



(51) International Patent Classification:

H04R 5/02 (2006.01) *H04S 1/00* (2006.01)

(21) International Application Number:

PCT/US2017/027820

(22) International Filing Date:

14 April 2017 (14.04.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive W., Houston, Texas 77070 (US).

(72) Inventor: **LIU, Wensen**; 16399 W Bernardo Dr., San Diego, California 92127 (US).

(72) Inventor (for US only): **BHARITKAR, Sunil**; 1501 Page Mill Road, Palo Alto, California 94304 (US).

(74) Agent: **HEWLETT-PACKARD COMPANY** et al.; Intellectual Property Administration, 3404 E. Harmony Road, Mail Stop 35, Fort Collins, Colorado 80528 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to the identity of the inventor (Rule 4.17(i))
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))

Published:

- with international search report (Art. 21(3))

(54) Title: CROSSTALK CANCELLATION FOR STEREO SPEAKERS OF MOBILE DEVICES

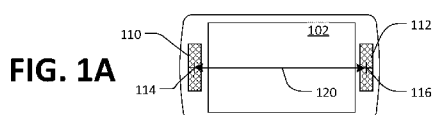


FIG. 1A

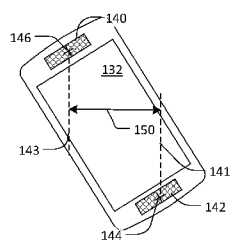


FIG. 1B

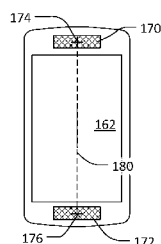


FIG. 1C

(57) Abstract: Described herein is a technology related to a mobile device with a pair of stereo speakers. The mobile device has an orientation detection system that detects the orientation of the mobile device and a crosstalk cancellation system that performs crosstalk cancellation with complementary cardioid beams in response to the detected orientation of the mobile device. The mobile device also has an audio system that emits complementary cardioid sound beams from the speakers.



CROSSTALK CANCELLATION FOR STEREO SPEAKERS OF MOBILE DEVICES

BACKGROUND

[0001] Mobile devices, such as smartphones, phablets, and tablet computers, are ubiquitous. Because they are mobile, these devices are small and thin. Despite that, most mobile devices are equipped with at least one audio speaker.

[0002] For many reasons, the mobile device is frequently the most convenient and personal music source. That is, people often play music or other high-quality audio using the speakers of their mobile device. Because of this, many mobile device have a pair of stereo speakers. These stereo speakers offer a much richer and immersive audio experience than was offered with a single speaker device.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] FIGs. 1A-1C illustrates several example mobile devices in various orientations in accordance with the technologies described herein.

[0004] FIG. 2 illustrates an example mobile device in a speaker-aligned landscape orientation in accordance with the technologies described herein.

[0005] FIG. 3 shows a component diagram of an example mobile device in accordance with the technologies described herein.

[0006] FIG. 4A is a simplified representation of omnidirectional acoustic beams formed by speakers of a mobile device. FIG. 4B is a simplified

representation of cardioid acoustic beam formed by speakers of an example mobile device in accordance with the technologies described herein.

[0007] FIG. 5 shows an example signal flow for forming cardioid beams in accordance with the technologies described herein.

[0008] FIG. 6 is a flowchart of an example process in accordance with the technologies described herein.

[0009] FIG. 7 is a flowchart of an example process to perform crosstalk cancellation for speaker-based spatial rendering in accordance with the technologies described herein.

[0010] FIG. 8 illustrates an example arrangements of virtual speakers employed with the crosstalk cancellation for speaker-based spatial rendering in accordance with the technologies described herein.

[0011] FIG. 9 is a flowchart of another example process to perform crosstalk cancellation for virtual speaker-based spatial rendering in accordance with the technologies described herein.

[0012] The Detailed Description references the accompanying figures. In the figures, the left-most digit(s) of a reference number identifies the figure in which the reference number first appears. The same numbers are used throughout the drawings to reference like features and components.

DETAILED DESCRIPTION

[0013] Described herein are technologies to facilitate a wider stereo image of a pair of stereo speakers of a mobile device and, in some instances, a reduction in crosstalk between the speakers. While mobile devices may have a pair of stereo speakers, a human listener rarely experiences an actual

stereo effect because of the limited spacing between the speakers of the devices and uncertainty of the orientation of those speakers relative the listener's ears. Unlike large stereo systems at home, mobile devices are small, and the stereo speakers are usually micro-speakers with a limited spacing therebetween.

[0014] Indeed, the mobile device has a relatively small stereo image and produces crosstalk. Stereo speakers produce a stereo image that has a "sweet spot" formed by focal point between the speakers. The sweet spot is where a person is fully capable of hearing the full stereo audio mix. Usually, stereo speakers' positions and the sweet spot form an equilateral triangle. Consequently, the sweet spot is five to eight inches away for most mobile devices. This is not a comfortable or convenient listening position.

[0015] Another common issue with stereo speakers is audible crosstalk. With stereo headphones, the sounds from left and right channels are isolated from each other. Unlike that, the sound from each speaker of a pair stereo speakers of a mobile device travels to both ears of the human listener. To reduce the audible crosstalk, the listener may have to hold the mobile device very closely to their face.

[0016] The stereo speakers of the technologies described herein form complementary cardioid beams under defined conditions. An example of a defined condition includes a calculation of the present horizontal spacing between the speakers of the mobile device based on a determination of the relative orientation (e.g., portrait vs. landscape) of the device. Alternatively or in addition, the technologies described herein utilizes a crosstalk canceller for virtual speaker-based spatial rendering.

[0017] With the technology described herein, a user of a mobile device with stereo mini-speakers may listen to the sounds emitted from the mini-speakers with a wider stereo image than is achieved without this technology. Alternatively or in addition, the user hears the stereo image with a minimal crosstalk between the stereo speakers.

[0018] Unlike stereo home theater where speakers are placed on the left and right side relative to the listener, a user can position or hold mobile devices in either portrait or landscape orientations. The mobile devices can also be positioned somewhere in between the two orientations.

[0019] FIGs. 1A-1C shows three examples of how a mobile device may be handled or positioned. These example scenarios represent mobile devices in differing orientations. The mobile devices shown here are smartphones and/or tablets, which are examples of a mobile device with stereo speakers in accordance with the technology described herein.

[0020] The FIG. 1A includes an example mobile device 102 in the landscape mode or orientation. More specifically, the device 102 is shown in speaker-aligned landscape orientation. The mobile device 102 includes a pair of stereo speakers: a nominally left speaker 110 and a nominally right speaker 112. The left speaker 110 has a cross 114 marking its center point, and the right speaker 112 has a cross 116 marking its center point. Double-headed arrow 120 represents a horizontal spacing or distance between the pair of speakers.

[0021] As depicted, the horizontal spacing may be measured from the center point of each speaker when the stereo speakers are designed with

direct front porting. In other implementations, the horizontal spacing may be measured from the center of the porting edge of each speaker.

[0022] As depicted, the speakers 110 and 112 of the mobile device 102 are positioned on the edge of the device, symmetrically along the device's longest axis, and away from the shortest axis. With this arrangement, the horizontal spacing (as represented by the arrow 120) is the maximum horizontal spacing achievable.

[0023] A maximum horizontal spacing is between five and ten inches is common for mobile devices. Dimensions of the particular mobile device that incorporates the technology described herein determines the actual maximum horizontal spacing. In some implementations, the value of the maximum horizontal spacing is stored in the memory of the mobile device.

[0024] The determined horizontal distance between the pair of stereo speakers is a maximum horizontal distance between the speakers when the detected orientation of the mobile device is a speaker-aligned landscape. The speaker-aligned landscape is an orientation of the mobile device where each of the speakers of the pair of stereo speakers is aligned on a substantially horizontal line with a maximum horizontal distance therebetween.

[0025] The FIG. 1B includes an example mobile device 132 in a mode or orientation that is somewhere in between landscape and portrait. The mobile device 132 includes a pair of stereo speakers: a nominally left speaker 140 and a nominally right speaker 142. The left speaker 140 has a cross 144 marking its center point, and the right speaker 142 has a cross 146 marking its center point.

[0026] A vertical dashed line 141 projects from the cross 144 of the left speaker 140. Another vertical dashed line 143 projects from the cross 146 of the right speaker 142. Double-headed arrow 150 represents a horizontal spacing or distance between vertical dashed lines of the speakers.

[0027] As depicted in FIG. 1B, the mobile device 132 is between horizontal (i.e., landscape) and vertical (i.e., portrait) orientation. As such, the horizontal spacing (as represented by the double-headed arrow 150) is less than the maximum horizontal spacing but more than zero.

[0028] The FIG. 1C includes an example mobile device 162 in a portrait mode or orientation. More specifically, the device 162 is shown in speaker-aligned portrait orientation. The mobile device 162 includes a pair of stereo speakers: a nominally left speaker 170 and a nominally right speaker 172. The left speaker 170 has a cross 174 marking its center point, and the right speaker 172 has a cross 176 marking its center point.

[0029] A vertical dashed line 180 connects cross 174 of the left speaker 170 to the cross 176 of the right speaker 172. Because the mobile device 162 is in a portrait orientation, there is zero horizontal spacing between the speakers. When in the portrait orientation as shown by FIG. 1C, the stereo image is lost. No speaker is actually on the left or the right.

[0030] The determined horizontal spacing between the pair of the stereo speakers is a minimum horizontal spacing (e.g., zero) between the speakers when the detected orientation of the mobile device is a speaker-aligned portrait. The speaker-aligned portrait is an orientation of the mobile device where each speaker of the stereo speakers is aligned on a substantially vertical line with no horizontal spacing therebetween.

[0031] For illustration purposes, the signal from the top speaker is treated as the nominal left channel while the signal from the bottom speaker is treated as a nominal right channel when the mobile device is in portrait orientation. Different processing schemes should be applied depending on the device orientation.

[0032] It is common for a mobile device comes equipped with three-dimensional (i.e., 3D) accelerometer and gyroscope. Consequently, the mobile device is capable of sensing its orientation relative to gravity.

[0033] FIG. 2 depicts an example orientation of another mobile device. FIG. 2 shows an example mobile device 202 in a speaker-aligned landscape orientation. The mobile device 202 has stereo speakers that are not aligned along its longest axis. The mobile device 202 includes a pair of stereo speakers: a nominally left speaker 210 and a nominally right speaker 212. The left speaker 210 has a cross 214 marking its center point, and the right speaker 212 has a cross 216 marking its center point. Double-headed arrow 220 represents a horizontal spacing or distance between the pair of speakers.

[0034] As depicted, the mobile device 202 is skewed relative to top and bottom of the drawing page. However, the double-headed arrow 220 is shown as being horizontal relative to the top and bottom of the drawing page. Indeed, the horizontal spacing shown in Fig. 2 is the maximum horizontal spacing achievable in this type of mobile device. Consequently, this orientation is called landscape. Herein, the label of the landscape is designated when the mobile device is oriented in such a way that the pair of speakers has their maximum horizontal spacing therebetween. Conversely, the label of the portrait is designated when the mobile device is oriented in such a way that

the pair of speakers has their minimum (e.g., zero) horizontal spacing therebetween.

[0035] With an implementation disclosed herein, the detected orientation of the mobile device ranges from a speaker-aligned landscape to a speaker-aligned portrait. The speaker-aligned landscape is an orientation of the mobile device where each speaker of the stereo speakers is aligned on a substantially horizontal line with a maximum horizontal distance therebetween. The speaker-aligned portrait is an orientation of the mobile device where each speaker of stereo speakers is aligned on a substantially vertical line with no horizontal distance therebetween.

[0036] Herein, a substantially horizontal line is plus or minus five degrees from a line that is parallel to the gravitational horizon. Likewise, a substantially vertical line is plus or minus five degrees from a line that is perpendicular to the gravitational horizon.

[0037] FIG. 3 shows an example of a mobile device 310 that implements the technologies described herein. As depicted, the example mobile device 310 includes a housing 305, at least one processor 312, an audio storage system 314, an orientation detector 316, a memory 320, and an audio system that includes a pair of stereo speakers.

[0038] The audio storage system 314 is a primary or secondary memory that stores audio files 330 with stereo channels: nominally left channel 332 and nominally right channel 334. Alternatively, the audio storage system 314 may be connected to a communications network and stream music therefrom.

[0039] The audio system includes a nominally left speaker 340 and a nominally right speaker 342. By default, the left speaker 340 emits an audible

sound signal based on receiving the nominally left channel 332. Similarly, by default, the right speaker 342 emits an audible sound signal based on receiving the nominally right channel 334.

[0040] The memory 320 stores sets of processor-executable instructions, such as a threshold calculator 322, a beamformer 324, and a virtual speaker-based stereo widening and crosstalk canceller 326. Alternatively, these functional components may be implemented as firmware or hardware only. Although depicted separately, the audio storage system 314 may be part of the memory 320.

[0041] The threshold calculator 322 calculates a threshold frequency that is based on a horizontal spacing between the speakers. To calculate this spacing, the threshold calculator 322 obtains the maximum horizontal spacing between the speakers, which is stored in the memory 320 and determines the orientation of the device with the aid of the orientation detector 316.

[0042] When the example mobile device 310 plays a music file (such as file 330) the frequencies below the calculated threshold frequency, the beamformer 324 produces complementary user-directed cardioid beams rather than their omnidirectional beams. These cardioid beams are formed for the frequencies below the calculated threshold frequency. This occurs when there is at least some horizontal spacing between the speakers.

[0043] For all frequencies above the threshold frequency, a virtual speaker based spatial rendering and crosstalk canceller 326 performs speaker-based crosstalk cancellation. The crosstalk canceller is discussed below.

[0044] FIG. 4A shows an example scenario 400 of a simplified depiction of acoustic beams formed by a pair of stereo mini-speakers from a mobile

device. For this example scenario 400, presume that the each of the left mini-speaker 410 and the right speaker 412 in a stereo pair have the same performance characteristics. As depicted, each speaker emits its own omnidirectional acoustic beam 420 and 422 from each micro-speaker.

[0045] A top-plan view of a user's head 430 is position in the sweet spot of the stereo image. If the user's head is further back from the speakers, then the user does not the full stereo effect. Regardless, the user is able to hear crosstalk. That is, the user can hear the acoustic beam 422 from the right speaker 412 in the left ear, and the user can hear the acoustic beam 420 from the left speaker 410 in-the right ear.

[0046] FIG. 4B shows an example scenario 440 of a simplified depiction of acoustic beams formed by a pair of stereo mini-speakers of a mobile device in accordance with the technologies described herein. For this example scenario 440, presume that the each of the left mini-speaker 450 and the right speaker 452 in a stereo pair have the same performance characteristics. As depicted, each speaker emits its own complementary user-directional cardioid acoustic beam 460 and 462 from each micro-speaker. The complementary cardioid sound beams are directed toward the opposite direction of the other channel and create a null toward the other channel.

[0047] A top-plan view of a user's head 470 is position in the sweet spot of the stereo image. This sweet spot (and the overall stereo image) is wider than with the omnidirectional beams. As depicted in FIG. 4B, the original left channel mostly reaches listener's left side and minimizes to the right side; likewise, the same effect applies to the original right signal. Thus the stereo image is widened, and crosstalk is reduced.

[0048] An implementation of the technology described herein form a complementary cardioid beam pattern like that shown in FIG. 4B by applying “+Left/-Left with delay” signal to “Left/Right” speakers. This generates the left cardioid beam 420 for left speaker 410. Likewise, applying “+Right/-Right with delay” signal to “Right/Left” speakers generates the right cardioid beam 422 for right speaker 422.

[0049] The threshold frequency (i.e., aliasing frequency) is calculated by dividing a speed of sound c by twice the determined horizontal spacing d . The calculated threshold frequency or aliasing frequency (f_a) is based on the following equation:

$$f_a = \frac{c}{2d}$$

Equation 1

[0050] The time-delay of the other stereo channel includes an opposite signal of the other stereo channel with a calculated time delay.

[0051] Equation 1 below provides more details. For that equation, $L_1(t)$ and $R_1(t)$ are the original signals of left and right channels. The signals applied to Left/Right speakers are as follows

$$\begin{aligned} L_2(t) &= L_1(t) - R_1(t - t_d) \\ R_2(t) &= R_1(t) - L_1(t - t_d) \end{aligned}$$

Equation 2

Where delay $t_d = d/c$ is the time lapse of sound traveling between stereo speakers, d is calculated horizontal spacing of the speakers, c is the speed of sound; $-L_1(t-t_d)$ and $-R_1(t-t_d)$ are the opposite signals with delay, respectively.

[0052] FIG. 5 illustrates a flow diagram of an example signal process 500 of the forming of a complementary acoustic cardioid beam in accordance with the technology described herein. The example process 500 utilizes equation 1 above.

[0053] Block 510 represents the original signal of the left channel: $L_1(t)$. Block 512 represents the original signal of the right channel: $R_1(t)$.

[0054] Block 514 represents the opposite ($-L_1(t-t_d)$) of the left channel with a delay for combining with the right channel of block 512. Block 516 represents the opposite ($-R_1(t-t_d)$) of the right channel with a delay for combining with the left channel of block 510.

[0055] Block 520 represents the modified left signal emitted by the left speaker to form the left complementary acoustic cardioid beam (like that shown by beam 460 in FIG. 4B). Block 522 represents the modified right signal emitted by the right speaker to form the right complementary acoustic cardioid beam (like that shown by beam 462 in FIG. 4B).

[0056] Without applying heavy signal processing, the two opposite orientated cardioid beams enlarge the stereo image, especially for the frequencies lower than the aliasing or threshold frequency.

[0057] FIG. 6 is a flow diagram illustrating an example process 600 that implements the techniques described herein. The example process 600 may be performed by, for example, mobile device 310. For illustration purpose, the

example process 600 is described as being performed by an example mobile device.

[0058] At 610, the example mobile device detects the orientation of the example mobile device.

[0059] At 612, the example mobile device, based on the detected orientation, calculates a horizontal distance between a pair of stereo speakers of the example mobile device. Block 614 is that calculated horizontal spacing (*d*).

[0060] At 620, the example mobile device obtains an audio file, signal, or stream, such as music files. The file or stream may be acquired from a storage system and/or a memory. The audio file has a pair of stereo channels: a nominally left channel and a nominally right channel.

[0061] At 622, the example mobile device performs a frame-based fast Fourier transform (FFT) on the audio file, signal, or stream. The frequencies are changing from a portion of signal (e.g., frame) to another portion of signal. In other words, audio frequency in one frame may be less than aliasing frequency while the audio frequency in the next frame may be greater than the aliasing frequency.

[0062] At 624, the example mobile device calculating a threshold frequency based, at least in part, on a determined horizontal distance (e.g., block 614) between the pair of speakers. At 624, the example mobile device determines whether the audio signal includes frequencies below the calculated threshold frequency.

[0063] At block 630, for the frequencies that fall below the calculated threshold frequency, the example mobile device performs crosstalk

cancellation with complementary cardioid beams. For this example process 600, this is accomplished by modifying each of the stereo channels to introduce a composite signal of each channel and a time delay from the other channel. Then, the example mobile device emits the complementary cardioid sound beams from the pair of speakers. One of the speakers produces sound based on one of the modified stereo channels while the other of the speakers produces sound based on the other of the modified stereo channels.

[0064] The actions of block 630 is performed when the calculated horizontal spacing is greater than zero. If the example mobile device is positioned in a portrait orientation, the horizontal spacing is zero. Thus, block 630 is not performed.

[0065] At block 632, for the frequencies the fall above the calculated threshold frequency, the example mobile device performs some other crosstalk cancellation. For example, the example mobile device may perform the crosstalk canceller for speaker-based spatial rendering

[0066] FIG. 7 shows a flowchart of an example process 700 for a crosstalk canceller for virtual speaker-based spatial rendering approach in accordance with a crosstalk matrix inversion. A crosstalk matrix is composed of inter-aural responses associated with the inter-aural cues that are loudspeaker angle dependent. Inter-aural responses include ipsilateral and contralateral responses. The listening position forms the angle of the loudspeakers relative to the listener.

[0067] The matrix can be constructed in the time-domain based on the convolution matrix of the ipsilateral and contralateral responses; or, in the frequency domain based on the Fourier transforms of the ipsilateral and

contralateral responses. In short, the approach renders surround sound images with mobile device's stereo speakers. This approach is employed for frequencies above the threshold frequency. Accordingly, at least implementation of the technology described herein is a hybrid approach for performing crosstalk cancellation.

[0068] At 710, the example mobile device obtains original left/right stereo channels (L_1 , R_1).

[0069] At 712, the example mobile device up-mixes the original stereo channels to five virtual surround sound channels: L_s , L_2 , C , R_2 , R_s .

[0070] FIG. 8 shows an example scenario 800 of the arrangement of the five virtual speakers if they actually existed physical space about a listener 830 of an example device 810 with a pair of physical stereo speakers: 812 and 814.

[0071] As depicted, virtual speaker 820 is L_s (Left Surround), virtual speaker 822 is Left, virtual speaker 824 is Center, virtual speaker 826 is Right and virtual speaker 828 is R_s (Right Surround). When the device is in landscape orientation, the Left/Right virtual channels may be chosen at $\pm 30^\circ$ while the Left Surround/Right Surround channels may be at $\pm 110^\circ$ directions relative to the human head median plane.

[0072] At 714, the example mobile device down-mixes the virtual surround sound channels. It does this by applying appropriate HRTFs (Head Related Transfer Functions) to synthesis back to stereo signals (L_3 , R_3) to be playback with stereo speakers 716 and 718.

[0073] The orientation data is used to generate new HRTFs for immersive rendering. For example, in portrait orientation, the crosstalk canceller uses

different angles for the filters as opposed to landscape orientation. Furthermore, as the human head rotates or translates relative to viewing screen interpolation of the crosstalk cancellation filters and the HRTFs can be used to stabilize the audio image perceived from the speakers.

[0074] The HRTFs (Head Related Transfer Functions) are applied to the five virtual speakers. The following are examples of which:

H_{CL} and H_{CR} are the HRTFs for center speaker to left and right ears, respectively;

H_{LL} and H_{LR} are the HRTFs for left speaker to left and right ears, respectively;

H_{RL} and H_{RR} are the HRTFs for right speaker to left and right ears, respectively;

H_{LSL} and H_{LSR} are the HRTFs for left-surround speaker to left and right ears, respectively;

H_{RSL} and H_{RSR} are the HRTFs for right-surround speaker to left and right ears, respectively.

[0075] This assumes that a human head is symmetrical and inline in the middle of the stereo speakers. Due to symmetrical nature, the above ten HRTFs can be reduced to 5 HRTFS as follows:

$$H_C = H_{CL} = H_{CR}$$

$$H_F = H_{LL} = H_{RR}$$

$$H_{FX} = H_{LR} = H_{RL}$$

$$H_S = H_{LSL} = H_{RSR}$$

$$H_{SX} = H_{LSR} = H_{RSL}$$

Equations 3

Where:

$H_c(f)$ – HRTF applied to the center channel, both L and R;

$H_F(f)$ - HRTF applied to the front speaker to adjacent ear;

$H_{FX}(f)$ - HRTF applied to the front speaker to opposite ear;

$H_S(f)$ - HRTF applied to the surround speaker to adjacent ear;

$H_{SX}(f)$ - HRTF applied to the surround speaker to opposite ear.

[0076] These Head Related Transfer Functions incorporate appropriate Inter-aural Time Difference (ITD) and Inter-aural Intensity Difference (IID) components based on virtual speakers' location relative to listener's ear.

[0077] FIG. 9 shows the signal processing flow 900 of this up-mix and down-mix processing.

[0078] At 910, the example mobile device obtains the original stereo channels: L_1 and R_1 .

[0079] At 912, the example mobile device up-mixes to produce five virtual signals: L_S (Left Surround), Left (L), Center (C), Right (R), and R_S (Right Surround).

[0080] At 914, the example mobile device performs HRTF processing.

[0081] At 916, the example mobile device down mixes to synthesized stereo signals: L_3 and R_3 . A physical left speaker 918 of the example mobile device emits sounds based on the L_3 signal. A physical right speaker 920 of the example mobile device emits sounds based on the R_3 signal.

[0082] For the obtained audio signal with frequencies both above and below the calculated threshold frequency, both forms of the crosstalk cancellation are performed. That is, the crosstalk cancellation with complementary cardioid beams is performed for frequencies below the threshold frequency and the crosstalk cancellation with speaker-based rendering is performed for frequencies above the threshold frequency.

[0083] Moreover, the word "exemplary" is used herein to mean serving as an example, instance, or illustration. Any aspect or design described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other aspects or designs. Rather, use of the word exemplary is intended to present concepts and techniques in a concrete fashion. The term "techniques" or "technologies," for instance, may refer to devices, apparatuses, systems, methods, articles of manufacture, and/or computer-readable instructions as indicated by the context described herein.

[0084] As used in this application, the term "or" is intended to mean an inclusive "or" rather than an exclusive "or." That is, unless specified otherwise or clear from context, "X employs A or B" is intended to mean any of the

natural inclusive permutations. That is, if X employs A; X employs B; or X employs both A and B, then "X employs A or B" is satisfied under any of the foregoing instances. In addition, the articles "a" and "an" as used in this application and the appended claims should generally be construed to mean "one or more," unless specified otherwise or clear from context to be directed to a singular form.

[0085] These processes are illustrated as a collection of blocks in a logical flow graph, which represents a sequence of operations that can be implemented in mechanics alone or a combination with hardware, software, and/or firmware. In the context of software/firmware, the blocks represent instructions stored on computer-readable storage media that, when executed by at least one processor, perform the recited operations.

[0086] Note that the order in which the processes are described is not intended to be construed as a limitation, and any number of the described process blocks can be combined in any order to implement the processes or an alternate process. Additionally, individual blocks may be deleted from the processes without departing from the spirit and scope of the subject matter described herein.

[0087] The term "computer-readable media" includes computer-storage media. For example, computer-storage media may include, but are not limited to, magnetic storage devices (e.g., hard disk, floppy disk, and magnetic strips), optical disks (e.g., compact disk (CD) and digital versatile disk (DVD)), smart cards, flash memory devices (e.g., thumb drive, stick, key drive, and SD cards), and volatile and non-volatile memory (e.g., random access memory (RAM), read-only memory (ROM)).

CLAIMS

What is claimed is:

1. A method comprising:

based on a detected orientation of a mobile device, determining a horizontal spacing between a pair of the stereo speakers of the mobile device;

calculating a threshold frequency based, at least in part, on the determined horizontal spacing between the pair of speakers;

obtaining an audio signal having a pair of stereo channels;

determining that the audio signal includes frequencies below the calculated threshold frequency;

in response to the determination that the audio signal includes frequencies below the calculated threshold frequency, performing crosstalk cancellation with complementary cardioid beams, which includes:

modifying each of the stereo channels to introduce a composite signal of each channel and a time delay from the other channel;

emitting complementary cardioid sound beams from the pair of speakers, wherein one of the speakers produces sound based on one of the modified stereo channels while the other of the speakers produces sound based on the other of the modified stereo channels.

2. A method of claim 1, wherein the determined horizontal spacing between the pair of the stereo speakers is a maximum horizontal spacing between the speakers when the detected orientation of the mobile device is a speaker-aligned landscape, wherein the speaker-aligned landscape is an

orientation of the mobile device where the stereo speakers are aligned on a substantially horizontal line with a maximum horizontal spacing therebetween.

3. A method of claim 1, wherein the threshold frequency is calculated by dividing a speed of sound by twice the determined horizontal spacing.

4. A method of claim 1, wherein the time-delay of the other stereo channel includes an opposite signal of the other stereo channel with a calculated time delay.

5. A method of claim 1, wherein the modifying includes calculating a time delay that is calculated by the determined horizontal spacing the stereo speakers divided by a speed of sound.

6. A method of claim 1, wherein the complementary cardioid sound beams are directed toward the opposite direction of the other channel and create a null toward the other channel.

7. A method of claim 1, in response to the determination that the audio signal includes frequencies above the calculated threshold frequency, performing crosstalk cancellation with virtual speaker-based rendering which includes for frequencies above the threshold frequency.

8. A mobile device comprising:
a housing;

a pair of stereo speakers in or on the housing;

an orientation detection system in the housing that is configured to detect an orientation of the mobile device;

an audio system in the housing that is configured to obtain an audio signal having a pair of stereo channels;

a crosstalk cancellation system in the housing that is configured to perform crosstalk cancellation with complementary cardioid beams based on the audio signal and in response to a detected orientation of the mobile device,

wherein the audio system is further configured to emit complementary cardioid sound beams from the pair of speakers.

9. A mobile device of claim 8, wherein an orientation of the mobile device ranges from a speaker-aligned landscape to a speaker-aligned portrait, wherein the speaker-aligned landscape is an orientation of the mobile device where each speaker of the stereo speakers are aligned on a substantially horizontal line with a maximum horizontal spacing therebetween and the speaker-aligned portrait is an orientation of the mobile device where each of the speakers of the pair of the stereo speakers are aligned on a substantially vertical line with no horizontal spacing therebetween.

10. A mobile device of claim 8, wherein the crosstalk cancellation performance includes:

based on a detected orientation, determine a horizontal spacing between the pair of the stereo speakers of the mobile device;

calculate a threshold frequency based, at least in part, on the

determined horizontal spacing between the pair of speakers;

determine that the audio signal includes frequencies below the calculated threshold frequency;

in response to the determination that the audio signal includes frequencies below the calculated threshold frequency, modify each of the stereo channels to introduce a composite signal of each channel and a time delay from the other channel,

wherein one of the speakers are configured to produce sound based on one of the modified stereo channels while the other of the speakers produces sound based on the other of the modified stereo channels.

11. A mobile device of claim 10, wherein the threshold frequency is calculated by dividing a speed of sound by twice the determined horizontal spacing.

12. A non-transitory machine-readable storage medium encoded with instructions executable by a processor, the machine-readable storage medium comprising:

obtaining an audio signal having a pair of stereo channels;

obtaining a threshold frequency;

determining that the audio signal includes frequencies below the calculated threshold frequency;

in response to the determination that the audio signal includes frequencies below the threshold frequency, performing crosstalk cancellation with complementary cardioid beams of the frequencies of the audio signal that below the threshold frequency;

performing crosstalk cancellation based on matrix inversion, wherein a matrix includes ear-related signals for frequencies of the audio signal that are above the threshold frequency;

emitting audio from a pair of stereo speakers of a mobile device, wherein the emitted audio is based, at least in part, on the crosstalk cancellations performed on the obtained audio signal.

13. A non-transitory machine-readable storage medium of claim 12, wherein the obtaining the threshold frequency includes calculating the threshold frequency based, at least in part, on a detected orientation of the mobile device.

14. A non-transitory machine-readable storage medium of claim 12, wherein the threshold frequency is calculated by dividing a speed of sound by twice the determined horizontal spacing.

15. A non-transitory machine-readable storage medium of claim 12, wherein the performance of the crosstalk cancellation with complementary cardioid beams includes:

modifying each of the stereo channels to introduce a time-delay of the other stereo channel;

emitting complementary cardioid sound beams from the pair of speakers, wherein one of the speakers produces sound based on one of the modified stereo channels while the other of the speakers produces sound based on the other of the modified stereo channels.

FIG. 1A

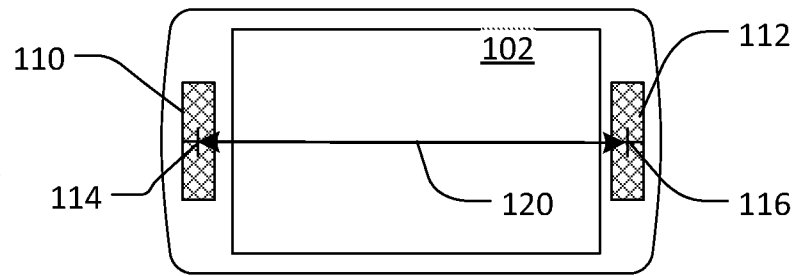


FIG. 1B

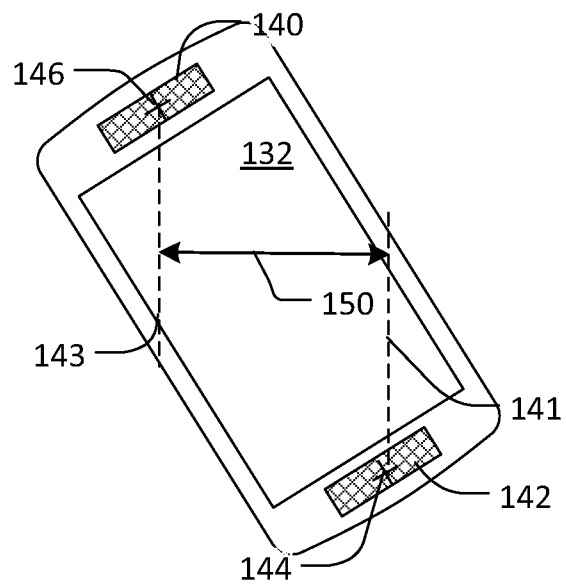
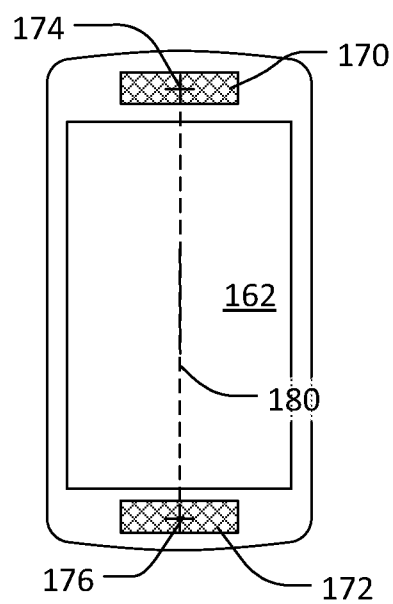
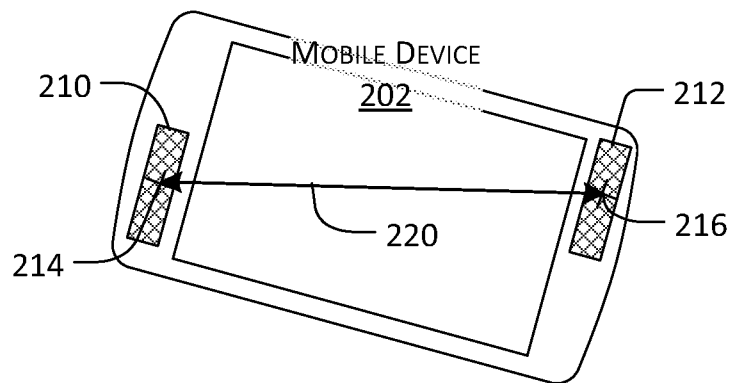
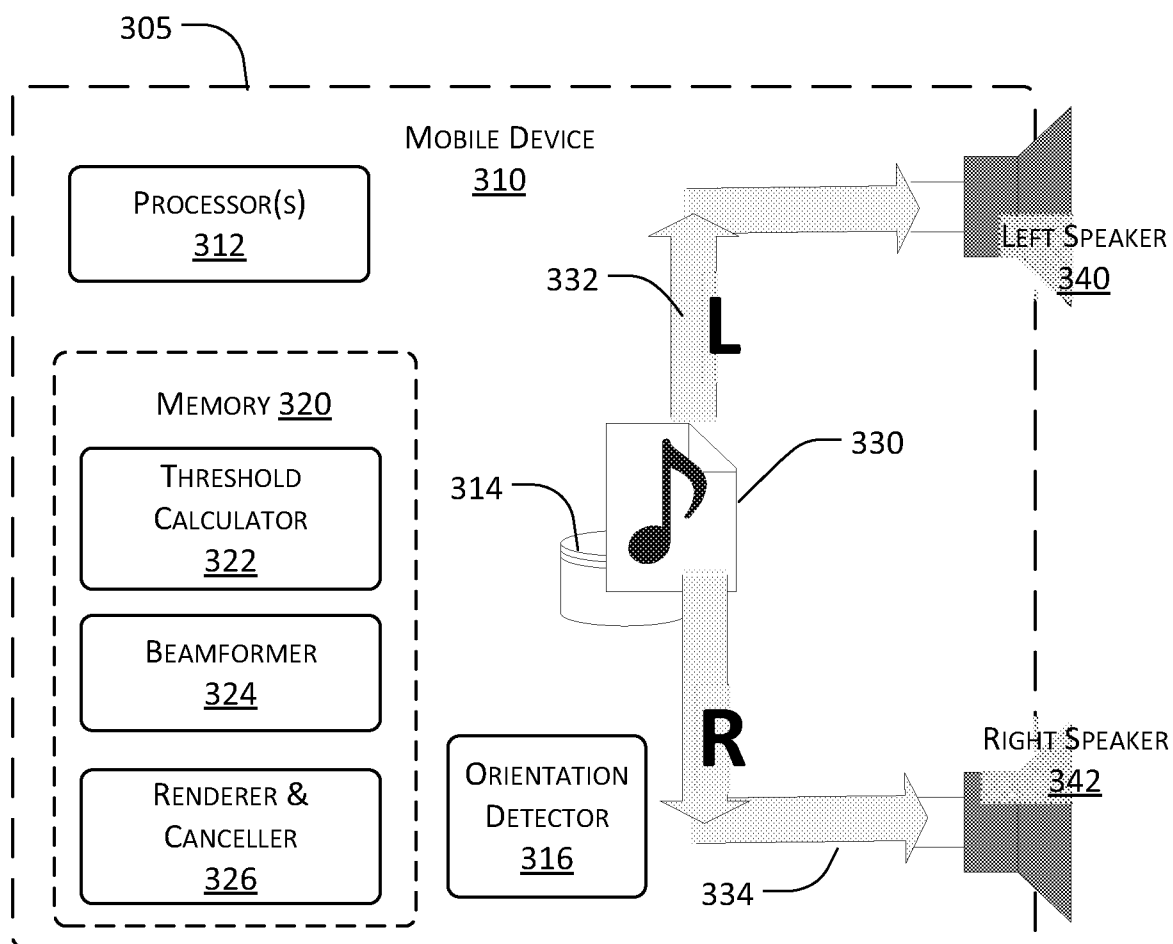
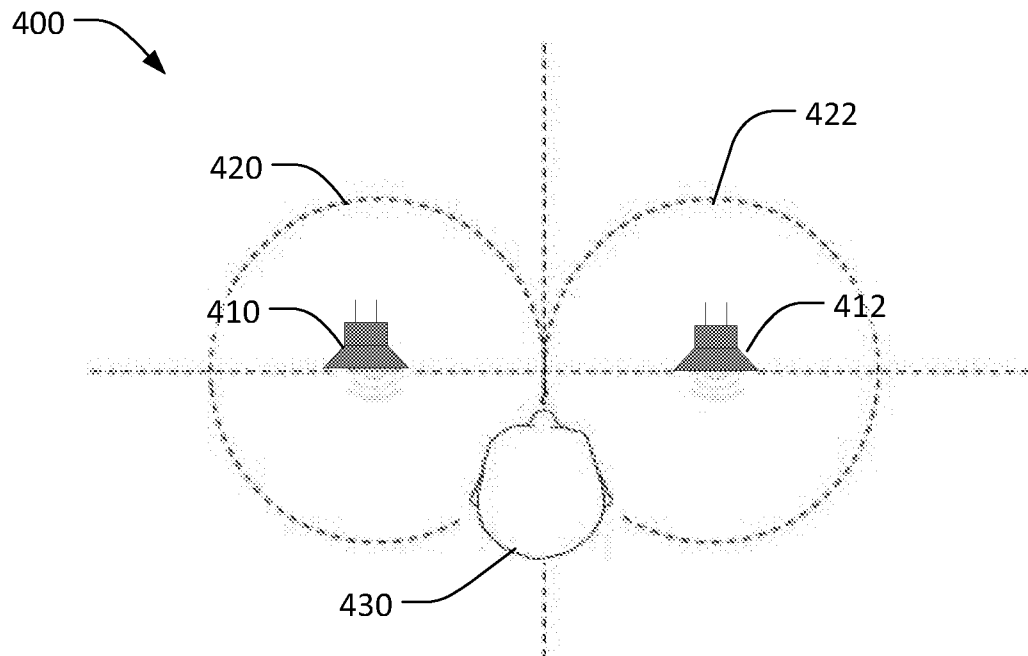
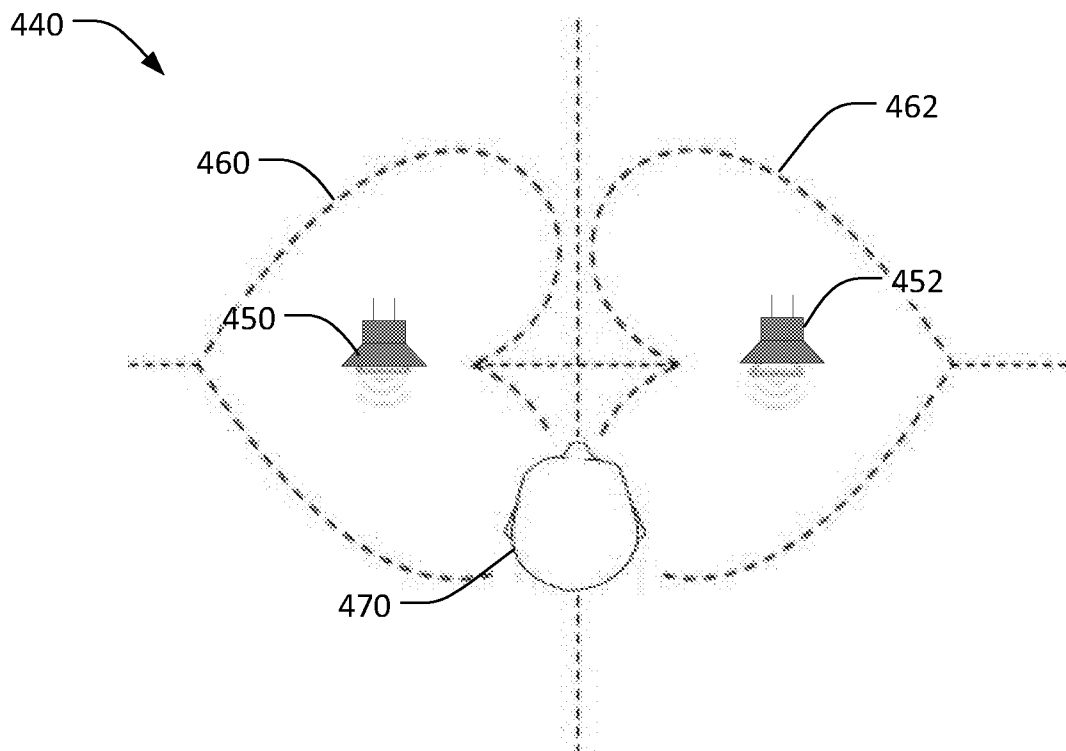
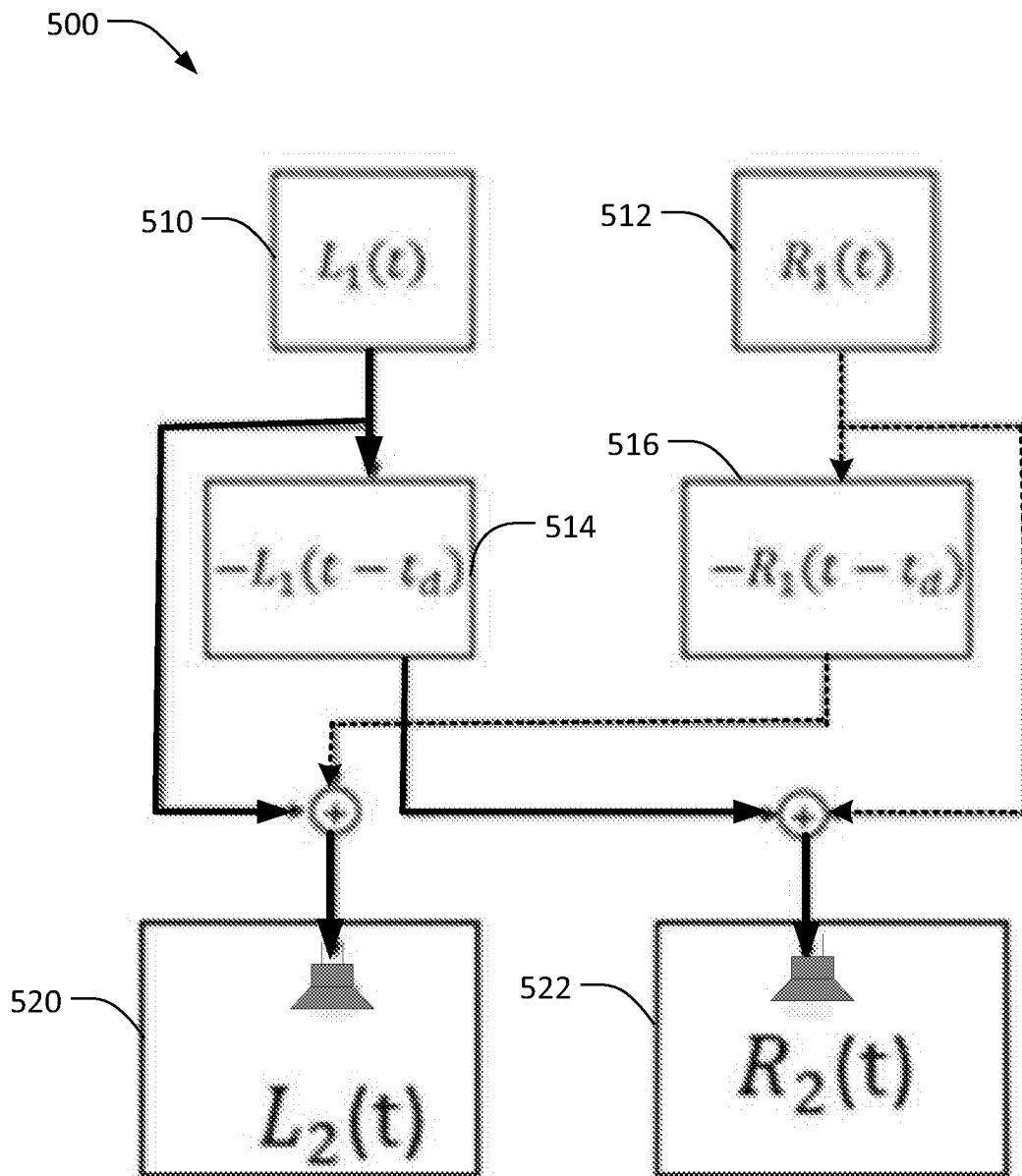


FIG. 1C



**FIG. 2****FIG. 3**

**FIG. 4A****FIG. 4B**

**FIG. 5**

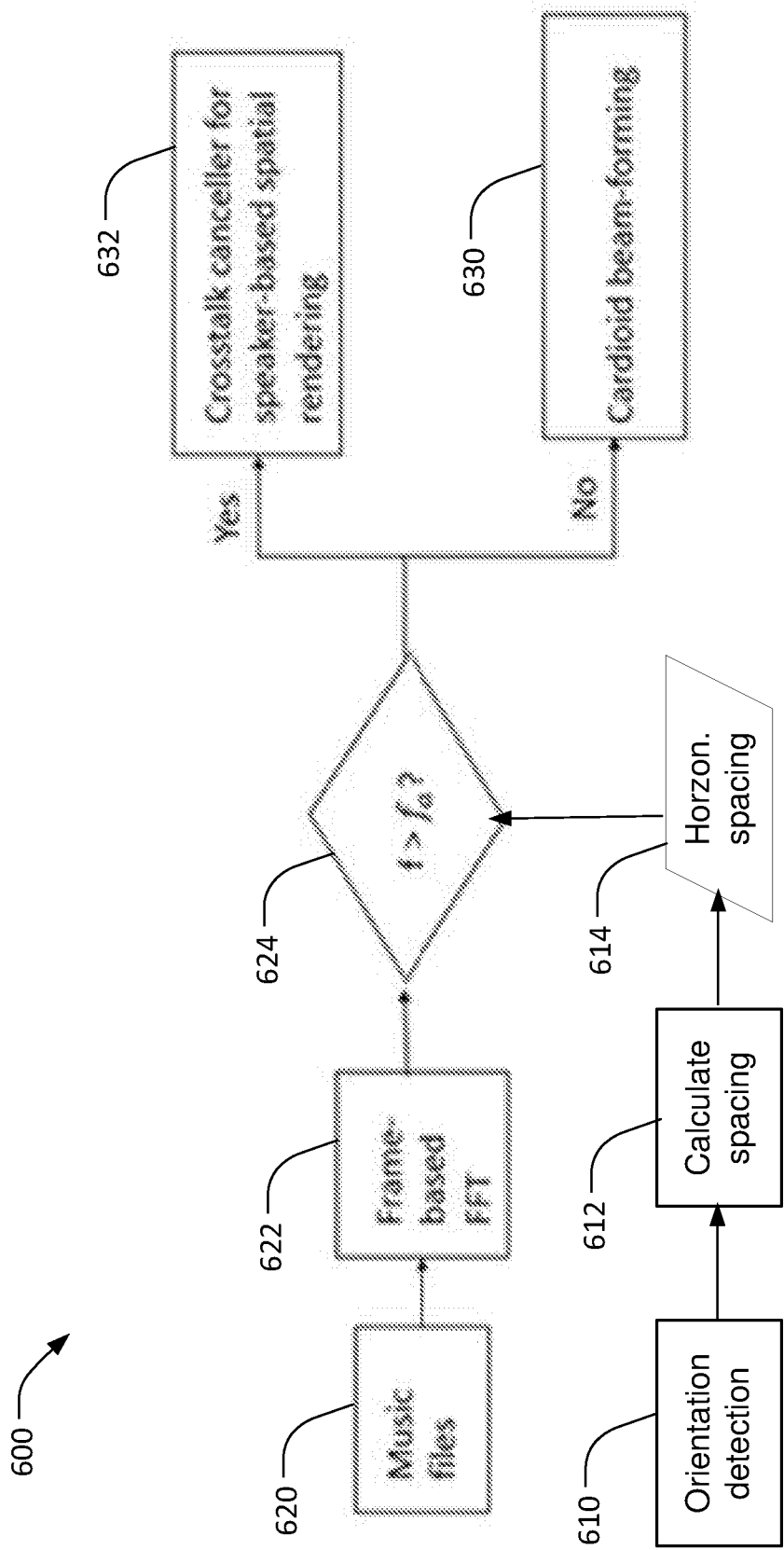
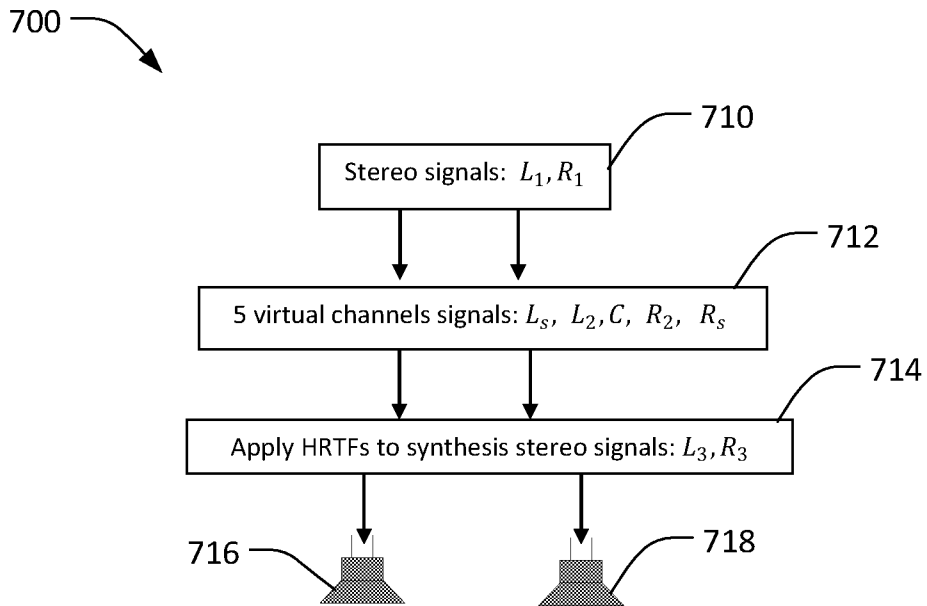
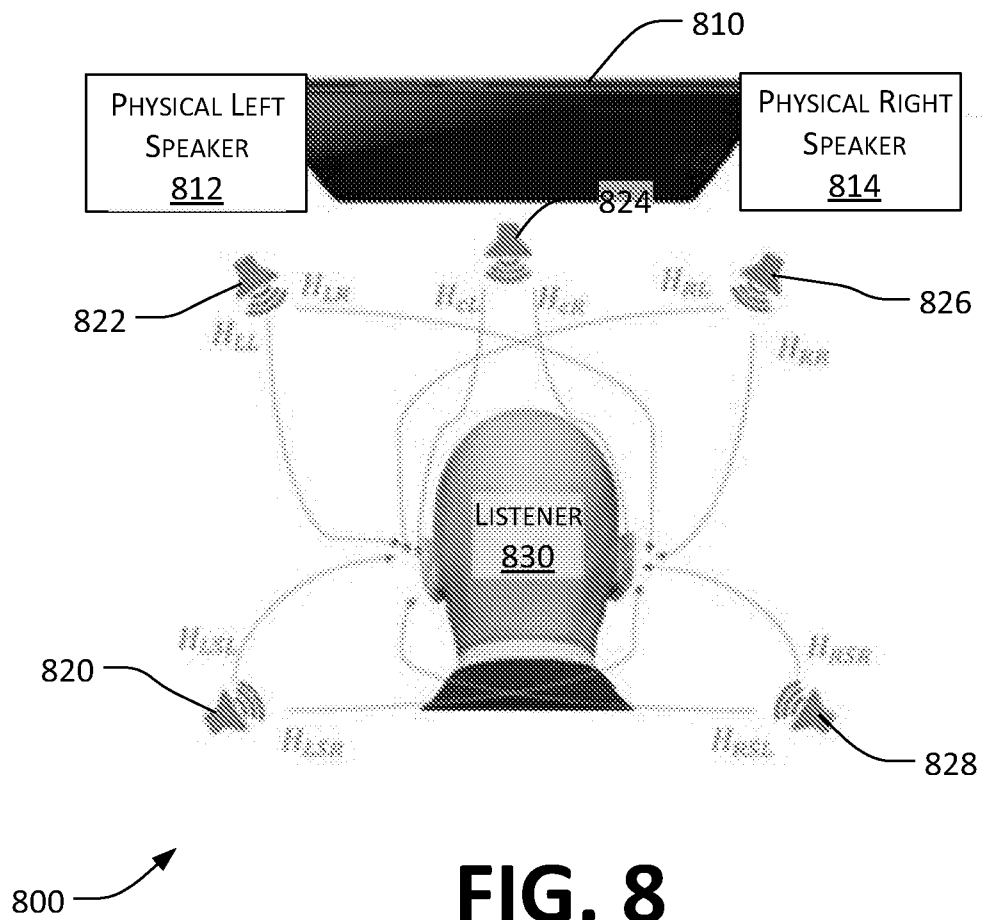
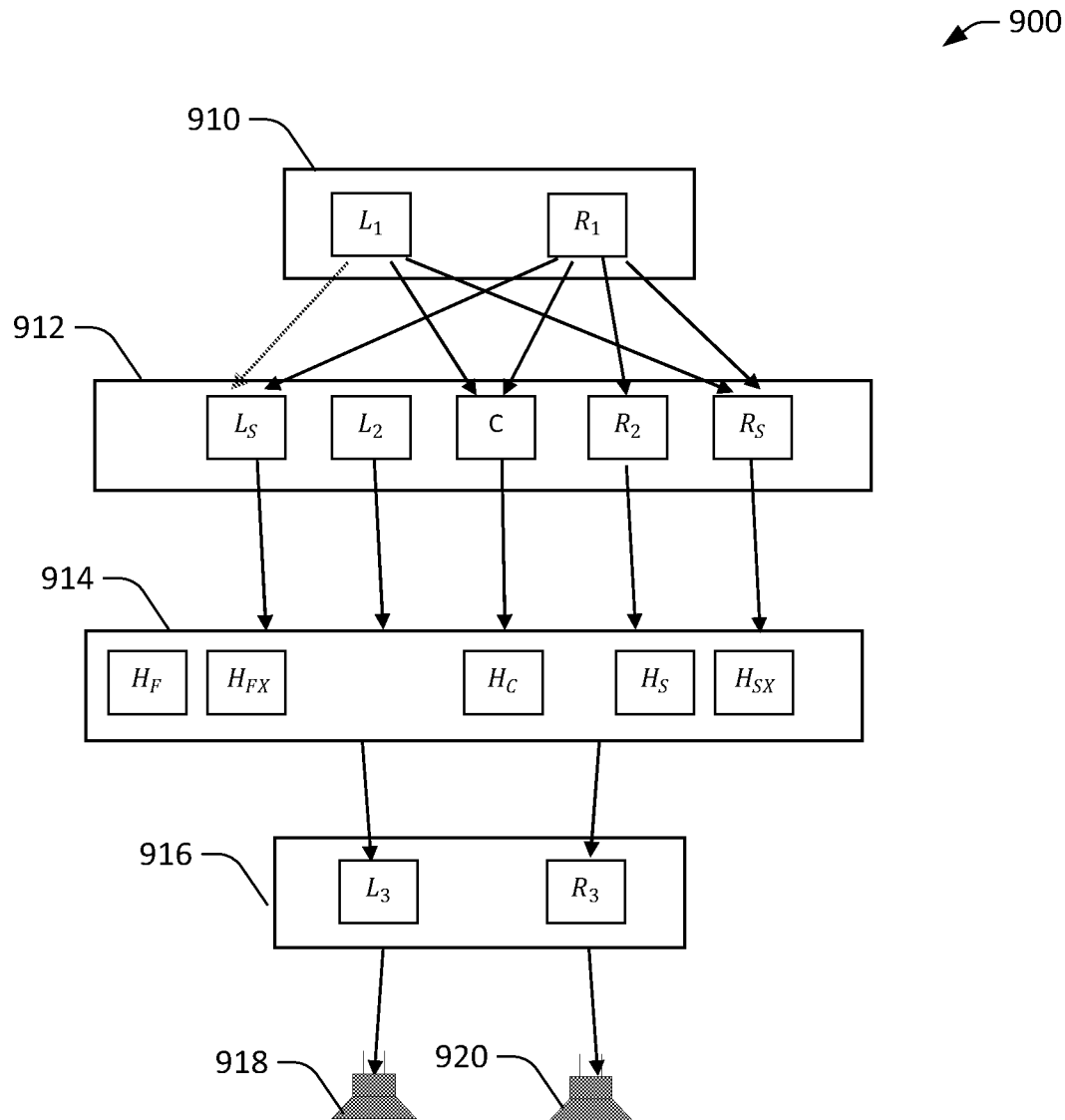


FIG. 6

**FIG. 7****FIG. 8**

**FIG. 9**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2017/027820

A. CLASSIFICATION OF SUBJECT MATTER		
H04R 5/02 (2006.01) H04S 1/00 (2006.01) According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
H04R 5/00-5/04, H04S 1/00, 7/00		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
PatSearch (RUPTO internal), USPTO, PAJ, K-PION, Esp@cenet, Information Retrieval System of FIPS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2016/033358 A1 (DOLBY LABORATORIES LICENSING CORPORATION) 03.03.2016, paragraphs [0003], [0034]-[0036], [0038]-[0040], [0051], [0074], [0080], [0084]	1-15
Y	JP 2010016573 A (KAWAMURA AKIO) 21.01.2010, abstract, paragraphs [0004]-[0007], fig. 2	1-15
Y	US 5870484 A (HAL GREENBERGER) 09.02.1999, col.11, lines 19-47, col.18, lines 1-30	1-7, 10-11, 13-14
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family	
"A" document defining the general state of the art which is not considered to be of particular relevance		
"E" earlier document but published on or after the international filing date		
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)		
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search	Date of mailing of the international search report	
18 December 2017 (18.12.2017)	21 December 2017 (21.12.2017)	
Name and mailing address of the ISA/RU: Federal Institute of Industrial Property, Berezhkovskaya nab., 30-1, Moscow, G-59, GSP-3, Russia, 125993 Facsimile No: (8-495) 531-63-18, (8-499) 243-33-37	Authorized officer T. Mukhina Telephone No. (499) 240-25-91	