US 20200100047A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2020/0100047 A1**

Lyren (43) **Pub. Date: Mar. 26, 2020**

(54) **ALTERING A GRAPHICAL REPRESENTATION TO INDICATE SOUND WILL EXTERNALLY LOCALIZE AS BINAURAL SOUND**

(71) Applicant: **Philip Scott Lyren**, Bangkok (TH)

(72) Inventor: **Philip Scott Lyren**, Bangkok (TH)

(21) Appl. No.: **16/695,113**

(22) Filed: **Nov. 25, 2019**

Related U.S. Application Data

(63) Continuation of application No. 16/354,185, filed on Mar. 15, 2019, now Pat. No. 10,499,184, which is a continuation of application No. 16/059,049, filed on Aug. 9, 2018, now Pat. No. 10,299,065, which is a continuation of application No. 15/898,587, filed on Feb. 18, 2018, now Pat. No. 10,063,990.

Publication Classification

(51) **Int. Cl.**

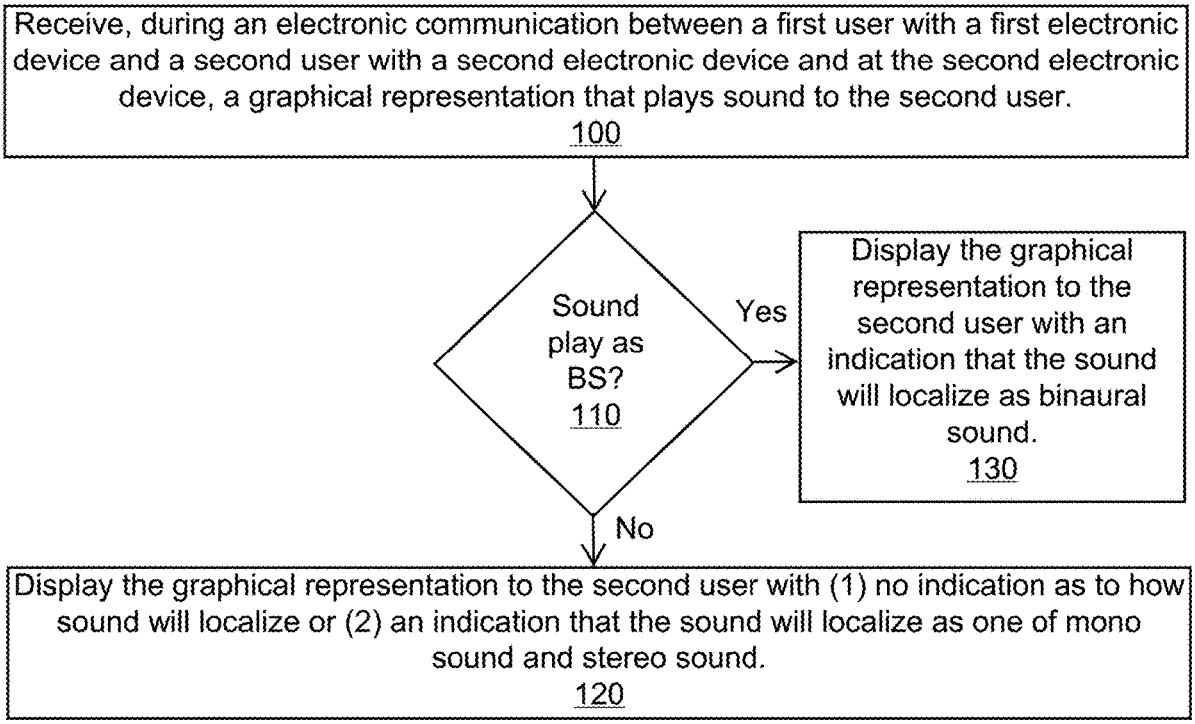
H04S 7/00	(2006.01)
H04M 1/725	(2006.01)
H04M 1/60	(2006.01)
G06F 3/0481	(2006.01)
G06T 11/60	(2006.01)

(52) **U.S. Cl.**

CPC ***H04S 7/308*** (2013.01); ***H04M 1/72547*** (2013.01); *H04R 5/02* (2013.01); ***G06F 3/0481*** (2013.01); ***G06T 11/60*** (2013.01); ***H04M 1/6066*** (2013.01)

(57) **ABSTRACT**

A method alters an appearance of an emoji based on whether a listener is wearing headphones or earphones. During an electronic communication, an electronic device of the user receives the emoji and sound to play to the user. The electronic device determines whether the user is wearing headphones or earphones and alters an appearance of the emoji based on this determination so the user knows that the sound will externally localize as binaural sound.



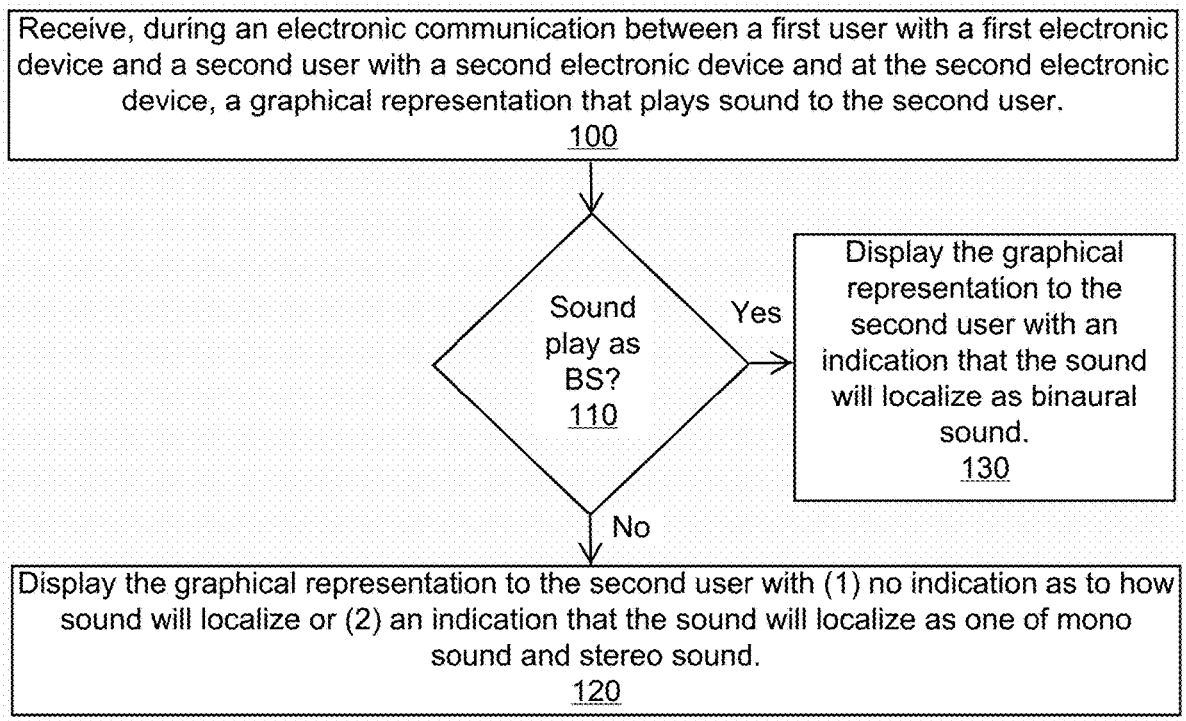


Figure 1

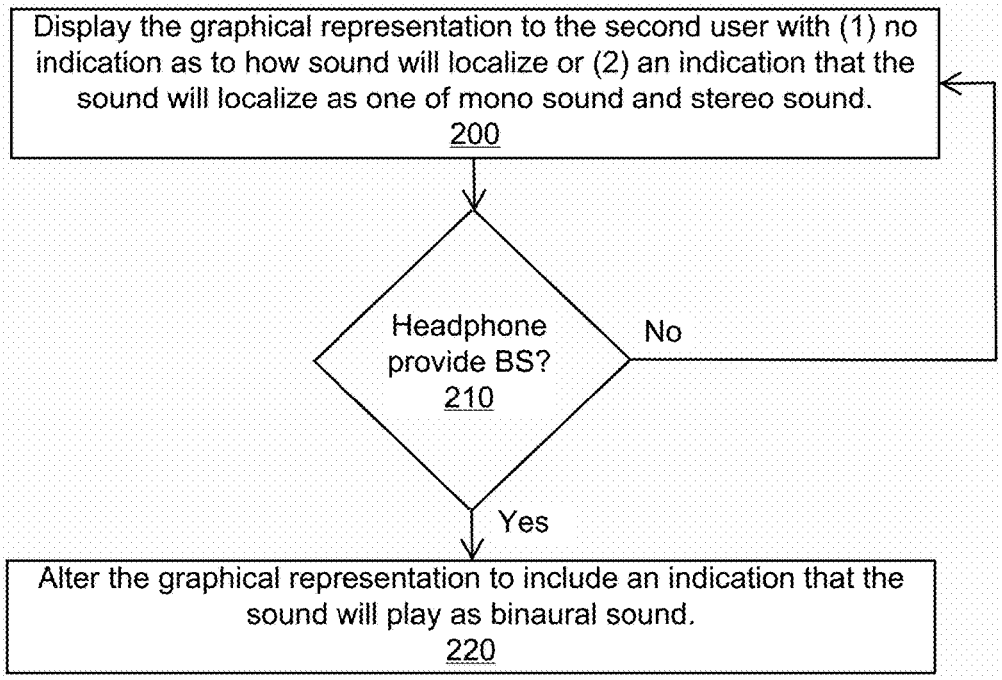


Figure 2

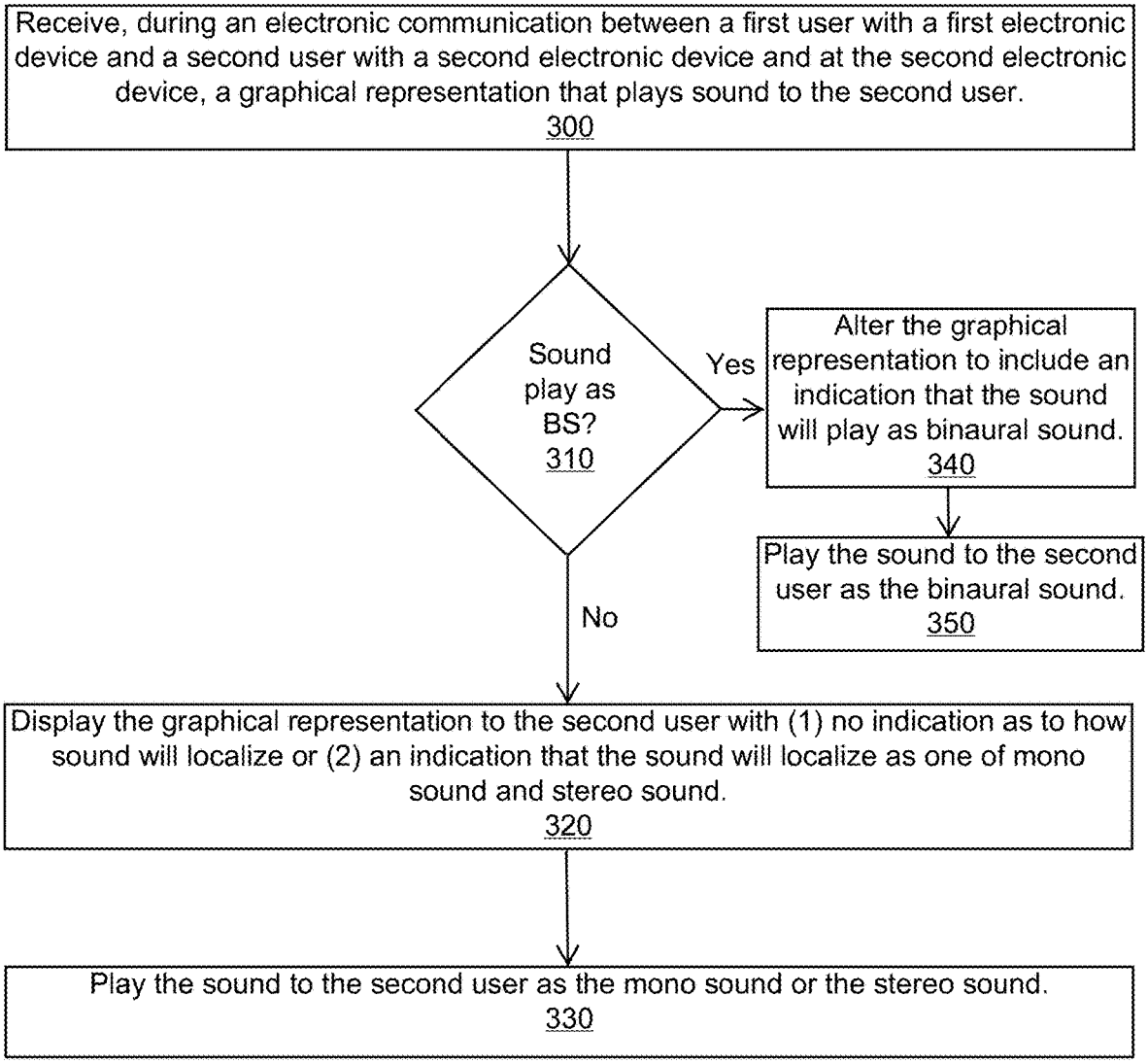

Receive, during an electronic communication between a first user with a first electronic device and a second user with a second electronic device and at the second electronic device, a graphical representation that plays sound to the second user.
300
Sound play as BS?
310
Yes
Alter the graphical representation to include an indication that the sound will play as binaural sound.
340
Play the sound to the second user as the binaural sound.
350
No
Display the graphical representation to the second user with (1) no indication as to how sound will localize or (2) an indication that the sound will localize as one of mono sound and stereo sound.
320
Play the sound to the second user as the mono sound or the stereo sound.
330

Figure 3

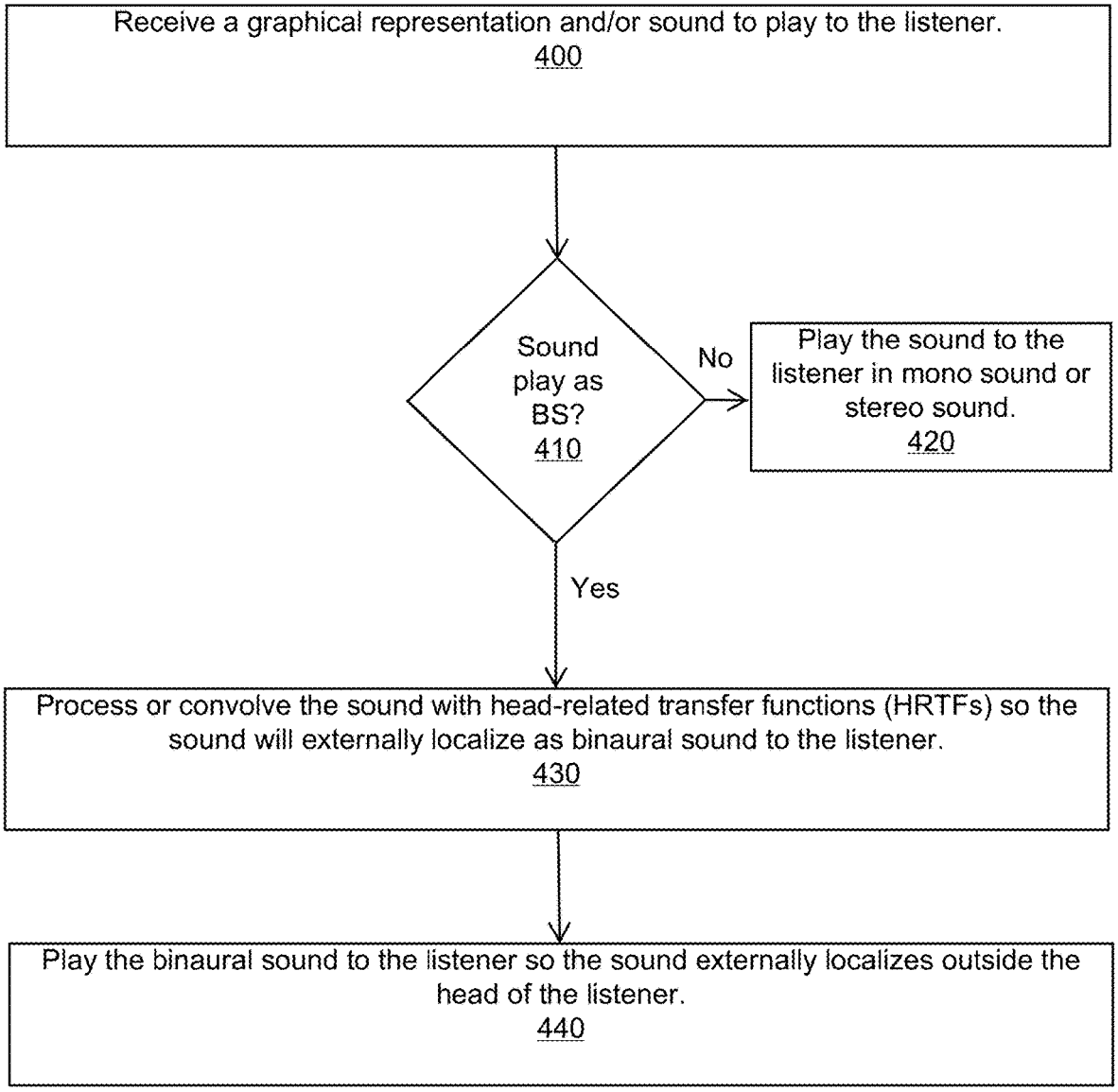

Receive a graphical representation and/or sound to play to the listener.
400
Sound play as BS?
410
No
Play the sound to the listener in mono sound or stereo sound.
420
Yes
Process or convolve the sound with head-related transfer functions (HRTFs) so the sound will externally localize as binaural sound to the listener.
430
Play the binaural sound to the listener so the sound externally localizes outside the head of the listener.
440

Figure 4

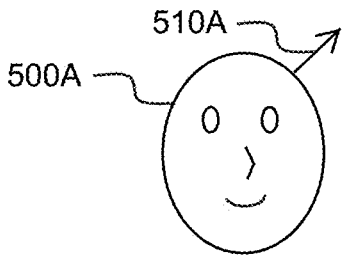


Figure 5A

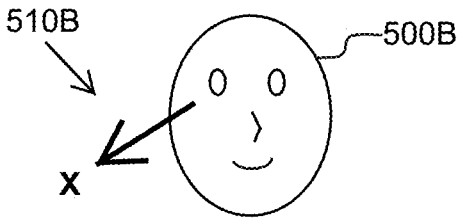


Figure 5B

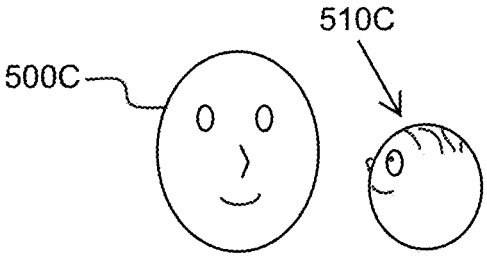


Figure 5C

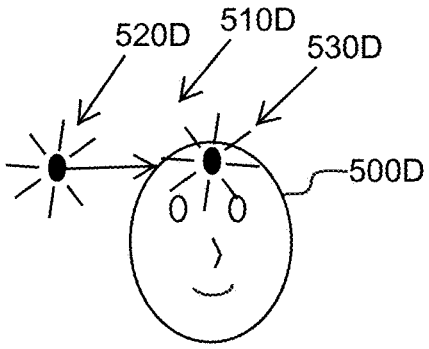


Figure 5D

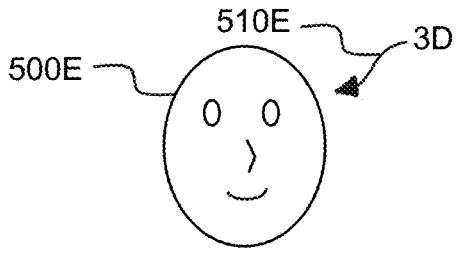


Figure 5E

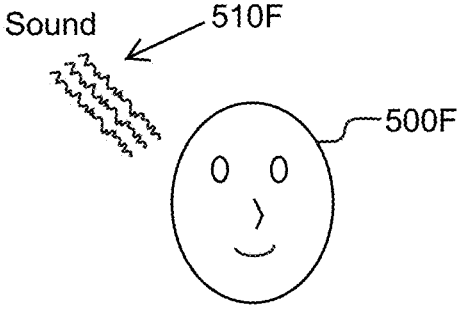


Figure 5F

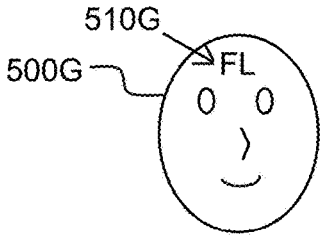


Figure 5G

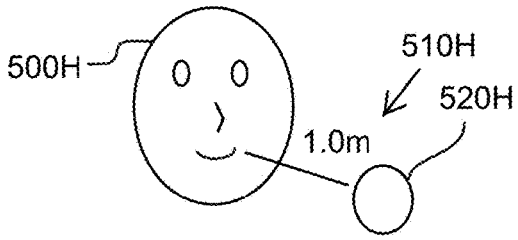


Figure 5H

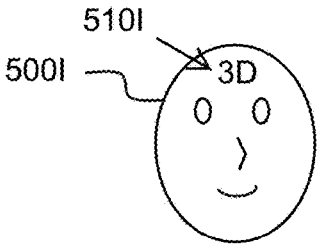


Figure 5I

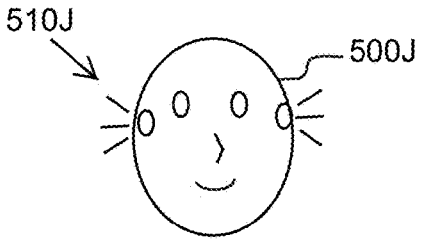


Figure 5J

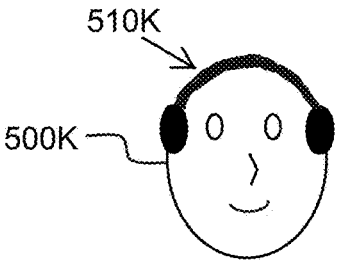


Figure 5K

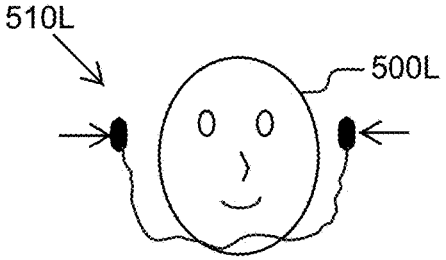


Figure 5L

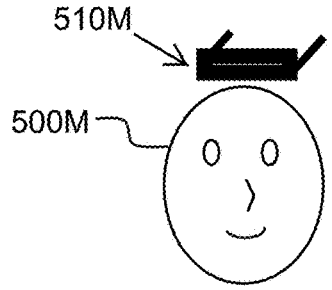


Figure 5M

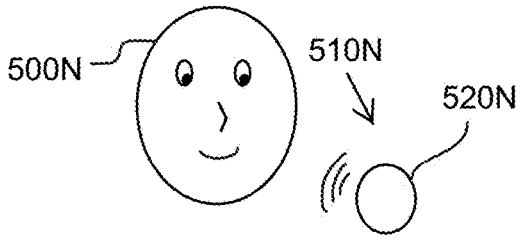


Figure 5N

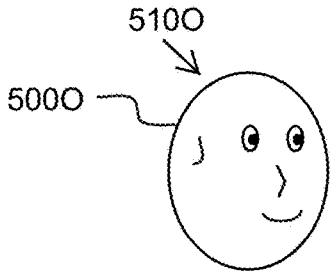


Figure 5O

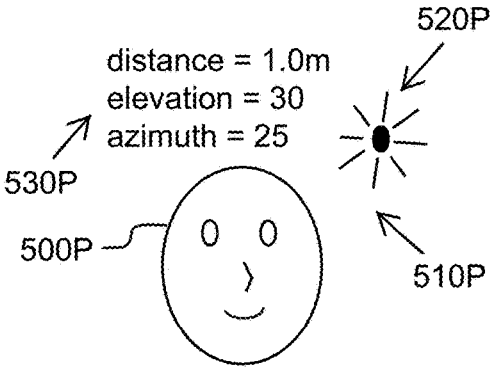


Figure 5P

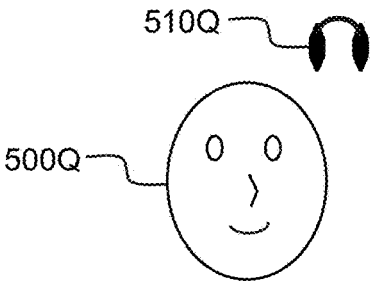


Figure 5Q

Figure 5R

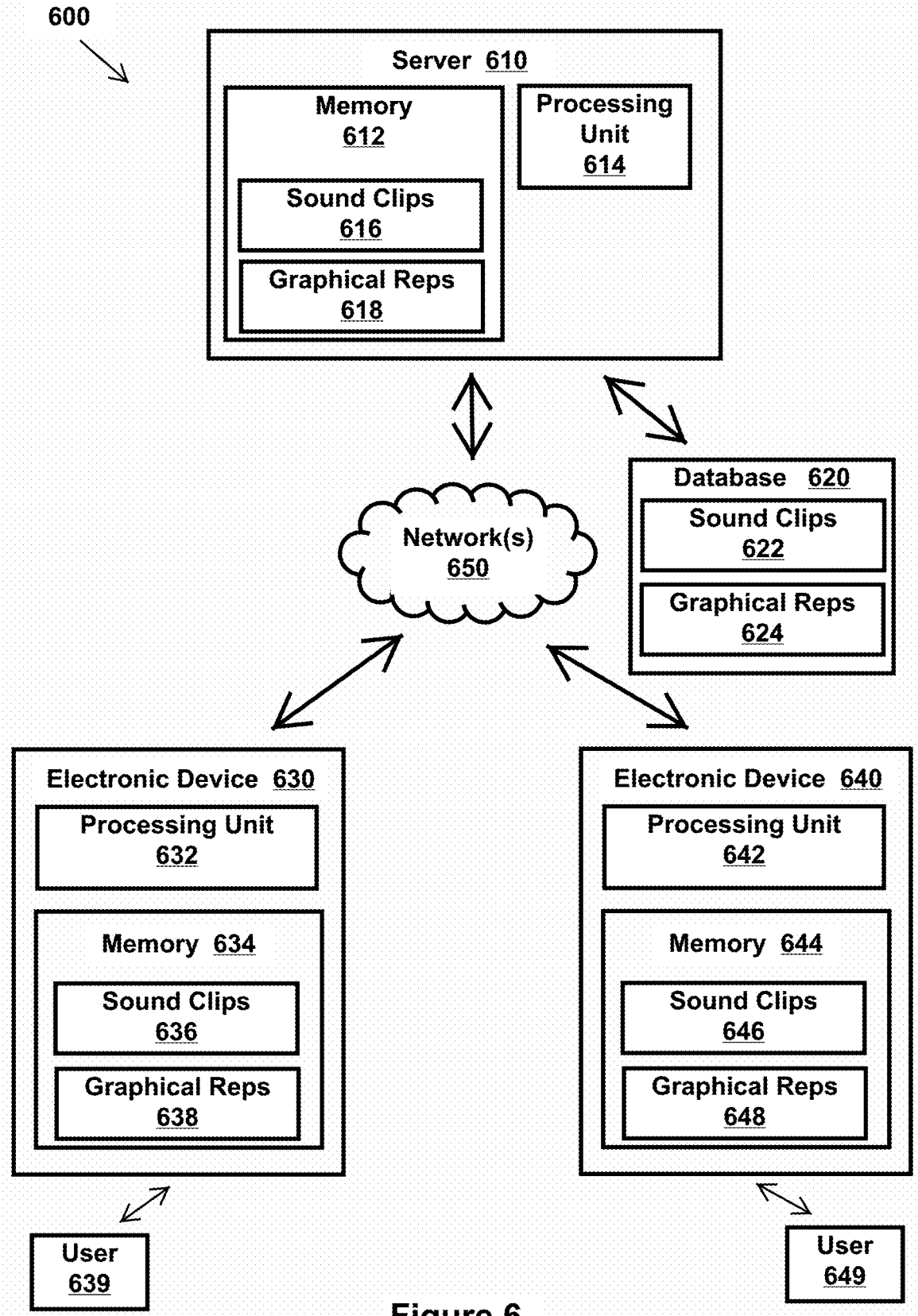

600
Server 610
Memory 612
Sound Clips 616
Graphical Reps 618
Processing Unit 614
Network(s) 650
Database 620
Sound Clips 622
Graphical Reps 624
Electronic Device 630
Processing Unit 632
Memory 634
Sound Clips 636
Graphical Reps 638
Electronic Device 640
Processing Unit 642
Memory 644
Sound Clips 646
Graphical Reps 648
User 639
User 649

Figure 6

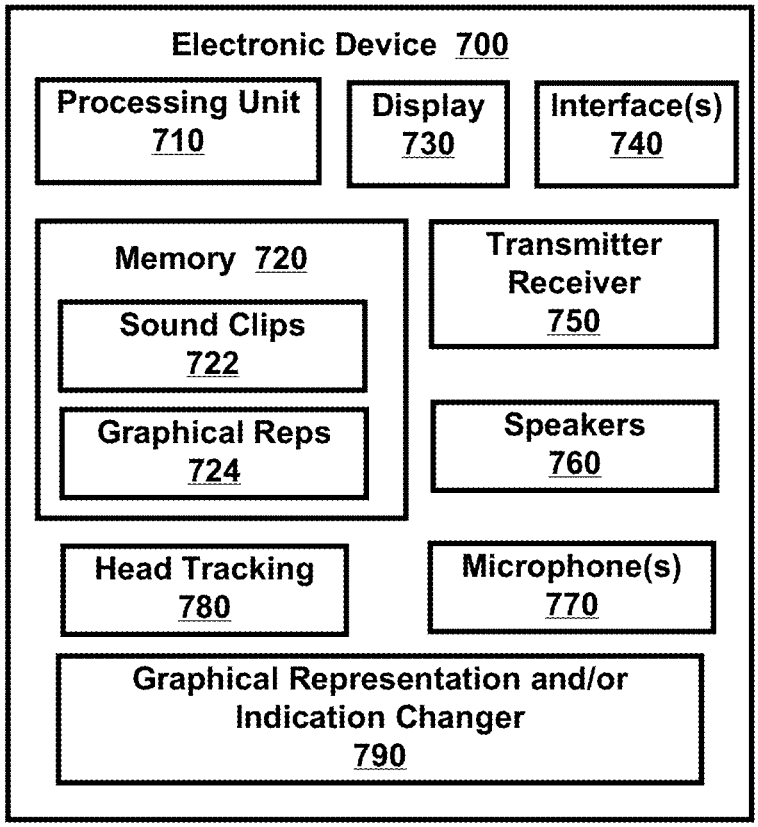

Electronic Device 700
Processing Unit 710
Display 730
Interface(s) 740
Memory 720
Sound Clips 722
Graphical Reps 724
Transmitter Receiver 750
Speakers 760
Head Tracking 780
Microphone(s) 770
Graphical Representation and/or Indication Changer 790

Figure 7

ALTERING A GRAPHICAL REPRESENTATION TO INDICATE SOUND WILL EXTERNALLY LOCALIZE AS BINAURAL SOUND

BACKGROUND

[0001] Three-dimensional (3D) sound localization offers people a wealth of new technological avenues to not merely communicate with each other but also to communicate with electronic devices, software programs, and processes.

[0002] As this technology develops, challenges will arise with regard to how sound localization integrates into the modern era. Example embodiments offer solutions to some of these challenges and assist in providing technological advancements in methods and apparatus using 3D sound localization.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] FIG. **1** is a method that determines how to display a graphical representation to a listener based on whether sound associated with the graphical representation will localize as binaural sound or one of mono sound and stereo sound in accordance with an example embodiment.

[0004] FIG. **2** is a method to alter a graphical representation being displayed to a listener based on whether the listener will hear sound with headphones or earphones as binaural sound in accordance with an example embodiment.

[0005] FIG. **3** is a method to alter an appearance of a graphical representation being displayed to a listener based on whether sound associated with the graphical representation will play to the listener as binaural sound or one of mono sound and stereo sound in accordance with an example embodiment.

[0006] FIG. **4** is a method to convolve sound into binaural sound when the listener will hear the sound as binaural sound in accordance with an example embodiment.

[0007] FIG. **5**A is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0008] FIG. **5**B is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0009] FIG. **5**C is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0010] FIG. **5**D is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0011] FIG. **5**E is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0012] FIG. **5**F is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0013] FIG. **5**G is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0014] FIG. **5**H is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0015] FIG. **5**I is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0016] FIG. **5**J is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0017] FIG. **5**K is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0018] FIG. **5**L is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0019] FIG. **5**M is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0020] FIG. **5**N is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0021] FIG. **5**O is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0022] FIG. **5**P is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0023] FIG. **5**Q is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0024] FIG. **5**R is a graphical representation with an indication of binaural sound in accordance with an example embodiment.

[0025] FIG. **6** is an example computer system in accordance with an example embodiment.

[0026] FIG. **7** is an example of an electronic device in accordance with an example embodiment.

SUMMARY

[0027] One example embodiment is a method or apparatus that includes graphical representations that when activated play sound in binaural sound to a listener. When displayed to the listener, the graphical representations show an indication that the sound will externally localize as binaural sound to the listener or localize as mono or stereo sound to the listener.

[0028] Example embodiment include altering the graphical representation being displayed based on whether the listener will hear the binaural sound as binaural sound or one of mono sound or stereo sound.

[0029] Other example embodiments are discussed herein.

DETAILED DESCRIPTION

[0030] Binaural sound or three-dimensional (3D) sound externally localizes away from a head of the listener, unlike stereo or mono sound that localizes inside the head of the listener or localizes to a physical speaker. Thus, when a listener hears binaural sound, a source or location of the sound occurs outside the head of the listener even though this location may be in empty space or space not occupied with a sound speaker.

[0031] Binaural sound has many technical challenges and problems, especially when users exchange or play binaural sound during an electronic communication. Example embodiments offer solutions to these challenges and problems.

[0032] One problem with binaural sound is the listener during an electronic communication with another person may not know that the sound being provided from the other

person is binaural sound as opposed to being stereo sound or mono sound. If the listener is not wearing headphones or earphones, then the binaural sound will not externally localize to the listener and hence the effect of binaural sound is lost. In order to externally localize binaural sound, the listener must wear headphones, earphones, or the like since binaural sound requires two speakers placed at the ears of the listener. Alternatively, the binaural sound must be provided to the listener through multiple speakers that provide a sweet spot where the binaural sound can be heard.

[0033] Consider an example in which two users engage in an electronic communication and exchange sound clips in binaural sound, such as voice messages, emojis that play binaural sound when activated, voices in binaural sound, etc. For example, a first user sends a second user an emoji that when activated plays a voice in binaural sound to the second user. When the second user receives the emoji, he or she is not aware that the voice is in binaural sound and hence does not wear headphones or earphones. When the second user activates the emoji and hears the voice, the voice does not externally localize and the effects of binaural sound are lost.

[0034] Example embodiments solve this problem and others. By way of example, an example embodiment provides a graphical representation with an indication that the sound the listener will hear is in binaural sound. This indication will inform the listener to wear headphones or earphones before listening to the sound from the graphical representation.

[0035] Another problem is that users may take off or put on headphones or earphones during an electronic communication, and this action will determine how the users hear the sound. The electronic device providing the binaural sound may also lose the ability to provide the binaural sound to the listener. Graphical representations sent during the electronic communication in these instances, however, will not accurately show how the sound will play to the user.

[0036] Consider an example in which two users engage in an electronic communication and exchange graphical representations that when activated play binaural sound. For example, the first user sends the second user an animated, talking, three-dimensional (3D) emoji that speaks to the second user with a voice that is supposed to localize as binaural sound outside of the head of the second user. The second user, however, is not wearing headphones or earphones, activates the emoji and hears the sound as mono sound that plays through a speaker in a smartphone. Since the sound does not externally localize, the second user is unimpressed with the sound and believes it did not work correctly. The second user is unaware that the sound did not localize because the second user failed to wear headphones or earphones.

[0037] Other problems exist because a first user or first electronic device sending sound to a second electronic device of a second user does not know whether the sound will play to the second user as mono sound, stereo sound, or binaural sound.

[0038] This situation becomes problematic when a graphical representation accompanies the sound. The graphical representation cannot accurately show the format of the sound (e.g., mono, stereo, or binaural) or instructions for the sound when the first user or first electronic device does not know how the sound will be played to the second user. For example, the second user may not have or be wearing headphones and hence unable to hear externalization of binaural sound. As another example, HRTFs of the second user may not be available, and hence the sound cannot be convolved into binaural sound. As another example, the electronic device of the second user may have a default setting in which all sounds are played as mono sound or stereo sound. Playing of binaural sound is set to "off." Alternatively, this default setting may be set to play all sounds (or certain sounds) in binaural sound. As yet another example, even if the first user believes or knows the sound will externally localize as binaural sound to the second user, the first user may not know a location of the SLP to the second user.

[0039] Example embodiments solve these problems and others. An example embodiment changes or alters the graphical representation being displayed to the listener based on whether the sound will externally localize as binaural sound or localize as stereo or mono sound. An appearance of the graphical representation changes in real-time to correspond with whether the sound will externally localize as binaural sound or localize as stereo or mono sound. In this way, the graphical representations shows an accurate, real-time representation as to how sound will localize to the listener. For example, the graphical icon is updated or changed to show how sound will play to the listener even if the listener takes an action changing how the sound plays (e.g., the listener disconnects the headphones, convolution of the sound is no longer possible, the listener removes the headphones, etc.).

[0040] FIG. **1** is a method that determines how to display a graphical representation to a listener based on whether sound associated with the graphical representation will localize as binaural sound or one of mono sound and stereo sound in accordance with an example embodiment.

[0041] Block **100** states receive, during an electronic communication between a first user with a first electronic device and a second user with a second electronic device and at the second electronic device, a graphical representation that plays sound to the second user.

[0042] For example, two or more users talk and/or message each other over one or more networks, such as the Internet, cellular network, etc. The users transmit and/or exchange graphical representations to each other during the electronic communication. For instance, the first user sends the second user the graphical represent, and/or the second user sends the first user a graphical representation.

[0043] The graphical representation can include or be associated with sound. For example, sound plays to the user when the user, an electronic device, or a software program activates the graphical representation or the sound associated with the graphical representation.

[0044] Consider an example in which two users execute a mobile messaging software application. The first user sends the second user an animated emoji (or animoji) that when activated or executed says "Hello" to the second user.

[0045] Block **110** makes a determination as to whether the sound will play to the second user as binaural sound (BS) or one of mono sound and stereo sound. If the answer to this determination is "no" then flow proceeds to block **120**, and if the answer to this determination is "yes" then flow proceeds to block **130**.

[0046] By way of example, sound can localize to the listener as mono sound or stereo sound when the sound is not convolved and played to the listener with headphones, earphones, etc. Mono sound and stereo sound can also

externally localize to speakers, such as speakers in a smartphone, stereo speakers in a room, etc. Alternatively, sound externally localizes to the listener when the sound is convolved into or captured as binaural sound or 3D sound. Binaural sound externally localizes outside or away from the head of the listener and is not required to localize to a physical object, such as a speaker. For instance, binaural sound can externally localize one or more meters away from a person at a location in empty space (e.g., where no speaker exists or no physical or tangible object exists). Binaural sound can also localize to physical objects that do not have an electronic speaker, such as localizing to a wall or a chair. Sound can also localize as a mix of binaural, mono, and stereo sounds, such as sound commencing as binaural sound then transitioning to stereo sound to the listener.

[0047] The determination of how sound will localize to the listener depends on one or more of several factors. Examples of these factors include, but are not limited to, whether the listener is wearing headphone, earphones, or another electronic device that provides binaural sound, whether the listener is hearing the sound from multiple speakers that provide binaural sound (such as two speakers with cross-talk cancellation), whether the sound was or can be convolved with head-related transfer functions (HRTFs) or other transfer functions or impulse responses, whether the listener authorized binaural sound to play to the listener, and whether the listener is located at a geographical location (indoor or outdoor) that allows binaural sound.

[0048] For example, if the listener is not wearing headphones, earphones, or another electronic device that provides binaural sound, then the sound will localize as mono sound or stereo sound. Alternatively, the listener may be wearing such an electronic device but this electronic device is off, muted, or not in communication with the electronic device providing the sound.

[0049] As another example, if the sound is mono sound or stereo sound and not subsequently convolved with HRTFs or other sound localization information (SLI), then the sound will not externally localize as binaural sound. For instance, a user receives a graphical representation and sound recorded in mono or stereo sound. The sound is intended to play to the user as binaural sound, but the electronic device providing the sound does not have HRTFs of the user in order to convolve the sound into binaural sound.

[0050] As another example, a user receives a graphical representation and sound captured or recorded in binaural sound. The user is wearing headphones and can hear the sound as binaural sound. The electronic device of the user, however, is set to only play sound as mono or stereo sound. For instance, a permission for binaural sound is set to off, or the user has not authorized binaural sound to play with the current executing software application.

[0051] As another example, a user is at a location that restricts or prohibits the use of binaural sound. For instance, the user is located in a school, library, government building, or airplane that prohibits listeners from communicating with binaural sound.

[0052] Block **120** states display the graphical representation to the second user with (1) no indication as to how sound will localize or (2) an indication that the sound will localize as one of mono sound and stereo sound.

[0053] By way of example, the graphical representation displays to the user with no indication as to how or where sound will localize. For instance, graphical representations display with a default understanding that sounds will be provided as mono sound or stereo sound. In this instance, an additional indication is not required since the listener knows that unless otherwise instructed or informed, the sound associated with the graphical representation will play as mono sound, stereo sound, or sound that does not externally localize as binaural sound.

[0054] By way of example, the graphical representation displays to the user with an indication as to how or where sound will localize. For instance, graphical representations display with an indication that sound the listener is hearing or is going to hear will be mono sound or stereo sound. This indication can include visual or audio information that informs or instructs the user that the sound will localize as one of mono sound and stereo sound.

[0055] Consider an example in which indication visually informs the user of how the sound will localize or be provided using one or more of text, color, brightness, shading, flashing or dimming of light, words, symbols, drawings, characters, photos, images, or other indicia or other examples discussed herein. For instance, the indication includes the word "stereo" or "mono" a symbol that the user recognizes or associates with stereo sound or mono sound.

[0056] Block **130** states display the graphical representation to the second user with an indication that the sound will localize as binaural sound.

[0057] By way of example, the graphical representation displays to the user with an indication as to how or where sound will localize as binaural sound. For instance, graphical representations display with an indication that sound the listener is hearing or is going to hear will be binaural sound. This indication can include visual or audio information that informs or instructs the user that the sound will localize as binaural sound.

[0058] Consider an example in which the indication visually informs the user of how the sound will localize or be provided using one or more of text, color, brightness, shading, flashing or dimming of light, words, symbols, drawings, characters, photos, images, or other indicia or examples discussed herein. For instance, the indication includes the word "3D" or a symbol that the user recognizes or associates with binaural sound.

[0059] In an example embodiment, the indication forms part of the graphical representation. For example, the indication is inside, on, or touching the graphical representation. Alternatively, the indication is separate or away from the graphical representation. For example, the indication simultaneously displays as a symbol or word separated from the graphical representation.

[0060] FIG. **2** is a method to alter a graphical representation being displayed to a listener based on whether the listener will hear sound with headphones or earphones as binaural sound in accordance with an example embodiment.

[0061] Block **200** states display the graphical representation to the second user with (1) no indication as to how sound will localize or (2) an indication that the sound will localize as one of mono sound and stereo sound.

[0062] Examples of block **200** are discussed in connection with block **120** and elsewhere herein.

[0063] Block **210** makes a determination as to whether headphones, earphones, or another electronic device capable of providing binaural sound to the listener are connected and/or available.

[0064] For example, this determination includes determining one or more of whether the electronic device (headphones, earphones, etc.) is powered on, whether the electronic device or sound to it is paused or muted, whether the electronic device is connected to and/or in electronic communication with the electronic device providing the sound, and whether the listener is wearing the electronic device.

[0065] By way of further example, this determination includes determining whether the electronic device (headphones, earphones, etc.) is connected to the electronic device providing the sound. For instance, determine whether the headphones or earphones are plugged into an audio port of a laptop, desktop, smartphone, or another electronic device. For instance, determine whether the headphones or earphones are in wireless communication with the electronic device providing the sound. This communication can include connecting, handshaking, or pairing through radio signals or radio frequency, infrared signals, or other wireless connectivity (e.g., electronic devices exchanging data via Bluetooth). As yet another example, the determination includes whether the headphones or earphones are connected to the internet, cellular network, or other wireless network providing the sound.

[0066] Consider an example in which the headphones form part of or are used in conjunction with a head mounted display (HMD). The HMD or headphones include a sensor that senses when the HMD is mounted to the head of the listener.

[0067] As another example, a HMD, headphones, or another wearable electronic device include hardware that measures head orientation, head movement, or head tracking (e.g., one or more of an accelerometer, gyroscope, magnetometer, or inertial sensor). This hardware monitors or tracks movement to determine when the electronic device is moving and worn by the listener.

[0068] As another example, headphones, earphones, or other electronic device may be off, muted, or busy and hence not available to provide binaural sound to the listener. For instance, a DSP in the electronic device is not available to convolve the sound with HRTFs into binaural sound when the sound is requested so the sound is provided to the listener as stereo or mono sound. As another example, the electronic device may not currently have access to a network location where the HRTFs are located or other data necessary to convolve the sound into binaural sound. For instance, wireless internet access is temporarily unavailable.

[0069] Consider an example in which the electronic device is or includes a wearable electronic device that tracks eye movement or head movement to determine whether the listener is currently wearing the electronic device. The electronic device may be powered on but binaural sound is not available to the listener if the listener is not wearing the electronic device.

[0070] Consider an example in which the headphones, earphones, or another electronic device wirelessly communicate with the electronic device transmitting or providing the sound. This electronic device providing the sound does not receive a signal and hence knows the headphones or earphones are out of range, not powered on, or otherwise unavailable.

[0071] If the answer to the determination in block **210** is "no" then flow proceeds to block **200**.

[0072] If the answer to the determination in block **210** is "yes" then flow proceeds to block **220**.

[0073] Block **220** states alter the graphical representation to include an indication that the sound will play as binaural sound.

[0074] The graphical representation is provided, changed, or altered to include an indication that the sound will externally localize to the listener as binaural sound.

[0075] For example, an example embodiment updates the graphical representation to include the indication. For instance, upon detecting that the headphones, earphones, or other electronic device are available to provide binaural sound to the listener, the indication or the graphical representation with the indication displays on an electronic device to the listener. In this way, the listener can visually see an indication showing sound is or will be provided as binaural sound.

[0076] Consider an example in which an electronic device stores in memory two or more versions of graphical representations. One version shows no indication how sound will localize, and this version has a default understanding to the listener that sounds will be provided as mono or stereo sound. Another version shows a visual indication that sound will be provided as binaural sound or 3D sound.

[0077] Consider an example during an electronic communication in which an electronic device of first user sends an electronic device of a second user a graphical representation shaped as the word "WOW." The second user wears headphones, and settings of the second electronic device are set to enable and provide binaural sound to the second user. When the second electronic device receives the graphical representation, the second electronic device automatically adds and displays a symbol or text to the "WOW" that the represents to the second user that the sound will play as binaural sound.

[0078] Consider the example in which the first user sends the second user the graphical representation shaped as the word "WOW." While the second user wears wireless earphones, a symbol of "3D" displays immediately below the word "WOW" so the second user visually knows the sound will be in 3D. When the second user removes his or her earphones, the symbol of "3D" immediately disappears from the display since the sound can no longer externally localize to the second user as binaural sound without the earphones. When the second user reconnects or puts the earphones back on, the symbol of "3D" immediately reappears back onto the display.

[0079] Consider an example during an electronic communication in which a first user sends a second user a talking 3D animated, moving emoji that says "How are you?". This emoji appears on the display of the smartphone of the second user and flashes which indicates to the second user that the emoji has a sound clip or voice message waiting to be heard. The letters "3D" also flash with the emoji and indicate that the sound is available as 3D sound. The second user dons headphones, activates the emoji, and hears the recorded audio saying "How are you?" that externally localizes as binaural sound in empty space that is one meter away from the head of the second user.

[0080] Examples for block **220** are also discussed in connection with block **130** and elsewhere herein.

[0081] FIG. **3** is a method to alter an appearance of a graphical representation being displayed to a listener based on whether sound associated with the graphical representa-

tion will play to the listener as binaural sound or one of mono sound and stereo sound in accordance with an example embodiment.

[0082] Block **300** states receive, during an electronic communication between a first user with a first electronic device and a second user with a second electronic device and at the second electronic device, a graphical representation that plays sound to the second user.

[0083] For example, the electronic device of the first user transmits the graphical representation and/or sound to the electronic device of the second user during the electronic communication between the two users.

[0084] Block **310** makes a determination as to whether the sound will play to the second user as binaural sound or one of mono sound and stereo sound.

[0085] This determination includes example embodiments discussed in connection with blocks **110** and **210**.

[0086] If the sound will play as mono sound or stereo sound, then flow proceeds to block **320**. This block states display the graphical representation to the second user with (1) no indication as to how sound will localize or (2) an indication that the sound will localize as one of mono sound and stereo sound.

[0087] Examples of block **320** are discussed in connection with blocks **120** and **200** and elsewhere herein.

[0088] Block **330** states play the sound to the second user as the mono sound or the stereo sound.

[0089] For example, one or more speakers in an electronic device play the sound to the second user as mono sound or stereo sound. For instance, the sound emanates from a speaker in a smartphone and into a room where the listener is located.

[0090] If the sound will play as binaural sound then flow proceeds to block **340**. This block states alter the graphical representation to include an indication that the sound will play as binaural sound.

[0091] Examples of block **340** are discussed in connection with block **220** and elsewhere herein.

[0092] Block **350** states play the sound to the second user as the binaural sound.

[0093] For example, speakers located at, in, or near the ears of the second user play the sound to the second user as the binaural sound.

[0094] Consider an example in which one or more electronic devices alter, change, update, or provide an appearance of a graphical representation based on how the sound is or will be provided to the listener or how the listener is or will localize the sound. For example, two or more user exchange or transmit graphical representations to electronic devices of each other. These graphical representations play sound to the users.

[0095] In this example, a first electronic device of a first user transmits the graphical representation and/or the sound to a second electronic device of the second user.

[0096] The first and second electronic devices have, for example, wireless transmitter/receivers that enable them to wireless transmit the graphical representations and/or sound.

[0097] The second electronic device determines whether the sound can be provided to the second user as binaural sound. For example, the second electronic device determines one or more of (1) whether the headphones or earphones are powered on, (2) whether the headphones or earphones are in communication with the first electronic device providing the graphical representation and/or sound, (3) whether the second user is wearing the headphones or earphones, (4) whether the sound is binaural sound (e.g., captured as binaural sound or already convolved into binaural sound), and (5) whether HRTFs to convolve the sound are available (e.g., stored in memory or otherwise retrievable).

[0098] When the sound cannot be provided as binaural sound, the second electronic device displays the graphical representation with one or more of no indication of how the sound will localize, with an indication that the sound will localize as mono or stereo sound, with an indication that the sound will localize to an external physical speaker, or with an indication that the sound will localize inside the head of the second user.

[0099] When the sound can be or will be provided as binaural sound, the second electronic device displays the graphical representation with one or more of an indication of how the sound will localize, with an indication that the sound will localize as binaural sound or 3D sound, or with an indication that the sound will localize outside the head of the second user (e.g., to a location in empty space or a location with no speaker).

[0100] The second electronic device displays the graphical representation and/or indication to visually inform the second user how the sound is being provided or will be provided. This information enables the second user to quickly determine if the sound will be heard in a conventional manner (e.g., as mono or stereo) or in an unconventional manner in binaural sound. In this way, the second user will not be startled or surprised to hear binaural sound. Additionally, the second user will be informed to wear headphones or earphones before listening to the sound accompanying the graphical representation so the effects of external localization of binaural sound are not lost.

[0101] An example embodiment alters or changes the graphical representation, indication, and/or information being displayed to the second user based on the determination of the how the sound is being provided or will be provided to the second user. As one example, an appearance of the graphical representation changes to show the second user that the sound is binaural sound. As another example, the graphical representation is unchanged but the indication is added to show the second user that the sound is binaural sound. As another example, the indication is added to the graphical representation to show the second user that the sound is binaural sound. As another example, a new or different graphical representation is displayed to the second user to show the second user that the sound is binaural sound. For instance, one graphical representation is swapped or changed with another, different graphical representation. As another example, a second or additional graphical representation is added to the display to show the second user that the sound is binaural sound.

[0102] In some instances, the ability to provide binaural sound may change during the electronic communication. In this instance, graphical representation and/or indication alters or changes to reflect this change in real-time.

[0103] Consider an example in which sound is being provided or will be provided to the listener as binaural sound, and the graphical representation and/or indication visually shows this information to the listener. Thereafter, the binaural sound is no longer being provided to the listener or capable of being provided to the listener (e.g., the listener removes the headphones or the earphones, mutes sound to them, disconnects them, or powers them off). This determi-

nation causes a change in how the graphical representation and/or indication is being displayed to the listener. For example, the electronic device removes the indication showing sound as being binaural sound. Removal of the indication and/or change to the graphical representation occurs in real-time in response to the change in how the sound is being provided or will be provided to the listener.

[0104] An example embodiment expedites processing of the electronic communication and saves processing resources. For example, a digital signal processor (DSP) in the electronic device of the user convolves the sound with head-related transfer functions (HRTFs) to change the sound into binaural sound when a determination is made that the sound can be provided to the listener in binaural sound. For instance, in response to determining that the listener is wearing headphones or earphones and hence can hear binaural sound, the electronic device processes the sound with the HRTFs. By contrast, processing resources are saved when the electronic device determines that the listener cannot hear binaural sound. For instance, the electronic device does not convolve the sound in response to determining that the second user is not wearing the headphones or the earphones. In this way, the sound is not unnecessarily processed into binaural sound before a determination is made as to whether the listener can actually hear binaural sound.

[0105] Consider an example in which a first user sends a graphical representation and sound to an electronic device of a second user. The electronic device of the second user determines that binaural sound cannot be provided to the second user (e.g., the second user is not wearing headphones or earphones or HRTFs are not available). At this point in time, the sound is not convolved, which saves processing resources. Further, the graphical representation is displayed as received, and the sound plays to the second user as mono or stereo sound. Subsequently, the second electronic device determines that binaural sound can play to the second user. At this point in time, the electronic device changes or alters the graphical representation or the information being displayed to the second user to include an indication that sound will play to the second user as binaural sound. The electronic device convolves the sound with HRTFs, changes the sound into binaural sound, and plays the binaural sound to the second user upon activation of the graphical representation.

[0106] As discussed herein, example embodiments include various ways to alter the graphical representation to notify the listener that the sound will be binaural sound. An appearance of the graphical representation and/or indication being displayed to provide the indication that visually informs the listener that the sound will externally localize as the binaural sound away from the head of the listener are changed or altered by one or more of the following: (1) adding a symbol "3D" to the display, to the graphical representation and/or to the indication, (2) adding one or more words that when read by the listener provide the indication that visually informs the listener that the sound will externally localize as the binaural sound away from the head of the listener, (3) flashing the graphical representation and/or the indication, (4) changing a brightness of the graphical representation and/or the indication, (5) turning on and off the graphical representation and/or indication, (6) changing a shape or size of the graphical representation and/or indication (7) changing the graphical representation and/or indication from 2D to 3D or from 3D to 2D, and (8) executing other examples are discussed herein.

[0107] The electronic device also removes the indication when the determination is made that the sound is no longer or can no longer be provided as binaural sound. For example, the electronic device removes the "3D" (or another graphical representation and/or indication discussed herein) being displayed in response to determining that the headphones or the earphones of the listener are not connected to the second electronic device, not powered on, not connected to the network, not being worn by the listener, muted or silenced, etc.

[0108] FIG. **4** is a method to convolve sound into binaural sound when the listener will hear the sound as binaural sound in accordance with an example embodiment.

[0109] Block **400** states receives a graphical representation and/or sound to play to the listener.

[0110] For example, an electronic device receives the graphical representation and/or sound during an electronic communication. The electronic device receives the sound from memory, from another electronic device, from a network, from a server, as streaming audio, etc.

[0111] Block **410** makes a determination as to whether the sound will play to the second user as binaural sound.

[0112] This determination includes example embodiments discussed including those discussed in connection with blocks **110**, **210**, and **310**.

[0113] If the answer to the determination in block **410** is "no" then flow proceeds to block **420** that states play the sound to the listener in mono sound or stereo sound.

[0114] For example, the sound plays through headphones or earphones as mono sound or stereo sound. As another example, the sound plays through one or more speakers in a wearable electronic device (WED) or a handheld portable electronic device (HPED) as mono sound or stereo sound. For instance, the sound emanated from a speaker in a smartphone or speaker in a laptop.

[0115] If the answer to the determination in block **410** is "yes" then flow proceeds to block **430** that states process or convolve the sound with head-related transfer functions (HRTFs) so the sound will externally localize as binaural sound to the listener.

[0116] For example, a processor (such as a DSP) processes or convolves the sound with one or more of head-related transfer functions (HRTFs), head-related impulse responses (HRIRs), room impulse responses (RIRs), room transfer functions (RTFs), binaural room impulse responses (BRIRs), binaural room transfer functions (BRTFS), interaural time delays (ITDs), interaural level differences (ITDs), and a sound impulse response.

[0117] One example embodiment processes or convolves the sound with the HRTFs before the electronic communication commences or before a point in time when the sound is requested for play during the electronic communication. For example, process the sound before the graphical representation is activated. This expedites playing of the sound to the listener since the listener does not have to wait while a processor processes or convolves the sound into binaural sound. Furthermore, the processor (such as the DSP) can be devoted to other tasks instead of convolving the sound into binaural sound.

[0118] Another example embodiment processes or convolves the sound during the electronic communication or at a point in time when the sound is requested for play in the

electronic communication. For example, the DSP convolves the sound when the graphical representation activates.

[0119] Sound includes, but is not limited to, one or more of stereo sound, mono sound, binaural sound, computer-generated sound, sound captured with microphones, and other sound. Furthermore, sound includes different types including, but not limited to, music, background sound or background noise, human voice, computer-generated voice, and other naturally occurring or computer-generated sound.

[0120] When the sound is recorded or generated in mono sound or stereo sound, convolution changes the sound to binaural sound. For example, one or more microphones record a human person speaking in mono sound or stereo sound, and a processor processes this sound with filters to change the sound into binaural sound.

[0121] The processor or sound hardware processing or convolving the sound can be located in one or more electronic devices or computers including, but not limited to, headphones, smartphones, tablet computers, electronic speakers, head mounted displays (HMDs), optical head mounted displays (OHMDs), electronic glasses (e.g., glasses that provide augmented reality (AR)), servers, portable electronic devices (PEDs), handheld portable electronic devices (HPEDs), wearable electronic devices (WEDs), and other portable and non-portable electronic devices. These electronic devices can also be used to execute example embodiments.

[0122] In one example embodiment, the DSP is located in the electronic device of the second user. In other example embodiments, the DSP is located in other electronic devices, such as a server or in the first electronic device of the first user.

[0123] The DSP processes or convolves stereo sound or mono sound with a process known as binaural synthesis or binaural processing to provide the sound with sound localization cues (ILD, ITD, and/or HRTFs) so the listener externally localizes the sound as binaural sound or 3D sound.

[0124] HRTFs can be obtained from actual measurements (e.g., measuring HRIRs and/or BRIRs on a dummy head or human head) or from computational modeling. HRTFs can also be general HRTFs (also known as generic HRTFs) or customized HRTFs (also known as individualized HRTFs). Customized HRTFs are specific to an anatomy of a particular listener. Each person has unique sets or pairs of customized HRTFs based on the shape of the ears or pinnae, head, and torso.

[0125] An example embodiment models the HRTFs with one or more filters, such as a digital filter, a finite impulse response (FIR) filter, an infinite impulse response (IIR) filter, etc. Further, an ITD can be modeled as a separate delay line.

[0126] When the binaural sound is not captured (e.g., on a dummy head or human head), the captured sound is convolved with sound localization information (SLI). This information includes one or more of HRTFs, HRIRs, BRTFs, BRIRs, ILDs, ITDs, and/or other information discussed herein. By way of example, SLI are retrieved, obtained, or received from memory, a database, a file, an electronic device (such as a server, cloud-based storage, or another electronic device in the computer system or in communication with a PED providing the sound to the user through one or more networks), etc. Instead of being retrieved from memory, this information can also be calculated in real-time.

[0127] A central processing unit (CPU), processor (such as a DSP), or microprocessor processes and/or convolves the sound with the SLI, such as a pair of head related transfer functions (HRTFs), ITDs, and/or ILDs so that the sound will localize to a zone, area, or sound localization point (SLP). For example, the sound localizes to a specific point (e.g., localizing to point (r, θ, ϕ)) or a general location or area (e.g., localizing to far-field location (θ, ϕ) or near-field location (θ, ϕ)). As an example, a lookup table that stores a set of HRTF pairs includes a field/column that specifies the coordinates associated with each pair, and the coordinates indicate the location for the origination of the sound. These coordinates include a distance (r) or near-field or far-field designation, an azimuth angle (θ), and/or an elevation angle (ϕ).

[0128] The complex and unique shape of the human pinnae transforms sound waves through spectral modifications as the sound waves enter the ear. These spectral modifications are a function of the position of the source of sound with respect to the ears along with the physical shape of the pinnae that together cause a unique set of modifications to the sound called head related transfer functions or HRTFs. A unique pair of HRTFs (one for the left ear and one for the right ear) can be modeled or measured for each position of the source of sound with respect to a listener as the customized HRTFs.

[0129] A HRTF is a function of frequency (f) and three spatial variables, by way of example (r, θ, ϕ) in a spherical coordinate system. Here, r is the radial distance from a recording point where the sound is recorded or a distance from a listening point where the sound is heard to an origination or generation point of the sound; θ (theta) is the azimuth angle between a forward-facing user at the recording or listening point and the direction of the origination or generation point of the sound relative to the user; and ϕ (phi) is the polar angle, elevation, or elevation angle between a forward-facing user at the recording or listening point and the direction of the origination or generation point of the sound relative to the user. By way of example, the value of (r) can be a distance (such as a numeric value) from an origin of sound to a recording point (e.g., when the sound is recorded with microphones) or a distance from a SLP to a head of a listener (e.g., when the sound is generated with a computer program or otherwise provided to a listener).

[0130] When the distance (r) is greater than or equal to about one meter (1 m) as measured from the capture point (e.g., the head of the person) to the origination point of a sound, the sound attenuates inversely with the distance. One meter or thereabout defines a practical boundary between near-field and far-field distances and corresponding HRTFs. A "near-field" distance is one measured at about one meter or less; whereas a "far-field" distance is one measured at about one meter or more. Example embodiments are implemented with near-field and far-field distances.

[0131] The coordinates for external sound localization can be calculated or estimated from an interaural time difference (ITD) of the sound between two ears. ITD is related to the azimuth angle according to, for example, the Woodworth model that provides a frequency independent ray tracing methodology. The coordinates (r, θ, ϕ) for external sound localization can also be calculated from a measurement of an orientation of and a distance to the face of the person when a head related impulse response (HRIR) is captured.

[0132] The coordinates can also be calculated or extracted from one or more HRTF data files, for example by parsing

known HRTF file formats, and/or HRTF file information. For example, HRTF data is stored as a set of angles that are provided in a file or header of a file (or in another predetermined or known location of a file or computer readable medium). The data can include one or more of time domain impulse responses (FIR filter coefficients), filter feedback coefficients, and an ITD value. This information can also be referred to as "a" and "b" coefficients. By way of example, these coefficients are stored or ordered according to lowest azimuth to highest azimuth for different elevation angles. The HRTF file can also include other information, such as the sampling rate, the number of elevation angles, the number of HRTFs stored, ITDs, a list of the elevation and azimuth angles, a unique identification for the HRTF pair, and other information. The data can be arranged according to one or more standard or proprietary file formats, such as AES69, and extracted from the file.

[0133] The coordinates and other HRTF information are calculated or extracted from the HRTF data files. A unique set of HRTF information (including r, θ, ϕ) is determined for each unique HRTF.

[0134] The coordinates and other HRTF information are also stored in and retrieved from memory, such as storing the information in a look-up table. The information is quickly retrieved to enable real-time processing and convolving of sound using HRTFs and hence improves computer performance of execution of binaural sound.

[0135] The SLP represents a location where a person will perceive an origin of the sound. For an external localization, the SLP is away from the person (e.g., the SLP is away from but proximate to the person or away from but not proximate to the person). The SLP can also be located inside the head of the person (e.g., when the sound is provided as mono sound or stereo sound). Sound can also switch between externally localizing and internally localizing, such as appearing to move and pass through a head of a listener.

[0136] SLI can also be approximated or interpolated based on known data or known SLI, such as SLI for other coordinate locations. For example, a SLP is desired to localize at coordinate location (2.0 m, 0°, 40°), but HRTFs for the location are not known. HRTFs are known for two neighboring locations, such as known for (2.0 m, 0°, 35°) and (2.0 m, 0°, 45°), and the HRTFs for the desired location of (2.0 m, 0°, 40°) are approximated from the two known locations. These approximated HRTFs are provided to convolve sound to localize at the desired coordinate location (2.0 m, 0°, 40°).

[0137] Sound is convolved either directly in the time domain with a finite impulse response (FIR) filter or with a Fast Fourier Transform (FFT). For example, an electronic device convolves the sound to one or more SLPs using a set of HRTFs, HRIRs, BRIRs, or RIRs and provides the person with binaural sound.

[0138] In an example embodiment, convolution involves an audio input signal and one or more impulse responses of a sound originating from various positions with respect to the listener. The input signal is a limited length audio signal (such as a pre-recorded digital audio file or sound clip) or an ongoing audio signal (such as sound from a microphone or streaming audio over the Internet from a continuous source). The impulse responses are a set of HRIRs, BRIRs, RIRs, etc.

[0139] Convolution applies one or more FIR filters to the input signals and convolves the input signals into binaural audio output or binaural stereo tracks. For example, the input signals are convolved into binaural audio output that is specific or individualized for the listener based on one or more of the impulse responses to the listener.

[0140] The FIR filters are derived binaural impulse responses. Alternatively or additionally, the FIR filters are obtained from another source, such as generated from a computer simulation or estimation, generated from a dummy head, retrieved from storage, computed based on known impulse responses captured from people, etc. Further, convolution of an input signal into binaural output can include sound with one or more of reverberation, single echoes, frequency coloring, and spatial impression.

[0141] Processing of the sound also includes calculating and/or adjusting an interaural time difference (ITD), an interaural level difference (ILD), and/or other aspects of the sound in order to alter the cues and artificially alter the point of localization. Consider an example in which the ITD is calculated for a location (θ, ϕ) with discrete Fourier transforms (DFTs) calculated for the left and right ears. The ITD is located at the point for which the function attains its maximum value, known as the argument of the maximum or arg max as follows:

$$ITD = \operatorname{argmax}(\tau)\sum_{n} d_{l,\theta,\phi}(n)\cdot d_{r,\theta,\phi}(n+\tau).$$

[0142] Subsequent sounds are filtered with the left HRTF, right HRTF, and/or ITD so that the sound localizes at (r, θ, ϕ). Such sounds include filtering stereo and monaural sound to localize at (r, θ, ϕ). For example, given an input signal as a monaural sound signal s(n), this sound is convolved to appear at (θ, ϕ) when the left ear is presented with:

$$s_l(n)=s(n-ITD)\cdot d_{l,\theta,\phi}(n);$$

and the right ear is presented with:

$$s_r(n)=s(n)\cdot d_{r,\theta,\phi}(n).$$

[0143] Consider an example in which a dedicated digital signal processor (DSP) executes frequency domain processing to generate real-time convolution of monophonic sound to binaural sound.

[0144] By way of example, a continuous audio input signal x(t) is convolved with a linear filter of an impulse response h(t) to generate an output signal y(t) as follows:

$$y(\tau) = x(\tau)\cdot h(\tau) = \int_0^{\infty} x(\tau - t)\cdot h(t)\cdot dt.$$

[0145] This reduces to a summation when the impulse response has a given length N and the input signal and the impulse response are sampled at t=iDt as follows:

$$y(i) = \sum_{j=0}^{N-1} x(i-j)\cdot h(j).$$

[0146] Execution time of convolution further reduces with a Fast Fourier Transform (FFT) algorithm and/or Inverse Fast Fourier Transform (IFFT) algorithm.

[0147] Consider another example of binaural synthesis in which recorded or synthesized sound is filtered with a binaural impulse response (e.g., HRIR or BRIR) to generate a binaural output sound to the person. The input sound is preprocessed to generate left and right audio streams that are mapped to one or more sound sources or sound localization points (known as SLPs). These streams are convolved with a binaural impulse response for the left ear and the right ear to generate the left and right binaural output sound signal. The output sound signal is further processed depending on a final destination. For example, a cross-talk cancellation algorithm is applied to the output sound signal when it will be provided through loudspeakers or applying artificial binaural reverberation to provide 3D spatial context to the sound.

[0148] Block **440** states play the binaural sound to the listener so the sound externally localizes outside the head of the listener.

[0149] The sound plays to the listener as binaural sound that externally localizes away from or outside of the head of the listener. For example, headphones or earphones provide this sound at one or more sound localization points (SLPs) discussed herein.

[0150] Consider an example in which a listener receives sound to play, such as a recorded voice message, a sound clip, or streaming audio. The electronic device receiving the sound makes a determination as to whether the sound can play to the listener as binaural sound. For example, the listener can hear binaural sound when the headphones or earphones are powered on, being worn, and receiving the sound. In response to this determination, the electronic device processes or convolves the sound from mono sound or stereo sound into binaural sound and plays the binaural sound to the listener.

[0151] An example embodiment saves processing resources. For example, the electronic device starts and stops convolution of the sound based on the determination of whether the binaural sound can be provided to the listener. For instance, the electronic device continues to convolve the sound as long as the listener hears the sound as binaural sound. The electronic device stops convolving the sound when the listener is no longer able to hear binaural sound. At this time, the listener may still be able to hear the sound as mono sound or stereo sound.

[0152] Consider an example in which the listener listens to streaming audio while wearing a WED (e.g., headphones, earphones, a head mounted display, or electronic glasses that provide augmented reality (AR) images). During the time that the listener wears the WED, a processor convolves the sound into binaural sound and plays the binaural sound to the listener. When the listener removes the WED, the output location of the sound switches to one or more speakers that emit the sound into the room or location of the listener, as opposed to emitting the sound directly into the ears of the listener through headphones or earphones. At this time, the listener is no longer able to externally localize the sound as binaural sound since he or she no longer wears the WED. The sound, however, does not stop or is not interrupted. Instead, the sound continues to play to the listener (e.g., a speaker in the WED, a speaker in a smartphone, speakers on the floor of a room, etc.). When the listener removes the WED, the processor ceases or stops convolving the sound into binaural sound since the listener is no longer able to hear binaural sound at this time. This process save processing resources since, for example, the DSP is no longer required to process sound with HRTFs.

[0153] Consider an example in which the listener wears headphones that communicate with a smartphone that provides audio to the headphones. While the listener wears the headphones, the listener hears the audio as binaural sound. For instance, a DSP in the smartphone processes the sound into binaural sound that plays to the listener. The binaural sound, however, ceases to play to the listener when the listener takes the headphones off, turns them off, or unplugs or physically or wirelessly disconnects them from the smartphone. At this moment in time, the output of the sound switches from being provided through the headphones to being provided through one or more speakers in the smartphone. The audio is not interrupted as the listener continues to hear the sound. In response to this determination, a DSP stops convolving the sound into binaural sound since the listener is not able to hear binaural sound from the small speakers in the smartphone. In this way, the sound is not unnecessarily convolved. When the listener commences to listen to the sound again through the headphones, convolution of the sound resumes. Thus, for example, the action of connecting and disconnecting the headphones, causes convolution to start and to stop in response to the connecting and the disconnecting.

[0154] Graphical representations include, but are not limited to, emoji, emoticons, animoji, icons, stickers, folders, documents, files, text or words, pictures, images, and other visible indicia that display on, thru, or with an electronic device. Furthermore, these graphical representations can be two-dimensional (2D), three-dimensional (3D), virtual reality (VR) images, augmented reality (AR) images, static or non-moving, moving, and other types of images.

[0155] The indication enables a user to see that sound associated with or corresponding to the graphical representation will externally localize as binaural sound or one of mono sound or stereo sound. For example, the user knows in advance that the sound will externally localize as binaural sound as opposed to hearing the sound as stereo sound or mono sound that internally localizes inside a head of the user. Before the user actually hears the sound associated with the graphical representation, the user knows that the sound is binaural sound and intended to externally localize outside of the head of the user.

[0156] The indication can also inform or instruct the user to wear headphones, earbuds, earphones, or another electronic device that provides binaural sound. Binaural sound cannot accurately localize outside a head of the listener unless the listener wears such an electronic device. Since the listener knows in advance that the sound is binaural sound, he or she will don headphones or earphones before listening to the sound if the listener is not already wearing such a device.

[0157] The indication can be part of the graphical representation itself. For example, the indication forms part of the structure or body of the graphical representation. The indication can also attach to the graphical representation or touch, supplement, compliment, or complete the graphical representation. Alternatively, the indication is separate from or apart from the graphical representation. For example, the indication and the graphical representation are two different and/or separate objects that are simultaneously displayed to the user.

[0158] By way example, the indication is or uses color, light, brightness, shading, or another enhancement or indication to visibly show the user that sound associated with the graphical representation will be played in and heard by the listener as binaural sound. Further examples of indications are shown in FIGS. **5**A-**5**R.

[0159] In an example embodiment, a sound file, sound clip, streaming sound, a recording, or other type of sound associates with or corresponds to a graphical representation. Binaural sound plays to the listener when the graphical representation activates.

[0160] In an example embodiment, a user, a listener, a program or software application, or an electronic device activates the graphical representation and/or causes the binaural sound to play to the listener.

[0161] For example, the listener interacts with a user interface and provides a command or instruction to play the sound upon receiving the graphical representation. For instance, the first user performs one or more actions that include, but are not limited to, clicking or activating an icon, emoji, graphical representation, or other indicia that represents a sound clip, sound file, streaming sound, or recording, selecting the sound from a menu (such as a dropdown menu), selecting the sound from a folder or file (such as a folder or file being displayed to the first user), providing a body gesture (such as a hand gesture or hand movement indicating a desire to play the sound), providing head movement or eye movement (such as the listener moving his or her head in a certain direction or pattern to indicate selection of the sound), providing a voice command (such as the listener speaking an instruction at a natural language user interface), or taking another action to have the sound played to the listener.

[0162] As another example, the sound automatically plays. For instance, the sound plays when the listener receives the graphical representation, opens the software program providing the graphical representation, or views the graphical representation on a display.

[0163] As another example, the sound plays when a sender of the sound (e.g., another user in an electronic communication with the listener) activates the sound or designates when the sound plays.

[0164] Binaural sound is provided to the listener through one or more electronic devices including, but not limited to, one or more of headphones, earphones, earbuds, bone conduction devices, or other electronic devices with speakers at, in, or near the ears of the listener. Binaural sound can be processed for crosstalk cancellation and provided through speakers separate or away from the listener (e.g., dipole stereo speakers). Electronic devices in communication with headphones, earphones, and earbuds can provide binaural sound to the listener (e.g., a smartphone in wireless communication with earphones).

[0165] Various types of electronic devices can include or be in communication with speakers to provide binaural sound to listeners. Examples of these electronic devices include, but are not limited to, wearable electronic glasses, smartphones, head mounted displays (HMDs), optical head mounted displays (OHMDs), wearable electronic devices (WEDs), portable electronic devices (PEDs), handheld portable electronic devices (HPEDs), laptop computers, tablet computers, desktop computers, and other electronic devices.

[0166] From the point-of-view of the listener, the sound originates or emanates from an object, point, area, or direction. This location for the origin of the sound is the sound localization point (SLP). By way of example, the SLP can be an actual point in space (e.g., an empty point in space 1-2 meters away from the head of the listener) or a point on or at a physical or virtual object (e.g., a mouth or head of an augmented reality (AR) or virtual reality (VR) image). The SLP does not have to be so precise since humans are not always able to localize sound to a particle point. As such, the SLP can also be a specific or general area (e.g., a location next to and on the right side of the listener) or a specific or general direction from where the sound originates to the listener (e.g., a location several meters behind the listener).

[0167] When binaural sound is provided to the listener, the listener will hear the sound as if it originates from the sound source, the source of sound, or the SLP. The sound, however, does not originate from the sound source since the sound source or SLP may be an inanimate object with no electronics or an animate object with no electronics. Alternatively, the sound source or SLP has electronics but does not have the capability to generate sound (e.g., the sound source has no speakers or sound system). As yet another example, the sound source or SLP has speakers and the ability to provide sound but is not providing sound to the listener. In each of these examples, the listener perceives the sound to originate from the sound source or SLP, but the sound source or SLP does not produce the sound. Instead, the sound is altered or convolved and provided to the listener so the sound appears to originate from the sound source or SLP.

[0168] In an example embodiment, at least a portion of the sound associated with, corresponding to, or provided from the graphical representation externally localizes away from the head of the listener in empty space (e.g., where no physical or tangible object exists) or occupied space. For example, the sound externally localizes proximate or near the listener, such as localizing within a few meters of the listener. For instance, the SLP where the listener localizes the sound is stationary or fixed in space (e.g., fixed in space with respect to the user, fixed in space with respect to an object in a room, fixed in space with respect to an electronic device, fixed in space with respect to another object or person).

[0169] Consider an example in which two users message and/or talk to each other in an electronic communication via a mobile messaging application. A HPED of the first user sends a HPEP of the second user a graphical representation, here a 3D sound emoji or animoji. This emoji appears on the display of the HPED of the second user as an animated or cartoon character. This character wears headphones that flash or change brightness. When the second user sees the headphones, he or she knows the emoji is informing the second user to wear headphones or earphones before listening to the sound associated with the emoji. The second user dons headphones, and the emoji (here the character) speaks. A voice of the emoji externally localizes as binaural sound to the listener at a SLP outside of the head of the second user. Once the user dons the headphones, the headphones disappear since this indication served its function to inform the user that sound would be binaural sound.

[0170] An example embodiment displays the graphical representation with an indication visually informing a listener a location where binaural sound associated with the graphical representation will externally localize to the listener.

[0171] The indication shows the user the location of the sound source or SLP where the binaural sound will originate to the listener. This location can be a physical or virtual object, a point, an area, or a direction. Further, the indication can provide the location with a precise or general direction of the SLP and/or a precise or general distance to the SLP.

[0172] In an example embodiment, the indication provides a precise, exact, or clearly identifiable location where the sound will originate to the listener. For example, the indication displays, provides, or points to a SLP that is a precise point or area that the listener can identify or see. For example, the indication indicates a physical or virtual object where the sound will externally localize. When the listener hears the sound, the sound indeed originates from the physical or virtual object identified by the identification.

[0173] In an example embodiment, the indication provides a general point, general area, or direction where the sound will originate to the listener. For example, the indication provides a heading or direction (e.g., North, South, East, West, etc.). As another example, the indication provides general direction or area (e.g., sound will originate to your left, to your right, in front of you, over there, behind you, etc.).

[0174] The indication can also provide a precise or general distance to the location. For example, upon seeing the indication, the listener knows the SLP will be one meter away or two meters away. As another example, upon seeing the indication, the listener knows the SLP will be near to the listener (e.g., within 3 meters) or far from the listener (e.g., greater than five meters, greater than 10 meters, or farther). Alternatively, the indication identifies whether the SLP will be "near-field" (which is sound that originates from a location that is within one meter from a head of the listener) or "far-field" (which is sound that originates from a location that is greater than one meter from the head of the listener).

[0175] As another example, the listener knows the specific direction from where the sound originates but not the precise distance to the sound. As another example, the listener knows a specific distance but not the location. For instance, the listener hears a voice and knows the voice originates about 2 meters behind the left side of the head of the listener.

[0176] By way of example, the SLP can be an actual point in space (e.g., an empty point in space 1-2 meters away from the head of the listener) or a point on a physical or virtual object (e.g., a mouth or head of an augmented reality (AR) or virtual reality (VR) image). The SLP does not have to be so precise since humans are not always able to localize sound to a particle point. As such, the SLP can also be a general area (e.g., a location next to and on the right side of the listener) or a general direction from where the sound originates to the listener (e.g., a location several meters behind the listener).

[0177] In an example embodiment, the listener hears the sound associated with the graphical representation as binaural sound that externally localizes away from the head of the listener to the location indicated by indication.

[0178] The listener hears the sound at the location that coincides with or matches with the location provided by the indication. In this way, the listener knows in advance where the binaural sound will externally localize to the listener before the listener hears the sound.

[0179] Consider an example in which the graphical representation is an emoji that includes a talking animated animal head or human head. When a listener clicks on or activates the emoji, the head talks and the listener hears the voice as binaural sound that externally localizes about one meter away from the listener. The emoji includes an indication showing that this voice will externally localize in front of and to the right of the head of the listener. For instance, the voice is convolved with head-related transfer functions (HRTFs) having spherical coordinates (distance r=1.0 m, elevation ϕ=0°, azimuth θ=30°). The listener activates the emoji and hears the voice originate from spherical coordinates (1.0, 0°, 30°) which coincides with the location indicated by the indication.

[0180] Consider another example in which the graphical representation is an emoji that displays on a HPED of the listener. The emoji appears in the center of the display of the HPED, and a flashing dot (or other flashing indicia) appears above and to the left of the emoji. This dot indicates the location where the binaural sound will localize with respect to the listener. In this example, the emoji represents or symbolizes the head of the listener, and the dot represents a relative location where the SLP will be. As shown on the display of the HPED, the SLP (shown as the flashing dot) is in front of and to the left of the emoji. As such, the SLP for the sound that the listener will hear will also occur in front of and to the left of the head of the listener.

[0181] An example embodiment displays the graphical representation that informs a listener to wear headphones or earphones before listening to binaural sound in accordance with an example embodiment. The indication visually informs the listener to wear headphones or earphones so the listener knows in advance that sound associated with the graphical representation will externally localize as binaural sound to the listener when the sound plays to the listener. In this way, the effects of externally localization are not lost, and the listener is able to enjoy 3D sound as opposed to hearing the sound as stereo sound or mono sound.

[0182] By way of example, the indication uses one or more of color, light, shading, brightness, or flashing to inform the listener to wear headphones or earphones before listening to the sound. As another example, the indication uses words or text to inform the user. For instance, the indication includes the words "headphones" or an acronym or symbol (e.g., "HP" or "3D") to signify the listener should wear headphones or earphones. As another example, the graphical representation includes a sign, symbol, emblem, artwork, image, or other indicia showing the listener to wear headphones or earphones. For instance, the graphical representation includes a head wearing headphones, about to wear headphones, or putting on headphones. Movements or actions of the graphical representation instruct the listener on what movements or actions the listener should do as well (e.g., movements of the graphical representation donning headphones shows the listener to don headphones).

[0183] During an electronic communication, one or more users exchange graphical representations, indications, and/or sound with each other.

[0184] By way of example, a computer or electronic device generates the sound (computer-generated sound), or microphones capture and record the sound to be sent. For example, one or more microphones capture the sound as mono sound or stereo sound when the first user speaks a message to the second user. As another example, the first electronic device or a server in communication with the first electronic device includes a plurality pre-recorded or previously generated sounds.

[0185] The sound can be stored in memory of an electronic device, obtained from memory of an electronic device (such as a computer or server), and/or transmitted or streamed over one or more networks.

[0186] Consider an example in which the first electronic device executes a mobile software messaging application that includes hundreds or thousands of sound clips or sound files. The first electronic device obtains or has access to these sound clips or sound files and can send them to other users of the mobile software messaging application.

[0187] Consider an example in which the first electronic device obtains the sound when the first user speaks into microphones in the first electronic device or in communication with the first electronic device. The microphones records the voice of the first user as he or she records a message or sound to be played to the second user.

[0188] The first electronic device transmits the sound and a graphical representation associated with or corresponding to the sound over one or more wired or wireless networks (e.g., a cellular network, the internet, etc.). For example, the first electronic device includes a wireless transmitter/receiver that sends the sound and graphical representation.

[0189] Consider an example in which the first user commands or instructs the sound clip to play to the second user during an electronic communication between the first and second users. In response to this command or instruction, the first electronic device transmits the sound clip and a 3D moving emoji to the second electronic device.

[0190] In another example embodiment, a server or another electronic device transmits the sound and/or graphical representation to the second electronic device. Consider an example in which the first and second users talk or message each other with a mobile messaging software application. The application executes on the electronic devices and one or more servers. When the first user clicks on a 3D sound emoji, this action causes one of the servers to transmit the 3D emoji and sound to the second electronic device.

[0191] The second electronic device receives the sound and the graphical representation from the first electronic device or another electronic device (e.g., a server) in communication with the first electronic. For example, the second electronic device includes a wireless transmitter/receiver that receives the sound and graphical representation over one or more networks.

[0192] A processor or sound hardware processes or convolves the sound with head-related transfer functions (HRTFs) or other SLI so the sound will externally localize as binaural sound to the listener.

[0193] In an example embodiment, the electronic device display the graphical representation with one or more of the following: (1) an indication that the sound will externally localize as binaural sound to the second user, (2) an indication of a location where the sound will externally localize as binaural sound to the second user, and (3) an indication informing the second user to wear headphones or earphones before listening to the sound.

[0194] The graphical representation displays one or more indications discussed herein in accordance with example embodiments. These indications include those discussed in connection with FIGS. **5**A-**5**R.

[0195] FIGS. **5**A-**5**R show a plurality of graphical representations with one or more indications in accordance with example embodiments. By way of example, the graphical representations **500**A-**500**R are shown as faces, such as a face of an emoji, emoticon, etc. Such faces can have many shapes and forms, such as human faces, cartoon character faces, animal faces, animated faces, etc.

[0196] Example embodiments are not limited to graphical representations that include faces, such as those shown in FIGS. **5**A-**5**R. Graphical representations can have many sizes, shapes, and forms (e.g., people, faces, characters, animals, and objects).

[0197] Furthermore, these graphical representations are shown as two-dimensional but can also be three-dimensional (3D). Further, the graphical representations can be static, such as a 2D or 3D emoji that do not move or change facial expressions. Alternatively, the graphical representations can be dynamic, such as 2D or 3D emoji that move, talk, change facial expressions, rotate, etc. Further yet, graphical representations in accordance with example embodiments can be presented as AR images and VR images.

[0198] The graphical representations **500**A-**500**R include or are associated with a sound, such as a sound clip, a sound file, a recorded voice message, streaming audio, etc. The sound can play for a short period of time (e.g., less than one second, one second, two seconds, etc.). For example, the sound is a voice saying "Hello" or "Hahahaha" or "Thank you" or another short audio message. As another example, the sound is a computer-generated "Beep" or phone ringing or explosion sound. The sound can play for longer periods of time (e.g., ten seconds, thirty seconds, one minute, several minutes, etc.). For example, the sound is a recorded message from a user during an electronic communication between two users.

[0199] By way of example, the sound plays when the listener activates the graphical representation or another action occurs that initiates or activates playing of the sound. For example, a first user sends a second user an emoji shaped like a heart. This heart appears on a display of an electronic device of the second user. When the second user clicks on the heart, a voice in binaural sound says "I love you" to the second user.

[0200] With example embodiments, the graphical representations can represent or symbolize the listener or source of sound (depending on what instructions or understandings are provided to the listener and/or users). The graphical representations and indications are displayed to the listener on a display of a WED, PED, HPED, HMD, or other electronic device discussed herein. The electronic device and display are not shown in FIGS. **5**A-**5**R for ease of illustration. Further, these figures are shown from the point-of-view of the listener looking at the display and/or interacting with the electronic device.

[0201] As explained herein, the graphical representation and/or indication can be altered or changed in response to determining how the sound will play to the listener or how the listener will hear the sound (e.g., hear the sound as binaural sound or one of mono sound or stereo sound). For example, the indication is added or removed from the graphical representation in response to this determination.

[0202] Consider an example in which the graphical representation is originally sent or displayed as **500**A-**500**R (e.g., without the indication). Upon determining how the listener will hear the sound or how the sound will localize, the graphical representation is altered to include the indication **510**A-**510**R. The indications can be added and removed as discussed herein.

[0203] FIG. **5**A shows a graphical representation **500**A with an indication **510**A shown as an arrow or pointer. One end of the arrow connects to the head of the graphical representation, and another end of the arrow points to a location where binaural sound will externally localize to the listener. Indication **510**A points upward and to one side of the face of the graphical representation. This location shows the listener where the binaural sound will externally localize to the listener when the binaural sound plays to the listener.

[0204] Consider an example in which the graphical representation **500**A displays on a display in front of the listener and represents a person or character that will talk to the listener. For example, the listener holds a smartphone or wears a HMD or WED, and the display shows the graphical representation **500**A in front of the face of the listener. In this instance, the arrow points to a location in space that is upward and to the right of the listener. When the listener activates the graphical representation, the sound externally localizes to the location indicated with the indication **510**A (here, upward and to the right of the face of the listener).

[0205] Consider an example in which the arrow is colored (e.g., blue or another color) and/or flashes or changes brightness. When users sees this arrow, they know that sound associated with the corresponding graphical representation will externally localize as binaural sound. This arrow appears on other emoji or emoticons. Users recognize the arrow as an indication or message that the sound will be in binaural sound.

[0206] The indications can thus serve as a way to visually inform users that the sound associated with the graphical representation will be binaural sound. Users learn the recognize the indication as a symbol for binaural sound. When a listener sees the arrow, he or she immediately knows in advance that the sound will be binaural sound and externally localize, as opposed to mono sound or stereo sound that internally localizes inside a head of the listener.

[0207] FIG. **5**B shows a graphical representation **500**B with an indication **510**B. The indication includes a pointer or arrow having one end near or adjacent to eyes of the head of the graphical representation, and another end of the arrow points to a location in empty space (marked with an "X") that is next to one side of the head or the face of the graphical representation. This location at the "X" indicates where binaural sound will externally localize to the listener. The arrow also indicates to the listener to look in this direction since this direction is where the sound will externally localize once played. The indication thus shows the listener which direction to look or to face so the listener looks at the SLP when the binaural sound activates and plays to the listener.

[0208] Consider an example in which the graphical representation **500**B displays through a HMD that the listener wears. The graphical representation does not initially display the indication **510**B. Binaural sound will externally localize at a SLP having spherical coordinate location (1.0, 0.0°, −30°) with respect to the head of the listener as an origin. The sound is about to play, but the listener is looking in a different direction, such as looking at an object at (3.0, 10°, +45°). The listener is thus not current looking at or facing the SLP. In response, the HMD flashes the indication **510**B on the display. The indication informs the listener that binaural sound is about to play. The indication also informs the listener to move his or her head in a direction show by the arrow since the sound will externally localize. The listener moves his or her head until the "X" is in his or her field-of-view. When this occurs, the HMD removes the indication **510**B and plays the binaural sound as an animated 3D VR talking human character.

[0209] FIG. **5**C shows a graphical representation **500**C with an indication **510**C shown as head of a person. The indication and the graphical representation together show where the binaural sound will externally localize to the listener before or while the sound plays to the listener.

[0210] Consider an example in which a first user and a second user talk or exchange talking graphical representations during an electronic communication. The first user sends the second user a talking emoji shown as **500**C which displays to the second user on a HPED. This emoji is an animated head that looks like or represents the first user. So, the face of the first user appears on the display of the HPED of the second user and faces the second user as shown in FIG. **5**C. The indication **510**C also appears on this display and represents the head of the second user. So, the second users sees himself or herself (indication **510**C) talking to the first user (graphical representation **500**C). The relative position of graphical representation **500**C to indication **510**C clearly shows that the first user is located in front of and to the left of the second user. In other words, the relative position of the two heads on the display of the HPED show where the SLP will be for the second user.

[0211] FIG. **5**D shows a graphical representation **500**D with an indication **510**D shown as source of binaural sound (a black dot with lines emanating from the black dot). As shown with an arrow, the source of binaural sound moves from a first location **520**D outside a head of a person to a second location **530**D inside a head of the person.

[0212] The indication **510**D provides the listener with a variety of different valuable information. Consider the example in which the graphical representation **500**D represents or symbolizes the head of the listener and is shown on a display of an HPED to the listener. First, the indication shows the listener that the sound will be binaural sound since the location **520**D is physically located outside of the head of the listener. Second, the indication shows a location of where this binaural sound will initially localize to the listener. As shown, the binaural sound initially has a SLP to a right side of the face of the listener. The indication shows transition or movement of the binaural sound. The SLP starts at **520**D and moves along the arrow to **530**D. Here, the sound initially localizes as binaural sound and then moves to internally localize as stereo or mono sound.

[0213] FIG. **5**E shows a graphical representation **500**E with an indication **510**E that includes a curved pointer or arrow. One end of the arrow points to the head or ears of the face of the graphical representation, and another end of the arrow includes a "3D" that indicates three-dimensional sound.

[0214] The indication **510**E shows the listener that the sound will or is localizing in binaural sound. The indication also a direction to the source of the sound that is located at the "3D." Furthermore, a size of the "3D" can indicate a distance to the source of the sound. For example, different font sizes represent different distances. For instance, a "3D" with a larger size indicates the source of sound is closer than a "3D" with a smaller size.

[0215] FIG. **5**F shows a graphical representation **500**F with an indication **510**F that shows the sound will externally localize to the listener as binaural sound. The indication

includes the word "Sound" and sound waves entering from one side of the head of the graphical representation.

[0216] Consider an example in which the graphical representation **500**F represents or symbolizes the head of the listener. In this instance, the source of sound originates from a right side located above the head of the listener.

[0217] FIG. **5**G shows a graphical representation **500**G with an indication **510**G that shows the sound will externally localize to the listener as binaural sound. The indication includes the acronym or letters "FL" that stand for "front left." Based on this indication, the listener expects the source of sound to be in front of the face of the listener and to his or her left.

[0218] The location of the source of binaural sound can appear inside the body of the graphical representation with words, text, symbols, images, or other indicia that indicate a direction and/or distance to the source of the sound. For example, instead of "FL" the indication includes a compass heading (such as North, South, East, or West) or a coordinate location (such as coordinate location in rectangular coordinates, polar coordinates, or spherical coordinates).

[0219] FIG. **5**H shows a graphical representation **500**H with an indication **510**H. The indication includes a SLP or source of sound **520**H with respect to the graphical representation **510**H. The indication shows a distance (1.0 meter) from the graphical representation to the source of sound. This distance shows the listener that the binaural sound will externally localize one meter away from the head of the listener.

[0220] In an example embodiment, the indication remains displayed with the graphical representation. For instance, while the graphical representation displays to the listener, the indication simultaneously displays to the listener. In another example embodiment, the indication displays for a temporary period of time with the graphical representation. For instance, the indication initially displays with the graphical representation to notify or inform the user of the existence and location of the source of the binaural sound. The indication then disappears while the graphical representation continues to display to the listener while the sound plays to the listener.

[0221] FIG. **5**I shows a graphical representation **500**I with an indication **510**I that shows the sound will externally localize to the listener as binaural sound. The indication includes the acronym or letters "3D" that stand for three-dimensional. Based on this indication, the listener expects the sound to be 3D sound or binaural sound that externally localizes to the listener.

[0222] Consider an example embodiment in which the indication is instead "Mono" or "Stereo" or another symbol or word to visually indicate that the sound will localize as or be provided as mono sound or stereo sound.

[0223] FIG. **5**J shows a graphical representation **500**J with an indication **510**J that shows the face with ears and sound (shown as three lines) emanating into the ears. This information informs the listener to wear headphones and informs the listener that sound is or will be binaural sound that externally localizes away from the head of the listener.

[0224] FIG. **5**K shows a graphical representation **500**K with an indication **510**K that includes headphones on the head of the face of the graphical representation. This indication informs the listener to wear or don headphones before listening to the sound.

[0225] Consider an example in which a first user sends a second user a 3D image that talks to the second user. When the second user activates the 3D image, a sound clip plays to the listener as binaural sound that externally localizes away from the head of the second user. The 3D image is or represents the first user. For instance, this image looks like or resembles the first user. This 3D image includes the first user wearing headphones. When the second user sees the image of the first user wearing headphones, the second user is reminded or instructed to wear headphones before listening to the sound clip. When the user puts on headphones and the sound clip begins to play, the headphones disappear from the 3D image.

[0226] FIG. **5**L shows a graphical representation **500**L with an indication **510**L that instructs a listener to wear earphones or earbuds before listening to the sound associated with the graphical representation. This indication informs the listener to wear or don headphones before listening to the sound. Specifically, the indication shows the graphical representation putting on earphones or shows earphones being placed on the head of the graphical representation. This information reminds or instructs the listener to also put on earphones.

[0227] Consider an example in which a display of user simultaneously displays many graphical representations. Graphical representations with indications indicating 3D sound or binaural sound quickly show the user which ones of the graphical representations are in binaural sound or will play as binaural sound. Those graphical representations without such an indication are in stereo or mono sound or will play as stereo or mono sound. Providing such indications with emojis, for example, enables a listener to quickly visually determine a format of emojis with sound (e.g., a format of stereo or mono sound versus a format of binaural sound).

[0228] FIG. **5**M shows a graphical representation **500**M with an indication **510**M that instructs a listener to wear a head mounted display (HMD) or other wearable device (e.g., a smartphone that attaches to the head of the listener via a head mount). This indication informs the listener to wear or don the HMD before listening to the sound. Specifically, the indication shows the graphical representation putting on the HMD. This information reminds or instructs the listener to also put on his or her HMD.

[0229] FIG. **5**N shows a graphical representation **500**N with an indication **510**N shown as a source of the binaural sound. The indication includes a circle that represents the source of sound or SLP from where the listener will hear the binaural sound. Three lines indicate sound emanating from the circle to the listener.

[0230] The graphical representation **500**N also includes eyes that are looking toward the indication **520**N. The direction of gaze of the graphical representation shows the listener the direction of the source of sound.

[0231] Consider an example in which two users send emojis with 3D sound to each other during an electronic communication. The graphical representation **500**N is a 3D talking emoji of the first user that was sent to the HPED of the second user. Before activating the emoji, the second user sees eyes of the graphical representation looking forward but to one side. This same side includes source of sound emitting sound. This information instructs the second user that the sound will be 3D sound or binaural sound. The information also instruct the second user of the location where the second

user will hear the sound. When the second user activates the emoji, its mouth moves and begins to talk to the second user. The indication **520**N vanishes from the display, and the second user sees the graphical representation **500**N talking.

[0232] FIG. **5**O shows a graphical representation **500**O with an indication **510**O shown as eyes and face of the graphical representation looking in a particular direction. This direction indicates the location of the binaural sound to the listener.

[0233] The face of the graphical representation **500**O is rotated to the right from the point-of-view of the listener. This information instructs the listener that the listener should look to his or her right. Alternatively or additionally, this information instructs the listener that the source of sound or SLP will be to the right of the head of the listener.

[0234] FIG. **5**O shows that the looking direction or gaze of the eyes can indicate the location of the source of sound or SLP. The direction of the face or looking direction of the face can also indicate this information. For example, the head of the graphical representation **500**O rotates to a right to indicate the SLP will be to the right, rotates left to indicate the SLP will be to the left, rotates up to indicate the SLP will be up, etc.

[0235] The graphical representation **500**O also includes eyes that are looking toward the indication **520**O. The direction of gaze of the graphical representation shows the listener the direction of the source of sound.

[0236] FIG. **5**P shows a graphical representation **500**P with an indication **510**P that includes a source of sound or SLP **520**P and information **530**P about the location of the SLP. The information shows a coordinate location where the SLP will or does exist for the listener. This SLP has a spherical coordinate location of (1.0 m, 30°, 25°).

[0237] In FIG. **5**P, the indication **510**P shows the precise or exact coordinate location where the sound will play to the listener. The SLP **520**P is also positioned at this same location. So both the coordinate location (shown at **530**P) and the SLP **520**P show the listener where the sound will play to the listener.

[0238] FIG. **5**Q shows a graphical representation **500**Q with an indication **510**Q that includes headphones. These headphones serve one or more of several functions. First, the headphones instruct the listener to put on headphones. Second, the headphones instruct the listener that the sound the listener will hear will be binaural sound since proper external sound localization of binaural sound requires headphones or earphones. Third, the headphones show a location where the binaural sound will externally localize to the listener. As shown in FIG. **5**Q, the sound will originate above and in front of the listener since this is the location of the headphones with respect to the head of the graphical representation.

[0239] FIG. **5**R shows a graphical representation **500**R with an indication **510**R that includes a facial expression of the graphical representation. The facial expression (shown as a sad face) indicates to the listener that the listener is not ready to hear the sound associated with the graphical representation.

[0240] Consider an example in which emojis with sad faces represent that the listener is not wearing headphones or earphones. When the listener dons headphones or earphones, the emoji with the sad face changes to a happy face. This change and the happy face signify to the listener that the listener is ready to hear the sound as binaural sound.

[0241] FIG. **6** is an example computer system **600** in accordance with an example embodiment.

[0242] The computer system **600** includes one or more of a server **610**, a database **620**, an electronic device **630**, and an electronic device **640** in communication over one or more networks **650**. User **639** is with or uses electronic device **630**, and user **649** is with or uses electronic device **640**. For illustration, a single server **610**, a single database **620**, two electronic devices **630** and **640**, and two users **639** and **649** are shown, but example embodiments can include a plurality of servers, databases, electronic devices, and users.

[0243] Server **610** includes a memory **612** and a processing unit **614**. The memory **621** includes sound clips **616** and graphical representations or graphical reps **618** of the sound clips. The server **610** couples to or communicates with the database **620** that includes sound clips **622** and graphical representations or graphical reps **624**.

[0244] Electronic device **630** includes a processing unit **632** and memory **634** with sound clips **636** and graphical representations or graphical reps **638**. User **639** interacts with or uses electronic device **630**.

[0245] Electronic device **640** includes a processing unit **642** and memory **644** with sound clips **646** and graphical representations or graphical reps **648**. User **649** interacts with or uses electronic device **630**.

[0246] FIG. **7** is an example of an electronic device **700** in accordance with an example embodiment.

[0247] The electronic device **700** includes a processor or processing unit **710**, memory **720** with sound clips **722** and graphical representations or graphical reps **724**, a display **730**, one or more interfaces **740**, a wireless transmitter/receiver **750**, speakers **760**, one or more microphones **770**, head tracking **780** (such as one or more of an inertial sensor, accelerometer, gyroscope, and magnetometer), and a graphical representation and/or indication changer (e.g., software, hardware, and/or program instructions discussed in figures herein that alter or change the graphical representation and/or indication in response to determining how the sound is or will play to the listener).

[0248] Memory includes computer readable medium (CRM). Examples of an interface include, but are not limited to, a network interface, a graphical user interface, a natural language user interface, a natural user interface, a phone control interface, a reality user interface, a kinetic user interface, a touchless user interface, an augmented reality user interface, and/or an interface that combines reality and virtuality.

[0249] Sound clips include sound files, sounds, recorded messages (such as voice messages or other recorded sound), computer-generated sounds, and other sound discussed herein. For example, users can record, exchange, and/or transmit sound clips or sounds. These sound include sending streaming sounds or sounds in real-time during an electronic communication.

[0250] The processor or processing unit includes a processor and/or a digital signal processor (DSP). For example, the processing unit includes one or more of a central processing unit, CPU, digital signal processor (DSP), microprocessor, microcontrollers, field programmable gate arrays (FPGA), application-specific integrated circuits (ASIC), etc. for controlling the overall operation of memory (such as random access memory (RAM) for temporary data storage, read only memory (ROM) for permanent data storage, and firmware).

[0251] Consider an example embodiment in which the processing unit includes both a processor and DSP that communicate with each other and memory and perform operations and tasks that implement one or more blocks of the flow diagram discussed herein. The memory, for example, stores applications, data, programs, sound clips, algorithms (including software to implement or assist in implementing example embodiments) and other data.

[0252] For example, a processor or DSP executes a convolving process with the retrieved HRTFs or HRIRs (or other transfer functions or impulse responses) to process sound clips so that the sound is adjusted, placed, or localized for a listener away from but proximate to the head of the listener. For example, the DSP converts mono or stereo sound to binaural sound so this binaural sound externally localizes to the user. The DSP can also receive binaural sound and move its localization point, add or remove impulse responses (such as RIRs), and perform other functions.

[0253] For example, an electronic device or software program convolves and/or processes the sound captured at the microphones of an electronic device and provides this convolved sound to the listener so the listener can localize the sound and hear it. The listener can experience a resulting localization externally (such as at a sound localization point (SLP) associated with near field HRTFs and far field HRTFs) or internally (such as monaural sound or stereo sound).

[0254] The memory stores HRTFs, HRIRs, BRTFs, BRIRs, RTFs, RIRs, or other transfer functions and/or impulse responses for processing and/or convolving sound. The memory can also store instructions for executing one or more example embodiments. Further, the memory can store the sound, graphical representations, and other information and instructions discussed herein.

[0255] The electronic device provides sound to the users through one or more speakers. Alternatively or in addition to the speakers, the electronic device can communicate with headphones, earphones, earbuds, bone conduction devices, or another electronic device that provides sound to the user.

[0256] The networks include one or more of a cellular network, a public switch telephone network, the Internet, a local area network (LAN), a wide area network (WAN), a metropolitan area network (MAN), a personal area network (PAN), home area network (HAM), and other public and/or private networks. Additionally, the electronic devices need not communicate with each other through a network. As one example, electronic devices couple together via one or more wires, such as a direct wired-connection. As another example, electronic devices communicate directly through a wireless protocol, such as Bluetooth, near field communication (NFC), or other wireless communication protocol.

[0257] By way of example, a computer and an electronic device include, but are not limited to, handheld portable electronic devices (HPEDs), wearable electronic glasses, electronic or smart watches, wearable electronic devices (WEDs), smart earphones or hearables, electronic devices with cellular or mobile phone capabilities or subscriber identification module (SIM) cards, desktop computers, servers, portable computers (such as tablet and notebook computers), smartphones, head mounted displays (HMDs), optical head mounted displays (OHMDs), headphones, and other electronic devices with a processor or processing unit, a memory, a DSP.

[0258] Example embodiments are not limited to HRTFs but also include other sound transfer functions and sound impulse responses including, but not limited to, head related impulse responses (HRIRs), room transfer functions (RTFs), room impulse responses (RIRs), binaural room impulse responses (BRIRs), binaural room transfer functions (BRTFs), headphone transfer functions (HPTFs), etc.

[0259] Example embodiments can be executed with one or more integrated circuits that are specifically customized, designed, or configured to execute one or more blocks discussed herein. For example, the electronic devices include a specialized or custom processor or microprocessor or semiconductor intellectual property (SIP) core or digital signal processor (DSP) with a hardware architecture optimized for convolving sound and executing one or more example embodiments.

[0260] Consider an example in which the HPED (including headphones) includes a customized or dedicated DSP that executes one or more blocks discussed herein (including processing and/or convolving sound into binaural sound for sound clips). Such a DSP has a better power performance or power efficiency compared to a general-purpose microprocessor and is more suitable for a HPED or WED due to power consumption constraints of the HPED or WED. The DSP can also include a specialized hardware architecture, such as a special or specialized memory architecture to simultaneously fetch or pre-fetch multiple data and/or instructions concurrently to increase execution speed and sound processing efficiency and to quickly correct errors while sound externally localizes to the user. By way of example, streaming sound data (such as sound data in a telephone call or software game application) is processed and convolved with a specialized memory architecture (such as the Harvard architecture or the Modified von Neumann architecture). The DSP can also provide a lower-cost solution compared to a general-purpose microprocessor that executes digital signal processing and convolving algorithms. The DSP can also provide functions as an application processor or microcontroller. The DSP can also prefetch sound clips and other sound from memory to expedite convolution.

[0261] Consider an example in which a customized DSP includes one or more special instruction sets for multiply-accumulate operations (MAC operations), such as convolving with transfer functions and/or impulse responses (such as HRTFs, HRIRs, BRIRs, et al.), executing Fast Fourier Transforms (FFTs), executing finite impulse response (FIR) filtering, and executing instructions to increase parallelism.

[0262] Consider another example in which sound clips, graphical representations, and/or HRTFs (or other transfer functions or impulse responses) are stored or cached in the DSP memory or local memory relatively close to the DSP to expedite binaural sound processing.

[0263] As used herein, an "emoji" is a graphical representation that includes images, symbols, or icons sent between users in electronic communications (such as text messages, e-mail, and social media) to express an emotional attitude of the writer, convey information, or communicate an message. Emojis can provide sound when activated or executed.

[0264] As used herein, "headphones" or "earphones" include a left and right over-ear ear cup, on-ear pad, or in-ear monitor (IEM) with one or more speakers or drivers for a left and a right ear of a wearer. The left and right cup, pad, or

IEM may be connected with a band, connector, wire, or housing, or one or both cups, pads, or IEMs may operate wirelessly being unconnected to the other. The drivers may rest on, in, or around the ears of the wearer, or mounted near the ears without touching the ears.

[0265] As used herein, the word "proximate" means near. For example, binaural sound that externally localizes away from but proximate to a user localizes within three meters of the head of the user.

[0266] As used herein, a "user" or a "listener" is a person (i.e., a human being). These terms can also be a software program (including an IPA or IUA), hardware (such as a processor or processing unit), an electronic device or a computer (such as a speaking robot or avatar shaped like a human with microphones in its ears or about six inches apart).

[0267] In some example embodiments, the methods illustrated herein and data and instructions associated therewith, are stored in respective storage devices that are implemented as computer-readable and/or machine-readable storage media, physical or tangible media, and/or non-transitory storage media. These storage media include different forms of memory including semiconductor memory devices such as DRAM, or SRAM, Erasable and Programmable Read-Only Memories (EPROMs), Electrically Erasable and Programmable Read-Only Memories (EEPROMs) and flash memories; magnetic disks such as fixed and removable disks; other magnetic media including tape; optical media such as Compact Disks (CDs) or Digital Versatile Disks (DVDs). Note that the instructions of the software discussed above can be provided on computer-readable or machine-readable storage medium, or alternatively, can be provided on multiple computer-readable or machine-readable storage media distributed in a large system having possibly plural nodes. Such computer-readable or machine-readable medium or media is (are) considered to be part of an article (or article of manufacture). An article or article of manufacture can refer to a manufactured single component or multiple components.

[0268] Blocks and/or methods discussed herein can be executed and/or made by a user, a user agent (including machine learning agents and intelligent user agents), a software application, an electronic device, a computer, firmware, hardware, a process, a computer system, and/or an intelligent personal assistant. Furthermore, blocks and/or methods discussed herein can be executed automatically with or without instruction from a user.

1.-20. (canceled)

21. A method comprising:

determining, with an electronic device, whether a user is wearing one of headphones and earphones;

displaying, with the electronic device, a graphical representation with a first appearance upon determining that the user is not wearing the one of the headphones and the earphones; and

displaying, with the electronic device, the graphical representation with a second appearance upon determining that the user is wearing the one of the headphones and the earphones, wherein the second appearance visually informs the user that upon activation of the graphical representation sound will externally localize as binaural sound away from a head of the user.

22. The method of claim **21** further comprising:

determining, with the electronic device, when the user removes the one of the headphones and the earphones; and

changing, in response to determining that the user removes the one of the headphones and the earphones, the graphical representation from having the second appearance to having the first appearance.

23. The method of claim **21** further comprising:

processing, with a digital signal processor (DSP) in the electronic device, the sound of the graphical representation to change the sound into the binaural sound in response to determining that the user is wearing the one of the headphones and the earphones; and

saving processing resources of the DSP by not processing the sound into the binaural sound in response to determining that the user is not wearing the one of the headphones and the earphones.

24. The method of claim **21** further comprising:

displaying, with the electronic device, the graphical representation with the second appearance that includes an indication showing where the sound will externally localize as the binaural sound away from the head of the user.

25. The method of claim **21**, wherein the first appearance of the graphical representation is altered by changing a color of the graphical representation being displayed to provide an indication that visually informs the user that the sound will externally localize as the binaural sound away from the head of the user.

26. The method of claim **21**, wherein the graphical representation with the second appearances includes "3D" to visually inform the user that the sound will externally localize as the binaural sound away from the head of the user.

27. The method of claim **21**, wherein the first appearance of the graphical representation is altered by adding one or more words to the graphical representation being displayed that when read by the user provides an indication that visually informs the user that the sound will externally localize as the binaural sound away from the head of the user.

28. A non-transitory computer-readable storage medium storing instructions that one or more electronic devices execute to perform a method comprising:

determining whether a user wears headphones;

displaying a graphical representation without an indication of binaural sound in response to determining that the user is not wearing the headphones; and

displaying the graphical representation with the indication of the binaural sound in response to determining that the user is wearing the headphones,

wherein the indication of the binaural sound visually shows the user that sound of the graphical representation will externally localize outside a head of the user as the binaural sound when the binaural sound of the graphical representation plays to the user through the headphones.

29. The non-transitory computer-readable storage medium of claim **28** further comprising:

processing, with a digital signal processor (DSP), the sound into the binaural sound in response to determining that the user wears the headphones; and

saving processing resources of the DSP by not convolving the sound into binaural sound in response to determining that the user is not wearing the headphones.

30. The non-transitory computer-readable storage medium of claim **28** further comprising:

adding "3D" to an appearance of the graphical representation as the indication of the binaural sound in response to determining that the user is wearing the headphones; and

removing the "3D" to the appearance of the graphical representation as the indication in response to determining that the user is not wearing the headphones, wherein the headphones include earphones.

31. The non-transitory computer-readable storage medium of claim **28** further comprising:

adding an appearance to the graphical representation in response to detecting the headphones connect with an electronic device via a wired or a wireless connection; and

removing the appearance from the graphical representation in response to detecting the headphones disconnected from the electronic device.

32. The non-transitory computer-readable storage medium of claim **28** further comprising:

providing the sound to the user as stereo sound or mono sound in response to determining that the user is not wearing the headphones; and

providing the sound to the user as the binaural sound in response to determining that the user is wearing the headphones.

33. The non-transitory computer-readable storage medium of claim **28** further comprising:

visually notifying the user that the sound from the graphical representation will externally localize as the binaural sound by flashing the graphical representation brighter and dimmer while the graphical representation is displayed, and wherein a wearable electronic device (WED) provides the binaural sound to the user.

34. A method comprising:

playing sound of graphical representation as mono sound or stereo sound upon determining that the user is not wearing a wearable electronic device (WED);

altering, in response to determining that the user is wearing the WED, an appearance of the graphical representation by adding a visual indication that shows the user that the sound will play as binaural sound that externally localizes outside a head of the user; and

displaying, with on a display, the graphical representation with the visual indication that shows the user that the sound will play as the binaural sound that externally localizes outside the head of the user.

35. The method of claim **34**, wherein the visual indication includes a term "3D".

36. The method of claim **34** further comprising:

removing the visual indication while the graphical representation is being displayed on the display in response to determining that the sound does not play as the binaural sound, wherein the WED is one of electronic glasses and a head mounted display (HMD).

37. The method of claim **34**, wherein the visual indication includes text informing the user that the sound will externally localize as the binaural sound.

38. The method of claim **34**, wherein the visual indication includes a change in color of the graphical representation to inform the user that the sound will externally localize as the binaural sound.

39. The method of claim **34**, wherein the visual indication includes a symbol that indicates the sound will externally localize as the binaural sound.

40. The method of claim **34** further comprising:

saving processing resources by not processing the sound into the binaural sound until after the WED determines whether the sound will play as the binaural sound or the mono or stereo sounds.

* * * * *