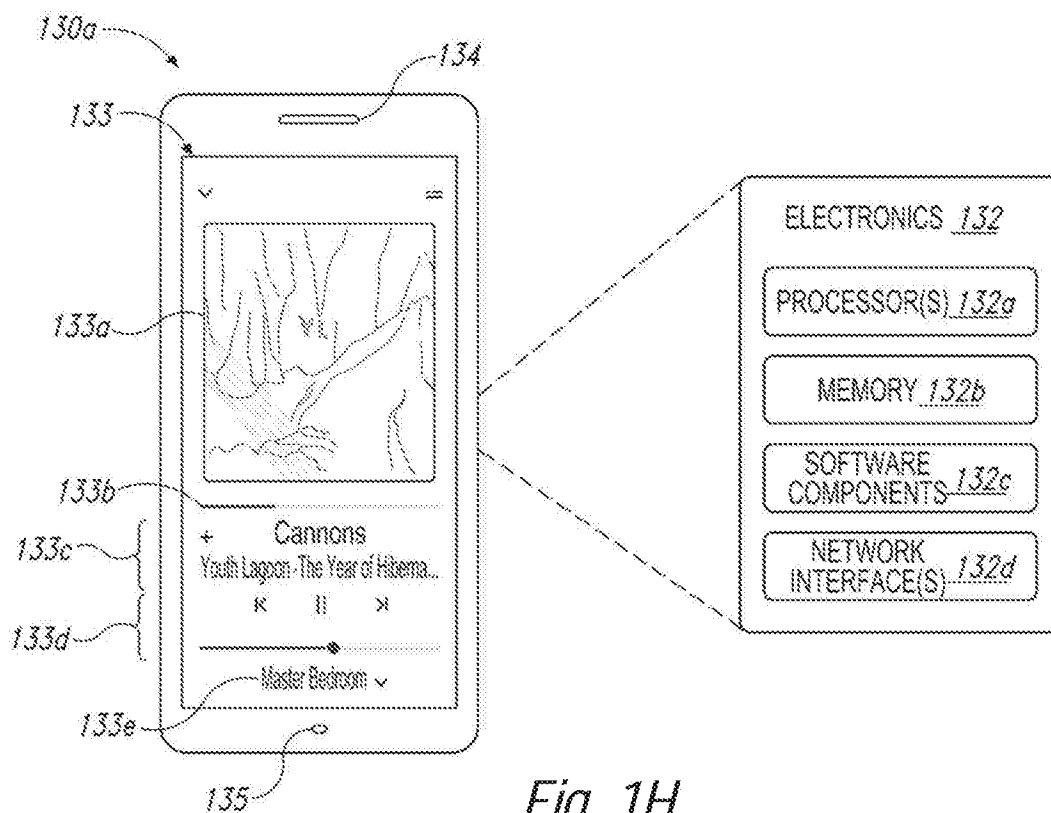


Fig. 1H



Fig. 1I

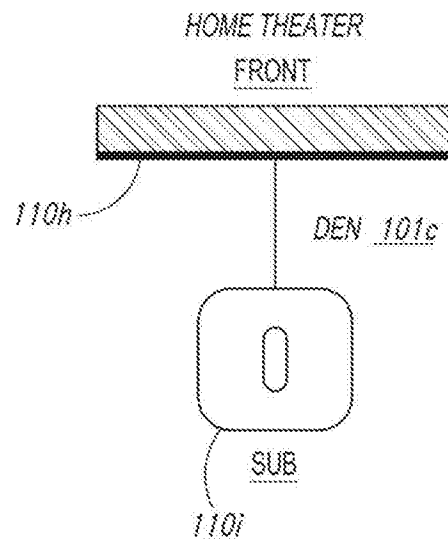


Fig. 1J

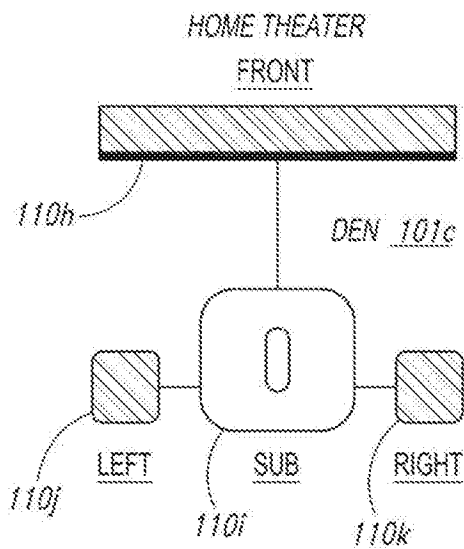


Fig. 1K

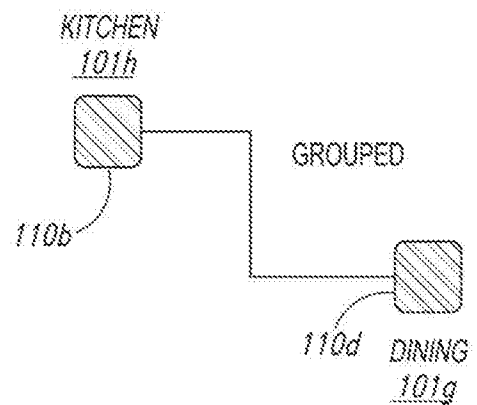


Fig. 1L

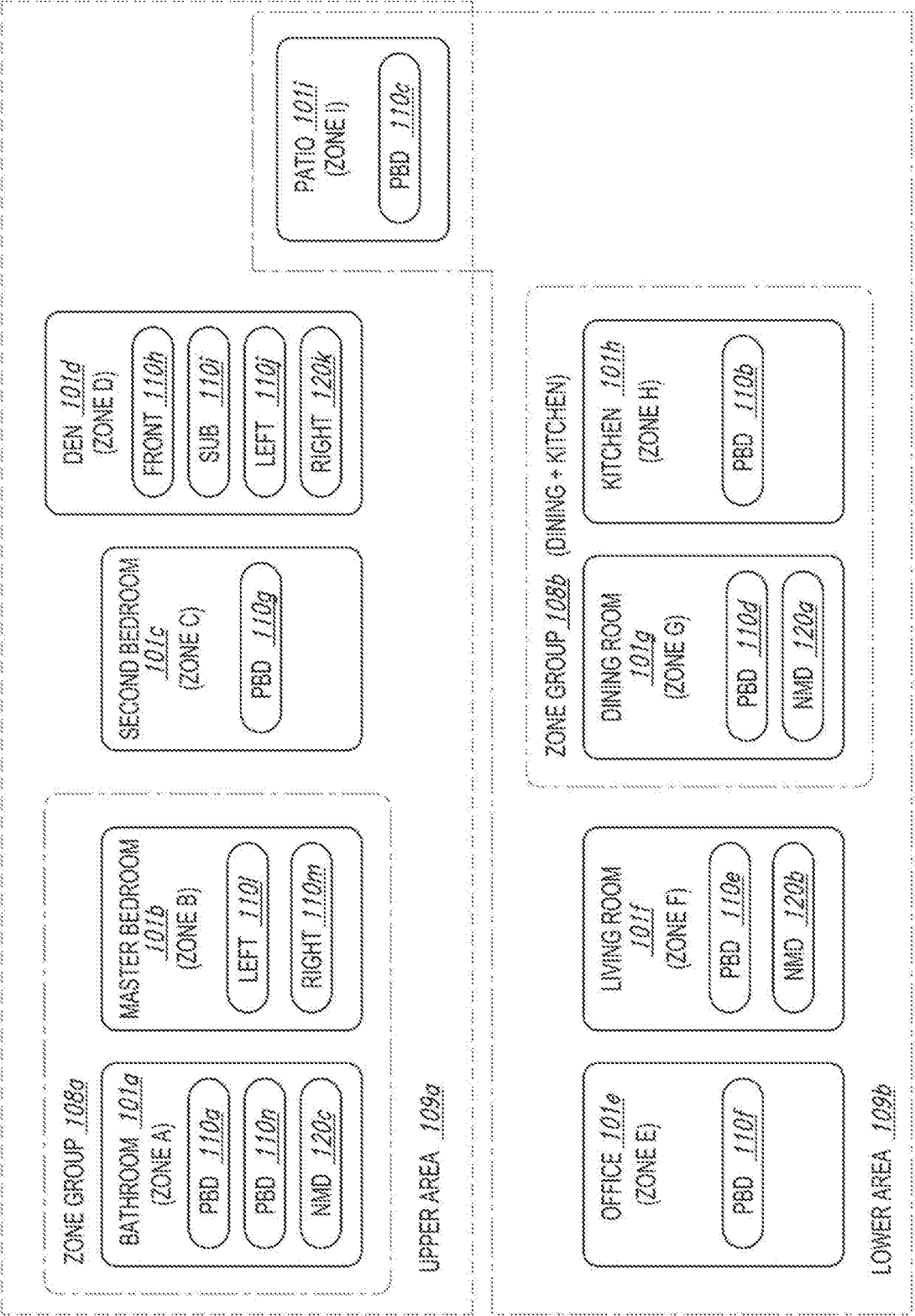


Fig. 1M

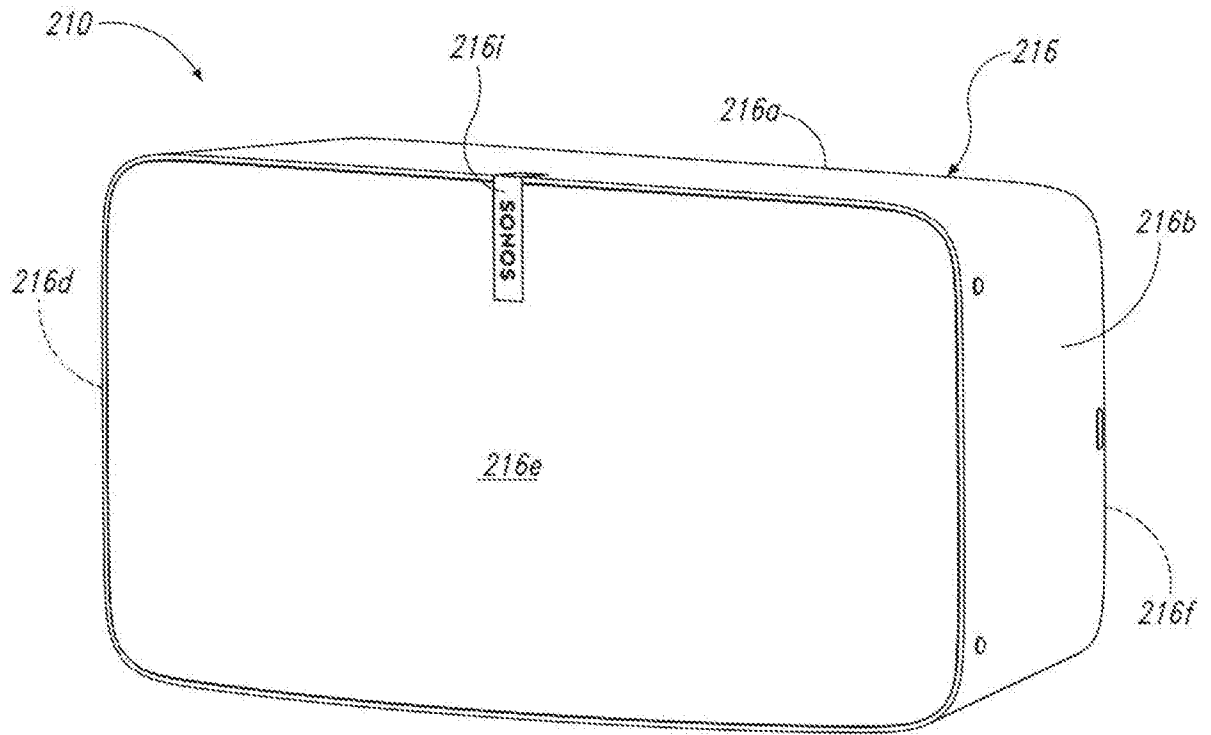


Fig. 2A

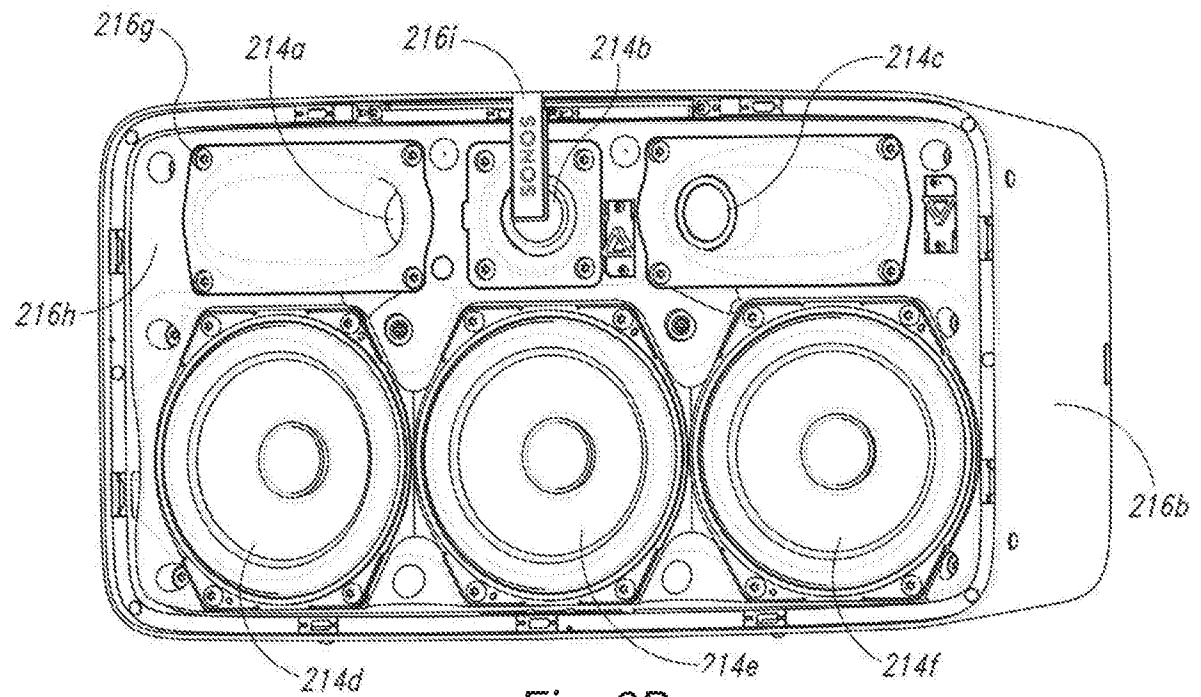


Fig. 2B

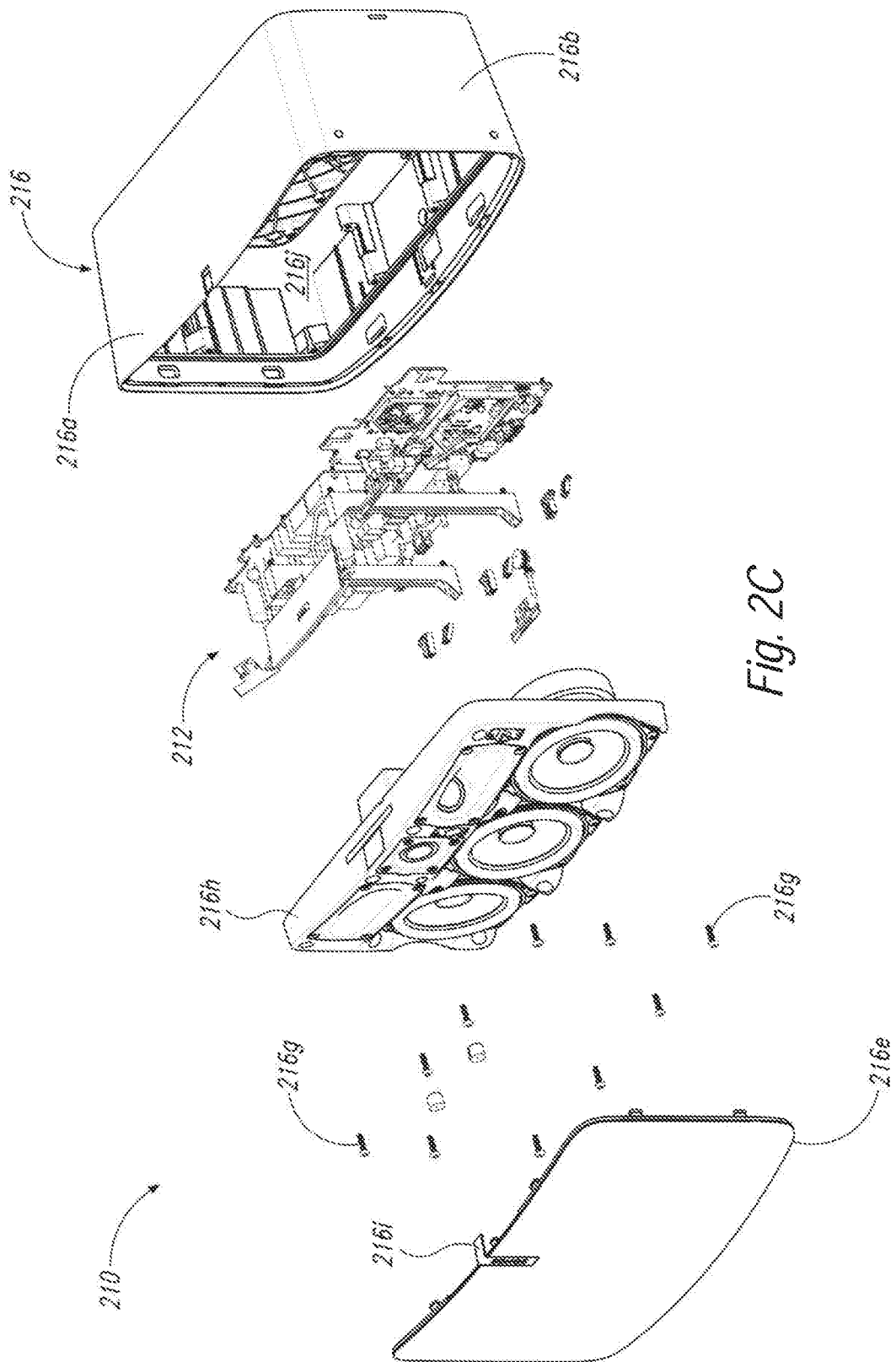
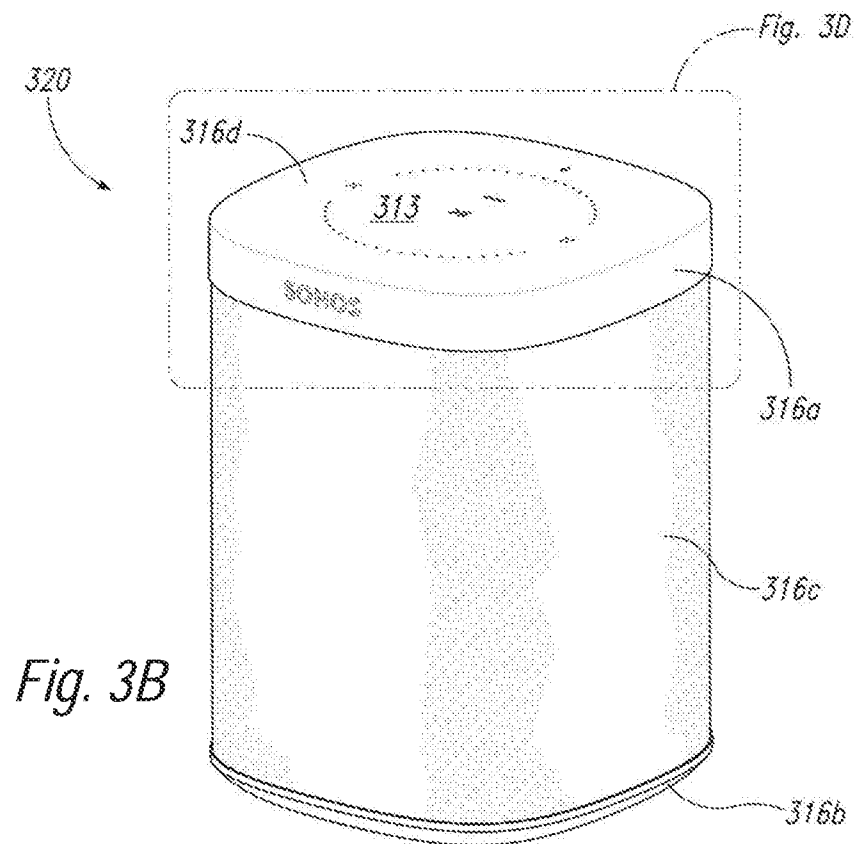
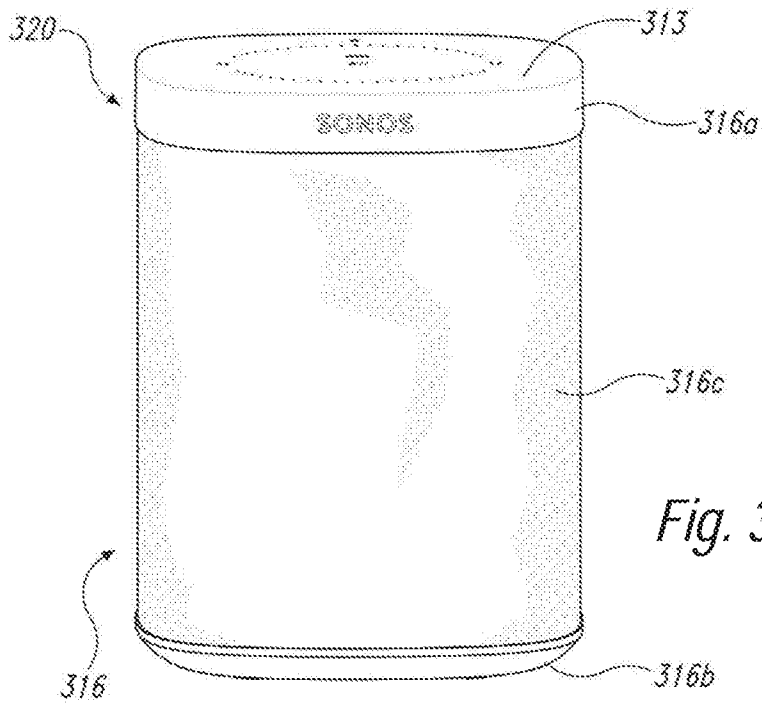


Fig. 2C



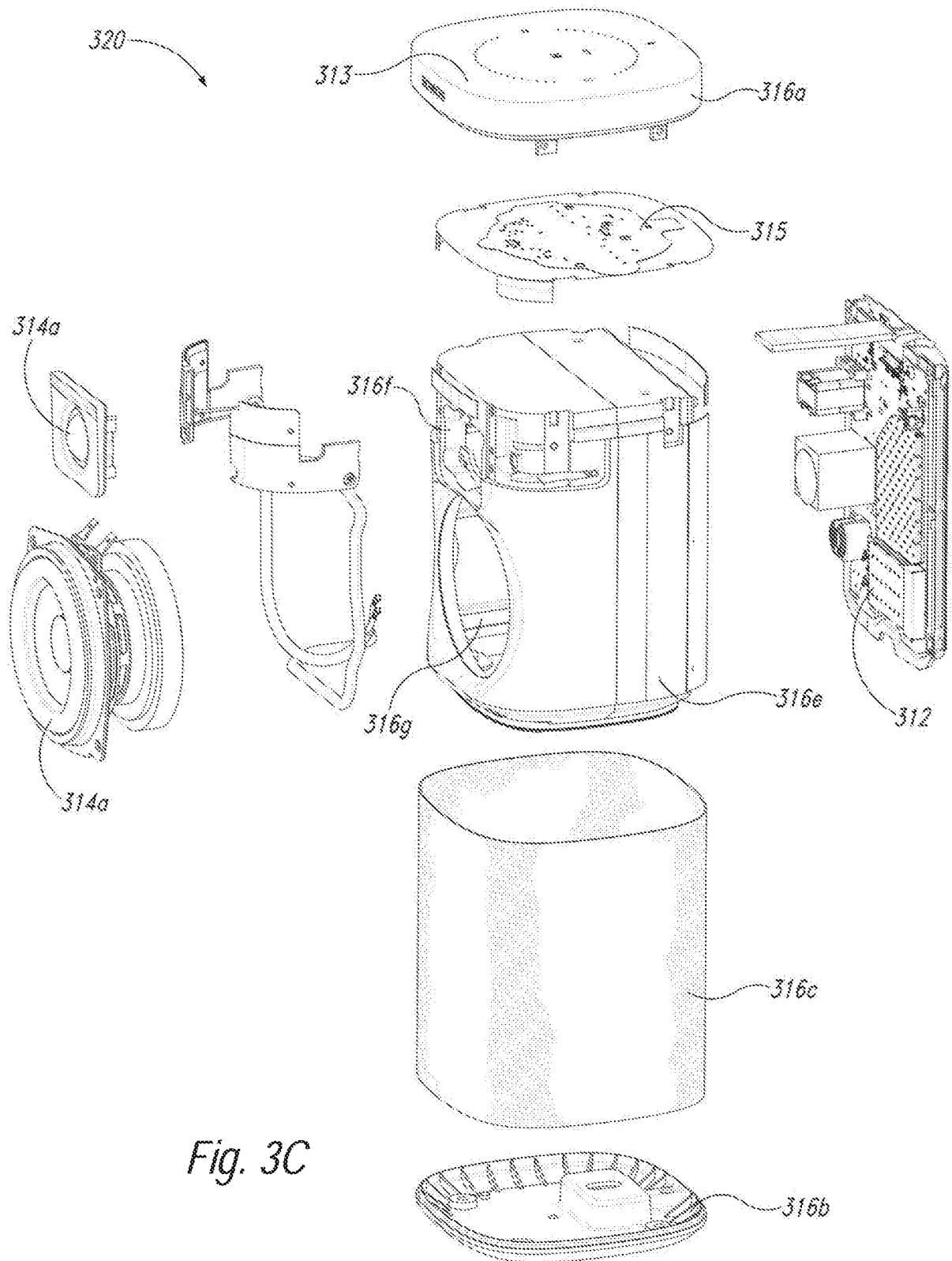


Fig. 3C

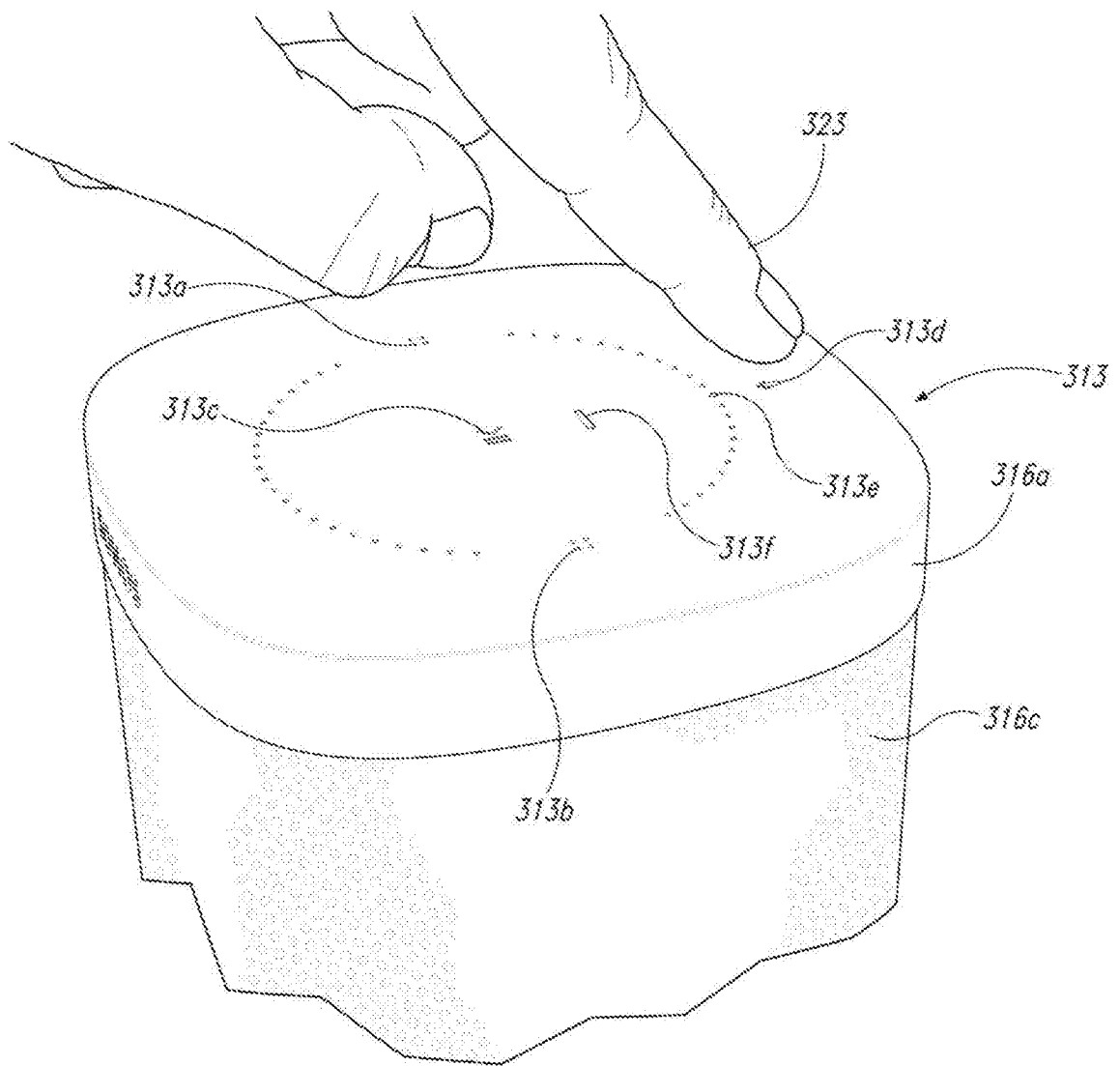


Fig. 3D

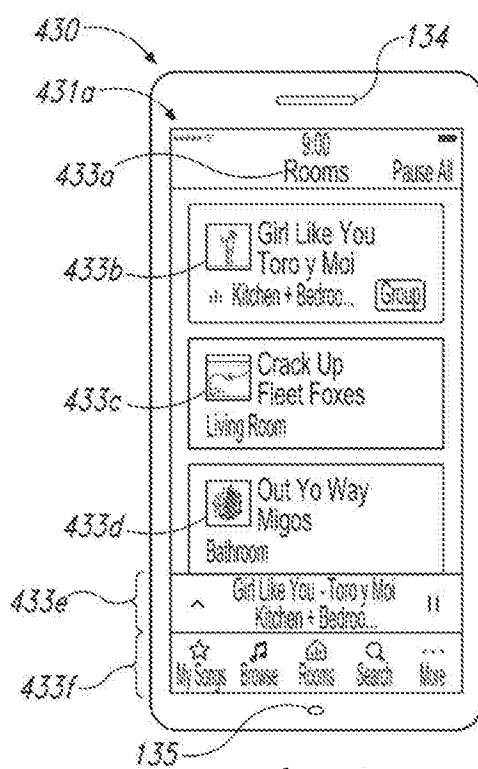


Fig. 4A

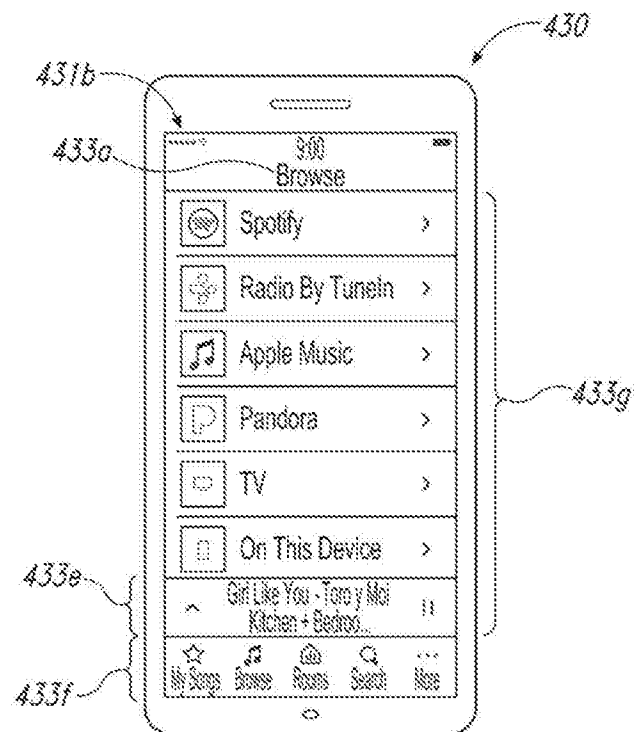


Fig. 4B

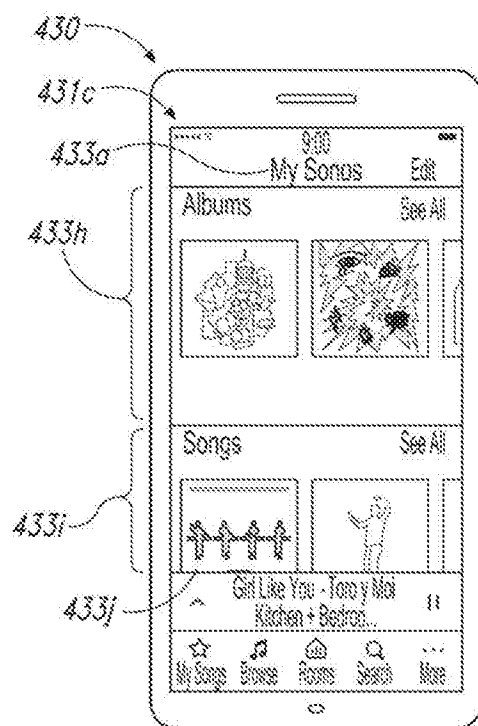


Fig. 4C

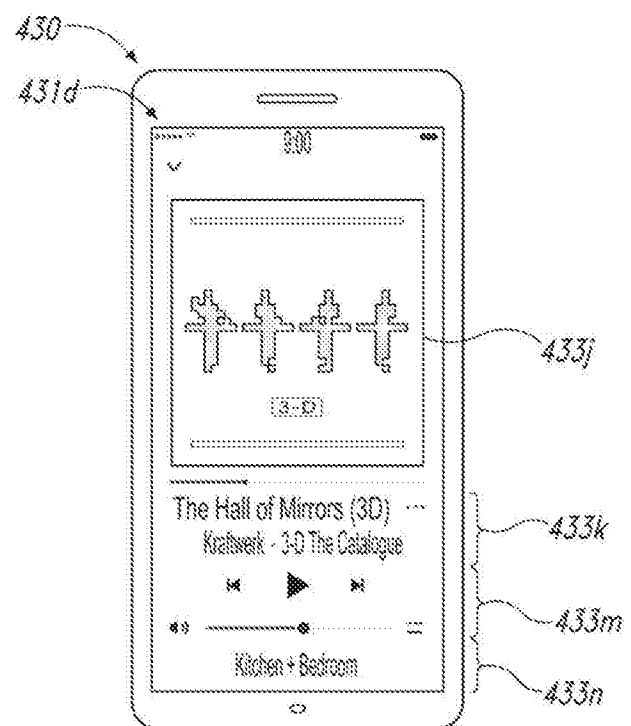


Fig. 4D

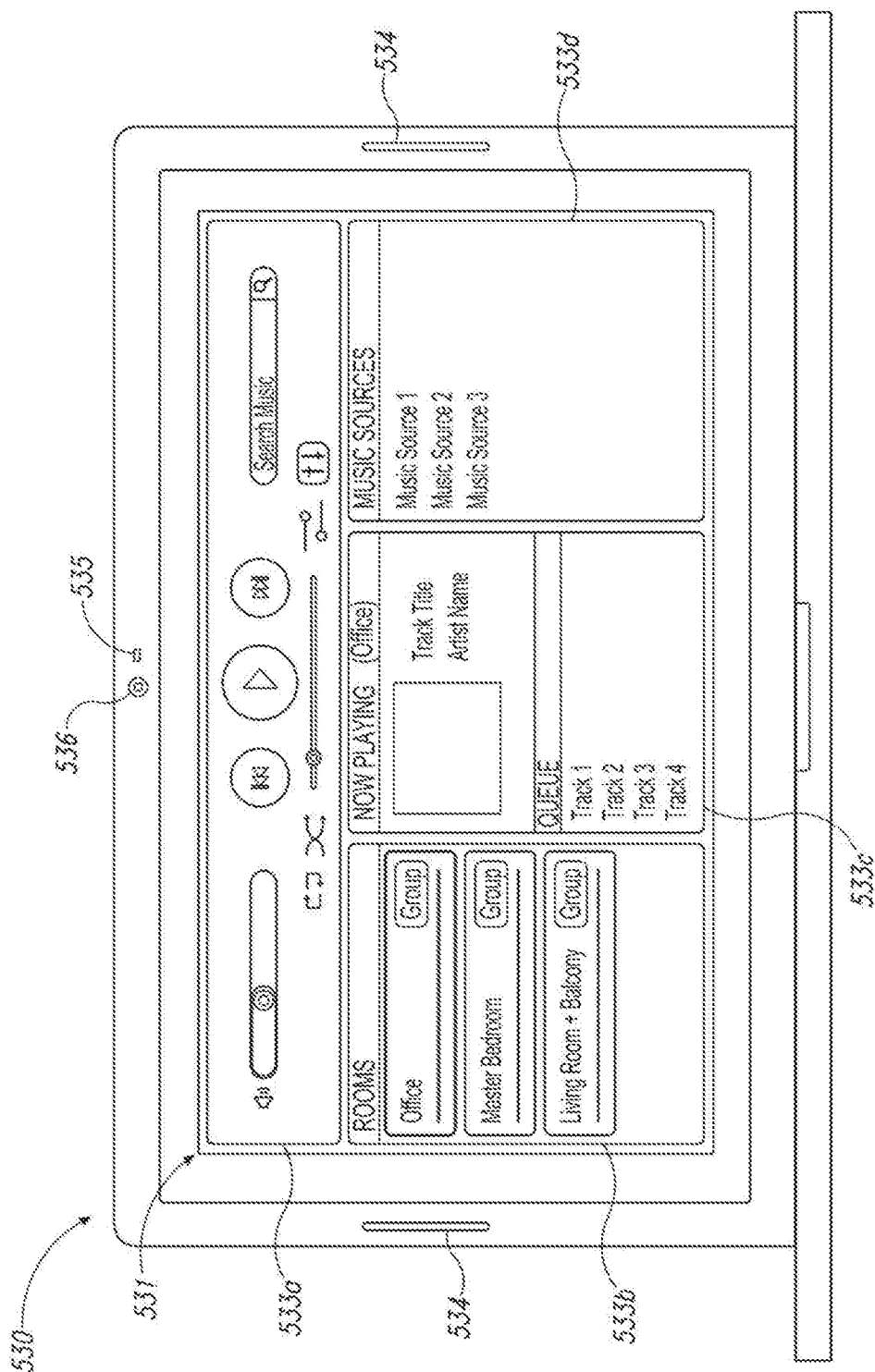


Fig. 5

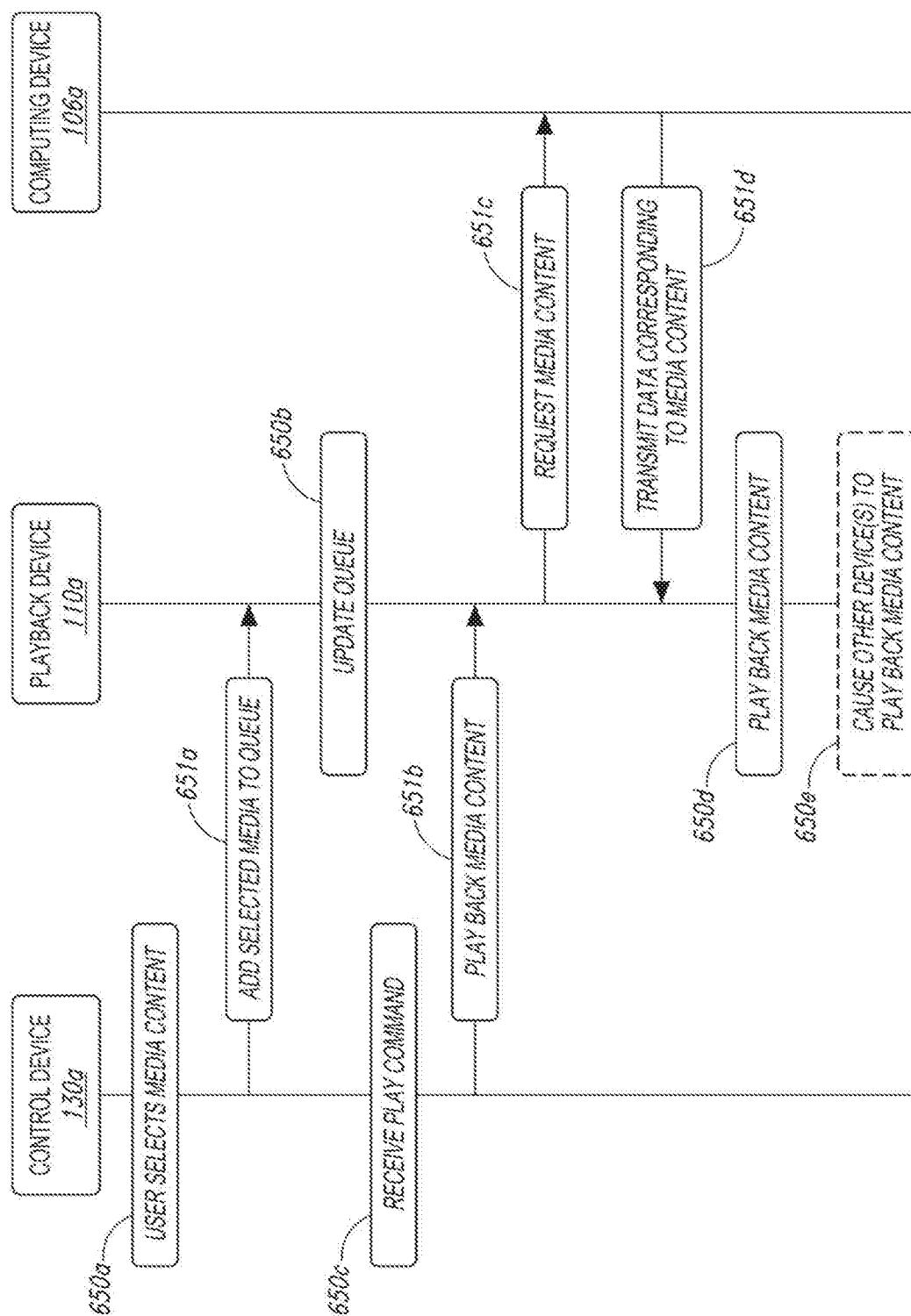


Fig. 6

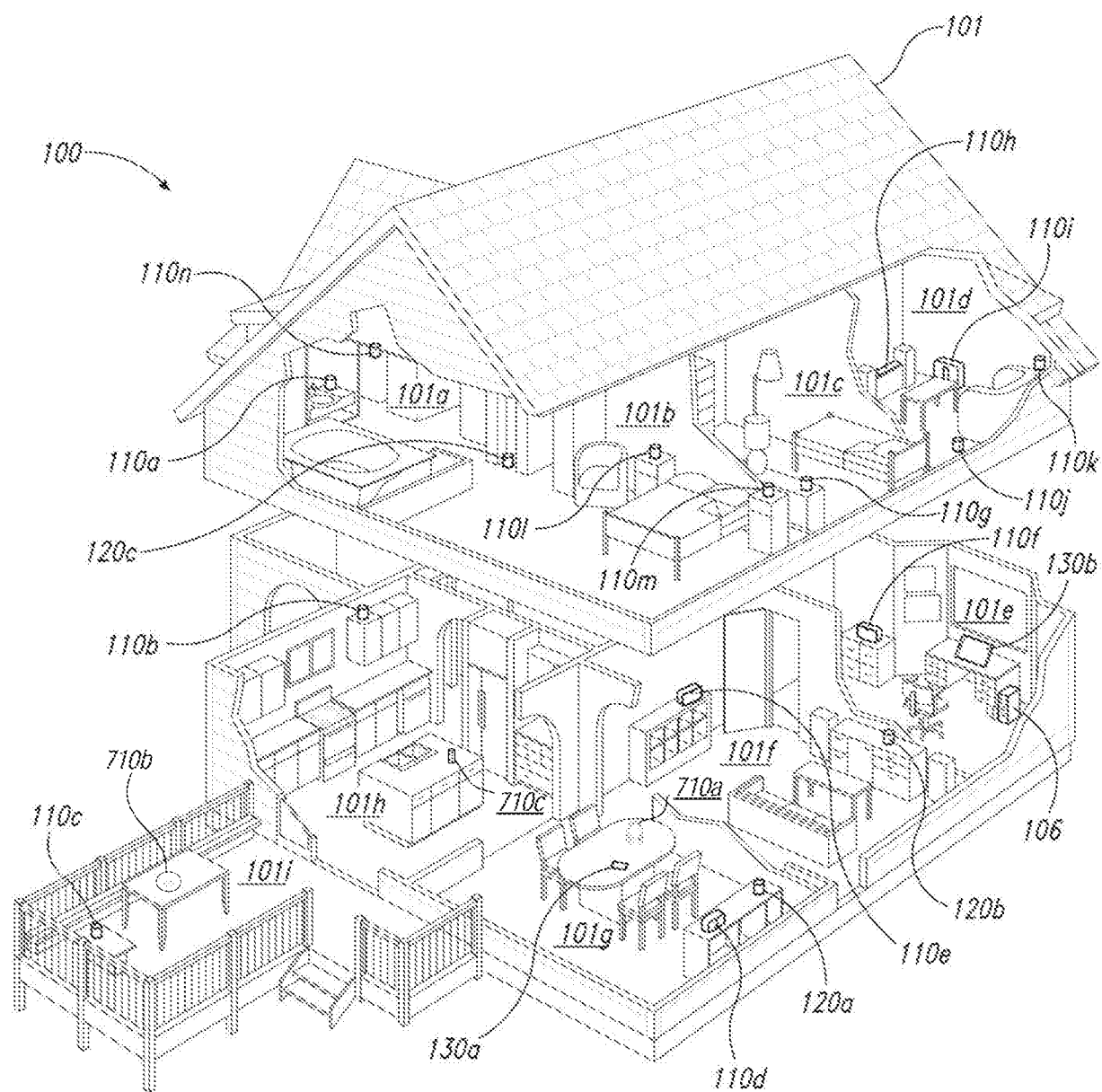


Fig. 7A

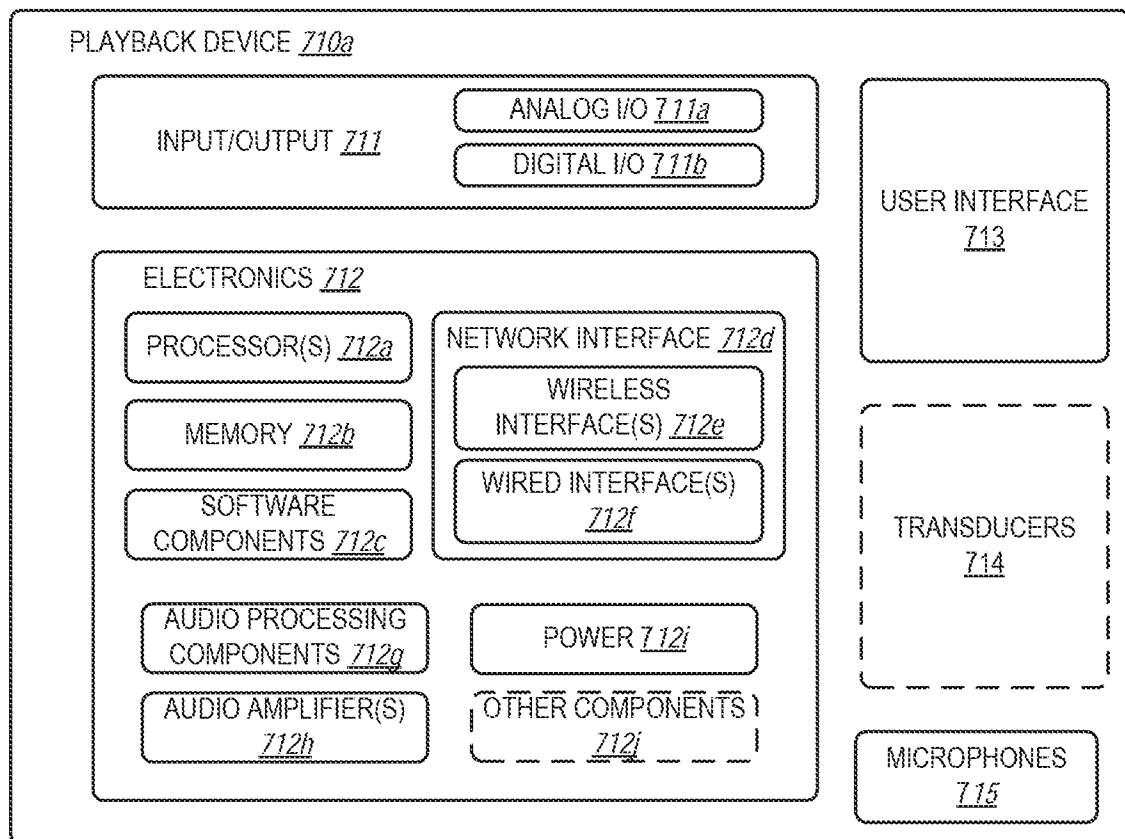


Fig. 7B

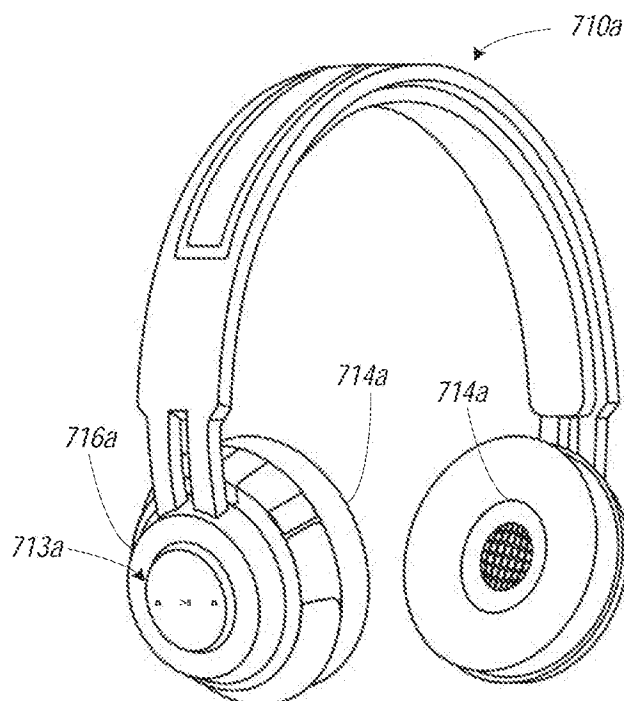


Fig. 7C

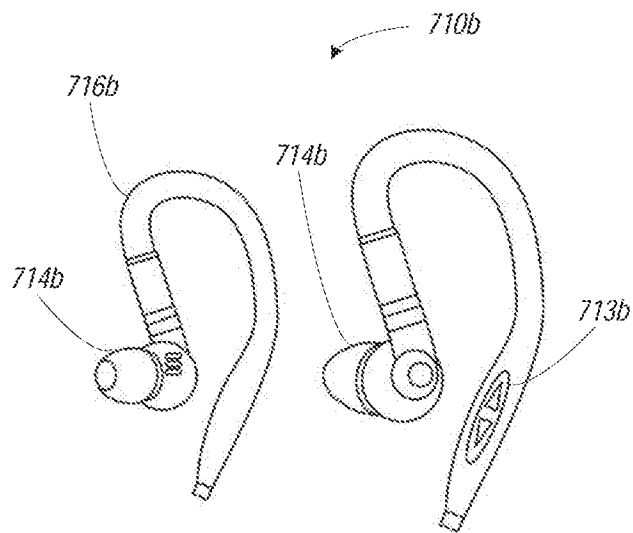


Fig. 7D

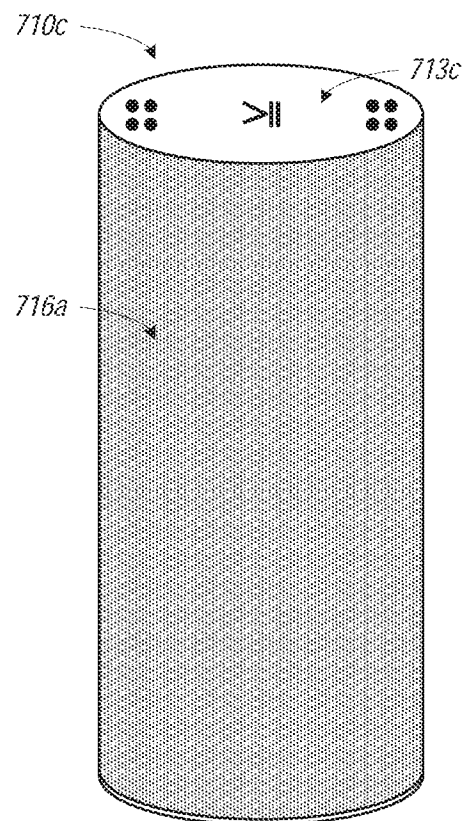


Fig. 7E

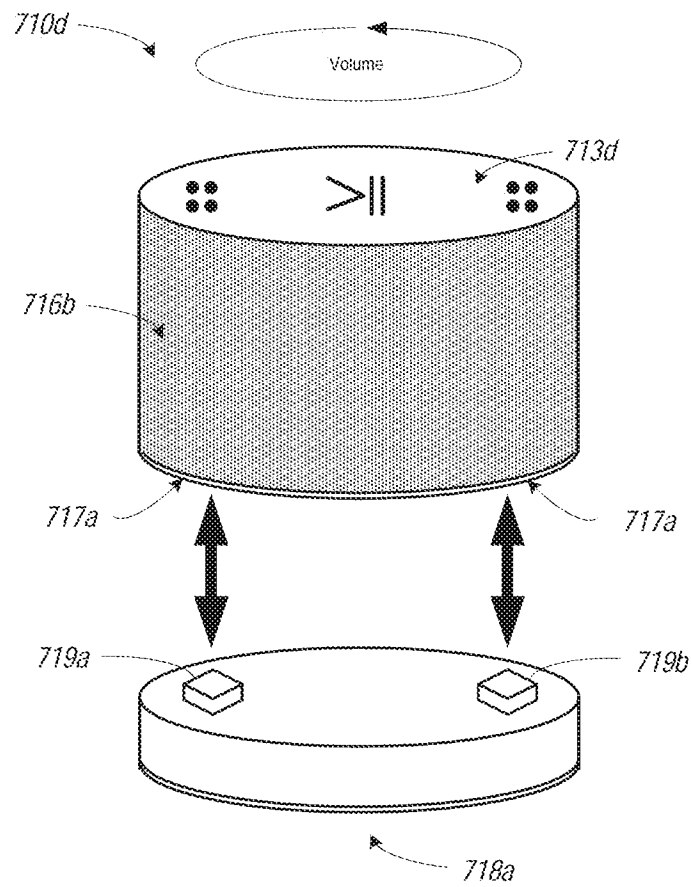


Fig. 7F

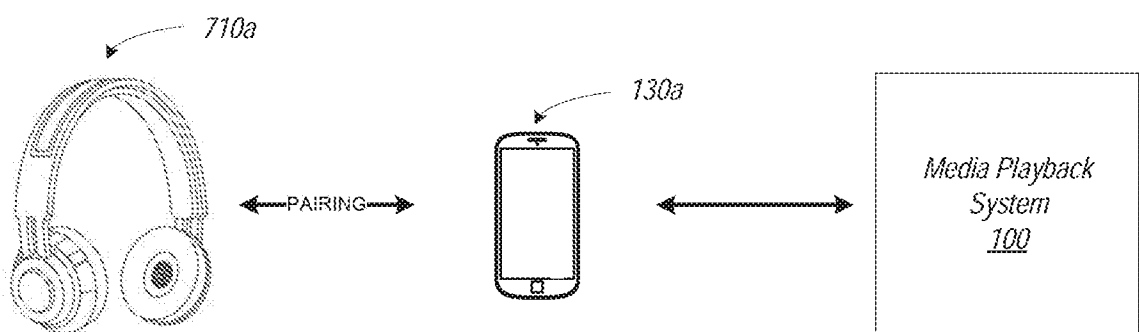


Fig. 7G

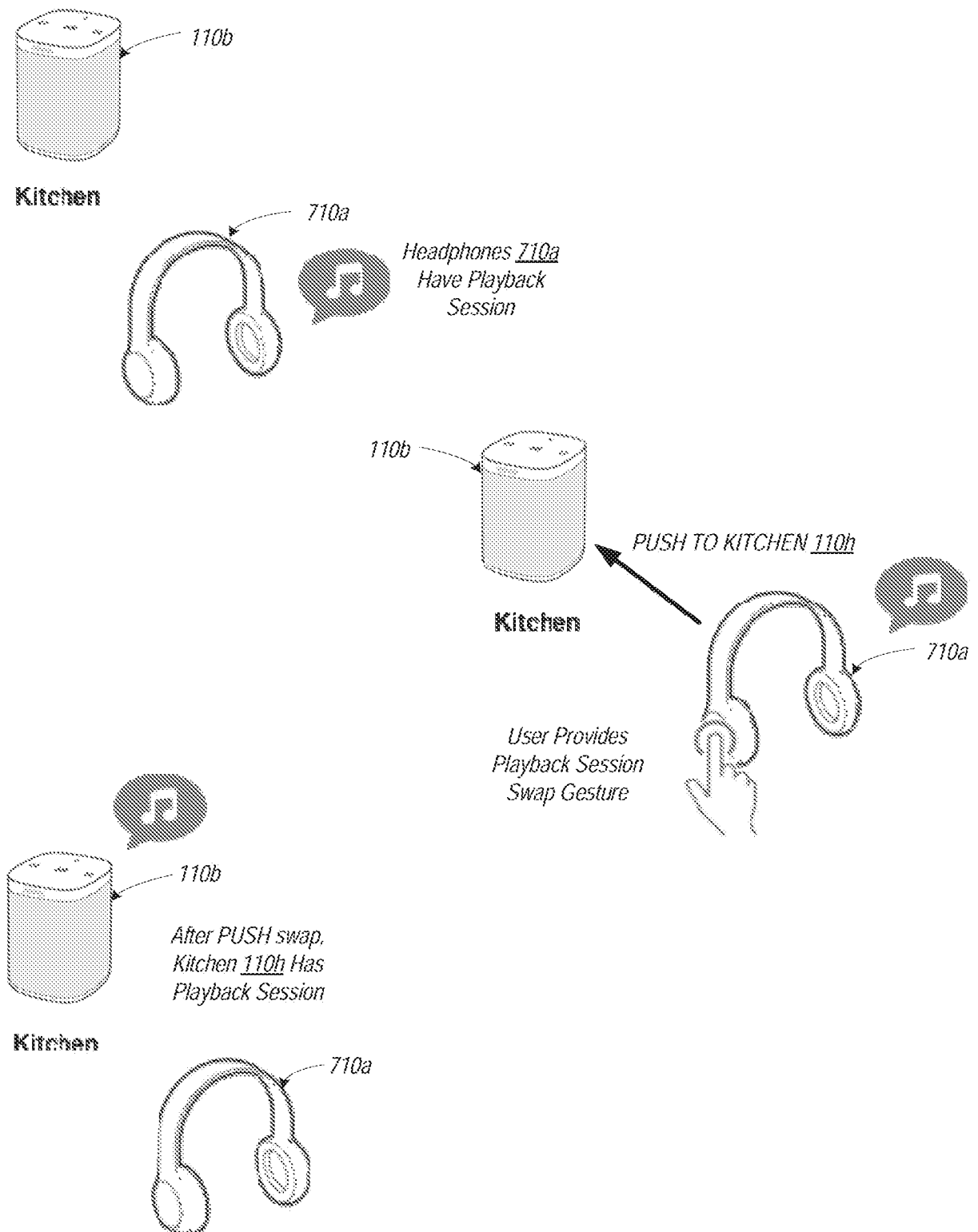


Fig. 8A

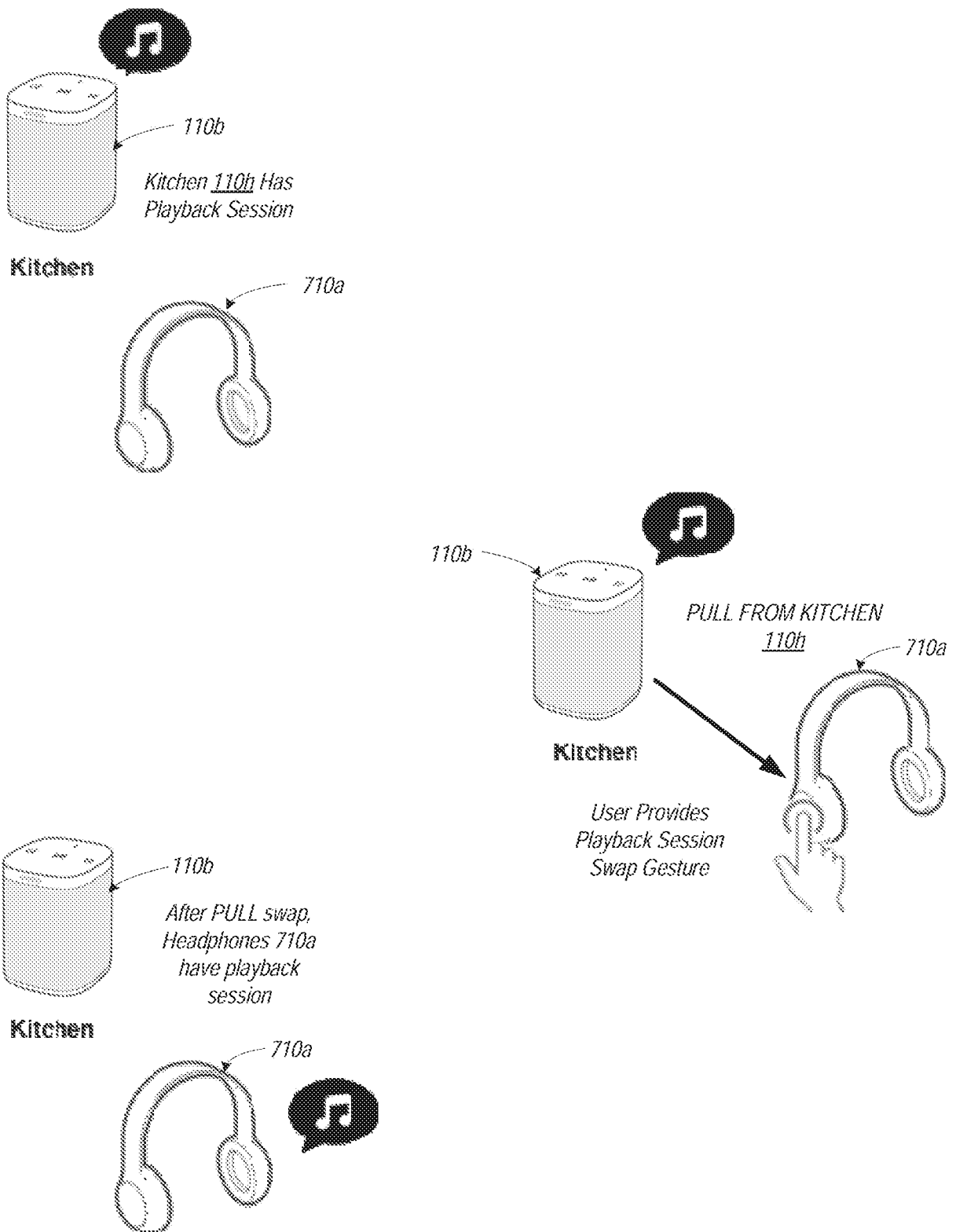


Fig. 8B

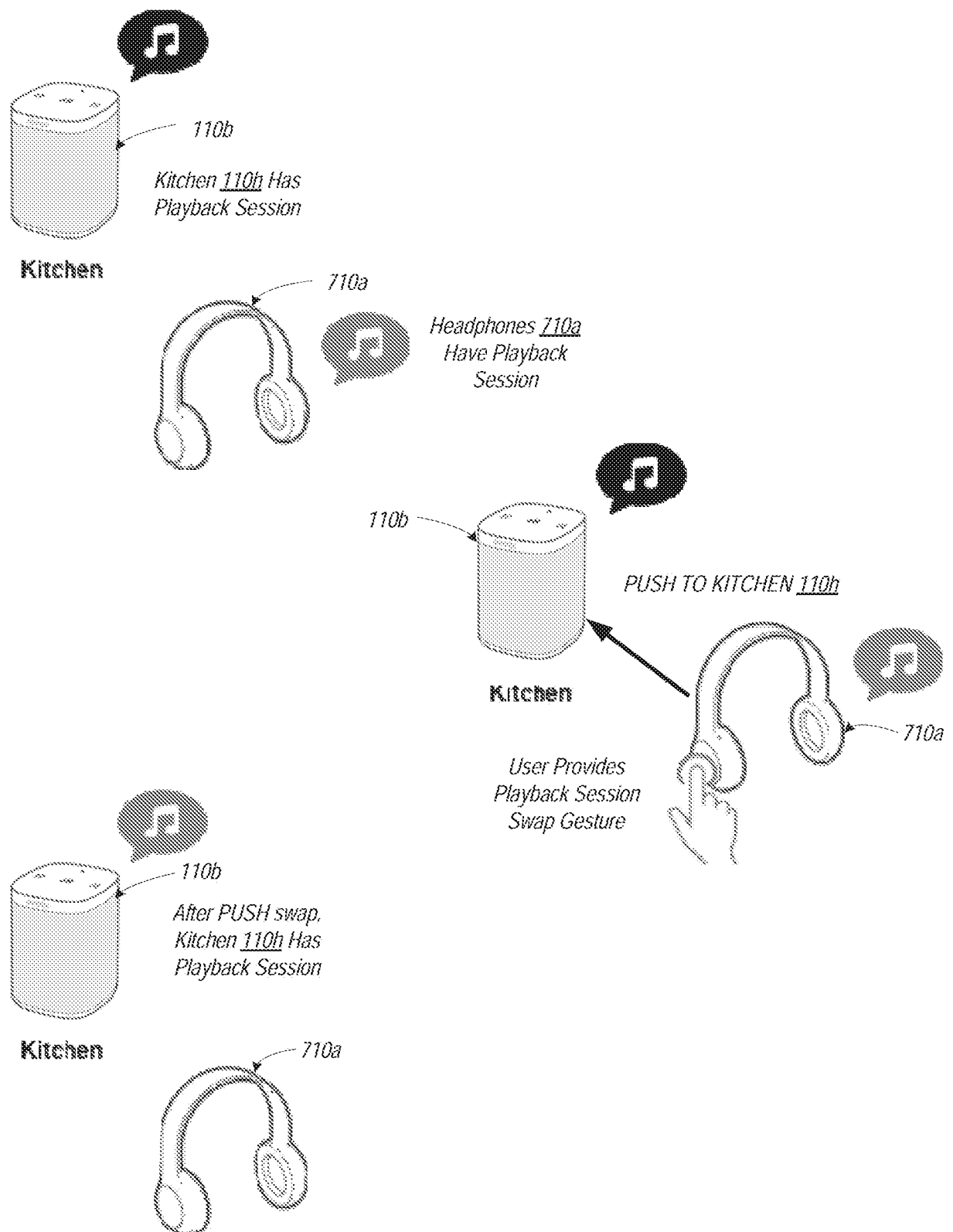
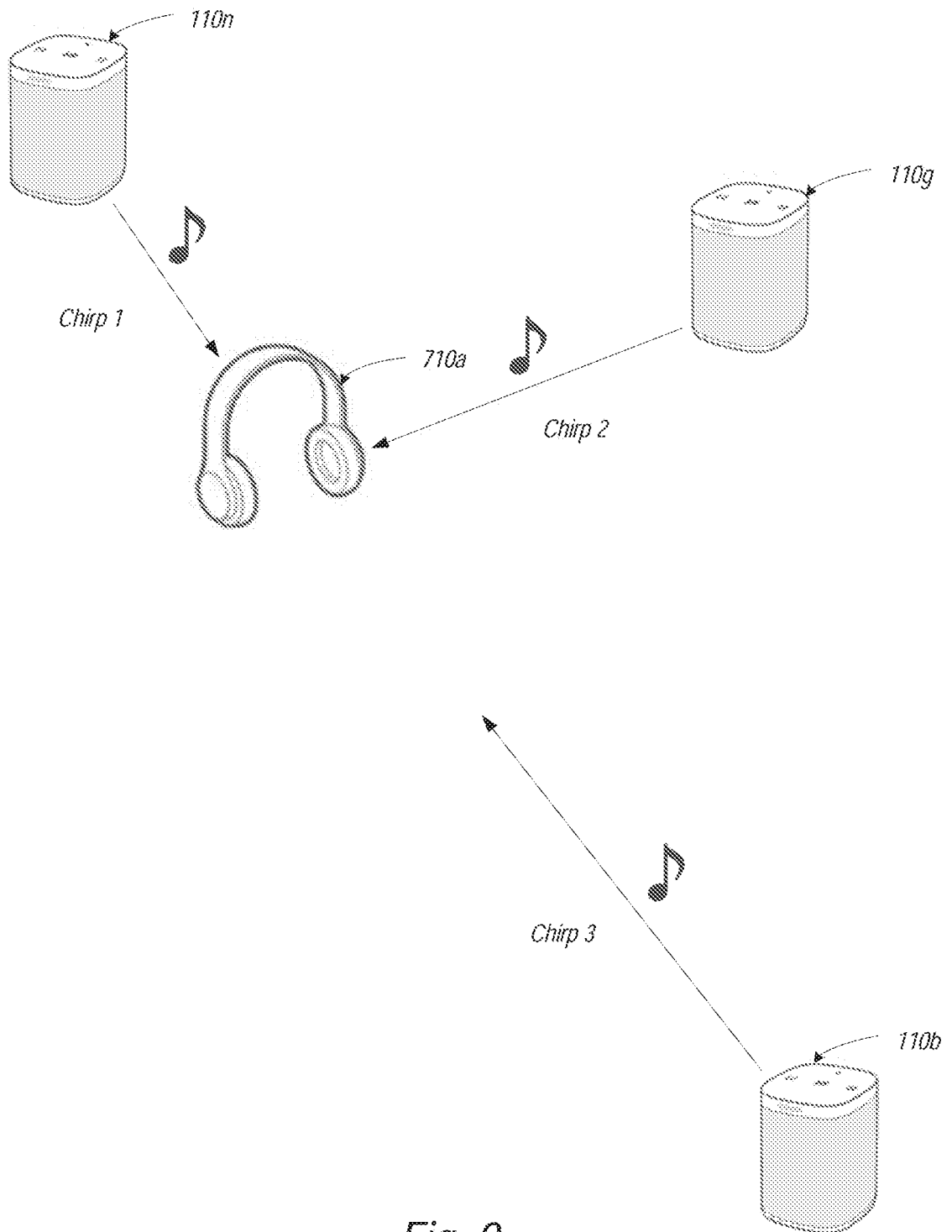


Fig. 8C



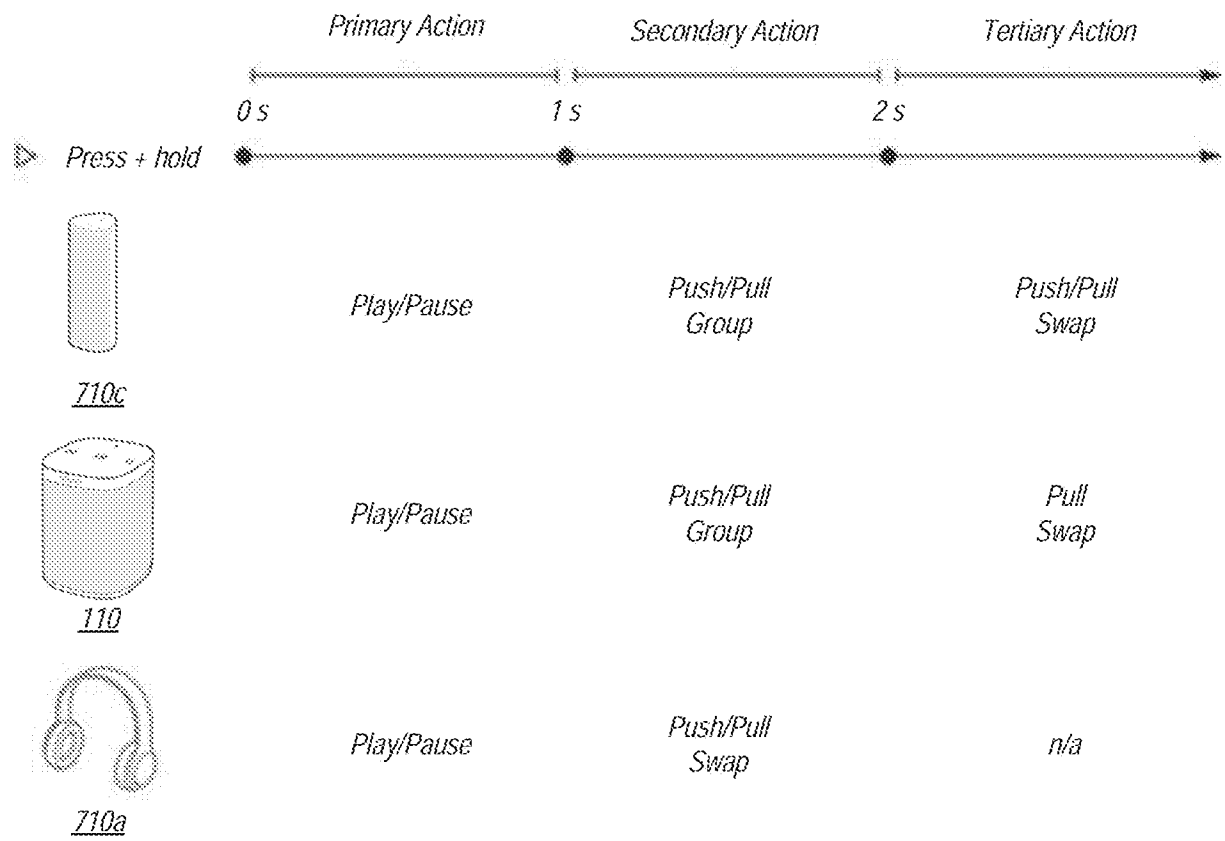


Fig. 10

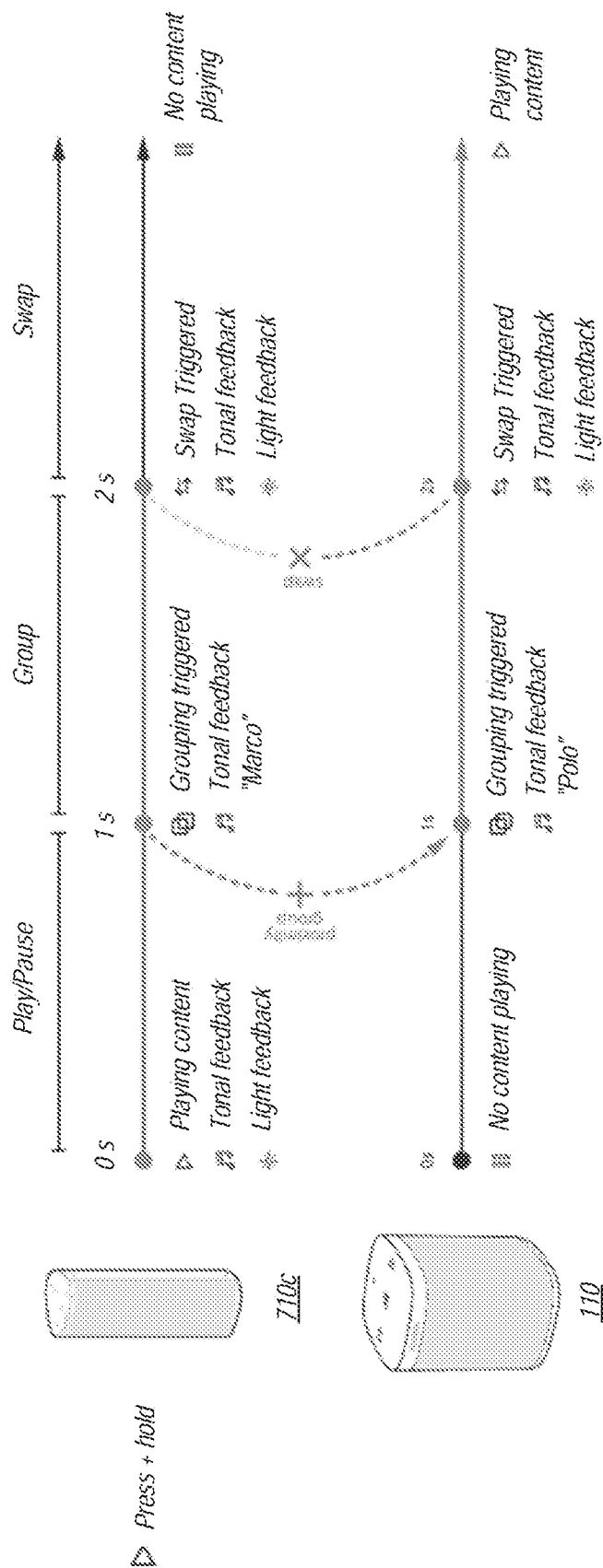


Fig. 11

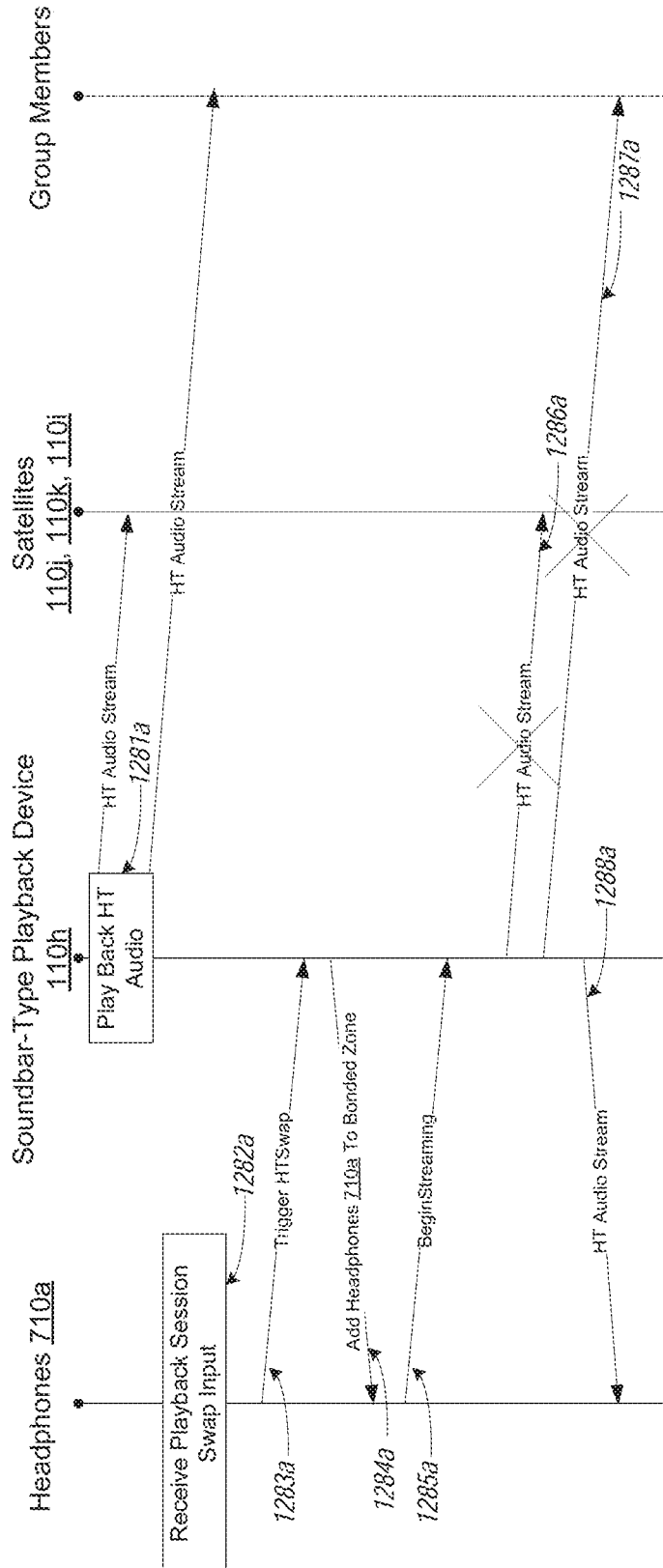


Fig. 12A

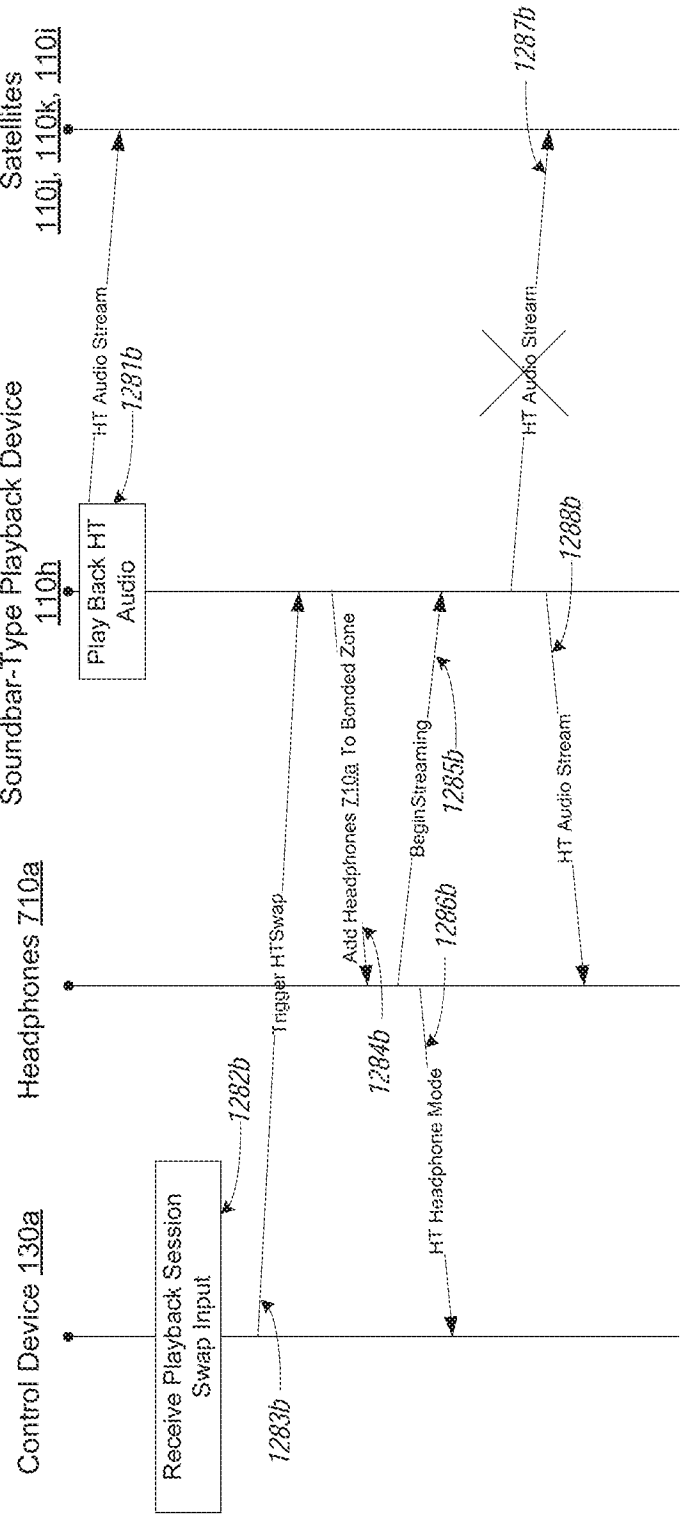
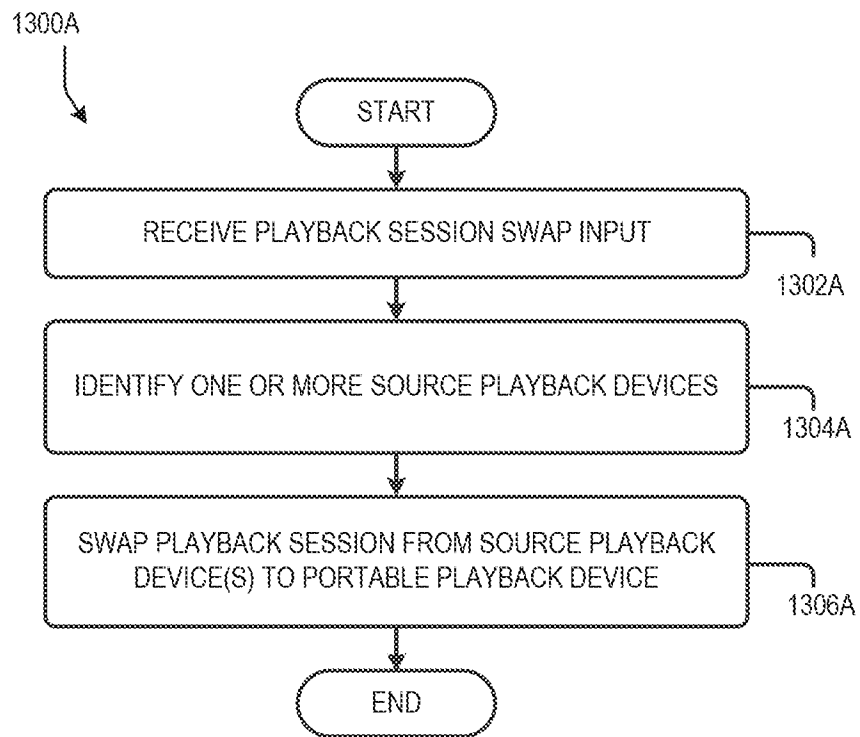
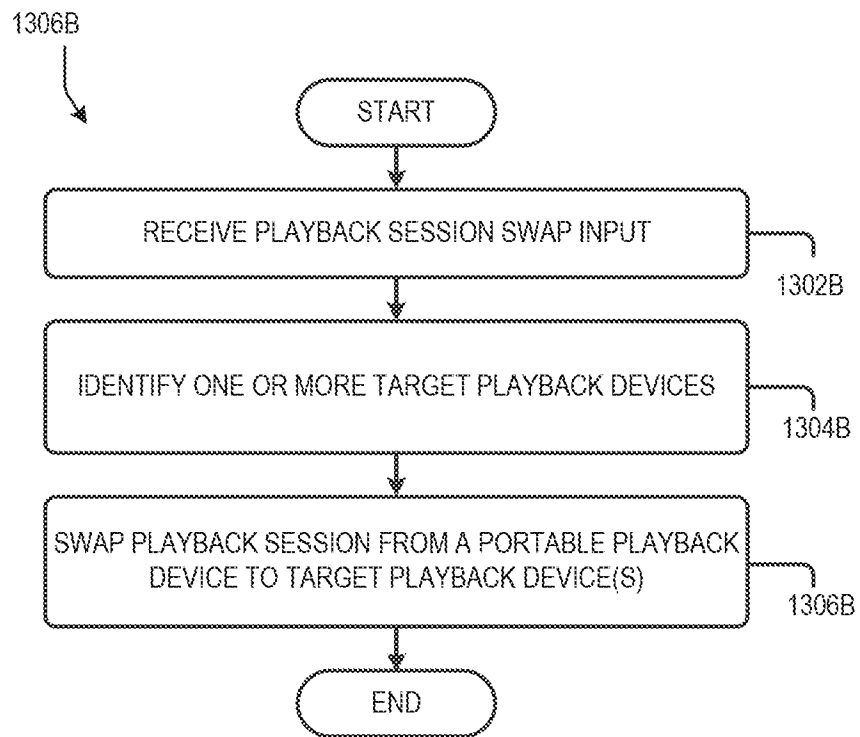
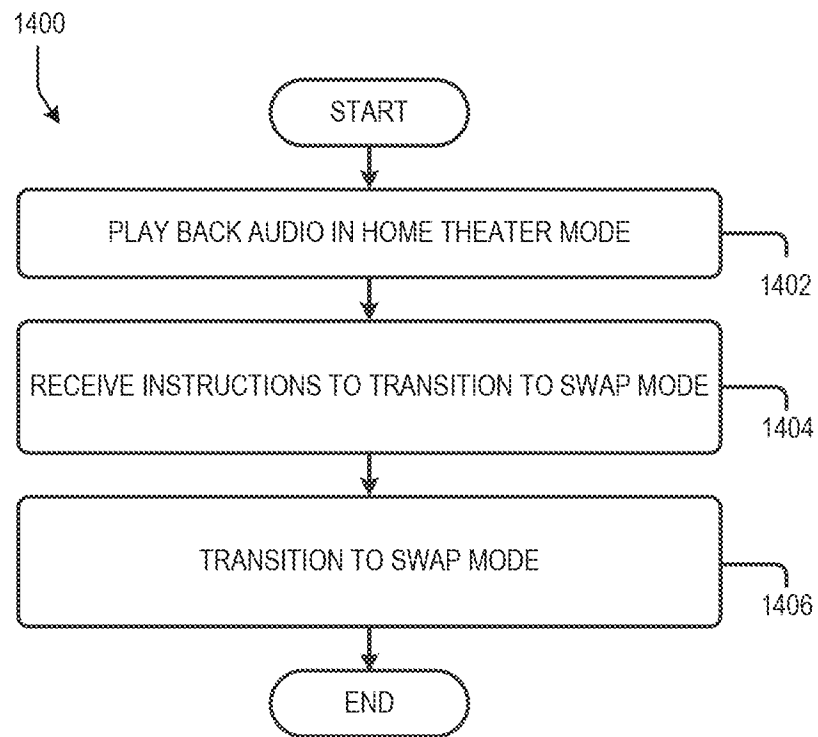
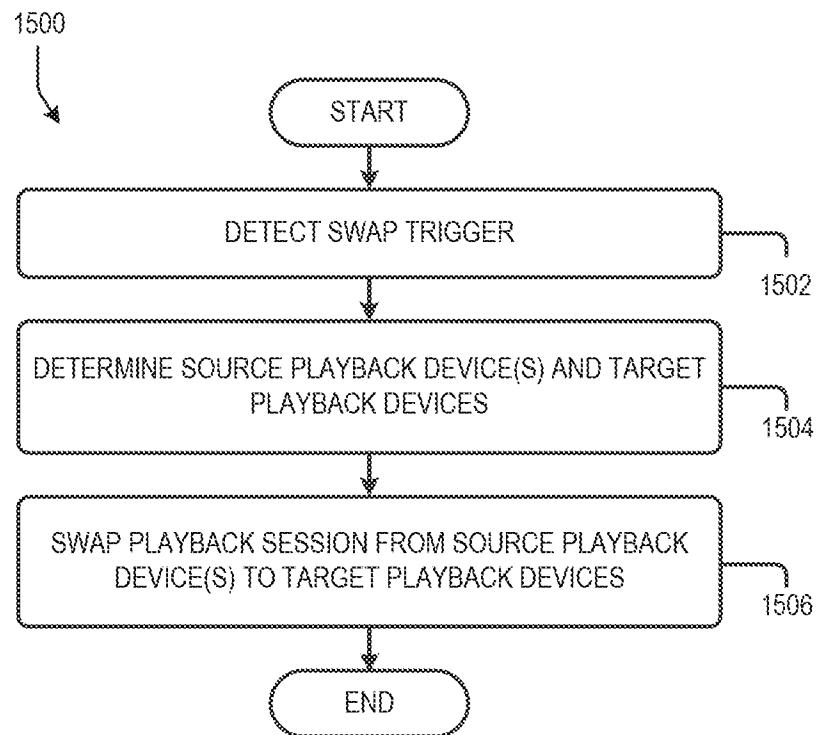


Fig. 12B

*Fig. 13A*

*Fig. 13B*

*Fig. 14*

*Fig. 15*

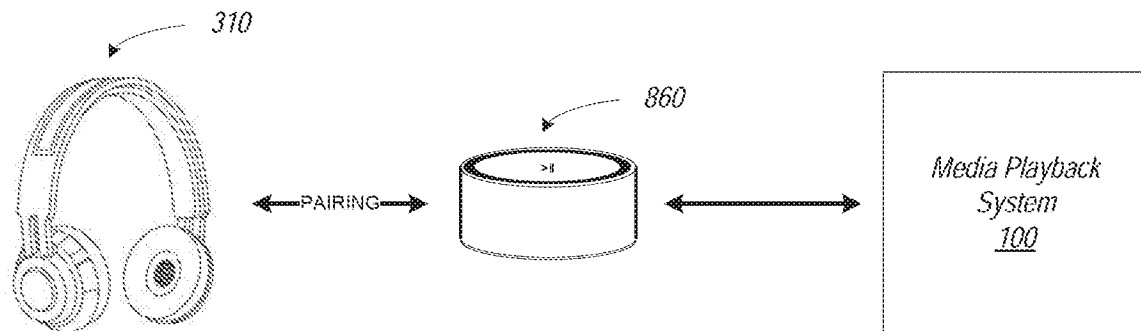


Fig. 16A

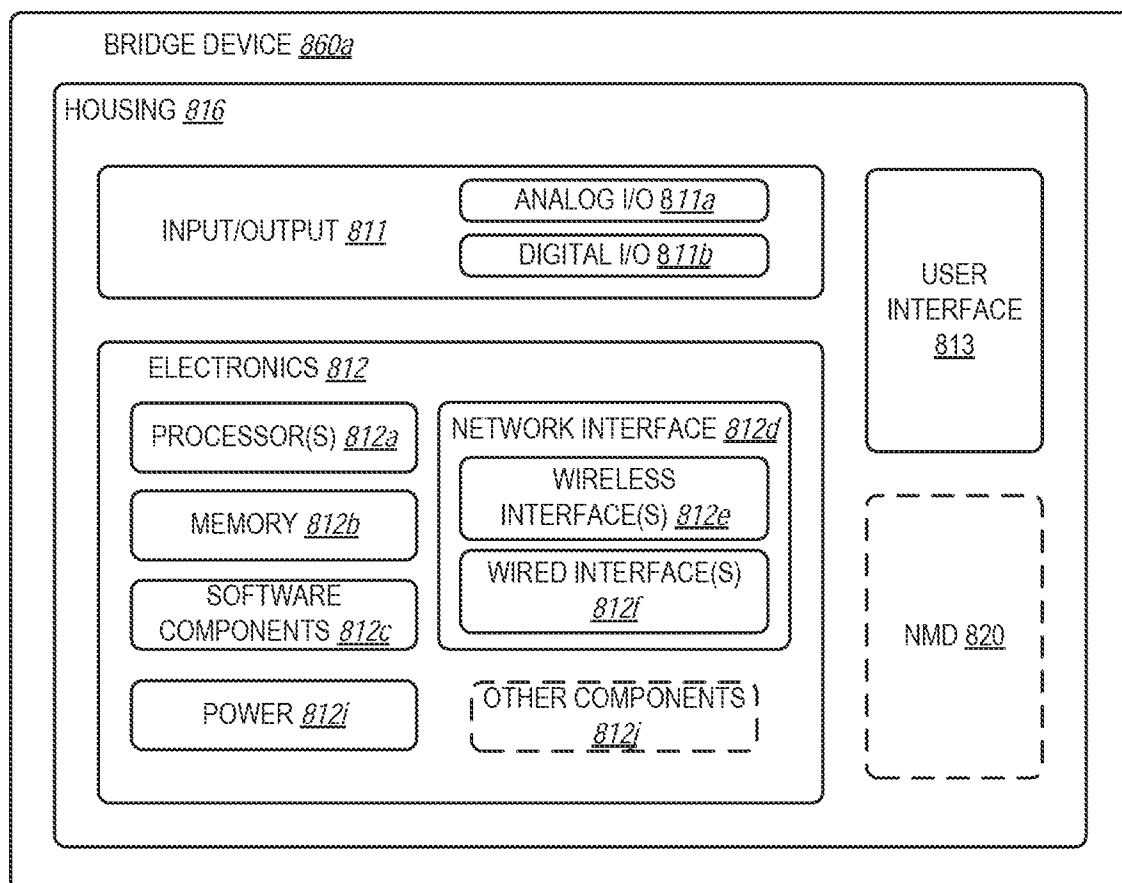


Fig. 16B

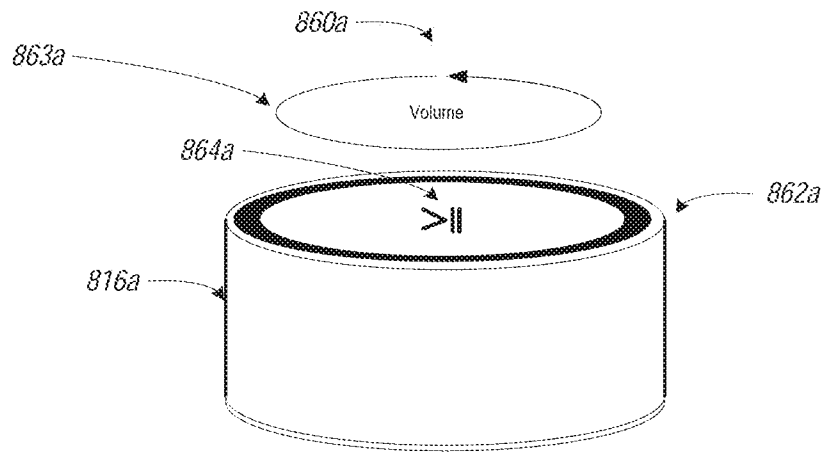


Fig. 16C

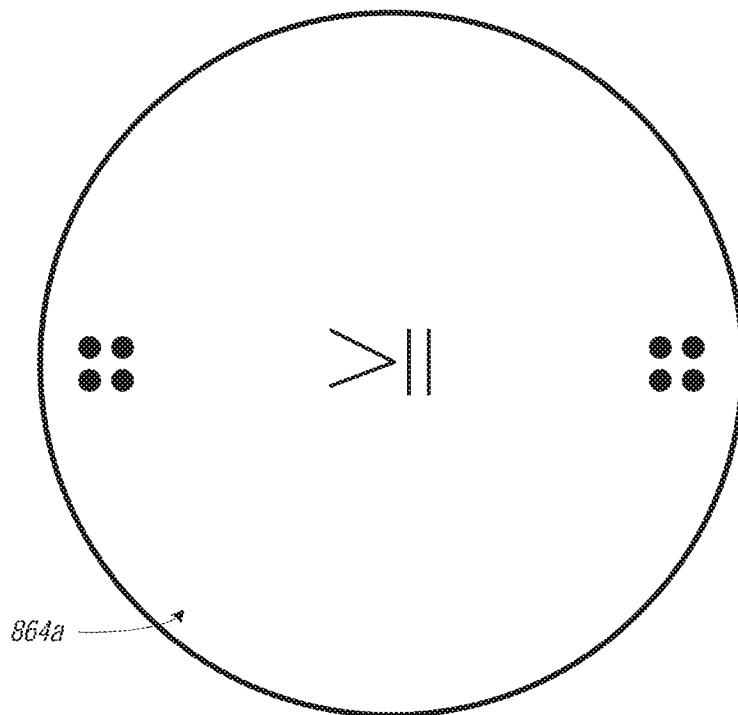


Fig. 16D

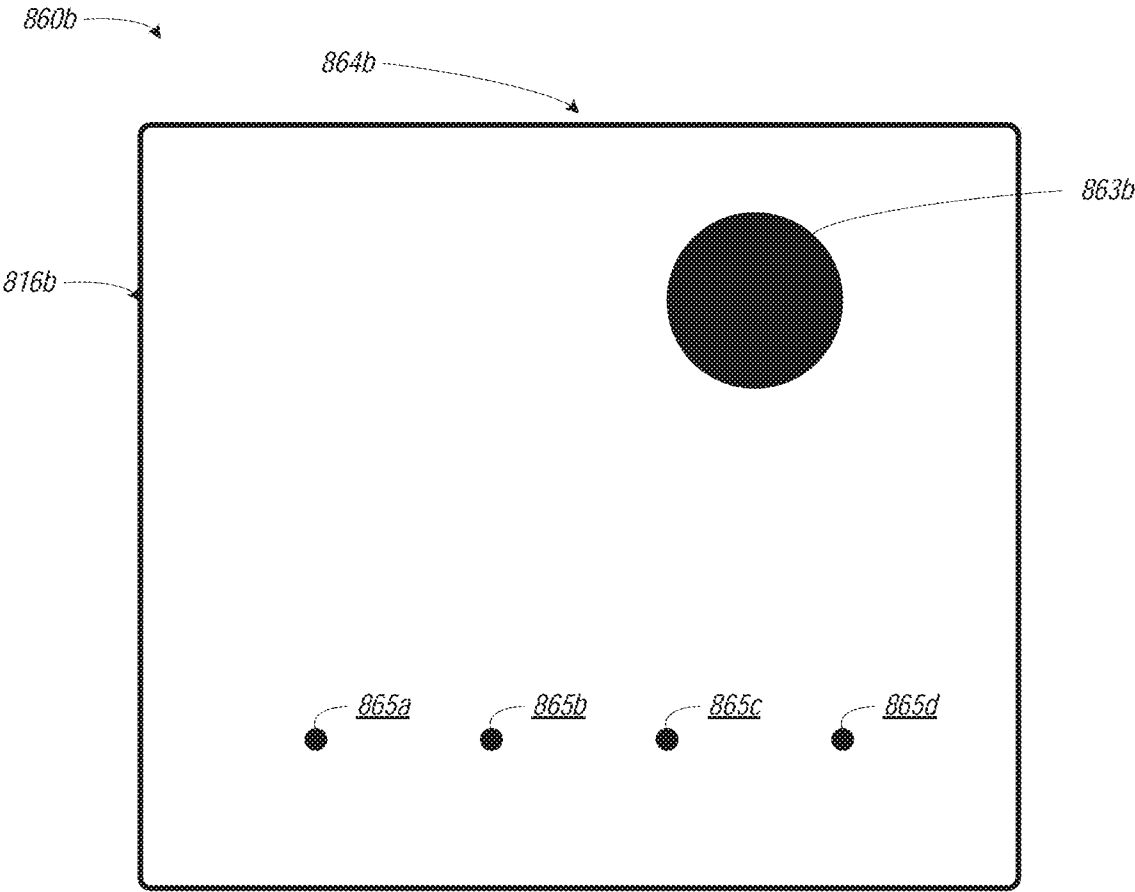


Fig. 16E

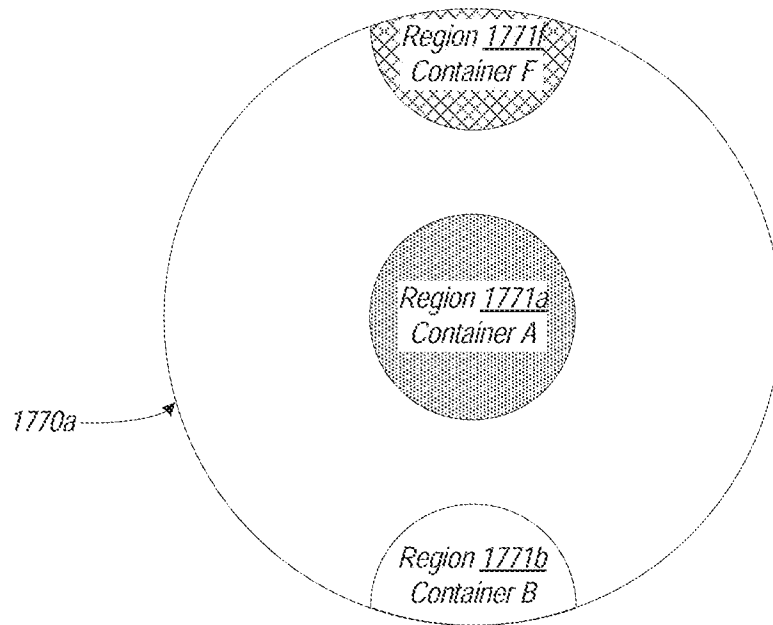


Fig. 17A

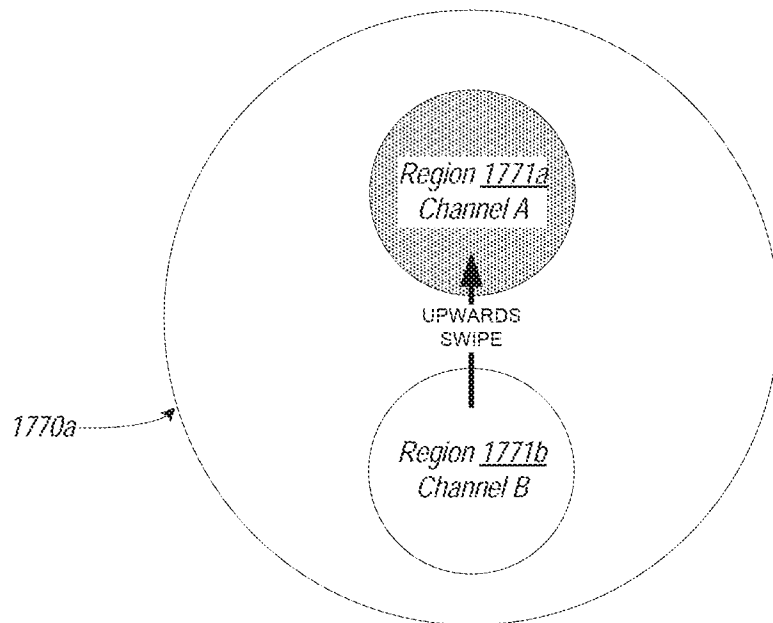


Fig. 17B

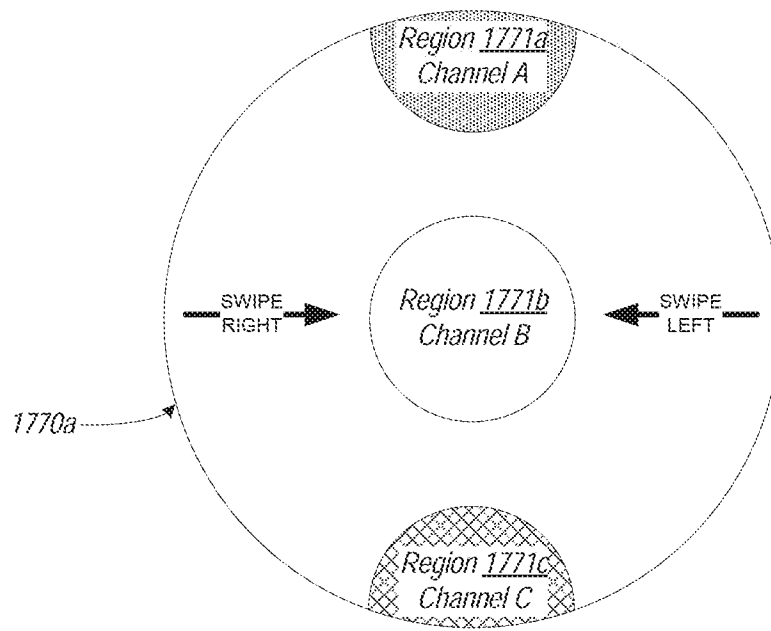


Fig. 17C

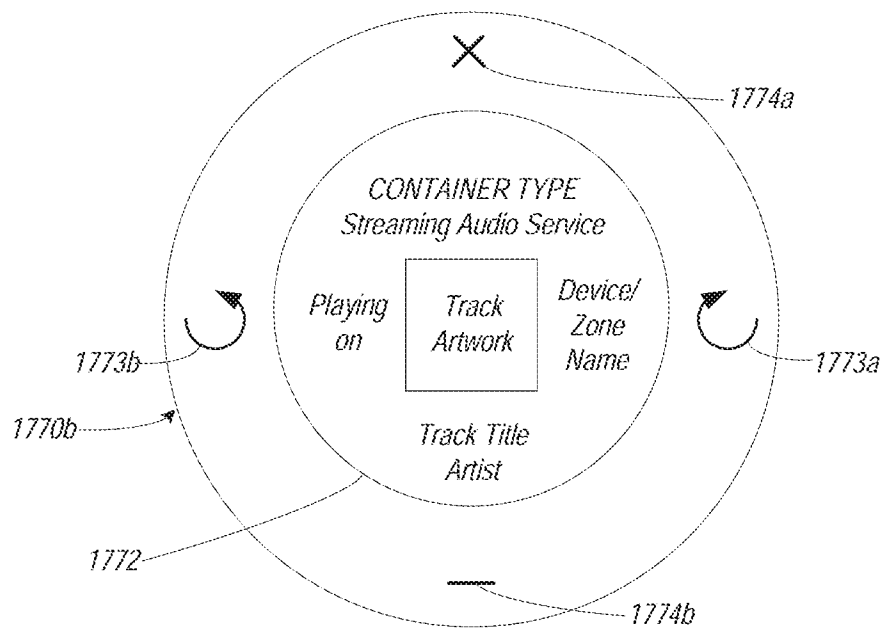


Fig. 17D

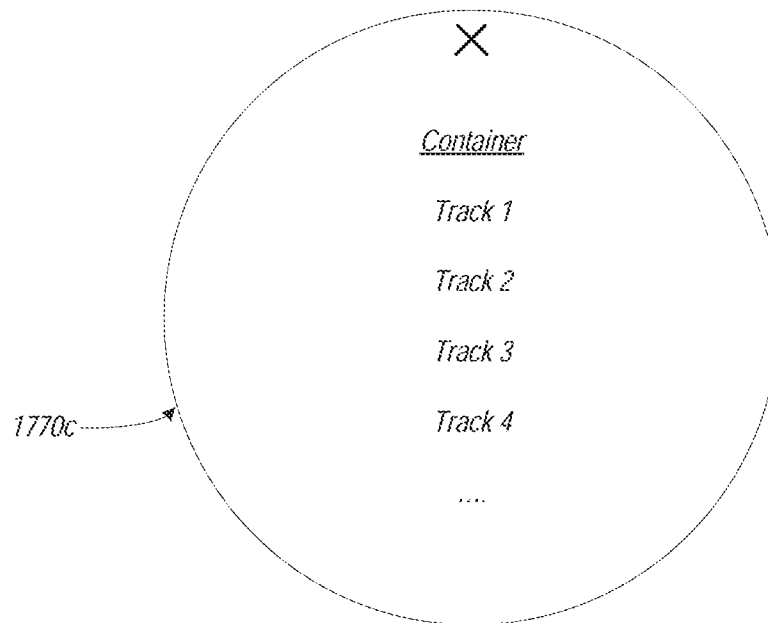


Fig. 17E

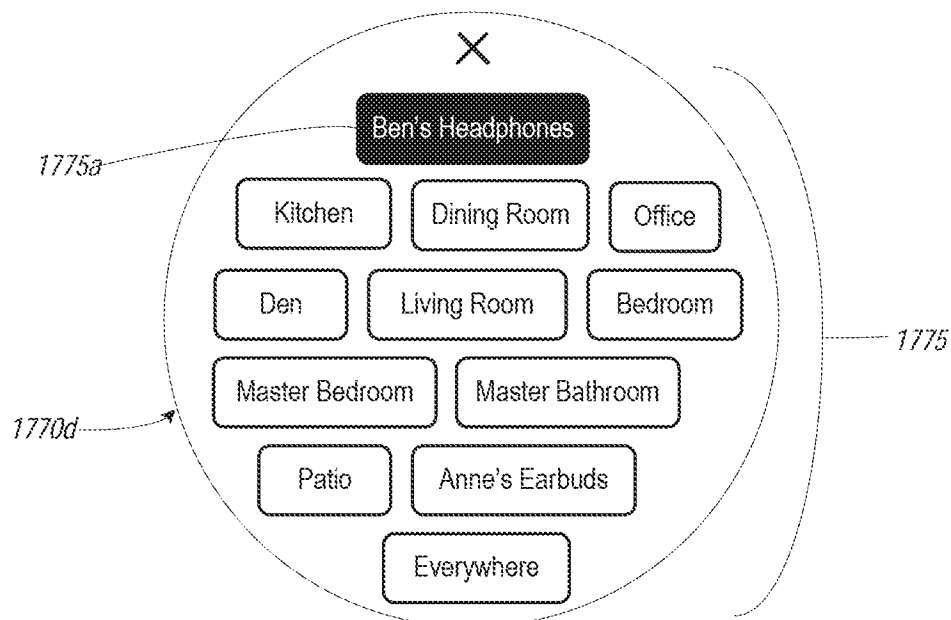


Fig. 17F

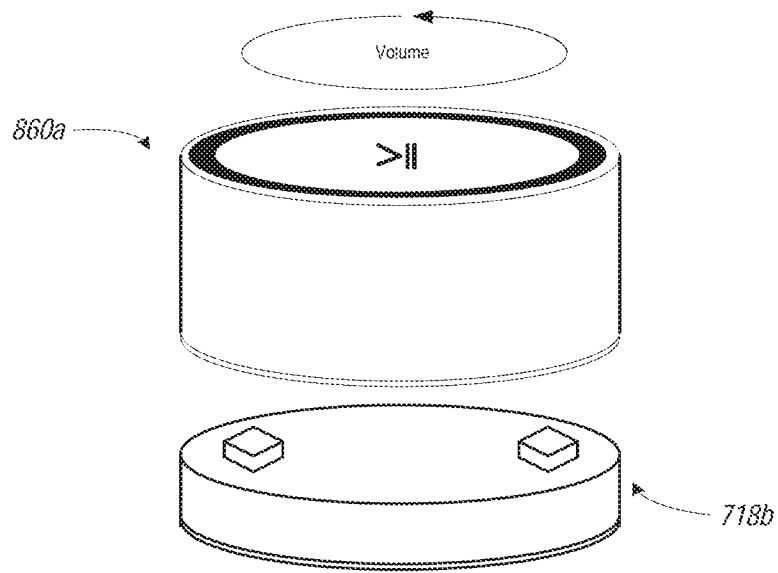


Fig. 18A

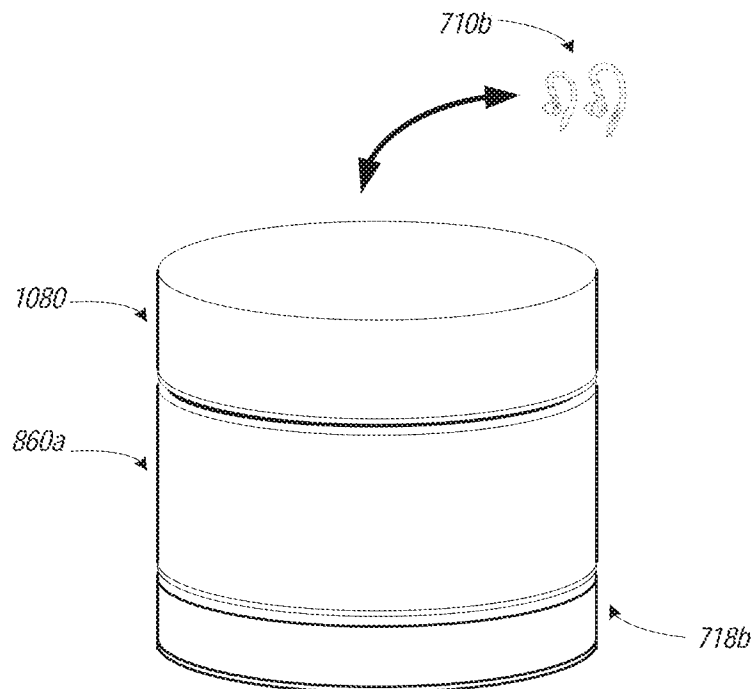


Fig. 18B

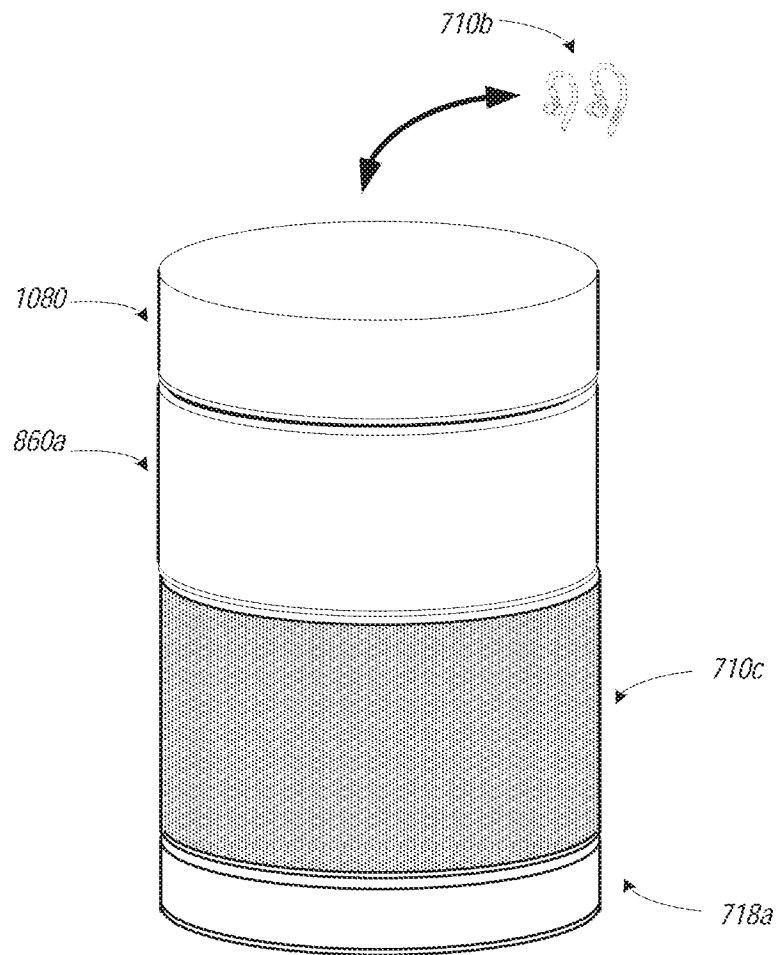


Fig. 18C

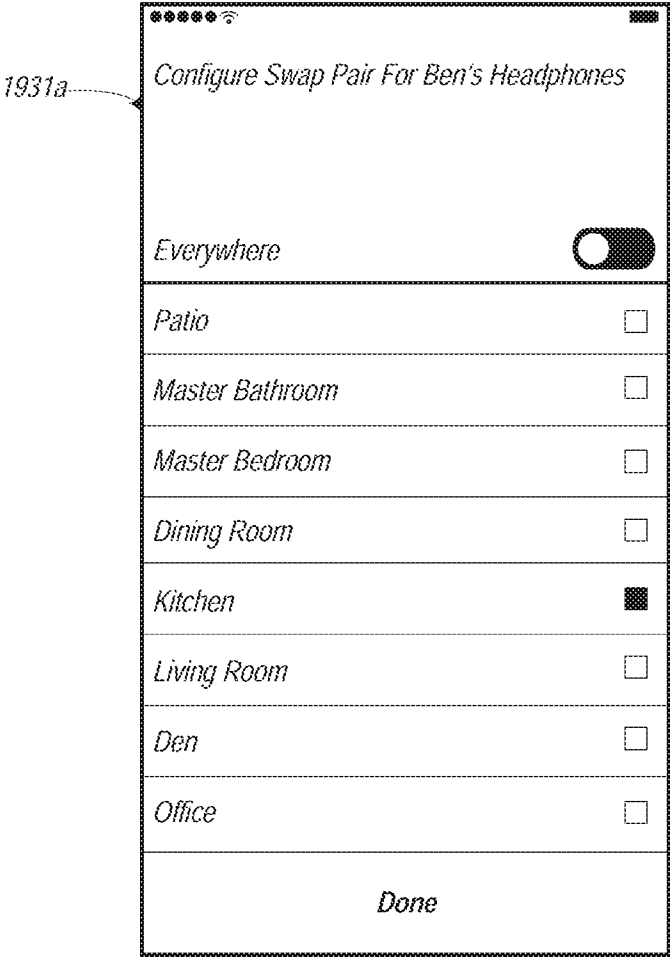
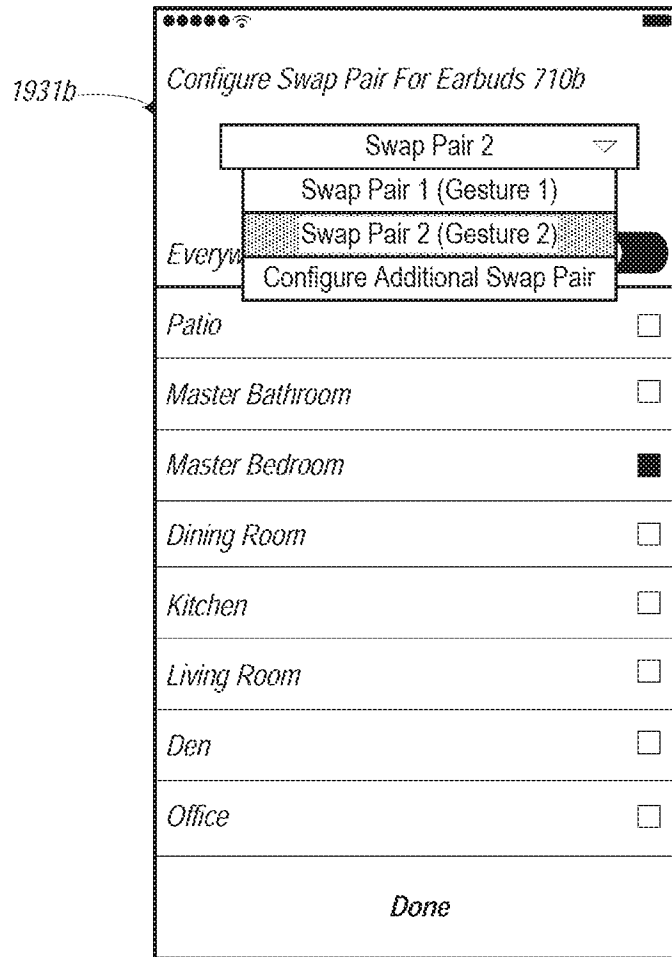
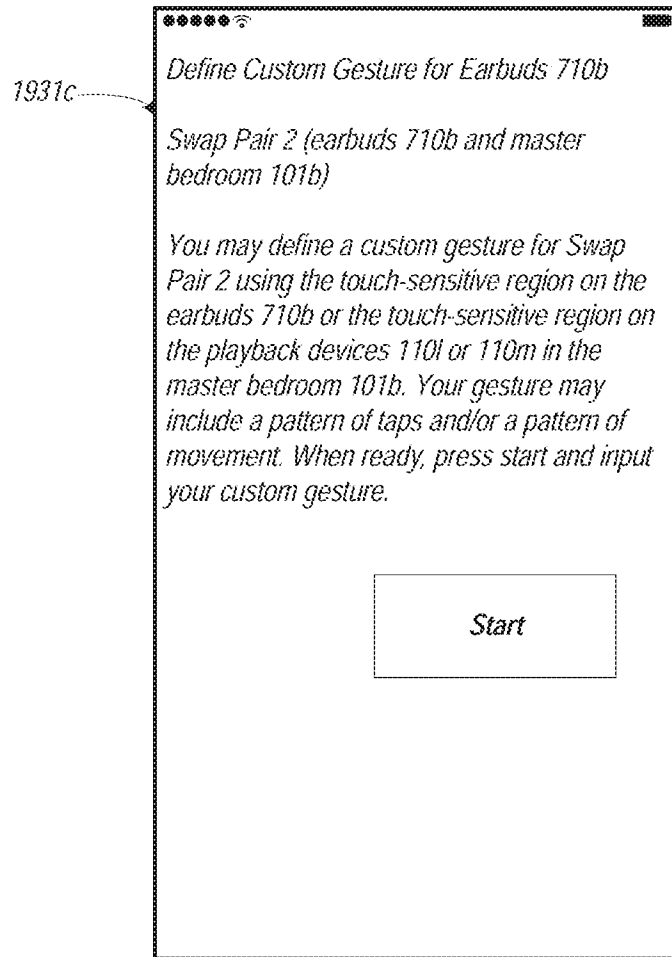


Fig. 19A

*Fig. 19B*

*Fig. 19C*

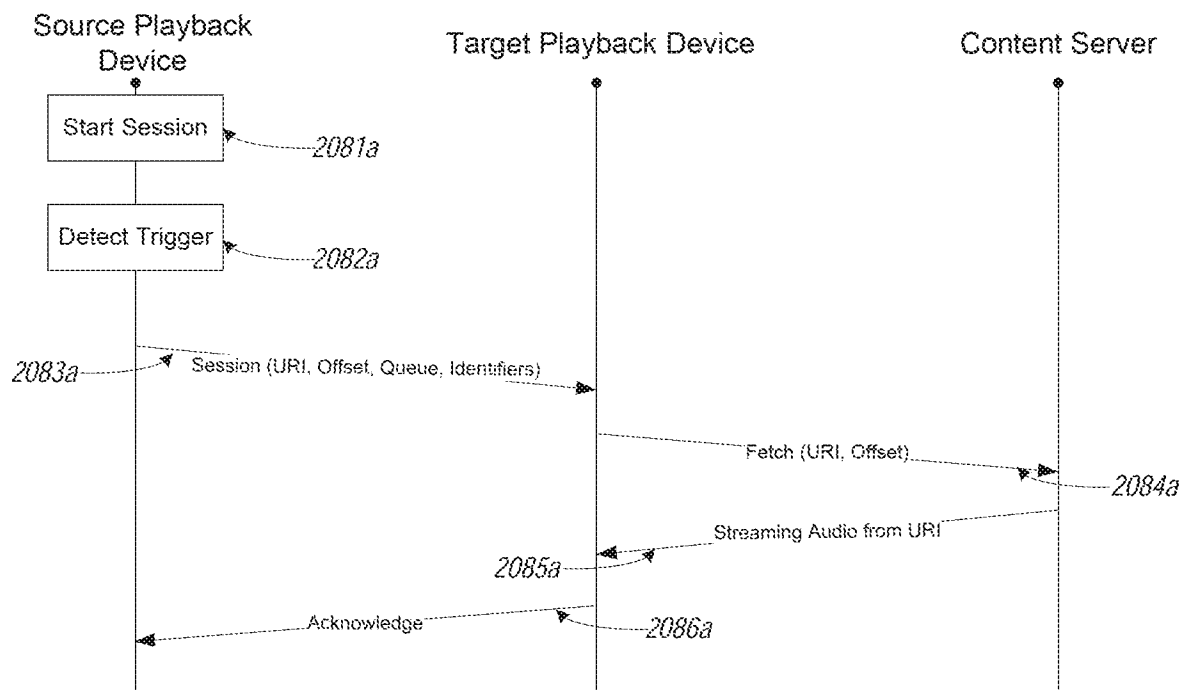


Fig. 20A

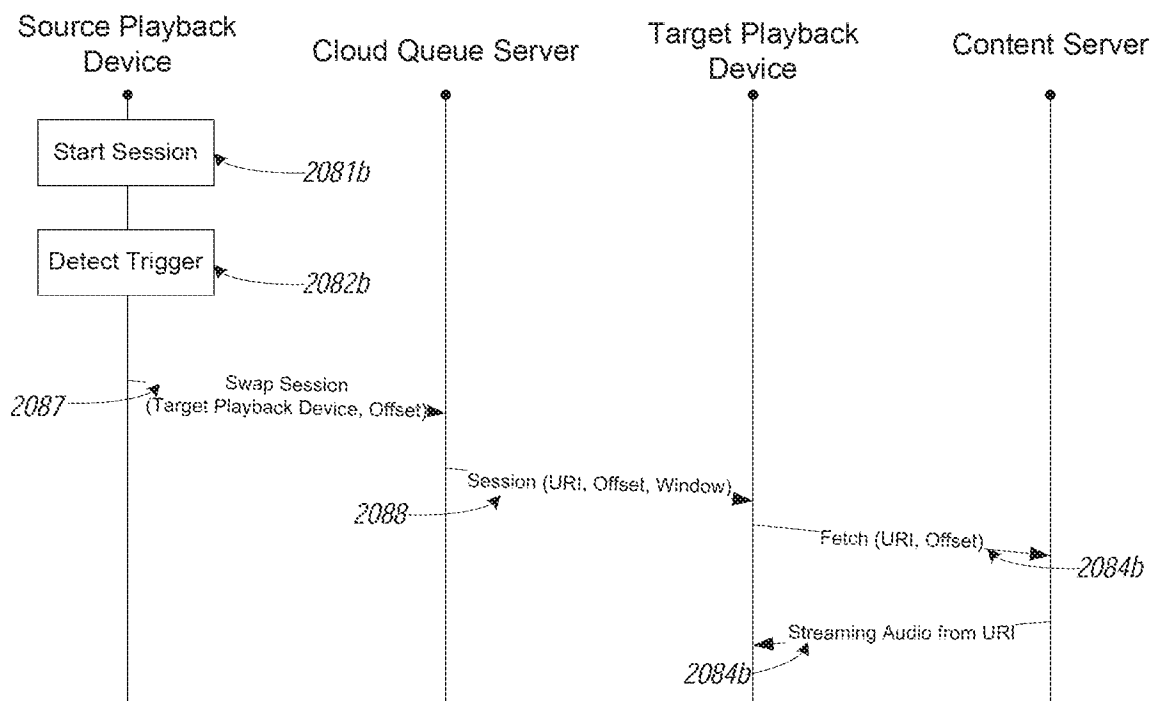
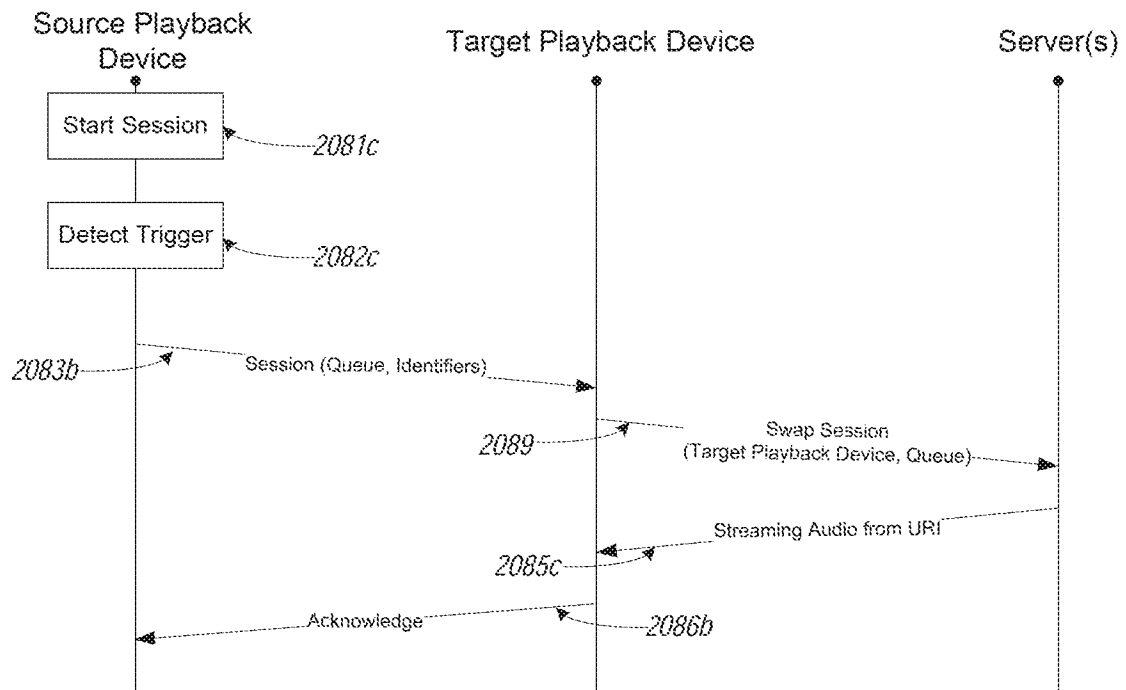
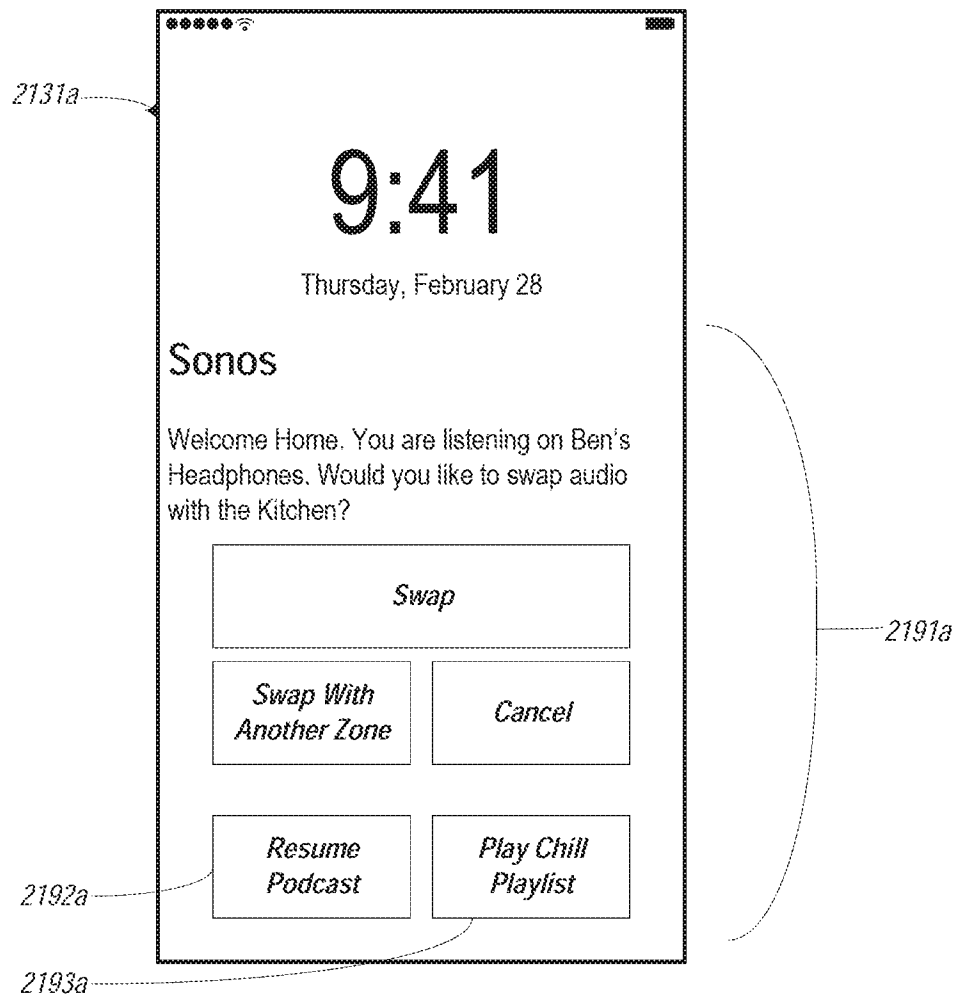
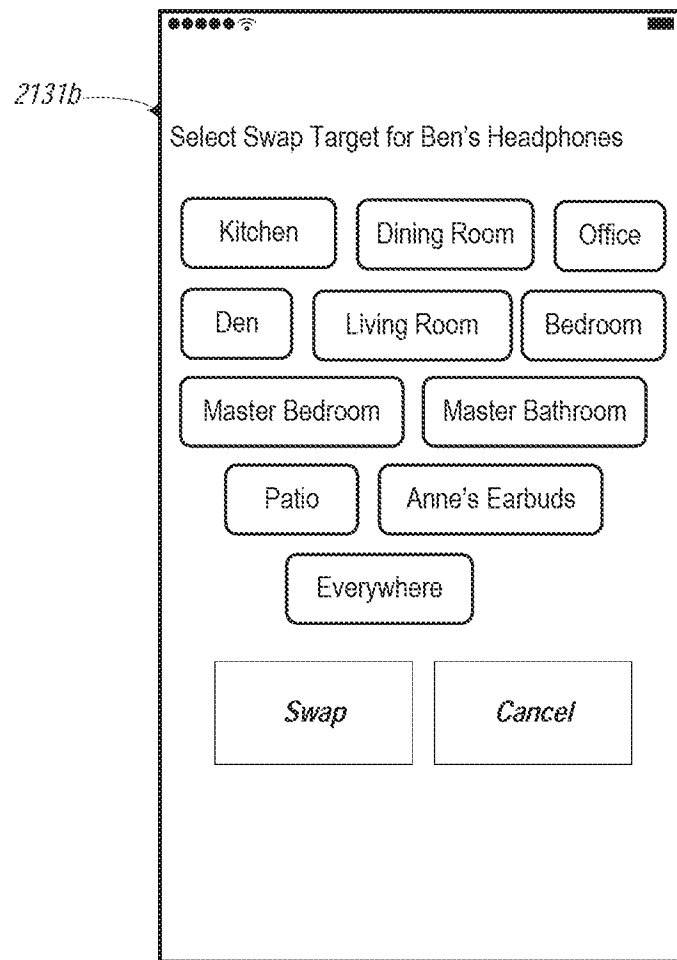
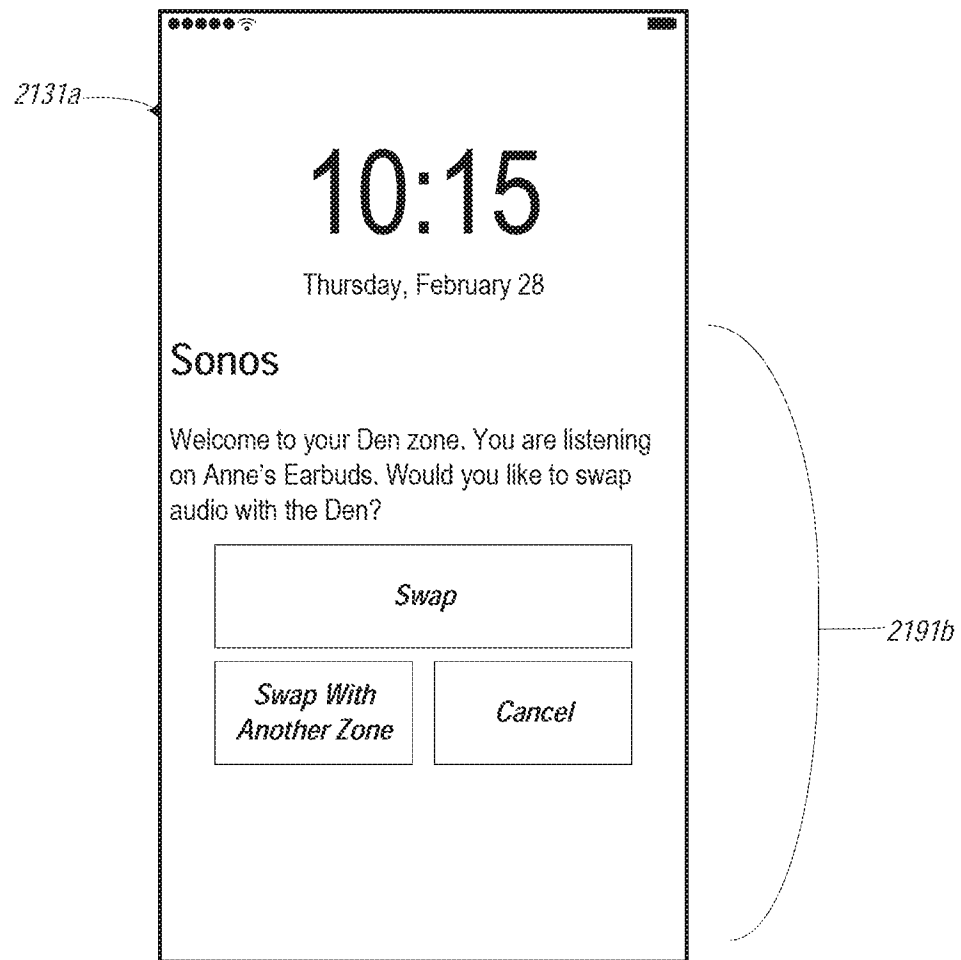


Fig. 20B

*Fig. 20C*

*Fig. 21A*

*Fig. 21B*

*Fig. 21C*

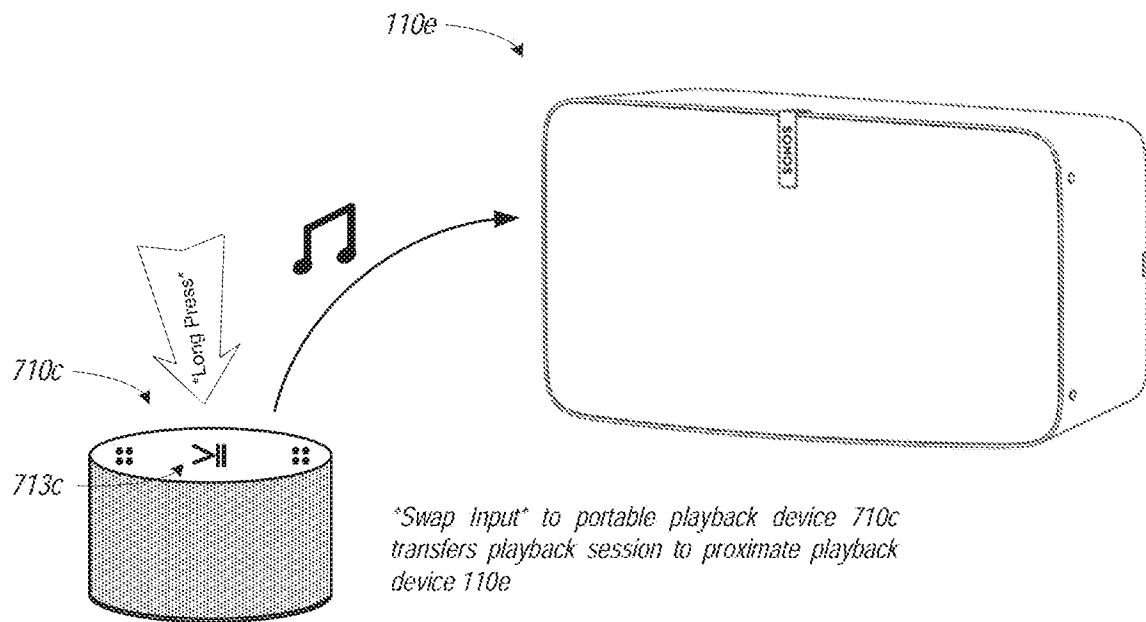


Fig. 22A

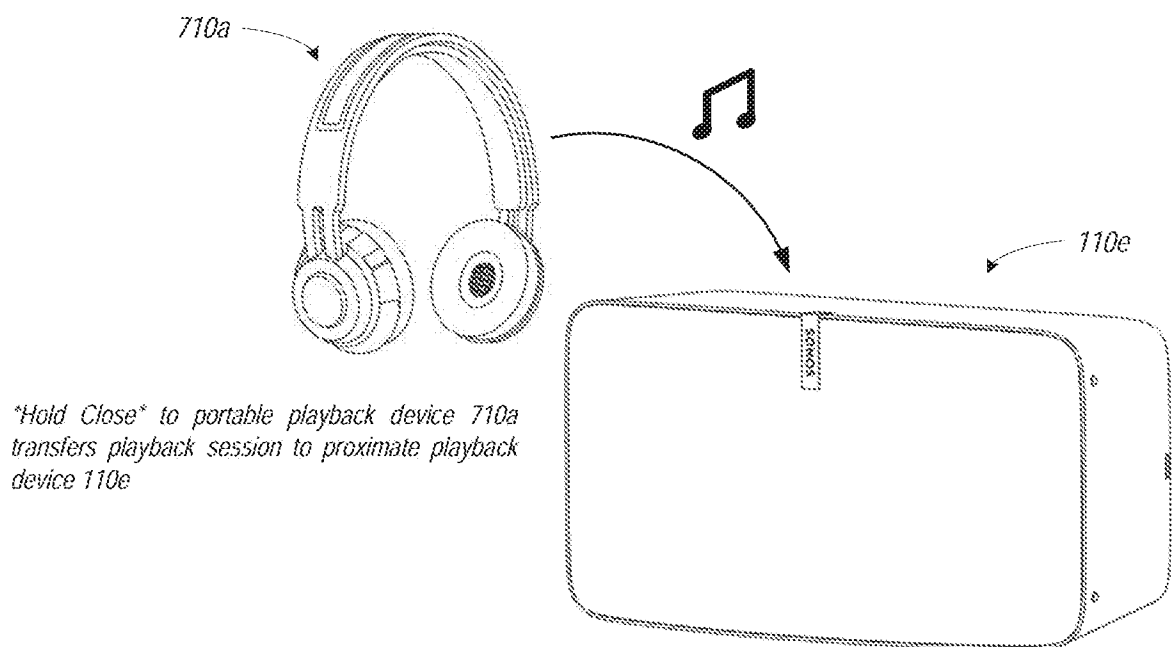


Fig. 22B

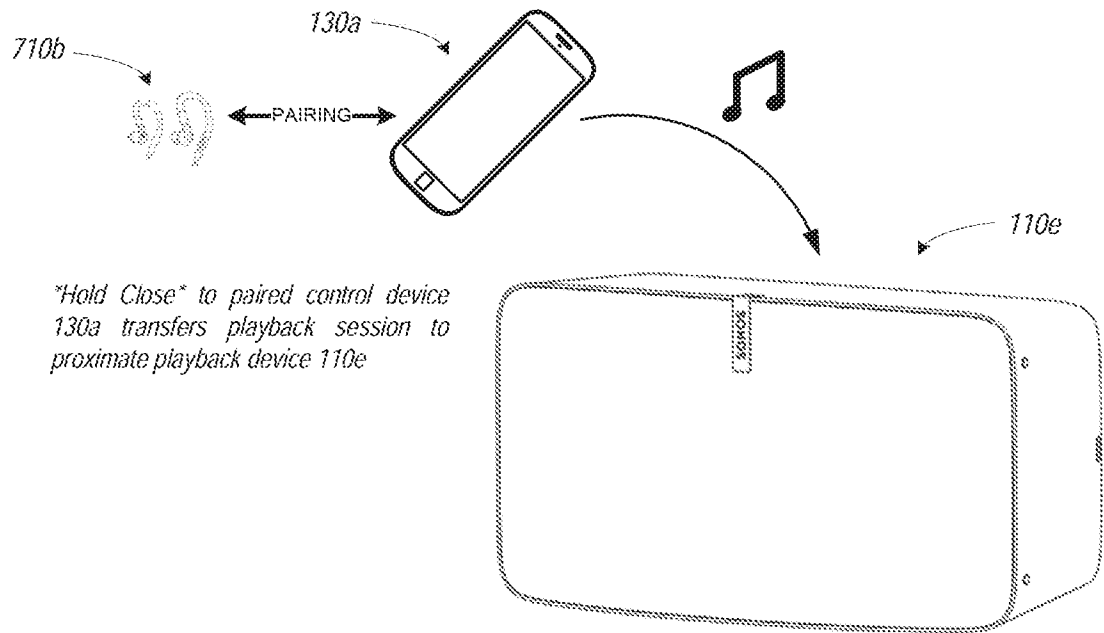


Fig. 22C

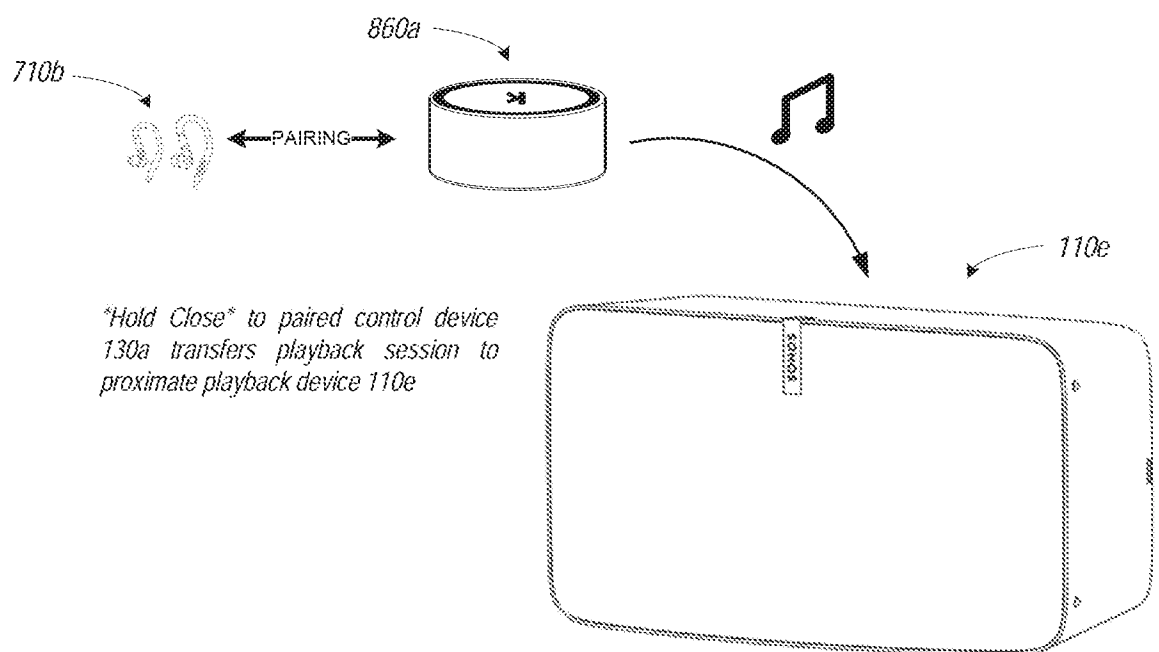


Fig. 22D

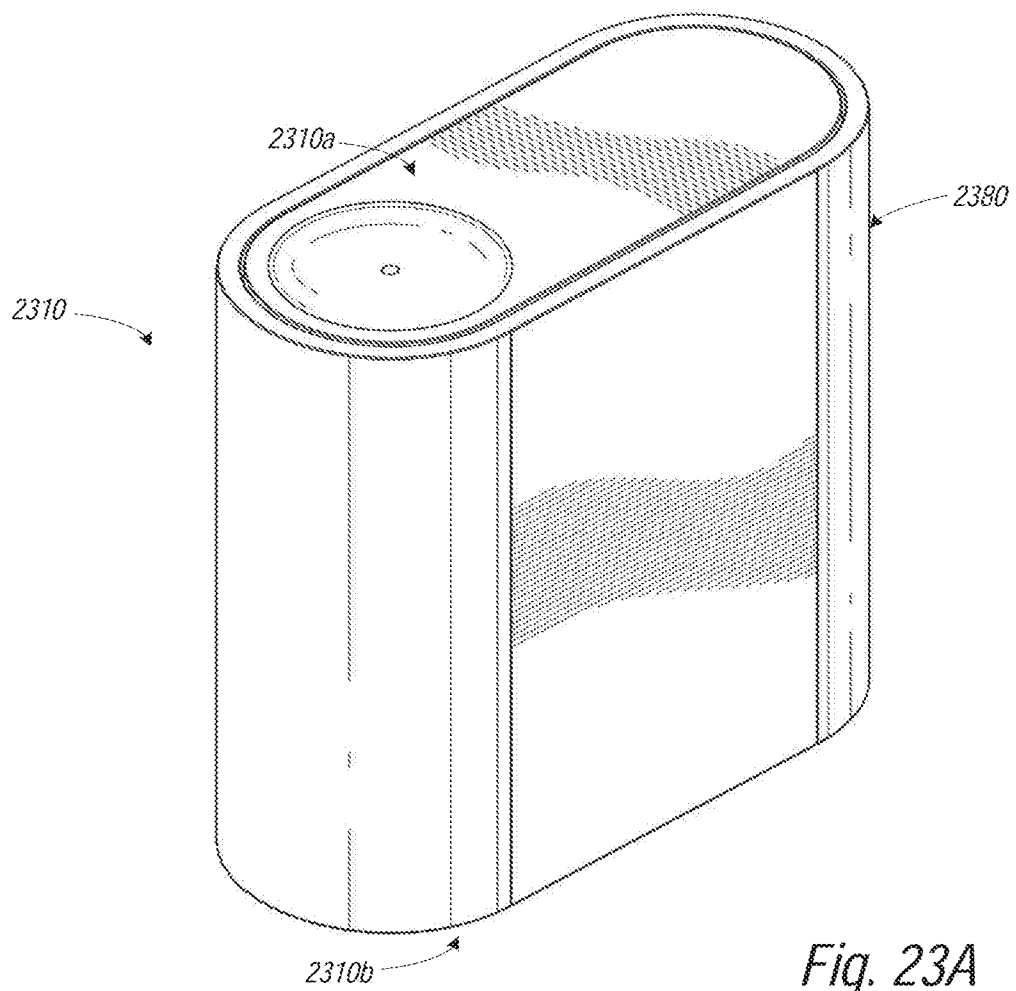
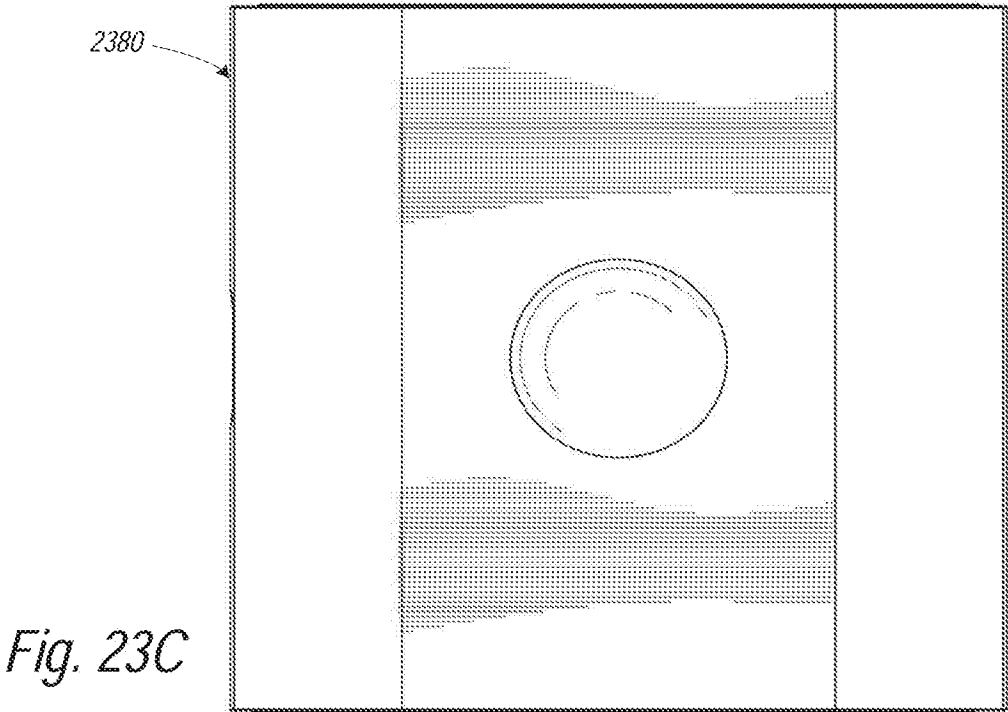
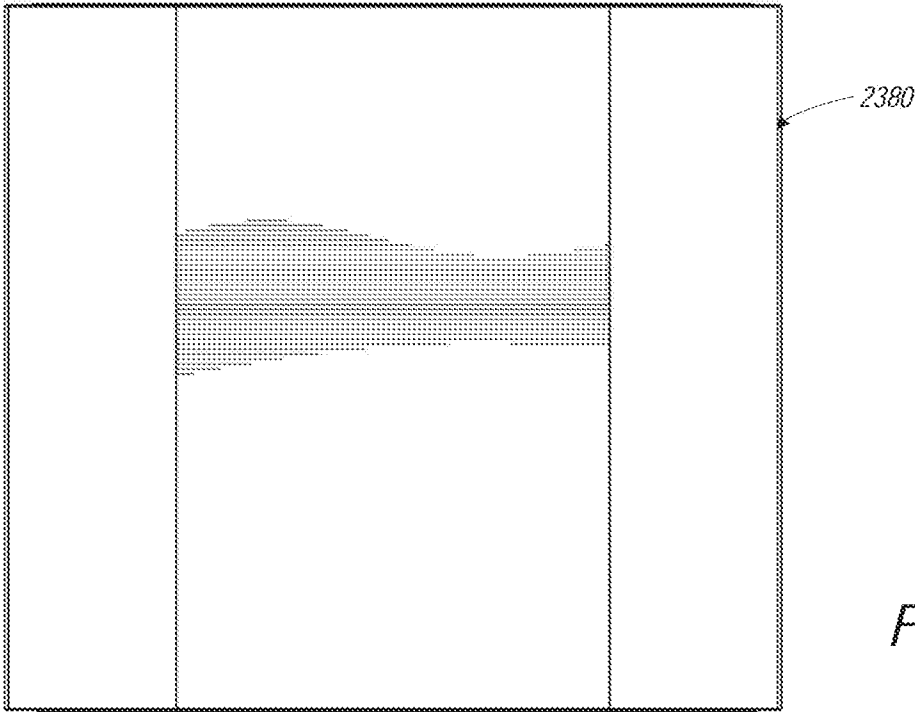


Fig. 23A



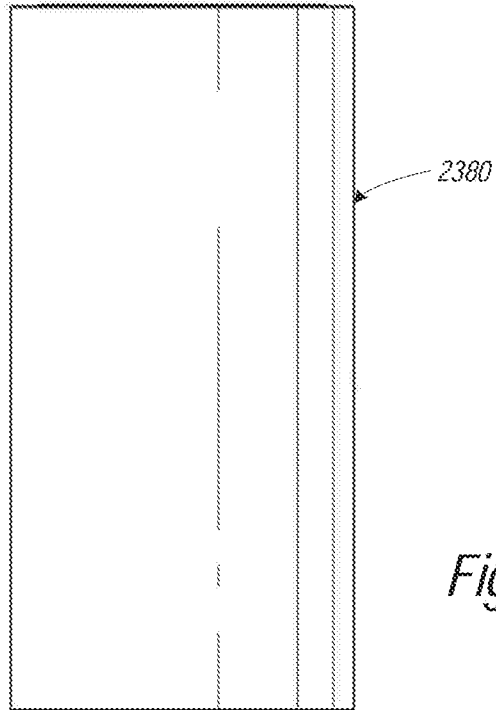


Fig. 23D

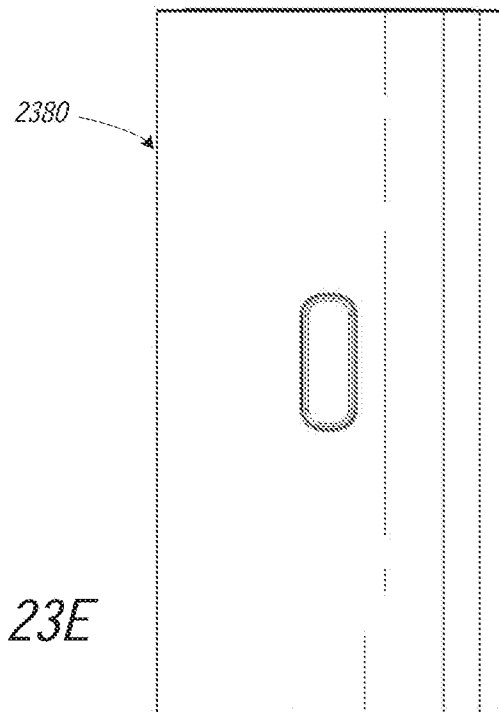


Fig. 23E

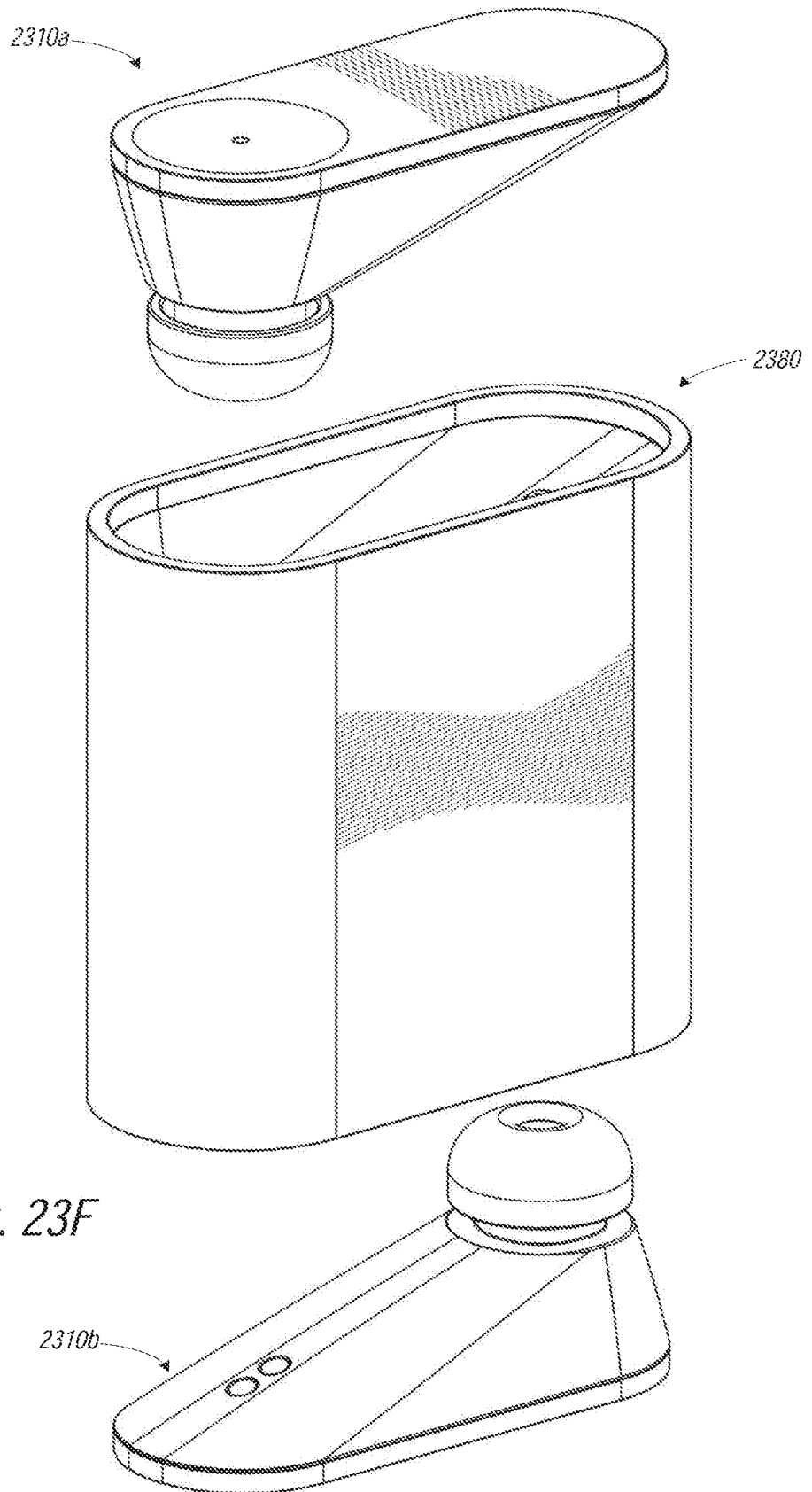


Fig. 23F

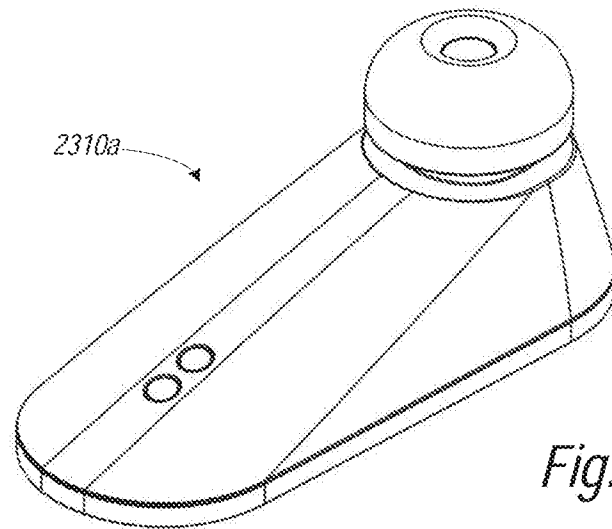


Fig. 23G

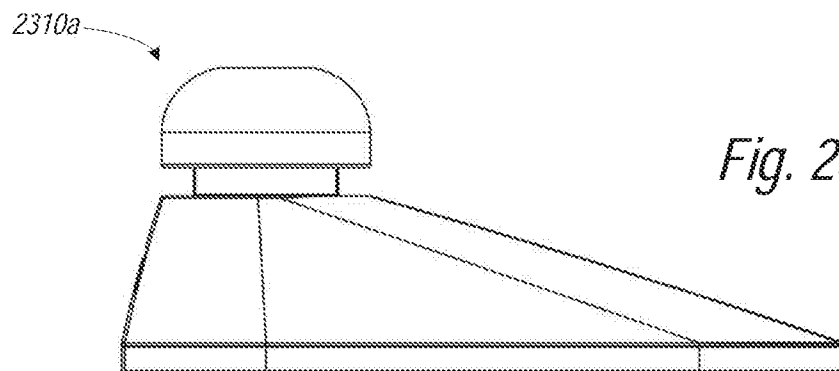


Fig. 23H

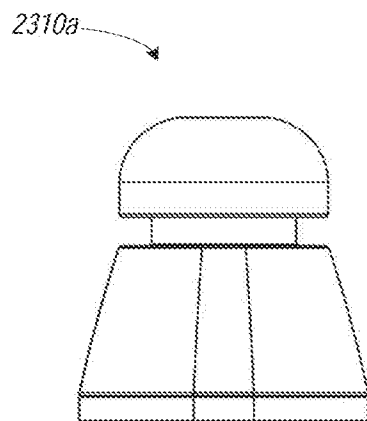
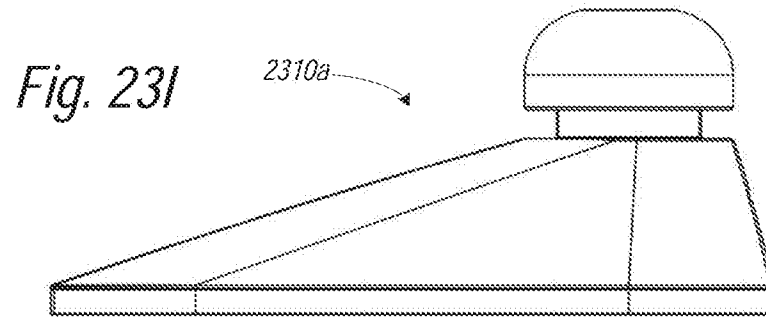


Fig. 23J

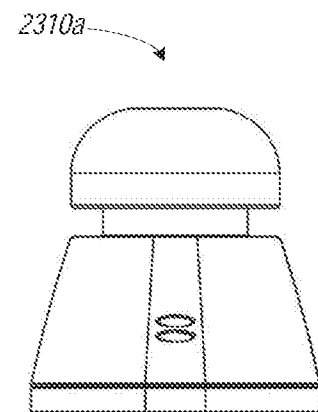


Fig. 23K

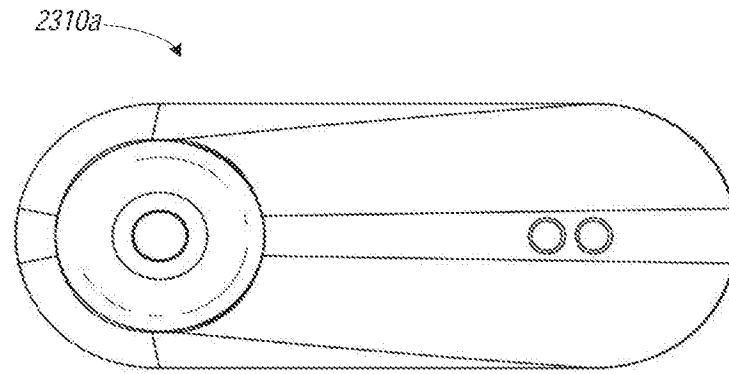


Fig. 23L

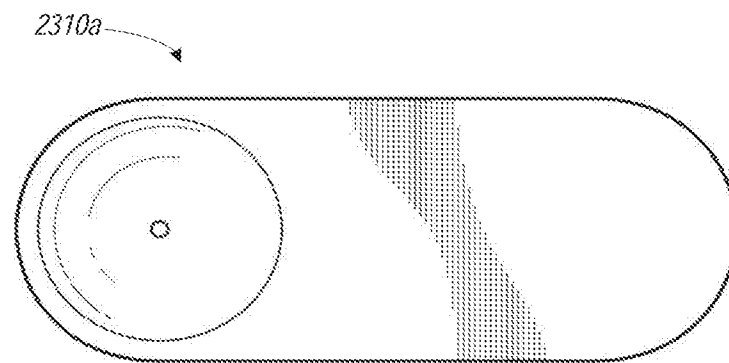


Fig. 23M

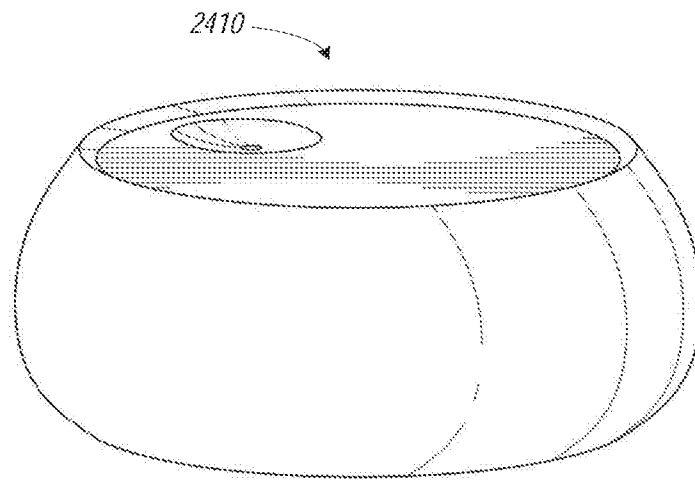


Fig. 24A

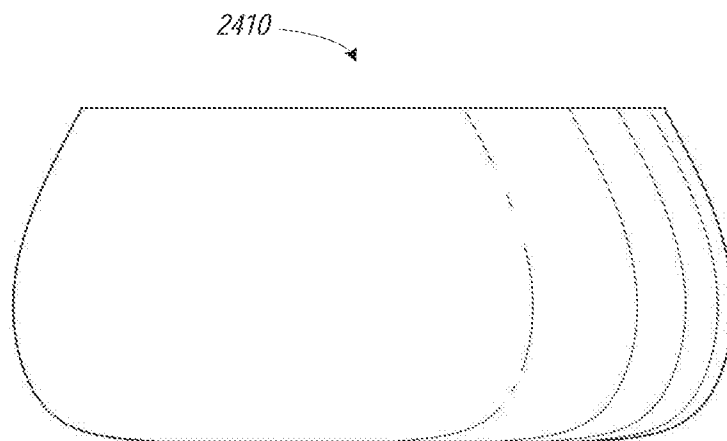


Fig. 24B

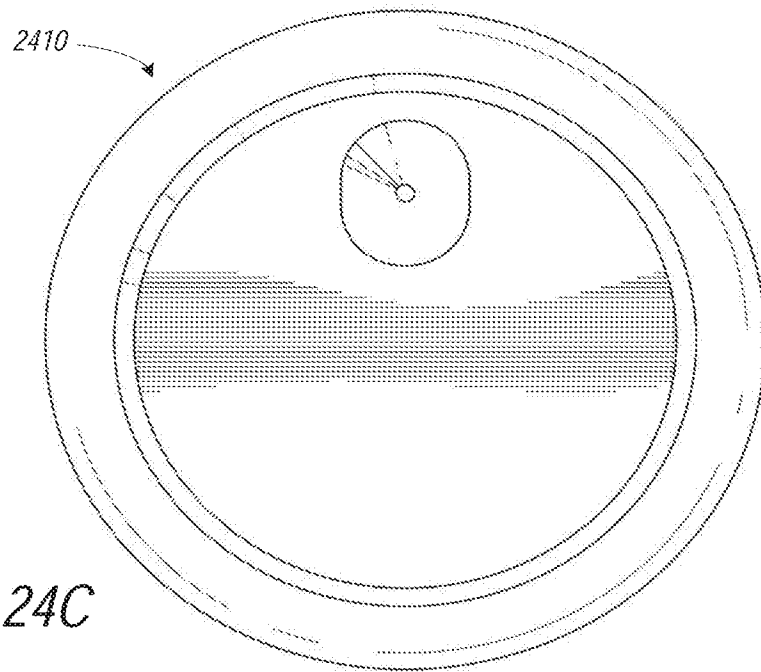


Fig. 24C

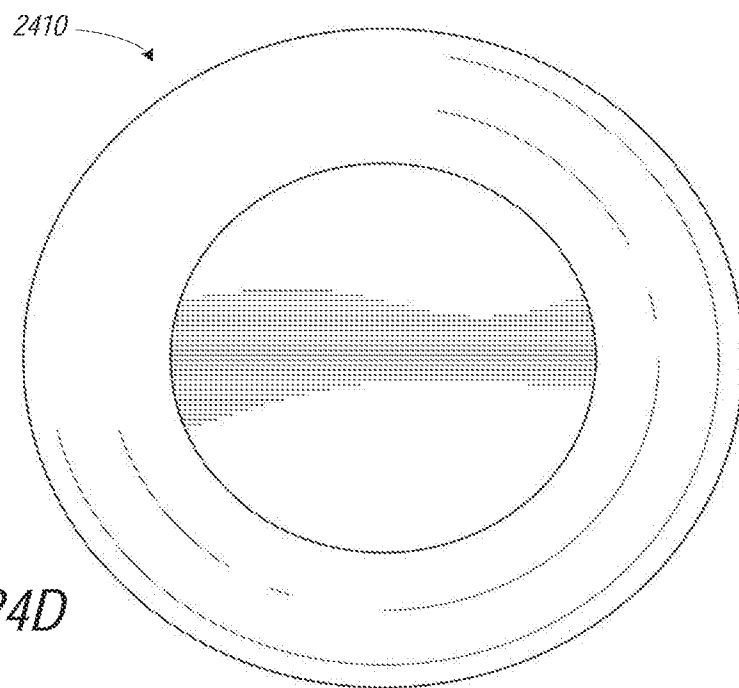


Fig. 24D

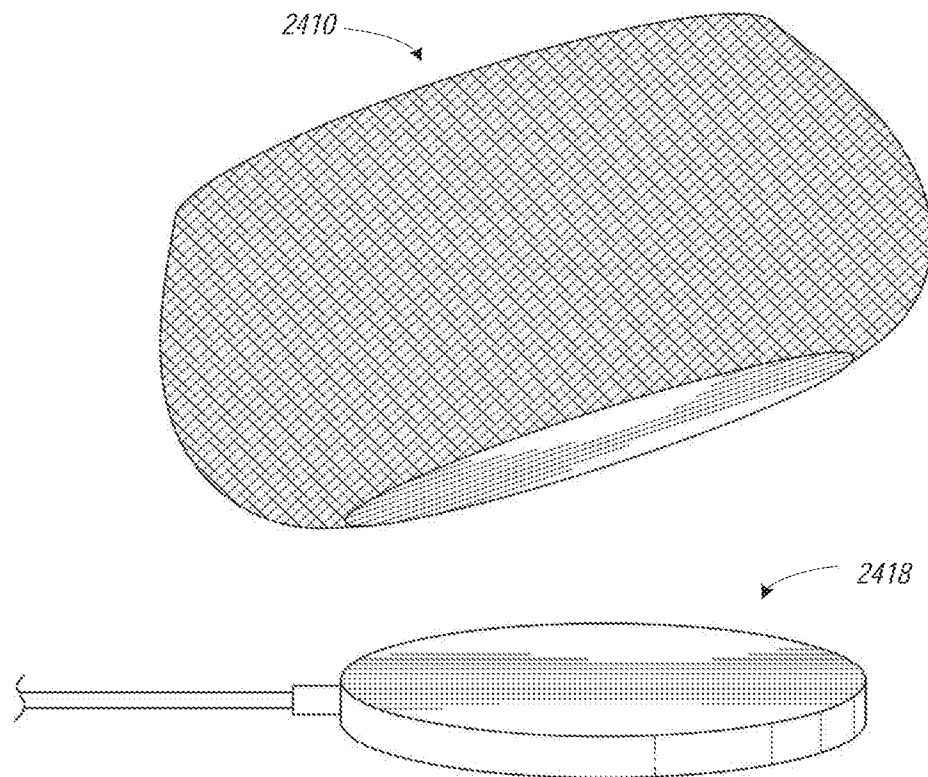


Fig. 24E

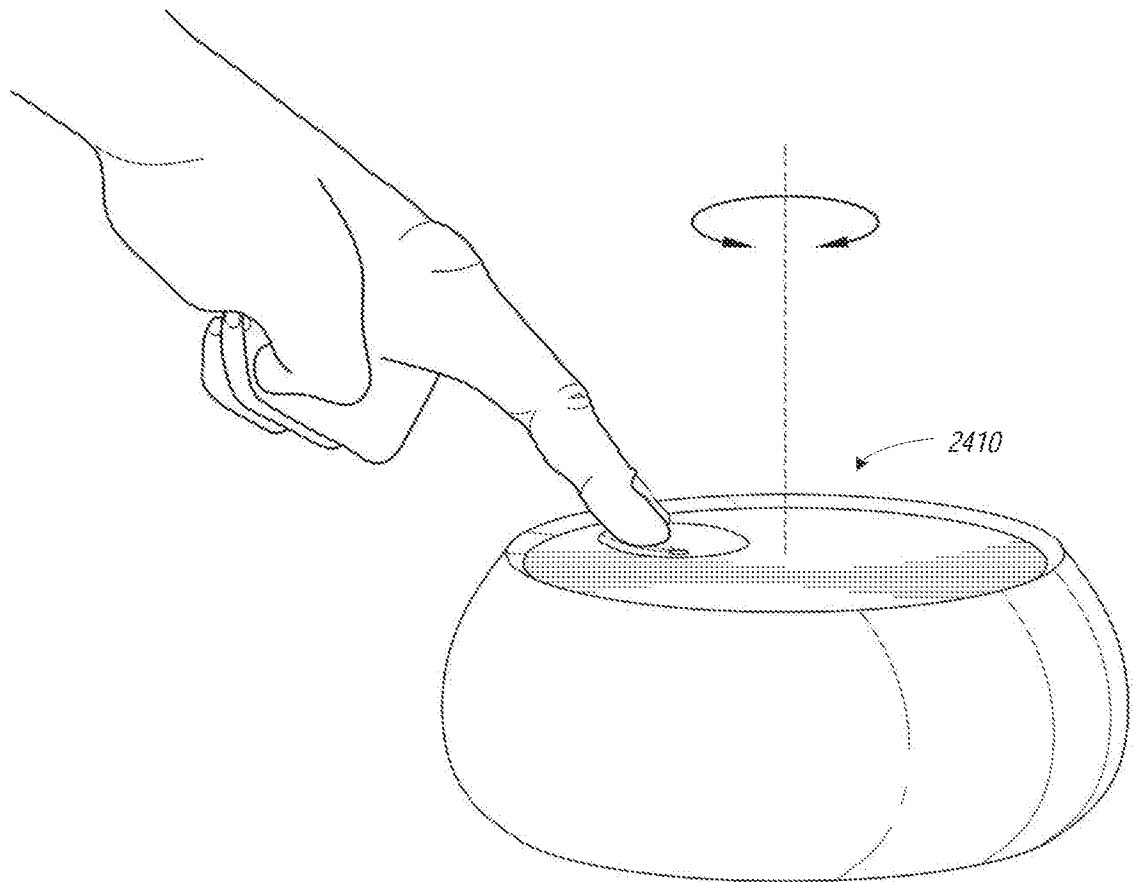


Fig. 24F

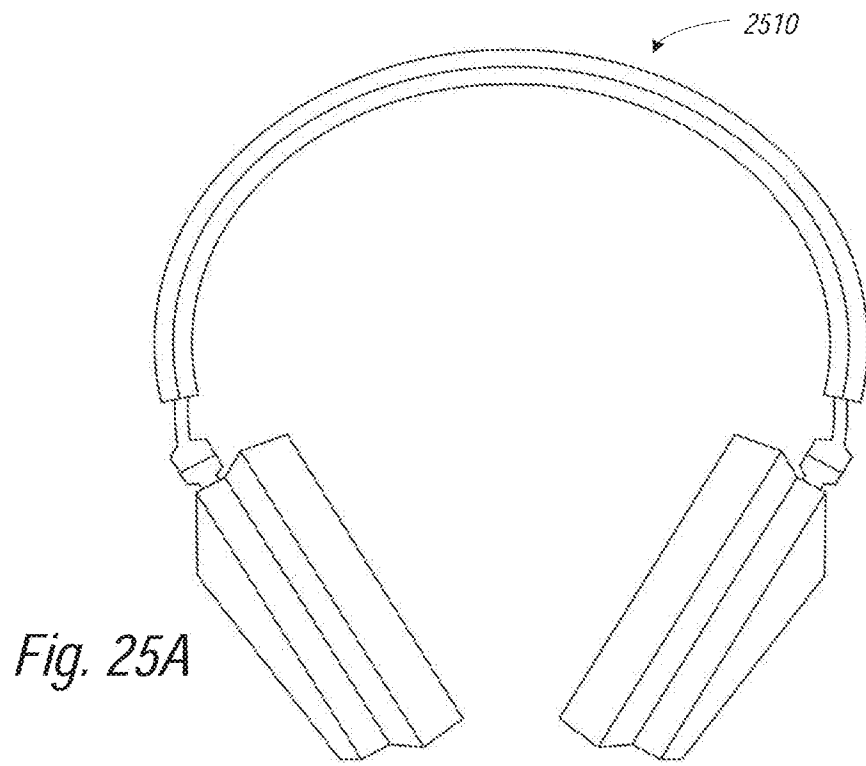


Fig. 25A

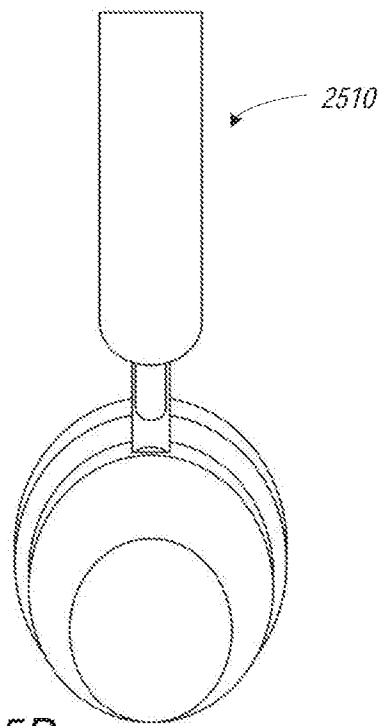


Fig. 25B

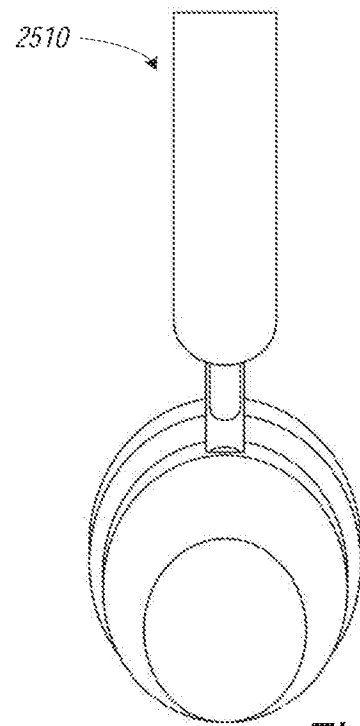
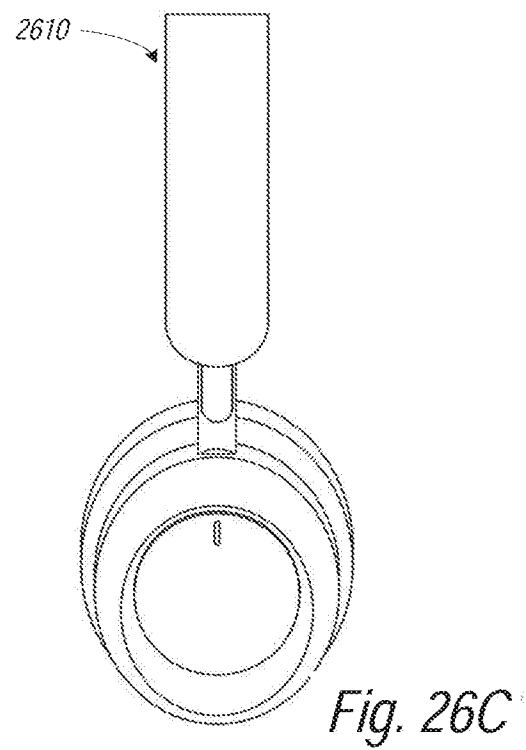
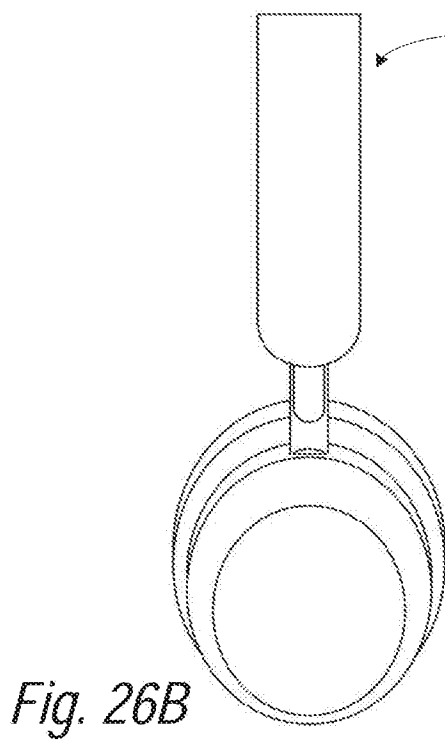
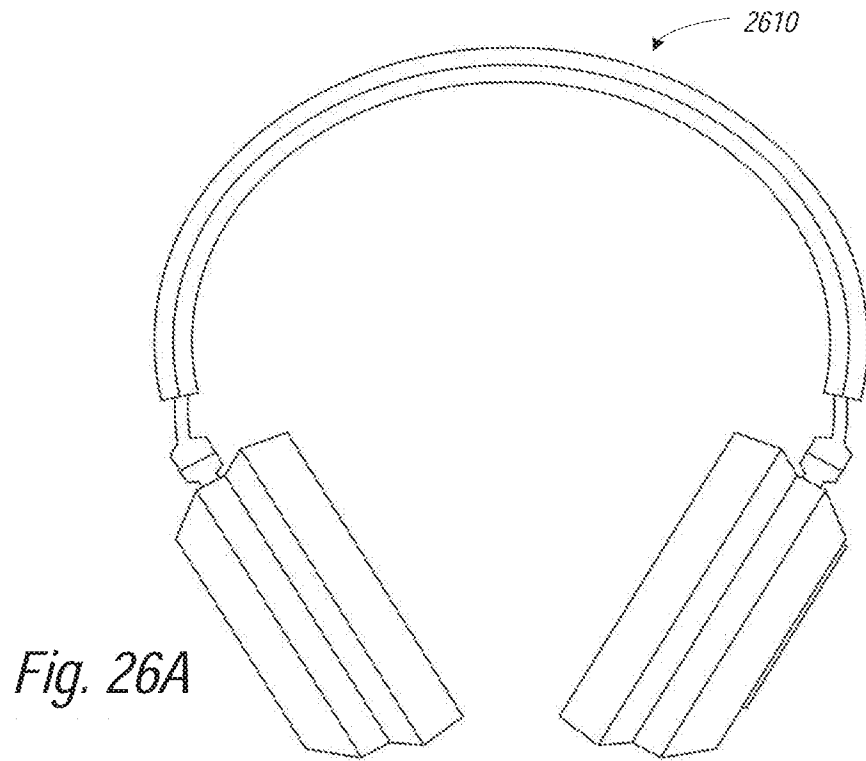


Fig. 25C



INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/020511

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04R27/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04R

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2018/041801 A1 (ZOUND INDUSTRIES INT AB [SE]) 8 March 2018 (2018-03-08) page 6, line 25 - page 9, line 5; figures 1, 2 page 13, line 6 - line 24 -----	1,2, 4-10, 13-22
X	US 2018/115844 A1 (LU YAO [US] ET AL) 26 April 2018 (2018-04-26) paragraph [0036]; figure 1 paragraphs [0034], [0048], [0060] paragraph [0051] paragraph [0057]; figures 5-9 paragraph [0061]; figure 10 paragraph [0045]; figure 2 ----- -/--	1,2,4, 6-11, 13-18



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

28 May 2020

Date of mailing of the international search report

08/06/2020

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Betgen, Benjamin

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2020/020511

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2014/115114 A1 (GARMARK STEN [SE] ET AL) 24 April 2014 (2014-04-24) paragraphs [0043], [0044] paragraphs [0050], [0078]; figure 2 -----	1-4,6, 8-10, 13-21
X	WO 2017/197187 A1 (GOOGLE LLC [US]) 16 November 2017 (2017-11-16) paragraph [0008] paragraph [0073] - paragraph [0078]; figure 3 paragraph [0089]; figure 4F -----	1-4,6, 8-10, 12-15, 18-21
X	US 2015/371529 A1 (DOLECKI ERIC E [US]) 24 December 2015 (2015-12-24) paragraph [0013] paragraph [0054] paragraph [0086]; figures 3A-3C paragraph [0105] - paragraph [0109]; figure 5 -----	1,2,4, 6-10, 12-21
A	US 10 001 965 B1 (LANG JONATHAN P [US] ET AL) 19 June 2018 (2018-06-19) figures 5, 6 -----	14
A	US 2016/286313 A1 (KOFMAN IGOR [US] ET AL) 29 September 2016 (2016-09-29) paragraph [0025]; figures 1A, 1B -----	14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2020/020511

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 2018041801	A1	08-03-2018	CN	109644306	A	16-04-2019
			EP	3504884	A1	03-07-2019
			US	2019179605	A1	13-06-2019
			WO	2018041801	A1	08-03-2018

US 2018115844	A1	26-04-2018	CN	107431860	A	01-12-2017
			EP	3272135	A1	24-01-2018
			US	2018115844	A1	26-04-2018
			WO	2016145235	A1	15-09-2016

US 2014115114	A1	24-04-2014	US	2014115114	A1	24-04-2014
			US	2017019441	A1	19-01-2017
			US	2019081997	A1	14-03-2019
			WO	2014064531	A1	01-05-2014

WO 2017197187	A1	16-11-2017	WO	2017197184	A1	16-11-2017
			WO	2017197186	A1	16-11-2017
			WO	2017197187	A1	16-11-2017

US 2015371529	A1	24-12-2015	US	2015371529	A1	24-12-2015
			US	2016378429	A1	29-12-2016
			WO	2015199927	A1	30-12-2015

US 10001965	B1	19-06-2018	US	10001965	B1	19-06-2018
			US	2018357039	A1	13-12-2018
			US	2020097250	A1	26-03-2020

US 2016286313	A1	29-09-2016	NONE			
