



- (51) International Patent Classification:
H04S 7/00 (2006.01) *H04R 5/02* (2006.01)
- (21) International Application Number:
PCT/FI2016/050432
- (22) International Filing Date:
15 June 2016 (15.06.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
14/743,144 18 June 2015 (18.06.2015) US
- (71) Applicant: **NOKIA TECHNOLOGIES OY** [FI/FI];
Karaportti 3, 02610 Espoo (FI).
- (72) Inventor: **LAITINEN, Mikko-Ville**; Melkonkatu 9 A 6,
00210 Helsinki (FI).
- (74) Agents: **NOKIA TECHNOLOGIES OY** et al.; Ari
Aarnio, IPR Department, Karakaari 7, 02610 Espoo (FI).
- (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, 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, KN, KP, KR, 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).

Published:
— with international search report (Art. 21(3))

(54) Title: BINAURAL AUDIO REPRODUCTION

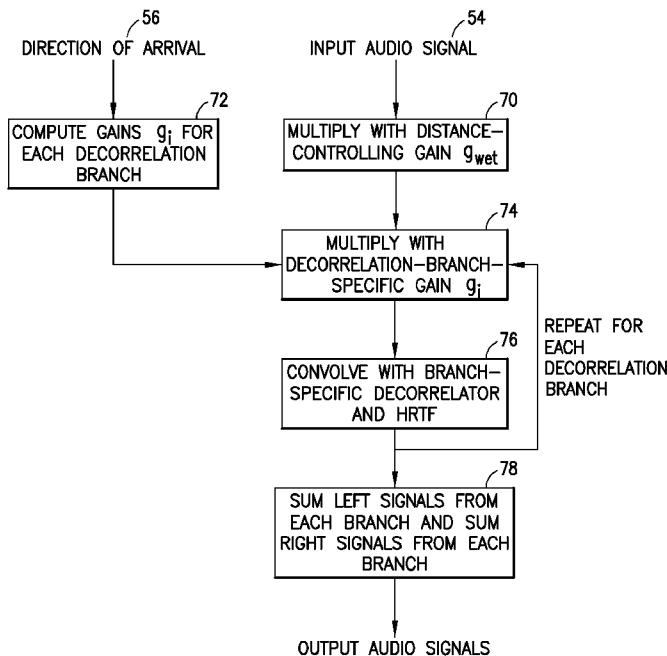


FIG.4

(57) Abstract: A method including providing an input audio signal in a first path and applying an interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path; providing the input audio signal in a second path, where the second path includes a plurality of filters and a respective amplifier for each filter, where the amplifiers are configured to be adjusted based upon the direction, and applying to an output from each of the filters a respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and combining the generated left signals to form a left output signal for a sound reproduction, and combining the generated right signals to form a right output signal for the sound reproduction.

WO 2016/203113 A1

Binaural Audio Reproduction

BACKGROUND

Technical Field

[0001] The exemplary and non-limiting embodiments relate generally to spatial sound reproduction and, more particularly, to use of decorrelators and head-related transfer functions.

Brief Description of Prior Developments

[0002] Spatial sound reproduction is known, such as which uses multi-channel loudspeaker setups, and such as which uses binaural playback with headphones.

SUMMARY

[0003] The following summary is merely intended to be exemplary. The summary is not intended to limit the scope of the claims.

[0004] In accordance with one aspect, an example method comprises providing an input audio signal in a first path and applying an interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path; providing the input audio signal in a second path, where the second path comprises a plurality of filters and a respective adjustable amplifier for each filter, where the amplifiers are configured to be adjusted based upon the direction, and applying to an output from each of the filters a respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and combining the generated left signals from the first and second paths to form a left output signal for a sound reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction..

[0005] In accordance with another aspect, an example embodiment is provided in an apparatus comprising a first audio signal path comprising an interpolated head-related transfer function (HRTF) pair applied to an input audio signal based upon
5 a direction configured to generate direction dependent first left and right signals in the first path; a second audio signal path comprising a plurality of: an adjustable amplifier configured to be adjusted based upon the direction; a filter for each adjustable amplifier, and a respective head-related
10 transfer function (HRTF) pair applied to an output from the filter, where the second path is configured to generate direction dependent second left and right signals for each filter in the second path, and where the apparatus is configured to combine the generated left signals from the first and second paths to
15 form a left output signal for a sound reproduction, and to combine the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

[0006] In accordance with another aspect, an example embodiment is provided in a non-transitory program storage device readable
20 by a machine, tangibly embodying a program of instructions executable by the machine for performing operations, the operations comprising: controlling, at least partially, a first audio signal path for an input audio signal comprising applying an interpolated head-related transfer function (HRTF) pair based
25 upon a direction to generate direction dependent first left and right signals in the first path; controlling, at least partially, a second audio signal path for the same input audio signal, where the second audio signal path comprises adjustable amplifiers configured to be set based upon the direction, applying outputs
30 from the amplifiers to respective filters for each of the amplifiers and applying to an output from each of the filters a respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and combining the generated left
35 signals from the first and second paths to form a left output

signal for a sound reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

BRIEF DESCRIPTION OF THE DRAWINGS

5 [0007] The foregoing aspects and other features are explained in the following description, taken in connection with the accompanying drawings, wherein:

[0008] Fig. 1 is a diagram illustrating an example apparatus;

10 [0009] Fig. 2 is a perspective view of an example of a headset of the apparatus shown in Fig. 1;

[0010] Fig. 3 is a diagram illustrating some of the functional components of the apparatus shown in Fig. 1;

[0011] Fig. 4 is a diagram illustrating an example method;

[0012] Fig. 5 is a diagram illustrating an example method; and

15 [0013] Fig. 6 is a diagram illustrating another example.

DETAILED DESCRIPTION OF EMBODIMENTS

20 [0014] Referring to Fig. 1, there is shown a front view of an apparatus 2 incorporating features of an example embodiment. Although the features will be described with reference to the example embodiments shown in the drawings, it should be understood that features can be embodied in many alternate forms of embodiments. In addition, any suitable size, shape or type of elements or materials could be used.

25 [0015] The apparatus 2 includes a device 10 and a headset 11. The device 10 may be a hand-held communications device which includes a telephone application, such as a smart phone for example. The device 10 may also comprise other applications including, for example, an Internet browser application, camera

application, video recorder application, music player and recorder application, email application, navigation application, gaming application, and/or any other suitable electronic device application. The device 10, in this example embodiment, 5 comprises a housing 12, a display 14, a receiver 16, a transmitter 18, a rechargeable battery 26, and a controller 20. The controller may comprise at least one processor 22, at least one memory 24, and software 28 in the memory 24. However, all of these features are not necessary to implement the features 10 described below. In an alternate example, the device 10 may be a home entertainment system, a computer such as used for gaming for example, or any suitable electronic device suitable to reproduce sound for example.

[0016] The display 14 in this example may be a touch screen 15 display which functions as both a display screen and as a user input. However, features described herein may be used in a display which does not have a touch, user input feature. The user interface may also include a keypad (not shown). The electronic circuitry inside the housing 12 may comprise a printed 20 wiring board (PWB) 21 having components such as the controller 20 thereon. The circuitry may include a sound transducer provided as a microphone and a sound transducer provided as a speaker and/or earpiece. The receiver 16 and transmitter 18 form a primary communications system to allow the apparatus 10 25 to communicate with a wireless telephone system, such as a mobile telephone base station for example.

[0017] The apparatus 10 is connected to a head tracker 13 by a link 15. The link 15 may be wired and/or wireless. The head tracker 13 is configured to track the position of a user's head. 30 In an alternate example, the head tracker 13 may be incorporated into the apparatus 10 and perhaps at least partially incorporated into the headset 11. Information from the head tracker 13 may be used to provide the direction of arrival 56 described below.

[0018] Referring also to Fig. 2, the headset 11 generally comprises a frame 30, a left speaker 32, and a right speaker 34. The frame 30 is sized and shaped to support the headset on a user's head. Please note that this is merely an example. As
5 another example, an alternative could be an in-ear headset or ear buds. The headset 11 is connected to the device 10 by an electrical cord 42. The connection may be a removable connection, such as with a removable plug 44 for example. In an alternate example, a wireless connection between the headset and
10 the device may be provided.

[0019] A feature as described herein is to be able to produce a perception of an auditory object in a desired direction and distance. The sound processed with features as described herein may be reproduced using the headset 11. Features as described
15 herein may use a normal binaural rendering engine together with a specific decorrelator engine. The binaural rendering engine may be used to produce the perception of direction. The decorrelator engine, consisting of several static decorrelators convolved with static head-related transfer functions (HRTF),
20 may be used to produce the perception of distance. Features may be provided with as little as two decorrelators. Any suitable number of decorrelators may be used, such as between 4-20 for example. Using more than about 20 might not be practical, since it increases computational complexity, and does not improve the
25 quality. However, there is no upper bound for the number of the decorrelators. The decorrelators may be any suitable filters which are configured to provide a decorrelator functionality. Each of the filters may be at least one of: a decorrelator, and a filter configured to provide a decorrelator functionality
30 wherein a respective signal is produced before applying the respective HRTF pair.

[0020] Head-related transfer functions (HRTF) are transfer functions measured in an anechoic chamber with the sound source at the desired direction and the microphones inside the ears.

There are a number of different ways to interpolate HRTFs. Creating interpolated HRTF filter pairs has been widely studied. For example, descriptions may be found in "Perceptual consequences of interpolating head-related transfer functions during spatial synthesis," by Elizabeth M. Wenzel and Scott H. Foster, in Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, pp. 102-105, October 1993; and "Interpolating between head - related transfer functions measured with low directional resolution," by Flemming Christensen, Henrik Møller, Pauli Minnaar, Jan Plogsties, and Søren Krarup Olesen, in Proceedings of the 107th AES Convention, New York, NY, USA, September 1999. For example, three HRTF pairs closest to the target direction may be selected from a HRTF database, and a weighted average of them may be computed separately for the left and the right ears. In addition, the corresponding impulse responses can be time-aligned before the averaging, and the inter-aural time differences (ITD) can be added after the averaging.

[0021] With features as described herein, the input signal may be convolved with these transfer functions, and the transfer functions are updated dynamically according to the head rotation of the user/listener. For example, if the auditory object is supposed to be in the front, and the listener turns her/his head to -30 degrees, the auditory object is updated to +30 degrees; thus remaining in the same position in the world coordinate system. As described below, a signal convolved with several static decorrelators convolved with static HRTFs causes ILD fluctuation, and the ILD fluctuation causes the externalized binaural sound. When the two engines are mixed in a suitable proportion, the result may provide a perception of an externalized auditory object in a desired direction.

[0022] Unlike past proposed use of decorrelators, and especially reverberators, for enhancing externalization, features as described herein propose use of a static

decorrelation engine comprising a plurality of static decorrelators. The input signal may be routed to each decorrelator after multiplication with a certain direction-dependent gain. The gain may be selected based on how close the relative direction of the auditory object is to the direction of the static decorrelator. As a result, interpolation artifacts, when rotating a listener's head, are avoided while still having some directionality for the decorrelated content; which was found to improve the quality. In addition, unlike proposed reverbetor-based methods, features as described herein do not cause a prominent perception of added reverberation.

[0023] Referring also to Fig. 3, a block diagram of an example embodiment is shown. The circuitry of this example is on the printed wiring board 21 of the device 10. However, in alternate example embodiments one or more of the components might be on the headset 11. In the example shown the components form a binaural rendering engine 50 and a decorrelator engine 52. An input audio signal 54 may be provided from a suitable source such as, for example, a sound recording stored in the memory 24, or from signals received by the receiver 16 by a wireless transmission. Please note that these are only examples. With features as described herein, any suitable signals can be used as an input, such as arbitrary signals for example. For example, input signals which could be used with features as described herein can include mono recordings of guitar, or speech, or any signals. In addition to the input audio signal, a direction of arrival indication of the sound is supplied to the two engines 50, 52 as indicated by 56. Thus, the inputs comprise one mono audio signal 54 and the relative direction of arrival 56.

[0024] In this example the path for the binaural rendering engine 50 includes a variable amplifier g_{dry} , and the path for the decorrelator engine 52 includes a variable amplifier g_{wet} . The gain provided by these amplifiers for the "dry" and the "wet" paths can be selected based on how "much" externalization is

desired. Basically, this affects the perceived distance of the auditory object. In practice, it has been noticed that good values include $g_{\text{dry}}=0.92$ and $g_{\text{wet}}=0.18$ for example. Please note that these are merely examples and should not be considered as limiting. As can be seen from the above, gain of the amplifiers can also be smaller than 1. Thus, "amplifying" is actually "attenuation" in that case.

[0025] The relative direction of arrival may be determined based on the desired direction in the world coordinate system, and the orientation of the head. The upper path of the diagram is a simply normal binaural rendering. A set of head-related transfer functions (HRTF) may be provided in a database in the memory 24, and the resulting HRTF may be interpolated based on the desired direction. Thus, for the first path provided by the engine 50, the input audio signal 54 may be convolved with the interpolated HRTF as indicated by 55. An HRTF is a transfer function that represents the measurement for one ear only (i.e. either the right ear only or the left ear only). The directionality requires both the right ear HRTF and the left ear HRTF. Thus, for a given direction, one requires an HRTF pair, and after interpolation 55 there are two paths. The direction of arrival 56 is introduced by the HRTF pair, and the HRTF filter comprises the respective pair.

[0026] The lower path in the block diagram of Fig. 3 shows the other engine 52 which forms a second different path from the first path of the first engine 50. The input audio signal 54 is routed to a plurality of decorrelators 58. The decorrelated signals are convolved with pre-determined HRTFs 68, which may be selected to cover the whole sphere around the listener. In one example, a suitable number of the decorrelator paths is twelve (12). However, this is merely an example. More or less than twelve decorrelators 58 may be provided, such as between about 6 and 20 for example.

[0027] Each decorrelator path has an adjustable amplifier g_1, g_2, \dots, g_i , located before its respective decorrelator 58. Gain of the amplifiers may be smaller than 1. Thus, amplifying is actually attenuation in that case. The amplifiers g_i are adjusted as computed by 60 which is based upon the direction of arrival signal 56. The gain g_i for each decorrelator path may be selected based on the direction of the source as follows

$$g_i = 0.5 + 0.5 (S_x D_{x,i} + S_y D_{y,i} + S_z D_{z,i})$$

where $\mathbf{s}=[S_x \ S_y \ S_z]$ is the direction vector of the source and $\mathbf{D}_i=[D_{x,i} \ D_{y,i} \ D_{z,i}]$ is the direction vector of the HRTF in the decorrelator path i . The decorrelators 58 can basically be any kind of decorrelator (e.g., different delays at different frequency bands).

[0028] In the example shown in Fig. 3, one input goes in and one output comes out from each decorrelator. These decorrelators may be designed in a nested structure so that one can have one block comprising all decorrelators and within this one block the same functionality can be provided. One could pre-convolve the decorrelator and the HRTF, and sum them together, after weighting them, based on the computed input gains (g_1-g_N). Then the input signal may be convolved with this filter. The output should be identical to the implementation shown in Fig. 3. In the case of a single source, Fig. 3 may be computationally the most efficient implementation.

[0029] In one example embodiment a pre-delay in the beginning of the decorrelator may be provided. Adding a pre-delay in the beginning of the decorrelator may be useful. The reason for the pre-delay is to mitigate the effect of the decorrelated signals to the perceived direction. This delay may be at least 2 ms for example. This is approximately the time instant when the summing localization ends and the precedence effect starts. As a result, the directional cues provided by the "dry" path dominate the perceived direction. The delay can be also less than 2 ms. The

optimal quality may be obtained using the value of at least 2 ms, but the method could be used with smaller values. For the first 2 ms after the first wavefront, the directions of the secondary wavefronts (whether they are real reflections or reproduced with loudspeakers or headphones or anything) affect the perceived direction. After 2 ms, the directions of the secondary wavefronts do not affect the perceived direction, they merely affect the perceived spaciousness and the apparent width of the sources. Hence, in order to minimally the perceived affect to the directions of the sources, the decorrelated paths may include this 2 ms delay. However, as noted above the method may work also with shorter delays. Nevertheless, adding the pre-delay is not required, especially since the decorrelators typically have some inherent delay, although it is potentially useful. For example, even a delay of 0 ms could be used because the decorrelators have some inherent delay (The decorrelators are essentially all pass filters, so they must have an impulse response longer than just one impulse). Thus, adding some additional delay, such as 2 ms, may be provided, but it is not required.

[0030] It should be noted that the number of decorrelator paths affects the suitable value for g_{wet} . In the end of the processing, the signals of the dry path and the wet paths are summed together as indicated by 62, yielding one signal 64 for left channel and one signal 66 for right channel. These signals can be reproduced using the speakers 32, 34 of the headphones 11. Furthermore, the ratio between g_{dry} and g_{wet} affects the perceived distance. Thus, controlling the amplifiers g_{dry} and g_{wet} can be used for controlling the perceived distance.

[0031] Features as described herein may be used in the field of spatial sound reproduction. In this field, the aim is to reproduce the perception of spatial aspects of a sound field. These include the direction, the distance, and the size of the

sound source, as well as properties of the surrounding physical space.

5 [0032] Human hearing perceives the spatial aspects using the two ears of the listener. So, if a suitable sound pressure signal is reproduced at the eardrums, the perception of spatial aspects should be as desired. Headphones are typically used for reproducing the sound pressure at the ears.

10 [0033] One would expect that recording the sound field using microphones inside the ears would provide good spatial cues. However, it does not allow the listener to rotate the head while listening. The lack of dynamic spatial cues is known to cause front-back confusions and lack of externalization. In addition, for example in virtual-reality applications, the listener has to be able to look around while having the perceived sound field static in the world coordinate system; which using microphones 15 inside the ears does not allow.

[0034] In theory, the binaural playback should produce a perception of an auditory object that is at the desired direction and distance. However, conventionally this does not typically 20 happen. The direction of the auditory object might be correct, but it is often perceived to be very close to the head or even inside the head (called internalization). This is contrary to the aim of a realistic, externalized, auditory object.

[0035] For head-related transfer functions (HRTF), in theory 25 the direction and the distance should match the measured ones. However, conventionally this does not happen, and instead, there is a perceived lack of externalization (the sound sources are perceived to be very close or inside the head). The reason for this lack of externalization is that the human hearing uses 30 direct-to-reverberant ratio (D/R ratio) as a cue for distance. Obviously, anechoic responses do not have these cues. As HRTF rendering cannot, in conventional practice, reproduce the sound pressure fully accurately to the ears, human hearing typically

interprets these sound sources as internalized or very close sources.

5 [0036] One solution to problems with HRTFs is to instead use binaural room impulse responses (BRIR). These are measured in a same way as HRTFs, but in a room. They provide externalization due to the presence of the D/R-ratio cues. However, there are some drawbacks. They always add the perception of reverberation of the room where they were measured; which is not typically desired. Second, the responses might be long which causes
10 computational complexity. Third, the perceived distance is locked to the distance where the responses where measured. If multiple distances are desired, all responses have to be measured at multiple distances, which can be time consuming, and the size of the database of the responses grows fast. Lastly, the
15 interpolation (when the listener rotates the head) between different responses can cause artifacts, such as changes in the timbre and a perception of frequency-changing comb filter. An alternative to BRIRs is to simulate the reflections and render them with HRTFs. However, the same problems are largely present
20 (the perception of added reverberation, interpolation artifacts, and computational complexity). Methods of adding reverberation to the HRTFs, and to use head tracking, suffer from the problems that were identified. Features as described herein may be used to avoid these problems.

25 [0037] The fluctuation of ILD is a process inside the auditory system. With features as described herein, audio signals may be created which cause this fluctuation of the ILDs. The fluctuation of inter-aural level differences (ILD) may be used for the perception of externalized binaural sound. This ILD
30 fluctuation is the reason why reverberation helps in externalization. Thus, it can also be assumed that reverberation itself is not necessarily needed for externalization; it is simply enough to cause proper ILD fluctuation. With features as

described herein, a method may be provided that can create this ILD fluctuation without unwanted side effects.

[0038] Similar problems are present in other fields of spatial audio, such as in systems capturing and reproducing sound fields. These systems also use decorrelation and reverberation strategies for improving externalization with binaural rendering. For example, the binaural implementation for directional audio coding (DirAC) uses decorrelators. However, the scope of these two techniques is different. With features as described herein, arbitrary mono signals may be positioned to desired directions and distances, whereas binaural DirAC attempts to recreate the perception of the sound field in the recording position using recorded B-format signals. Binaural DirAC also performs time-frequency analysis, extracts the "diffuse" (or "reverberant") components from the captured signals, and applies decorrelation on the extracted diffuse components. Features as described herein do not require such processing.

[0039] Referring also to Fig. 4, a diagram of an example method is shown. Fig. 4 generally corresponds to the "wet" signal path shown in Fig. 3. The input audio signal 54 and the direction of arrival 56 are provided. The input audio signal 54 is multiplied with a distance controlling gain g_{wet} as indicated by block 70. Gains g_i are computed for each decorrelation branch as indicated by block 72. As indicated by block 74, the output from multiplication 70 is multiplied with a decorrelation-branch-specific gain g_i , and convolved with a branch-specific decorrelator 58 and HRTF 68. The output from the branches are then summed as indicated by 78 and 62 in Fig. 3.

[0040] The method improves the typical binaural rendering by providing externalization which is much better, repeatable, and adjustably correct than conventional methods. In addition, this is achieved without a prominent perception of added reverberation. Importantly, the method was found not to cause

any interpolation artifacts for the decorrelated signal path. The interpolation artifacts are avoided because the decorrelated signals are statically reproduced from the same directions. Only the gain for each decorrelator is changed, and this may be
5 changed smoothly. As the decorrelator outputs are mutually incoherent, changing the levels of the input signal for them does not cause significant timbre changes; preventing interpolation artifacts for the wet signal path.

[0041] In addition, the method is relatively efficient
10 computationally. Only the decorrelators are somewhat heavy to compute. Moreover, if the method is a part of a spatial sound processing engine that uses decorrelators and HRTFs anyway, the processing is computationally very efficient; only a few multiplications and additions are required.

[0042] Although the perception of added reverberation might
15 not be fully avoided, especially if the source is desired to be very far away, audio sources which are very far are rarely completely anechoic. In addition, the level of perceived reverberation is assumed to be significantly lower than with
20 typical solutions.

[0043] In virtual-reality (VR) applications, the sound is typically reproduced using headphones. The reason for this is that the video is reproduced using head-mounted displays. As the video is seen by only one individual at a time, it makes
25 sense that also the audio is heard by only that individual. In addition, as VR content may have visual and auditory content all around the subject, loudspeaker reproduction would require setups with large number of loudspeakers. Thus, headphones are the logical option for spatial-sound reproduction in such
30 applications.

[0044] Spatial audio is often delivered in multi-channel format (such as 5.1 or 7.1 audio for example). Thus, there is a need for a system that can render these signals using headphones so

that they are perceived as if they were reproduced in a good listening room with a corresponding loudspeaker setup. Such a system can be implemented using the features as described herein. The input to the system can include the multi-channel audio signals, the corresponding loudspeaker directions, and the head-orientation information. The head orientation is typically obtained automatically from a head-mounted display. The loudspeaker setup is often available in the metadata of the audio file, or it can be pre-defined.

10 **[0045]** Each audio signal of the multi-channel file may be positioned to the direction determined by the loudspeaker setup. Moreover, when the subject rotates her/his head, these directions may be rotated accordingly; in order to keep them in the same positions in the world coordinate system. The auditory objects may be positioned to suitable distances. When these features of auditory reproduction are combined with head-tracked stereoscopic visual reproduction, the result is very natural perception of the reproduced world around. The output of the system is an audio signal for each channel of the headphones. 15 These two signals can be reproduced with normal headphones. Other use cases can easily be derived for the VR context. For example, the features could be used for positioning auditory objects to arbitrary directions and distances in real time. The directions and the distances could be obtained from the VR rendering engine. 25

[0046] With features as described herein, single monophonic sources may be processed separately. Obviously, these monophonic sources may realize a multi-channel signal when put together, but it is not required in the method. They can be fully independent sources. This is unlike conventional processes where either multi-channel signals (e.g., 5.1 or stereo) are processed, or somehow combined processed signals are processed. 30

[0047] Features as described herein also proposes to enhance externalization by applying fixed decorrelators. This may be used to avoid any interpolation artifacts when the system is combined with head tracking (which requires to rotate auditory objects as a function head orientation). This is unlike conventional methods where there is no specific processing of signals for head tracking; the directions of the sources are simply rotated. Thus, conventionally all components of the processing require rotation, and this rotation needs interpolation, which potentially causes artifacts. With features as described herein, these interpolation artifacts are avoided by not rotating decorrelated components and, instead, having fixed decorrelators with direction-dependent input gains.

[0048] Features as described herein do not require decreasing the coherence between loudspeaker channels of multi-channel audio files. Instead, features may comprise decreasing the coherence between resulting headphone channels. Moreover, mono audio files may be used instead of multi-channel audio files. Conventional methods do not take head tracking into account and, thus, direct interpolation would be required in the case of head tracking. Features as described herein, on the other hand, provide an example system and method to take the head tracking into account, and to avoid interpolation by having the fixed decorrelators.

[0049] In one type of conventional system, the aim is to extract multiple auditory objects from a stereo downmix and to render all these objects with headphones. Decorrelation is needed in this context in case there are more independent components in the same time-frequency tile than there are downmix signals. In this case the decorrelator creates incoherence to reflect the perception of multiple independent sources. Features as described herein does not need to include this kind of processing. It simply aims to render single audio signals by decreasing the resulting inter-aural coherence in order to

enhance externalization. Features as described herein also use multiple decorrelators, and each output is convolved with a dedicated HRTF. Each auditory object may be processed separately. These features create a better perception of envelopment, and the decorrelated signal path has a perceivable direction. These properties yield a perception of higher audio quality.

[0050] An example method comprises providing an input audio signal in a first path and convolving with an interpolated first head-related transfer function (HRTF) based upon a direction; providing the input audio signal in a second path, where the second path comprises a plurality of branches comprising respective decorrelators in each branch and an amplifier in each branch adjusted based upon the direction, and applying to a respective output from each of the decorrelators respective second head-related transfer functions (HRTF); and combining outputs from the first and second paths to form a left output signal and a right output signal.

[0051] The method may further comprise selecting a first gain to be applied to the input audio signal at a start of the first path and a second gain to be applied to the input audio signal at a start of the second path based upon a desired externalization. The method may further comprise selecting respective different gains to be applied to the input audio signal before the decorrelators. The respective different gains may be selected based, at least partially, upon the direction. The decorrelators may be static decorrelators and where the second head-related transfer function (HRTF) are static HRTF. Outputs from the first path may comprise a left output signal and a right output signal from the first head-related transfer function (HRTF), and where the outputs from the second path comprise a left output signal and a right output signal from each of the second head-related transfer functions (HRTF).

[0052] An example apparatus may comprise a first audio signal path comprising an interpolated first head-related transfer function (HRTF) configured to convolute the input audio signal based upon a direction; a second audio signal path comprising a plurality of branches, each branch comprising: an adjustable amplifier configured to be adjusted based upon the direction; a decorrelator, and a respective second head-related transfer function (HRTF), where the apparatus is configured to combine outputs from the first and second paths to form a left output signal and a right output signal.

[0053] The first audio signal path may comprise a first variable amplifier before the first head-related transfer function (HRTF), where the second audio signal path comprises a second variable amplifier before the decorrelators, and the apparatus comprises an adjuster to adjust a desired externalization by based upon adjusting the first and second variable amplifiers. The apparatus may further comprise a selector connected to the adjustable amplifiers, where the adjuster is configured to adjust the adjustable amplifiers based, at least partially, upon the direction. The decorrelators may be static decorrelators and where the second head-related transfer function (HRTF) are static HRTF. The first head-related transfer function (HRTF) may be configured to generate a first path left output signal and a first path right output signal, and where each of the second head-related transfer functions (HRTF) are configured to generate a second path left output signal and a second path right output signal.

[0054] An example non-transitory program storage device may be provided, such as memory 24 for example, readable by a machine, tangibly embodying a program of instructions executable by the machine for performing operations, the operations comprising controlling, at least partially, first outputs from a first audio signal path from an input audio signal comprising convolving with an interpolated first head-related transfer function (HRTF)

based upon a direction; controlling, at least partially, second outputs from a second audio signal path from the same input audio signal, where the second audio signal path comprises branches, comprising amplifying the input audio signal in each branch based upon the direction, decorrelating by a decorrelator and applying to a respective output from each of the decorrelators a respective second head-related transfer function (HRTF) filtering; and combining the outputs from the first and second audio signal paths to form a left output signal and a right output signal.

[0055] The operations may further comprise selecting a first gain to be applied to the input audio signal at a start of the first path and a second gain to be applied to the input audio signal at a start of the second path based upon a desired externalization. The operations may further comprise selecting respective different gains to be applied to the input audio signal before the decorrelators. The respective second head-related transfer function (HRTF) filtering may comprise use of static head-related transfer function (HRTF) filters. The operations may further comprise outputs from the first path comprising a left first path output signal and a right first path output signal from the first head-related transfer function (HRTF), and where the outputs from the second path comprise a left second path output signal and a right second path output signal from each of the second head-related transfer function (HRTF) filtering.

[0056] Any combination of one or more computer readable medium(s) may be utilized as the memory. The computer readable medium may be a computer readable signal medium or a non-transitory computer readable storage medium. A non-transitory computer readable storage medium does not include propagating signals and may be, for example, but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device, or any suitable

combination of the foregoing. More specific examples (a non-exhaustive list) of the computer readable storage medium would include the following: an electrical connection having one or more wires, a portable computer diskette, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM or Flash memory), an optical fiber, a portable compact disc read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any suitable combination of the foregoing.

10 [0057] An example apparatus may be provided comprising means for providing an input audio signal in a first path and applying an interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path as indicated by block 80; means
15 for providing the input audio signal in a second path as indicated by block 82, where the second path comprises a plurality of filters and a respective adjustable amplifier for each filter, where the amplifiers are configured to be adjusted based upon the direction, and means for applying to an output
20 from each of the filters a respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and combining the generated left signals from the first and second paths as indicated by block 84 to form a left output signal for
25 a sound reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

[0058] In one example embodiment, for the dry path shown in Fig 3, there a HRTF database may be provided containing 36 HRTF
30 pairs. Using the HRTF database and the direction of arrival, the method may create one interpolated HRTF pair (such as using Vector Base Amplitude Panning (VBAP) so it is a weighted sum of three HRTF pairs selected by the VBAP algorithm). The input signal may be convolved with this one interpolated HRTF pair.

For the wet path, there another HRTF database may be provided containing 12 HRTF pairs. These HRTF pairs are fixed to the different branches of the wet path (i.e., HRTF1, HRTF2, ..., HRTF12). For this example embodiment the input signal is always convolved with all these HRTF pairs after the gains and the decorrelators. The HRTF database of the wet path may be a subset of the HRTF database of the dry path in order to avoid having multiple databases. However, from the algorithm point of view, it could equally well be a completely different database.

5
10 [0059] In the examples described above, HRTF pairs have been mentioned. It is a transfer function which is transformed from head related impulse responses (HRIRs). Direction dependent impulse response measurements for each ear can be obtained on an individual or using a dummy head for example. A database can be
15 formed with HRTFs, as also mentioned above. In alternative embodiments, one could introduce localization cues rather than introducing the entire HRTF pairs. These localization cues can be extracted from respective HRTF pairs. Put another way, an HRTF pair can possess these direction dependent localization
20 cues already. So, the method could process input signals to introduce desired directionalities in order to simulate the effect of HRTF pairs. A mapping table could contain these localization cues as a function of direction. The method may be used with "simplified" HRTFs containing only the localization
25 cues, such as interaural time difference (ITD) and interaural intensity difference (ILD). Thus, HRTFs referred to herein may comprises these "simplified" HRTFs. Adding ITD and frequency-dependent ILD is a form of HRTF filtering, although a very simple form. Related to the HRTF pairs, these HRTFs may be obtained
30 using measurements by measuring right and left ear impulse responses as a function of sound source position relative to the head position where direction dependent HRTF pairs are obtained from measurements. The HRTF pairs may be obtained by numerical models (simulations). Simulated HRIR or HRTF pairs would work
35 equally well as the measured ones. Simulated HRIR or HRTF pairs

might even be better due to absence of the potential measurement noise and errors.

[0060] Fig. 3 presents an example implementation using a block diagram for simplicity. The first and second path (dry and wet) are basically trying to form respective ear signals for sound reproduction. The functionality of the blocks shown in Fig. 3 could be drawn in other ways. Basically the exact shape of Fig. 3 is not essential for the method/functionality. This would have one interpolation (or panning) computation and two convolutions for the dry path, and 12 decorrelations and 24 convolutions for the wet path. And in the end, all 13 signals would be summed from the left ear and all 13 signals would be summed for the right ear. In the case of multiple simultaneous sources (e.g., 10), other kinds of implementations can be more efficient. One example implementation has fixed HRTFs. The dry signal path (using VBAP) may create three weighted signals with routing to HRTF pairs computed with VBAP. This process is repeated for all sources. The wet signal path creates 12 weighted signals. This process is repeated for each source and the signals are summed together. The decorrelation can be applied once to all signals (i.e., 12 decorrelations). In the end, the dry and the wet signals from all the sources are summed together for the corresponding HRTF and convolved with corresponding HRTF pairs. Thus, the HRTF filtering is performed only once (but potentially for many HRTF pairs if the sources are at different directions).

[0061] It should be noted that the output of both implementations described above would be identical. In which order one performs different operations affects the computation efficiency, but the output is the same. The operations (convolution, sum, and multiplication) are linear, so they can be freely rearranged without changing the output.

[0062] In virtual-reality (VR) applications, the sound is typically reproduced using headphones, and the video is reproduced using a head-mounted display. As the video is seen

by only one individual at a time, it makes sense that also the audio be heard by only that individual. In addition, as VR content may have visual and auditory content all around the subject, a loudspeaker reproduction would require setups with
5 large number of loudspeakers. Thus, headphones are the logical option for spatial-sound reproduction in such applications.

[0063] Spatial audio is often delivered in multi-channel format (such as 5.1 or 7.1 audio). Features as described herein may render these signals using headphones so that they are perceived
10 as if they were reproduced in a good listening room with a corresponding loudspeaker setup. The input to the system may be the multi-channel audio signals, the corresponding loudspeaker directions, and the head-orientation information. The head orientation may be obtained automatically from the head-mounted
15 display. The loudspeaker setup is often available in the metadata of the audio file, or it can be pre-defined.

[0064] Referring also to Fig. 6, an example for rendering multi-channel audio files, such as for VR for example, is shown. Each loudspeaker signal (1, 2, ... N) has a binaural renderer 100.
20 Each binaural renderer 100 may be as shown in Fig. 3 for example. Thus, Fig. 6 illustrates an embodiment having plurality of the devices shown in Fig. 3. The input to each binaural renderer 100 includes the respective audio signal $102_1, 102_2, \dots, 102_N$, and a rotational direction signal $104_1, 104_2, \dots, 104_N$. The rotational
25 direction signals $104_1, 104_2, \dots, 104_N$ are determined based upon a channel direction signal $106_1, 106_2, \dots, 106_N$ and a head direction signal 108. The left and right outputs from the binaural renderers 100 are summed at 110 and 112 to form the left headphone signal 64 and the right headphone signal 66.

[0065] Features as described herein may be used to position each audio signal of the multi-channel file to the channel direction similar to determined by the loudspeaker setup. Moreover, when the subject rotates her/his head, these directions may be rotated accordingly in order to keep them in
30

the same positions in the world coordinate system. The auditory objects may also be positioned to suitable distances. When these features of auditory reproduction are combined with head-tracked stereoscopic visual reproduction, the result is very natural perception of the reproduced world around. The output of the system is an audio signal for each channel of the headphones. These two signals can be reproduced with normal headphones.

[0066] Also, other use cases can easily be derived for the present invention in the VR context. For example, features could be used for positioning auditory objects to arbitrary directions and distances in real time. The directions and the distances could be obtained from the VR rendering engine.

[0067] Referring also to Fig. 5, an example method may comprise providing an input audio signal in a first path and applying an interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path as indicated by block 80; providing the input audio signal in a second path as indicated by block 82, where the second path comprises a plurality of filters and a respective adjustable amplifier for each filter, where the amplifiers are configured to be adjusted based upon the direction, and applying to an output from each of the filters a respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and combining the generated left signals from the first and second paths as indicated by block 84 to form a left output signal for a sound reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

[0068] The method may further comprise selecting respective different gains to be applied by the amplifiers to the input audio signal before the filters. The filters may be static decorrelators and the head-related transfer functions (HRTF) pairs of the second path may be static HRTF pairs. The method

may further comprise setting the adjustable amplifiers in the second path at different settings relative to one another based upon the direction. Applying the interpolated head-related transfer function (HRTF) pair to the input audio signal in the first path may comprise convolving the interpolated head-related transfer function (HRTF) pair to the input audio signal in the first path based upon the direction. The method may be applied to a plurality of respective multi-channel audio signals as shown in Fig. 6 as the input audio signal at a same time, and where a plurality of left signals and right signals from the respective multi-channel audio signals are combined for the sound reproduction.

[0069] An example apparatus may comprise a first audio signal path comprising an interpolated head-related transfer function (HRTF) pair applied to an input audio signal based upon a direction configured to generate direction dependent first left and right signals in the first path; a second audio signal path comprising a plurality of: an adjustable amplifier configured to be adjusted based upon the direction; a filter for each adjustable amplifier, and a respective head-related transfer function (HRTF) pair applied to an output from the filter, where the second path is configured to generate direction dependent second left and right signals for each filter in the second path, and where the apparatus is configured to combine the generated left signals from the first and second paths to form a left output signal for a sound reproduction, and to combine the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

[0070] The apparatus may further comprise a selector connected to the adjustable amplifiers, where the adjuster is configured to adjust the adjustable amplifiers to different respective settings based, at least partially, upon the direction. The filters may be static decorrelators and where the head-related transfer function (HRTF) pairs of the second audio signal path

are static. The first audio signal path may be configured to convolve the interpolated head-related transfer function (HRTF) pair to the input audio signal based upon the direction. The apparatus comprises a plurality of pairs of the first and second paths as illustrated by Fig. 6, and where the apparatus is configured to apply a respective multi-channel audio signal to a respective one of the pairs of the first and second paths as the input audio signal at a same time, and where a plurality of left signals and right signals from the respective multi-channel signals are combined for the sound reproduction.

[0071] An example apparatus may be provided in a non-transitory program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine for performing operations, the operations comprising: controlling, at least partially, a first audio signal path for an input audio signal comprising applying an interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path; controlling, at least partially, a second audio signal path for the same input audio signal, where the second audio signal path comprises adjustable amplifiers configured to be set based upon the direction, applying outputs from the amplifiers to respective filters for each of the amplifiers and applying to an output from each of the filters a respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and combining the generated left signals from the first and second paths to form a left output signal for a sound reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

[0072] Features as described above have been primarily described with regard to headset sound reproduction. However, features could also to used for non-headset reproduction

including loudspeaker playback for example. A feature of the method as described herein is to avoid the interpolation artifacts when the head of a user is rotated. In the case of the loudspeaker playback that is not an issue since there is no head tracking in loudspeaker playback, but there is no reason why it could not be applied to the loudspeaker playback. Thus, the method can be easily adapted to loudspeaker playback. The interpolated HRTFs (in the dry path) may be replaced by loudspeaker-based positioning (such as amplitude panning, ambisonics, or wave-field synthesis), and the fixed HRTFs (in the wet path) may be replaced by actual loudspeakers.

[0073] It should be understood that the foregoing description is only illustrative. Various alternatives and modifications can be devised by those skilled in the art. For example, features recited in the various dependent claims could be combined with each other in any suitable combination(s). In addition, features from different embodiments described above could be selectively combined into a new embodiment. Accordingly, the description is intended to embrace all such alternatives, modifications and variances which fall within the scope of the appended claims.

CLAIMS

What is claimed is:

1. A method comprising:

5 providing an input audio signal in a first path and applying an interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path;

10 providing the input audio signal in a second path, where the second path comprises a plurality of filters and a respective adjustable amplifier for each filter, where the amplifiers are configured to be adjusted based upon the direction, and applying to an output from each of the filters a respective head-related transfer function (HRTF)
15 pair to generate direction dependent second left and right signals for each filter in the second path; and

20 combining the generated left signals from the first and second paths to form a left output signal for a sound reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

2. A method as in claim 1, further comprising, based upon a desired externalization, selecting a first gain to be applied to the input audio signal at a start of the first path and a second
25 gain to be applied to the input audio signal at a start of the second path.

3. A method as in any one of claims 1 and 2, further comprising selecting respective different gains to be applied by the amplifiers to the input audio signal before the filters.

4. A method as in claim 3, where the respective different gains are selected based, at least partially, upon the direction.

5. A method as in any one of claims 1 to 4, where the filters are static decorrelators and where the head-related transfer functions (HRTF) pairs of the second path are static HRTF pairs.

6. A method as in any one of claims 1 to 5, further comprising setting the adjustable amplifiers in the second path at different settings relative to one another based upon the direction.

7. A method as in any one of claims 1 to 6, where applying the interpolated head-related transfer function (HRTF) pair to the input audio signal in the first path comprising convolving the interpolated head-related transfer function (HRTF) pair to the input audio signal in the first path based upon the direction.

8. A method as in any one of claims 1 to 7, where the method is applied to a plurality of respective audio signals as the input audio signal at a same time, and where a plurality of left signals and right signals from the respective audio signals are combined for the sound reproduction.

9. An apparatus comprising:

20 a first audio signal path comprising an interpolated head-related transfer function (HRTF) pair applied to an input audio signal based upon a direction configured to generate direction dependent first left and right signals in the first path;

25 a second audio signal path comprising a plurality of:

an adjustable amplifier configured to be adjusted based upon the direction;

a filter for each adjustable amplifier, and

a respective head-related transfer function (HRTF) pair applied to an output from the filter,

5 where the second path is configured to generate direction dependent second left and right signals for each filter in the second path, and

where the apparatus is configured to combine the generated left signals from the first and second paths to form a left output signal for a sound reproduction, and to combine the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

10

10. An apparatus as in claim 9, where the first audio signal path comprises a first variable amplifier before the first head-related transfer function (HRTF) pair, where the second audio signal path comprises a second variable amplifier before the filters, and the apparatus comprises an adjuster to adjust a desired externalization based upon adjusting the first and second variable amplifiers.

15

11. An apparatus as in any one of claims 9 and 10, further comprising a selector connected to the adjustable amplifiers, where the adjuster is configured to adjust the adjustable amplifiers to different respective settings based, at least partially, upon the direction.

20

12. An apparatus as in any one of claims 9 to 11, where the filters are static decorrelators and where the head-related transfer function (HRTF) pairs of the second audio signal path are static .

25

13. An apparatus as in any one of claims 9 to 12, where the first audio signal path is configured to convolve the interpolated head-related transfer function (HRTF) pair to the input audio signal based upon the direction.

30

14. An apparatus as in any one of claims 9 to 13, where the apparatus comprises a plurality of pairs of the first and second paths, and where the apparatus is configured to apply a respective multi-channel audio signal to a respective one of the
5 pairs of the first and second paths as the input audio signal at a same time, and where a plurality of left signals and right signals from the respective multi-channel signals are combined for the sound reproduction.

15. A non-transitory program storage device readable by a
10 machine, tangibly embodying a program of instructions executable by the machine for performing operations, the operations comprising:

controlling, at least partially, a first audio signal path for an input audio signal comprising applying an
15 interpolated head-related transfer function (HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path;

controlling, at least partially, a second audio signal path for the same input audio signal, where the second audio
20 signal path comprises adjustable amplifiers configured to be set based upon the direction, applying outputs from the amplifiers to respective filters for each of the amplifiers and applying to an output from each of the filters a
25 respective head-related transfer function (HRTF) pair to generate direction dependent second left and right signals for each filter in the second path; and

combining the generated left signals from the first and second paths to form a left output signal for a sound
30 reproduction, and combining the generated right signals from the first and second paths to form a right output signal for the sound reproduction.

16. A non-transitory program storage device as in claim 15, where the operations further comprise, based upon a desired externalization, selecting a first gain to be applied to the input audio signal at a start of the first path and a second gain to be applied to the input audio signal at a start of the second path.

17. A non-transitory program storage device as in any one of claims 15 and 16, where the operations further comprise selecting respective different gains to be applied to the input audio signal by the amplifiers before the decorrelators.

18. A non-transitory program storage device as in any one of claims 15 to 17, where the respective second head-related transfer function (HRTF) filtering comprises use of static head-related transfer function (HRTF) filters.

19. A non-transitory program storage device as in claim 18, where the operations further comprise outputs from the first path comprising a left first path output signal and a right first path output signal from the first head-related transfer function (HRTF), and where the outputs from the second path comprise a left second path output signal and a right second path output signal from each of the second head-related transfer function (HRTF) filtering.

20. A non-transitory program storage device as in any one of claims 15 to 19, where the operations further comprises the input audio signal comprising a plurality of respective multi-channel signals being controlled at a same time, and where a plurality of left signals and right signals from the respective multi-channel signals are combined for the sound reproduction.

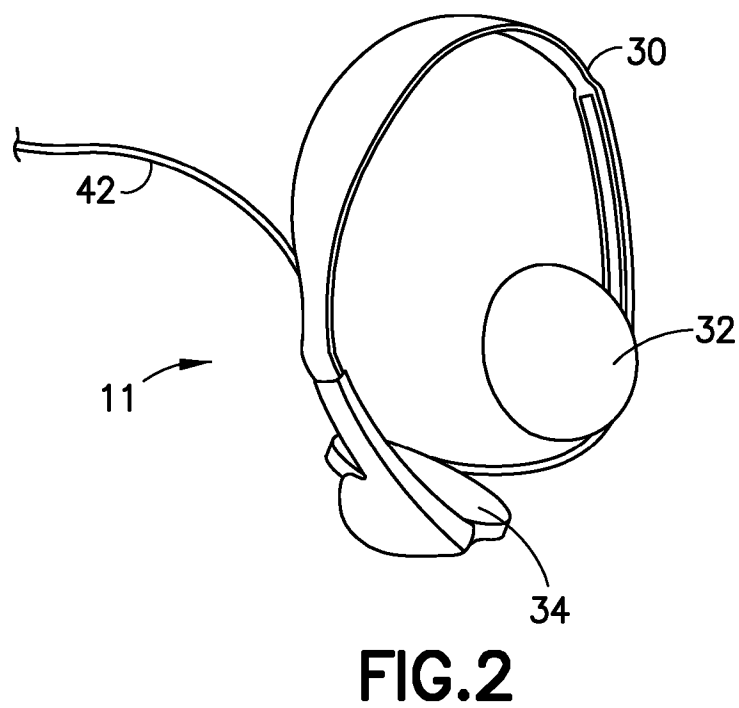
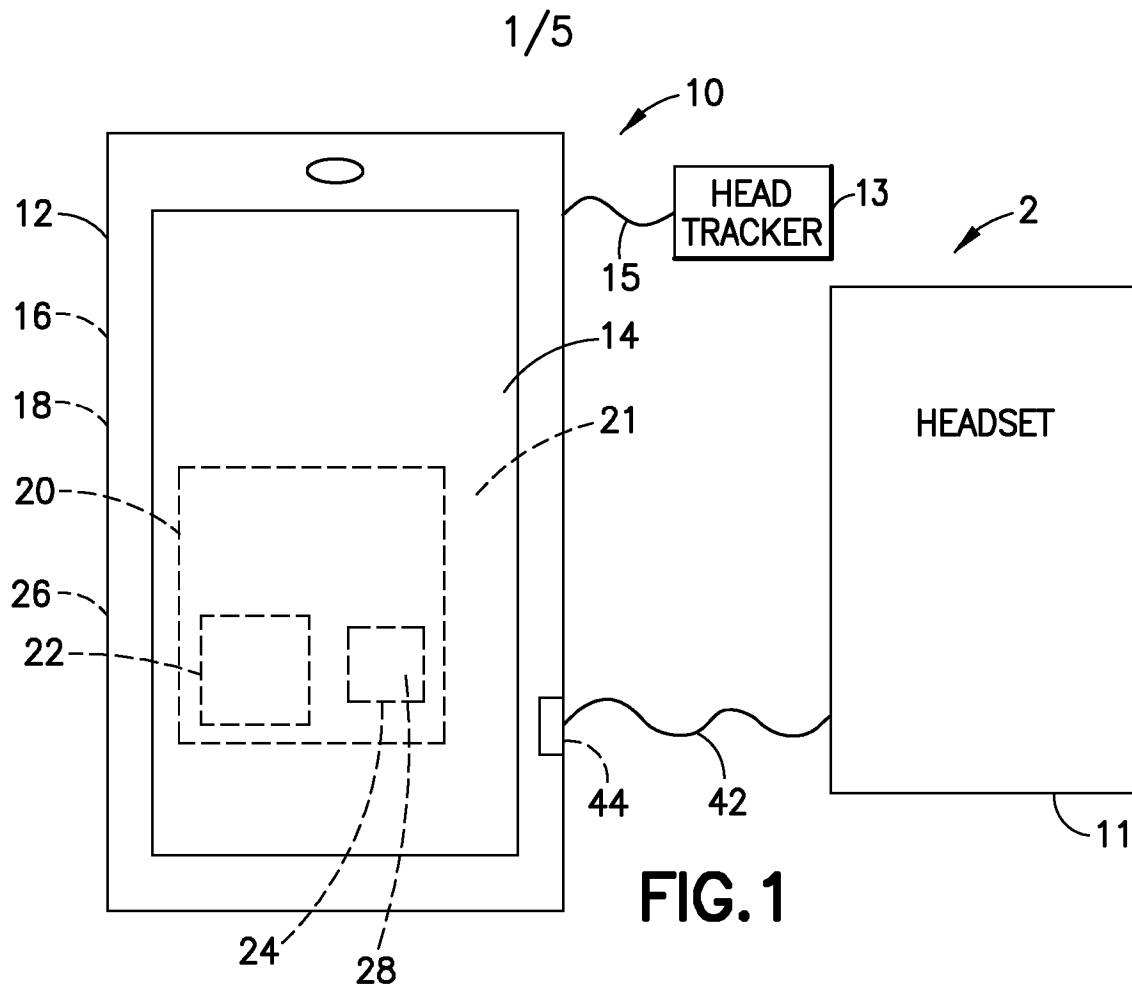
21. An apparatus comprising:

means for providing an input audio signal in a first path and applying an interpolated head-related transfer function

(HRTF) pair based upon a direction to generate direction dependent first left and right signals in the first path;

means for providing the input audio signal in a second path, where the second path comprises a plurality of filters and a respective adjustable amplifier for each filter, where
5 the amplifiers are configured to be adjusted based upon the direction; and

means for applying to an output from each of the filters a respective head-related transfer function (HRTF) pair to
10 generate direction dependent second left and right signals for each filter in the second path, and combining the generated left signals from the first and second paths to form a left output signal for a sound reproduction, and
combining the generated right signals from the first and
15 second paths to form a right output signal for the sound reproduction.



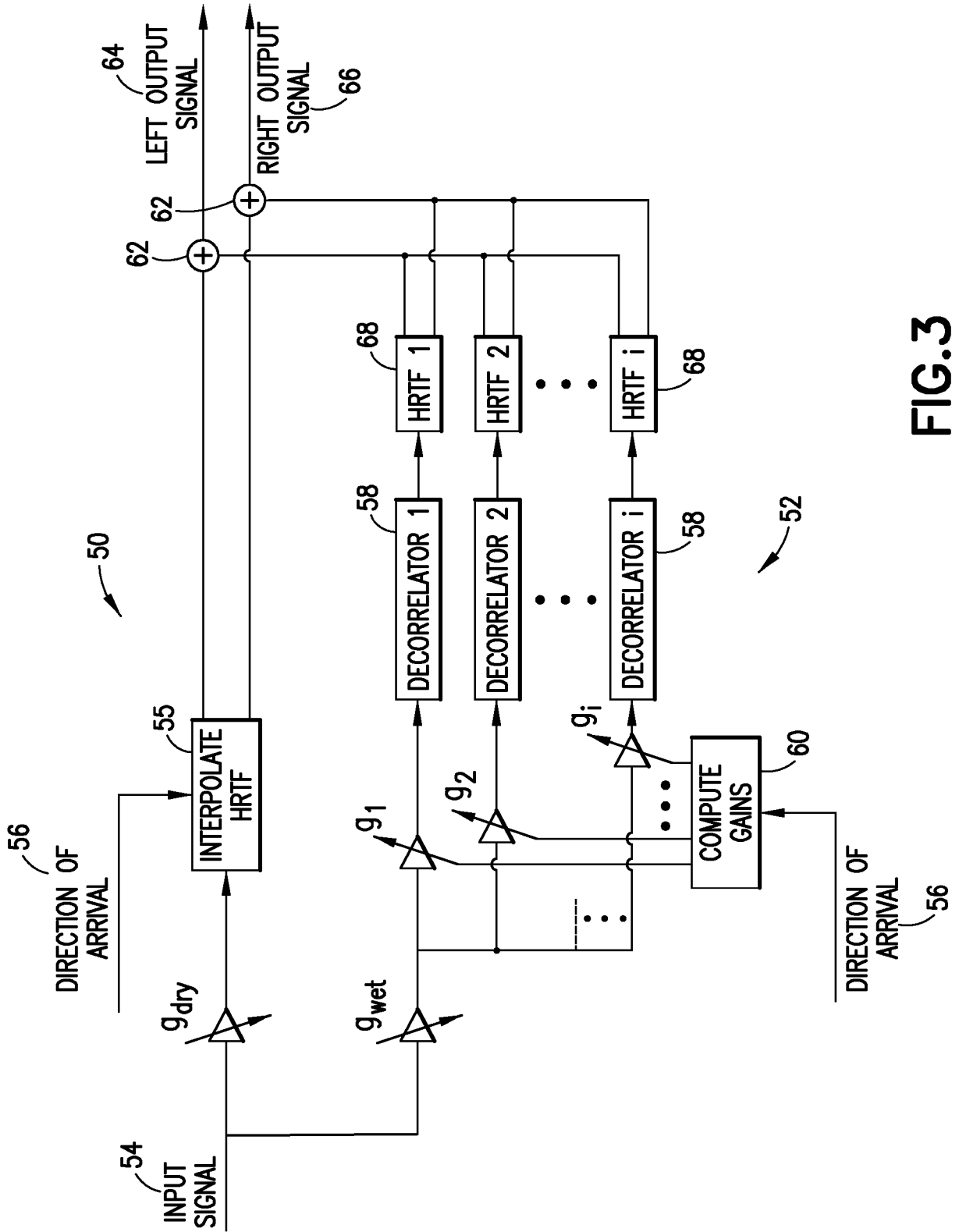


FIG.3

3/5

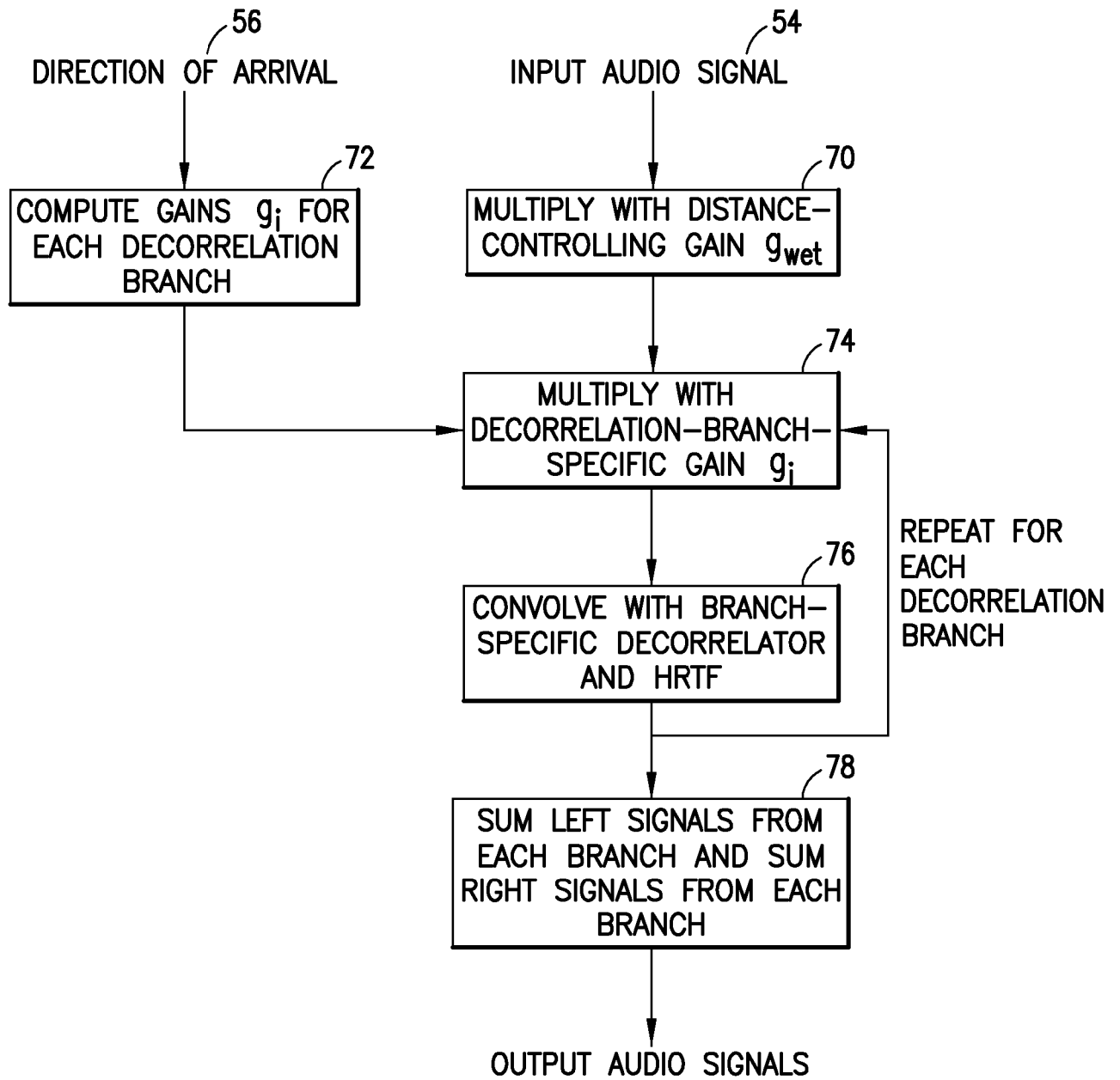
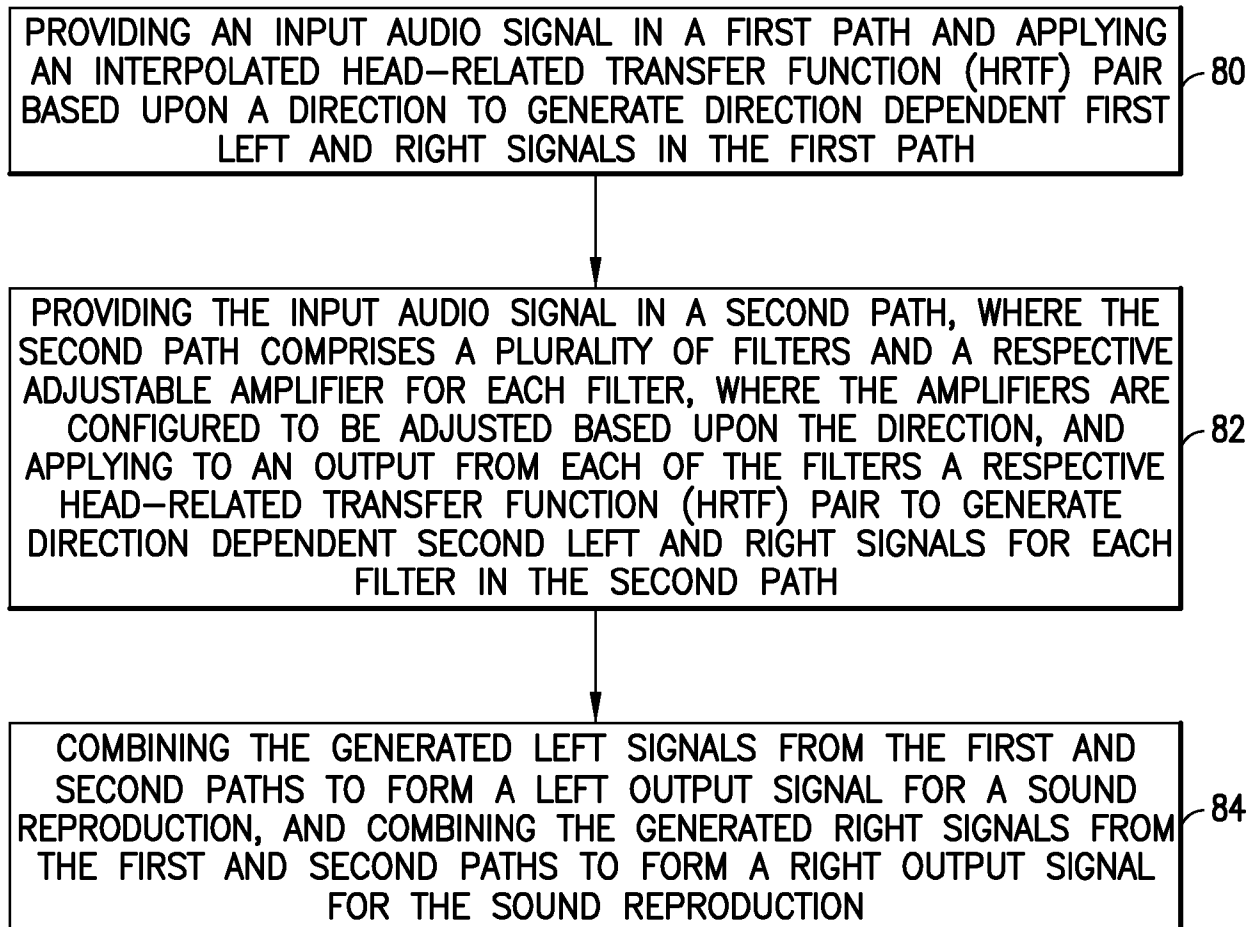


FIG.4

4/5

**FIG.5**

5/5

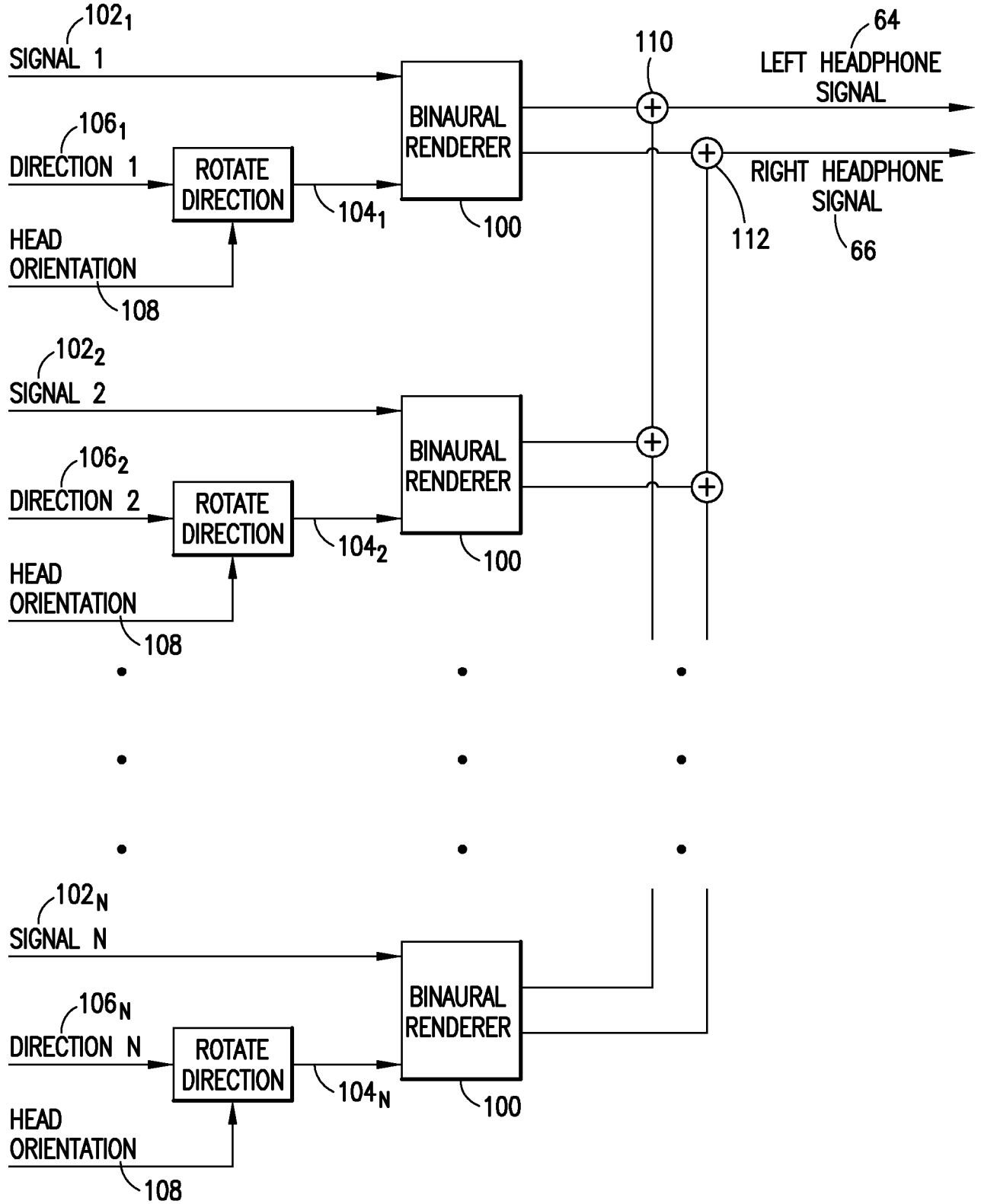


FIG.6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2016/050432

A. CLASSIFICATION OF SUBJECT MATTER

See extra sheet

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)

IPC: H04R, H04S

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

FI, SE, NO, DK

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

EPO-Internal, WPIAP, Internet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011299707 A1 (MEYER LAURENS [AU]) 08 December 2011 (08.12.2011) abstract; paragraphsn [0039], [0045]; figs 2-3	1-4, 6-11 and 13-21
A	WO 2014036121 A1 (DOLBY LAB LICENSING CORP [US]) 06 March 2014 (06.03.2014) abstract; page 33, rows 6-10; page 41, rows 9-29	1-21
A	US 2009252356 A1 (GOODWIN MICHAEL M [US] et al.) 08 October 2009 (08.10.2009) abstract	1-21
A	WO 2007112756 A2 (UNIV AALBORG [DK]) 11 October 2007 (11.10.2007) abstract	1-21

 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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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

24 August 2016 (24.08.2016)

Date of mailing of the international search report

01 September 2016 (01.09.2016)

Name and mailing address of the ISA/FI
Finnish Patent and Registration Office
P.O. Box 1160, FI-00101 HELSINKI, Finland
Facsimile No. +358 9 6939 5328

Authorized officer
Janne Nummela
Telephone No. +358 9 6939 500

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FI2016/050432

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009046864 A1 (MAHABUB JERRY [US] et al.) 19 February 2009 (19.02.2009) abstract	1-21

INTERNATIONAL SEARCH REPORT
Information on Patent Family Members

International application No.
PCT/FI2016/050432

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
US 2011299707 A1	08/12/2011	US 9332372 B2	03/05/2016
		TW 201215179 A	01/04/2012
		WO 2011154270 A1	15/12/2011
.....			
WO 2014036121 A1	06/03/2014	CN 104604257 A	06/05/2015
		CN 104604257 B	25/05/2016
		EP 2891338 A1	08/07/2015
		HK 1205845 A1	24/12/2015
		JP 2015530825 A	15/10/2015
		US 2015223002 A1	06/08/2015
.....			
US 2009252356 A1	08/10/2009	US 8374365 B2	12/02/2013
		CN 101828407 A	08/09/2010
		CN 101828407 B	16/12/2015
		CN 101884065 A	10/11/2010
		CN 101884065 B	10/07/2013
		CN 101889307 A	17/11/2010
		CN 101889307 B	23/01/2013
		CN 101981811 A	23/02/2011
		CN 101981811 B	23/10/2013
		CN 102272840 A	07/12/2011
		CN 102783187 A	14/11/2012
		CN 102783187 B	03/08/2016
		CN 105376673 A	02/03/2016
		EP 2272169 A2	12/01/2011
		EP 2382631 A2	02/11/2011
		GB 201006663 D0	09/06/2010
		GB 2466172 A	16/06/2010
		GB 2466172 B	06/03/2013
		GB 201006666 D0	09/06/2010
		GB 2467247 A	28/07/2010
		GB 2467247 B	29/02/2012
		GB 201006665 D0	09/06/2010
		GB 2467668 A	11/08/2010
		GB 2467668 B	07/12/2011
		SG 172862 A1	29/08/2011
		SG 182561 A1	30/08/2012
SG 187503 A1	28/02/2013		
SG 10201500753Q A	29/04/2015		
TW 201143483 A	01/12/2011		
TW I528841 B	01/04/2016		
US 2009252341 A1	08/10/2009		
US 8204237 B2	19/06/2012		
US 2008205676 A1	28/08/2008		

INTERNATIONAL SEARCH REPORT
Information on Patent Family Members

International application No.
PCT/FI2016/050432

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
		US 8345899 B2	01/01/2013
		US 2007269063 A1	22/11/2007
		US 8379868 B2	19/02/2013
		US 2008031462 A1	07/02/2008
		US 8619998 B2	31/12/2013
		US 2009092259 A1	09/04/2009
		US 8712061 B2	29/04/2014
		US 2009103749 A1	23/04/2009
		US 8934640 B2	13/01/2015
		US 2008232617 A1	25/09/2008
		US 9014377 B2	21/04/2015
		US 2008175394 A1	24/07/2008
		US 9088855 B2	21/07/2015
		US 2011188660 A1	04/08/2011
		US 9247369 B2	26/01/2016
		US 2009110204 A1	30/04/2009
		US 2014270281 A1	18/09/2014
		WO 2009046223 A2	09/04/2009
		WO 2009046460 A2	09/04/2009
		WO 2009052444 A2	23/04/2009
		WO 2009146047 A2	03/12/2009
		WO 2010080854 A2	15/07/2010
		WO 2011093793 A1	04/08/2011
.....			
WO 2007112756 A2	11/10/2007	EP 2005793 A2	24/12/2008
		US 2009052703 A1	26/02/2009
.....			
US 2009046864 A1	19/02/2009	US 9197977 B2	24/11/2015
		CN 101960866 A	26/01/2011
		CN 101960866 B	25/09/2013
		CN 102440003 A	02/05/2012
		CN 102440003 B	27/01/2016
		CN 103716748 A	09/04/2014
		EP 2119306 A2	18/11/2009
		EP 2356825 A1	17/08/2011
		JP 2010520671 A	10/06/2010
		JP 5285626 B2	11/09/2013
		JP 2012506673 A	15/03/2012
		JP 5694174 B2	01/04/2015
		JP 2013211906 A	10/10/2013
		US 2010246831 A1	30/09/2010
		US 8520873 B2	27/08/2013
		US 2014064494 A1	06/03/2014
		US 9271080 B2	23/02/2016
		WO 2008106680 A2	04/09/2008

INTERNATIONAL SEARCH REPORT
Information on Patent Family Members

International application No.
PCT/FI2016/050432

Patent document cited in search report	Publication date	Patent family members(s)	Publication date
		WO 2010048157 A1	29/04/2010
.....			

CLASSIFICATION OF SUBJECT MATTER

IPC
H04S 7/00 (2006.01)
H04R 5/02 (2006.01)