



(12) 发明专利申请



(10) 申请公布号 CN 118160031 A

(43) 申请公布日 2024. 06. 07

(21) 申请号 202280072317.9

(22) 申请日 2022.10.19

(30) 优先权数据

21204641.1 2021.10.26 EP

(85) PCT国际申请进入国家阶段日

2024.04.26

(86) PCT国际申请的申请数据

PCT/EP2022/078998 2022.10.19

(87) PCT国际申请的公布数据

W02023/072684 EN 2023.05.04

(71) 申请人 皇家飞利浦有限公司

地址 荷兰艾恩德霍芬

(72) 发明人 J·G·H·科庞

(74) 专利代理机构 永新专利商标代理有限公司
72002

专利代理师 刘瑜

(51) Int.Cl.

G10K 15/12 (2006.01)

H04S 1/00 (2006.01)

H04S 7/00 (2006.01)

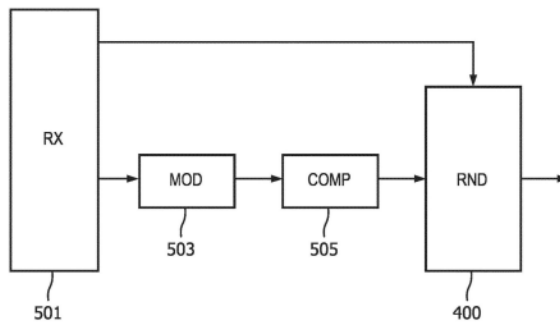
权利要求书2页 说明书29页 附图8页

(54) 发明名称

一种音频装置及其操作方法

(57) 摘要

一种音频装置,包括:接收机(501),其接收音频数据和包括用于针对环境的混响参数的数据的元数据。修改器(503)生成用于第一混响参数的修改后的第一参数值,所述第一混响参数是混响延迟参数或者混响衰减率参数。补偿器(505)响应于第一混响参数的修改,来生成用于第二混响参数的修改后的第二参数值。第二混响参数指示声学环境中的混响能量。渲染器(400)通过使用元数据渲染音频数据来生成音频输出信号,并且具体地,混响渲染器(407)响应于第一修改后的参数值和第二修改后的参数值,根据音频信号中的至少一个音频信号为至少一个音频输出信号生成至少一个混响信号分量。补偿可以提供改进的感知混响,同时允许灵活的适配。



1. 一种音频装置,包括:

接收机(501),所述接收机被布置为接收音频数据以及针对所述音频数据的元数据,所述音频数据包括针对表示环境中的音频源的多个音频信号的数据,并且所述元数据包括用于针对所述环境的混响参数的数据;

修改器(503),所述修改器被布置为通过修改第一混响参数的初始第一参数值来生成修改后的第一参数值,所述第一混响参数是来自由混响延迟参数以及混响衰减率参数组成的组中的参数;

补偿器(505),所述补偿器被布置为响应于所述第一混响参数的修改,通过修改针对第二混响参数的初始第二参数值来生成修改后的第二参数值,所述第二混响参数包括于所述元数据中并且指示声学环境中的混响能量;

渲染器(400),所述渲染器被布置为通过使用所述元数据渲染所述音频数据来生成音频输出信号,所述渲染器包括混响渲染器(407),所述混响渲染器被布置为响应于所述第一修改后的参数值以及所述第二修改后的参数值,并且根据所述音频信号中的至少一个音频信号为至少一个音频输出信号生成至少一个混响信号分量。

2. 根据权利要求1所述的装置,其中,所述补偿器(505)包括用于扩散混响的模型,所述模型取决于所述第一混响参数以及所述第二混响参数,并且所述补偿器被布置为响应于所述模型来确定所述修改后的第二参数值。

3. 根据前述任一权利要求所述的装置,其中,所述第一混响参数是混响衰减率。

4. 根据权利要求3所述的装置,其中,所述补偿器(505)被布置为修改所述第二参数值,以减少由于所述第一混响参数的修改而导致的针对所述混响衰减率的幅度参考的改变。

5. 根据权利要求4所述的装置,其中,所述补偿器(505)被布置为修改所述第二参数值,使得对于所述第一混响参数的修改,针对所述混响衰减率的幅度参考实质上不改变。

6. 根据前述任一权利要求所述的装置,其中,所述第一混响参数是指示针对所述环境中的混响的传播时间延迟的混响延迟参数。

7. 根据前述任一权利要求所述的装置,其中,所述第二混响参数指示在由所述第一混响参数指示的传播时间延迟之后的所述声学环境中的混响能量。

8. 根据权利要求6或7所述的装置,其中,所述补偿器(505)被布置为确定所述修改后的第二参数值以减少第一混响能量测量和第二混响能量测量之间的差,所述第一混响能量测量是在由所述修改后的第一参数值表示的修改后的延迟之后的混响能量,并且是使用所述修改后的延迟值以及所述修改后的第二参数值根据混响模型来确定的;并且所述第二混响能量测量是在所述修改后的延迟之后的混响能量,并且是使用初始延迟值以及所述初始第二参数值根据所述混响模型来确定的。

9. 根据权利要求8所述的装置,其中,所述补偿器(505)被布置为确定所述修改后的第二混响参数值,使得所述第一混响能量测量和所述第二混响能量测量实质上相同。

10. 根据权利要求6至9中任一项所述的装置,其中,所述补偿器(505)被布置为修改所述第二参数值,以针对超过由所述修改后的第一参数值指示的延迟的延迟,减少作为时间的函数的混响幅度的差。

11. 根据前述任一权利要求所述的装置,其中,所述第二参数表示相对于所述环境中的总发射声音的扩散混响声音的电平。

12. 根据前述权利要求1至10中任一项所述的装置, 其中, 所述第二混响参数表示在所述环境中针对声音传播的直接响应的能量等于所述环境中的混响能量的距离。

13. 根据前述权利要求1-10中任一项所述的装置, 其中, 所述第一混响参数是所述元数据的所述混响参数中的一个。

14. 根据前述任一权利要求所述的装置, 其中, 所述渲染器被布置为根据所述第二参数值确定所述至少一个混响信号分量的电平增益。

15. 一种操作音频装置的方法, 包括:

接收音频数据以及针对所述音频数据的元数据, 所述音频数据包括针对表示环境中的音频源的多个音频信号的数据, 并且所述元数据包括用于针对所述环境的混响参数的数据;

通过修改第一混响参数的初始第一参数值来修改第一参数值, 所述第一混响参数是来自自由混响延迟参数和混响衰减率参数组成的组中的参数;

响应于所述第一混响参数的修改, 通过修改针对第二混响参数的初始第二参数值来生成修改后的第二参数值, 所述第二混响参数包括于所述元数据中并且指示声学环境中的混响能量;

通过使用所述元数据渲染所述音频数据来生成音频输出信号, 所述渲染包括响应于所述第一修改后的参数值以及所述第二修改后的参数值, 并且根据所述音频信号中的至少一个音频信号为至少一个音频输出信号生成至少一个混响信号分量。

16. 一种计算机程序产品, 包括计算机程序代码单元, 当所述程序在计算机上运行时, 所述计算机程序代码单元适应于执行权利要求15的所有所述步骤。

一种音频装置以及其操作方法

技术领域

[0001] 本发明涉及一种用于生成音频输出信号的装置和方法,并且特别地但是不排除地,用于生成包括模拟环境的混响特性的扩散混响信号分量的音频输出信号,作为例如虚拟现实体验的部分。

背景技术

[0002] 近年来,随着不断开发和引入新的服务以及利用和消费视听内容的新方式,基于这样的视听内容的体验的种类和范围已经实质上增加。特别地,正在开发许多空间和交互的服务、应用和体验,以给用户更多参与和沉浸式的体验。

[0003] 这样的应用的示例为虚拟现实 (VR)、增强现实 (AR) 和混合现实 (MR) 应用,它们正迅速成为主流,其许多解决方案都是针对消费者市场的。多个标准化机构也正在制定多种标准。这样的标准化活动正在积极为VR/AR/MR系统的各个方面(包括例如流式传输、广播、渲染等)制定标准。

[0004] VR应用倾向于提供与用户处于不同世界/环境/场景中相对应的用户体验,而AR(包括混合现实MR)应用倾向于提供与用户处于当前环境中相对应的用户体验,但是添加了附加信息或者虚拟对象或者信息。因此,VR应用倾向于提供完全沉浸式的合成生成的世界/场景,而AR应用倾向于提供覆盖在用户物理所在的真实场景中的部分合成的世界/场景。然而,术语经常互换使用,并且具有高度重叠。在下文中,术语“虚拟现实/VR”将被用来表示虚拟现实和增强/混合现实两者。

[0005] 作为示例,日益流行的服务以这样的方式提供图像和音频:用户能够主动地和动态地与系统交互以改变渲染的参数,使得这将适应用户位置和定向的移动和改变。在许多应用中,非常吸引人的特征是改变观看者的有效观看位置和观看方向的能力,例如允许观看者在所呈现的场景中移动和“环顾四周”。

[0006] 这样的特征可以特别地允许向用户提供虚拟现实体验。这可以允许用户(相对)自由地在虚拟环境中移动,并且动态地改变他的位置和他正在看的地方。典型地,这样的虚拟现实应用基于场景的三维模型,该模型被动态评估以提供具体的请求视图。这种方法在例如用于计算机和控制台的游戏应用(例如,第一人称射击类别)中是公知的。

[0007] 特别是对于虚拟现实应用,还期望所呈现的图像是三维图像,典型地使用立体显示器呈现。事实上,为了优化观看者的沉浸感,用户典型地优选于将呈现的场景体验为三维场景。事实上,虚拟现实体验应该优选地允许用户选择他/她自己相对于虚拟世界的位置、视点和时刻。

[0008] 除了视觉渲染之外,大多数VR/AR应用进一步提供了对应的音频体验。在许多应用中,音频优选地提供空间音频体验,其中,音频源被感知到从对应于视觉场景中对对象的位置的位置到达。因此,音频和视频场景优选地被感知为一致的,并且两者提供了完整的空间体验。

[0009] 例如,许多沉浸式体验是通过虚拟音频场景来提供的,该虚拟音频场景是使用双

声道音频渲染技术的耳机再现所生成的。在许多场景中,这样的耳机再现可以基于头部跟踪,使得可以响应于用户的头部移动来进行渲染,这高度增加了沉浸感。

[0010] 对于许多应用来说,一个重要的特征是如何生成和/或分发能够提供音频环境的自然和真实感知的音频。例如,当针对虚拟现实应用生成音频时,重要的是不仅生成所期望的音频源,而且还对音频源进行修改以提供音频环境的真实感知,包括衰减、反射、染色等。

[0011] 对于房间声学,或者更一般地环境声学,环境的墙壁、地板、天花板、对象等的声波反射导致声源信号的延迟和衰减(典型地依赖于频率)版本经由不同路径到达收听者(即,VR/AR系统的用户)。该组合效应可以通过脉冲响应来建模,该脉冲响应在下文中可以被称为房间脉冲响应(RIR)(尽管该术语暗示了以房间形式的声学环境的具体用途,但是它倾向于相对于声学环境更一般地使用,无论这是否与房间相对应)。

[0012] 如图1所示,房间脉冲响应典型地由直接声音组成,直接声音取决于声源到收听者的距离,然后是表征房间的声学属性的混响部分。房间的大小和形状、声源和收听者在房间中的位置以及房间表面的反射属性都对该混响部分的特性起着作用。

[0013] 混响部分可以分成两个时间区域,通常是重叠的。第一区域包括所谓的早期反射,它表示在到达收听者之前声源在房间内的墙壁或者障碍物上的隔离反射。随着时间滞后/(传播)延迟的增加,在固定时间间隔中存在的反射的数量增加,并且路径可以包括二次或者更高阶反射(例如,反射可以来自几面墙壁或者墙壁和天花板两者等)。

[0014] 混响部分中的第二区域是其中这些反射的密度增加到人脑无法再隔离的程度的部分。这个区域典型地被称为扩散混响、后期混响或者混响尾。

[0015] 混响部分包含提示,该提示给听觉系统关于源的距离、以及房间的大小和声学属性的信息。混响部分的能量相对于消声部分的能量在很大程度上确定了声源的感知距离。最早反射的电平和延迟可以提供关于声源离墙有多近的提示,并且通过人体测量的滤波可以加强对特点的墙、地板或者天花板的评估。

[0016] (早期)反射的密度有助于感知房间的大小。反射的能量电平下降60dB所需的时间(由混响时间 T_{60} 指示)是针对反射在房间中消散速度的常用测量。混响时间提供了房间声学属性的信息;例如,特别是墙壁是否反射性很强(例如,浴室)或者是否有很大的声音吸收(例如,有家具、地毯和窗帘的卧室)。

[0017] 此外,当RIR是双声道房间脉冲响应(BRIR)的一部分时,RIR可以依赖于用户的人体测量属性,因为RIR被头部、耳朵和肩膀滤波;即,头部相关脉冲响应(HRIR)。

[0018] 由于后期混响中的反射不能被收听者区分和隔离,因此它们通常用例如使用反馈延迟网络的参数混响器来模拟和参数化表示,如在公知的Jot混响器中。

[0019] 对于早期反射,与入射方向和距离相关的延迟是人们提取关于房间和声源的相对位置的信息的重要提示。因此,对早期反射的模拟必须比后期混响的更加明确。因此,在有效的声学渲染算法中,早期反射的模拟与后期混响的模拟是不同的。针对早期反射的公知方法是镜像房间边界中的每个边界的声源,以生成表示反射的虚拟声源。

[0020] 对于早期反射,用户和/或声源相对于房间的边界(墙壁、天花板、地板)的位置是相关的,而对于后期混响,房间的声学响应是扩散的,并且因此在整个房间中趋于更加均匀。这使得后期混响的模拟通常比早期反射在计算上更高效。

[0021] 由房间定义的后期混响的两个主要属性是表示高于给定电平的时间的脉冲响应

的幅度和斜率的参数。在自然房间中,这两个参数都倾向于具有强烈的频率依赖性。

[0022] 传统上用于指示对应于扩散混响的脉冲响应的幅度和斜率的参数的示例包括已知的 T_{60} 值和混响电平/能量。最近提出了幅度电平的其他指示(例如,指示扩散混响能量和总发射源能量之间的比率的具体参数)。

[0023] 这样的已知的方法倾向于提供混响的高效描述,这允许在渲染侧准确地再现环境的混响特性。然而,尽管这些方法在寻求准确地渲染环境中的混响时倾向于是有利的,但是在一些场景中,它们倾向于是次优的,并且特别是倾向于相对不灵活的。典型地,适配和修改处理和/或所产生的混响分量倾向于是困难的,并且特别是在不降低(感知到的)音频质量和/或要求比优选的计算资源更多的情况下。

[0024] 因此,用于渲染针对环境的混响音频的改进方法将是有利的。特别地,允许改进的操作、增加的灵活性、减少的复杂度、促进的实现方式、改进的音频体验、改进的音频质量、减少的计算负担、改进的对变化位置的适用性、针对虚拟/混合/增强现实应用的改进的性能、改进的用于扩散混响的感知提示、增加的和/或促进的适应性、增加的处理灵活性、增加的渲染侧定制和/或改进的性能和/或操作的方法将是有利的。

发明内容

[0025] 因此,本发明寻求单独地或者以任何组合优选地减轻、减缓或者消除上述缺点中的一个或多个。

[0026] 根据本发明的一方面,提供了一种音频装置,包括:接收机,其被布置为接收音频数据和针对音频数据的元数据,音频数据包括针对表示环境中的音频源的多个音频信号的数据,并且元数据包括用于针对环境的混响参数的数据;修改器,其被布置为通过修改第一混响参数的初始第一参数值来生成修改后的第一参数值,第一混响参数是来自由混响延迟参数和混响衰减率参数组成的组中的参数;补偿器,其被布置为响应于第一混响参数的修改,通过修改针对第二混响参数的初始第二参数值来生成修改后的第二参数值,第二混响参数包括于元数据中并且指示声学环境中的混响能量;渲染器,其被布置为通过使用元数据渲染音频数据来生成音频输出信号,渲染器包括混响渲染器,其被布置为响应于第一修改后的参数值和第二修改后的参数值,并且根据音频信号中的至少一个音频信号为至少一个音频输出信号生成至少一个混响信号分量。

[0027] 本发明可以提供对包括混响分量的音频的改进和/或促进的渲染。本发明可以在许多实施例和场景中生成更自然的发声(扩散)混响信号,从而提供声学环境的改进感知。音频输出信号和混响信号分量的渲染通常可以以减少的复杂度和减少的计算资源需求来生成。

[0028] 该方法可以提供对处理和/或渲染的音频的改进的、增加的和/或促进的灵活性和/或适配。在许多应用和实施例中,这样的适配可以通过修改参数值来执行适配而被实质上促进。特别地,在许多情况下,算法、过程和/或渲染操作可以不改变,而是可以简单地通过修改参数值来实现所要求的适配。还可以通过基于混响延迟参数和/或混响衰减率参数如何改变而修改第二混响参数(其指示声学环境中的混响能量)来进一步促进对混响输出和/或处理的适配或者修改。

[0029] 修改混响延迟参数和/或混响衰减率参数可以提供特别高效和有利的对混响的操

作和适配,并且第二混响参数可以针对该修改自动补偿。这可以自动减少或者移除对混响延迟参数和/或混响衰减率参数的修改的意外影响。例如,它可以减少适配的感知影响和/或可以例如提供更一致和/或和谐的音频信号输出。

[0030] 该方法允许通过相对较少的参数高效地表示声学环境中的扩散混响声音。

[0031] 在许多实施例中,该方法可以允许独立于源和/或收听者位置来生成扩散混响信号。这可以允许对于其中位置改变的动态应用(例如,对于许多虚拟现实和增强现实应用)高效地生成扩散混响信号。

[0032] 音频装置可以在单个设备或者单个功能单元中被实现,或者可以跨不同设备或者功能分布。例如,音频装置可以被实现为解码器功能单元的一部分,或者可以分布为在解码器侧执行的一些功能元件以及在编码器侧执行的其他元件。

[0033] 补偿器可以被布置为响应于修改后的第一参数值与初始第一参数值之间的差而生成修改后的第二参数值。

[0034] 在许多实施例中,渲染器包括用于渲染针对音频信号的直接路径分量和/或早期反射分量的另一渲染器,并且渲染器可以被布置为响应于直接路径分量、早期反射分量和至少一个混响信号的组合而生成输出信号。

[0035] 混响渲染器可以是扩散混响渲染器。混响渲染器可以是参数混响渲染器,例如反馈延迟网络(FDN)混响器,并且具体是Jot混响器。

[0036] 元数据可以用于音频信号/音频源和/或环境。

[0037] 根据本发明的可选特征,补偿器包括用于扩散混响的模型,该模型依赖于第一混响参数和第二混响参数,并且补偿器被布置为响应于该模型来确定修改后的第二参数值。

[0038] 该方法可以提供用于生成反应频率依赖性的扩散混响信号的特别高效的操作系统。

[0039] 该模型可以是算术函数/等式/或者函数/等式的集合。

[0040] 根据本发明的可选特征,第一混响参数是混响衰减率。

[0041] 本发明可以提供改进的性能和/或操作。它可以促进和/或改进适配和灵活性,并且可以允许增加对所渲染混响的控制。混响衰减率参数可以提供特别高效的适配,并且可以特别允许对环境中混响的感知属性的实际适配。

[0042] 混响衰减率参数可以例如是 T_{60} (或者更一般地是 T_{xx} ,其中,xx可以是任何合适的整数)参数。

[0043] 根据本发明的可选特征,补偿器被布置为修改第二参数值,以减少由于第一混响参数的修改而导致的针对混响衰减率的幅度参考的改变。

[0044] 这可以允许特别有利的适配,并且可以允许非常高效但典型地是低复杂度的补偿。

[0045] 幅度参考可以是混响衰减率和第二参数的函数。

[0046] 根据本发明的可选特征,补偿器被布置为修改第二参数值,使得对于第一混响参数的修改,针对混响衰减率的幅度参考实质上不变。

[0047] 这可以允许特别有利的操作和/或性能。

[0048] 根据本发明的可选特征,第一混响参数是指示针对环境中的混响的传播时间延迟的混响延迟参数。

[0049] 本发明可以提供改进的性能和/或操作。它可以促进和/或改进适配和灵活性,并

且可以允许增加对所渲染的混响的控制。混响延迟参数可以提供特别高效的适配,并且可以特别允许对环境中混响的感知属性的实际适配。

[0050] 混响延迟参数具体地可以是预延迟参数。

[0051] 传播时间延迟可以指示在房间中波传播中与参考事件的时间偏移。典型地,参考事件是在音频源处的声音能量的发射,但是在一些情况下/实施例中可以是直接路径响应。更具体地,它可以指示房间脉冲响应中的滞后。在许多实施例中,它可以指示偏移时间,对于该偏移时间,计算指示声学环境中的混响能量的第二混响参数。该值可以通过分析由混响参数表示的房间脉冲响应来进行选择。例如,传播时间延迟可以指示在源处的发射和信号(即,早期反射后的声音)的扩散后期混响部分的开始之间的延迟,并且可以以秒为单位指定,或者传播时间延迟可以指示从其扩散的房间响应中的滞后,即,来自所有方向的入射电平相同和房间中所有位置的电平相似。

[0052] 根据本发明的可选特征,第二混响参数指示在由第一混响参数指示的传播时间延迟之后的声学环境中的混响能量。

[0053] 这可以允许特别有利的操作和/或性能。

[0054] 根据本发明的可选特征,补偿器被布置为确定修改后的第二参数值以减少第一混响能量测量和第二混响能量测量之间的差,第一混响能量测量是在由修改后的第一参数值表示的修改后的延迟之后的混响能量,并且是使用修改后的延迟值和修改后的第二参数值根据混响模型来确定的;以及第二混响能量测量是在修改后的延迟之后的混响能量,并且是使用初始延迟值和初始第二参数值根据混响模型来确定的。

[0055] 这可以允许特别有利的操作和/或性能。在许多情况下,它可以允许减少混响延迟参数的修改对所渲染的混响的感知效果。

[0056] 根据本发明的可选特征,补偿器被布置为确定修改后的第二混响参数值,使得第一混响能量测量和第二混响能量测量实质上相同。

[0057] 这可以允许特别有利的操作和/或性能。在许多场景下,它可以允许减少的或者甚至实质上消除混响延迟参数的修改对所渲染的混响的感知效果。

[0058] 根据本发明的可选特征,补偿器被布置为修改第二参数值,以针对超过由修改后的第一参数值指示的延迟的延迟,减少作为时间函数的混响幅度的差。

[0059] 这可以允许特别有利的操作和/或性能。在许多场景中,它可以允许减少混响延迟参数的修改对所渲染的混响的感知效果。

[0060] 在许多实施例中,混响渲染器被布置为生成至少一个混响信号分量,以仅包括与超过由第一修改后的混响参数指示的传播延迟时间的传播延迟相对应的贡献。

[0061] 在一些实施例中,混响渲染器被布置为生成至少一个混响信号分量,以仅包括在超过由第一修改后的混响参数指示的传播延迟时间的时间处对应于房间脉冲响应的一部分的贡献。

[0062] 根据本发明的可选特征,第二参数表示相对于环境中的总发射声音的扩散混响声音的电平。

[0063] 这可以提供特别有利的操作和/或性能。

[0064] 在许多实施例中,第二参数表示相对于环境中的总发射能量的扩散混响声音的能量。

[0065] 扩散混响信号与总信号的关系/比率还可以被称为扩散混响信号电平与总信号电平比率或者扩散混响电平与总电平比率或者发射源能量与扩散混响能量比率(或者其变型/排列)。

[0066] 根据本发明的可选特征,第二混响参数表示在环境中声音传播的直接响应的能量等于环境中混响的能量的距离。

[0067] 这可以提供特别有利的操作和/或性能。

[0068] 第二混响参数可以是临界距离参数。

[0069] 在一些实施例中,第二参数表示用于针对环境的房间脉冲响应在给定的确定时间/滞后处的幅度。

[0070] 根据本发明的可选特征,第一混响参数是元数据的混响参数中的一个。

[0071] 根据本发明的可选特征,渲染器被布置为根据第二参数值确定至少一个混响信号分量的电平增益。

[0072] 这可以在许多场景中提供混响信号分量的高效并且有利的生成。电平增益例如可以是确定/设置/控制混响信号分量的电平的增益/比例因子。

[0073] 这可以提供特别有利的操作和/或性能。

[0074] 根据本发明的一方面,提供了一种操作音频装置的方法,包括:接收音频数据和针对音频数据的元数据,音频数据包括针对表示环境中的音频源的多个音频信号的数据,并且元数据包括用于针对环境的混响参数的数据;通过修改第一混响参数的初始第一参数值来修改第一参数值,第一混响参数是来自自由混响延迟参数和混响衰减率参数组成的组中的参数;响应于第一混响参数的修改,通过修改针对第二混响参数的初始第二参数值来生成修改后的第二参数值,第二混响参数包括于元数据中并且指示声学环境中的混响能量;通过使用元数据渲染音频数据来生成音频输出信号,渲染包括响应于第一修改后的参数值和第二修改后的参数值,并且根据音频信号中的至少一个音频信号为至少一个音频输出信号生成至少一个混响信号分量。

[0075] 本发明的这些和其他方面、特征和优点将从下面描述的实施例中变得明显,并且将参照实施例进行说明。

附图说明

[0076] 将参考附图仅通过示例的方式来描述本发明的实施例,在附图中,

[0077] 图1示出了房间脉冲响应的示例;

[0078] 图2示出了房间脉冲响应的示例;

[0079] 图3示出了虚拟现实系统的元件的示例;

[0080] 图4示出了根据本发明的一些实施例的用于生成音频输出的渲染器的示例;

[0081] 图5示出了根据本发明的一些实施例的用于生成音频输出的音频装置的示例;

[0082] 图6示出了房间脉冲响应的示例;

[0083] 图7示出了用于房间脉冲响应的幅度和累积能量的示例;

[0084] 图8示出了房间脉冲响应的混响部分的示例;

[0085] 图9示出了房间脉冲响应的混响部分的示例;

[0086] 图10示出了房间脉冲响应的混响部分的示例;

- [0087] 图11示出了房间脉冲响应的混响部分的示例;
- [0088] 图12示出了房间脉冲响应的混响部分的示例;
- [0089] 图13示出了参数混响器的示例;以及
- [0090] 图14示出了混响器的示例。

具体实施方式

[0091] 下面的描述将集中于用于虚拟现实应用的音频处理和渲染,但是可以理解,所描述的原理和概念可以用于许多其他应用和实施例中。

[0092] 允许用户在虚拟世界中移动的虚拟体验正变得越来越流行,并且正在开发服务以满足这样的需求。

[0093] 在一些系统中,VR应用可以通过例如独立设备在本地提供给观看者,该独立设备不使用或者甚至无法访问任何远程VR数据或者处理。例如,诸如游戏控制台之类的设备可以包括用于存储场景数据的存储装置、用于接收/生成观看者姿势的输入端、以及用于从场景数据生成对应图像的处理器。

[0094] 在其他系统中,VR应用可以在远离观看者的地方实现和执行。例如,用户本地的设备可以检测/接收移动/姿势数据,该移动/姿势数据被发送到远程设备,该远程设备处理数据以生成观看者姿势。然后,远程设备可以基于描述场景的场景数据针对用户姿势生成合适的视图图像和对应的音频信号。然后,视图图像和对应的音频信号被发送到观看者本地的设备,在该设备处呈现视图图像和对应的音频信号。例如,远程设备可以直接生成由本地设备直接呈现的视频流(典型地是立体声/3D视频流)和对应的音频流。因此,在这样的示例中,除了发送移动数据和呈现接收到的视频数据之外,本地设备可以不执行任何VR处理。

[0095] 在许多系统中,功能可以跨本地设备和远程设备分布。例如,本地设备可以处理接收到的输入和传感器数据,以生成连续发送到远程VR设备的用户姿势。然后,远程VR设备可以生成对应的视图图像和对应的音频信号,并且将这些视图图像和音频信号发送到本地设备以进行呈现。在其他系统中,远程VR设备不直接生成视图图像和对应的音频信号,但是可以选择相关的场景数据并且将其发送给本地设备,然后本地设备可以生成所呈现的视图图像和对应的音频信号。例如,远程VR设备可以识别最近的捕获点并且提取对应的场景数据(例如,对象源及其位置元数据的集合)并且将其发送给本地设备。然后,本地设备可以处理接收到的场景数据以生成用于具体的当前用户姿势的图像和音频信号。典型地,用户姿势将与头部姿势相对应,并且对用户姿势的引用典型地可以等效地被认为对应于对头部姿势的引用。

[0096] 在许多应用中,特别是对于广播服务,源可以以独立于用户姿势的场景的图像(包括视频)和音频表示的形式发送或者流式传输场景数据。例如,对应于特定虚拟房间范围内的音频源的信号和元数据可以被发送或者流式传输到多个客户端。然后,各个客户端可以本地合成对应于当前用户姿势的音频信号。类似地,源可以发送音频环境的一般描述,包括描述环境中的音频源和环境的声学特性。然后可以例如使用双声道渲染和处理在本地生成音频表示并且将其呈现给用户。

[0097] 图3示出了其中远程VR客户端设备301例如经由网络305(例如,互联网)与VR服务器303联系的VR系统的这样的示例。服务器303可以被布置为同时支持潜在大量的客户端设

备301。

[0098] 例如,VR服务器303可以通过发送包括图像数据形式的图像表示的图像信号来支持广播体验,该图像表示可以由客户端设备用于本地合成与适当的用户姿势(姿势指代的是位置和/或定向)对应的视图图像。类似地,VR服务器303可以发送场景的音频表示,允许针对用户姿势本地合成音频。具体地,当用户在虚拟环境中移动时,更新被合成并且呈现给用户的图像和音频以反映用户在(虚拟)环境中的当前(虚拟)位置和定向。

[0099] 在许多应用中,例如图3的应用中,因此可以期望对场景建模并且生成高效的图像和音频表示,该高效的图像和音频表示可以被高效地包括于数据信号中,然后数据信号可以被发送或者流式传输到各种设备,该各种设备可以针对不同于捕获姿势的姿势本地合成视图和音频。

[0100] 在一些实施例中,表示场景的模型可以例如存储在本地,并且可以在本地使用以合成适当的图像和音频。例如,房间的音频模型可以包括在房间中可以听到的音频源的属性的指示以及房间的声学属性。然后可以使用该模型数据来合成用于具体位置的适当音频。

[0101] 如何表示音频场景以及如何使用这种表示来生成音频是个关键问题。旨在向收听者提供自然和逼真效果的音频渲染典型地包括声学环境的渲染。对于许多环境,这包括表示和渲染环境中(例如,房间中)存在的扩散混响。这样的扩散混响的渲染和表示已经被发现对环境的感知有显著的影响,例如对音频是否被感知为表示自然和逼真的环境有显著的影响。在下面,将描述用于表示音频场景和渲染音频(以及特别是扩散混响音频)的有利方法。

[0102] 将参考包括如图4所示的渲染器400的音频装置来描述该方法。音频装置被布置为生成表示声学环境中的音频的音频输出信号。具体地,音频装置可以生成表示由在具有多个音频源并且具有给定声学属性的虚拟环境中移动的用户感知的音频的音频。每个音频源由表示来自音频源的声音的音频信号以及可以描述音频源特性的元数据(例如,针对音频信号提供电平指示)来表示。另外,提供元数据来表征声学环境。

[0103] 渲染器400包括用于每个音频源的路径渲染器401。每个路径渲染器401被布置为生成表示从音频源到收听者的直接路径的直接路径信号分量。直接路径信号分量是基于收听者和音频源的位置生成的,并且可以通过取决于距离和例如音频源在到用户的具体方向(例如,对于非全向源)上的相对增益,缩放潜在地依赖于频率的音频源的音频信号来具体生成直接信号分量。

[0104] 在许多实施例中,渲染器401还可以基于位于源位置和用户位置之间的遮挡物或者衍射(虚拟)元件来生成直接路径信号。

[0105] 在许多实施例中,路径渲染器401还可以针对各个路径生成另外的信号分量,其中,这些路径包括一个或多个反射。例如,这可以通过评估墙壁、天花板等的反射来完成,如技术人员所知。可以将直接路径和反射路径分量组合成针对每个路径渲染器的单个输出信号,并且因此可以针对每个音频源生成表示直接路径和早期/分立反射的单个信号。

[0106] 在一些实施例中,针对每个音频源的输出音频信号可以是双声道信号,并且因此每个输出信号可以包括左耳和右耳(子)信号两者。

[0107] 来自路径渲染器401的输出信号被提供给组合器403,该组合器403组合来自不同

路径渲染器401的信号以生成单个组合信号。在许多实施例中,可以生成双声道输出信号,并且组合器可以执行对来自路径渲染器401的各个信号的组合(例如,加权组合(即,来自路径渲染器401的所有右耳信号可以相加在一起以生成组合的右耳信号,并且来自路径渲染器401的所有左耳信号可以相加在一起以生成组合的左耳信号))。

[0108] 路径渲染器和组合器可以以任何适当的方式实现,典型地包括作为用于在适当的计算资源(例如,微控制器、微处理器、数字信号处理器、或者包括支持电路系统(诸如存储器之类的)的中央处理单元)上进行处理的可执行代码。可以理解,多个路径渲染器可以实现为并行功能单元,例如一组专用处理单元,或者可以实现为针对每个音频源的重复操作。典型地,针对每个音频源/信号执行相同的算法/代码。

[0109] 除了各个路径音频分量之外,渲染器400还被布置为生成表示环境中的扩散混响的信号分量。扩散混响信号在具体示例中是如下生成的:通过将源信号组合成下混频信号,并且然后将混响算法应用于下混频信号以生成扩散混响信号。

[0110] 图4的音频装置包括下混频器405,其接收多个声源(典型地是混响器正在针对其模拟扩散混响的声学环境中的所有源)的音频信号,并且将它们组合成下混频。因此,下混频反映了环境中生成的所有声音。针对个体音频信号的系数/加权可以例如被设置为反映对应声源的电平。

[0111] 下混频被馈送到混响渲染器/混响器407,该混响渲染器/混响器407被布置为基于下混频生成扩散混响信号。具体地,混响器407可以是参数混响器,例如Jot混响器。混响器407耦合到向其馈送扩散混响信号的组合器403。然后,组合器403继续将扩散混响信号与表示各个路径的路径信号组合,以生成组合音频信号,该组合音频信号表示由收听者感知的环境中的组合声音。

[0112] 在示例中,渲染器是音频装置的部分,其被布置为接收音频数据以及针对环境的元数据,并且基于所接收的数据渲染表示环境的至少一部分的音频。图5示出了这样的装置的示例,并且将参考图4和图5的示例描述用于基于接收到的音频数据和元数据生成音频输出信号(具体地,混响信号分量)的方法。具体地,图5的音频装置可以与图3的客户端设备301或者是其一部分相对应。

[0113] 图5的音频装置包括接收机501,其被布置为从一个或多个源接收数据。该源可以是用于提供数据的任何合适的源,并且可以是内部源或者外部源。接收机501可以包括用于接收/取回数据的所要求的功能,例如无线电功能、网络接口功能等。

[0114] 接收机501可以从任何合适的源并且以任何合适的形式接收数据,包括例如作为音频信号的部分。数据可以从内部源或者外部源接收。接收机401可以例如被布置为经由到内部源的网络连接、无线电连接或者任何其他合适的连接来接收房间数据。在许多实施例中,接收机可以从诸如本地存储器之类的本地源接收数据。在许多实施例中,接收机501可以例如被布置为从诸如本地RAM或者ROM存储器之类的本地存储器中取回房间数据。在具体示例中,接收机501可以包括用于接合到网络305以便从VR服务器303接收数据的网络功能。

[0115] 接收机501可以以任何合适的方式实现,包括例如使用分立的或者专用的电子设备。接收机501例如可以被实现为诸如专用集成电路(ASIC)之类的集成电路。在一些实施例中,电路可以被实现为编程处理单元,例如,在适当的处理器(例如,中央处理单元、数字信号处理单元或者微控制器等)上运行的固件或者软件。可以理解,在这样的实施例中,处理

单元可以包括板上或者外部存储器、时钟驱动电路系统、接口电路系统、用户接口电路系统等。这样的电路系统还可以被实现为处理单元、集成电路和/或分立电子电路系统的部分。

[0116] 接收到的数据包括用于表示环境中的音频源的多个音频信号的音频数据。具体地音频数据包括多个音频信号,其中,音频信号中的每个表示一个音频源(并且因此音频信号描述来自音频源的声音)。

[0117] 此外,接收机501接收针对音频源和/或环境的元数据。

[0118] 针对个体音频信号/源的元数据可以包括针对音频源的(相对)信号电平指示,其中,信号电平指示可以指示由音频信号表示的声源的电平/能量/幅度。针对源的元数据还可以包括指示来自声源的声音辐射的方向性的方向性数据。针对音频信号的方向性数据可以例如描述增益模式,并且可以具体地描述距音频源的位置在不同方向上的针对音频源的相对增益/能量密度。元数据还可以包括其他数据,例如音频源的标称、开始或者当前(或者可能静态)位置的指示。

[0119] 接收机501进一步接收指示声学环境的元数据。具体地,接收机501接收包括描述环境的混响属性的混响参数的元数据。特别地,元数据可以包括混响衰减率参数的指示,并且潜在地还可以包括混响延迟参数的指示。元数据还可以包括指示混响的能量/电平的混响能量参数。

[0120] 例如房间脉冲响应(RIR)的扩散混响属性可以由经由参数数据通信给渲染器的参数来表示。

[0121] 至少部分地描述环境的混响的参数是混响延迟参数。混响延迟参数可以指示来自音频源的混响的延迟。具体地,混响延迟参数可以具体地指示RIR的混响部分的开始时间(在RIR中)。

[0122] 在许多实施例中,元数据可以包括扩散混响信号应该何时开始的指示,即,它可以指示与扩散混响信号相关联的时间延迟。时间延迟指示可以具体为预延迟的形式。

[0123] 预延迟可以表示RIR中的延迟/滞后,并且可以被定义为早期反射和扩散、后期混响之间的阈值。由于该阈值典型地作为从(或多或少)分立反射到完全干扰的高阶反射的混合的平滑转换的部分发生,因此可以使用适当的评估/决策过程来选择适当的阈值。该确定可以基于RIR的分析自动进行,或者基于房间维度和/或材料属性计算。

[0124] 替代地,可以选择固定阈值,例如,进入RIR的80ms。预延迟可以用秒、毫秒或者样本指示。在下面的描述中,假定预延迟被选择为在混响实际上扩散之后的点处。然而,如果不是这样的话,所描述的方法仍然可以充分工作。

[0125] 因此,预延迟指示从源发射的开始到扩散混响响应的开始。例如,如图6中示出的示例,如果源在 t_0 (例如 $t_0=0$)开始发射,则直接声音在 $t_1>t_0$ 到达用户,第一反射在 $t_2>t_1$ 到达用户,并且早期反射和扩散混响之间的定义的阈值在 $t_3>t_2$ 到达用户。然后,预延迟为 t_3-t_0 。预延迟可以被认为反映了针对扩散混响的开始时的传播延迟。

[0126] 在许多实施例中,混响延迟参数(例如,以预延迟的形式)可以包括于元数据中。然而,在其他实施例中,它可以是预先确定的或者固定的参数。例如,比特流可以根据定义标准预延迟的适当音频标准或者规范,参照该标准或者规范可以给出其他混响参数(例如,衰减率或者混响能量参数)。

[0127] 至少部分地描述环境的混响的另一参数是混响衰减率参数。混响衰减率参数可以

指示针对环境的混响的电平降低率,并且具体地可以指示RIR的混响部分的电平降低率。具体地,混响衰减率参数可以指示RIR的混响部分的斜率。

[0128] 混响衰减率参数可以指示作为时间/滞后/延迟的函数的混响的电平变化,并且可以具体地指示作为延迟/时间的函数的混响(并且具体地是RIR的混响部分)的衰减/减少电平。在一些实施例中,混响衰减率参数可以是指示每时间单位(例如,每秒)混响响应减少的平均分贝(dB)的参数,或者是针对描述线性幅度或者能量域(例如, $2^{-\gamma t}$)中的电平衰减的指数方程的指数系数。

[0129] 混响衰减率参数可以在不同实施例之间变化。在许多实施例中,其可以例如是本领域技术人员已知的 T_{60} 、 T_{30} 或者 T_{20} 参数。这些参数指示混响能量衰减60dB(分别为30、20dB)所需的时间。例如,由与能量衰减曲线(EDC)下降60dB相对应的时间表示,由以下积分方程给出:

$$[0130] \quad EDC(t) = \int_t^{t_{\max}} RIR(x) dx$$

[0131] 其中, $t_{\max}$ 可以是 $t_{\max} = \infty$ 或者房间脉冲响应($RIR(t)$)消失在RIR的噪声底限中的点。

[0132] 至少部分地描述环境的混响的另一参数是指示声学环境中的混响能量的混响参数,并且其具体地可以指示RIR的混响部分的能量。这样的参数还可以被称为混响能量参数。混响能量参数可以例如被给出为相对于总源能量的混响能量、临界距离、相对于总源能量的混响幅度等。

[0133] 在许多实施例中,环境的混响(并且具体地是RIR的(扩散)混响部分)可以通过混响延迟参数、混响衰减率参数和混响能量参数的组合来表征。这样的参数集合可以描述混响何时开始、混响的电平的时间进程以及混响的总电平。可以接收一个、多个或者所有这些参数作为元数据的部分。

[0134] 可以用接收到的混响参数控制渲染的音频的混响部分来渲染接收到的音频数据,从而产生以与环境混响分量对应的混响分量生成的输出音频信号。然而,图5的音频装置进一步包括允许在本地适应和定制混响的功能。在图5的音频装置中,这是通过包括允许在用于控制渲染器400的混响渲染之前修改混响延迟参数和/或混响衰减率参数的功能来实现的。

[0135] 在图5的音频装置中,接收机501耦合到渲染器400,并且接收到的音频数据被直接馈送到渲染器400。然而,元数据不直接馈送到渲染器400,而是首先馈送到修改器503,该修改器503被布置为修改作为混响延迟参数或者混响衰减率参数的第一混响参数(在一些情况下,这两种参数都可以被修改)。

[0136] 因此,第一混响参数最初可以具有给定参数值,并且该参数值可以由修改器503修改为(不同的)修改后的参数值。例如,对于混响延迟参数,可以将初始延迟值修改为修改后的延迟值,该修改后的延迟值典型地可以是更小或者更大延迟(尽管在一些实施例中,修改器503可以是非对称的,并且只能增加延迟或者只能减少延迟)。

[0137] 替代地或者另外地,对于混响衰减率参数,可以将初始衰减率值修改为修改后的衰减率值,该修改后的衰减率值典型地可以是更小或者更大衰减率/梯度(尽管在一些实施

例中,修改器503可以是非对称的,并且可以仅能够增加衰减率或者仅能够降低衰减率)。

[0138] 参数值的修改可以是完全自动的,并且由装置本身取决于例如当前操作条件来确定。例如,取决于可用的计算资源,通过修改器503改变混响延迟参数,可以动态地改变分别由路径渲染器401和混响渲染器407处理的RIR的量。在其他实施例和应用中,修改可以响应于用户输入,并且事实上用户可以直接控制混响参数的修改。例如,如果用户期望较小的混响体验,则用户输入可以允许混响衰减率参数被修改为与较高衰减率相对应的参数值,并且因此混响会更快地消失。可以理解,用于修改的许多其他原因、方法和目的是可能的,并且所描述的方法不依赖于用于修改混响参数的具体背景或者方法。

[0139] 发明人已经认识到,虽然通过修改描述RIR的混响部分的第二混响参数来修改渲染(并且具体地是适配和定制混响渲染)的这样的方法可以是高效和有利的,但是它并非在所有场景中都是最佳的,并且在许多场景中可以导致被感知为不理想的音频渲染。例如,在许多场景中,它可能会引入假象、质量退化、感知失真、和/或RIR不同部分之间的不平衡。

[0140] 本发明人还已经认识到,通过引入修改混响参数的补偿,可以减轻或者潜在地甚至实质上移除许多缺点,该混响参数指示环境中混响的能量(混响能量参数),并且具体地,指示RIR的混响部分的能量/电平。补偿基于对混响延迟和/或衰减率参数的修改,并且具体地基于针对第一混响参数的修改后的参数值与第一参数的原始值之间的差。特别地,补偿来自接收到的元数据的混响能量参数可以导致与修改后的混响参数的改进的一致性,并且可以允许例如感知到更自然的发声混响和总体音频体验。

[0141] 因此,图5的装置包括补偿器505,该补偿器505被布置为响应于第一混响参数的修改,通过修改针对第二混响参数的混响值来生成修改后的第二混响参数值,其中,第二混响参数被提供为元数据的一部分,并且其中,第二混响参数是指示声学环境中混响的能量的混响能量参数。

[0142] 例如,补偿器505可以被布置为适配混响能量参数以反映对于修改后的混响延迟参数,如果RIR或多或少被渲染为扩散混响而不是路径反射,则能量可以改变。作为另一示例,对于混响衰减率参数的改变,混响能量参数可以被改变以对不同衰减率的能量进行归一化。

[0143] 在元数据中,不同应用和比特流中的不同参数可以用于指示扩散混响的能量。典型地,RIR的扩散部分的能量倾向于由单个参数指示。然而,在一些情况下,可以使用多个参数,既可以作为替代参数也可以作为组合。能量指示可以是频率相关的。

[0144] 因此,在不同的实施例中,由补偿器修改的具体混响能量参数也可以不同。在下文中,将描述一些特别有利的混响能量参数:

[0145] 混响电平/能量典型地与直接声音有其主要的心理声学相关性。两者之间的电平差是声源和用户(或者RIR测量点)之间距离的指示。较大的距离将导致更多的直接声音衰减,而后期混响的电平保持相同(在整个房间中是相同的)。类似地,对于方向性依赖于用户相对于源的位置的源,当用户围绕源移动时,方向性影响直接响应,但是不影响混响的电平。

[0146] 因此,通常可以有利地不指示相对于直接声音的混响电平,而是可以使用与房间内的源和用户位置无关的更一般的属性。

[0147] 在一些实施例中,混响能量参数可以是指示相对于环境中的总发射声音的扩散混

响声音的电平参数。混响能量参数可以指示扩散混响信号与总信号的比率,即,扩散与源的比率DSR,其可以用来将用户接收到的扩散混响能量的量或者源的电平表示为该源的总发射能量的比率。可以以如下方式来表示,即,对于要渲染的信号和对应的元数据(例如,预增益)的电平校准适当地调节扩散混响能量。

[0148] 以这种方式表示它可以确保值独立于环境中收听者和源的绝对位置和定向,独立于用户相对于源的相对位置和定向,并且反之亦然,独立于用于渲染混响的具体算法,并且可以确保与系统中使用的信号电平存在有意义的联系。

[0149] 如将在后面描述的,对于这样的混响能量参数,所描述的示例性渲染可以计算下混频系数,该下混频系数考虑方向性模式以在源信号之间施加正确的相对电平,以及DSR以在混响器407的输出上实现正确的电平。

[0150] DSR可以表示发射源能量与扩散混响属性之间的比率,例如具体地,扩散混响信号的能量或者(初始)电平。

[0151] 描述将主要集中在指示相对于总能量的扩散混响能量的DSR上:

$$[0152] \quad DSR = \frac{\text{扩散混响能量}}{\text{总发射能量}}$$

[0153] 在下文中这将被称为DSR(扩散-源比率)。

[0154] 可以理解,比率和反比可以提供相同的信息,即,任何比率都可以表示为反比。因此,扩散混响信号与总信号的关系可以用反映扩散混响声音电平的值除以反映总发射声音的值的分数来表示,或者等效地用反映总发射声音的值除以反映扩散混响声音电平的分数的分数来表示。还可以理解,可以引入估计值的各种修改,例如,可以应用非线性函数(例如,对数函数)。

[0155] 这样的方法可以与目前的标准建议相一致。在针对MPEG-I音频征集提案(CfP)的准备工作中,已经定义了编码器输入格式(EIF)(MPEG输出文档N19211中的3.9节,“MPEG-I 6DoF Audio Encoder Input Format”,MPEG130)。EIF通过预延迟和直接-扩散比率(DDR)来定义混响电平。尽管与名称不相符,但它被定义为发射源能量与预延迟后扩散混响能量的比率($DDR=DSR$)。

[0156] 扩散混响能量可以被认为从扩散区段开始由房间响应产生的能量,例如,它可以是从由预延迟指示的时间到无穷远的RIR的能量。注意,房间的后续激励将累加到混响能量,所以这典型地只能通过具有Dirac脉冲的激励来直接测量。替代地,它可以从测量的RIR导出。

[0157] 混响能量表示扩散场空间中单个点的能量,而不是在整个空间上积分的能量。

[0158] 对上述的特别有利的替代方案是使用指示相对于环境中总发射声音的能量的扩散声音的初始幅度的DSR。具体地,DSR可以指示由预延迟指示的时间处的混响幅度。

[0159] 预延迟处的幅度可以是预延迟时或者紧接预延迟后的房间脉冲响应的最大激励。例如,在预延迟后的5ms、10ms、20ms或50ms内。在特定范围内选择最大激励的原因在于,在预延迟时间,房间脉冲响应可以巧合地处于响应的低部分。在总体趋势是幅度衰减的情况下,在预延迟后短间隔内的最大激励典型地也是整个扩散混响响应的最大激励。

[0160] 使用指示初始幅度(在例如10毫秒的间隔内)的DSR使得将DSR映射到许多混响算法中的参数更容易并且更鲁棒。因此,在一些实施例中,DSR可以如下给出:

[0161]
$$DSR = \frac{\text{扩散混响初始幅度}}{\text{总发射能量}}$$

[0162] 在一些实施例中,混响能量参数可以表示针对环境的房间脉冲响应在预先确定时间的幅度。如在上述示例中,幅度可以作为相对幅度(例如,相对于总发射能量)给出,和/或预先确定时间可以是RIR的扩散混响部分的初始化的开始时间。

[0163] DSR中的参数相对于相同的源信号电平参考来表示。

[0164] 例如,这可以通过在特定已知条件(例如,源与麦克风之间的距离和源的方向性模式)内用麦克风测量(或者模拟)感兴趣房间的RIR来实现。源应该向房间发射经过校准的量的能量,例如,已知能量的Dirac脉冲。

[0165] 测量设备中的电转换和模拟到数字转换的校准系数可以被测量或者从规范中导出。它还可以根据RIR中的直接路径响应来计算,该直接路径响应可以根据源的方向性模式和源-麦克风距离来预测。直接响应在数字域中具有特定的能量,并且表示发射能量乘以麦克风方向的方向性增益和距离增益,该距离增益可以依赖于麦克风表面相对于半径等于源-麦克风距离的总球体表面积。

[0166] 两个元件应使用相同的数字电平参考。例如。全刻度1kHz正弦与100dB SPL相对应。

[0167] 从RIR中测量扩散混响能量,并且用校准因子对其进行补偿,得到与已知发射能量相同域的适当能量。加上发射的能量,就可以计算出合适的DSR。

[0168] 参考距离可以指示应用于信号的距离增益为0dB的距离,即,不应该应用增益或者衰减来补偿距离。然后可以通过考虑相对于参考距离的实际距离来计算由路径渲染器401应用的实际距离增益。

[0169] 表示距离对声音传播的影响是参考给定距离执行的。使距离加倍使能量密度(每表面单位的能量)减少6dB。使距离减半使能量密度(每表面单位的能量)增加6dB。

[0170] 为了确定给定距离处的距离增益,必须知道对应于给定电平的距离,因此可以确定当前距离的相对变化,即,以便确定密度减小或者增加了多少。

[0171] 忽略空气中的吸收,并且假设没有反射或者遮挡元件存在,源的发射能量在以源位置为中心的任何半径的任何球体上都是恒定的。与实际距离相对应的表面与参考距离的比率指示能量的衰减。在渲染距离d处的线性信号幅度增益可以表示为:

[0172]
$$g_{\text{dist}} = \sqrt{\frac{A_{\text{ref}}}{A_{\text{rend}}}} = \sqrt{\frac{4\pi r_{\text{ref}}^2}{4\pi d^2}} = \frac{r_{\text{ref}}}{d}$$

[0173] 其中, r_{ref} 为参考距离。

[0174] 例如,如果参考距离为1米并且渲染距离为2米,则这导致约6dB的信号衰减(或者-6dB的增益)。

[0175] 总发射能量指示可以表示声源发射的总能量。典型地,声源向所有方向辐射,但是在所有方向上都相等。在源周围的球体上能量密度的积分可以提供总的发射能量。在扬声器的情况下,通常可以根据施加在终端上的电压和描述阻抗、能量损失和电能转化为声压波的扬声器系数来计算发射能量。

[0176] 在一些实施例中,混响能量参数可以表示环境中针对声音传播的直接响应的能量

等于环境中混响能量的距离。例如,这样的参数可以是临界距离参数。

[0177] 临界距离可以被认为/定义为从源到(潜在的标称/虚拟/理论)点(或者音频接收机(例如,麦克风))的距离,在该点处直接响应的能量等于混响响应的能量。在方向性变化的情况下,该距离可以取决于接收机相对于源的方向而变化。

[0178] 混响声音的能量或多或少地不依赖于房间中的源和接收机位置。早期反射仍然与位置有关,但是越进入RIR,电平对位置的依赖性越小。由于这种属性,存在源的直接声音与相同源的混响声音同样响亮/具有相同的电平的距离。

[0179] 扩散混响在整个房间内具有均匀的电平,而不管音频源的位置如何。直接路径响应的电平非常依赖于麦克风/观察者/收听者的位置和源之间的距离。音频源的直接响应电平的衰减作为其到麦克风的距离的函数是非常明确定义的。因此,在音频源和麦克风之间的距离经常被用来表示临界距离。在该距离处,音频源的直接响应已经衰减到与(恒定)混响电平相同电平。临界距离是本领域技术人员已知的声学属性。

[0180] 在图5的方法中,装置因此可以允许用补偿器修改特定混响元数据参数(延迟和衰减率),然后调整相关联的混响能量元数据。补偿可以例如使得混响能量元数据和其他元数据参数之间的关系根据适当的算法、准则和测量保持与原始相似。然后将修改/补偿后的混响参数馈送给渲染器,其中,混响信号分量的渲染基于修改后的混响参数值而不是原始值。

[0181] 在许多实施例中,混响器400可以具体地被布置为根据第二参数值确定至少一个混响信号分量的电平增益。例如,由渲染器执行以生成混响信号分量的路径/信号处理可以包括设置混响信号分量的能量电平的增益/比例因子。例如,渲染器400可以包括能量归一化函数,紧随其后(或者在其之前)是施加到混响信号分量(或者从其生成的输入音频信号)的可变增益。可变增益可以设置混响信号分量的总电平。渲染器400可以被布置为从修改/补偿后的第二参数值确定可变增益的增益。

[0182] 在许多实施例中,补偿器505包括用于扩散混响的模型,其中,该模型基于混响参数。补偿器505可以被布置为基于该混响模型确定新值,并且具体地可以修改参数,使得对于修改后的参数的模型评估提供期望的结果,该结果典型地可以从初始参数值确定。例如,可以确定补偿后的混响能量参数值,使得可以从原始参数值的模型确定的参数或者测量,对于修改后的混响衰减率参数和/或混响延迟参数和补偿后的混响能量参数的组合来说是不变的(或者以期望的方式改变)。例如,这样的测量可以是RIR的直接路径分量的能量(或者初始时间间隔的能量,该初始时间间隔例如是直到混响开始的时间/延迟)与混响部分的能量之间的能量/电平比率。作为另一示例,测量可以是初始参考幅度。

[0183] 在混响元数据包括衰减率(例如, T_{60} 、 T_{30} 、 T_{20})和混响能量指示(例如,DSR)的比特流中,能量指示必须显式或者隐式地与混响响应/RIR的特定选择相关。这典型地涉及在RIR中的特定滞后/延迟处开始并且在RIR中持续到足够远,其中,响应幅度已经衰减到足够接近RIR中的噪声底限(可以由数字表示的分辨率引起的噪声或者由测量或者测量设备引入的噪声)。由于混响的典型指数衰减性质,针对混响能量的主要定义点典型地是能量测量的开始滞后,其与上述的预延迟参数相对应。

[0184] 预延迟值可以与其他混响元数据一起提供,但是还可以通过应用中使用的混响能量指示的定义来暗示。

[0185] 一般的数学方程可以典型地用作扩散混响幅度包络的简单模型。指数函数典型地

很好地匹配衰减幅度包络:

$$[0186] \quad AA = A_0 \cdot e^{-\alpha(t-t_{pre})}$$

[0187] 对于 $t \geq t_{pre}$ = 预延迟, 并且 $\alpha = \frac{\ln(10^{-\frac{60}{T60}})}{T60}$ (由T60控制的衰减因子), 以及 A_0 为预延迟时 (t_{pre}) 的幅度。因此, 在这种情况下, 混响延迟参数可以由预延迟给出, 混响衰减率参数由T60值给出, 并且混响能量参数由预延迟时 (t_3) 的幅度给出。

[0188] 计算这样的函数的累积能量, 它将渐近地接近一些最终能量值, 如图7所指示的。

[0189] 典型地, 扩散混响作为时间的函数是相当稀疏的 (许多值低于指数函数给出的幅度指示), 并且为了从上述方程确定混响的能量, 典型地包括补偿, 通常简称为比例因子。

[0190] 事实上, 从数学模型开始, 用模型计算的能量典型地与混响能量成正比。因此, 在没有 (经验推导的) 校正的情况下, 它通常不是预测混响能量的合适模型。然而, 该比例可以用于在没有任何校正的情况下计算针对预延迟或者 T_{60} 的修改的能量调整因子。该模型可以利用从预延迟到无穷大的积分 (因为模型中不包括噪声底限) 来计算混响能量, 并且可以解析求解 (使用 $t = \frac{n}{f_s}$):

$$[0191] \quad G_{corr} \cdot E_{pre} = \int_{n=n_{pre}}^{\infty} A_0^2 \cdot e^{-2\alpha \frac{n-n_{pre}}{f_s}} = \left[-\frac{A_0^2 \cdot f_s}{2\alpha} \cdot e^{-2\alpha \frac{(n-n_{pre})}{f_s}} \right]_{n_{pre}}^{\infty} = \frac{A_0^2 \cdot f_s}{2\alpha}$$

[0192] 其中, G_{corr} 表示将模型能量映射到混响能量的校正因子, A_0 表示在 $t = t_{pre}$ (预延迟) 时的初始混响幅度, 并且 E_{pre} 表示预延迟之后的混响能量。

[0193] 该模型可以例如用于确定修改之前和修改之后的模型能量预测之间的比率, 然后可以对混响能量参数进行适配以反映该改变, 例如, 它可以简单地由相同的比率来补偿。

[0194] 在一些实施例中, 修改器503可以具体地被布置为修改指示环境/RIR中混响的传播时间延迟的混响延迟参数。具体地, 修改器503可以被布置为修改预延迟。预延迟典型地用于指示RIR的扩散混响部分的开始。因此, 预延迟可以指示RIR由扩散混响支配的时间 (延迟), 并且因此典型地由扩散混响渲染器 (例如, Jot混响器) 渲染的部分。因此, 预延迟典型地由渲染器使用以指示由扩散混响渲染功能而不是由路径渲染器渲染RIR的哪个部分。在图4的示例中, 预延迟用于指示分别由混响器407和路径渲染器401渲染的RIR的时刻。

[0195] 在一些实施例中, 修改器403可以被布置为在渲染之前修改预延迟 (无论是默认值或者还是由接收到的元数据指示的值)。这可以修改由扩散混响渲染器407建模了多少RIR以及由路径渲染器401渲染了多少RIR。如图8和图9所示, 其示出了RIR的扩散混响部分, 修改前的预延迟 t_{pre} 可以修改到新的值 t_{rend} , 其可以比原始值 t_{pre} 早 (图8) 或者晚 (图9)。

[0196] 在一些实施例中, 可以例如手动地执行这样的修改以实现期望的感知效果。例如, 路径渲染器可以倾向于提供更准确的渲染, 并且用户可以例如通过修改预延迟来调整所渲染音频的质量。

[0197] 然而, 在一些实施例中, 修改可以是自动的。例如, 路径渲染倾向于比使用参数混响器的扩散混响渲染对计算资源的要求显著地更高。在一些实施例中, 修改器可以被布置为确定设备的计算负载和/或确定用于渲染的可用计算资源量 (用于确定这样的测量的许多方法对于本领域技术人员来说是已知的)。修改器可以被布置为响应于可用计算资源修改混响延迟参数/预延迟。特别地, 它可以针对增加的可用资源量而增加延迟, 并且针对减

少的可用资源量而减少延迟。例如,延迟(修改)可以是可用计算资源的单调递减函数。

[0198] 除了渲染器配置之外,还可以出于其他原因改变预延迟参数,例如,将元数据转码为要求与隐式预延迟值或者具有特定滤波器长度的共同信令HRTF对齐的不同格式。

[0199] 因此,包括扩散混响渲染的渲染器可以从与元数据指示的预延迟(或者默认/标称预延迟)不同的滞后渲染扩散混响。因此,所要求的混响能量将不同于接收到的元数据指示的混响能量,这导致混响效果/体验不同于元数据所期望的混响效果/体验。在许多情况下,这种差距可以是显著的。

[0200] 在所描述的方法中,补偿器505可以调整元数据的混响能量参数以表示感知上相似的混响能量元数据,对于该混响能量元数据,调整后的预延迟与渲染延迟(或者以其他方式的目标延迟)相对应。调整可以使得具有更新后的预延迟的混响能量表示与原始混响能量元数据相似的混响效果/体验。例如,在图8和图9中,灰色区域指示扩散混响器应提供的混响能量。这与从预延迟 t_{pre} 到无穷远的RIR不同。在图8中,能量元数据值太低,混响渲染不能在较早的滞后处开始(虚线三角形)。在图9中,能量元数据值太高,无法在较晚的滞后处开始渲染(虚线三角形)。

[0201] 在许多实施例中,修改器505可以被布置为修改混响能量参数,使得在修改混响延迟参数之后的RIR的部分期间混响的能量/幅度/电平被认为是混响部分,并且其具体地将由混响渲染器渲染,在使用由参数指示的初始延迟和能量以及当使用修改后的延迟和能量确定时将是相似的或者甚至相同的。

[0202] 具体地,在许多实施例中,补偿器505可以被布置为确定修改后的混响能量参数值,使得其减少在第一混响能量测量和第二混响能量测量之间的差。对于从修改后的延迟值开始的混响确定两个能量测量,并且使用相同的模型确定两个能量测量,例如具体地,先前引入的指数下降混响模型。然而,第一测量是通过使用针对混响延迟参数和混响能量参数的修改后的参数值评估模型来确定的,而第二测量是通过使用针对混响延迟参数和混响能量参数的初始(修改/补偿之前)参数值评估模型来确定的。补偿器505可以具体地设置修改后的混响能量参数值,使得这些能量相等,从而使得修改后的延迟之后的混响能量将与原始值一致。

[0203] 因此,可以将第一混响能量测量确定为由修改后的混响延迟参数表示的修改后的延迟之后的混响能量。它可以使用修改后的延迟值和修改后的混响能量参数从混响模型确定。第一混响能量测量可以指示使用修改后的值计算的修改后的延迟之后的混响能量。

[0204] 第二混响能量测量还可以被确定为由修改后的混响延迟参数表示的修改后的延迟之后的混响能量。它还可以从相同的混响模型中确定,但是通过使用初始延迟值和初始混响能量参数。第二混响能量测量可以指示使用初始值计算的修改后的延迟之后的混响能量。

[0205] 在许多实施例中,补偿器505可以被布置为修改混响能量参数,使得其减少(或者甚至移除)作为时间函数的混响幅度的差,以用于在修改后的延迟(具体地,指示由混响渲染器渲染的RIR的部分的渲染延迟)之后的混响。

[0206] 如前所述,混响渲染器典型地被布置为生成混响信号分量,以仅包括与超过由修改后的延迟指示的传播延迟时间的传播延迟对应的贡献。混响渲染器可以具体地实现RIR中在修改后的延迟时间之后的部分。

[0207] 作为使用先前提供的指数模型的具体示例,可以认为,如果自初始未修改的预延迟和之后的混响的能量与模型能量(G_{corr})成比例,那么自修改后的预延迟的混响的能量将以相同的方式成比例(即,指示稀疏性所要求的补偿可以是相同的)。

$$[0208] \quad G_{corr} \cdot E_{render} = \int_{n=n_{render}}^{\infty} A_0^2 \cdot e^{-2\alpha \frac{n-n_{pre}}{f_s}} = \left[-\frac{A_0^2 \cdot f_s}{2\alpha} \cdot e^{-2\alpha \frac{(n-n_{pre})}{f_s}} \right]_{n_{render}}^{\infty}$$

$$[0209] \quad G_{corr} \cdot E_{render} = \frac{A_0^2 \cdot f_s}{2\alpha} \cdot e^{-2\alpha \frac{(n_{render}-n_{pre})}{f_s}}$$

[0210] 其中, $\alpha = \frac{\ln(10^{-\frac{60}{20}})}{T_{60}}$ 和 E_{render} 是基于模型计算出的能量测量(并且索引pre总体上用于指示修改前的初始值,并且索引render用于指示修改后的值)。

[0211] 可以利用这些方程计算能量转换因子,其将混响能量元数据从对应于初始预延迟的值缩放到对应于修改后的预延迟(也被称为渲染延迟)的值,并且仍然描述相同的混响特性:

$$[0212] \quad G_{conv} = \frac{E_{render}}{E_{pre}} = \frac{\frac{A_0^2 \cdot f_s}{2\alpha} \cdot e^{-2\alpha \frac{(n_{render}-n_{pre})}{f_s}}}{\frac{A_0^2 \cdot f_s}{2\alpha}} = e^{-2\alpha \frac{(n_{render}-n_{pre})}{f_s}}$$

[0213] 从方程可以看出,当 $n_{render} > n_{pre}$ 时,转换因子小于1;并且当 $n_{pre} > n_{render}$ 时,转换因子大于1。

[0214] 例如,在使用 DSR_{render} 计算混响渲染的配置之前,可以用它来补偿DSR参数:

$$[0215] \quad DSR_{render} = DSR_{metadata} * G_{conv}$$

[0216] 在一些实施例中,修改器可以被布置为修改混响衰减率,例如 T_{60} 值。例如,在许多实施例中,这可以是期望的,以便通过修改感知到的混响量来修改环境的感知体验。例如,它可以由用户手动修改以提供修改后的感知,例如具体地提供不同的艺术效果。

[0217] 然而,修改衰减率还可以影响混响能量。 T_{60} 越短,则混响能量越少,因为它对应于更快的衰减。

[0218] 此外,改变的衰减率不仅可以影响预延迟之后混响响应的衰减率,而且典型地还影响预延迟之前的衰减,并且因此影响与混响能量指示相关联的预延迟滞后处的初始混响响应幅度。这可以通过图10、图11和图12示出,其示出了修改/补偿之前的混响能量参数指示的能量(由灰色三角形指示)与所期望的渲染条件(即,对于修改后的衰减参数)不匹配的情况。在图10中,未修改的混响能量参数将具有太高的值,无法以较短的衰减时间渲染混响(虚线三角形)。在图11中,未修改的混响能量参数将具有太低的值,无法以较长的衰减时间渲染混响(虚线三角形)。

[0219] 在图5的系统中,补偿器可以补偿混响能量参数以指示可以与修改后的混响衰减率参数值相对应的修改后的能量电平。对于增加的衰减率可以减少指示的能量值和/或对于减小的衰减率可以增加指示的能量值。

[0220] 在许多实施例中,补偿器505可以被布置为修改混响能量参数值,以减少由于第一混响参数的修改而导致的混响衰减率的幅度参考(图12中的 A_{00}) 的改变,并且具体地,可以寻求将该参考幅度维持为实质上不改变。

[0221] 幅度参考是混响衰减率和混响能量参数的函数,并且例如可以被认为是RIR在 $t=0$ 时的值,该值导致RIR的扩散混响部分(即,预延迟之后的RIR)的衰减率和能量电平,如衰减率和混响能量指示所指示的。

[0222] 这典型地可以导致混响能量参数被修改以与修改后的衰减率相对应,类似于原始混响能量元数据如何与原始衰减率相对应。

[0223] 作为具体示例,修改器503可以改变 T_{60} 值以修改房间特性,并且作为响应,以DSR的形式修改混响能量参数。基于例如先前呈现的用于混响的模型,可以确定应该如何调整DSR。典型地,当 T_{60} 改变时,在扩散混响的预延迟时间/开始时的幅度 A_0 也改变,如图12所示。因此,可以认为对DSR有双重影响,一个直接来自混响期间改变的衰减,并且另一个是改变的衰减对RIR直到预延迟的影响,并且因此对混响部分开始时的幅度 A_0 的影响。

[0224] A_0 的改变可以由预延迟之前改变的衰减率的影响来确定。典型地,RIR的早期部分非常依赖于RIR的测量或者建模中使用的源和接收机位置。例如,这导致早期衰减,当源和接收机相对较近时,该早期衰减导致RIR的早期部分中更陡的衰减。

[0225] 在为扩散混响建模调整混响参数方面,忽略这些方面并且假定RIR在其整个长度上具有一致的衰减率通常是有益的。这与源和接收机相对距离较远很匹配。

[0226] 为此,该方法可以基于针对衰减线在 $t=t_0$ 时的参考幅度,如图12所示。

$$[0227] \quad A_{00} = A_0 * 10^{\left(\frac{t_{pre}-t_0}{T_{60}} * \frac{60}{20}\right)}$$

[0228] 其中,典型地 $t_0=0$ 。

[0229] 接下来,针对修改后的混响延迟参数的修改后的 A_0 值(A_r)可以利用修改后的 T_{60} 参考的 T_{60_r} 来计算。

$$[0230] \quad A_r = A_{00} * 10^{-\left(\frac{t_{pre}-t_0}{T_{60_r}} * \frac{60}{20}\right)}$$

[0231] 或者,放在一起

$$[0232] \quad A_r = A_0 * 10^{\left(\left(\frac{T_{60_r}-T_{60}}{T_{60} * T_{60_r}}\right) * \frac{60}{20} * (t_{pre}-t_0)\right)}$$

[0233] 然后,针对混响能量的转换因子变成:

$$[0234] \quad G_{conv} = \frac{E_{render}}{E_{pre}} = \frac{A_r^2 \alpha}{\alpha_r A_0^2}$$

[0235] 其中, $\alpha_r = \frac{\ln\left(10^{-\frac{60}{20}}\right)}{T_{60_r}}$ 。进一步简化为:

$$[0236] \quad G_{conv} = \frac{T_{60_r}}{T_{60}} * 10^{\left(2 * \left(\frac{T_{60_r}-T_{60}}{T_{60} * T_{60_r}}\right) * \frac{60}{20} * (t_{pre}-t_0)\right)}$$

[0237] 转换增益通过乘法施加,类似于混响延迟参数的修改的情况。

[0238] 当 T_{60} 依赖于频率时,转换增益依赖于频率。

[0239] 在上述示例中,混响能量参数的补偿通过确定线性转换或者补偿因子并且将其应用于以DSF参数形式的混响能量参数来简单地实现。

[0240] 对于混响能量参数,例如,作为临界距离或者幅度参数,可以使用类似的方法。

[0241] 例如,如果混响能量参数是临界距离参数,这也意味着特定的预延迟,根据该预延迟计算混响响应能量。因此,可以应用相同的转换。例如:

[0242] $E_{pre} = E_{cd}$

[0243] $E_{rend} = E_{pre} * G_{conv} = E_{cd} * G_{conv}$

[0244] 其中, E_{cd} 是临界距离处直接响应的能量, E_{pre} 是从与临界距离元数据相关联的预延迟中测量的混响能量, 并且 E_{rend} 表示来自渲染延迟的混响能量。

[0245] 在混响能量参数以幅度表示的示例中, 例如初始混响能量幅度与源能量 (或者总能量或者源幅度) 的比率, 取增益的平方根, 这是本领域技术人员所已知的。

[0246] $DSR_{render} = DSR_{metadata} * \sqrt{G_{conv}}$

[0247] 如果混响延迟参数和混响衰减率参数两者都改变, 则可以组合补偿。例如, 可以例如简单地通过将上述不同参数所指示的转换增益相乘来进行组合。

[0248] 下面将更详细地描述图4和图5所示方法的各种实施例的具体方面。

[0249] 渲染器407可以通过生成各个音频源的下混频, 并且然后将该信号应用到参数混响器, 例如, 图13的Jot混响器, 来具体地生成混响, 其中, 参数混响器是基于混响参数建立的。

[0250] 该方法可以基于将混响过程应用于如前所述和如图14所示的下混频信号。下混频系数可以被确定并且与下混频中该音频信号的加权相对应。下混频系数可以是生成下混频信号的加权组合中的音频信号的加权。因此, 当组合音频信号以生成下混频信号 (在许多实施例中是单声道信号) 时, 下混频系数可以是音频信号的相对加权, 例如, 它们可以是加权求和的加权。

[0251] 下混频系数可以基于接收到的扩散混响信号与总信号的比率, 即, 扩散与源的比率DSR。

[0252] 响应于所确定的总发射能量指示来进一步确定系数, 该总发射能量指示指示从音频源发射的总能量。虽然DSR典型地对于一些音频信号 (并且典型地对所有音频信号) 是公共的, 但是总发射能量指示典型地是针对每个音频源的。

[0253] 总发射能量指示典型地指示归一化的总发射能量, 并且可以独立于信号内容, 完全由诸如方向性模式和参考距离之类的源属性定义。相同的归一化可以应用于所有音频源以及直接和反射路径分量。因此, 总发射能量指示可以是相对于针对其他音频源/信号的总发射能量指示或者相对于各个路径分量或者相对于音频信号的全刻度采样值的相对值。

[0254] 当与DSR组合时, 总发射能量指示可以针对每个音频源提供下混频系数, 其反映对来自该音频源的扩散混响声音的相对贡献。因此, 将下混频系数确定为DSR和总发射能量指示的函数可以提供反映对扩散声音的相对贡献的下混频系数。因此, 使用下混频系数来生成下混频信号可以产生反映环境中生成的总声音的下混频信号, 其中, 声源中的每一个被适当地加权并且声学环境被准确地建模。

[0255] 在许多实施例中, 作为DSR和总发射能量指示的函数的下混频系数与响应于混响器属性的缩放相组合可以提供这样的下混频系数: 其反映扩散混响声音相对于对应路径信号分量的适当相对电平。

[0256] 总发射能量可以根据针对音频源接收的元数据来确定。

[0257] 接收到的元数据可以包括针对每个源的信号参考电平, 其提供音频电平的指示。信号参考电平典型地是标准化的值或者相对值, 其提供相对于其他音频源或者相对于标准化参考电平的信号参考电平的指示。因此, 信号参考电平典型地可以不指示针对源的绝对

声音电平,而是指示相对于其他音频源的相对电平。

[0258] 在具体示例中,信号参考电平可以包括以参考距离形式的指示,该参考距离提供将应用于音频信号的距离衰减为0dB的距离。因此,对于等于参考距离的在音频源和收听者之间的距离,可以使用接收到的音频信号而不用任何与距离相关的缩放。对于小于参考距离的距离,衰减较小,并且因此在确定收听位置处的声音电平时应该应用高于0dB的增益。对于高于参考距离的距离,衰减较高,并且因此在确定收听位置处的声音电平时应该应用高于0dB的衰减。等同地,对于音频源和收听位置之间的给定距离,将对与较高参考距离相关联的音频信号应用比与较短参考距离相关联的音频信号更高的增益。由于音频信号典型地被归一化以表示有意义的参考距离或者利用整个动态范围(例如,喷气发动机和板球两者都将由利用所使用的数据字的整个动态范围的音频信号来表示),所以参考距离针对具体音频源提供信号参考电平的指示。

[0259] 在该示例中,信号参考电平进一步由被称为预增益的参考增益指示。针对每个音频源提供参考增益,并且提供在确定所渲染的音频电平时应该应用于音频信号的增益。因此,预增益可以用于进一步指示不同音频源之间的电平变化。

[0260] 元数据进一步可以包括指示来自音频信号表示的声源的声音辐射的方向性的方向性数据。针对每个音频源的方向性数据可以指示在距音频源不同的方向上相对于信号参考电平的相对增益。方向性数据可以例如提供定义每个方向上的增益的对来自音频源的辐射模式的完整函数或者描述。作为另一示例,可以使用简化的指示,例如指示预先确定模式的单个数据值。作为又一示例,方向性数据可以针对不同方向间隔的范围(例如,球体的区段)提供各个增益值。

[0261] 因此,元数据与音频信号一起可以允许生成音频电平。具体地,路径渲染器可以通过将增益应用于音频信号来确定针对直接路径的信号分量,其中,增益是预增益、根据音频源和收听者之间的距离和参考距离确定的距离增益、以及在从音频源到收听者的方向上的方向性增益的组合。

[0262] 关于扩散混响信号的生成,元数据用于基于针对音频源的信号参考电平和方向性数据来确定针对音频源的(归一化的)总发射能量指示。

[0263] 具体地,可以通过在所有方向上积分方向性增益(例如,在以音频源位置处为中心的球体表面上积分)来生成总发射能量指示,并且由信号参考电平对其进行缩放,并且具体地由距离增益和预增益缩放。

[0264] 然后利用DSR处理确定出的总发射能量指示,以生成下混频系数。

[0265] 然后使用下混频系数来生成下混频信号。具体地,可以将下混频信号生成为音频信号的组合,并且具体地说是求和,其中,每个音频信号由针对对应音频信号的下混频系数加权。

[0266] 下混频典型地被生成为单信号,然后该单信号被馈送到混响器,该混响器继续生成扩散混响信号。

[0267] 应当注意,尽管由路径渲染器401对各个路径信号分量的渲染和生成例如关于确定距离增益和方向性增益是位置相关的,但是扩散混响信号的生成可以独立于源和收听者两者的位置。

[0268] 可以基于信号参考电平和方向性数据来确定总发射能量指示,而不考虑源和收听

者的位置。具体地,针对源的预增益和参考距离可以用于确定与源的标称距离处的非方向性相关的信号参考电平(标称距离对于所有音频信号/源都是相同的),并且该标称距离相对于例如音频信号的全刻度样本进行归一化。对于归一化球面,例如对于在参考距离处的球面,可以对所有方向上的方向性增益执行积分。因此,总发射能量指示将独立于源和收听者位置(反映扩散混响声音在诸如房间之类的环境中趋于均匀)。然后将总发射能量指示与DSR组合,以生成下混频系数(在许多实施例中,还可以考虑诸如混响器的参数的其他参数)。由于DSR也与位置无关,就像下混频和混响处理一样,因此可以在不考虑源和收听者的具体位置的情况下生成扩散混响信号。

[0269] 这样的方法可以提供高性能并且自然地发声的音频感知,而不用要求过多的计算资源。它可以特别地适合于例如虚拟现实应用,其中,用户(以及源)可以在环境中移动,并且因此收听者(以及可能的一些或者全部音频源)的相对位置可以动态地改变。

[0270] 混响器可以通过考虑针对音频源的方向性数据来确定总发射能量指示。应当注意,在确定可能具有变化的源方向性的针对源的扩散混响信号时,重要的是使用总发射能量,而不仅仅是信号电平或者信号参考电平。例如,考虑源方向性对应于方向性系数为1的非常窄的波束,并且对于所有其他方向的系数为0(即,能量只在非常窄的波束中发送)。在这种情况下,发射的源能量可以非常类似于音频信号和信号参考电平的总能量,因为这表示总能量。如果替代地考虑具有相同能量和信号参考电平但是具有全向方向性的音频信号的另一源,则该源的发射能量将远高于音频信号能量和信号参考电平。因此,在两个源同时活动的情况下,全向源的信号在扩散混响信号中(并且因此在下混频中)应当表现得比特定源强得多。

[0271] 发射的能量可以通过对围绕音频源的球体表面上的能量密度进行积分来确定。忽略距离增益,即,在距离增益为0dB的半径(即,半径对应于参考距离)的表面上积分,总发射能量指示可以由以下确定:

$$[0272] \quad E = \iint_S (g \cdot p \cdot x)^2 dS$$

[0273] 其中,g是方向性增益函数,p是与音频信号/源相关联的预增益,并且x指示音频信号本身的电平。

[0274] 由于p与方向无关,也可以将它从积分中移除。类似地,信号x与方向无关(方向性增益反映了这种变化)。可以稍后对其相乘,因为:

$$[0275] \quad g_1^2 * x^2 + g_2^2 * x^2 + g_3^2 * x^2 + \dots = (g_1^2 + g_2^2 + g_3^2 + \dots) * x^2$$

[0276] 并且因此积分变得与信号无关。

[0277] 下面将更详细地描述用于确定该积分的一种具体方法。

[0278] 期望对球面上的方向性增益进行积分。

$$[0279] \quad E_{scale} = \iint_S g^2 dS$$

[0280] 使用半径等于参考距离(r)的球体意味着距离增益是0dB,并且因此可以忽略距离增益/衰减。

[0281] 在这个示例中选择球体是因为它提供了有利的计算,但是相同的能量可以从包围源位置的任何形状的任何封闭表面确定。只要在积分中使用适当的距离增益和方向性增

益,并且考虑有效表面正对着源位置(即,法向量与源位置一致)。

[0282] 表面积分应该定义小表面dS。因此,用方位角(a)和仰角(e)两个参数定义球体提供了这样做的维度。用坐标系作解,我们得到:

$$[0283] \quad f(a, c, r) = r \cdot \cos(e) \cdot \cos(a) \cdot u_x + r \cdot \cos(e) \cdot \sin(a) \cdot u_y + r \cdot \sin(e) \cdot u_z$$

[0284] 其中, u_x 、 u_y 和 u_z 是坐标系的单位基向量。

[0285] 小表面dS是球面相对于两个参数的偏导数交叉乘积的幅度,乘以每个参数的微分:

$$[0286] \quad dS = |f_a \times f_e| da de$$

[0287] 导数确定在感兴趣点处与球体相切的向量。

$$[0288] \quad f_a = -r \cdot \cos(e) \cdot \sin(a) \cdot u_x + r \cdot \cos(e) \cdot \cos(a) \cdot u_y + 0 \cdot u_z$$

$$[0289] \quad f_e = -r \cdot \sin(e) \cdot \cos(a) \cdot u_x - r \cdot \sin(e) \cdot \sin(a) \cdot u_y + r \cdot \cos(e) \cdot u_z$$

[0290] 导数的交叉乘积是垂直于两者的向量。

$$[0291] \quad f_a \times f_e = (r^2 \cdot \cos(e) \cdot \cos(a) \cdot \cos(e) + 0 \cdot \sin(e) \cdot \sin(a)) \cdot u_x + (-r^2 \cdot \cos(e) \cdot \sin(a) \cdot \cos(e) + r^2 \cdot \cos(e) \cdot \sin(a) \cdot \sin(e) \cdot \sin(a)) \cdot u_y + (r^2 \cdot \cos(e) \cdot \sin(a) \cdot \sin(e) \cdot \sin(a) + r^2 \cdot \cos(e) \cdot \cos(a) \cdot \sin(e) \cdot \cos(a)) \cdot u_z$$

$$[0292] \quad = r^2 \cdot \cos^2(e) \cdot \cos(a) \cdot u_x + r^2 \cdot \cos^2(e) \cdot \sin(a) \cdot u_y + (r^2 \cdot \cos(e) \cdot \sin(e) \cdot \sin^2(a) + r^2 \cdot \cos(e) \cdot \sin(e) \cdot \cos^2(a)) \cdot u_z$$

$$[0293] \quad = r^2 \cdot \cos^2(e) \cdot \cos(a) \cdot u_x + r^2 \cdot \cos^2(e) \cdot \sin(a) \cdot u_y + (r^2 \cdot \cos(e) \cdot \sin(e) \cdot (\sin^2(a) + \cos^2(a))) \cdot u_z$$

$$[0294] \quad = r^2 \cdot \cos^2(e) \cdot \cos(a) \cdot u_x + r^2 \cdot \cos^2(e) \cdot \sin(a) \cdot u_y + r^2 \cdot \cos(e) \cdot \sin(e) \cdot u_z$$

[0295] 交叉乘积的幅度是由向量 f_a 和 f_e 跨越的平行四边形的表面积,并且因此也是球体上的表面积:

$$[0296] \quad |f_a \times f_e| = \sqrt{(r^2 \cdot \cos^2(e) \cdot \cos(a))^2 + (r^2 \cdot \cos^2(e) \cdot \sin(a))^2 + (r^2 \cdot \cos(e) \cdot \sin(e))^2}$$

$$[0297] \quad = \sqrt{r^4 \cdot \cos^4(e) \cdot \cos^2(a) + r^4 \cdot \cos^4(e) \cdot \sin^2(a) + r^4 \cdot \cos^2(e) \cdot \sin^2(e)}$$

$$[0298] \quad = \sqrt{r^4 \cdot \cos^4(e) \cdot (\cos^2(a) + \sin^2(a)) + r^4 \cdot \cos^2(e) \cdot \sin^2(e)}$$

$$[0299] \quad = \sqrt{r^4 \cdot \cos^4(e) + r^4 \cdot \cos^2(e) \cdot \sin^2(e)}$$

$$[0300] \quad = \sqrt{r^4 \cdot \cos^2(e) \cdot (\cos^2(e) + \sin^2(e))}$$

$$[0301] \quad = \sqrt{r^4 \cdot \cos^2(e)}$$

$$[0302] \quad = \text{abs}(r^2 \cdot \cos(e)) = r^2 \cdot \cos(e) \text{ 当 } e = [-0.5\pi, 0.5\pi]$$

[0303] 结果是:

$$[0304] \quad dS = r^2 \cdot \cos(e) \cdot da \cdot de$$

[0305] 其中,前两个项定义了归一化的表面积,并且基于区段da和de的大小乘以da和de,就变成了实际的表面。然后,表面上的二重积分可以用方位角和仰角表示。表面dS如上所述用a和e表示。这两个积分可以在方位角 $= 0 \cdots 2\pi$ (内积分) 和仰角 $= -0.5\pi \cdots 0.5\pi$ (外积分) 上执行。

$$[0306] \quad E_{scale} = \int_{-0.5\pi}^{0.5\pi} \int_0^{2\pi} g^2(a, e) \cdot r^2 \cdot \cos(e) \cdot da \cdot de$$

[0307] 其中, $g(a, e)$ 是作为方位角和仰角的函数的方向性。因此, 如果 $g(a, e) = 1$, 结果应该是球的表面。(用解析法求积分作为证明, 结果如预期的 $4\pi r^2$)。

[0308] 在许多实际实施例中, 方向性模式可以不作为可积函数提供, 但是例如作为分立的采样点集合提供。例如, 每个采样的方向性增益与方位角和仰角相关联。典型地, 这些示例将表示球体上的网格。处理这一问题的一种方法是将积分转化为求和, 即, 可以执行分立积分。在该示例中, 积分可以实现为球面上方向性增益可用的点的求和。这给出了 $g(a, e)$ 的值, 但是要求正确地选择 da 和 de , 从而它们就不会由于重叠或者间隙而导致大的误差。

[0309] 在其他实施例中, 方向性模式可以被提供为空间中有限数量的非均匀间隔的点。在这种情况下, 方向性模式可以在感兴趣的方位角和仰角范围内被插值和均匀地重新采样。

[0310] 替代解决方案可以是假定 $g(a, e)$ 在其定义点附近是常数, 并且用局部解析法求解积分。例如, 对于小方位角和仰角范围。例如, 相邻定义点之间的中间点。这使用了上面的积分, 但是 a 和 e 的范围不同, 并且 $g(a, e)$ 假定为常数。

[0311] 实验表明, 通过直接求和, 即使方向性的分辨率相当粗糙, 误差也很小。此外, 误差与半径无关。对于 10 个点之间的方位角的线性间距, 以及仰角的 10 个线性间隔点, 导致相对误差为 -20dB。

[0312] 如上所表示的积分, 提供了以球半径进行缩放的结果。因此, 它以参考距离进行缩放。这种对半径的依赖是因为我们没有考虑到两个不同半径之间的“距离增益”的反向影响。如果半径加倍, 则“流”过固定表面面积 (例如, 1cm^2) 的能量会降低 6dB。因此, 可以说积分应该考虑距离增益。然而, 积分是在参考距离进行的, 该参考距离被定义为距离增益反映在信号中的距离。换言之, 由参考距离指示的信号电平不被包括作为被积分值的缩放, 但是由随参考距离变化的被执行积分的表面积反映 (因为积分是在半径等于参考距离的球体上执行的)。

[0313] 因此, 如上所述的积分反映音频信号能量缩放因子 (包括任何预增益或者类似的校准调整), 因为音频信号表示半径等于参考距离 (没有方向性增益) 的球体上固定表面积处的正确信号回放能量。

[0314] 这意味着如果参考距离较大, 在不改变信号的情况下, 总信号能量缩放因子也较大。这是因为对应的信号表示声源, 该声源比具有相同信号能量但是在较小的参考距离处的声源相对响亮。

[0315] 换言之, 通过在半径等于参考距离的球面上执行积分, 自动地考虑由参考距离提供的信号电平指示。更高的参考距离将导致更大的表面积, 并且从而导致更大的总发射能量指示。具体地, 在距离增益为 1 的距离处直接执行积分。

[0316] 上面的积分得到归一化到所使用的表面单位和用于指示参考距离 r 的单位的值。如果参考距离 r 以米表示, 那么积分的结果以 m^2 为单位提供。

[0317] 为了将估计出的发射能量值与信号相关, 应该以对应于信号的表面单位表示。由于信号的电平表示了参考距离处应该为用户播放的电平, 因此人耳的表面积可能更适合。在参考距离处, 这个表面相对于整个球体表面, 将与人将感知到的源能量的部分相关。

[0318] 因此, 表示针对音频信号中的全刻度样本归一化的发射源能量的总发射能量指示可以通过以下指示:

$$[0319] \quad E_{norm} = \frac{E_{dir,r} * p^2}{S_{ear}}$$

[0320] 其中, $E_{dir,r}$ 指示通过对半径等于参考距离的球体表面上的方向性增益进行积分而确定的能量, p 是预增益, 并且 S_{ear} 是归一化缩放因子 (将确定的能量与人耳的面积相关)。

[0321] 利用表征空间的扩散声学特性的DSR以及从方向性、预增益和参考距离元数据得出的计算出的发射源能量, 可以计算对应的混响能量。

[0322] DSR典型地可以用其两个分量使用的相同参考电平来确定。这可能与总发射能量指示相同或者可以不同。无论如何, 当这样的DSR与总发射能量指示组合时, 当使用由上述积分确定的总发射能量时, 所得到的混响能量也表示为针对音频信号中的全刻度样本进行归一化的能量。换言之, 所考虑的所有能量本质上都归一化到相同的参考电平, 以便它们可以直接组合而不要求电平调整。具体地, 所确定的总发射能量可以直接与DSR一起使用, 以针对从每个源生成的扩散混响生成电平指示, 其中, 该电平指示直接指示相对于针对其他音频源的扩散混响和相对于各个路径信号分量的适当电平。

[0323] 作为具体示例, 针对不同源的扩散混响信号分量的相对信号电平可以通过将DSR乘以总发射能量指示直接获得。

[0324] 在所描述的系统中, 不同音频源对扩散混响信号的贡献的适配至少部分地通过适配用于生成下混频信号的下混频系数来执行。因此, 可以生成下混频系数, 使得来自每个音频源的扩散声音的相对贡献/能量电平反映针对源的确定的扩散混响能量。

[0325] 作为具体示例, 如果DSR指示初始幅度电平, 则可以确定下混频系数与DSR乘以总发射能量指示成比例 (或者相等)。如果DSR指示能量电平, 则可以确定下混频系数与DSR的平方根乘以总发射能量指示成比例 (或者相等)。

[0326] 作为具体示例, 用于多个输入信号中具有索引 x 的信号提供适当调整的下混频系数 d_x 可以通过以下计算:

$$[0327] \quad d_x = p \cdot \sqrt{E_{norm,x} * DSR}$$

[0328] 其中, p 表示预增益, 并且 $E_{norm,x} = \frac{E_{scale}}{S_{ear}}$ 为在预增益之前信号 x 的归一化发射源能量。DSR表示扩散混响能量与发射源能量之比。当下混频系数 d_x 被应用于输入信号 x 时, 所得到的信号表示信号电平, 该信号电平在被具有单位能量的混响响应的混响器滤波时, 相对于信号 x 的直接路径渲染以及相对于其他源 $j \neq x$ 的直接路径和扩散混响能量, 针对信号 x 提供正确的扩散混响能量。

[0329] 可替代地, 可以根据以下计算下混频系数 d_x :

$$[0330] \quad d_x = E_{norm,x} * DSR$$

[0331] 其中, $E_{norm,x} = \frac{E_{dir,r} * p^2}{S_{ear}}$ 表示针对信号 x 的归一化发射源能量, 并且DSR表示扩散混响能量与初始混响响应幅度之比。当下混频系数 d_x 应用于输入信号 x 时, 所得到的信号表示对应于扩散混响信号的初始电平的信号电平, 并且可以由具有以幅度1开始的混响响应的混响器处理。因此, 混响器的输出相对于信号 x 的直接路径渲染以及相对于其他源 $j \neq x$ 的直接路径和扩散混响能量, 为信号 x 提供正确的扩散混响能量。

[0332] 在许多实施例中, 通过将DSR与总发射能量指示相组合来部分地确定下混频系数。

无论DSR指示总发射能量与扩散混响能量还是用于扩散混响响应的初始幅度的关系,通常需要进行进一步适配下混频系数以适配所使用的具体混响器算法,该算法对信号进行缩放,使得混响处理器的输出反映期望的能量或者初始幅度。例如,在混响算法中,当输入电平保持相同时,反射的密度对产生的混响能量具有很强的影响。作为另一示例,混响算法的初始幅度可以不等于其激励的幅度。因此,可能需要特定于算法或者特定于算法和配置的调整。这可以包括在下混频系数中,并且典型地对所有源都是通用的。对于一些实施例,这些调整可以应用于下混频或者包括于混响器算法中。

[0333] 一旦生成了下混频系数,就可以例如通过直接加权组合或者求和来生成下混频信号。

[0334] 所述方法的优点在于可以使用常规的混响器。例如,混响器407可以由反馈延迟网络实现,例如在标准Jot混响器中实现。

[0335] 如图13所示,反馈延迟网络的原理是使用具有不同延迟的一个或多个(典型地多于一个)反馈环路。输入信号(在本例中是下混频信号)被馈送到环路,在该环路中信号以适当的反馈增益被反馈。通过组合环路中的信号来提取输出信号。因此,信号以不同的延迟连续重复。使用互质的延迟和在环路之间混合信号的反馈矩阵可以创建类似于真实空间中混响的模式。

[0336] 反馈矩阵中元素的绝对值必须小于1,以实现稳定的、衰减的脉冲响应。在许多实现方式中,附加增益或者滤波器包括在环路中。这些滤波器可以代替矩阵控制衰减。使用滤波器的好处在于,对于不同的频率,衰减响应可以是不同的。

[0337] 在混响器的输出是双声道渲染的一些实施例中,估计出的混响可以分别由用于左耳和右耳的平均HRTF(头部相关传递函数)滤波,以便产生左声道混响信号和右声道混响信号。当HRTF在用户周围的球体上在以均匀间隔的间隔的多于一个距离处可用时,人们可以理解,使用具有最大距离的HRTF集合来生成用于左耳和右耳的平均HRTF。使用平均HRTF可以基于/反映混响是各向同性的并且来自所有方向的考虑。因此,可以使用所有HRTF的平均值,而不是包括针对给定方向的一对HRTF。可以针对左耳执行一次平均和针对右耳执行一次平均,并且可以使用得到的滤波器来处理用于双声道渲染的混响器的输出。

[0338] 在一些情况下,混响器本身可以引入输入信号的染色,导致不具有如DSR所描述的期望的输出扩散信号能量的输出。因此,该过程的效果也可能是均衡的。该均衡可以基于解析地确定为混响器操作的频率响应的逆的滤波器来执行。在一些实施例中,可以使用诸如线性回归、线拟合之类的机器估计学习技术来估计传递函数。

[0339] 在一些实施例中,相同的方法可以均匀地应用于整个频带。然而,在其他实施例中,可以执行频率相关处理。例如,所提供的元数据参数中的一个或多个可以是频率相关的。在这样的示例中,装置可以被布置为将信号划分成对应于频率依赖性的不同频带,并且可以在频带中的每一个中单独地执行如前所述的处理。

[0340] 具体地,在一些实施例中,扩散混响信号与总信号之比DSR是频率相关的。例如,可以为分立频带/频段的范围提供不同的DSR值,或者可以根据频率提供DSR。在这样的实施例中,装置可以被布置为生成反映DSR的频率相关性的频率相关下混频系数。例如,可以生成用于各个频带的下混频系数。类似地,因此可以生成频率相关下混频和扩散混响信号。

[0341] 对于频率相关DSR,在其他实施例中,下混频系数可以由滤波器来补充,该滤波器

对音频信号进行滤波,作为下混频生成的部分。作为另一示例,DSR效应可以被分离成频率无关(宽带)分量和频率相关分量,该频率无关(宽带)分量用于生成频率无关下混频系数,该系数用于在生成下混频信号时缩放各个音频信号,该频率相关分量可以被应用到下混频,例如通过将频率相关滤波器应用到下混频。在一些实施例中,这样的滤波器可以与进一步的染色滤波器组合,例如作为混响器算法的部分。图7示出了相关(u, v)和染色(h_L, h_R)滤波器的示例。这是专用于双声道输出的反馈延迟网络,称为Jot混响器。

[0342] 因此,在一些实施例中,DSR可以包括频率相关分量部分和非频率相关分量部分,并且下混频系数可以依赖于非频率相关分量部分(并且独立于频率相关部分)来确定。然后,可以基于频率相关分量部分来适配下混频的处理,即,可以依赖于频率相关部分来适配混响器。

[0343] 在一些实施例中,来自音频源中的一个或多个的声音辐射的方向性可以是频率相关的,并且在这样的场景中,可以生成频率相关的总发射能量,当与DSR(其可以是频率相关的或者独立的)组合时,可以得出频率相关下混频系数。

[0344] 这可以例如通过在分立频带中执行单独处理来实现。与针对频率相关DSR的处理相反,方向性的频率相关性典型地必须在下混频信号生成之前(或者作为其部分)执行。这反映了典型地要求频率相关的下混频来包括方向性的频率相关效应,因为这些效应对于不同的源典型地是不同的。在积分之后,可能净效应在频率上有显著的变化,即,针对给定源的总发射能量指示可以具有实质性的频率相关性,而这对于不同的源是不同的。因此,由于不同的源典型地具有不同的方向性模式,因此对于不同的源的总发射能量指示也典型地具有不同的频率相关性。

[0345] 下面将描述可能的方法的具体示例。提供表征空间的扩散声学属性,并且根据方向性、预增益和参考距离元数据确定发射源能量的DSR,这允许计算对应的期望混响能量。例如,这可以确定为:

[0346] $E_{\text{norm}} * \text{DSR}$

[0347] 当用于计算DSR的组件使用相同的参考电平(例如,与信号的全刻度相关)时,当使用如上针对发射源能量计算的 E_{norm} 时,所得到的混响能量也将是针对PCM信号中的全刻度样本归一化的能量,并且因此对应于可以应用于对应输入信号以在所使用的信号表示中提供正确混响电平的扩散混响的脉冲响应(IR)的能量。

[0348] 这些能量值可以用于在混响算法之前确定混响算法、下混频系数或者下混频滤波器的配置参数。

[0349] 存在不同的方式来生成混响。基于反馈延迟网络(FDN)的算法(例如,Jot混响器)是合适的低复杂度方法。可替代地,噪声序列可以被成形为具有适当的(频率相关的)衰减和频谱形状。在这两个示例中,原型IR(具有至少适当的T60)可以被调整,使得其(频率相关的)电平被校正。

[0350] 可以调整混响器算法,使得它们产生具有单位能量的脉冲响应(或者DSR的单位初始幅度可以与初始幅度相关),或者混响器算法可以包括其自身的补偿,例如在Jot混响器的染色滤波器中。可替代地,可以通过(潜在地频率相关的)调整来修改下混频,或者可以修改由系数处理器507产生的下混频系数。

[0351] 补偿可以通过在没有任何这样的调整,但是应用了所有其他配置(例如适当的混

响时间 (T60) 和反射密度 (例如, FDN 中的延迟值)) 的情况下生成脉冲响应并且测量该 IR 的能量来确定。

$$[0352] \quad E_{protRev} = \sum_{n=0}^{N_{max}} h_{rev}(n)$$

[0353] 补偿可以是该能量的逆。为了包括在下混频系数中, 典型地应用平方根。例如:

$$[0354] \quad d_x = \sqrt{\frac{E_{norm,x} * DSR}{E_{protRev}}}$$

[0355] 在许多其他实施例中, 补偿可以从配置参数得出。例如, 当 DSR 相对于初始混响幅度时, 可以从其配置导出第一反射。根据定义, 相关滤波器是能量保持的, 并且彩色滤波器也可以设计成如此。

[0356] 假设染色滤波器没有净增强或者衰减, 混响器可以例如导致依赖于 T60 和最小延迟值 minDelay 的初始幅度 (A_0):

$$[0357] \quad A_0 = 10^{\frac{-3 * \text{minDelay}}{T60 * fs}} * \sqrt{T60}$$

[0358] 预测混响能量还可以启发式地进行。

[0359] 作为扩散混响能量的一般模型, 可以考虑指数函数 $A(t)$:

$$[0360] \quad \alpha = \frac{\ln\left(10^{-\frac{60}{T60}}\right)}{-T60}$$

$$[0361] \quad A(t) = A_0 * e^{-\alpha(t-t_3)}$$

[0362] 对于 $t \geq t_3 = \text{预延迟}$ 。衰减因子 α 是由 T60 控制的延迟因子, 并且 A_0 是预延迟的幅度。

[0363] 计算这样的函数的累积能量, 它会逐渐接近某个最终能量值。最终能量值与 T60 几乎完全成线性关系。

[0364] 线性关系的因子依赖于函数 A 的稀疏性 (每第 2 个值设为 0 约等于能量的一半)、初始值 A_0 (能量与 A_0^2 线性缩放) 和采样率 (随 fs 的改变线性缩放)。使用 T60、反射密度 (由 FDN 延迟得出) 和采样率, 可以用这样的函数可靠地建模扩散尾。针对模型的 A_0 可以如上所示计算, 以与 FDN 的 A_0 相等。

[0365] 当生成宽带 T60 值在 0.1-2s 范围内的多个参数混响时, IR 的能量与模型接近线性。实际能量与指数方程模型平均值之间的缩放因子由 FDN 响应的稀疏性确定。这种稀疏性在 IR 结束时变得较少, 但是在开始时影响最大。通过对上面多个延迟值的配置的测试发现, 模型缩减因子与 FDN 中配置的延迟之间的最小差之间存在近似线性关系。

[0366] 例如, 对于 Jot 混响器的特定实现方式, 这可以大约是缩放因子 SF, 其通过以下计算:

$$[0367] \quad SF = 7.0208 * \text{MinDelayDiff} + 214.1928$$

[0368] 模型的能量由 $t=0$ 到无穷远积分计算。这可以解析地进行并且得出:

$$[0369] \quad \text{revModelEnergy} = \frac{A_0^2 * fs}{2 * \alpha}$$

[0370] 组合以上, 我们得到了以下混响能量的预测。

$$[0371] \quad revPredEnergy = \frac{A_0^2 * fs}{2 * alpha} * \frac{1}{0.8776 * MinDelayDiff + 26.7741}$$

[0372] 可以理解,为了清楚起见,上述描述参照不同的功能电路、单元和处理器描述了本发明的实施例。然而,显而易见的是,在不损害本发明的情况下,可以使用不同功能电路、单元或处理器之间的任何适当的功能分配。例如,示出由单独的处理器或者控制器执行的功能可以由相同的处理器或者控制器执行。因此,对具体功能单元或者电路的引用仅被视为对用于提供所述功能的适当单元的引用,而不是对严格的逻辑或者物理结构或者组织的指示。

[0373] 本发明可以以任何合适的形式实现,包括硬件、软件、固件或者这些的任何组合。本发明可任选地至少部分地实现为运行在一个或多个数据处理器和/或数字信号处理器上的计算机软件。本发明实施例的元件和组件可以以任何合适的方式在物理上、功能上和逻辑上实现。事实上,功能可以在单个单元中、多个单元中或者作为其他功能单元的部分来实现。因此,本发明可以在单个单元中实现,或者可以在物理上和功能上分布在不同的单元、电路和处理器之间。

[0374] 虽然已经结合一些实施例描述了本发明,但是本发明并不旨在局限于本文阐述的具体形式。而是,本发明的范围仅受所附权利要求书的限制。另外地,尽管特征看起来可以结合特定实施例来描述,但本领域技术人员将认识到根据本发明可以组合所描述的实施例的各种特征。在权利要求中,术语“包括”并不排除其他元件或者步骤的存在。

[0375] 此外,虽然被单独列出,但是多个单元、元件、电路或者方法步骤可以由例如单个电路、单元或者处理器来实现。另外地,尽管各个特征可以包括于不同的权利要求中,但是这些特征可以有利地组合,并且包括于不同的权利要求中并不意味着特征的组合是不可行和/或有利的。而且,在一个类别的权利要求中包括特征并不意味着对这一类别的限制,而是指示该特征在适当时等同地适用于其他权利要求类别。此外,权利要求中的特征的次序并不意味着特征必须以任何具体的次序工作,特别是方法权利要求中的单个步骤的次序并不意味着必须以这种次序执行步骤。而是,这些步骤可以以任何适当的次序执行。另外,单数引用并不排除复数。因此,对“一(a)”、“一个(an)”、“第一”、“第二”等的引用并不排除复数。在权利要求中提供的附图标记仅仅是作为澄清的示例,不应被解释为以任何方式限制权利要求的范围。

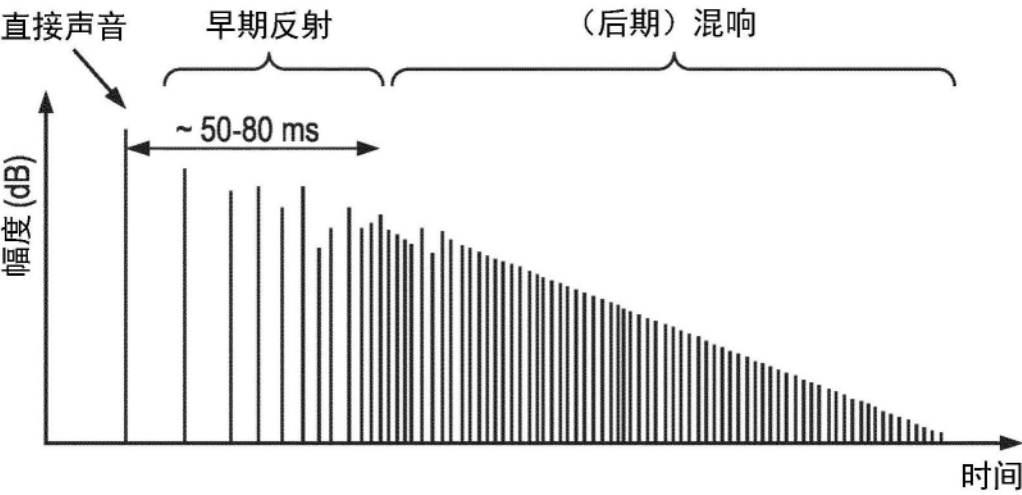


图1

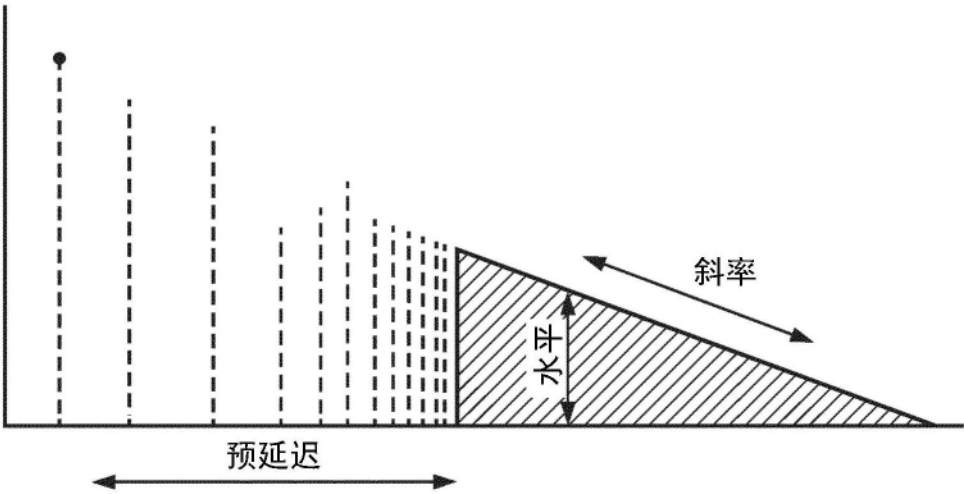


图2

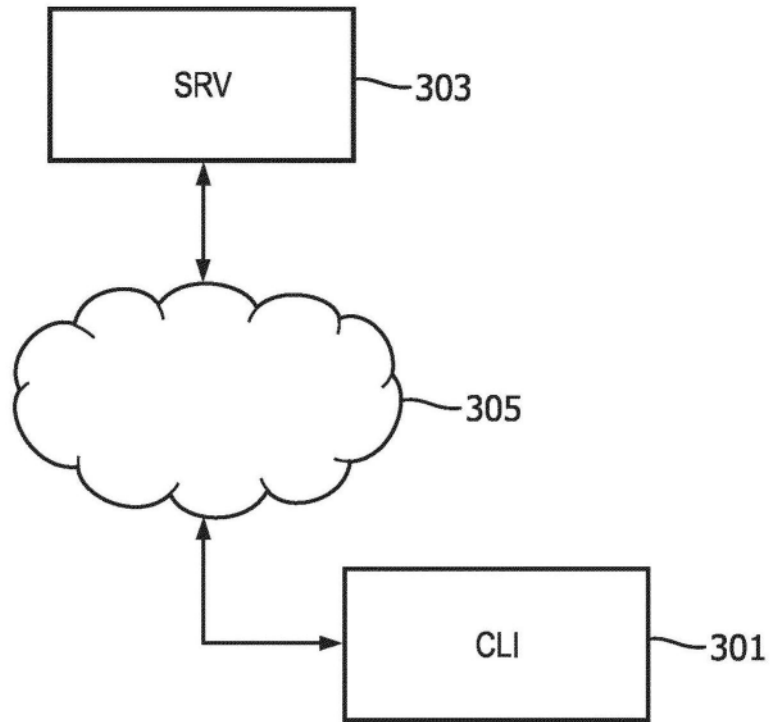


图3

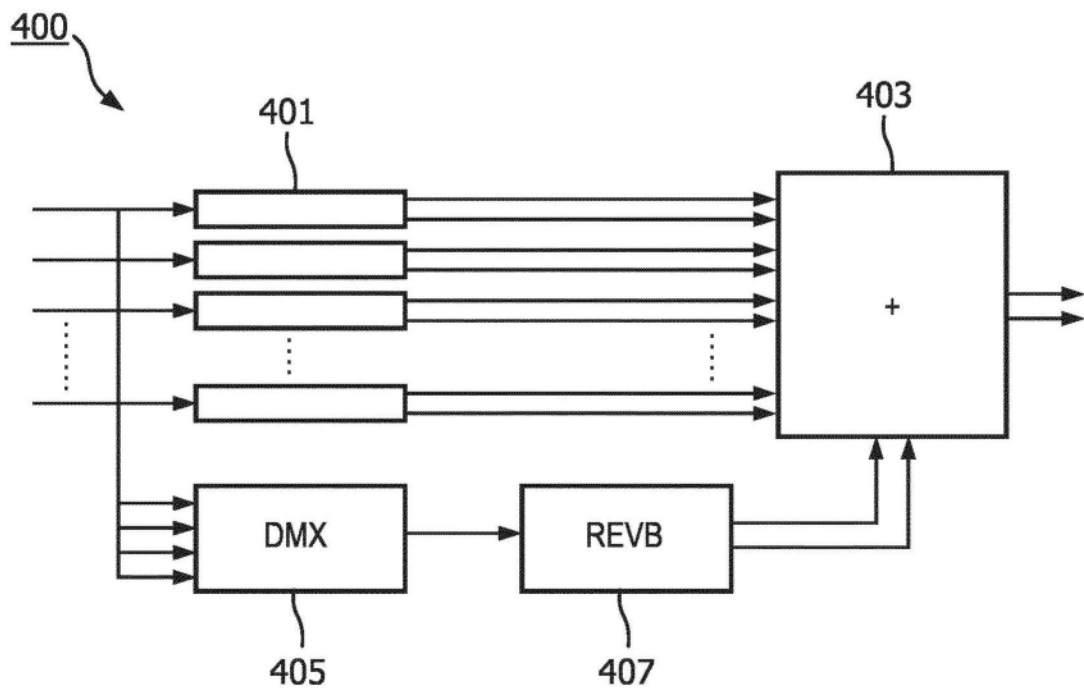


图4

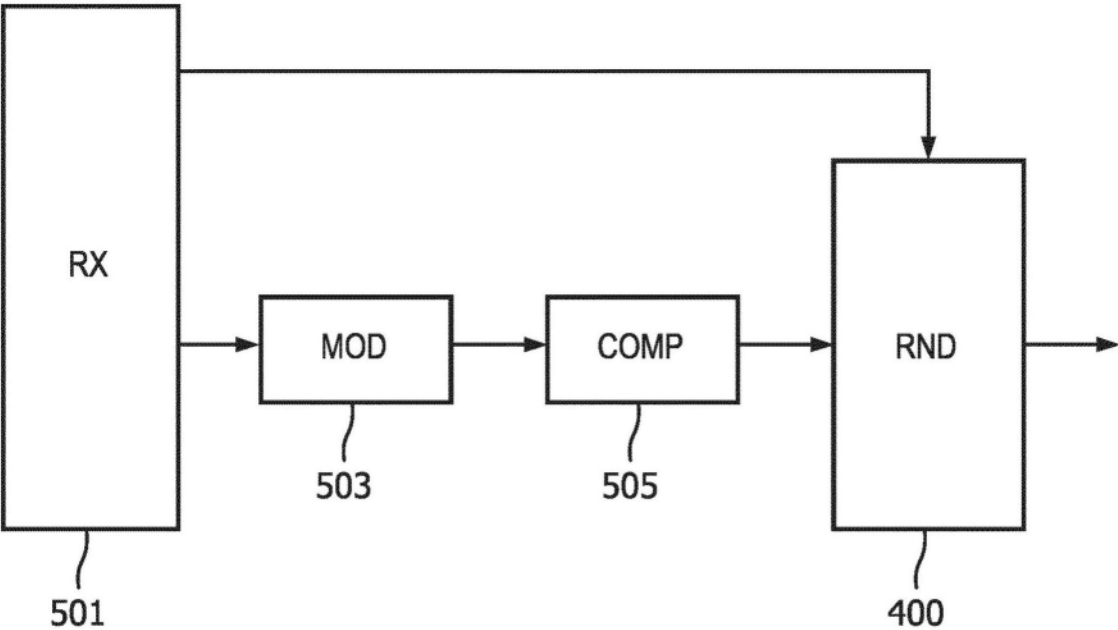


图5

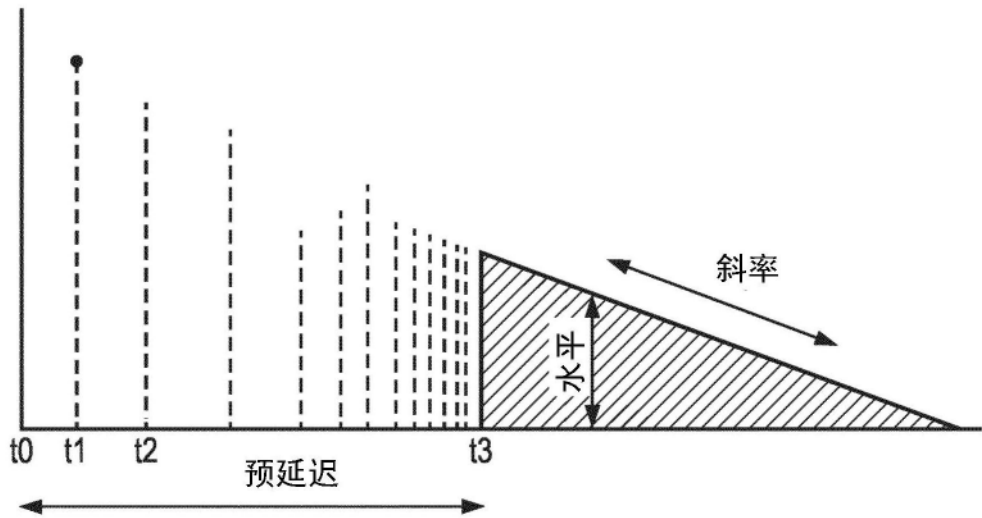


图6

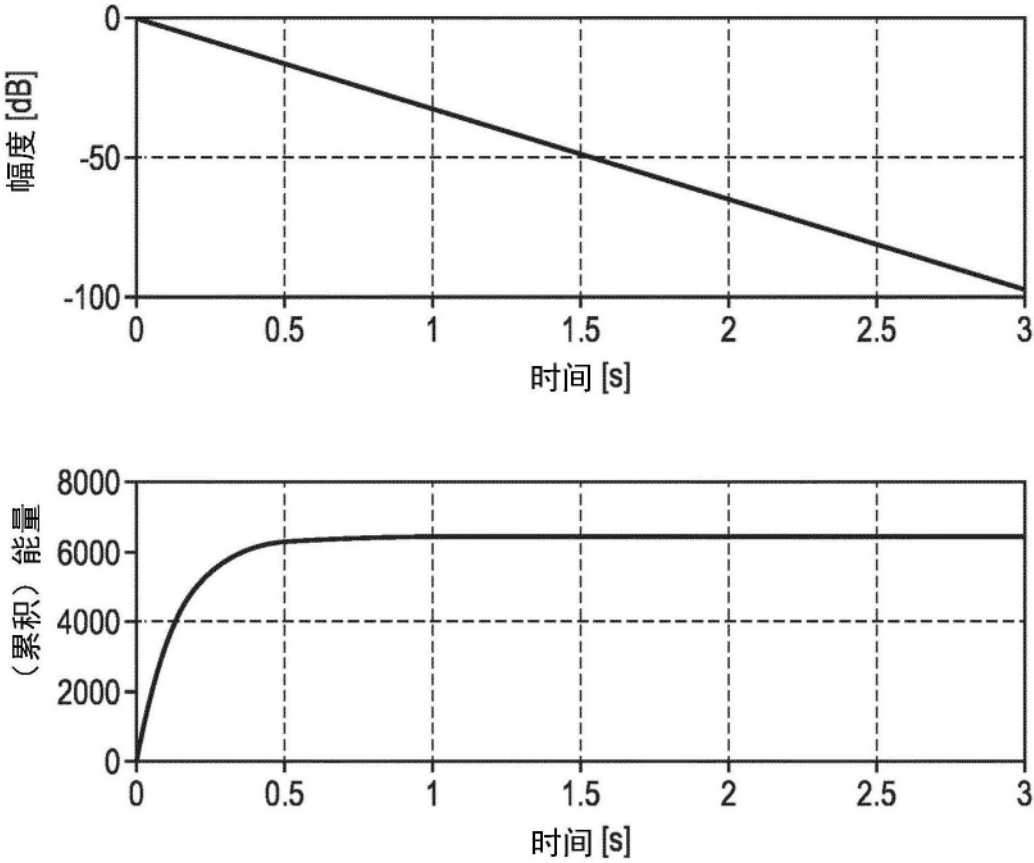


图7

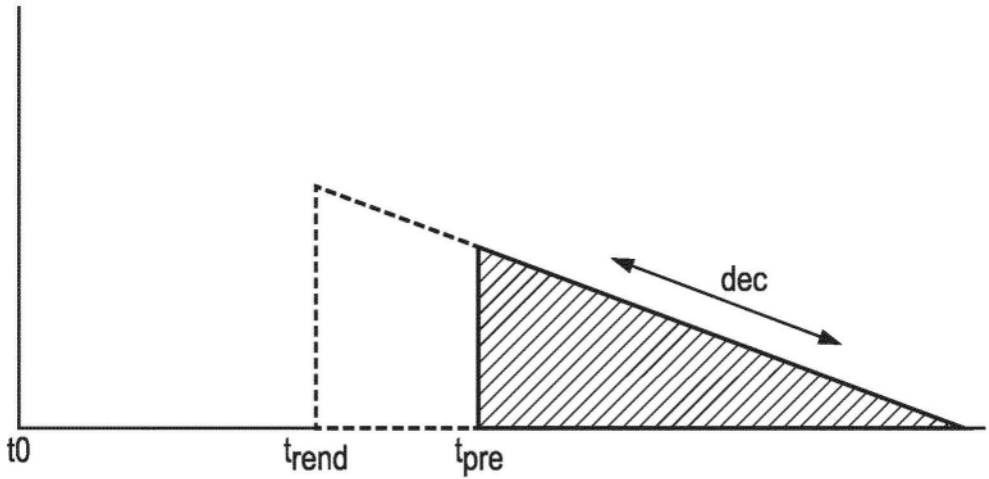


图8

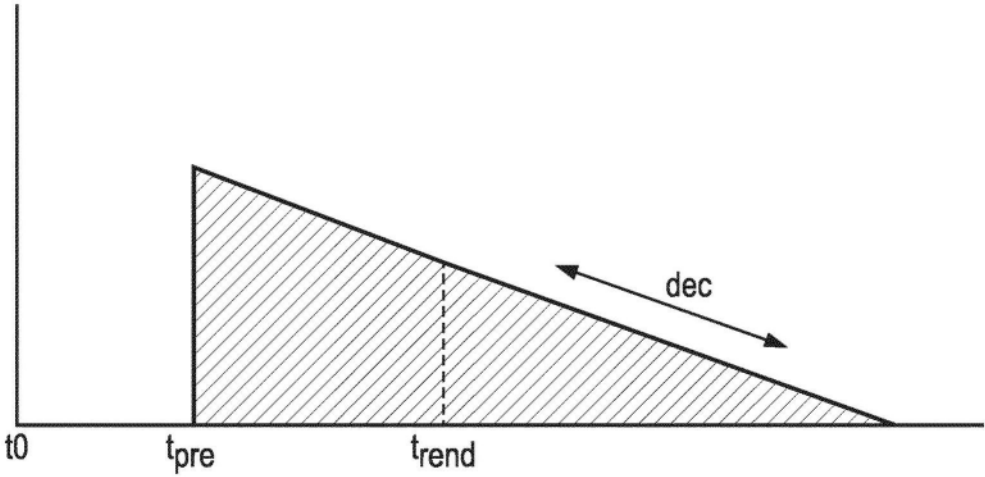


图9

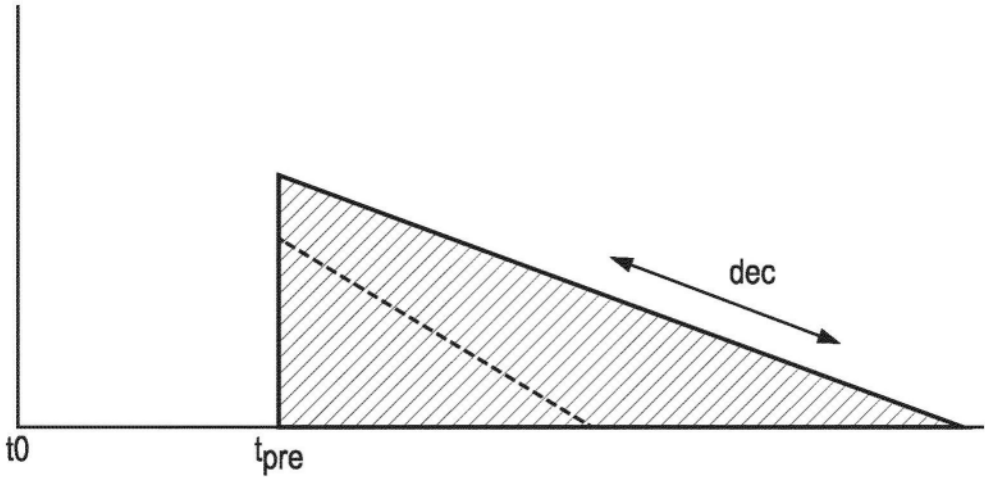


图10

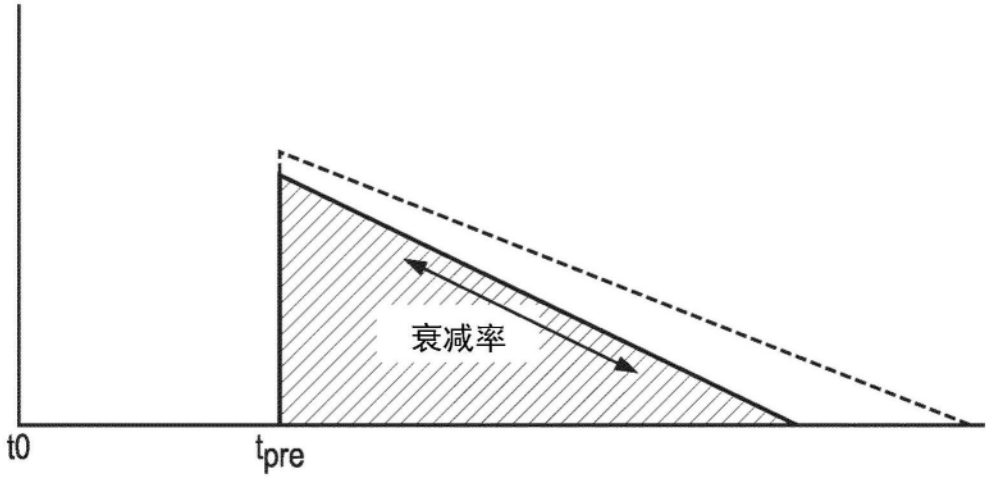


图11

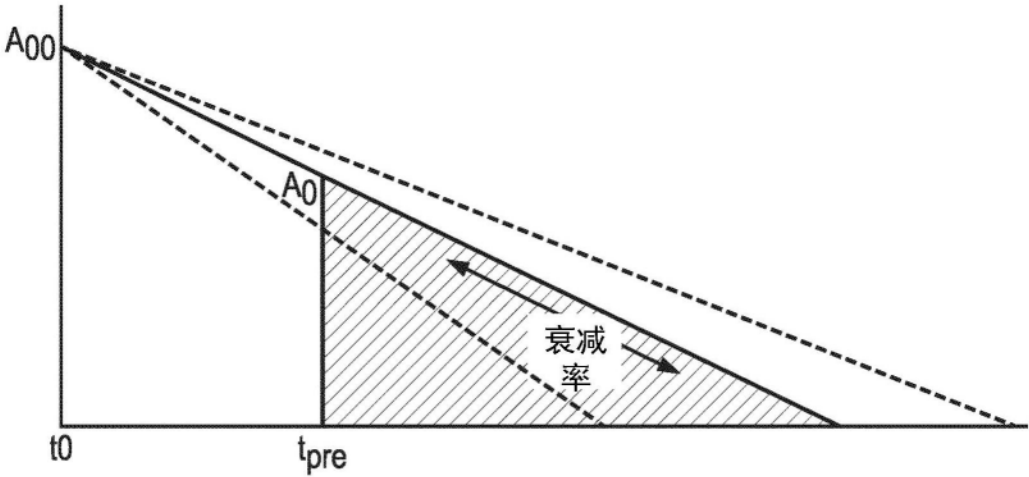


图12

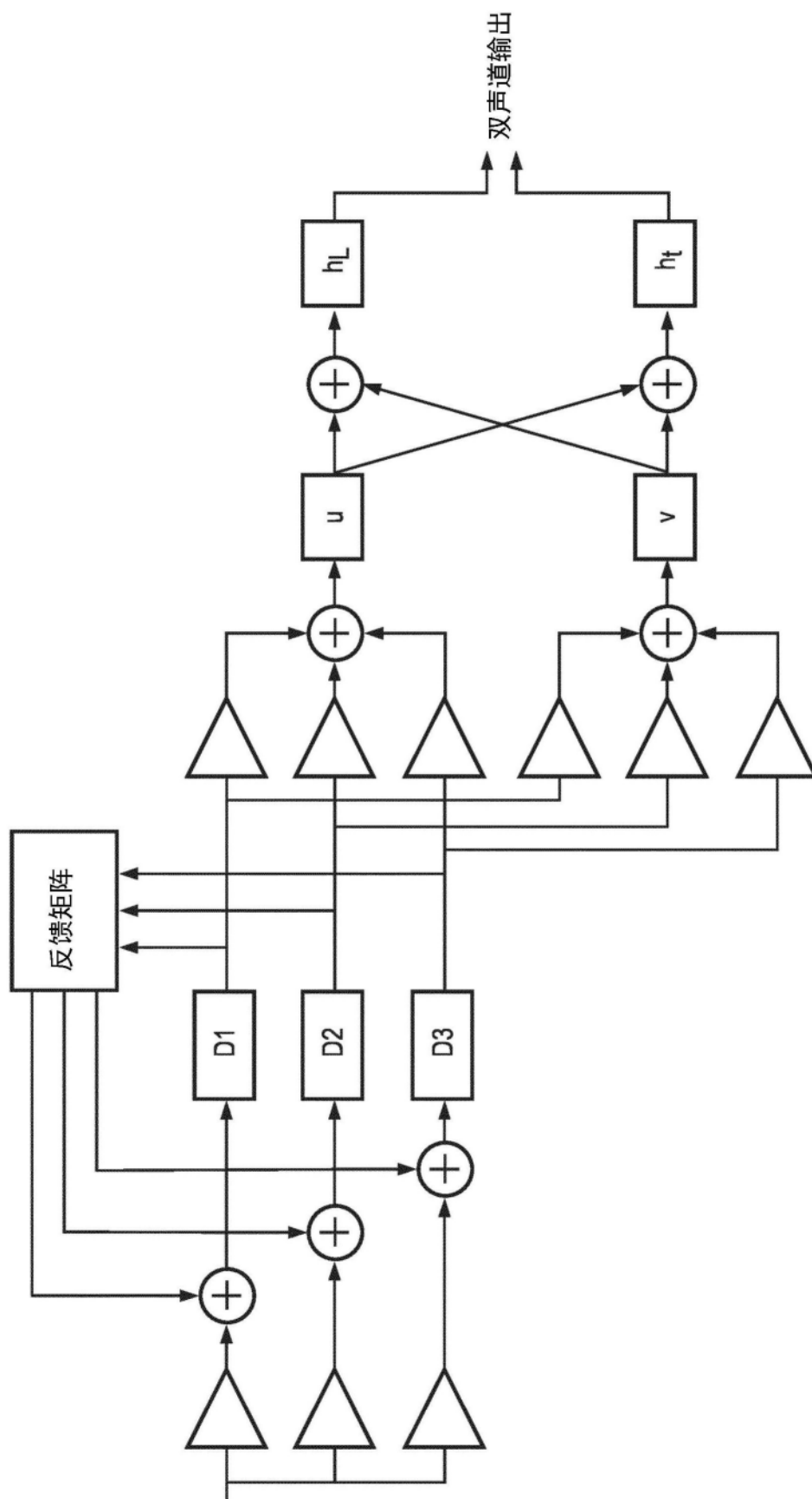


图13

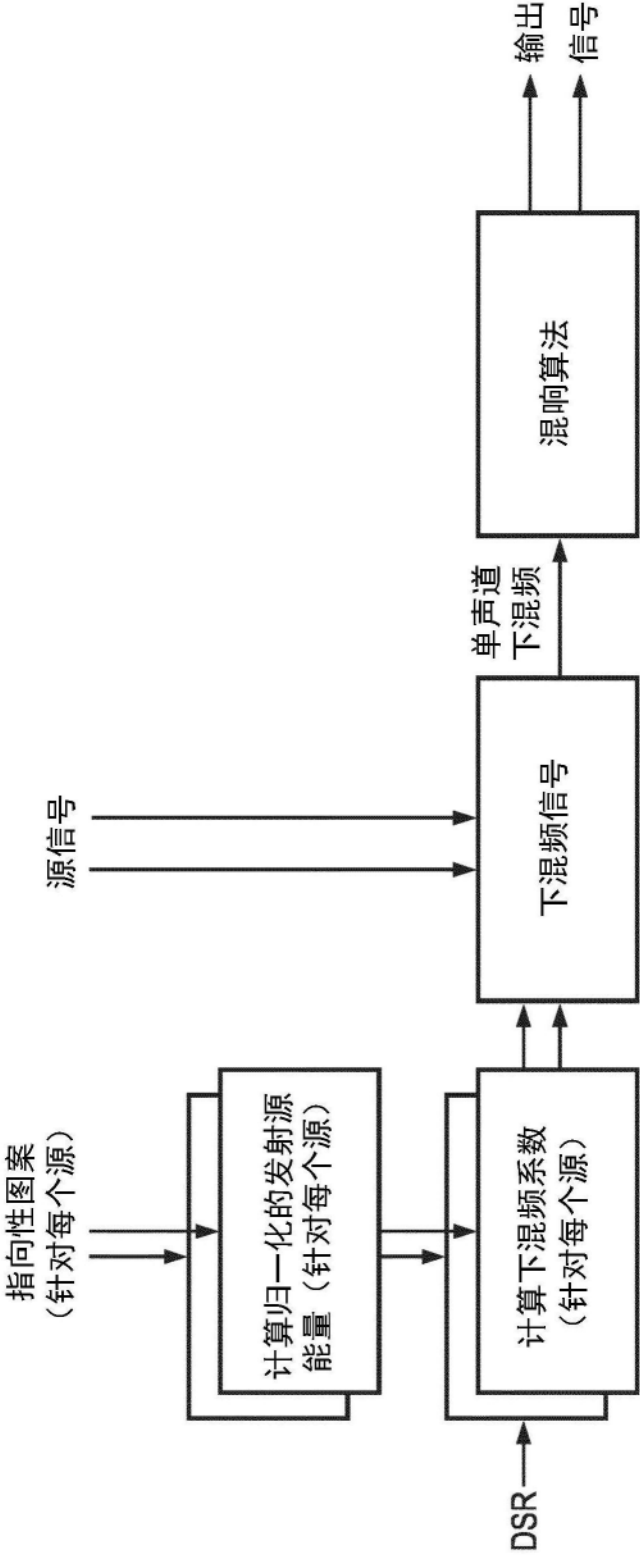


图14