



(12) 发明专利

(10) 授权公告号 CN 102804814 B

(45) 授权公告日 2015. 09. 23

(21) 申请号 201080065614. 8

简·阿比尔高·佩德森

(22) 申请日 2010. 09. 28

(74) 专利代理机构 北京安信方达知识产权代理

(30) 优先权数据

有限公司 11262

PA201000251 2010. 03. 26 DK

代理人 周靖 郑霞

(85) PCT国际申请进入国家阶段日

(51) Int. Cl.

2012. 09. 20

H04S 5/00(2006. 01)

(86) PCT国际申请的申请数据

(56) 对比文件

PCT/EP2010/064369 2010. 09. 28

WO 0162045 A1, 2001. 08. 23,

(87) PCT国际申请的公布数据

WO2011/116839 EN 2011. 09. 29

WO 0162045 A1, 2001. 08. 23,

CN 101401456 A, 2009. 04. 01,

US 2008298597 A1, 2008. 12. 04, 全文.

(73) 专利权人 邦及欧路夫森有限公司

审查员 郭倩茜

地址 丹麦斯楚厄

(72) 发明人 帕特里克·詹姆斯·赫加蒂

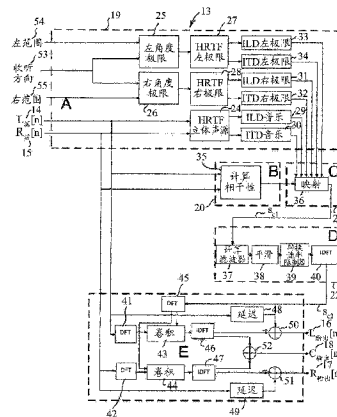
权利要求书3页 说明书16页 附图24页

(54) 发明名称

多通道声音重放方法和设备

(57) 摘要

本发明涉及用于选择通过一个或多个补充声音重放换能器进行重放的听觉信号分量的方法, 所述一个或多个补充声音重放换能器, 例如扬声器, 被放置在一对主要声音重放换能器之间, 例如在立体声扬声器装备中的左扬声器和右扬声器之间或在环绕声扬声器装备中的毗邻的扬声器之间, 所述方法包括以下步骤: (i) 指定所述补充声音重放换能器中的一个被定位在其内或待被定位在其内的方位角范围, 以及收听方向; (ii) 基于所述方位角范围和所述收听方向, 分别确定左右两耳间声强差极限和左右两耳间时间差极限; (iii) 将一对输入信号提供至所述一对主要声音重放换能器; (iv) 预处理所述输入信号中的每个, 由此提供一对已预处理的输入信号; (v) 确定在所述已预处理的信号之间的随频率而变的两耳间声强差和两耳间时间差; 以及 (vi) 将所述输入信号的分别具有在所述左右两耳间声强差极限之间的区间内的两耳间声强差和在所述左右两耳间时间差极限之间的区间内的两耳间时间差的那些信号分量提供至相应的补充声音重放换能器。本发明还涉及用于实施以上方法的设备以及这些设备的系统。



1. 一种用于选择在具有一个或多个补充声音重放换能器的扬声器装备中进行重放的听觉信号分量的方法,所述一个或多个补充声音重放换能器被放置在一对主要声音重放换能器之间,所述方法包括以下步骤:

(i) 指定所述补充声音重放换能器中的一个被定位在其内或待被定位在其内的方位角范围;

(ii) 基于所述方位角范围,根据在所述方位角范围的每个极限角度的源的自由场两耳脉冲响应分别确定左右两耳间声强差极限和左右两耳间时间差极限;

(iii) 将一对输入信号提供至所述一对主要声音重放换能器;

(iv) 预处理相应于理想的立体声收听的两耳脉冲响应与用于所述一对主要声音重放换能器的所述输入信号中的每个,由此提供一对已预处理的信号;

(v) 确定所述已预处理的信号之间的、随频率而变的两耳间声强差和两耳间时间差;以及

(vi) 将所述输入信号的分别具有在所述左右两耳间声强差极限之间的区间内的两耳间声强差和在所述左右两耳间时间差极限之间的区间内的两耳间时间差的那些信号分量提供至相应的补充声音重放换能器。

2. 根据权利要求 1 所述的方法,其中收听方向被指定为所述扬声器装备的听觉旋转。

3. 根据权利要求 1 所述的方法,其中具有在所述极限外部的两耳间声强差和两耳间时间差的那些信号分量被提供至所述主要声音重放换能器。

4. 根据权利要求 1 所述的方法,其中具有在所述极限外部的两耳间差的那些信号分量被作为输入信号提供至用于实施根据权利要求 1 所述的方法的装置。

5. 根据权利要求 1 所述的方法,其中所述两耳脉冲响应包括头部相关的传递函数。

6. 根据权利要求 1 所述的方法,还包括确定所述一对输入信号之间的相干性,并且其中所述信号分量在被提供至所述一个或多个补充声音重放换能器之前按所述相干性加权。

7. 根据权利要求 4 所述的方法,其中相对于收听者的正面方向以及因此由预处理装置进行的分别的处理由所述收听者选择。

8. 根据权利要求 4 所述的方法,其中相对于收听者的正面方向以及因此由预处理装置进行的分别的处理借助于被附接于收听者的头部跟踪装置进行控制。

9. 一种用于选择在具有一个或多个补充声音重放换能器 (56) 的扬声器装备中进行重放的听觉信号分量的设备,所述一个或多个补充声音重放换能器 (56) 被放置在一对主要声音重放换能器 (2、3) 之间,所述设备包括:

(i) 指定装置 (53、54、55),其用于指定所述补充声音重放换能器 (56) 中的一个被定位在其内或待被定位在其内的方位角范围;

(ii) 确定装置 (25、26、27、28、31、32、33、34),其基于所述方位角范围根据在所述方位角范围的每个极限角度的源的自由场两耳脉冲响应分别确定左右两耳间声强差极限和左右两耳间时间差极限;

(iii) 左和右输入终端 (14、15),其将一对输入信号提供至所述一对主要声音重放换能器 (2、3);

(iv) 预处理装置 (24),其用于预处理相应于理想的立体声收听的两耳脉冲响应与被提供在所述左和右输入终端 (14、15) 上用于所述一对主要声音重放换能器的所述输入信

号中的每个,由此提供一对已预处理的信号;

(v) 确定装置 (24),其用于确定所述已预处理的信号之间的随频率而变的两耳间声强差和两耳间时间差;以及

(vi) 信号处理装置 (22、23),其用于将所述输入信号的分别具有在所述左右两耳间声强差极限之间的区间内的两耳间声强差和在所述左右两耳间时间差极限之间的区间内的两耳间时间差的那些信号分量提供至补充输出终端 (18) 以提供至相应的补充声音重放换能器 (56)。

10. 根据权利要求 9 所述的设备,其中具有在所述极限外部的两耳间声强差和两耳间时间差的那些信号分量被提供至所述主要声音重放换能器 (2、3)。

11. 根据权利要求 9 所述的设备,其中具有在所述极限外部的两耳间差的那些信号分量被作为输入信号提供至另一根据权利要求 9 或 10 所述的设备。

12. 根据权利要求 9 所述的设备,其中所述预处理装置 (24) 是头部相关的传递函数装置。

13. 根据权利要求 9 所述的设备,还包括确定所述一对输入信号 (14、15) 之间的相干性的相干性确定装置 (35),并且其中所述输入信号 (14、15) 的所述信号分量在通过所述补充输出终端 (18) 提供至所述一个或多个补充声音重放换能器 (56) 之前按所述输入信号 (14、15) 之间的通道间相干性加权。

14. 根据权利要求 9 所述的设备,其中相对于收听者的正面方向以及因此由所述预处理装置 (24) 进行的分别的处理由所述收听者选择。

15. 根据权利要求 9 所述的设备,其中相对于收听者的正面方向以及因此由所述预处理装置 (24) 进行的分别的处理借助于被附接于收听者的头部跟踪装置或其他的用于确定所述收听者相对于声音重放换能器的装备的取向的装置进行控制。

16. 根据权利要求 9-15 中任一项所述的设备,其中所述补充声音重放换能器是扬声器。

17. 根据权利要求 9-15 中任一项所述的设备,其中所述主要声音重放换能器是在立体声扬声器装备中的左扬声器和右扬声器。

18. 根据权利要求 16 所述的设备,其中所述主要声音重放换能器是在立体声扬声器装备中的左扬声器和右扬声器。

19. 根据权利要求 9-15 中任一项所述的设备,其中所述主要声音重放换能器是在环绕声扬声器装备中的毗邻的扬声器。

20. 根据权利要求 16 所述的设备,其中所述主要声音重放换能器是在环绕声扬声器装备中的毗邻的扬声器。

21. 根据权利要求 9-15 中任一项所述的设备,其中所述指定装置是键盘或触摸屏。

22. 根据权利要求 16 所述的设备,其中所述指定装置是键盘或触摸屏。

23. 一种用于选择在具有一个或多个补充声音重放换能器 (56) 的扬声器装备中进行重放的听觉信号分量的系统,所述一个或多个补充声音重放换能器 (56) 被放置在一对主要声音重放换能器 (2、3) 之间,所述系统包括根据前述权利要求 9 至 15 中任一项所述的设备中的至少两个,其中所述设备中的第一个被提供有第一左和右输入信号 (14、15),并且其中所述设备中的所述第一个在左输出终端 (16)、右输出终端 (17) 和补充输出终端 (18) 上

提供输出信号,在所述补充输出终端(18)上的输出信号被提供至补充声音重放换能器,并且在所述左输出终端和所述右输出终端上的输出信号被分别提供至随后的根据前述权利要求9至15中任一项所述的设备的各自的输入信号,由此输出信号被提供至多个补充声音重放换能器(56)中的相应的补充声音重放换能器。

多通道声音重放方法和设备

技术领域

[0001] 本发明大体上涉及通过扬声器装备进行声音重放的领域,并且更具体地涉及用于获得在宽的收听区域上被重放的声音的稳定的听觉空间知觉的方法和系统。又更具体地,本发明涉及在受限的环境(例如汽车座舱)中使用的这样的方法和系统。

[0002] 发明背景

[0003] 立体声是流行的空间音频重放形式。立体声信号可以通过原位的立体声话筒录音或通过混合多个单声道信号来产生,这在现代流行音乐中是典型的。这种类型的材料通常意图使用对称地布置的匹配的扬声器对进行重放,如在 ITU-R BS. 1116[1997] 和 ITU-R BS. 775-1[1994] 中提出的。

[0004] 如果满足上文的推荐,那么收听者将感知至少在扬声器之间扩展的听觉场景,如在 Bregman[1994] 中描述的,包括各种虚拟源、幻像。如果 ITU 推荐中的一个或多个不被满足,那么后果可能是听觉场景的劣化,见例如 Bech[1998]。

[0005] 非常典型的是,在汽车中收听立体声的材料。最现代化的汽车在交付时就配备有出厂时已安装好的由立体声声源例如 CD 播放器和 2 个或更多个扬声器组成的音响系统。

[0006] 然而,当将汽车收听环境与 ITU 推荐比较时,与理想条件之间通常将存在以下偏离:

[0007] (i) 收听位置是错误的;

[0008] (ii) 扬声器位置是错误的;

[0009] (iii) 具有接近扬声器的大的反射表面。

[0010] 至少由于这些原因,听觉场景的保真度在汽车中典型地被劣化。

[0011] 将理解,虽然在本说明书中反复地参考在汽车中的音频重放,但是在下文中描述的本发明的原理的使用和本发明的系统和方法的具体的实施环境不限于汽车音频重放,而是还可以在多种其他的收听情况中获得应用。

[0012] 将是有利的是,尽管具有上文提到的距理想的收听条件的偏离,但是该重放系统和方法将能够获得具有高保真度的音频重放。

[0013] 听觉重放基本包括两个知觉方面:(i) 声音环境中的声源的音色的重放,以及(ii) 声音环境的空间属性的重放,例如获得声音环境中的声源的稳定的定位的能力以及获得对环境中的单个声源的空间延伸或宽度的正确的感知的能力。这些方面以及表征这些的具体的感知属性二者可能通过在封闭空间(例如汽车的座舱)中的音频重放遭受劣化。

[0014] 发明概述

[0015] 本部分将初步比较和对比在汽车收听环境与在自由场中的轴上和离轴环境中的立体声重放。在该比较之后,根据人耳的耳间传递函数,对汽车收听环境中的听觉场景的劣化进行分析。在该介绍之后,将给出本发明的主要原理的概述,根据该原理提供方法和相应的立体向多-单通道转换器(stereo to multi-mono converter)设备,借助于该方法和设备,听觉场景的听觉部件的定位可以独立于收听位置进行。

[0016] 本发明的实施环境将在本发明的详细描述中描述,该部分将还包括通过在 Matlab

软件的辅助下对模拟的其输出的分析,对根据本发明的立体向多-单通道转换器的实施环境的性能进行评价。

[0017] 理想的立体声收听环境

[0018] 双通道立体声(将在下文被称为立体声(stereo))是一种通过两个声源重放空间听觉场景的方式。Blauert[1997]作出了在术语声音和听觉之间的以下的区分:

[0019] 声音是指作为事件的特征的物理现象(例如声波、声源或声信号)。

[0020] 听觉是指被收听者感知的(例如听觉图像或场景)。

[0021] 这种区分也将在本说明书中应用。

[0022] Blauert[1997]将空间听觉定义为听觉事件的地点和声音事件的物理特性之间的关系。

[0023] 对于立体声信号的扬声器重放来说,收听者和声源的在水平面中的理想的相对位置在ITU-R BS. 1116[1997]和ITU-R BS. 775-1[1994]中描述,并且在图1中图形地示出,图1图示了扬声器和收听者的对于立体声信号的重放的理想的布置。

[0024] 收听者应当位于具有最小值 $d_l = d_r = d_{lr} = 2$ 米的等边三角形的顶点处。扬声器应当被分别地放置在其他两个顶点处。这些扬声器应当在频率响应和功率响应的方面匹配。距墙壁的最小距离应当是1米。距天花板的最小距离应当是1.2米。

[0025] 在本说明书中,小写字母变量将被用于时间域信号,例如 $x[n]$,并且大写字母变量将被用于频率域表示,例如 $X[k]$ 。

[0026] 声音信号 $l_{\text{耳}}[n]$ ($l_{\text{ear}}[n]$)和 $r_{\text{耳}}[n]$ ($r_{\text{ear}}[n]$)涉及两耳并且将在本说明书全文中被采用以意指在收听者的耳道的入口处测量的那些信号。Hammershøj和Møller[1996]显示,所有的定位所需要的方向信息在这些信号中可获得。两耳的信号之间的差异的属性被称为两耳间。参照图1,考虑其中仅具有由信号 $l_{\text{源}}[n]$ ($l_{\text{source}}[n]$)馈送的一个声源的情况。在这种情况下,左耳被称为和声源在身体的同侧,因为其和声源在相对于 0° 方位或正中线的同一个半球中, $h_{\text{LL}}[n]$ 是 $l_{\text{source}}[n]$ 和 $l_{\text{ear}}[n]$ 之间的传输路径的脉冲响应。相似地,右耳被称为在身体的对侧的,并且 $h_{\text{RL}}[n]$ 是 $l_{\text{source}}[n]$ 和 $r_{\text{ear}}[n]$ 之间的传输路径的脉冲响应。在理想的情况下 $\Theta_L = \Theta_R = 30^\circ$ 。

[0027] 如果本环境是用于自由场中的点源,那么这些脉冲响应、或频率域中的头部相关的传递函数(HRTF),将含有关于由躯干、头部和耳廓(外耳)导致的衍射、散射、干涉和共振效应的信息并且在表征源和收听者的相对位置的方式上有差异。在本发明中使用的HRTF是来自CIPIC Interface Laboratory[2004]数据库,并且具体地用于具有小的耳廓的KEMAR®头部和躯干模拟器。然而,将理解,头部相关的传递函数的其他的实例也可以根据本发明被使用,例如来自真实的人耳的、来自人造的人耳(人造的头部)的,以及甚至模拟的HRTF的头部相关的传递函数。

[0028] 这些信号的频率域表示使用离散傅里叶变换DFT来计算,如在以下的六个等式中以公式表示的,这些等式在Oppenheim和Schäfer[1999,页561]中共同被称为傅里叶分析等式。

$$[0029] \quad L_{\text{ear}}[k] = \sum_{n=0}^{N-1} l_{\text{ear}}[n] e^{j(2\pi/N)kn}$$

$$[0030] \quad R_{ear}[k] = \sum_{n=0}^{N-1} r_{ear}[n] e^{j(2\pi/N)kn}$$

$$[0031] \quad L_{source}[k] = \sum_{n=0}^{N-1} l_{source}[n] e^{j(2\pi/N)kn}$$

$$[0032] \quad R_{source}[k] = \sum_{n=0}^{N-1} r_{source}[n] e^{j(2\pi/N)kn}$$

$$[0033] \quad H_{LL}[k] = \sum_{n=0}^{N-1} h_{LL}[n] e^{j(2\pi/N)kn}$$

$$[0034] \quad H_{LR}[k] = \sum_{n=0}^{N-1} h_{LR}[n] e^{j(2\pi/N)kn}$$

[0035] 左耳和右耳之间的差异通过以下的等式中定义的耳间传递函数 $H_{IA}[k]$ 描述：

$$[0036] \quad H_{IA}[k] = \frac{L_{source}[k] \cdot H_{LL}[k]}{L_{source}[k] \cdot H_{LR}[k]}$$

[0037] 两耳听觉系统是指在两耳的信号上操作以产生被感知的空间感的过程的集合。所评价的基本的指标是两耳间的声强差 ILD 和两耳间的时间差 ITD。这些量在下文定义。

[0038] ILD 是指与平均声压级相关的 $L_{耳}[k]$ ($L_{ear}[k]$) 和 $R_{耳}[k]$ ($R_{ear}[k]$) 之间的相异度。ILD 由 $H_{IA}[k]$ 的幅度定量地描述。

[0039] ITD 是指与它们的时间关系相关的 $L_{ear}[k]$ 和 $R_{ear}[k]$ 之间的相异度。ITD 由 $H_{IA}[k]$ 的相位延迟定量地描述。在具体的频率处的相位延迟是除以频率后的负的展开的相位。

[0040] 对于其中 $L_{source}[k]$ 和 $R_{source}[k]$ 二者都存在的情况，耳间传递函数由以下的等式给出：

$$[0041] \quad H_{IA}[k] = \frac{L_{source}[k] \cdot H_{LL}[k] + R_{source}[k] \cdot H_{RL}[k]}{L_{source}[k] \cdot H_{LR}[k] + R_{source}[k] \cdot H_{RR}[k]}$$

[0042] 如果传输路径是线性的和时不变的，即 LTI，那么它们的脉冲响应可以被独立地确定，并且 $H_{IA}[k]$ 可以通过上文的等式中的叠加确定。

[0043] 信号的功率谱密度是其自相关的傅里叶变换。 $l_{source}[n]$ 和 $r_{source}[n]$ 的功率谱密度在频率域内可以计算为谱与其复共轭的乘积，如在以下的等式中示出的：

$$[0044] \quad P_L[k] = L_{source}[k] \cdot L_{source}[k]^*$$

$$[0045] \quad P_R[k] = R_{source}[k] \cdot R_{source}[k]^*$$

[0046] 交功率谱密度是两个信号之间的互相关的傅里叶变换。 $l_{source}[n]$ 和 $r_{source}[n]$ 的交功率谱密度在频率域内可以计算为 $L_{源}[k]$ ($L_{source}[k]$) 与 $R_{源}[k]$ ($R_{source}[k]$) 的复共轭的乘积，如在以下的等式中示出的：

$$[0047] \quad P_{LR}[k] = L_{source}[k] \cdot R_{source}[k]^*$$

[0048] $l_{source}[n]$ 和 $r_{source}[n]$ 之间的相干性是两个信号之间的相似性的表示并且取在 0 至 1 之间的值。其通过两个信号的功率谱密度和它们的交功率谱密度计算。相干性可以在频率域内使用以下的等式 (6) 被计算。容易示出，如果单一的数据块被使用，那么 $C_{LR} = 1$ ，并且因此在多个正在被分析的信号块上计算 C_{LR} 。

$$[0049] \quad C_{LR}[k] = \frac{|P_{LR}^2[k]|}{P_L[k] \cdot P_R[k]}$$

[0050] 要求的是, $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 共同是平稳的随机过程。根据 Shanmugan 和 Breipohl[1988], 这意指, 自相关和联合分布应当是随时间迁移不变的。

[0051] 当 $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 是相干的并且不具有 ILD 或 ITD 时, 并且假设自由场条件以及头部和躯干对称, 那么 $H_{IA}[k]$ 的幅度和相位 = 0, 如图 2 中所示的。在某个频率的正的 ILD 将意指在该频率的 $l_{\text{source}}[n]$ 的较高的水平。相似地, 在某个频率的正的 ITD 将意指该频率在 $l_{\text{source}}[n]$ 中较早发生。

[0052] 正常并且健康的听觉系统在这样的条件下的输出是单一的听觉图像, 也被称为幻像, 其以两个源之间的弧形段上的 0 度方位线为中心。诸如到达每个耳的声音是相同的环境也被称为影响双耳的 (diotic)。相似地, 如果具有小的 ILD 和 / 或 ITD 差异, 那么单一的听觉图像将仍然被感知。该图像的在两个源之间的位置通过 ITD 和 ILD 被确定。这种现象被称为相加定位 (Blauert[1997, 页 209]), ILD 和 ITD 指标被“相加”, 导致单一的知觉事件。这形成作为产生空间听觉场景的方式的立体声的基础。

[0053] 如果 ITD 超出约 1ms, 相应于约 0.34m 的距离, 那么听觉事件将被定位在最早的源处。这被称为第一声波靠前定律。因此, 在初始的声音中, 仅是在 1ms 内到达人耳的声音对于在立体声中的定位是关键的。这是关于源和室边界之间的距离的 ITU 推荐的一个原因。如果延迟被进一步增加, 那么第二听觉事件将作为第一听觉事件的回声被感知。

[0054] 真实的立体声音乐信号可以具有任何数量的分量, 其 $C_{LR}[k]$ 范围作为时间的函数在 0 至 1 之间。当 $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 被立体声音乐信号驱动时, 两耳听觉系统的输出是发生在两个源之间的听觉场景, 其范围和本质取决于立体声音乐信号之间的关系。

[0055] 离轴收听环境

[0056] 在前述的关于理想的立体声收听环境的段落中, 已经考虑了相对于立体声声源对称地定位的收听位置。即, 收听者位于所谓的“最有效点”的中心处, 其是收听室中的将发生最优的空间声音重放的区域。取决于源、收听位置和室边界之间的距离, “最有效点”的有效区域将变化, 但是其将是有限的。由于该原因, 典型的是, 某些收听者在离轴位置。离轴收听位置的一个实例在图 3 中示出。

[0057] 在以下的分析中, 再次地假设在自由场中的点源和对称的 HRTF。

[0058] 参照图 3, 明显的是, 从两个声源到每个相应的耳的传播路径具有不同的长度, $d_1 < d_r$ 。汽车收听环境中的典型的距离约是 $d_1 = 1\text{m}$, $d_r = 1.45\text{m}$ 并且 $d_{lr} = 1.2\text{m}$ 。因为 $d_r - d_1 = 0.45\text{m}$, 所以使用第一声波靠前定律具有直接的问题, 后果是听觉场景的大部分向左声源坍塌 (collapse)。除这之外, 角度 Θ_L 和 Θ_R 不再是相等的并且因此两耳的脉冲响应将不再是相等的, 即 $h_{LL}[n] \neq h_{RR}[n]$ 并且 $h_{LR}[n] \neq h_{RL}[n]$ 。如果角度被估计为 $\Theta_L = 25^\circ$ 并且 $\Theta_R = 35^\circ$ 并且两耳的脉冲响应被修改以模拟近似的路径长度差的延迟和衰减, 那么 $H_{IA}[k]$ 的幅度和相位是如图 4 中所示的。

[0059] 与在轴上收听位置不同, 当 $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 被相同的信号驱动时, 在这种情况下, 听觉图像不可能位于收听者的正前方, 而是将很有可能向左“倾斜”, 或甚至完全地坍塌至左源的位置。音色将也被影响, 因为 ITD 偏移将创建梳状滤波器, 如可以在图 4 中示出的 ILD 图的较大的峰值中看到的。对于真实的立体声音乐信号, 听觉场景将很有可能不被

精确地重放,因为相加定位不再基于意图的两耳间指标。如果仅具有一个收听者,那么这些效应可以通过使用例如由 Tokuno, Kirkeby, Nelson 和 Hamada[1997] 描述的方法去卷积进行校正。

[0060] 大多数真实的立体声收听环境不同于上文描述的理想的情况。真实的扬声器由于制造公差不可能具有完全匹配的频率响应和功率响应。此外,扬声器在真实的收听室中的位置可能接近障碍物和反射表面,其可能引入影响 H_{TA} 的幅度和相位的依赖于频率的传播路径。如提到的,ITU 推荐意图减少这样的效应。

[0061] 虽然本发明可以在许多不同的环境中应用,但是特别地,在汽车座舱中的立体声重放将在以下的段落中被详细地描述。

[0062] 汽车中的收听环境

[0063] 汽车的立体声环境和“理想的”立体声环境之间的差异中的某些将在下文简要地描述。

[0064] 当电动的活塞扬声器被使用时,也典型的是,多个换能器被用于重放音频频谱(20Hz 至 20kHz)。这样做的一个原因是作为频率的函数的由活塞辐射的声压的增加的方向性。其对于离轴收听有重大意义,如上文提到的。当波的传播在活塞(扬声器膜)上发生时,在高频率处,这种类型的扬声器的锥体也停止像活塞一样运动,从而产生失真。这种现象被称为圆锥体分裂。

[0065] 扬声器典型地被安装在铁格(grill)后方,在车身中的各种空腔内部。据此,声音可以运动经过多个谐振系统。扬声器也将可能激励辐射另外的声音的其他的振动系统,例如门饰板。源可以接近座舱的边界并且其他的较大的反射表面可以在距源 0.34m 内。这将导致反射在直达声的 1ms 内到达,从而影响定位。与右信号相比,在左信号的源的路径中可以具有不同的障碍物(例如由于仪表组和方向盘,仪表板不是对称的)。座椅中的声音吸收材料例如罩和泡棉在空间上不均匀地分布。在近似 65 至 400Hz 之间的低频率,交通装置座舱中的声场包括或多或少将被阻尼的各种模式。

[0066] 结果是, $l_{ear}[n]$ 和 $r_{ear}[n]$ 分别将是 从换能器经过座舱至相应的耳的多个传输路径的叠加。

[0067] 这种情况由于以下事实进一步复杂化,即对于所有的驾驶者和乘客而言不具有固定的收听位置,并且代替地,收听区域的概念被使用。收听区域坐标系在图 5 中示出。

[0068] “收听区域”是收听者的耳朵最可能在其中并且因此回放系统的行为在其中是最关键的空间区域。在汽车中就座的驾驶者的位置被很好地在文献中记载,见例如 Parkin, Mackay 和 Cooper[1995]。通过将由 Parkin 等人提出的对于 95% 的观测数据与在 ITU-T P. 58[1996] 中推荐的头部几何构型组合,以下的收听窗应当包括大多数驾驶者的耳。参照图 6 中示出的汽车收听的实施例。

[0069] 驾驶者的收听区域(在图 6 中被表示为围绕收听者的头部的矩形)距原点的近似的距离是 $d_l = 1m$, $d_r = 1.45m$ 并且 $d_{lr} = 1.2m$ 。在驾驶者的收听区域和乘客的收听区域的中心之间的近似的距离是 $d_{listeners} = 0.8m$ 。

[0070] 在汽车的“收听区域”中的四个位置的耳间传递函数已经根据使用人造头部进行的测量被计算。图 7 示出了在位置 1(在驾驶者的收听窗口的后方)和在位置 2(在驾驶者的收听窗口的前方)的 H_{TA} 。图 8 示出了在位置 3(在乘客的收听窗口的后方)和在位置

4(在乘客的收听窗口的前方)的 H_{IA} 。

[0071] 这些图揭示了在四个不同的收听位置之间的大的幅度差和相位差。在多于一个位置处修正这些差异是不可能的,并且在其他的位置处,去卷积可能甚至增加差异并且引入其他的可听到的人为因素,例如预振铃。要点是,在本环境中,去卷积不是用于定位的劣化的符合实际的解决方案。

[0072] 立体向多-单通道转换

[0073] 前述的分析表明了离轴收听位置在立体声重放下如何改变耳间传递函数。听觉场景将在其上如意图地被感知的较小的收听区域是作为空间声音重放的方式的立体声的限制。对于该问题的解决方案由 Pedersen 在 EP1260119B1 中提出。

[0074] 在以上的文献中提出的解决方案由以下组成:多个声音信号从立体声信号中获取,使得这些信号中的每个可以通过放置在立体声信号由上文描述的理想立体声装备重放时所产生的那些幻像源位置处的一个或多个扬声器进行重放。这种立体向多-单通道转换意图将幻像源转变为实声源,由此使它们的地点独立于收听位置。分析立体声信号,并且它们的各个频率分量的方位地点根据通道间幅度和相位差以及通道间相干性估计。

[0075] 在以上的背景下,本发明的一个目的是提供一种不仅在所选择的优选的收听位置处而且更一般地在收听室的较大部分,特别地但是不排他地,在汽车的座舱内创建所给出的听觉场景的令人满意的重放的方法和相应的系统或设备。

[0076] 以上的和其他的目的和优点根据本发明通过提供立体向多-单通道转换方法和相应的设备或系统来获得,根据这些,分布在听觉场景上并且构成听觉场景的幻像源的地点从两耳信号 $l_{ear}[n]$ 和 $r_{ear}[n]$ 被估计。为了确定哪个扬声器应当重放立体声信号的每个单独的分量,每个扬声器被分配待覆盖的方位角的一个范围,该范围可以与重放系统中的扬声器的数量成反比。根据在方位角的同一个范围上的头部相关的传输函数计算的 ILD 和 ITD 极限被分配给每个扬声器。立体声信号的每个分量被扬声器重放,其 ILD 和 ITD 极限与特定的信号分量的 ILD 和 ITD 一致。如上文提到的,立体声信号之间的较高的通道间相干性对于幻像源的发生是必需的,并且因此整个过程仍然根据这种相干性进行调整。

[0077] 与在上文提到的 EP1260119B1 中描述的最初的立体声向多-单通道系统和方法相比,本发明获得对幻像源的位置的更好的预测,普通的收听者将通过从经过 HRTF 处理之后的左信号和右信号而不是从用于在一般的立体声装备中的扬声器重放的左信号和右信号获取 ITD、ILD 和相干性来感知该位置,即幻像源的预测是基于两耳信号。在本发明中使用的基于两耳信号对幻像源的最可能的位置的预测具有非常重要的结果,即幻像源的在空间中任何地点的定位(即不仅被约束于在一般的立体声的装备中的在收听者的前方且在左扬声器和右扬声器之间的部分)可以发生,在该预测之后具体的信号分量可以被引导至放置在收听区域周围的任何地点处的扬声器。

[0078] 在根据本发明的系统和方法的具体的实施环境中,结合有头部跟踪设备,使得头部跟踪设备可以传感收听者的头部的取向并且以使得收听者的头部的正面方向相应于由多个扬声器重放的听觉场景的正面方向的方式改变用于每个单独的扬声器的各自的信号的处理。这种效果根据本发明由头部跟踪装置提供,所述头部跟踪装置与提供用于设置左角度限制装置和右角度限制装置的控制信号的收听者相关联,例如在本发明的详细描述中示出的。

[0079] 虽然本发明的说明书将着重于应用三个扬声器（左、中央和右扬声器）的立体声向多-单通道系统和方法的实施环境，但是根据本发明的原理可以将系统和方法调整为其他数量的扬声器，例如调整为放置为围绕收听者的在经过他的耳朵的水平面中的五个扬声器，如从家庭使用的环绕声系统或汽车中的扬声器装备已知的。一个这种类型的实施环境将在本发明的详细描述中描述。

[0080] 根据本发明的第一方面，因此提供了一种用于选择用于重放的听觉信号分量的方法，所述方法借助于放置在一对主要声音重放换能器（例如在立体声扬声器装备中的左扬声器和右扬声器或在环绕声扬声器装备中的毗邻的扬声器）之间的一个或多个补充的声音重放换能器，例如扬声器，所述方法包括以下步骤：

[0081] (i) 指定方位角范围和收听方向，在所述方位角范围内，所述补充声音重放换能器中的一个被定位或待被定位在其中；

[0082] (ii) 基于所述方位角范围和所述收听方向，分别确定左右两耳间声强差极限和左右两耳间时间差极限；

[0083] (iii) 将一对输入信号提供至所述一对主要声音重放换能器；

[0084] (iv) 预处理所述输入信号中的每个，由此提供一对已预处理的输入信号；

[0085] (v) 确定在所述已预处理的信号之间的随频率而变的两耳间声强差和两耳间时间差；以及

[0086] (vi) 将所述输入信号的具有在所述左右两耳间声强差极限之间的区间内的两耳间声强差和在所述左右两耳间时间差极限之间的区间内的两耳间时间差的那些信号分量提供至相应的补充声音重放换能器。

[0087] 根据本发明的方法的具体的实施环境，具有在所述极限外部的两耳间声强差和两耳间时间差的那些信号分量被分别提供至所述左主要声音重放换能器和所述右主要声音重放换能器。

[0088] 根据本发明的方法的另一个具体的实施环境，具有在所述极限外部的两耳间的差值的那些信号分量被作为输入信号提供至用于实施上述方法的装置。

[0089] 根据本发明的方法的一个具体的实施环境，所述预处理装置是头部相关的传递函数装置，即预处理装置的输入通过相应于真实的人的头部相关函数（HRTF）、人造头部的头部相关传递函数或模拟的头部相关函数的函数被处理。

[0090] 根据本发明的方法的目前优选的具体的实施环境，所述方法还包括确定所述一对输入信号之间的相干性，并且其中所述信号分量在提供至所述一个或多个补充声音重放换能器之前按所述相干性加权。

[0091] 根据本发明的方法的又一具体的实施环境，相对于收听者的正面方向以及因此通过所述预处理装置（例如头部相关传递函数）的分别的处理由所述收听者选择。

[0092] 根据本发明的方法的一个具体的实施环境，相对于收听者的正面方向以及因此通过所述预处理装置（例如头部相关传递函数）的分别的处理借助于附接于收听者的头部跟踪装置进行控制。

[0093] 根据本发明的第二方面，还提供用于选择用于重放的听觉信号分量的设备，所述设备借助于放置在一对主要声音重放换能器（例如在立体声扬声器装备中的左扬声器和右扬声器或在环绕声扬声器装备中的毗邻的扬声器）之间的一个或多个补充声音重放换

能器,例如扬声器,其中所述设备包括:

[0094] (i) 指定装置,例如键盘或触摸屏,其用于指定所述补充声音重放换能器中的一个被定位在其中或待被定位在其中的方位角范围,并且用于指定收听方向;

[0095] (ii) 确定装置,其基于所述方位角范围和所述收听方向分别确定左右两耳间声强差极限和左右两耳间时间差极限;

[0096] (iii) 左和右输入终端,其将一对输入信号提供至所述一对主要声音重放换能器;

[0097] (iv) 预处理装置,其用于预处理分别提供在所述左和右输入终端上的所述输入信号中的每个,由此提供一对已预处理的输入信号;

[0098] (v) 确定装置,其用于确定在所述已预处理的输入信号之间的随频率而变的两耳间声强差和两耳间时间差;以及

[0099] (vi) 信号处理装置,其用于将所述输入信号的具有在所述左右两耳间声强差极限之间的区间内的两耳间声强差和在所述左右两耳间时间差极限之间的区间内的两耳间时间差的那些信号分量提供至补充输出终端以提供至相应的补充声音重放换能器。

[0100] 根据本发明的设备的一个实施环境,具有在所述极限外部的两耳间声强差和两耳间时间差的那些信号分量被分别提供至所述左主要声音重放换能器和所述右主要声音重放换能器。

[0101] 根据本发明的另一个实施环境,具有在所述极限外部的两耳间的差值的那些信号分量被作为输入信号提供至上文指定的设备,由此将可能的是,装备包括放置在围绕收听者的地点处的多个补充换能器的较大的系统。例如,在包括前左、前中、前右、后左和后右的主要扬声器的环绕声扬声器装备中,根据本发明的系统可以例如分别向放置在前左和后左的主要扬声器之间和放置在前右和后右的主要扬声器之间的扬声器提供信号。多种其他的扬声器布置可以利用本发明的原理进行装备,并且这样的装备将全部落入本发明的范围内。

[0102] 根据本发明的优选的实施环境,所述预处理装置是头部相关传递函数装置。

[0103] 根据本发明的又一并且目前也优选的实施环境,所述设备包括确定所述一对输入信号之间的相干性的相干性确定装置,并且所述输入信号的信号分量在通过所述输出终端提供至所述一个或多个补充声音重放换能器之前按所述输入信号之间的通道间相干性加权。

[0104] 根据本发明的设备的又一实施环境,相对于收听者的正面方向以及因此通过所述预处理装置(例如头部相关传递函数)的分别的处理由所述收听者选择,例如使用合适的接口,例如键盘或触摸屏。

[0105] 根据本发明的设备的可选择实施环境,相对于收听者的正面方向以及因此通过所述预处理装置(例如头部相关传递函数)的分别的处理借助于附接于收听者的头部跟踪装置或其他用于确定所述收听者相对于声音重放换能器的装备的取向的装置进行控制。

[0106] 根据本发明的第三方面,提供用于选择用于重放的听觉信号分量的系统,所述系统借助于放置在一对主要声音重放换能器(例如在立体声扬声器装备中的左扬声器和右扬声器或在环绕声扬声器装备中的毗邻的扬声器)之间的一个或多个补充声音重放换能器,所述系统包括根据本发明的所述设备中的至少两个,其中所述设备中的第一个被提供

有第一左和右输入信号,并且其中所述第一设备提供在左输出终端、右输出终端和补充输出终端上的输出信号,在所述补充输出终端上的输出信号被提供至补充声音重放换能器,并且在所述左输出终端和右输出终端上的输出信号被分别提供至随后的根据本发明的设备的各自的输入信号,由此输出信号被提供至多个补充声音重放换能器中的分别的换能器。这样的系统的非限制性的实施例已经在上文描述。

[0107] 附图简述

[0108] 通过参照附图阅读本发明的实施环境的以下的详细描述,本发明将被更好地理解,在附图中:

[0109] 图 1 图示了用于立体声信号的重放的扬声器和收听者的理想的布置;

[0110] 图 2 示出了用于理想的立体声重放的作为频率的函数的 (a) 两耳间声强差 (ILD) 和 (b) 两耳间时间差。

[0111] 图 3 图示了相对于立体声扬声器对的离轴收听位置的情况;

[0112] 图 4 示出了用于离轴收听的作为频率的函数的 (a) 两耳间声强差 (ILD) 和 (b) 两耳间时间差;

[0113] 图 5 示出了收听区域坐标系和收听者的头部取向;

[0114] 图 6 图示了汽车收听环境;

[0115] 图 7 示出了 (a) 作为频率的函数的位置 1ILD, (b) 作为频率的函数的位置 1ITD, (c) 作为频率的函数的位置 2ILD 和 (d) 作为频率的函数的位置 2ITD;

[0116] 图 8 示出了对于汽车中收听的 (a) 作为频率的函数的位置 3ILD, (b) 作为频率的函数的位置 3ITD, (c) 作为频率的函数的位置 4ILD 和 (d) 作为频率的函数的位置 4ITD;

[0117] 图 9 示出了根据本发明的实施环境的立体向多-单通道转换器的框图,所述转换器包括分别用于左扬声器、中央扬声器和右扬声器的三个输出通道;

[0118] 图 10 示出了中央扬声器的地点和角度极限的实施例;

[0119] 图 11 示出了在收听方向旋转之后的中央扬声器的地点和角度极限;

[0120] 图 12 示出了 (a) $H_{IA \text{ 音乐}}(f)$ 的幅度、(b) $H_{IA \text{ 音乐}}(f)$ 的相位延迟。

[0121] 图 13 示出了 (a) $IDL_{\text{左极限}}$ 、(b) $ILD_{\text{右极限}}$ 、(c) $ITD_{\text{左极限}}$ 、和 (d) $ITD_{\text{右极限}}$;

[0122] 图 14 示出了对于《Bird on a Wire》的 512 个样本的块的左通道和右通道之间的相干性;

[0123] 图 15 示出了在 -10° 和 $+10^\circ$ 的源的 ILD 阈值以及 $H_{IA \text{ music}}(f)$ 的幅度。

[0124] 图 16 示出了 $ILD_{\text{音乐}}$ 到滤波器的映射;

[0125] 图 17 示出了 $ILD_{\text{音乐}}$ 到滤波器的映射;

[0126] 图 18 示出了在 -10° 和 $+10^\circ$ 的源的 ITD 阈值以及 $H_{IA \text{ 音乐}}(f)$ 的相位延迟。

[0127] 图 19 示出了 $ITD_{\text{音乐}}$ 到滤波器的映射;

[0128] 图 20 示出了 $ITD_{\text{音乐}}$ 到滤波器的映射;

[0129] 图 21 示出了 $H_{\text{中央}}(f)$ 的幅度;

[0130] 图 22 示出了 50Hz 正弦波的具有由随时间变化的滤波导致的不连续部的部分;

[0131] 图 23 示出了 $H_{\text{中央}}(f)$ 的 $1/3$ 八音度平滑的幅度;

[0132] 图 24 示出了对于两个毗邻的分析块的 $H_{\text{中央}}(f)$ 的幅度;

[0133] 图 25 示出了在转换速率极限之后的对于两个毗邻的分析块的 $H_{\text{中央}}(f)$ 的幅度;

[0134] 图 26 示出了 50Hz 正弦波的具有由转换速率极限导致的减少的不连续部的部分；
[0135] 图 27 示出了 $H_{\text{中央}}(k)$ 的脉冲响应；
[0136] 图 28 示出了 (a) 线性卷积的输出, 以及 (b) 循环卷积的输出；
[0137] 图 29 示出了具有零填充的 (a) 线性卷积的输出, 以及 (b) 循环卷积的输出；
[0138] 图 30 示出了中央扬声器的地点和角度极限, 其中收听方向在一对主要扬声器之间的角度范围的外部。

[0139] 发明的详细描述

[0140] 在下文, 描述了根据本发明的设备, 也称为立体向多-单通道转换器的具体的实施环境。结合本实施环境的详细描述, 具体的数值, 例如与扬声器装备中的各个角度相关的数值, 在文字和图中使用, 并且偶尔在各种数学表达式中使用, 但是将理解, 这样的具体的值将仅被理解为构成实施例, 并且其他的参数值也将被本发明覆盖。这种转换器的基本的功能性原理将参照图 9 中示出的示意性的框图被描述。虽然图 9 中示出的实施环境可扩展至 n 个扬声器, 并且可以被应用于以多于两个通道编码的听觉场景, 但是在下文中描述的实施环境提供了提取用于除了一般的立体声重放系统的左扬声器和右扬声器 (“主要” 扬声器) 之外的一个补充扬声器的信号。如图 11 中所示的, 在下文的详细描述中, 一个补充扬声器 56 通常被放置为相对于 0° 方位方向被旋转并且在收听者的正中平面。图 10 中示出的环境构成一个具体的实施例, 其中 $v_{\text{收听}} (v_{\text{listen}})$ 等于零度方位角。

[0141] 再次参照图 9, 根据本发明的该实施环境的立体向多-单通道转换器 (和相应的方法) 包括五个主要功能, 在框图中被标记为 A 至 E。

[0142] 在功能块 A 中, 对两耳信号执行计算和分析, 以确定在到来的立体声信号 $L_{\text{source}}[n]$ 和 $R_{\text{source}}[n]$ (参考数字分别为 14 和 15) 中的特定的信号分量是否可归于包括用于重放音频信号的补充扬声器 56 的给定的方位区间。这样的区间在图 10 和 11 中图示, 相应于中央扬声器 56。

[0143] 在本实施环境中, 输入信号 14、15 在 HRTF 立体声源块 24 中被转换为相应的两耳信号, 并且基于该两耳信号, 立体声输入信号 14、15 中的每个信号分量的两耳间声强差 (ILD) 和两耳间时间差 (ITD) 在被称为 ILD 音乐 29 和 ITD 音乐 30 的块中确定。在块 25 和 26 中, 左角度极限和右角度极限分别基于在终端 54 (左范围)、53 (收听方向) 和 55 (右范围) 的相应的输入信号被设置 (例如 10 和 11 图中所示的)。HRTF 的相应的值在 27 和 28 中确定。这些 HRTF 极限在块 31、32、33 和 34 中被转换为相应的两耳间声强差和两耳间时间差的极限。功能块 A (参考数字 19) 的输出是立体声信号 14、15 的每个信号分量的 ILD 和 ITD 29、30 以及右和左 ILD 和 ITD 极限 31、32、33、34。这些来自功能块 A 的输出信号被提供至功能块 C (参考数字 21) 中的映射函数, 如在下文中描述的。

[0144] 输入的立体声信号 14、15 另外被提供至功能块 B (参考数字 20), 功能块 B 计算输入的立体声信号 14、15 的左信号 14 和右信号 15 之间的通道间相干性。所得到的相干性被提供至块 C 中的映射函数。

[0145] 功能块 C (21) 将在功能 A (19) 和 B (20) 中计算出的两耳间差值和相干性映射到滤波器 D (22) 中, 该两耳间差值和通道间相干性将被用于提取将由中央扬声器重放的输入信号 $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ (14、15) 的那些分量。因此, 抽取的基本概念是, 将以高概率度导致幻像源在补充扬声器 56 被定位的位置处或附近被感知的立体声信号分量将被引导至补充

扬声器 56。“附近”的意思实际上通过块 A(19) 中限定的角度极限确定,并且形成幻像源的可能性通过块 20 中确定的左右通道间相干性确定。

[0146] 图 9 中示出的本发明的实施环境的基本功能在下文更详细地描述。特定的计算和图涉及其中信号被针对另外的一个扬声器提取的实施例,所述另外的一个扬声器放置在分别以 +/-30 度方位放置的左扬声器和右扬声器之间的零度方位,相应于传统的立体声扬声器装备的这种装备如在图 10 中示意性地示出的。左范围、收听位置和右范围输入信号 54、53、55 的相应的值在此选择为 -10 度、0 度、+10 度方位,相应于图 10 中示出的情况。

[0147] 功能 A:两耳信号的计算和分析

[0148] 第一步骤由通过将来自立体声信号源的输入的立体声信号 $l_{source}[n]$ 和 $r_{source}[n]$ 与在 -30° 的源的自由场两耳脉冲响应 ($h_{-30^\circ L}[n]$ 和 $h_{-30^\circ R}[n]$) 和 $+30^\circ$ 的源的自由场两耳脉冲响应 ($h_{+30^\circ R}[n]$ 和 $h_{+30^\circ L}[n]$) 卷积来计算耳的输入信号 $l_{ear}[n]$ 和 $r_{ear}[n]$ 组成。时域卷积典型地以公式表达为第一序列的每个样本与另一个第二序列的时间反向形式 (time reversed version) 的乘积的总和,如在以下的表达式中示出的:

$$[0149] \quad l_{ear}[n] = \sum_{k=-\infty}^{\infty} l_{source}[n] h_{-30degL}[n-k] + \sum_{k=-\infty}^{\infty} r_{source}[n] h_{+30degL}[n-k]$$

$$[0150] \quad r_{ear}[n] = \sum_{k=-\infty}^{\infty} r_{source}[n] h_{+30degR}[n-k] + \sum_{k=-\infty}^{\infty} l_{source}[n] h_{-30degR}[n-k]$$

[0151] 这些信号相应于在上文描述的理想立体声的情况下的耳输入信号。

[0152] 中央扬声器意图重放听觉场景的被定位在根据角度变量左范围、右范围和收听方向 (也被称为 v_{Lrange} (v_{Lrange})、 v_{Rrange} (v_{Rrange}) 和 v_{Listen}) 计算的左角度极限 v_{Llimit} 和右角度极限 v_{Rlimit} 之间的部分,如在以下等式中:

$$[0153] \quad v_{Llimit} = v_{Lrange} - v_{Listen}$$

$$[0154] \quad v_{Rlimit} = v_{Rrange} - v_{Listen}$$

[0155] 在本具体的实施例中, v_{Lrange} 、 v_{Rrange} 分别是 $-/+10$ 度,并且 v_{Listen} 是 0 度。

[0156] 如果回放系统含有多个扬声器,那么角度变量左范围、右范围和收听方向允许获得的听觉场景的取向和宽度被操纵。图 11 示出了其中收听方向不是零度方位,结果导致听觉场景向左旋转 (当与图 10 中的环境比较时) 的实施例。对这些变量的改变可以由收听者明确地作出,或可以是收听者位置跟踪器 (例如由收听者佩戴的头部跟踪器) 的结果。

[0157] 此外,在图 30 中示出了更普遍的情况,其中收听方向在包括补充扬声器 56 的角度范围的外部。虽然没有详细地描述,但是这种情况也被本发明覆盖。

[0158] 根据在 v_{Llimit} 度的源的自由场两耳脉冲响应 $h_{v_{Llimit}degL}[n]$ 和 $h_{v_{Llimit}degR}[n]$ 以及在 v_{Rlimit} 度的源的自由场两耳脉冲响应 $h_{v_{Rlimit}degL}[n]$ 和 $h_{v_{Rlimit}degR}[n]$ 计算在每种情况下的 ILD 极限和 ITD 极限。

[0159] 在本实施环境中,在功能 A 至 D 中的信号分析的其余部分在上文描述的信号的 N 个样本的块的频率域表示上操作。使用矩形窗口。在下文描述的实施例中 $N = 512$ 。

[0160] 耳输入信号、音乐信号和两耳脉冲响应 (对于在自由场中的 0° 的源;这种处理是针对中央扬声器) 的块的频率域表示使用 DFT 计算,如在以下的等式中以公式表达的:

$$[0161] \quad L_{ear}[k] = \sum_{n=0}^{N-1} l_{ear}[n] e^{j(2\pi/N)kn}$$

$$[0162] \quad R_{ear}[k] = \sum_{n=0}^{N-1} r_{ear}[n] e^{j(2\pi/N)kn}$$

$$[0163] \quad L_{source}[k] = \sum_{n=0}^{N-1} l_{source}[n] e^{j(2\pi/N)kn}$$

$$[0164] \quad R_{source}[k] = \sum_{n=0}^{N-1} r_{source}[n] e^{j(2\pi/N)kn}$$

$$[0165] \quad H_{v_{limit}degL}[k] = \sum_{n=0}^{N-1} h_{v_{limit}degL}[n] e^{j(2\pi/N)kn}$$

$$[0166] \quad H_{v_{limit}degR}[k] = \sum_{n=0}^{N-1} h_{v_{limit}degR}[n] e^{j(2\pi/N)kn}$$

$$[0167] \quad H_{v_{limit}degL}[k] = \sum_{n=0}^{N-1} h_{v_{limit}degL}[n] e^{j(2\pi/N)kn}$$

$$[0168] \quad H_{v_{limit}degR}[k] = \sum_{n=0}^{N-1} h_{v_{limit}degR}[n] e^{j(2\pi/N)kn}$$

[0169] 然后,三个耳间传递函数如以下示出的被计算:

$$[0170] \quad H_{Lleftlimit}[k] = \frac{H_{v_{limit}degL}[k]}{H_{v_{limit}degR}[k]}$$

$$[0171] \quad H_{Lrightlimit}[k] = \frac{H_{v_{limit}degL}[k]}{H_{v_{limit}degR}[k]}$$

$$[0172] \quad H_{LAnisic}[k] = \frac{L_{ear}[k]}{R_{ear}[k]}$$

[0173] 如上文提到的, $ILD_{leftlimit}$ 、 $ILD_{rightlimit}$ 和 $ILD_{音乐}$ 根据合适的传递函数的幅度计算。相似地, $ITD_{leftlimit}$ 、 $ITD_{rightlimit}$ 、和 $ITD_{音乐}$ 根据合适的传递函数的相位计算。

[0174] 每个 FFT 仓 k 的中心频率 f 根据 FFT 采样数和采样率计算。用于以下的实施例的音乐信号是《Bird on a Wire》的在该音乐开始之后的样本 $n = 2049 : 2560$ 。参照图 12, 示出了 $ILD_{音乐}$ 和 $ITD_{音乐}$ 。

[0175] 参照图 13(左图), 示出了 $ILD_{leftlimit}$ 和 $ILD_{rightlimit}$ 。

[0176] 这些 ILD 和 ITD 函数是图 9 中的功能块 C(参考数字 21) 的映射步骤的输入的一部分。

[0177] 功能 B: 信号之间的相干性的计算

[0178] $l_{source}[n]$ 和 $r_{source}[n]$ 之间的相干性, 如上文提到的, 取在 0 至 1 之间的值, 根据两个信号的功率谱密度和它们的交功率谱密度计算。

[0179] $l_{source}[n]$ 和 $r_{source}[n]$ 的功率谱密度在频率域内可以作为谱与其复共轭的乘积计算, 如以下示出的:

$$[0180] \quad P_{LL}[k] = L_{source}[k] \cdot L_{source}[k]^*$$

$$[0181] \quad P_{RR}[k] = R_{source}[k] \cdot R_{source}[k]^*$$

[0182] $l_{source}[n]$ 和 $r_{source}[n]$ 的交功率谱密度在频率域内可以作为 $L_{source}[k]$ 与 $R_{source}[k]$ 的复共轭的乘积计算, 如以下示出的:

$$[0183] \quad P_{LR}[k] = L_{source}[k] \cdot R_{source}[k]^*$$

[0184] 相干性在频率域内可以借助于以下等式计算：

$$[0185] \quad C_{LR}[f] = \frac{|P_{LR}^2|}{P_{LL} \cdot P_{RR}}$$

[0186] C_{LR} 在此示出的实施例中在 8 个块上计算。

[0187] 如果 $l_{source}[n] = r_{source}[n]$ ，那么 C_{LR} 将在所有的频率都等于 1。如果 $l_{source}[n]$ 和 $r_{source}[n]$ 是两个独立的随机信号，那么 C_{LR} 将在所有的频率都接近 0。对于音乐的块的 $l_{source}[n]$ 和 $r_{source}[n]$ 之间的相干性在图 14 中示出。

[0188] 功能 C：将两耳间差和相干性映射至滤波器

[0189] 本功能块将在功能 A 和 B 中计算的两耳间差和相干性映射到滤波器，所述滤波器将用于提取 $l_{source}[n]$ 和 $r_{source}[n]$ 的将由中央扬声器重分的分量。基本思想是，ILD、ITD 和通道间相干性函数对总体滤波器的贡献关于某些阈值被确定，所述阈值根据意图被扬声器覆盖的角度范围来确定。在下文中，中央扬声器被分配 -10 至 +10 度的角度范围。

[0190] 将 ILD 映射至滤波器幅度

[0191] ILD 阈值根据在 -10 和 +10 度的源的自由场的耳间传递函数确定。两种不同的计算 ILD 对最终的滤波器的贡献的方式在下文简要地描述。

[0192] 在第一映射途径中，任何具有在极限外部的幅度的频率仓都被衰减，如可以在图 15 中看到的。理想地，衰减应当是无限的。实际上，衰减被限于 A dB，在本实施例中为 30dB，以避免人为因素被过滤，例如点击。这些人为因素将在下文被进一步讨论。ILD 到滤波器的这种类型的映射在图 16 中示出。

[0193] 一个可选择的方法是简单地使用在 0 度的源的 $H_{IAff}[f]$ 和作为滤波器幅度的 H_{IA} 音乐 $[f]$ 之间的幅度差异的负的绝对值，如图 17 中所示的。以这种方式，如果 H_{IA} 音乐 $[f]$ 和 $H_{IAff}[f]$ 之间的差越大，那么 H_{IA} 音乐 $[f]$ 被衰减得越多。不具有在以上的方法中的硬阈值，并且因此某些分量将流入毗邻的扬声器中。

[0194] 将 ITD 映射至滤波器幅度

[0195] 如在之前的段落中描述的，ITD 阈值根据分别在 -10 和 +10 度的源的自由场的耳间传递函数来确定。再次地，两种用于包括 ITD 对最终的滤波器的贡献的方法在下文描述。

[0196] 将在 0 度的源的 $H_{IAff}[f]$ 和 H_{IA} 音乐 $[f]$ 之间的相位差使用中央扬声器的 ITD 阈值作图，如图 18 所示。

[0197] 第一“硬阈值”映射途径的结果是图 19 中示出的滤波器幅度。其中 ITD 在由分别在 -10 和 +10 度的自由场源装备的阈值的外部的所有的频率仓在本实施例中衰减 30dB。

[0198] 另一个途径是基于每个频率仓的与分别在 -30 和 +30 度的自由场源比较的延迟百分比计算在每个频率仓的衰减。例如，如果在某个频率的最大延迟是 16 个样本，并且音乐块的 ITD 是 4 个样本，那么其总延迟的百分比将是 25%。然后衰减可以是总数的 25%。即，如果被允许的总衰减是 30dB，那么相关的频率仓将被衰减 18dB。

[0199] 以这种方式设计的滤波器幅度的实施例在图 20 中示出。

[0200] 将相干性映射至滤波器幅度

[0201] 因为强度和时移对于相干信号最好地起作用，所以立体向多 - 单通道转换的操作应当优选地将 $l_{source}[n]$ 和 $r_{source}[n]$ 之间的相干性考虑在内。当这些信号完全不相干时，没有信号应当被发送至中央通道。如果信号是完全地相干的并且不具有 ILD 和 ITD，那

么理想地, $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 的整个内容应当被发送至中央扬声器, 并且没有任何内容应当被发送至左扬声器和右扬声器。

[0202] 相干性在本实施中被用作比例因子并且在以下的段落中描述。

[0203] 功能 D: 滤波器设计

[0204] 用于中央扬声器的基本的滤波器 $H_{\text{中央}}(f)$ 作为 ILD 滤波器、ITD 滤波器和相干性的乘积被计算, 如下文的等式中以公式表达的。重要的是注意, 这是线性相位滤波器, 每个频率仓的虚部被设置为 0, 因为其不期望将相移引入音乐中。

[0205] $H_{\text{center}}[f] = \text{ILDMAP}_{\text{center}}[f] \cdot \text{ITDMAP}_{\text{center}}[f] \cdot C_{\text{LR}}[f]$

[0206] 结果是具有如图 21 中示出的幅度的滤波器。

[0207] $H_{\text{center}}[f]$ 对于每个块被更新, 即其是时变滤波器。这种类型的滤波器引入了失真, 如果块之间的不连续部过于大的话, 该失真可以被听到。图 22 示出了这种情况的一个实施例, 其中在样本 400 至 900 之间的 50Hz 正弦波的部分中的不连续部可以被观察到。

[0208] 两种用于减少失真的方式在本实现中应用。

[0209] 首先, 交叉频率平滑被应用于 $H_{\text{center}}[f]$ 。这减少了在毗邻的频率仓的滤波器幅度中的急剧的改变。这种平滑通过用任一侧的 1/3 倍频程的幅度的平均值代替每个频率仓的幅度而实现, 导致图 23 中示出的滤波器。注意, y 轴的比例与图 21 相比发生了改变。

[0210] 转换速率极限从一个块至下一个也被应用于每个频率仓的幅度。图 24 示出了本块和前一个块的 $H_{\text{center}}[f]$ 。约 15dB 的幅度差可以在约 1kHz 和 10kHz 看到。

[0211] 这些差的幅度将导致听上去像鼠标点击的可听到的失真。转换速率极限使用条件逻辑语句被实施, 其一个实施例在以下的伪代码中给出。

[0212] 算法 1 (用于限制滤波器的转换速率的伪代码):

[0213] if new value > (old value+maximum positive change) then

[0214] new value = (old value+maximum positive change)

[0215] else

[0216] if new value < (old value-maximum negative change) then

[0217] new value = (old value-maximum negative change)

[0218] end if

[0219] end if

[0220] 选择最大的正的和负的变化值是在失真和具有滤波器之间的一种权衡, 所述滤波器足够迅速地反应以表示 $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 之间的关系的最重要的随时间变化的性质。在本实施例中根据经验确定值并且 1.2dB 被发现是可接受的。图 25 示出了使用该 1.2dB 转换速率极限的在本块和前一个块的 $H_{\text{center}}[f]$ 之间的变化。

[0221] 再次地考虑在约 1kHz 和 10kHz 的区域。清楚的是, 仅高至转换速率极限的差被维持。图 26 示出了 50Hz 正弦波的部分, 其中交叉频率平滑和转换速率极限被应用于时变滤波器。在图 22 中清楚地可见的不连续部被很大地减少。滤波器的增益在该频率也发生变化的事实也从正弦波的水平已经变化的事实可以看出。如上文提到的, 在表示源材料中的通道间关系的精确度和避免来自时变滤波器的人为因素之间存在权衡。

[0222] 如果等效于循环卷积的快速卷积将被使用, 那么滤波器必须被转换为它们的时域形式, 使得时域混叠可以被合适地控制 (这将在下文更彻底地描述)。

[0223] 离散傅里叶逆变换被简称为 IDFT, 并且由以下的等式给出, 并且被称为 $H_{\text{centre}}[k]$ 的傅里叶综合等式, 获得其脉冲响应。

$$[0224] \quad h_{\text{center}}[n] = \frac{1}{N} \sum_{k=0}^{N-1} H_{\text{center}}[k] e^{-j(2\pi/N)kn}$$

[0225] 因为 $H_{\text{centre}}[f]$ 是线性相位, 所以 $H_{\text{中央}}[n]$ 是 N 个样本长度的非因果有限脉冲响应 (FIR) 滤波器, 这意味着其超前第一样本。这种类型的滤波器可以通过应用 $N/2$ 个样本的延迟变为因果的, 如图 27 中所示的。注意, 滤波器是关于样本 $N/2+1$ 对称的。相位点值 (tap value) 已经被归一化, 仅为了作图的目的。

[0226] 功能 E: 计算用于每个扬声器的信号

[0227] 使用重叠保留 (overlap-save) 法的快速卷积

[0228] 在时间域中用于卷积两个序列的时间与 N^2 成比例, 其中 N 是最长的序列的长度。而在频率域中, 用于卷积两个序列 (它们的频率响应的乘积) 的时间与 $N \log N$ 成比例。这意味着, 对于比约 64 个样本长的序列来说, 频率域卷积在计算上更高效并且因此有短语快速卷积。在两个方法的输出中具有重要的差异, 即频率域卷积是循环的。在图 28 的粗实线中示出的曲线是图 27 中的长度 $N = 512$ 的滤波器与长度 $M = 512$ 的 500Hz 正弦波的时间域卷积的输出序列。注意, 256 个样本预振铃是使线性相位滤波器有因果关系的原因。在这种情况下, 输出序列是 $(N+M)-1 = 1023$ 个样本长。图 28 中示出的淡曲线是同一个滤波器和正弦波快速卷积的输出序列并且仅 512 个样本长。应当在样本 512 之后到来的样本已经被循环地移位和添加到样本 1 至 511 中, 这种现象被称为时域混叠。

[0229] 时域混叠可以通过在傅里叶变换之前零填充序列而避免, 并且这是返回至上文的关于功能块 D 的段落中提到的滤波器的时间域表示的原因。图 29 中的粗曲线是图 27 中的长度 $N = 512$ 的滤波器与长度 $M = 1024$ 的 500Hz 正弦波的时间域卷积的输出序列。在这种情况下, 输出序列是 $(N+M)-1 = 1535$ 个样本长。图 29 中的细曲线是零填充到长度 $N = 1024$ 个样本的同一滤波器和仍然具有长度 $M = 1024$ 的正弦波快速卷积的输出序列。在此输出序列是 1024 个样本长, 然而, 与上文的情况形成对比, 输出序列的在与零填充相同位置的部分, 即样本 512 至 1024, 与时间域卷积的输出是相同的。

[0230] 保存该部分并且通过将 512 个样本沿着正弦波向前移位来重复该过程被称为快速卷积的重叠保留法, 并且等效于时间域卷积, 除了另外的 256 个样本延迟产生了与过滤过程相关联的总延迟 $\text{Filter_delay} = 512$ 个样本之外。为了对这种技术进行充分的解释, 参考 Oppenheim 和 Schaffer [1999, 587 页]。

[0231] 输出信号的计算

[0232] 待被中央扬声器重放的信号 $C_{\text{output}}[n]$ 使用以下等式计算:

$$[0233] \quad l_{\text{filtered}}[n] = \left(\frac{1}{N} \sum_{k=0}^{N-1} H_{\text{center}}[k] \cdot L_{\text{source}}[k] e^{-j(2\pi/N)kn} \right)$$

$$[0234] \quad r_{\text{filtered}}[n] = \left(\frac{1}{N} \sum_{k=0}^{N-1} H_{\text{center}}[k] \cdot R_{\text{source}}[k] e^{-j(2\pi/N)kn} \right)$$

$$[0235] \quad c_{\text{output}}[n] = l_{\text{filtered}}[n] + r_{\text{filtered}}[n]$$

[0236] 待分别被左扬声器和右扬声器重放的信号然后通过分别从 $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 减去 $C_{\text{output}}[n]$ 被计算, 如在以下的等式中示出的。注意, $l_{\text{source}}[n]$ 和 $r_{\text{source}}[n]$ 被延迟以将

滤波器延迟计算在内。

$$[0237] \quad l_{\text{output}}[n] = Z^{-\text{filter_delay}} \cdot l_{\text{source}}[n] - l_{\text{filtered}}[n]$$

$$[0238] \quad r_{\text{output}}[n] = Z^{-\text{filter_delay}} \cdot r_{\text{source}}[n] - r_{\text{filtered}}[n]$$

[0239] 在其中 $r_{\text{source}}[n] = -l_{\text{source}}[n]$ 的具体的情况下, 信号是负相关的, 并且容易表明, 所有的输出信号都将是零。在这种情况下, 交功率谱密度的相位的绝对值 $P_{\text{LR}}[k]$ 将等于 $\pi \forall k$, 并且相干性 $C_{\text{LR}}[k]$ 将等于 $1 \forall k$ 。在以下的伪代码中的条件语句被应用以确保 $l_{\text{output}}[n] = l_{\text{source}}[n]$, $r_{\text{output}}[n] = -l_{\text{source}}[n]$ 并且 $C_{\text{output}}[n] = 0$ 。

[0240] 算法 2(用于处理负相关的信号的伪代码):

[0241] if $C_{\text{LR}}[k]=1$ AND $\frac{|\text{phase}(P_{\text{LR}}[k])|}{\pi} = 1$ then

[0242] $C_{\text{LR}}[k] = 0$

[0243] end if

[0244] 此外在 $l_{\text{source}}[n]$ 或 $r_{\text{source}}[n]$ 任一个无声的情况下, 那么 $C_{\text{LR}}[k]$ 应当是零。然而, 可以具有防止这发生的数值问题。在本实现中, 如果 $P_{\text{LL}}[k]$ 或 $P_{\text{RR}}[k]$ 中任一个的值下降到低于 -140dB , 那么 $C_{\text{LR}}[k]$ 被设置为零。

[0245] 参考文献

[0246] [1] Albert S. Bregman. Auditory Scene Analysis(听觉场景分析). The MIT Press, Cambridge, Massachusetts, 1994.

[0247] [2] Søren Bech. Spatial aspects of reproduced sound in small rooms(在小房间中的再现声音的空间方面). J. Acoust. Soc. Am., 103:434-445, 1998.

[0248] [3] Jens Blauert. Spatial Hearing(空间听觉). MIT Press, Cambridge, Massachusetts, 1994.

[0249] [4] D. Hammershøj and H. Møller. Sound transmission to and within the human ear canal(人类耳道内的声音传输). J. Acoust. Soc. Am., 100(1):408-427, 1996.

[0250] [5] CIPIC Interface Laboratory. The cipic hrtf database(CIPIC 头部反应传递函数数据库), 2004.

[0251] [6] Allan V. Oppenheim and Ronald W. Schaffer. Discrete-Time Signal Processing(离散时间信号处理). Prentice-Hall, Upper Saddle River, 199.

[0252] [7] H. Tokuno, O. Kirkeby, P. A. Nelson and H. Hamada. Inverse filter of sound reproduction systems using regularization(使用正则化的声音再现系统的逆滤波器). IEICE Trans. Fundamentals, E80-A(5):809-829, May 1997.

[0253] [8] S. Perkin, G. M. Mackay, and A. Cooper. How drivers sit in cars(司机如何在车里就坐). Accid. Anal. And Prev., 27(6):777-783, 1995.

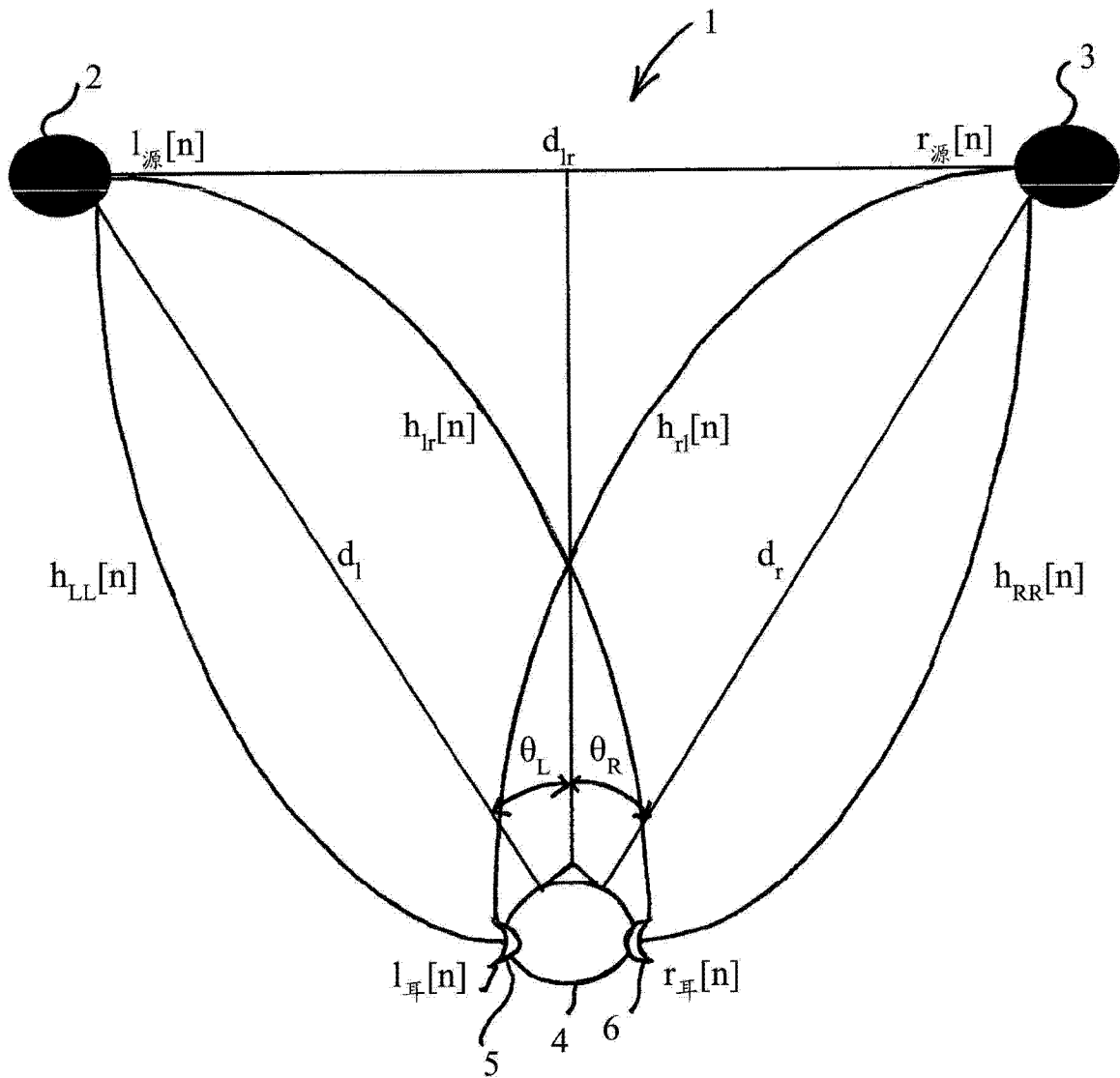


图 1

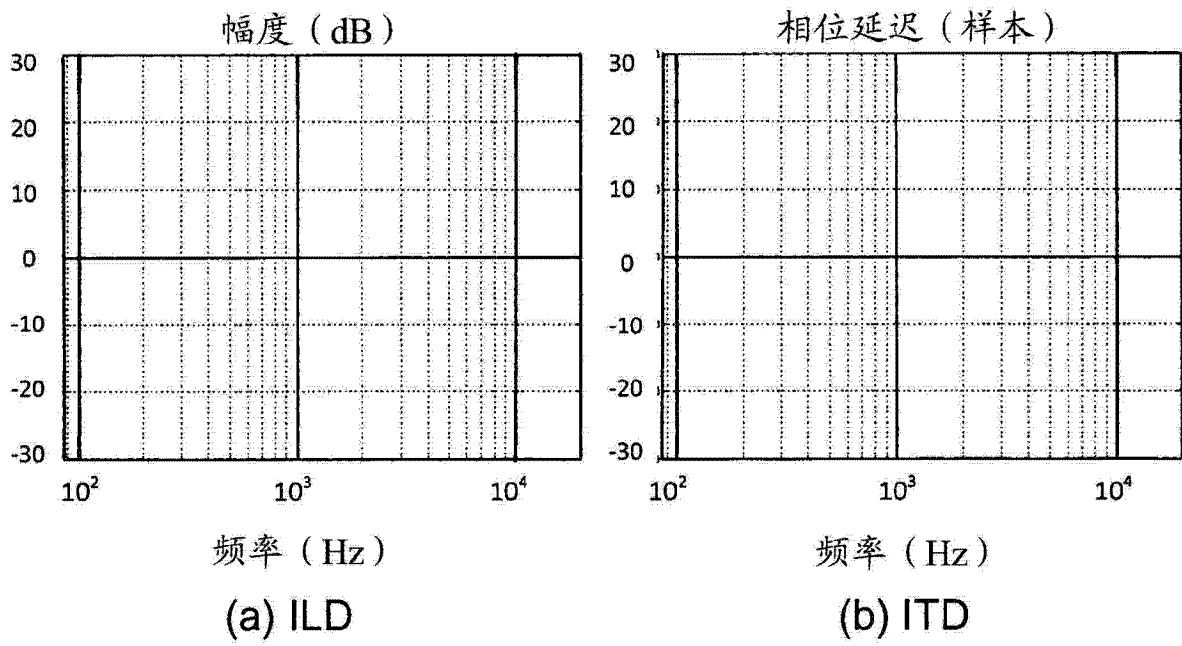


图 2

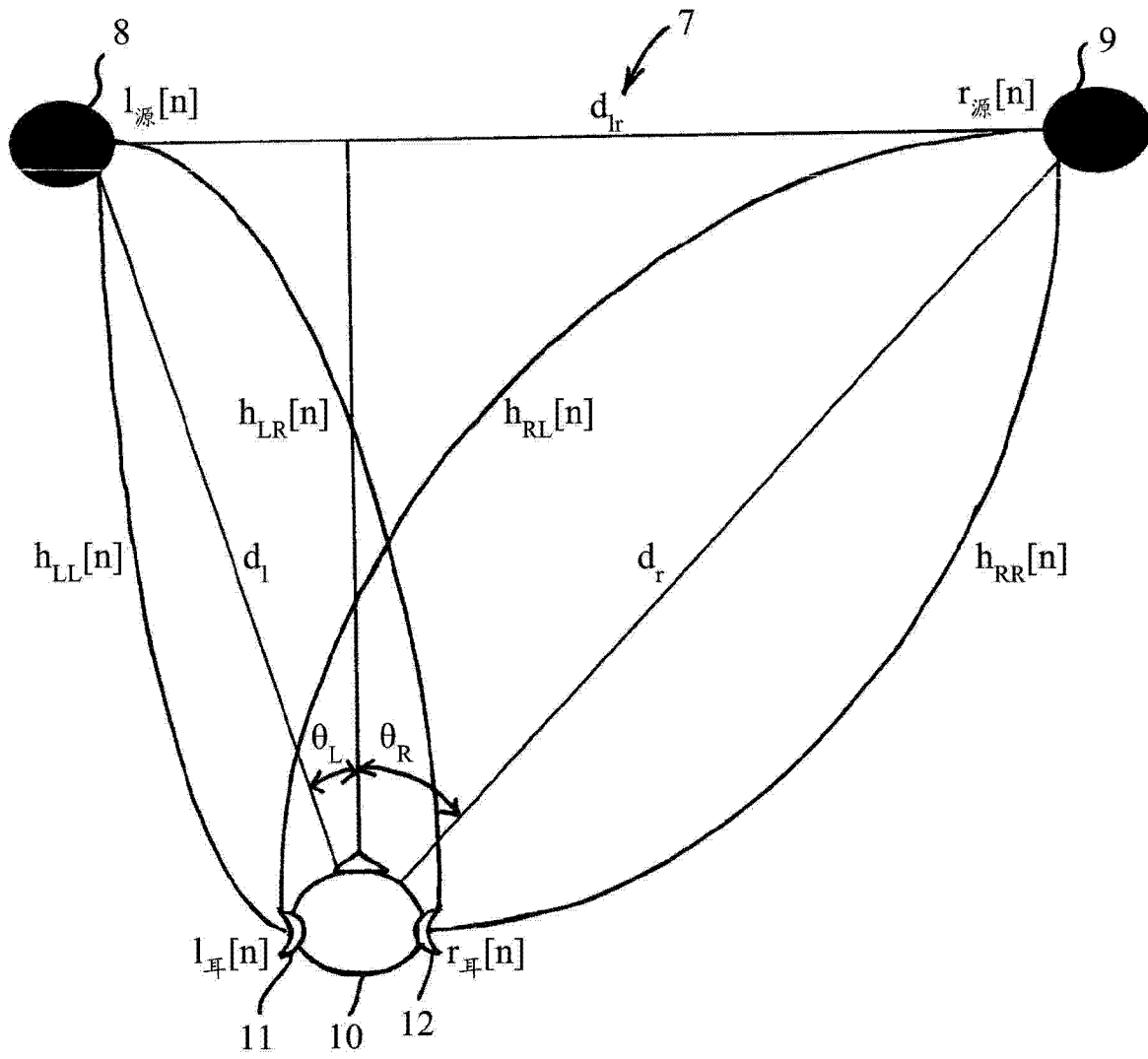


图 3

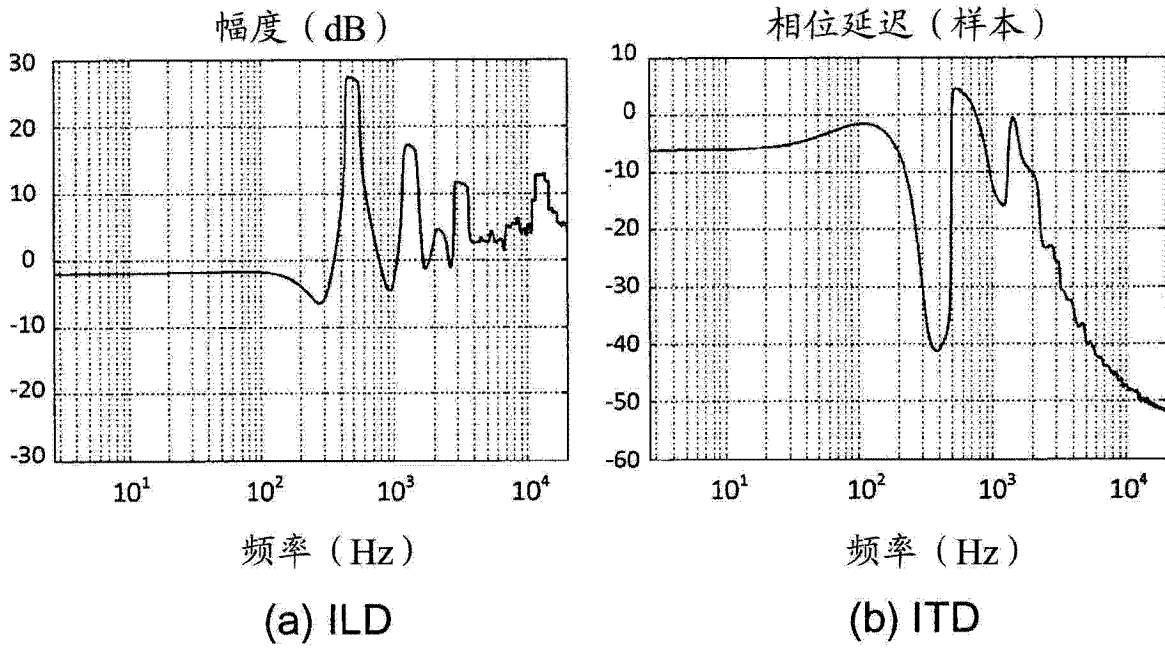


图 4

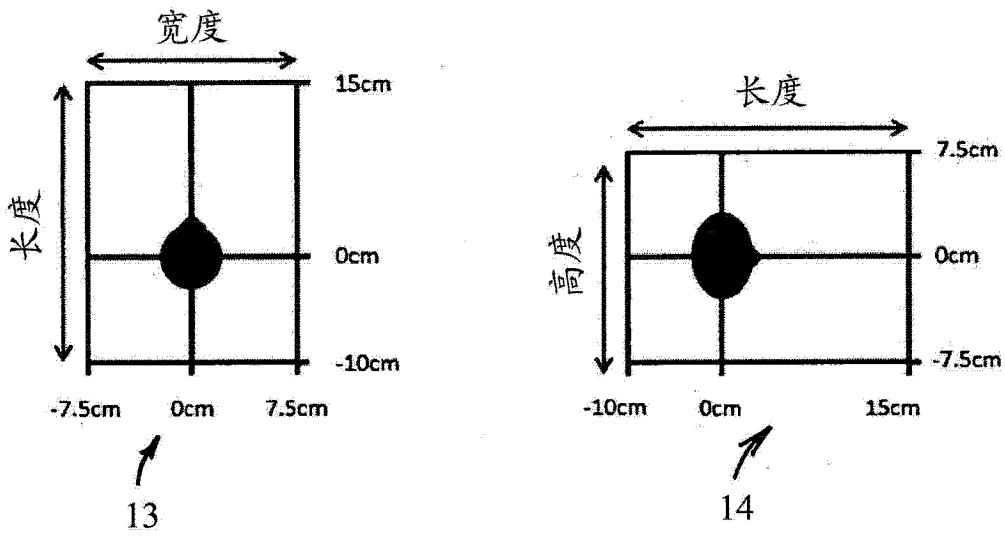


图 5

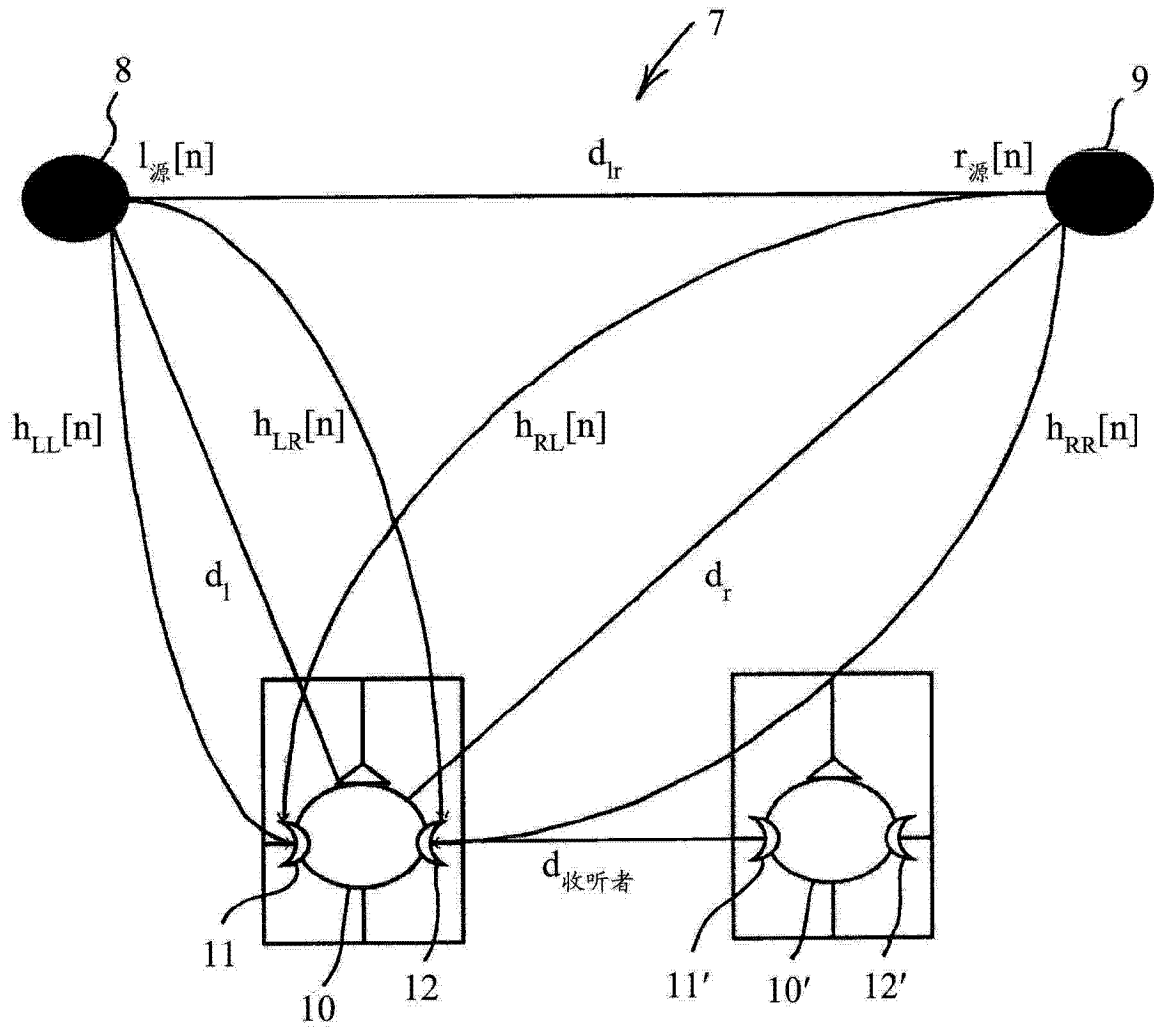


图 6

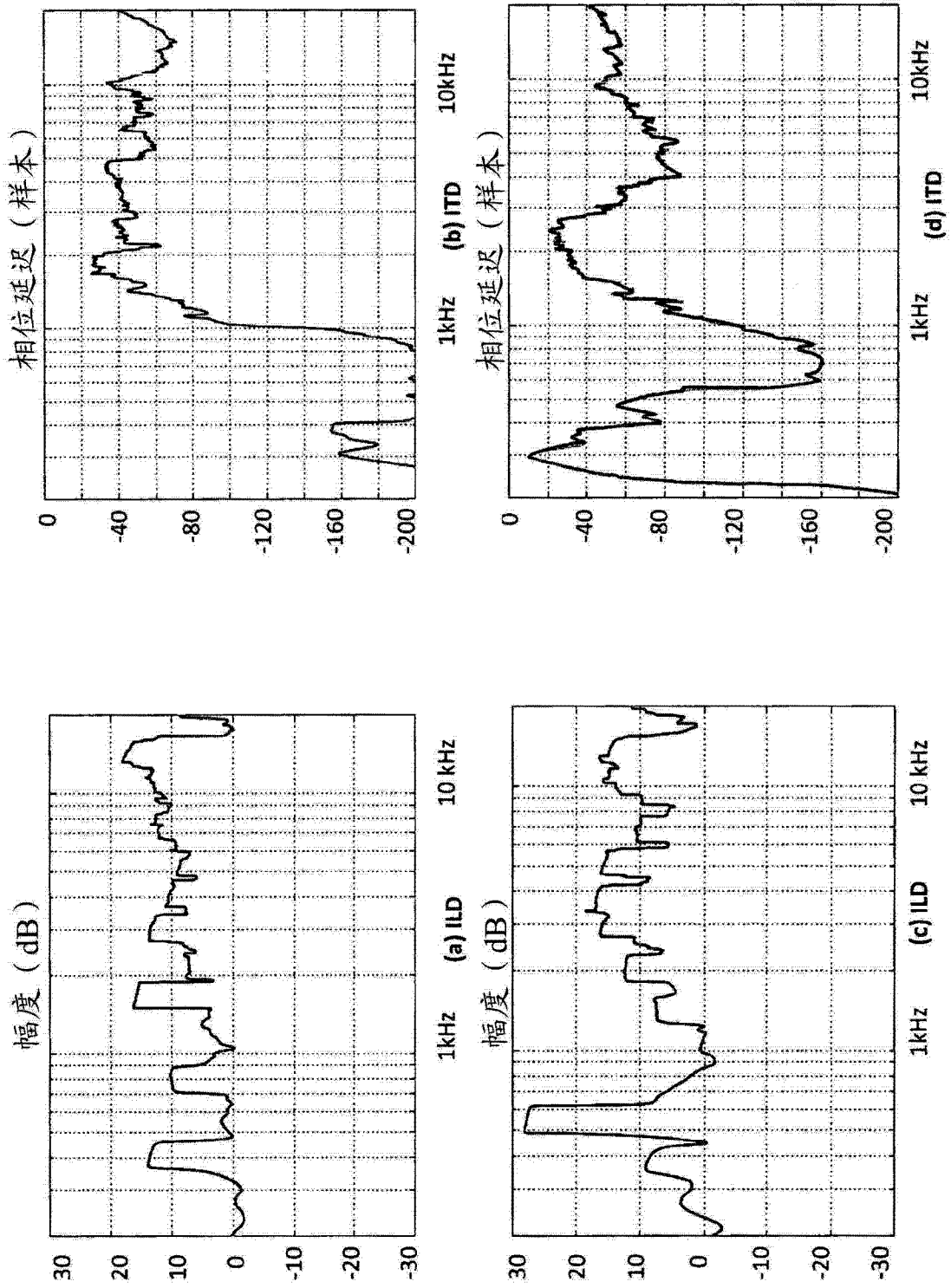


图 7

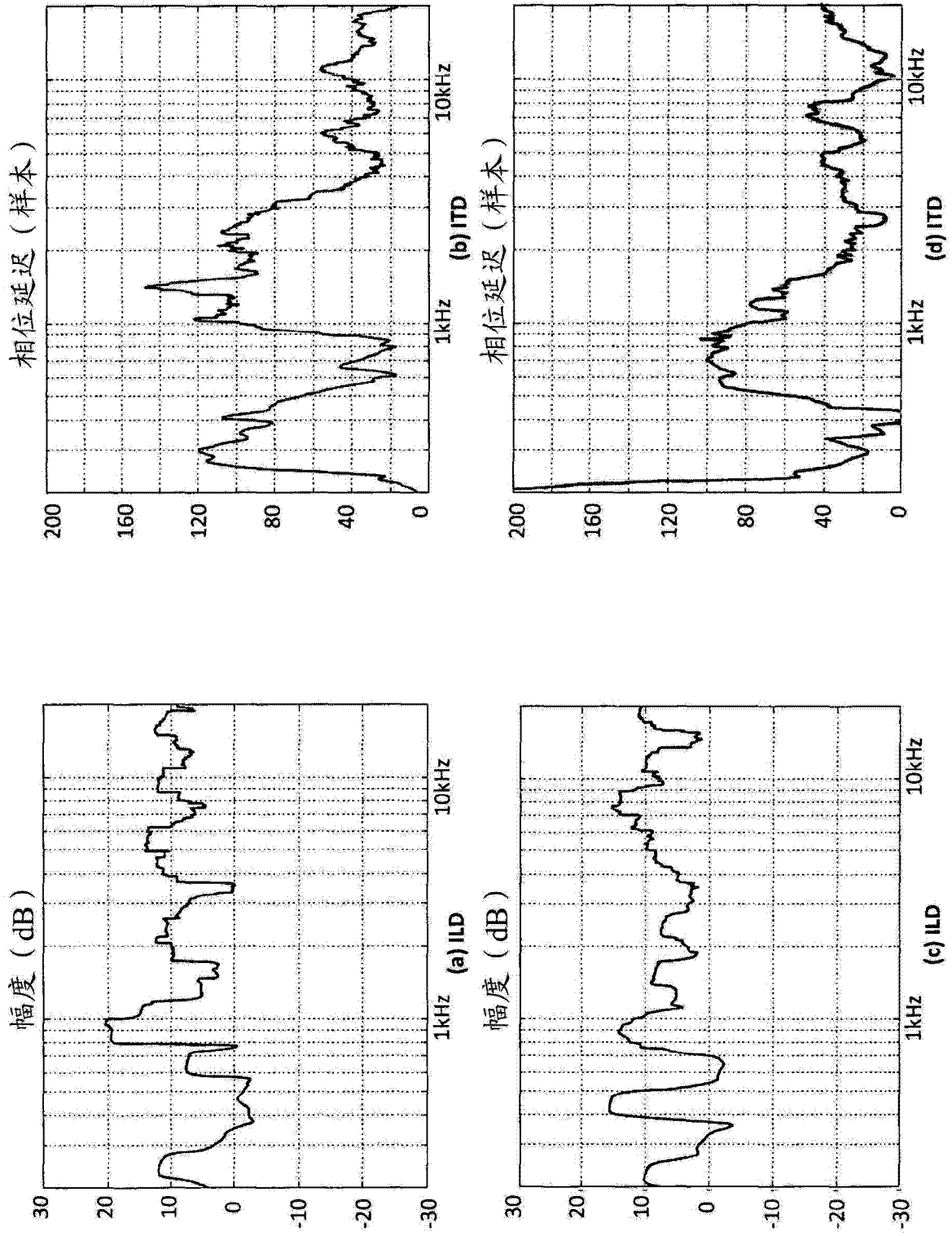


图 8

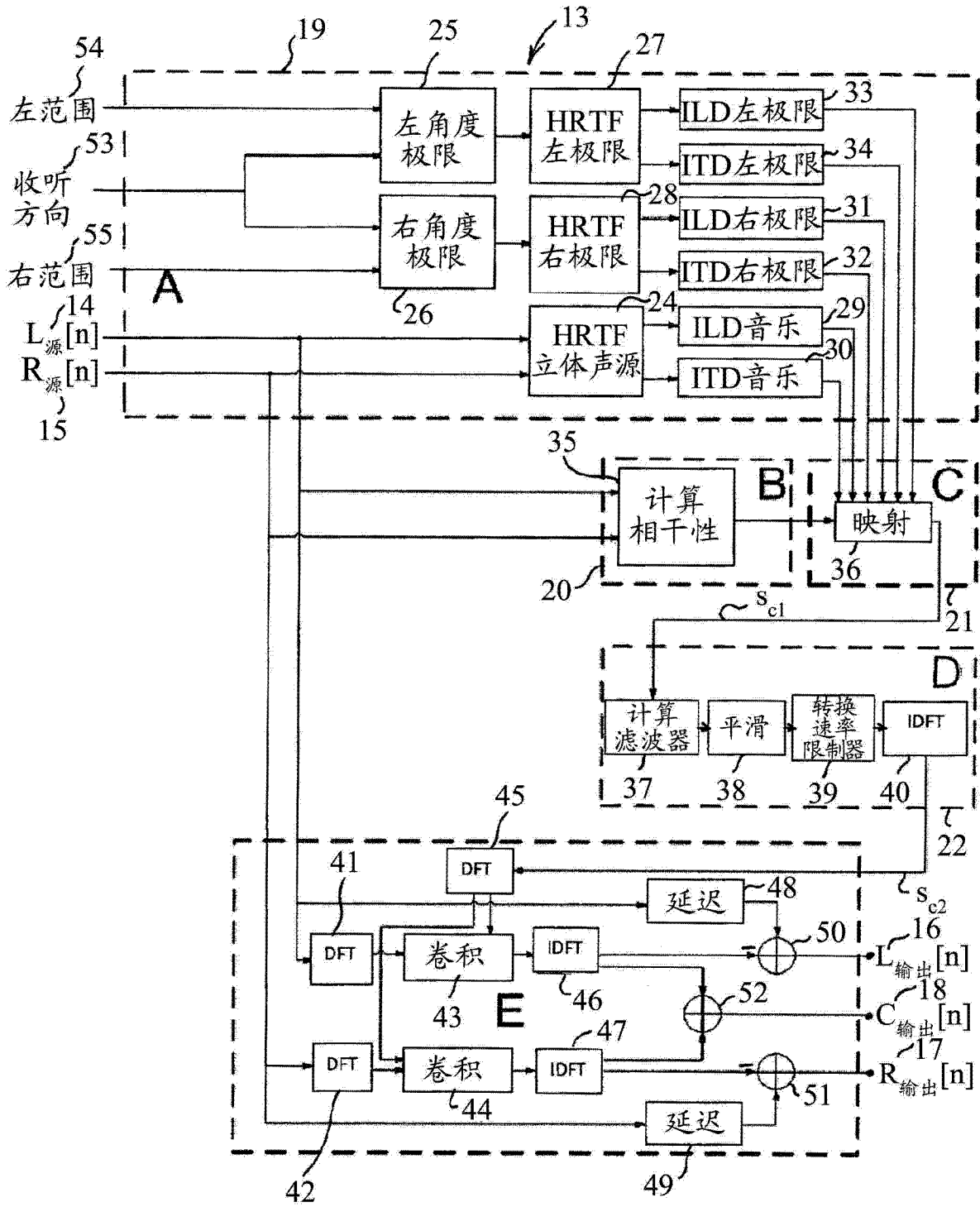


图 9

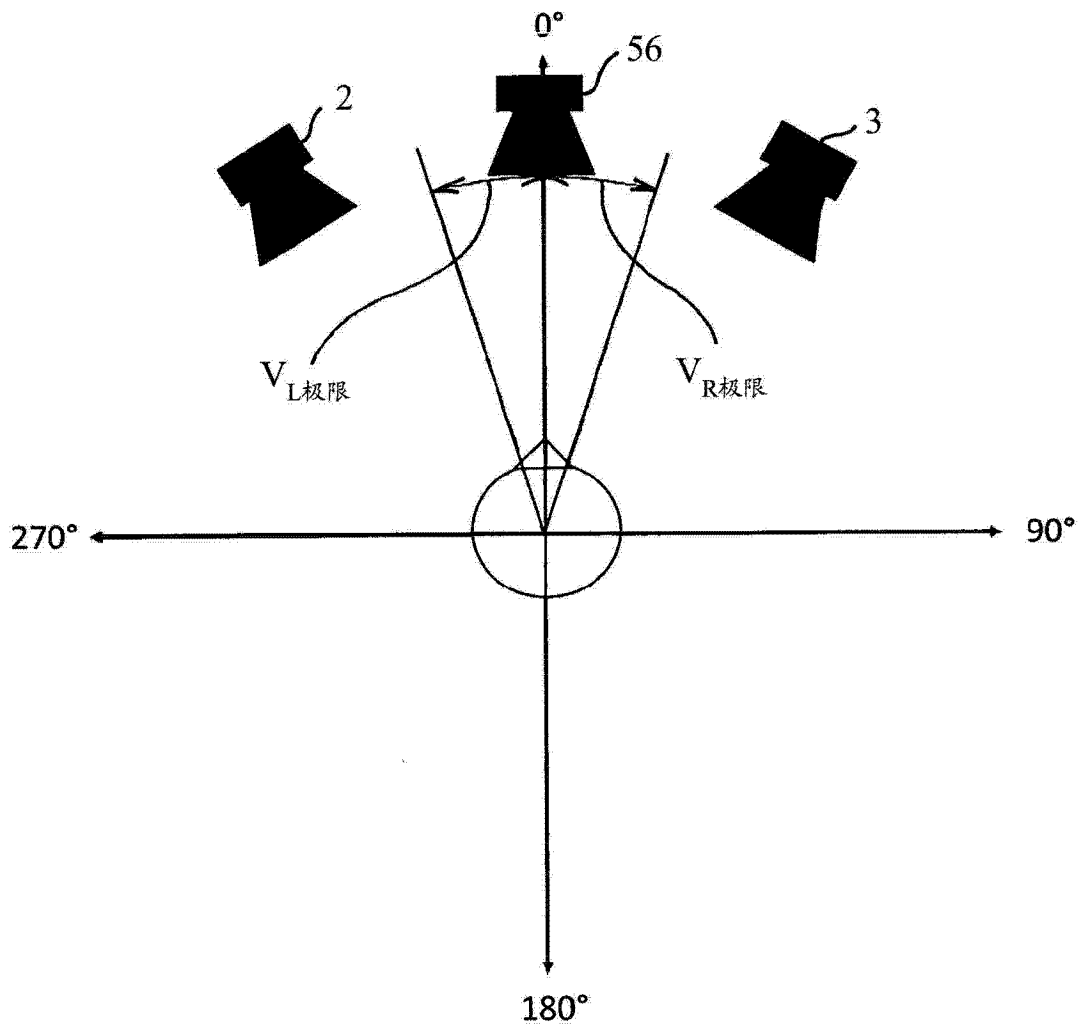


图 10

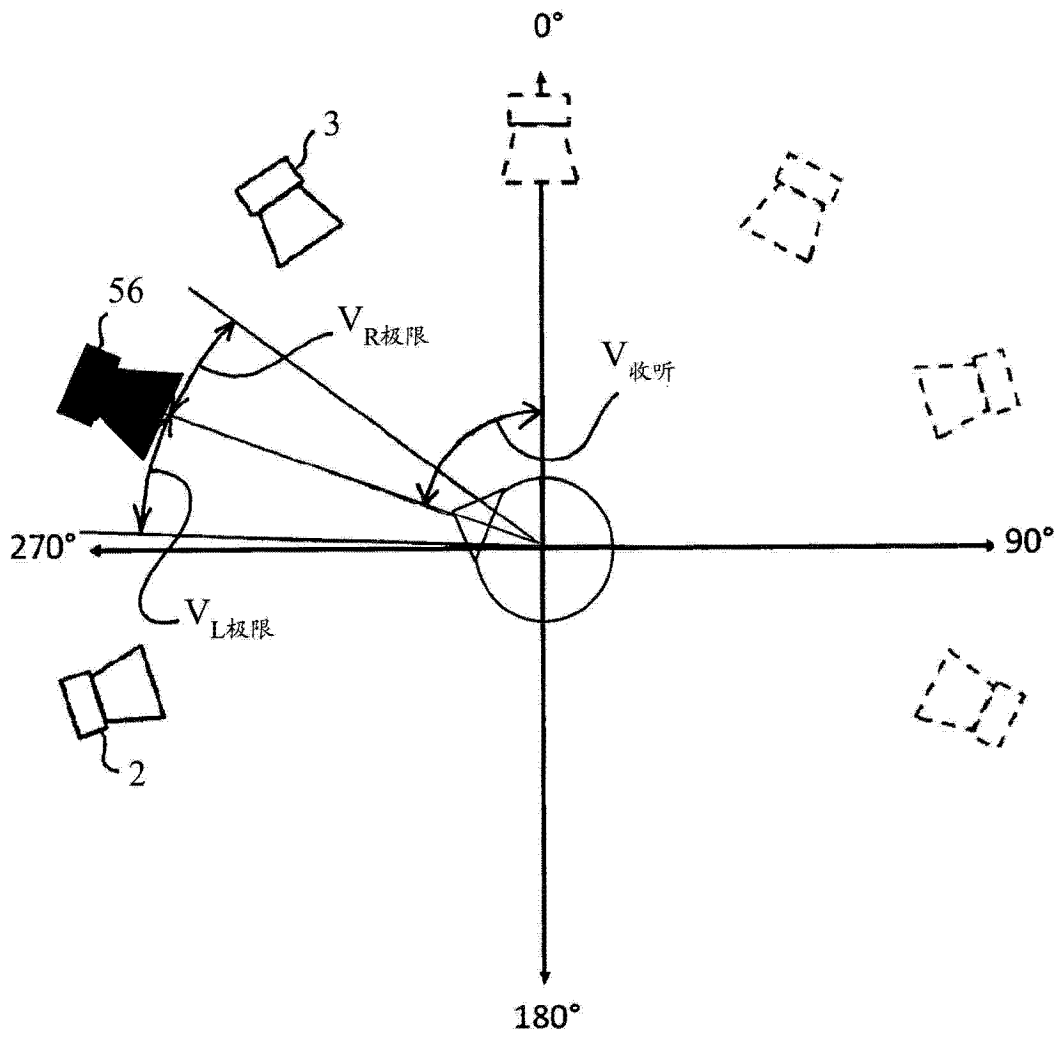


图 11

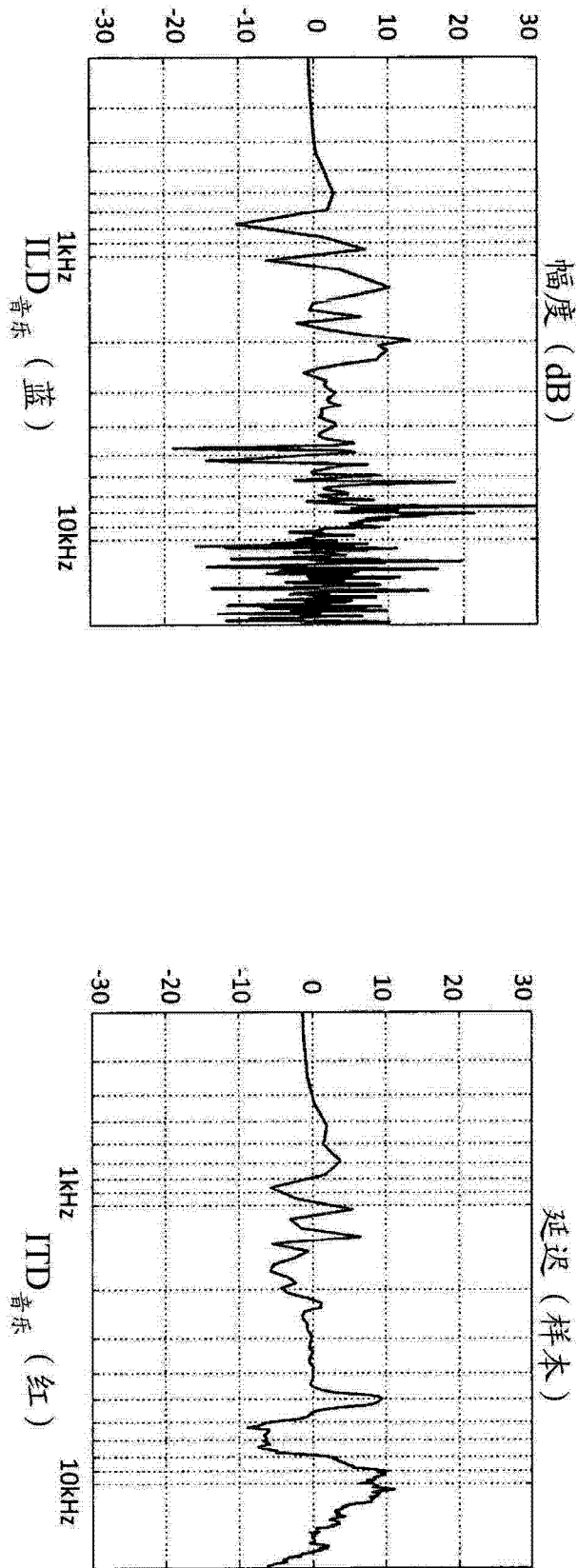


图 12

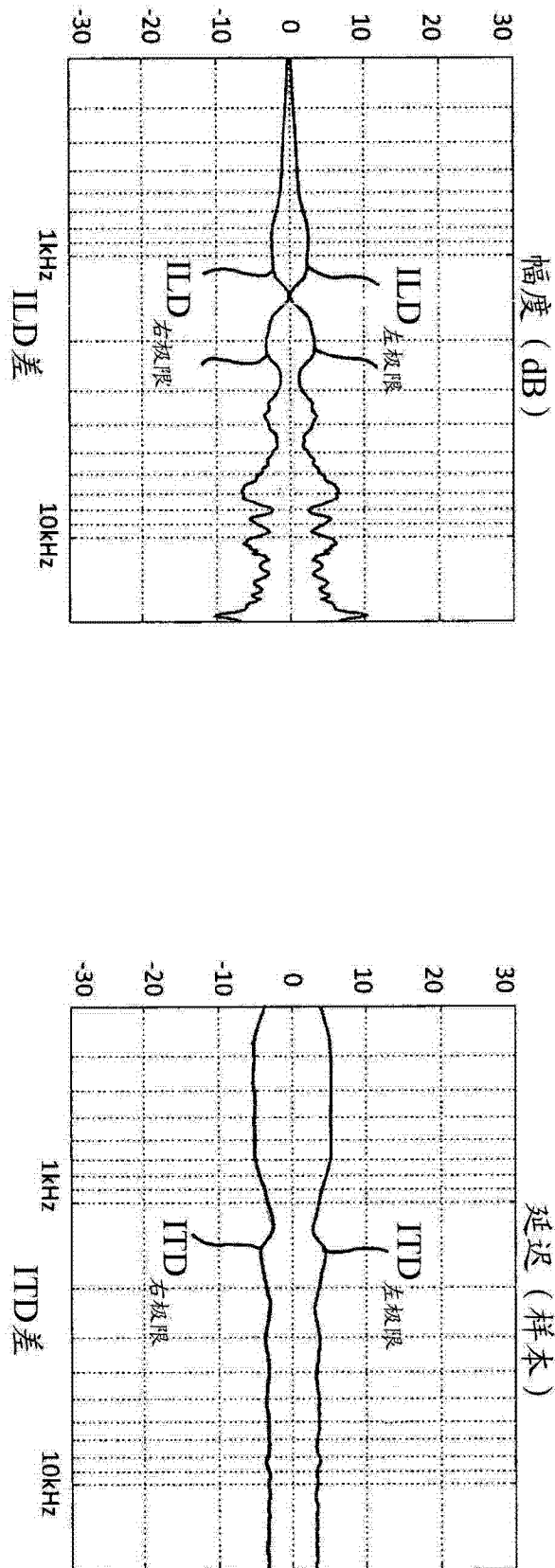


图 13

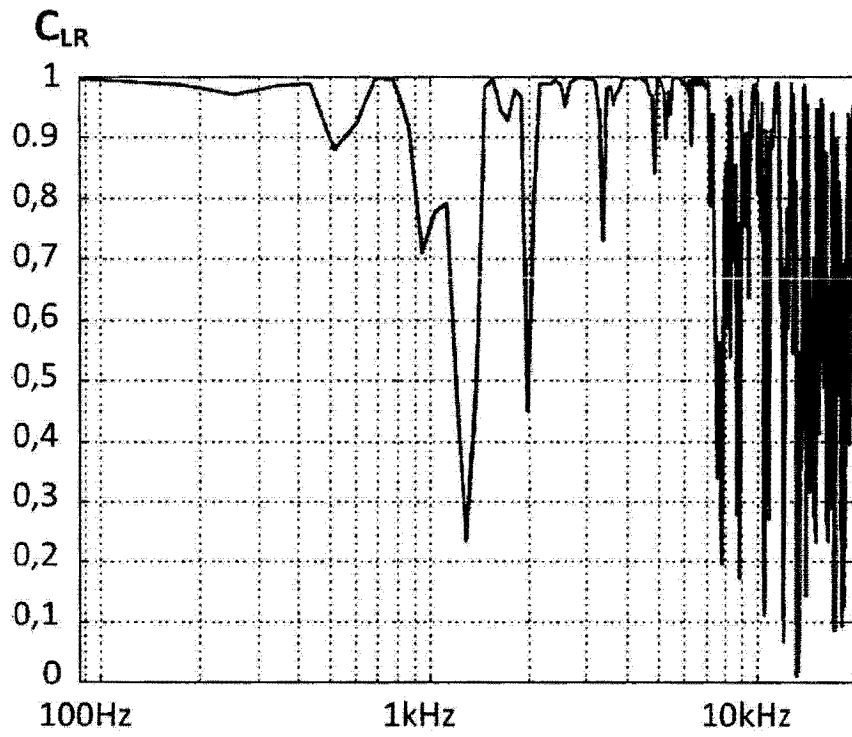


图 14

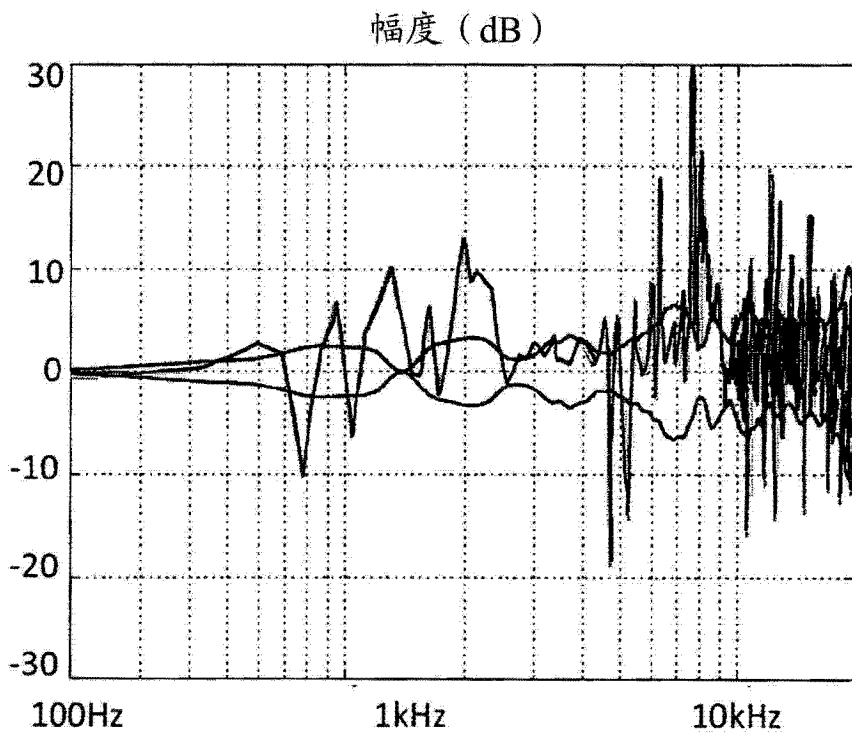


图 15

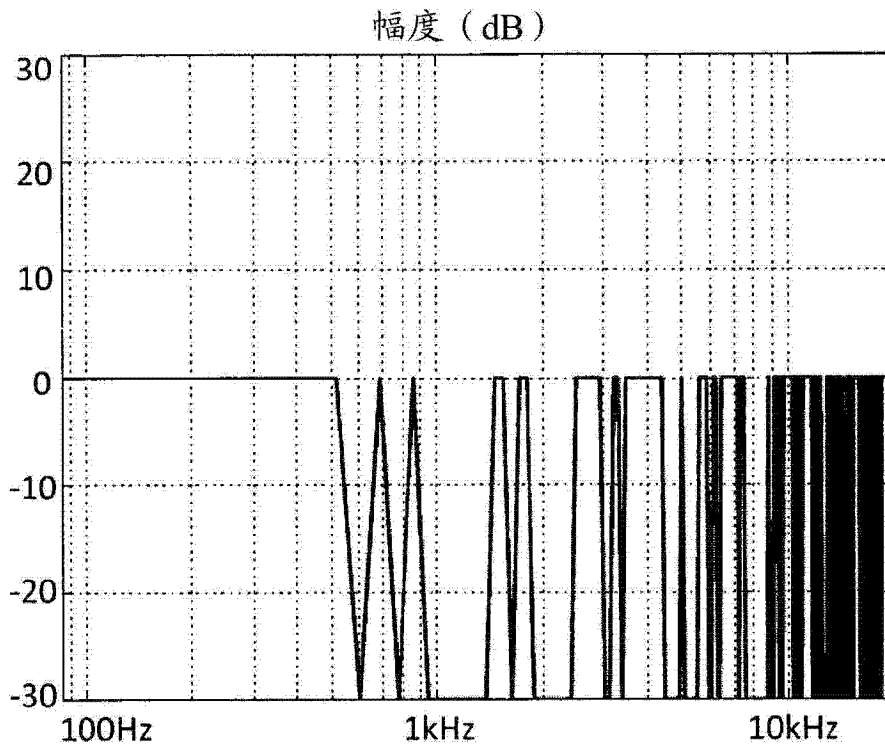


图 16

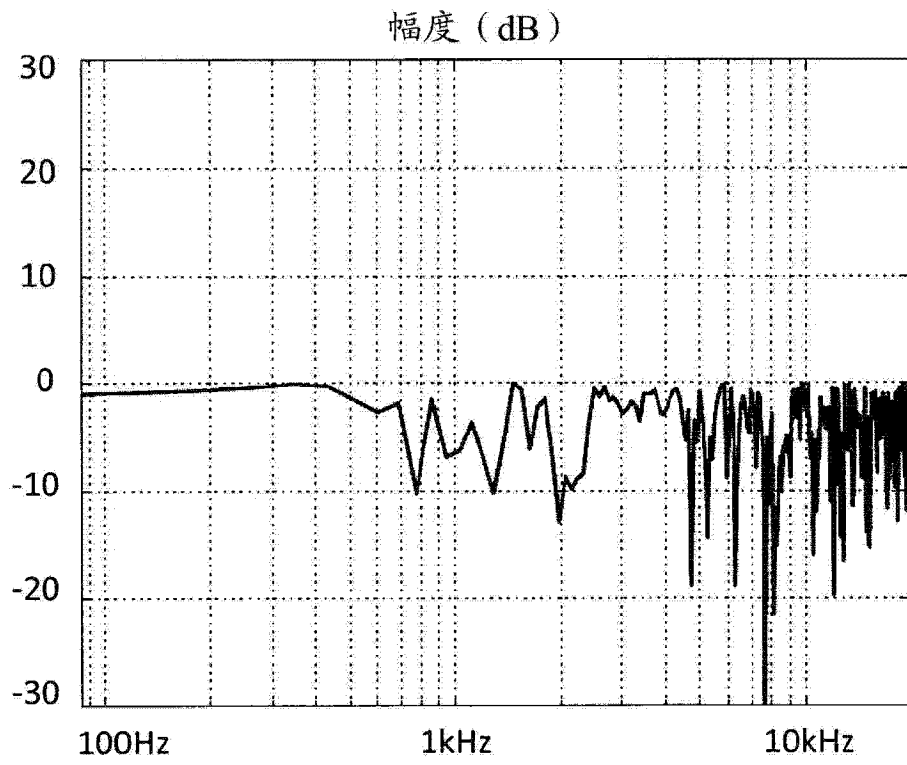


图 17

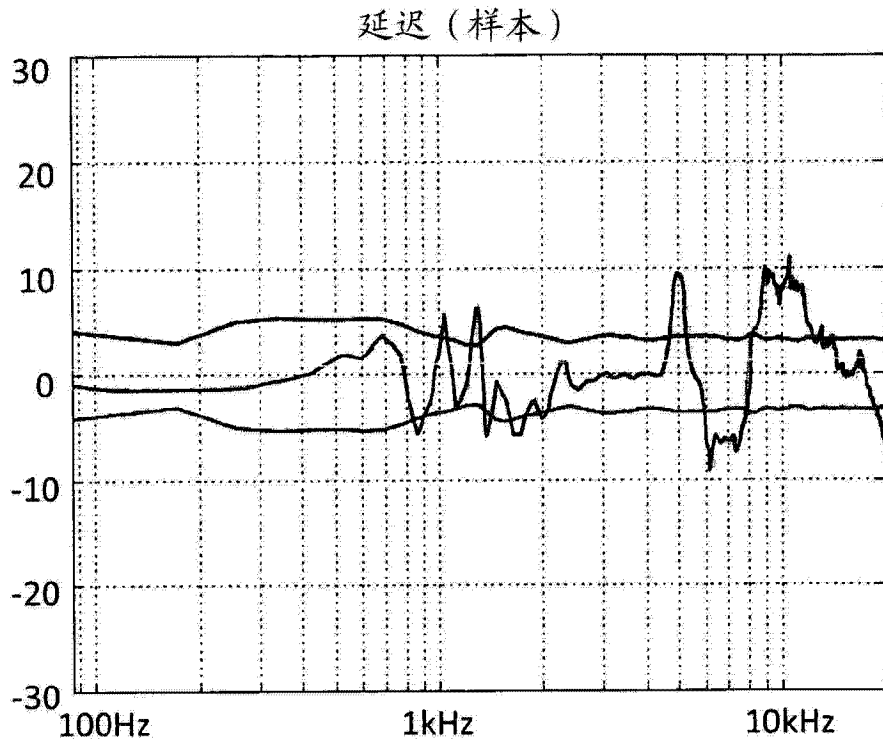


图 18

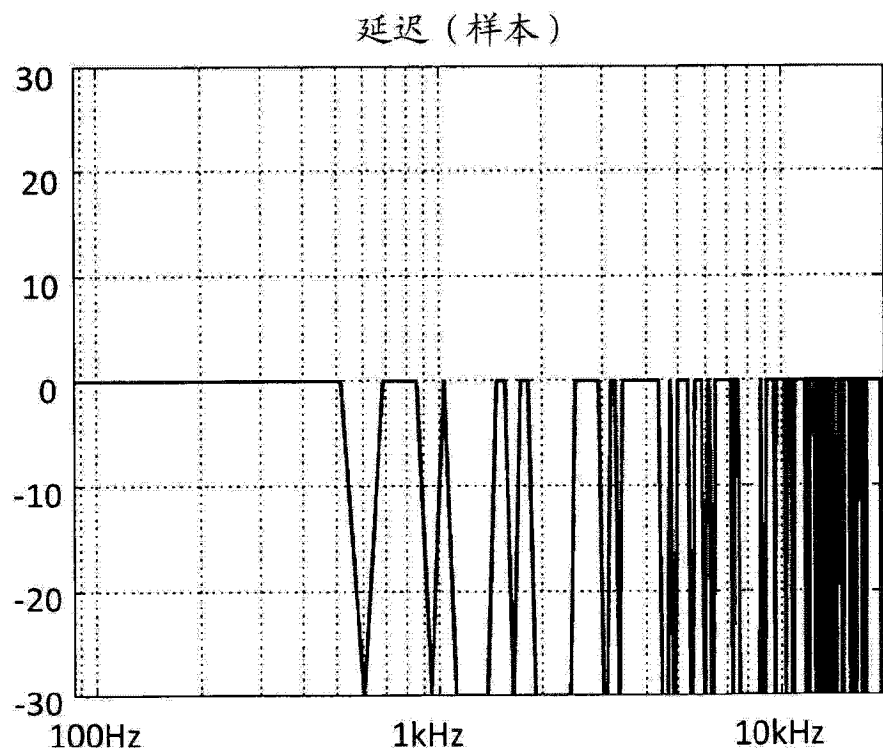


图 19

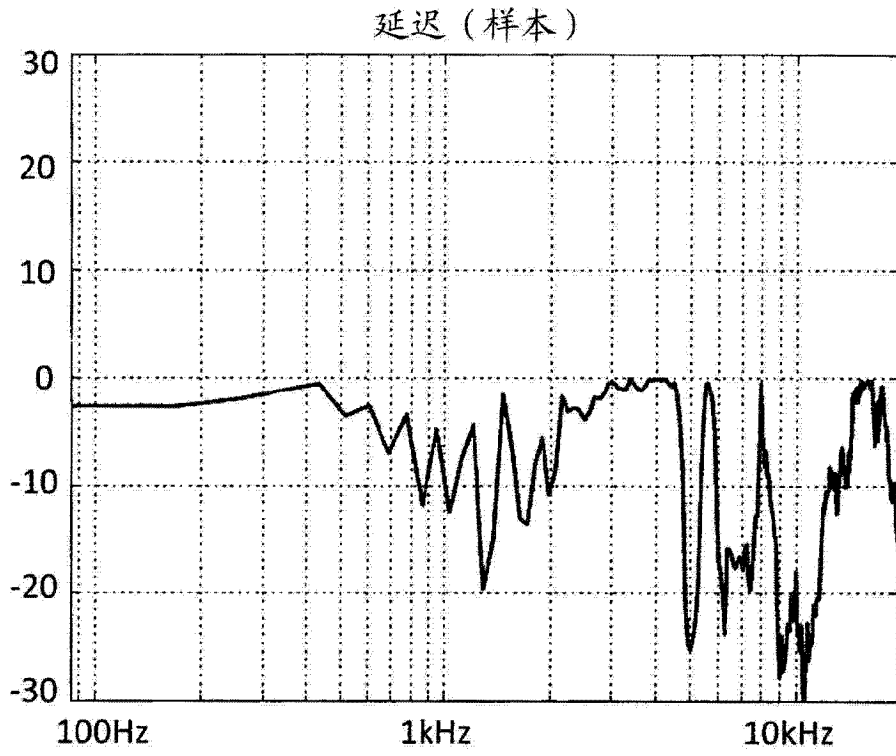


图 20

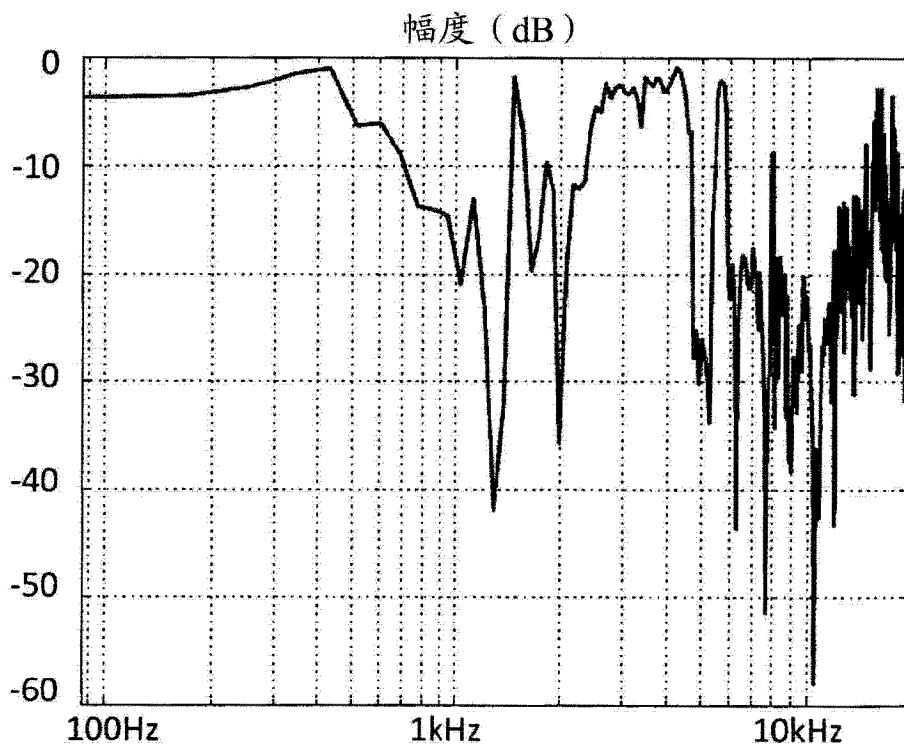


图 21

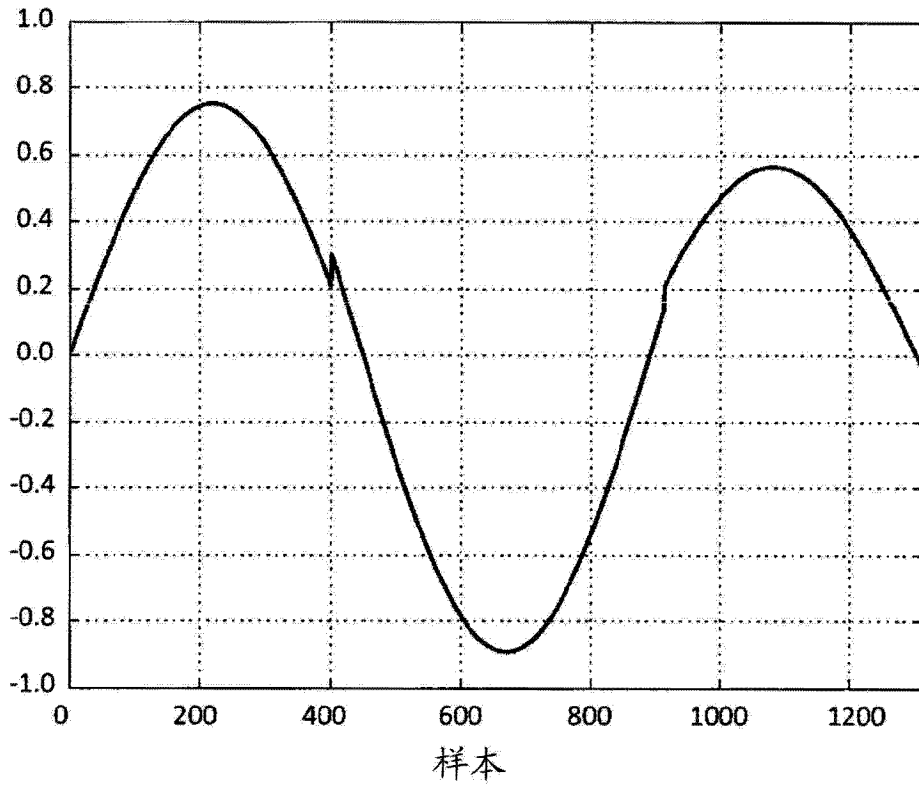


图 22

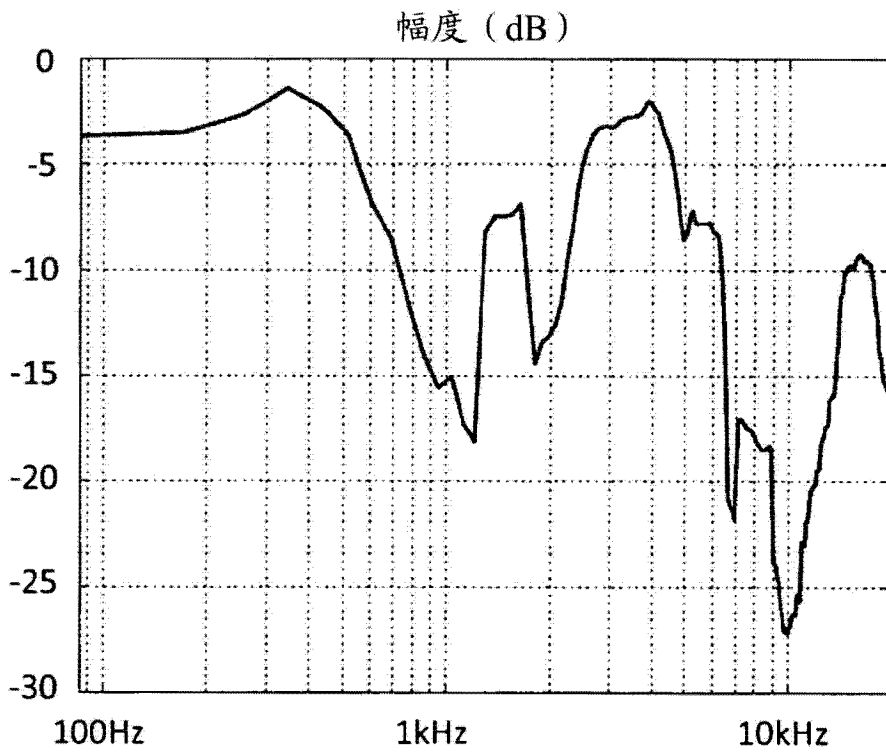


图 23

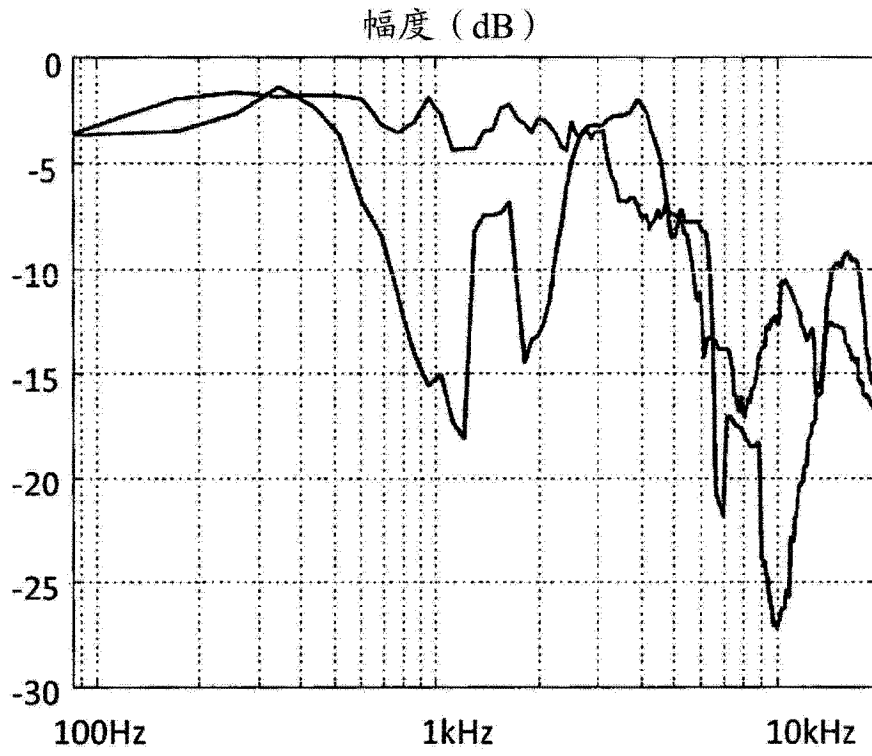


图 24

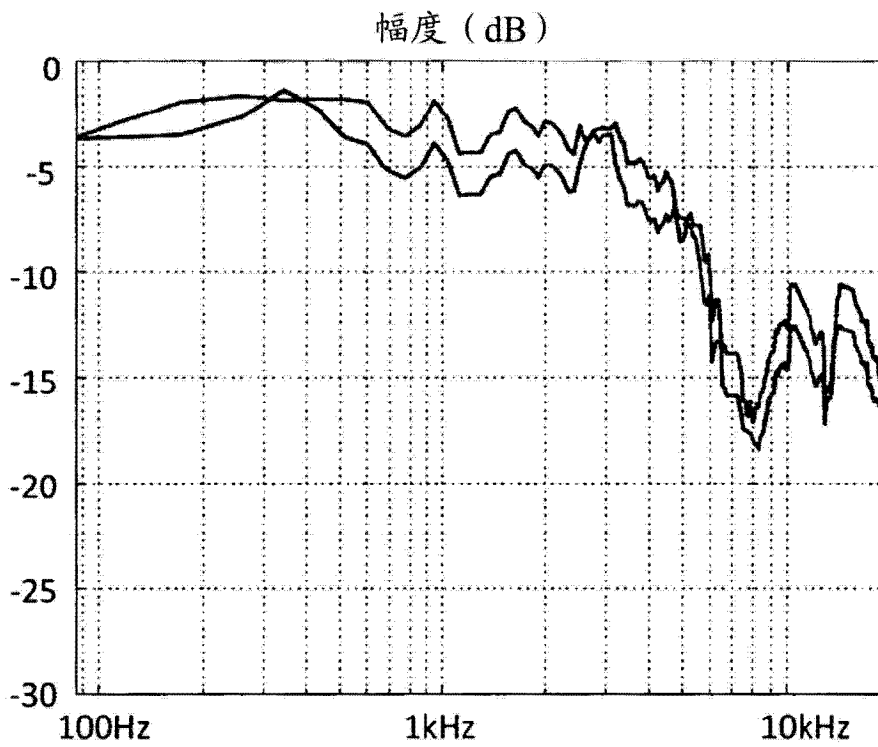


图 25

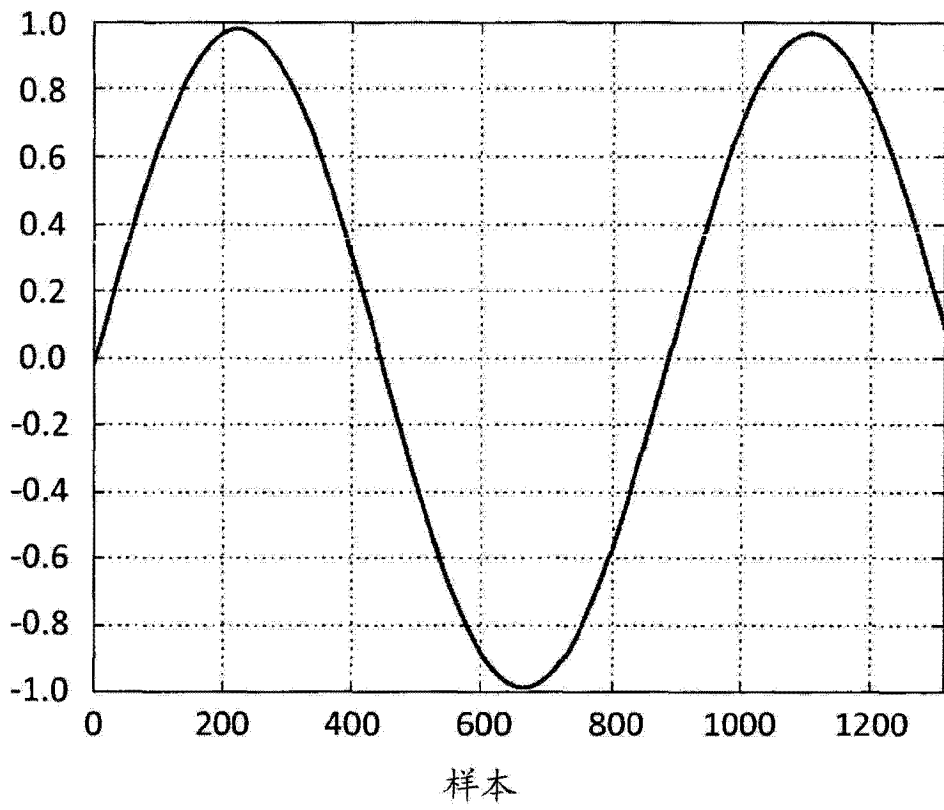


图 26

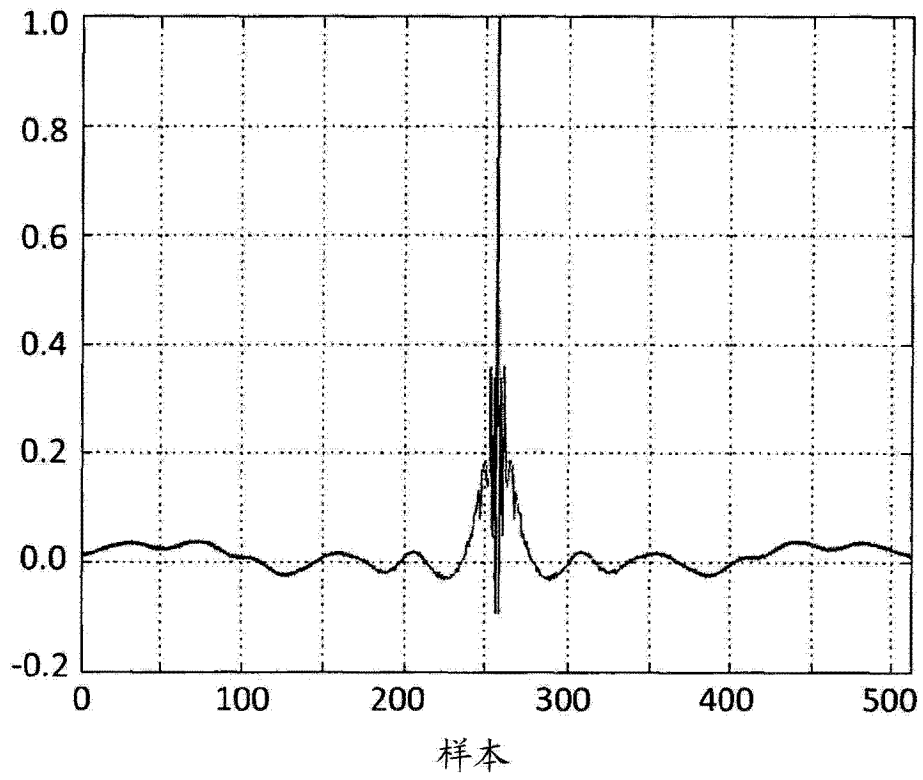


图 27

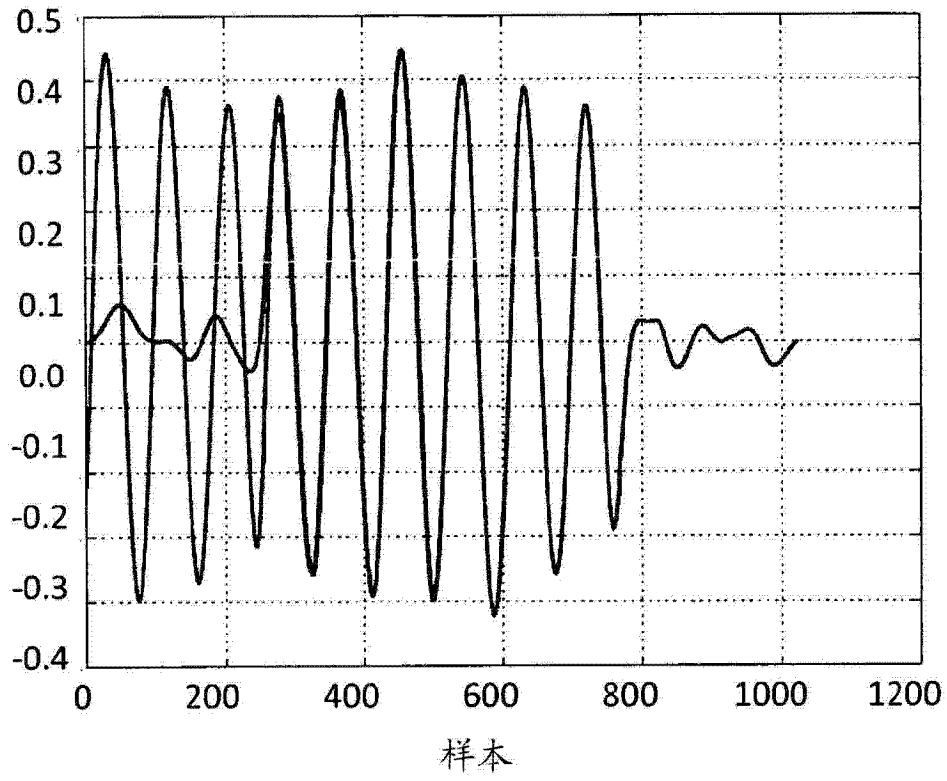


图 28

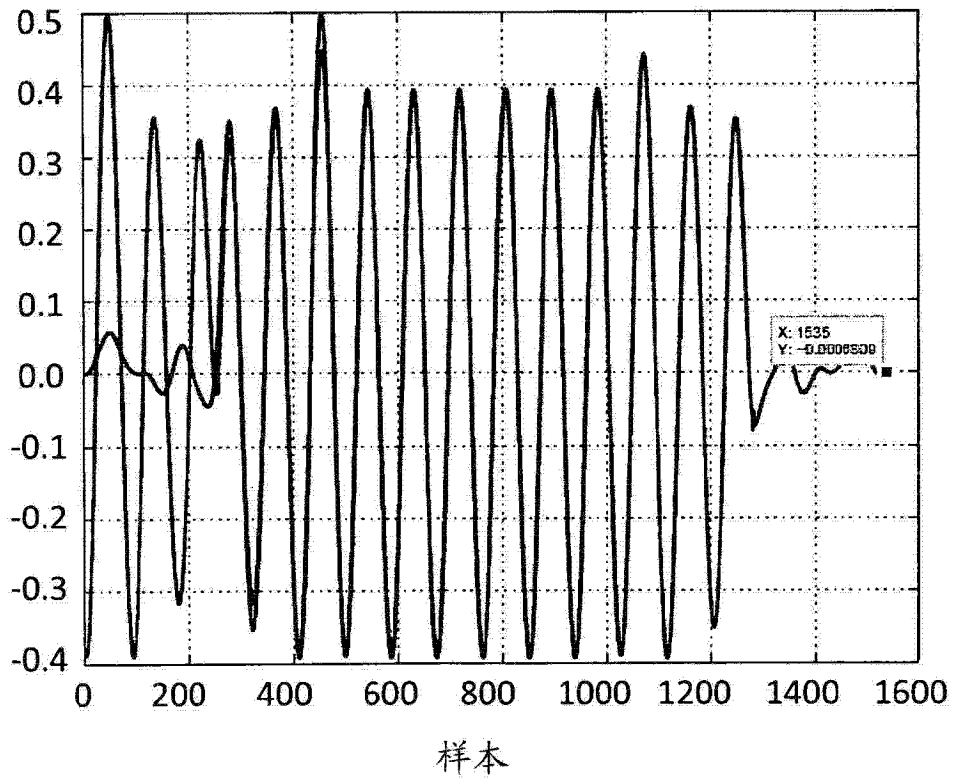


图 29

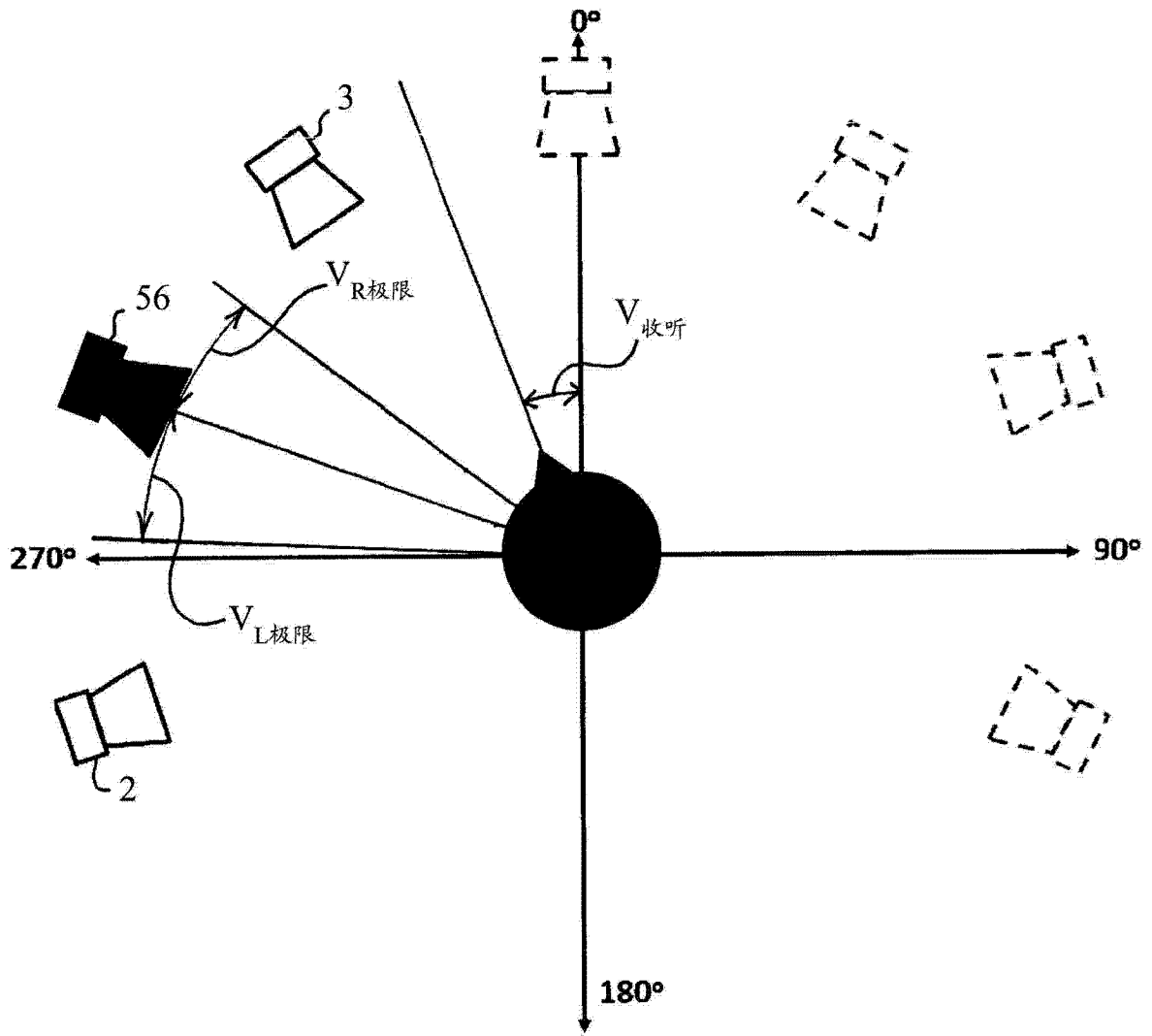


图 30