



(12)发明专利

(10)授权公告号 CN 107770718 B

(45)授权公告日 2020.01.17

(21)申请号 201711094044.5

(22)申请日 2014.12.18

(65)同一申请的已公布的文献号
申请公布号 CN 107770718 A

(43)申请公布日 2018.03.06

(30)优先权数据
61/923,579 2014.01.03 US
61/988,617 2014.05.05 US

(66)本国优先权数据
201410178258.0 2014.04.29 CN

(62)分案原申请数据
201480071993.X 2014.12.18

(73)专利权人 杜比实验室特许公司
地址 美国加利福尼亚

(72)发明人 颜冠杰 D·J·布里巴特

G·A·戴维森 R·威尔森

D·M·库珀 双志伟

(74)专利代理机构 中国国际贸易促进委员会专
利商标事务所 11038

代理人 宿小猛

(51)Int.Cl.
H04S 7/00(2006.01)
G10L 19/008(2013.01)
H04S 3/00(2006.01)

(56)对比文件
US 2011135098 A1,2011.06.09,
CN 103355001 A,2013.10.16,
CN 101843114 A,2010.09.22,
CN 103329576 A,2013.09.25,

审查员 陈嵘

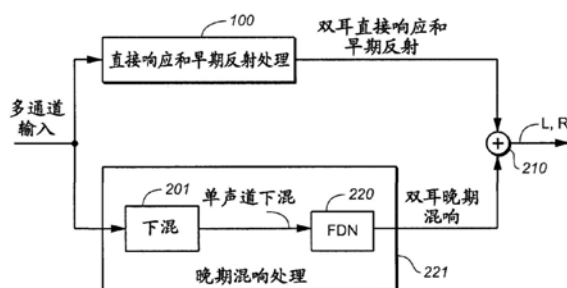
权利要求书2页 说明书27页 附图12页

(54)发明名称

响应于多通道音频通过使用至少一个反馈
延迟网络产生双耳音频

(57)摘要

本公开涉及响应于多通道音频通过使用至少一个反馈延迟网络产生双耳音频。在一些实施例中,提供用于响应多通道音频信号的通道产生双耳信号的虚拟化方法,这些虚拟化方法向各通道应用双耳房间脉冲响应(BRIR),包含通过使用至少一个反馈延迟网络(FDN)以向通道的下混应用公共晚期混响。在一些实施例中,输入信号通道在第一处理路径中被处理,以向各通道应用的用于该通道的单通道BRIR的直接响应和早期反射部分,并且,通道的下混在包含至少一个应用公共晚期混响的FDN的第二处理路径中被处理。典型地,公共晚期混响模仿单通道BRIR中的至少一些晚期混响部分的共同宏观属性。其它方面是被配置为执行方法的任何实施例的耳机虚拟化器。



1. 一种用于响应多通道音频输入信号的一组通道产生双耳信号的方法,包括:
向该组通道中的各通道应用双耳房间脉冲响应BRIR以由此产生经滤波的信号;和
组合经滤波的信号以产生双耳信号,

其中,向该组通道中的各通道应用BRIR包含通过使用晚期混响产生器(200)来响应于被断言至所述晚期混响产生器(200)的控制值向该组通道中的通道的下混应用公共晚期混响,其中,所述公共晚期混响模仿在该组通道中的至少一些通道上共享的单通道BRIR的晚期混响部分的共同宏观属性,以及

其中,所述下混是该组通道中的通道的立体声下混。

2. 根据权利要求1所述的方法,其中,向该组通道中的各通道应用BRIR包含向该组通道中的各通道应用该通道的单通道BRIR的直接响应和早期反射部分。

3. 根据权利要求1或2所述的方法,其中,所述晚期混响产生器(200)包括用于向该下混应用公共晚期混响的反馈延迟网络的群(203,204,205),其中该群中的各反馈延迟网络(203,204,205)向该下混的不同频带应用晚期混响。

4. 根据权利要求3所述的方法,其中,反馈延迟网络(203,204,205)中的每一个在滤波器组域中实现。

5. 根据权利要求1或2所述的方法,其中,所述晚期混响产生器(200)包括用于将公共晚期混响应用于该组通道中的通道的下混的单个反馈延迟网络(220),其中该反馈延迟网络(220)在时域中实现。

6. 根据权利要求1或2所述的方法,其中,所述共同宏观属性包括平均功率谱、能量衰变结构、模态密度和峰密度中的一个或多个。

7. 根据权利要求1或2所述的方法,其中,所述控制值中的一个或多个是频率相关的,和/或所述控制值之一是混响时间。

8. 一种用于响应多通道音频输入信号的一组通道产生双耳信号的系统,所述系统包括一个或多个处理器,用于:

向该组通道中的各通道应用双耳房间脉冲响应BRIR以由此产生经滤波的信号;和
组合经滤波的信号以产生双耳信号,

其中,向该组通道中的各通道应用BRIR包含通过使用晚期混响产生器(200)来响应于被断言至所述晚期混响产生器(200)的控制值向该组通道中的通道的下混应用公共晚期混响,其中,所述公共晚期混响模仿在该组通道中的至少一些通道上共享的单通道BRIR的晚期混响部分的共同宏观属性,以及

其中,该组通道中的通道的所述下混是该组通道中的通道的立体声下混。

9. 根据权利要求8所述的系统,其中,向该组通道中的各通道应用BRIR包含向该组通道中的各通道应用该通道的单通道BRIR的直接响应和早期反射部分。

10. 根据权利要求8或9所述的系统,其中,所述晚期混响产生器(200)包括被配置为向该下混应用公共晚期混响的反馈延迟网络的群(203,204,205),其中该群中的各反馈延迟网络(203,204,205)向该下混的不同频带应用晚期混响。

11. 根据权利要求10所述的系统,其中,反馈延迟网络(203,204,205)中的每一个在滤波器组域中实现。

12. 根据权利要求8或9所述的系统,其中,所述晚期混响产生器(200)包括在时域中实

现的反馈延迟网络(220),并且所述晚期混响产生器(200)被配置为在所述反馈延迟网络(220)中在时域中处理该下混以将公共晚期混响应用于所述下混。

13.根据权利要求8或9所述的系统,其中,所述共同宏观属性包括平均功率谱、能量衰变结构、模态密度和峰密度中的一个或多个。

14.根据权利要求8或9所述的系统,其中,所述控制值中的一个或多个是频率相关的,和/或所述控制值之一是混响时间。

15.一种用于响应多通道音频输入信号的一组通道产生双耳信号的设备,包括:
一个或多个处理器;以及

存储指令的一个或多个存储介质,所述指令在由所述一个或多个处理器执行时使得执行根据权利要求1-7中任一项所述的方法。

16.一种计算机可读存储介质,存储指令,所述指令在由一个或多个处理器执行时使得执行根据权利要求1-7中任一项所述的方法。

17.一种包括用于执行根据权利要求1-7中任一项所述的方法的部件的装置。

响应于多通道音频通过使用至少一个反馈延迟网络产生双耳音频

[0001] 本申请是申请号为201480071993.X、申请日为2014年12月18日、发明名称为“响应于多通道音频通过使用至少一个反馈延迟网络产生双耳音频”的发明专利申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求2014年4月29日提交的中国专利申请No.201410178258.0;2014年1月3日提交的美国临时申请No.61/923579;以及2014年5月5日提交的美国临时专利申请No.61/988617的优先权,这些申请中的每一个的全部内容通过引用并入这里。

技术领域

[0004] 本发明涉及用于如下这样的方法(有时称为耳机虚拟化方法)和系统,其响应于多通道输入信号通过对于音频输入信号的一组通道中的每一个通道(例如,对于所有通道)应用双耳房间脉冲响应(BRIR)而产生双耳信号。在一些实施例中,至少一个反馈延迟网络(FDN)向通道的下混应用下混BRIR的晚期混响部分。

背景技术

[0005] 耳机虚拟化(或双耳呈现)是一种旨在通过使用标准立体声耳机传输环绕声体验或身临其境的声场的技术。

[0006] 早期耳机虚拟化器在双耳呈现中应用头部相关传递函数(HRTF)以传送空间信息。HRTF是表征在无回声的环境中声音如何从空间中的特定点(声源位置)发送到收听者的两耳的一组方向和距离相关滤波器对。可在呈现的经HRTF滤波的双耳内容中感知诸如耳间时间差(ITD)、耳间水平差(ILD)、头部遮蔽效果、由于肩部和耳廓反射导致的谱峰和谱凹口的必要空间线索(cue)。由于人头部大小的约束,HRTF不提供足够的或鲁棒的关于超出大致1米的源距离的线索。作为结果,仅基于HRTF的虚拟化器通常不能实现良好的外在化(externalization)或感知距离。

[0007] 我们日常生活中的大多数的声音事件发生在混响环境中,在该环境中,除了通过HRTF被模型化的直接路径(从源到耳朵)以外,音频信号也通过各种反射路径到达收听者的耳朵。反射引入了对诸如距离、房间大小和空间的其它属性的听知觉深刻影响。为了在双耳呈现中传送该信息,除了直接路径HRTF中的线索以外,虚拟化器需要应用房间混响。双耳房间脉冲响应(BRIR)表征在特定声学环境中从空间中的特定点到收听者的耳朵的音频信号的变换。理论上,BRIR包含关于空间感知的所有声音线索。

[0008] 图1是被配置为向多通道音频输入信号的全频率范围通道(X_1 、 \dots 、 X_N)应用双耳房间脉冲响应(BRIR)的一种类型的常规耳机虚拟化器的框图。通道 X_1 、 \dots 、 X_N 中的每一个是与相对于假定的收听者的不同源方向(即,从相应的扬声器的假定位置到假定的收听者位置的直接路径的方向)对应的扬声器通道,并且,每个这种通道与用于相应的源方向的BRIR卷积。需要对于每个耳朵模拟来自各通道的声音路径。因此,在本文件的剩余部分中,术语

BRIR将指的是一个脉冲响应或者与左耳和右耳相关联的一对脉冲响应。因此,子系统2被配置为将通道 X_1 与 $BRIR_1$ (用于相应的源方向的BRIR) 卷积,子系统4被配置为将通道 X_N 与 $BRIR_N$ (用于相应的源方向的BRIR) 卷积,等等。各BRIR子系统(子系统2、 \dots 、4中的每一个)的输出是包含左通道和右通道的时域信号。BRIR子系统的左通道输出在加算元件6中被混合,并且BRIR子系统的右通道输出在加算元件8中被混合。元件6的输出是从虚拟化器输出的双耳音频信号的左通道L,元件8的输出是从虚拟化器输出的双耳音频信号的右通道R。

[0009] 多通道音频输入信号还可包含在图1中被标识为“LFE”通道的低频效果(LFE)或低音炮通道。以常规的方式,LFE通道不与BRIR卷积,而作为替代,在图1的增益级5中衰减(例如,衰减-3dB或更多),并且增益级5的输出(通过元件6和8)均等地混合到虚拟化器的双耳输出信号的各通道中。为了使级5的输出与BRIR子系统(子系统2、 \dots 、4)的输出时间对准,在LFE路径中可能需要附加的延迟级。作为替代方案,LFE通道可简单地被忽略(即,不通过虚拟化器被断言(assert)或者被处理)。例如,本发明的图2实施例(后面将描述)简单地忽略由此处理的多通道音频输入信号的任何LFE通道。许多消费者耳机不能精确地再现LFE通道。

[0010] 在一些常规的虚拟化器中,输入信号经受到变换到QMF(正交镜像滤波器)域中的时域到频域变换,以产生QMF域频率成分的通道。这些频率成分在QMF域中经受滤波(例如,在图1的子系统2、 \dots 、4的QMF域实现中),并且,得到的频率成分典型地然后变换回时域(例如,在图1的子系统2、 \dots 、4中的每一个的最后级中),使得虚拟化器的音频输出是时域信号(例如,时域双耳信号)。

[0011] 一般地,输入到耳机虚拟化器的多通道音频信号的各全频率范围通道被假定为指示从在相对于收听者的耳朵的已知位置处的声音源发射的音频内容。耳机虚拟化器被配置为向输入信号的每个这种通道应用双耳房间脉冲响应(BRIR)。各BRIR可分解成两个部分:直接响应和反射。直接响应是与声音源的到达方向(DOA)对应的、由于(声音源与收听者之间的)距离而以适当的增益和延迟被调整的并且可选地对于小距离随视差效果而增扩的HRTF。

[0012] BRIR的剩余部分模型化反射。早期反射通常是一次和二次反射,并且具有相对稀疏的时间分布。各一次或二次反射的微结构(例如,ITD和ILD)是重要的。对于稍晚反射(在入射到收听者之前从多于两个的表面反射的声音),回声密度随反射次数增加而增加,并且,各单次反射的微观属性变得难以观察。对于越来越晚的反射,宏观结构(例如,整个混响的空间分布、耳间相干性和混响延迟率)变得更重要。因此,反射可进一步分成两个部分:早期反射(early reflection)和晚期混响(late reverberation)。

[0013] 直接响应的延迟是距收听者的源距离除以声音的速度,并且其水平(在没有接近源位置的大的表面或墙壁的情况下)与源距离成反比。另一方面,晚期混响的延迟和水平一般对源位置不敏感。由于实际的考虑,虚拟化器可选择时间对准来自具有不同的距离的源的直接响应,并且/或者压缩它们动态范围。但是,BRIR内的直接响应、早期反射和晚期混响之间的时间和水平关系应被保持。

[0014] 典型的BRIR的有效长度在大多数的声学环境中延长到几百毫秒或更长。BRIR的直接应用需要与具有数以千计的抽头(tap)的滤波器卷积,这在计算上是昂贵的。另外,在没有参数化的情况下,为了实现足够的空间分辨率,将需要大的存储器空间以存储用于不同

的源位置的BRIR。最后的但同样重要的,声音源位置可随时间改变,并且/或者,收听者的位置和取向可随时间改变。这种移动的精确仿真需要时变BRIR脉冲响应。如果这样的时变滤波器的脉冲响应具有许多抽头,那么这种时变滤波器的适当的内插和应用可能是挑战性的。

[0015] 具有称为反馈延迟网络(FDN)的公知的滤波器结构的滤波器可被用于实现空间混响器,该空间混响器被配置为对于多通道音频输入信号的一个或更多个通道应用仿真混响。FDN的结构是简单的。它包含数个混响箱(例如,在图4中FDN中,包含增益元件 g_1 和延迟线 z^{-n_1} 的混响箱),每个混响箱具有延迟和增益。在FDN的典型实现中,来自所有混响箱的输出通过单一反馈矩阵被混合,并且矩阵的输出被反馈到混响箱的输入并与其求和。可对混响箱输出进行增益调整,并且,对于多通道或双耳回放可适当地重新混合混响箱输出(或它们的增益调整版本)。可通过具有紧凑的计算和存储器印迹的FDN产生和应用自然发声(sounding)混响。因此,FDN已被用于虚拟化器中以补充通过HRTF产生的直接响应。

[0016] 例如,市售的Dolby Mobile耳机虚拟化器包含具有基于FDN的结构混响器,该混响器可操作为对于五通道音频信号(具有左前、右前、中心、左环绕和右环绕通道)的各通道应用混响,并通过使用一组五个头部相关传递函数(“HRTF”)滤波器对的不同的滤波器对来对各混响通道进行滤波。Dolby Mobile耳机虚拟化器也可响应二通道音频输入信号而操作,以产生二通道“经混响的”双耳音频输出(已被应用了混响的二通道虚拟环绕声输出)。当经混响的双耳输出通过一对耳机被呈现和再现时,在收听者的耳膜处感知为来自位于左前、右前、中心、左后(环绕)和右后(环绕)位置的五个扬声器的经HRTF滤波的混响声音。虚拟化器上混经下混的二通道音频输入(没有使用与音频输入一起接收的任何空间线索参数)以产生五个上混音频通道,对于经上混的通道应用混响,并且下混五个经混响的通道信号以产生虚拟化器的二通道混响输出。在不同的HRTF滤波器对中对用于各上混通道的混响进行滤波。

[0017] 在虚拟化器中,FDN可被配置为实现一定的混响衰变时间(reverb decay time)和回声密度。但是,FDN缺少仿真早期反射的微观结构的灵活性。并且,在常规的虚拟化器中,FDN的调谐和配置主要是启发式的。

[0018] 不仿真所有反射路径(早期和晚期)的耳机虚拟化器不能实现有效的外在化。发明人认识到,使用试图仿真所有反射路径(早期和晚期)的FDN的虚拟化器在仿真早期反射和晚期混响两者并将两者应用于音频信号时通常只获得有限的成功。发明人还认识到,使用FDN但不具有适当地控制诸如混响衰变时间、耳间相干性和直接与晚期比的空间声学属性的能力的虚拟化器可实现某种程度的外在化,但代价是引入过量的音色失真和混响。

发明内容

[0019] 在第一类的实施例中,本发明是一种响应多通道音频输入信号的一组通道(例如,通道中的每一个或者全频率范围通道中的每一个)产生双耳信号的方法,包括以下的步骤:
(a) 对于该组通道中的每一通道应用双耳房间脉冲响应(BRIR)(例如,通过将该组通道中的每一通道和与所述通道对应的BRIR卷积),由此产生经滤波的信号(包含通过使用至少一个反馈延迟网络(FDN)以向该组通道中的通道的下混(例如,单音下混(monophonic downmix))应用公共晚期混响(common late reverberation));和(b)组合经滤波的信号以

产生双耳信号。典型地,FDN的群被用于向该下混应用公共晚期混响(例如,使得各FDN向不同的频带应用公共晚期混响)。典型地,步骤(a)包含向该组通道中的每一通道应用用于该通道的单通道BRIR的“直接响应和早期反射”部分的步骤,并且,公共晚期混响被产生以模仿单通道BRIR中的至少一些(例如,全部)的晚期混响部分的共同宏观属性(collective marco attribute)。

[0020] 用于响应多通道音频输入信号(或响应这种信号的一组通道)产生双耳信号的方法有时在这里被称为“耳机虚拟化”方法,并且,被配置为执行这种方法的系统有时在这里被称为“耳机虚拟化器”(或“耳机虚拟化系统”或“双耳虚拟化器”)。

[0021] 在第一类的典型的实施例中,在滤波器组域(例如,混合复正交镜像滤波器(HCQMF)域或正交镜像滤波器(QMF)域或可包含抽取(decimation)的另一变换或子带域)中实现FDN中的每一个,并且,在一些这种实施例中,通过控制用于应用晚期混响的各FDN的配置,控制双耳信号的频率相关空间声学属性。典型地,为了实现多通道信号的音频内容的高效的双耳呈现,通道的单音下混被用作FDN的输入。第一类的典型的实施例包括例如通过对反馈延迟网络断言控制值以设定所述反馈延迟网络的输入增益、混响箱(reverb tank)增益、混响箱延迟或输出矩阵参数中的至少一个来调整与频率相关属性(例如,混响衰变时间、耳间相干性、模态密度和直接与晚期比(direct-to-late ratio))对应的FDN系数的步骤。这使得能够实现声学环境的更好的匹配和更自然的发声输出。

[0022] 在第二类的实施例中,本发明是一种响应具有通道的多通道音频输入信号通过向输入信号的一组通道中的各通道(例如,输入信号的通道中的每一个或输入信号的全频率范围通道)应用双耳房间脉冲响应(BRIR)以产生双耳信号的方法,包括通过:在第一处理路径中处理该组通道中的各通道,该第一处理路径被配置为模型化并向所述各通道应用用于该通道的单通道BRIR的直接响应和早期反射部分;以及在第二处理路径(与第一处理路径并联)中处理该组通道中的通道的下混(例如,单音(单声道)下混),该第二处理路径被配置为模型化并向该下混应用公共晚期混响。典型地,公共晚期混响被产生以模仿单通道BRIR中的至少一些(例如,全部)的晚期混响部分的共同宏观属性。典型地,第二处理路径包含至少一个FDN(例如,对于多个频带的每一个有一个FDN)。典型地,单声道下混被用作由第二处理路径实现的各FDN的所有混响箱的输入。典型地,为了更好地模拟声学环境并产生更自然的发声双耳虚拟化,设置用于各FDN的宏观属性的系统控制的机构。由于大多数这种宏观属性是依赖于频率的,因此,典型地在混合复正交镜像滤波器(HCQMF)域、频域、域或另一滤波器组域中实现各FDN,并且,对于各频带使用不同或独立的FDN。在滤波器组域中实现FDN的主要益处是允许应用具有与频率相关的混响性能的混响。在各种实施例中,通过使用各种滤波器组(包含但不限于实数值或复数值正交镜像滤波器(QMF)、有限脉冲响应滤波器(FIR滤波器)、无限脉冲响应滤波器(IIR滤波器)、离散傅立叶变换(DFT)、(修正的)余弦或正弦变换、小波变换或交叠滤波器(cross-over filter))中的每一个,在宽范围的各种滤波器组域的任一个中实现FDN。在优选的实现中,使用的滤波器组或变换包含用以降低FDN处理的计算复杂性的抽取(例如,减少频域信号表示的采样率)。

[0023] 第一类(和第二类)的一些实施例实现以下特征中的一个或多个:

[0024] 1. 滤波器组域(例如,混合复正交镜像滤波器域)FDN实现或混合滤波器组域FDN实现和时域晚期混响滤波器实现,其例如通过提供改变在不同的带中的混响箱延迟以作为频

率的函数改变模态密度的能力,典型地允许对于各频带独立调整FDN的参数和/或设定(使得能够对频率相关声学属性进行简单和灵活的控制);

[0025] 2.为了在直接和晚期响应之间保持适当的水平和定时关系,用于(从多通道输入音频信号)产生在第二处理路径中处理的下混(例如,单音下混)信号的特定下混处理依赖于各通道的源距离和直接响应的操作。

[0026] 3.在第二处理路径中(例如,在FDN的群的输入或输出处)应用全通滤波器(APF),以在不改变得到的混响的频谱和/或音色的情况下引入相位差异和增大的回声密度;

[0027] 4.在复值、多比率结构中在各FDN的反馈路径中实现分数延迟(fractional delay),以克服与被量化为下采样因子网格的延迟有关的问题;

[0028] 5.在FDN中,通过使用基于各频带中的希望的耳间相干性设定的输出混合系数,混响箱输出直接线性混合到双耳通道中。可选地,混响箱到双耳输出通道的映射跨着频带交替,以在双耳通道之间实现经平衡的延迟。而且,可选地,向混响箱输出应用归一化因子以在保留分数延迟和总功率的同时均一化它们的水平;

[0029] 6通过设定各频带中的增益与混响箱延迟的适当的组合控制依赖于频率的混响衰变时间和/或模态密度,以对真实房间进行仿真;

[0030] 7.对于每个频带(例如,在相关处理路径的输入或输出处)应用一个标度因子,以:

[0031] 控制与真实房间匹配的频率相关直接与晚期比(DLR)(可使用简单模型以基于目标DLR和例如为T60的混响衰变时间计算需要的标度因子);

[0032] 提供低频衰减以减轻过量的组合伪像和/或低频杂声;和/或

[0033] 向FDN响应应用扩散场谱整形;

[0034] 8.实现用于控制诸如混响衰变时间、耳间相干性和/或直接与晚期比的晚期混响的必要频率相关属性的简单的参数模型。

[0035] 本发明的多个方面包括执行(或被配置为执行或支持执行)音频信号(例如,其音频内容由扬声器通道构成的音频信号和/或基于对象的音频信号)的双耳虚拟化的方法和系统。

[0036] 在另一类的实施例中,本发明是一种响应多通道音频输入信号的一组通道产生双耳信号的方法和系统,包括对于该组通道中的每一通道应用双耳房间脉冲响应(BRIR),由此产生经滤波的信号(包含通过使用单个反馈延迟网络(FDN)以向该组通道中的通道的下混应用公共晚期混响);和组合经滤波的信号以产生双耳信号。该FDN在时域中实现。在一些这样的实施例中,时域FDN包括:

[0037] 输入滤波器,具有被耦接以接收下混的输入,其中,该输入滤波器被配置用于响应于下混产生第一经滤波的下混;

[0038] 全通滤波器,被耦接和配置为响应于第一经滤波的下混产生第二经滤波的下混;

[0039] 混响应用子系统,具有第一输出和第二输出,其中,混响应用子系统包括一组混响箱,每一混响箱具有不同的延迟,并且其中混响应用子系统被耦接并配置用于响应于第二经滤波的下混产生第一未混合双耳通道和第二未混合双耳通道,在第一输出处断言第一未混合双耳通道并且在第二输出处断言第二未混合双耳通道;以及

[0040] 耳间互相关系数(IACC)滤波和混合级,被耦接到混响应用子系统,并且被配置用于响应于第一未混合双耳通道和第二未混合双耳通道产生第一混合双耳通道和第二混合

双耳通道。

[0041] 输入滤波器可被实现为(优选地作为两个滤波器的级联,该两个滤波器被配置用于)产生第一经滤波的下混,使得每个BRIR具有至少基本上匹配目标直接与晚期比(DLR)的直接与晚期比(DLR)。

[0042] 每个混响箱可被配置用于产生延迟信号,并且可包括混响滤波器(例如,被实现为架式型滤波器(shelf filter)),该混响滤波器被耦接和配置用于向在所述每个混响箱中传播的信号应用增益,使得延迟信号具有至少基本上匹配用于所述延迟信号的目标衰变增益的增益,旨在实现各BRIR的目标混响衰变时间特性(例如, T_{60} 特性)。

[0043] 在一些实施例中,第一未混合双耳通道领先于第二未混合双耳通道,混响箱包括被配置用于产生具有最短延迟的第一延迟信号的第一混响箱和被配置用于产生具有次最短延迟的第二延迟信号的第二混响箱,其中第一混响箱被配置用于向第一延迟信号应用第一增益,第二混响箱被配置用于向第二延迟信号应用第二增益,第二增益与第一增益不同,第二增益不同于第一增益,并且第一增益和第二增益的应用导致第一未混合双耳通道相对于第二未混合双耳通道衰减。典型的,第一混合双耳通道和第二混合双耳通道指示被重新居中(recenter)的立体声图像。在一些实施例中,IACC滤波和混合级被配置用于产生第一混合双耳通道和第二混合双耳通道,使得所述第一混合双耳通道和第二混合双耳通道具有至少基本上匹配目标IACC特性的IACC特性。

[0044] 本发明的典型的实施例提供用于支持由扬声器通道构成的输入音频和基于对象的输入音频两者的简单且统一的构架。在向作为对象通道的输入信号通道应用BRIR的实施例中,在各对象通道上执行的“直接响应和早期反射”处理假定由具有对象通道的音频内容的元数据指示的源方向。在向作为扬声器通道的输入信号通道应用BRIR的实施例中,在各扬声器通道上执行的“直接响应和早期反射”处理假定与扬声器通道对应的源方向(即,从相应的扬声器的假定位置到假定的收听者位置的直接路径的方向)。不管输入通道是对象通道还是扬声器通道,“晚期混响”处理都在输入通道的下混(例如,单音下混)上被执行,且不假定下混的音频内容的任何特定的源方向。

[0045] 本发明的其它方面是被配置为(例如,被编程为)执行本发明的方法的任何实施例的耳机虚拟化器、包含这种虚拟化器的系统(例如,立体、多通道或其它解码器)和存储用于实现本发明的方法的任何实施例的代码的计算机可读介质(例如,盘)。

附图说明

[0046] 图1是常规的耳机虚拟化系统的框图。

[0047] 图2是包含本发明的耳机虚拟化系统的实施例的系统的框图。

[0048] 图3是本发明的耳机虚拟化系统的另一实施例的框图。

[0049] 图4是包含于图3系统的典型实现中的一种类型的FDN的框图。

[0050] 图5是可通过本发明的虚拟化器的实施例实现的作为以Hz计的频率的函数的以毫秒计的混响衰变时间(T_{60})的曲线图,对于该虚拟化器,两个特定频率(f_A 和 f_B)中的每一个处的 T_{60} 的值被设定如下:在 $f_A=10\text{Hz}$ 时, $T_{60,A}=320\text{ms}$,在 $f_B=2.4\text{Hz}$ 时, $T_{60,B}=150\text{ms}$ 。

[0051] 图6是可通过本发明的虚拟化器的实施例实现的作为以Hz计的频率的函数的耳间相干性(Coh)的曲线图,对于该虚拟化器,控制参数 Coh_{\max} 、 Coh_{\min} 和 f_C 被设定为具有以下的

值: $\text{Coh}_{\max}=0.95$, $\text{Coh}_{\min}=0.05$, $f_c=700\text{Hz}$ 。

[0052] 图7是可通过本发明的虚拟化器的实施例实现的作为以Hz计的频率的函数的在源距离为1米的情况下的以dB计的直接与晚期比 (DLR) 的示图, 对于该虚拟化器, 控制参数 DLR_{1K} 、 $\text{DLR}_{\text{slope}}$ 、 DLR_{\min} 、 $\text{HPF}_{\text{slope}}$ 和 f_T 被设定为具有以下的值: $\text{DLR}_{1K}=18\text{dB}$, $\text{DLR}_{\text{slope}}=6\text{dB}/10$ 倍频率, $\text{DLR}_{\min}=18\text{dB}$, $\text{HPF}_{\text{slope}}=6\text{dB}/10$ 倍频率, $f_T=200\text{Hz}$ 。

[0053] 图8是本发明的耳机虚拟化系统的晚期混响处理子系统的另一实施例的框图。

[0054] 图9是包含于本发明的系统的一些实施例中的一种类型的FDN的时域实现的框图。

[0055] 图9A是图9的滤波器400的实现的示例的框图。

[0056] 图9B是图9的滤波器406的实现的示例的框图。

[0057] 图10是本发明的耳机虚拟化系统的实施例的框图, 其中晚期混响处理子系统221在时域中实现。

[0058] 图11是图9的FDN的元件422、423和424的实施例的框图。

[0059] 图11A是图11的滤波器500的典型实现的频率响应 (R1)、图11的滤波器501的典型实现的频率响应 (R2) 以及并联连接的滤波器500和501的响应的曲线图。

[0060] 图12是可通过图9的FDN的实现而获得的IACC特性 (曲线“ I ”) 以及目标IACC特性 (曲线“ I_T ”) 的示例的曲线图。

[0061] 图13是通过适当地将滤波器406、407、408和409中的每一个实现为架式型滤波器而可通过图9的FDN的实现而获得的 T_{60} 特性的曲线图。

[0062] 图14是通过适当地将滤波器406、407、408和409中的每一个实现为两个IIR滤波器的级联而可通过图9的FDN的实现而获得的 T_{60} 特性的曲线图。

具体实施方式

[0063] (表示法和术语)

[0064] 在整个本公开中 (包含在权利要求中), 在广义上使用表达方式“对”信号或数据执行操作 (例如, 对信号或数据滤波、缩放、变换或者应用增益), 以表示直接对信号或数据执行操作或者对信号或数据的经处理版本 (例如, 在执行操作之前已经受到初步滤波或预处理的信号的版本) 执行操作。

[0065] 在整个本公开中 (包含在权利要求中), 在广义上使用表达方式“系统”以表示装置、系统或子系统。例如, 实现虚拟化器的子系统可被称为虚拟化器系统, 并且, 包含这种子系统的系统 (例如, 响应多个输入产生X个输出信号的系统, 其中, 子系统产生输入中的M个输入, 并且, 从外部源接收其它的X-M个输入) 也可被称为虚拟化器系统 (或虚拟化器)。

[0066] 在整个本公开中 (包含在权利要求中), 在广义上使用表达方式“处理器”以表示可编程为或者 (例如, 通过软件或固件) 另外可被配置为对数据 (例如, 音频或视频或其它图像数据) 执行操作的系统或装置。处理器的例子包括场可编程门阵列 (或其它可配置的集成电路或芯片组)、被编程并且/或者另外被配置为对音频或其它声音数据执行流水线处理的数字信号处理器、可编程通用处理器或计算机、以及可编程微处理器芯片或芯片组。

[0067] 在整个本公开中 (包含在权利要求中), 在广义上使用表达方式“分析滤波器组”以表示如下这样的系统 (例如, 子系统), 其被配置为对时域信号应用变换 (例如, 时域到频域变换) 以在一组频带中的每一频带中产生指示时域信号的内容的值 (例如, 频率成分)。在整

个本公开中(包含在权利要求中),在广义上使用表达方式“滤波器组域”以表示通过变换或分析滤波器组产生的频率成分的域(例如,在其中处理这种频率成分的域)。滤波器组域的例子包含(但不限于)频域、正交镜像滤波器(QMF)域和混合复正交镜像滤波器(HCQMF)域。可通过分析滤波器组应用的变换的例子包含(但不限于)离散余弦变换(DCT)、修正离散余弦变换(MDCT)、离散傅立叶变换(DFT)和小波变换。分析滤波器组的例子包含(但不限于)正交镜像滤波器(QMF)、有限脉冲响应滤波器(FIR滤波器)、无限脉冲响应滤波器(IIR滤波器)、交叠滤波器和具有其它适当的多速率结构的滤波器。

[0068] 在整个本公开中(包含在权利要求中),术语“元数据”指的是与相应的音频数据(也包含元数据的位流的音频内容)分开且不同的数据。元数据与音频数据相关联,并指示音频数据的至少一个特征或特性(例如,对于音频数据或者由音频数据指示的对象的轨迹,已执行或者应执行什么类型的处理)。元数据与音频数据的相关联是时间同步的。因此,当前(最近接收或更新)的元数据可指示相应的音频数据同时具有被指示的特征并且/或者包含被指示类型的音频数据处理的结果。

[0069] 在整个本公开中(包含在权利要求中),使用术语“耦接”或“被耦接”以意味着直接或间接连接。因此,如果第一装置与第二装置耦接,那么该连接可以通过直接连接,或者通过经由其它装置和连接的间接连接。

[0070] 在整个本公开中(包含在权利要求中),以下的表达方式具有以下定义:

[0071] 扬声器和扩音器被同义使用以表示任何声音发射换能器。该定义包括实现多个换能器(例如,低音炮和高音喇叭)的扩音器;

[0072] 扬声器馈送:直接应用于扩音器的音频信号,或者要被应用于串行的放大器和扩音器的音频信号;

[0073] 通道(或“音频通道”):单音音频信号。这种信号可以典型地以等同于向希望或标称位置处的扩音器直接应用信号的方式被呈现。希望的位置可以是静止的(物理扩音器典型地是这种情况),或者可以是动态的。

[0074] 音频节目:一组的一个或多个音频通道(至少一个扬声器通道和/或至少一个对象通道),并且可选地,还包含相关联的元数据(例如,描述希望的空间音频表示的元数据);

[0075] 扬声器通道(或“扬声器馈送通道”):与指定扩音器(处于希望或标称位置)相关联或者与被限定的扬声器配置内的指定扬声器区域相关联的音频通道。扬声器通道以等同于向指定扩音器(处于希望或标称位置)或者向指定扬声器区域中的扬声器直接应用音频信号的方式被呈现。

[0076] 对象通道:指示由音频源(有时,称为音频“对象”)发出的声音的音频通道。典型地,对象通道确定参数音频源描述(例如,指示参数音频源描述的元数据被包含于对象通道中或者与对象通道一起被提供)。源描述可确定由源发出的声音(作为时间的函数)、作为时间的函数的源的表现位置(例如,3D空间坐标),并且可选地确定表征源的至少一个附加的参数(例如,表现源尺寸或宽度);

[0077] 基于对象的音频节目:音频节目,该音频节目包含一组的一个或多个对象通道(并且可选地还包含至少一个扬声器通道),并且,可选地还包含相关联的元数据(例如,指示发出由对象通道指示的声音的音频对象的轨迹的元数据或另外指示由对象通道指示的声音的希望的空间音频表示的元数据,或指示作为由对象通道指示的声音的源的至少一个

音频对象的元数据)；

[0078] 呈现:将音频节目转换成一个或更多个扬声器馈送的处理或将音频节目转换成一个或更多个扬声器馈送并且通过使用一个或更多个扩音器将扬声器馈送转换成声音的处理(在后一种情况下,呈现有时在这里被称为“通过”扩音器呈现)。可通过直接向希望的位置处的物理扩音器应用信号而(“在”希望的位置处)通常地(trivially)呈现音频通道,或者,可通过使用被设计为(对于收听者而言)基本上等同于这种通常呈现的各种虚拟化技术中的一种呈现一个或更多个音频通道。在后一种情况下,各音频通道可被转换成应用到在一般与希望的位置不同的已知位置的扩音器的一个或更多个扬声器馈送,使得响应馈送通过扩音器发出的声音将被感觉为是从希望的位置发出的。这种虚拟化技术的例子包括通过耳机的双耳呈现(例如,通过使用对于耳机配戴者仿真可达7.1环绕声通道的Dolby Headphone处理)和波场合成。

[0079] 这里,多通道音频信号是“x.y”或“x.y.z”通道信号的表示法指示信号具有“x”全频率扬声器通道(与标称位于假定的收听者的耳朵的水平面中的扬声器对应)、“y”LFE(或低音炮)通道,并且,还任选地具有“z”全频率头顶扬声器通道(与位于假定的收听者的头部上方、例如处于房间的天花板或附近的扬声器对应)。

[0080] 这里,表述“IACC”的通常含义指的是耳间互相关系数,其是音频信号到达收听者的耳朵的时间之间的差的量度,典型地由从第一值到中间值到最大值的范围中的数值指示,该第一值指示到达信号的幅值相等并且正好异相,中间值指示到达信号不具有相似性,最大值指示相同到达信号具有相同的幅值和相位。

[0081] 优选实施例的详细描述

[0082] 本发明的许多实施例在技术上是可能的。通过本公开本领域技术人员将明了如何实现这些实施例。将参照图2到14描述本发明的系统和方法的实施例。

[0083] 图2是包括本发明的耳机虚拟化系统的实施例的系统(20)的框图。耳机虚拟化系统(有时称为虚拟化器)被配置为向多通道音频输入信号的N个全频率范围通道(X_1, \dots, X_N)应用双耳房间脉冲响应(BRIR)。通道 X_1, \dots, X_N (可以是扬声器通道或对象通道)的每一个与相对于假定的收听者的特定的源方向和距离对应,并且,图2系统被配置为将每一这样的通道与用于相应的源方向和距离的BRIR卷积。

[0084] 系统20可以是解码器,其被耦接为接收编码音频节目并包含被耦接和配置为通过从该节目恢复N个全频率范围通道(X_1, \dots, X_N)而解码该节目并将它们提供给虚拟化系统的元件12、 \dots 、14和15(包含如所示的那样耦接的元件12、 \dots 、14、15、16和18)的子系统(图2未示出)。解码器可包含附加的子系统,其中的一些执行不与由虚拟化系统执行的虚拟化功能有关的功能,并且其中的一些可执行与虚拟化功能有关的功能。例如,后一些功能可包含从编码的节目提取元数据和将元数据提供给虚拟化控制子系统,该虚拟化控制子系统使用元数据以控制虚拟化器系统的元件。

[0085] 子系统12(与子系统15)被配置为将通道 X_1 与BRIR₁(用于相应的源方向和距离的BRIR)卷积,子系统14(与子系统15)被配置为将通道 X_N 与BRIR_N(用于相应的源方向的BRIR)卷积,并且对于N-2个其它的BRIR子系统中的一个也是诸如此类的。子系统12、 \dots 、14和15中的每一个的输出是包含左通道和右通道的时域信号。加算元件16和18与元件12、 \dots 、14和15的输出耦接。加算元件16被配置为组合(混合)BRIR子系统的左通道输出,并且,加算元件

18被配置为组合(混合)BRIR子系统的右通道输出。元件16的输出是从图2的虚拟化器输出的双耳音频信号的左通道L,并且,元件18的输出是从图2的虚拟化器输出的双耳音频信号的右通道R。

[0086] 从本发明的耳机虚拟化器的图2实施例与图1的常规的耳机虚拟化器的比较可清楚地看出本发明的典型实施例的重要特征。出于比较的目的,我们假定图1和图2系统被配置为使得,当对它们中的每一个断言同一多通道音频输入信号时,系统向输入信号的全频率范围通道 X_i 应用具有相同的直接响应和早期反射部分的BRIR $_i$ (即,图2的相关EBRIR $_i$)(但未必具有相同的成功度)。通过图1或图2系统应用的各BRIR $_i$ 可分解成两个部分:直接响应和早期反射部分(例如,通过图2的子系统12~14应用的EBRIR $_1$ 、 \cdots 、EBRIR $_N$ 部分中的一个)和晚期混响部分。图2实施例(和本发明的其它典型实施例)假定单通道BRIR的晚期混响部分BRIR $_i$ 可跨着源方向并因此跨着所有通道被共享,并因此向输入信号的所有全频率范围通道的下混应用相同的晚期混响(即,公共晚期混响)。该下混可以是所有输入通道的单音(单声道)下混,但作为替代,可以从输入通道(例如,从输入通道的子集)获得的立体声或多通道下混。

[0087] 更具体而言,图2的子系统12被配置为将通道 X_1 与EBRIR $_1$ (用于相应的源方向的直接响应和早期反射BRIR部分)卷积,子系统14被配置为将通道 X_N 与EBRIR $_N$ (用于相应的源方向的直接响应和早期反射BRIR部分)卷积,等等。图2的晚期混响子系统15被配置为产生输入信号的所有全频率范围通道的单声道下混,并将该下混与LBRIR(被下混的所有通道的公共晚期混响)卷积。图2虚拟化器的各BRIR子系统(子系统12、 \cdots 、14和15中的每一个)的输出包含(从相应的扬声器通道或下混产生的双耳信号的)左通道和右通道。BRIR子系统的左通道输出在加算元件16中组合(混合),并且,BRIR子系统的右通道输出在加算元件18中组合(混合)。

[0088] 假定在子系统12、 \cdots 、14和15中实现适当的水平调整和时间对准,加算元件(addition element)16可实现为简单地合计相应的左双耳通道采样(子系统12、 \cdots 、14和15的左通道输出),以产生双耳输出信号的左通道。类似地,同样假定在子系统12、 \cdots 、14和15中实现适当的水平调整和时间对准,加算元件18也可实现为简单地合计相应的右双耳通道采样(例如,子系统12、 \cdots 、14和15的右通道输出),以产生双耳输出信号的右通道。

[0089] 图2的子系统15可被以各种方式中的任一种实现,但典型地包括被配置为向对其断言的输入信号通道的单音下混应用公共晚期混响的至少一个反馈延迟网络。典型地,在子系统12、 \cdots 、14中的每一个应用它处理的通道(X_i)的单通道BRIR的直接响应和早期反射部分(EBRIR $_i$)的情况下,公共晚期混响被产生以模仿单通道BRIR(其“直接响应和早期反射部分”通过子系统12、 \cdots 、14被应用)中的至少一些(例如,全部)的晚期混响部分的共同宏观属性。例如,子系统15的一个实现具有与图3的子系统200相同的结构,该子系统200包含被配置为向对其断言的输入信号通道的单音下混应用公共晚期混响的反馈延迟网络的群(203、204、 \cdots 、205)。

[0090] 图2的子系统12、 \cdots 、14可被以各种方式中的任一种实现(在时域中或在滤波器组域中),任何特定应用的优选实现依赖于各种考虑(诸如(例如)性能、计算和存储)。在一个示例性实现中,子系统12、 \cdots 、14中的每一个被配置为将其断言的通道与对应于和该通道相关联的直接和早期响应的FIR滤波器卷积,其中增益和延迟被适当地设定为使得子系统

12、...、14的输出可简单且高效地与子系统15的那些输出组合。

[0091] 图3是本发明的耳机虚拟化系统的另一实施例的框图。图3实施例与图2类似,其中两个(左通道和右通道)时域信号从直接响应和早期反射处理子系统100被输出,并且两个(左通道和右通道)时域信号从晚期混响处理子系统200被输出。加算元件210与子系统100和200的输出耦接。元件210被配置为组合(混合)子系统100和200的左通道输出以产生从图3虚拟化器输出的双耳音频信号的左通道L,并组合(混合)子系统100和200的右通道输出以产生从图3虚拟化器输出的双耳音频信号的右通道R。假定在子系统100和200中实现了适当的水平调整和时间对准,元件210可实现为简单地合计从子系统100和200输出的相应的左通道采样以产生双耳输出信号的左通道,并简单地合计从子系统100和200输出的相应的右通道采样以产生双耳输出信号的右通道。

[0092] 在图3系统中,多通道音频输入信号的通道 X_i 被引向两个并行处理路径并在其中经受处理:一个处理路径通过直接响应和早期反射处理子系统100;另一个处理路径通过晚期混响处理子系统200。图3系统被配置为向各通道 X_i 应用 $BRIR_i$ 。各 $BRIR_i$ 可分解成两个部分:直接响应和早期反射部分(通过子系统100被应用)和晚期混响部分(通过子系统200被应用)。在操作中,直接响应和早期反射处理子系统100由此产生从虚拟化器输出的双耳音频信号的直接响应和早期反射部分,并且,晚期混响处理子系统(“晚期混响产生器”)200由此产生从虚拟化器输出的双耳音频信号的晚期混响部分。子系统100和200的输出(通过加算子系统210)被混合以产生典型地从子系统210向呈现系统(未示出)断言的双耳音频信号,在该呈现系统中,该信号经受双耳呈现以供耳机回放。

[0093] 典型地,当通过一对耳机呈现和再现时,从元件210输出的典型的双耳音频信号在收听者的耳膜被感知为来自处于宽范围的各种位置中的任一个的“N”个扩音器(这里 $N \geq 2$,且N典型地等于2、5或7)的声音,这些位置包含处于收听者前方、后方和上方的位置。在图3系统的操作中产生的输出信号的再现可给予收听者声音来自多于两个(例如,5个或7个)“环绕声”源的体验。这些源中的至少一些是虚拟的。

[0094] 直接响应和早期反射处理子系统100可被以各种方式中的任一种实现(在时域中或在滤波器组域中),其中任何特定应用的优选实现依赖于各种考虑(诸如(例如)性能、计算和存储)。在一个示例性实现中,子系统100被配置为将对其断言的各通道和对应于与该通道相关联的直接和早期响应的FIR滤波器卷积,其中增益和延迟被适当地设定为使得子系统100的输出可简单且高效地与子系统200的那些输出相组合(在元件210中)。

[0095] 如图3所示,晚期混响产生器200包含如所示的那样耦接的下混子系统201、分析滤波器组202、FDN群(FDN 203、204、...、和205)和合成滤波器组207。子系统201被配置为将多通道输入信号的通道下混到单声道下混,并且,分析滤波器组202被配置为向该单声道下混应用变换以将该单声道下混分成“K”个频带,这里,K是整数。对于FDN 203、204、...、205中的不同的一个断言各不同的频带中的滤波器组域值(从滤波器组202输出的)(这些FDN中的“K”个分别被耦接和被配置为向对其断言的滤波器组域值应用BRIR的晚期混响部分)。滤波器组域值优选在时间上被抽取以降低FDN的计算复杂度。

[0096] 原则上,(对于图3的子系统100和子系统201的)各输入通道可在其自身FDN(或FDN群)中被处理,以仿真其BRIR的晚期混响部分。尽管与不同的声源位置相关联的BRIR的晚期混响部分典型地在脉冲响应中的均方根差方面明显不同,但诸如它们的平均功率谱、它们

的能量衰变结构、模态密度和峰密度等的它们的统计属性常常是非常相似的。因此，一组BRIR的晚期混响部分典型地跨通道在感知上非常相似，因此能够使用一个共用FDN或FDN群（例如，FDN 203、204、…、205）以仿真两个或更多个BRIR的晚期混响部分。在典型的实施例中，使用一个这种共用FDN（或FDN群），并且，其输入包含从输入通道构建的一个或更多个下混。在图2的示例性实施例中，下混是所有输入通道的单声道下混（在子系统201的输出处被断言）。

[0097] 参照图2实施例，FDN 203、204、…、和205中的每一个在滤波器组域中被实现，并且被耦接和配置为处理从分析滤波器组202输出的值的不同频带，以产生各带的左混响信号和右混响信号。对于各带，左混响信号是滤波器组域值序列，并且右混响信号是另一滤波器组域值序列。合成滤波器组207被耦接和配置为向从FDN输出的2K个滤波器组域值序列（例如，QMF域频率成分）应用频域到时域变换，并将变换后的值组装成左通道时域信号（指示已应用晚期混响的单声道下混的音频内容）和右通道时域信号（也指示已应用晚期混响的单声道下混的音频内容）。这些左通道信号和右通道信号被输出到元件210。

[0098] 在典型的实施例中，FDN 203、204、…、和205中的每一个在QMF域中被实现，并且，滤波器组202将来自子系统201的单声道下混变换至QMF域（例如，混合复正交镜像滤波器（HCQMF）域），使得从滤波器组202对FDN 203、204、…、和205中的每一个的输入断言的信号是QMF域频率成分序列。在这样的实现中，从滤波器组202对FDN 203断言的信号是第一频带中的QMF域频率成分序列，从滤波器组202对FDN 204断言的信号是第二频带中的QMF域频率成分序列，并且，从滤波器组202对FDN 205断言的信号是第“K”个频带中的QMF域频率成分序列。当分析滤波器组202这样被实现时，合成滤波器组207被配置为向来自FDN的2K个输出QMF域频率成分序列应用QMF域到时域变换，以产生输出到元件210的左通道和右通道晚期混响时域信号。

[0099] 例如，如果在图3系统中 $K=3$ ，那么存在对于合成滤波器组207的6个输入（从FDN 203、204和205中的每一个输出的左和右通道，包含频域或QMF域采样）和来自207的两个输出（左和右通道，分别由时域采样构成）。在本例子中，滤波器组207典型地会实现为两个合成滤波器组：一个合成滤波器组被配置为产生从滤波器组207输出的时域左通道信号（对于其将断言来自FDN 203、204和205的3个左通道）；并且第二合成滤波器组被配置为产生从滤波器组207输出的时域右通道信号（对于其将断言来自FDN 203、204和205的3个右通道）。

[0100] 可选地，控制子系统209与FDN 203、204、…、205中的每一个耦接，并被配置为对FDN中的每一个断言控制参数，以确定通过子系统200应用的晚期混响部分（LBRIR）。在下文描述这种控制参数的例子。设想在一些实现中，控制子系统209可实时操作（例如，响应通过输入装置对其断言的用户命令），以实现由子系统200应用到输入通道的单音下混的晚期混响部分（LBRIR）的实时变化。

[0101] 例如，如果对于图2系统的输入信号是5.1通道信号（其的全频率范围通道按以下的通道次序：L、R、C、Ls、Rs），那么所有全频率范围通道具有相同的源距离，并且，下混子系统201可实现为如下的简单地合计全频率范围通道以形成单声道下混的下混矩阵：

[0102] $D = [1 \ 1 \ 1 \ 1 \ 1]$

[0103] 在全通滤波（在FDN 203、204、…、205中的每一个中在元件301中）之后，单声道下混以功率守恒的方式上混到4个混响箱：

$$[0104] \quad U = \begin{bmatrix} 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \\ 1/\sqrt{4} \end{bmatrix}$$

[0105] 作为替代方案(作为例子),可选择将左侧通道扫调(pan)到前两个混响箱,将右侧通道扫调到最后两个混响箱,并将中心通道扫调到所有混响箱。在这种情况下,下混子系统201实现为形成两个下混信号:

$$[0106] \quad D = \begin{bmatrix} 1 & 0 & 1/\sqrt{2} & 1 & 0 \\ 0 & 1 & 1/\sqrt{2} & 0 & 1 \end{bmatrix}$$

[0107] 在本例子中,对于混响箱的上混(在FDN 203、204、...、205中的每一个中)为:

$$[0108] \quad U = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

[0109] 由于存在两个下混信号,因此,全通滤波(在FDN 203、204、...、205中的每一个中的元件301中)需要被应用两次。会对于(L,Ls)、(R,Rs)和C的晚期混响引入差异,尽管它们均具有相同的宏观属性。当输入信号通道具有不同的源距离时,仍需要在下混处理中应用适当的延迟和增益。

[0110] 下面描述图3虚拟化器的子系统100和200以及下混子系统201的特定实现的考虑。

[0111] 通过子系统201实现的下混处理依赖于要被下混的各通道的(声音源与假定的收听者位置之间)源距离和直接响应的处理。直接响应的延迟 t_d 为:

$$[0112] \quad t_d = d/v_s$$

[0113] 这里,d是声音源与收听者之间的距离, v_s 是声音速度。并且,直接响应的增益与 $1/d$ 成比例。如果在具有不同的源距离的通道的直接响应的处理中保留这些规则,那么子系统201可实现所有通道的直下混,原因是晚期混响的延迟和水平一般对源位置不敏感。

[0114] 由于实际考虑,虚拟化器(例如,图3的虚拟化器的子系统100)可实现为时间对准具有不同的源距离的输入通道的直接响应。为了保留各通道的直接响应和晚期反射之间的相对延迟,具有源距离d的通道在与其它的通道下混之前应被延迟 $(d_{\max} - d)/v_s$ 。这里, d_{\max} 表示最大可能源距离。

[0115] 虚拟化器(例如,图3的虚拟化器的子系统100)也可实现为压缩直接响应的动态范围。例如,具有源距离d的通道的直接响应可通过 $d^{-\alpha}$ 而不是 d^{-1} 的因子被缩放,这里, $0 \leq \alpha \leq 1$ 。为了保留直接响应和晚期混响之间的水平差,下混子系统201可能需要实现为在具有源距离d的通道与其它的缩放通道下混之前通过 $d^{1-\alpha}$ 的因子缩放它。

[0116] 图4的反馈延迟网络是图3的FDN 203(或204或205)的示例性实现。虽然图4系统具有4个混响箱(分别包含增益级 g^i 和与增益级的输出耦接的延迟线 z^{-n_i}),但系统的变型(和在本发明的虚拟化器的实施例中使用的其它FDN)实现多于或少于四个的混响箱。

[0117] 图4的FDN包含输入增益元件300,与元件300的输出耦接的全通滤波器 (APF) 301、与APF 301的输出耦接的加算元件302、303、304和305、以及分别与元件302、303、304和305中的不同的一个的输出耦接的4个混响箱 (分别包含增益元件 g_k (元件306中的一个)、与其耦接的延迟线 z^{-M_k} (元件307中的一个) 和与其耦接的增益元件 $1/g_k$ (元件309中的一个), 这里, $0 \leq k-1 \leq 3$)。酉矩阵 (unitary matrix) 308与延迟线307的输出耦接,并被配置为将反馈输出断言到元件302、303、304和305中的每一个的第二输入。(第一和第二混响箱的) 两个增益元件309的输出被断言至加算元件310的输入,并且,元件310的输出被断言至输出混合矩阵312的一个输入。(第三和第三混响箱的) 另两个增益元件309的输出被断言至加算元件311的输入,并且,元件311的输出被断言至输出混合矩阵312的另一个输入。

[0118] 元件302被配置为向第一混响箱的输入添加与延迟线 z^{-n1} 对应的矩阵308的输出 (即,通过矩阵308应用来自延迟线 z^{-n1} 的输出的反馈)。元件303被配置为向第二混响箱的输入添加与延迟线 z^{-n2} 对应的矩阵308的输出 (即,通过矩阵308应用来自延迟线 z^{-n2} 的输出的反馈)。元件304被配置为向第三混响箱的输入添加与延迟线 z^{-n3} 对应的矩阵308的输出 (即,通过矩阵308应用来自延迟线 z^{-n3} 的输出的反馈)。元件305被配置为向第四混响箱的输入添加与延迟线 z^{-n4} 对应的矩阵308的输出 (即,通过矩阵308应用来自延迟线 z^{-n4} 的输出的反馈)。

[0119] 图4的FDN的输入增益元件300耦接为接收从图3的分析滤波器组202输出的变换后单音下混信号 (滤波器组域信号) 的一个频带。输入增益元件300向对其断言的滤波器组域信号应用增益 (缩放) 因子 G_{in} 。所有频带的 (通过图3的全部FDN 203、204、...、205实现的) 缩放因子 G_{in} 共同地控制晚期混响的谱整形和水平。在图3虚拟化器的所有FDN中设定输入增益 G_{in} 常常考虑以下的目标:

[0120] 匹配真实房间的应用于各通道的BRIR的直接与晚期比 (DLR) ;

[0121] 用于减轻过量梳状伪像和/或低频杂声的必要的低频衰减;和

[0122] 扩散场谱包络线的匹配。

[0123] 如果假定 (通过图3的子系统100被应用的) 直接响应在所有的频带中提供单一增益,那么通过将 G_{in} 设定如下可实现特定的DLR (功率比) :

[0124] $G_{in} = \sqrt{\ln(10^6) / (T60 * DLR)}$,

[0125] 这里, $T60$ 是定义为混响衰变60dB所花费的时间的混响衰变时间 (通过后面讨论的混响延迟和混响增益确定),并且“ \ln ”表示自然对数函数。

[0126] 输入增益因子 G_{in} 可依赖于正被处理的内容。这种内容依赖性的一个应用是确保各时间/频率段中的下混的能量等于正被下混的各个通道信号的能量的和,而不管在输入通道信号之间是否可能存在任何相关性。在这种情况下,输入增益因子可以是 (或者可乘以) 类似于或等于下式的项:

$$[0127] \quad \sqrt{\frac{\sum_i \sum_j x_i^2(j)}{\sum_j y^2(j)}}$$

[0128] 这里, i 是给定时间/频率片段或子带的所有下混采样上的索引, $y(i)$ 是片段的下混采样, $x_i(j)$ 是对下混子系统201的输入断言的输入信号 (对于通道 X_i)。

[0129] 在图4的FDN的典型的QMF域实现中,从全通滤波器 (APF) 301的输出断言至混响箱

的输入的信号是QMF域频率成分序列。为了产生更自然的发声FDN输出,APF 301被应用到增益元件300的输出以引入相位差异和增大的回声密度。作为替代方案,或者,附加地,一个或更多个全通延迟滤波器可被应用到:(图3的)下混子系统301的各个输入(在该输入在子系统201中下混并通过FDN被处理之前);或者在图4所示的混响箱前馈或后馈路径中(例如,除了各混响箱中的延迟线 z^{-M_k} 以外或者作为其替代);或FDN的输出(即,输出矩阵312的输出)。

[0130] 在实现混响箱延迟 z^{-n_i} 时,混响延迟 n_i 应是互质数,以避免混响模式在相同频率处对准。为了避免伪发声输出,延迟的和应足够大以提供足够的模态密度。但是,最短的延迟应足够短以避免晚期混响与BRIR的其它成分之间的过量时间间隙。

[0131] 典型地,混响箱输出首先扫调到左或右双耳通道。通常,被扫调到两个双耳通道的混响箱输出的集合在数量上相等且相互排斥。还希望平衡这两个双耳通道的定时。因此,如果具有最短延迟的混响箱输出前往一个双耳通道,那么具有次最短延迟的混响箱输出会前往另一通道。

[0132] 混响箱延迟可以在频带间不同,以作为频率的函数改变模态密度。一般地,较低频带需要更高的模态密度,因此需要更长的混响箱延迟。

[0133] 混响箱增益 g_i 的幅值和混响箱延迟联合地确定图4的FDN的混响衰减时间:

[0134] $T_{60} = -3n_i / \log_{10}(|g_i|) / F_{FRM}$

[0135] 这里, F_{FRM} 是滤波器组202(图3)的帧率。混响箱增益的相位引入分数延迟以克服与被量化到滤波器组的下混因子网格的混响箱延迟有关的问题。

[0136] 单一反馈矩阵308在反馈路径中在混响箱之间提供均匀的混合。

[0137] 为了均一化混响箱输出的水平,增益元件309向各混响箱的输出应用归一化增益 $1/|g_i|$,以在保留通过它们的相位引入的分数延迟的同时去除混响箱增益的水平影响。

[0138] 输出混合矩阵312(也被标识为矩阵 M_{out})是被配置为混合来自初始扫调的未混合双耳通道(分别为元件310和311的输出)以实现具有希望的耳间相干性的输出左和右双耳通道(在矩阵312的输出处断言的L和R信号)的 2×2 矩阵。未混合双耳通道在初始扫调之后接近不相关,原因是它们不包含任何共用混响箱输出。如果希望的耳间相干性是Coh,这里 $|Coh| \leq 1$,那么输出混合矩阵312可被定义为:

[0139]
$$M_{out} = \begin{bmatrix} \cos \beta & \sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}, \text{ 其中 } \beta = \arcsin(Coh) / 2$$

[0140] 由于混响箱延迟不同,因此,未混合双耳通道中的一个会经常领先于另一个。如果混响箱延迟和扫调的组合跨频带是相同的,那么会导致声音图像偏差。如果跨着频带交替扫调图案使得混合双耳通道在交替的频带中相互领先和尾随,那么可减轻该偏差。这可通过如下操作来实现,即将输出混合矩阵312实现为在奇数频带中(即,在第一频带(通过图3的FDN 203处理)和第三频带等中)具有在前面的段落中阐述的形式,并在偶数频带中(即,在第二频带(通过图4的FDN204处理)和第四频带等中)具有以下形式:

[0141]
$$M_{out,alt} = \begin{bmatrix} \sin \beta & \cos \beta \\ \cos \beta & \sin \beta \end{bmatrix}$$

[0142] 这里, β 的定义保持相同。应当注意, 矩阵312可实现为在所有频带的FDN中相同, 但是, 其输入的通道次序可对交替的频带被切换(即, 在奇数频带中, 元件310的输出可被断言至矩阵312的第一输入且元件311的输出可被断言至矩阵312的第二输入, 并且, 在偶数频带中, 元件311的输出可被断言至矩阵312的第一输入且元件310的输出可被断言至矩阵312的第二输入)。

[0143] 在频带(部分)重叠的情况下, 在其上矩阵312的形式交替的频率范围的宽度可增加(即, 它可对于每两个或三个连续的带交替一次), 或者, 上式中的 β 的值(对于矩阵312的形式)可被调整以确保平均相干值等于希望的值以补偿连续频带的谱重叠。

[0144] 如果在本发明的虚拟化器中以上限定的目标声学属性 T_{60} 、Coh和DLR对于各特定的频带的FDN是已知的, 那么FDN中的每一个(均具有图4所示的结构)可被配置为实现目标属性。具体而言, 在一些实施例中, 各FDN的输入增益 (G_{in})、混响箱增益和延迟 (g_i 和 n_i) 和输出矩阵 M_{out} 的参数可被设定(例如, 通过由图3的控制子系统209对其断言的控制值被设定), 以根据这里描述的关系实现目标属性。实际上, 通过具有简单的控制参数的模型设定频率相关属性常常足以产生匹配特定声学环境的自然发声晚期混响。

[0145] 下面描述可如何通过确定少量的频带中的每一个的目标混响衰减时间 (T_{60}) 来确定本发明虚拟化器的实施例的各特定频带的FDN的目标混响衰减时间 (T_{60})。FDN响应的水平随时间以指数的方式衰变。 T_{60} 与衰变因子 df (定义为单位时间上的dB衰减) 成反比:

$$[0146] \quad T_{60} = 60/df.$$

[0147] 衰变因子 df 依赖于频率, 并且, 一般在对数频率坐标上线性增加, 因此, 混响衰减时间也是频率的函数, 一般随频率增加而减小。因此, 如果确定(例如, 设定)两个频率点的 T_{60} 值, 那么对于所有频率的 T_{60} 曲线被确定。例如, 如果频率点 f_A 和 f_B 的混响衰变时间分别是 $T_{60,A}$ 和 $T_{60,B}$, 那么 T_{60} 曲线被定义为:

$$[0148] \quad T_{60}(f) = \frac{T_{60,A} T_{60,B} \log(f_B / f_A)}{T_{60,A} \log(f / f_A) - T_{60,B} \log(f / f_B)}$$

[0149] 图5示出可通过本发明的虚拟化器的实施例实现的 T_{60} 曲线的例子, 对于该曲线, 两个特定频率 (f_A 和 f_B) 中的每一个处的 T_{60} 的值被设定为: 在 $f_A = 10\text{Hz}$ 处, $T_{60,A} = 320\text{ms}$, 在 $f_B = 2.4\text{Hz}$ 处, $T_{60,B} = 150\text{ms}$ 。

[0150] 下面描述可如何通过设定少量的控制参数来实现本发明的虚拟化器的实施例的各特定频带的FDN的目标耳间相干性 (Coh) 的例子。晚期混响的耳间相干性 (Coh) 在很大程度上遵循扩散声场的图案。其可通过直至交越频率 f_c 的 sinc 函数以及在交越频率以上的常数被模型化。Coh曲线的简单模型为:

$$[0151] \quad \text{Coh}(f) = \begin{cases} \text{Coh}_{\min} + (\text{Coh}_{\max} - \text{Coh}_{\min}) \text{sinc}(f / f_c) & , f \leq f_c \\ \text{Coh}_{\min} & , f \geq f_c \end{cases}$$

[0152] 这里, 参数 Coh_{\min} 和 Coh_{\max} 满足 $-1 \leq \text{Coh}_{\min} < \text{Coh}_{\max} \leq 1$, 并且控制Coh的范围。最佳交越频率 f_c 依赖于收听者的头部尺寸。 f_c 太高导致内在化的声源图像, 而值太小导致声源图像分散或分离。图6是可通过本发明的虚拟化器的实施例实现的Coh曲线的例子, 对于该曲线, 控制参数 Coh_{\max} 、 Coh_{\min} 和 f_c 被设定为具有以下的值: $\text{Coh}_{\max} = 0.95$, $\text{Coh}_{\min} = 0.05$, $f_c = 700\text{Hz}$ 。

[0153] 下面描述可如何通过设定少量的控制参数来实现本发明的虚拟化器的实施例的各特定频带的FDN的目标直接与晚期比 (DLR) 的例子。单位为dB的直接与晚期比 (DLR) 一般在对数频率坐标上线性增加。它可通过设定 DLR_{1K} (在1KHz的DLR, 单位为dB) 和 DLR_{slope} (以每10倍频率的dB计) 被控制。但是, 较低频范围中的低DLR常常导致过量的梳状伪像。为了减轻该伪像, 添加两个修正机制以控制DLR:

[0154] 最小DLR底: DLR_{min} (以dB计); 和

[0155] 由过渡频率 f_T 和低于该频率的衰减曲线斜率 HPF_{slope} (以每10倍频率的dB计) 定义的高通滤波器。

[0156] 得到的单位是dB的DLR曲线被定义如下:

$$[0157] \quad DLR(f) = \max(DLR_{1K} + DLR_{slope} \log_{10}(f/1000), DLR_{min}) + \min(HPF_{slope} \log_{10}(f/f_T), 0)$$

[0158] 应当注意, 即使在相同的声学环境中, DLR也随源距离改变。因此, 这里, DLR_{1K} 和 DLR_{slope} 两者是对于诸如1米的标称源距离的值。图7是通过本发明的虚拟化器的实施例实现的对于1米源距离的DLR曲线的例子, 其中控制参数 DLR_{1K} 、 DLR_{slope} 、 DLR_{min} 、 HPF_{slope} 和 f_T 被设定为具有以下值: $DLR_{1K}=18\text{dB}$, $DLR_{slope}=6\text{dB}/10\text{倍频率}$, $DLR_{min}=18\text{dB}$, $HPF_{slope}=6\text{dB}/10\text{倍频率}$, $f_T=200\text{Hz}$ 。

[0159] 这里公开的实施例的变型例具有以下特征中的一个或多个:

[0160] 本发明的虚拟化器的FDN在时域中实现, 或者, 它们具有带有基于FDN的脉冲响应捕获和基于FIR的信号滤波的混合实现。

[0161] 本发明的虚拟化器实现为允许在执行下混步骤期间应用作为频率的函数的能量补偿, 该下混步骤产生用于晚期混响处理子系统的下混输入信号; 并且,

[0162] 本发明的虚拟化器实现为允许响应外部因素 (即, 响应控制参数的设定) 手动或自动控制被应用的晚期混响属性。

[0163] 对于其中系统延滞是关键的和由分析和合成滤波器组导致的延迟被禁止的应用, 本发明的虚拟化器的典型实施例的滤波器组域FDN结构可被变换至时域, 并且, 在虚拟化器的一类实施例中可在时域中实现各FDN结构。在时域实现中, 为了允许依赖频率的控制, 应用输入增益因子 (G_{in})、混响箱增益 (g_i) 和归一化增益 ($1/|g_i|$) 的子系统被具有类似的振幅响应的滤波器替代。输出混合矩阵 (M_{out}) 也被滤波器的矩阵替代。与其它的滤波器不同, 该滤波器的矩阵的相位响应是关键的原因, 其原因是功率守恒和耳间相干性可能受相位响应影响。时域实现中的混响箱衰变可能需要 (相对于它们在滤波器组域实现中的值) 稍微改变, 以避免作为共用因子共享滤波器组步幅。由于各种约束, 本发明的虚拟化器的FDN的时域实现的性能不能确切地匹配其滤波器组域实现的性能。

[0164] 下面参照图8描述本发明的虚拟化器的本发明的晚期混响处理子系统的混合 (滤波器组域和时域) 实现。本发明的晚期混响处理子系统的该混合实现是实现基于FDN的脉冲响应捕获和基于FIR的信号过滤的图4的晚期混响处理子系统的变型例。

[0165] 图8的实施例包含元件201、202、203、204、205和207, 它们与图3的子系统200的附图标记相同的元件相同。将不参照图8重复这些元件的以上描述。在图8实施例中, 单位脉冲产生器211被耦接为对分析滤波器组202断言输入信号 (脉冲)。实现为FIR滤波器的LBRIR滤

波器208(单声道入、立体声出)向从子系统201输出的单音下混应用适当的BRIR的晚期混响部分(LBRIR)。因此,元件211、202、203、204、205和207是到LBRIR滤波器208的处理侧链。

[0166] 每当要修正晚期混响部分LBRIR的设定时,脉冲产生器211操作以对元件202断言单位脉冲,并且,得到的来自滤波器组207的输出被捕获并且被断言至滤波器208(以设定滤波器208来应用由滤波器组207的输出确定的新LBRIR)。为了加速从LBRIR设定变化到新LBRIR生效的时间的时间流逝,新LBRIR的采样可在变得可用时开始替代旧LBRIR。为了缩短FDN的固有延滞,可以舍弃LBRIR的初始零。这些选项提供了灵活性,并允许混合实现提供潜在的性能提高(相对于由滤波器组域实现所提供的),但代价是来自FIR过滤的计算增加。

[0167] 对于系统延滞是关键的关键但计算能力较不关注的的应用,可使用侧链滤波器组域晚期混响处理器(例如,通过图8的元件211、202、203、204、...205和207实现)以捕获要由滤波器208应用的有效FIR脉冲响应。FIR滤波器208可实现该被捕获的FIR响应并且直接将其应用到输入通道的单声下混(在输入通道的虚拟化期间)。

[0168] 例如,通过利用可由系统的用户(例如,通过操作图3的控制子系统209)调整的一个或更多个预设,各种FDN参数以及作为结果的晚期混响属性可被手动调谐并随后硬接线到本发明的晚期混响处理子系统的实施例中。但是,给定晚期混响、其与FDN参数的关系以及修正其行为的能力的高级描述,各种方法被构想用于控制基于FDN的晚期混响处理器的各种实施例,包括(但不限于)以下方面:

[0169] 1. 最终用户可例如通过显示器上的(例如,通过图3的控制子系统209的实施例实现的)用户界面或使用(例如,通过图3的控制子系统209的实施例实现的)物理控件切换预设来手动控制FDN参数。以这种方式,最终用户可根据爱好、环境或内容调整房间仿真。

[0170] 2. 例如,通过与输入音频信号一起提供的元数据,要被虚拟化的音频内容的作者可提供与内容本身一起被传送的设定或希望的参数。这种元数据可被解析和使用(例如,通过图3的控制子系统209的实施例),以控制相关的FDN参数。因此,元数据可指示诸如混响时间、混响水平和直接与混响比等的性能,并且,这些性能可以是随时间改变的,并且可通过时变元数据被信令。

[0171] 3. 回放装置可通过使用一个或更多个传感器获知其位置或环境。例如,移动装置可使用GSM网络、全球定位系统(GPS)、已知的WiFi接入点或任何其它的位置服务,以确定装置处于哪里。随后,(例如,通过图3的控制子系统209的实施例)可使用指示位置和/或环境的数据,以控制相关的FDN参数。因此,可响应装置的位置修改FDN参数,以例如模拟物理环境。

[0172] 4. 关于回放装置的位置,可以使用云服务或社交媒体以得出消费者在某个环境中最常用的设定。另外,用户可与(已知)位置相关联地向云服务或社交媒体服务上载他们的当前的设定,以使得可用于其它用户或自身。

[0173] 5. 回放装置可包含诸如照相机、光传感器、麦克风、加速计、陀螺仪的其它传感器,以确定用户的活动和用户所处的环境,以优化用于该特定活动和/或环境的FDN参数。

[0174] 6. 可通过音频内容控制FDN参数。音频分类算法或手动注释的内容可指示音频段是否包含语音、音乐、声音效果、静音等。可根据这种标签调整FDN参数。例如,可对于对话减少直接与混响比,以改善对话可理解性。另外,可以使用视频分析以确定当前视频段的位置,并且,FDN参数可相应地被调整以更接近地仿真在视频中描述的环境;和/或

[0175] 7. 固态回放系统可使用与移动装置不同的FDN设定,例如,设定可以是与装置相关的。存在于起居室中的固态系统可仿真具有远隔的源的典型(相当混响)起居室方案,而移动装置可呈现更接近收听者的内容。

[0176] 本发明的虚拟化器的一些实现包含被配置为应用分数延迟以及整数采样延迟的FDN(例如,图4的FDN的实现)。例如,在一个这种实现中,分数延迟元件在各混响箱中与应用等于采样周期的整数的整数延迟的延迟线串联连接(例如,各分数延迟元件被定位在延迟线中的一个之后或者另外与其串联)。可通过与采样周期的分数对应的各频带中的相位偏移(单位复数乘法)来近似分数延迟。这里, f 是延迟分数, τ 是频带的希望的延迟, T 是频带的采样周期。在QMF域中应用混响的上下文中如何应用分数延迟是公知的。

[0177] 在第一类的实施例中,本发明是一种用于响应多通道音频输入信号的一组通道(例如,通道中的每一个或者全频率范围通道中的每一个)产生双耳信号的耳机虚拟化方法,包括以下的步骤:(a)向该组通道中的各通道应用双耳房间脉冲响应(BRIR)(例如,在图3的子系统100和200中,或者在图2的子系统12、...、14和15中,通过将该组通道中的各通道与和所述通道对应的BRIR进行卷积),由此产生经滤波的信号(例如,图3的子系统100和200的输出,或者图2的子系统12、...、14和15的输出),包含通过使用至少一个反馈延迟网络(例如,图3的FDN 203、204、...、205)以向该组通道中的通道的下混(例如,单音下混)应用公共晚期混响;和(b)组合经滤波的信号(例如,在图3的子系统210或图2的包含元件16和18的子系统)以产生双耳信号。典型地,FDN群被用于向下混应用公共晚期混响(例如,各FDN向不同的频带应用公共晚期混响)。典型地,步骤(a)包含向该组通道中的各通道应用该通道的单通道BRIR的“直接响应和早期反射”部分(例如,在图3的子系统100或图2的子系统12、...、14中)的步骤,并且,公共晚期混响被产生以模仿单通道BRIR中的至少一些(例如,全部)的晚期混响部分的共同宏观属性。

[0178] 在第一类的典型实施例中,在混合复正交镜像滤波器(HCQMF)域或正交镜像滤波器(QMF)域中实现FDN中的每一个,并且,在一些这种实施例中,通过控制用于应用晚期混响的各FDN的配置,控制双耳信号的频率相关空间声学属性(例如,使用图3的子系统209)。典型地,为了实现多通道信号的音频内容的高效双耳呈现,通道的单音下混(例如,由图3的子系统201产生的下混)被用作FDN的输入。典型地,下混处理基于各通道的源距离(即,通道的音频内容的假定源与假定的用户位置之间的距离)被控制并且依赖于与源距离对应的直接响应的处理,以便保留各BRIR的时间和水平结构(即,由一个通道的单通道BRIR的直接响应和早期反射部分确定的各BRIR,连同包含该通道的下混的公共晚期混响)。虽然要下混的通道可在下混期间以不同的方式时间对准和缩放,但用于各通道的BRIR的直接响应、早期反射和公共晚期混响部分之间的适当的水平和时间关系应得到保持。在使用单个FDN群以产生用于被进行下混(以产生下混)的所有通道的公共晚期混响部分的实施例中,需要在下混产生的过程中(向被进行下混的各通道)应用适当的增益和延迟。

[0179] 这类的典型实施例包括调整(例如,使用图3的控制子系统209)与频率相关属性(例如,混响衰减时间、耳间相干性、模态密度和直接与晚期比)对应的FDN系数的步骤。这使得能够实现声学环境的更好的匹配和更自然的发声输出。

[0180] 在第二类的实施例中,本发明是一种用于响应多通道音频输入信号通过向输入信号的一组通道中的各通道(例如,输入信号的通道中的每一个通道或输入信号的全频率

范围通道)应用双耳房间脉冲响应(BRIR)(例如,将各通道与相应的BRIR进行卷积)以产生双耳信号的方法,包括:在(例如,通过图3的子系统100或图2的子系统12、...、14实现的)第一处理路径中处理该组通道中的各通道,该第一处理路径被配置为模型化并向所述各通道应用该通道的单通道BRIR的直接响应和早期反射部分(例如,通过图2的子系统12、14或15应用的EBRIR);以及在与第一处理路径并行的(例如,通过图3的子系统200或图2的子系统15实现的)第二处理路径中处理该组通道中的通道的下混(例如,单音下混)。第二处理路径被配置为模型化并向该下混应用公共晚期混响(例如,通过图2的子系统15应用的LBRIR)。典型地,公共晚期混响模仿单通道BRIR中的至少一些(例如,全部)的晚期混响部分的共同宏观属性。典型地,第二处理路径包含至少一个FDN(例如,对于多个频带的每一个使用一个FDN)。典型地,单声道下混被用作由第二处理路径实现的各FDN的所有混响箱的输入。典型地,为了更好地仿真声学环境并产生更自然的发声双耳虚拟化,设置用于各FDN的宏观属性的系统控制的机构(例如,图3的控制子系统209)。由于大多数这种宏观属性是依赖于频率的,因此,典型地在混合复正交镜像滤波器(HCQMF)域、频域、域或另一滤波器组域中实现各FDN,并且,对于各频带使用不同的FDN。在滤波器组域中实现FDN的主要益处是允许应用具有频率相关的混响性能的混响。在各种实施例中,通过使用各种滤波器组(包括但不限于正交镜像滤波器(QMF)、有限脉冲响应滤波器(FIR滤波器)、无限脉冲响应滤波器(IIR滤波器)或交叠滤波器)中的任一种,在各种滤波器组域的任一个中实现FDN。

[0181] 第一类(和第二类)的一些实施例实现以下特征中的一个或多个:

[0182] 1. 滤波器组域(例如,混合复正交镜像滤波器域)FDN实现(例如,图4的FDN实现)或混合滤波器组域FDN实现和时域晚期混响滤波器实现(例如,参照图8描述的结构),其例如通过提供在不同的带中改变混响箱衰变以便作为频率的函数改变模态密度的能力,典型地允许独立调整各频带的FDN的参数和/或设定(这使得能够简单灵活地控制频率相关声学属性);

[0183] 2. 特定下混处理,其被用于(从多通道输入音频信号)产生在第二处理路径中处理的下混(例如,单音下混)信号,依赖于各通道的源距离和直接响应的处理,以便在直接和晚期响应之间保持适当的水平和定时关系。

[0184] 3. 在第二处理路径中(例如,在FDN群的输入或输出处)应用全通滤波器(例如,图4的APF 301),以在不改变得到的混响的波谱和/或音色的情况下引入相位差异和增大的回声密度;

[0185] 4. 在复值、多比率结构中在各FDN的反馈路径中实现分数延迟,以克服与被量化为下采样因子网格的延迟有关的问题;

[0186] 5. 在FDN中,通过使用基于各频带中的希望的耳间相干性设定的输出混合系数,混响箱输出直接线性混合到双耳通道中(例如,通过图4的矩阵312)。可选地,混响箱到双耳输出通道的映射跨着频带交替,以在双耳通道之间实现平衡延迟。还可选地,向混响箱输出应用归一化因子以在保留分数延迟和总功率的同时均匀化它们的水平;

[0187] 6. 通过设定各频带中的增益与混响箱延迟的适当组合来(例如,通过使用图3的控制子系统209)控制依赖于频率的混响衰变时间,以模拟真实房间;

[0188] 7. (例如,在相关处理路径的输入或输出处)对于每个频带(例如,通过图4的元件306和309)应用一个标度因子,以完成以下过程:

[0189] 控制与真实房间匹配的频率相关直接与晚期比 (DLR) (可使用简单模型以基于目标 DLR 和例如为 T60 的混响衰减时间计算需要的标度因子) ;

[0190] 提供低频衰减以减少过量的组合伪信号;和/或

[0191] 向 FDN 响应应用扩散场谱整形;

[0192] 8. (例如,通过图3的控制子系统209) 实现用于控制诸如混响衰变时间、耳间相干性和/或直接与晚期比的晚期混响的基本频率相关属性的简单的参数模型。

[0193] 在一些实施例(例如,对于其中系统延滞是关键的和由分析和合成滤波器组导致的延迟被禁止的应用)中,本发明的系统的典型实施例的滤波器组域 FDN 结构(例如,每个频带中的图4的 FDN) 被在时域中实现的 FDN 结构(例如,图10的 FDN 220,其可如图9中所示地实现) 替代。在本发明的系统的时域实施例中,为了允许依赖频率的控制,应用输入增益因子 (G_{in})、混响箱增益 (g_i) 和归一化增益 ($1/|g_i|$) 的滤波器组域实施例的子系统被时域滤波器 (和/或增益元件) 替代。典型滤波器组域实现的输出混合矩阵(例如,图4的输出混合矩阵 312) 被(在典型时域实施例中) 时域滤波器的输出集合(例如,图9的元件424的图11实现的元件500至503) 替代。不同于典型时域实施例的其它滤波器,滤波器的此输出集合的相位响应典型地是关键(这是因为功率守恒和耳间相关性可能受相位响应影响)。在一些时域实施例中,混响箱延迟相对于它们的在对应的滤波器组域实现中的值改变(例如,稍微改变), (例如,以避免共享作为共用因子的滤波器组步幅)。

[0194] 除了图3的系统的元件202-207在图10的系统中被在时域中实现的单个 FDN 220 替代(例如,图10的 FDN 220 可如同图9的 FDN 那样被实现) 之外,图10是类似于图3的本发明的耳机虚拟化系统的实施例的框图。在图10中,两个(左通道和右通道)时域信号被从直接响应和早期反射处理系统100输出,并且两个(左通道和右通道)时域信号被从晚期混响处理系统221输出。加算元件210被耦接到子系统100和200的输出。元件210被配置为组合(混合) 子系统100和221的左通道输出以产生从图10的虚拟化器输出的双耳音频信号的左通道L, 并且组合(混合) 子系统100和221的右通道输出以产生从图10的虚拟化器输出的双耳音频信号的右通道R。假定在子系统100和221中实现了适当的水平调整和时间对准,元件210可被实现为简单地合计从子系统100和221输出的对应的左通道采样以产生双耳输出信号的左通道,并且简单地合计从子系统100和221输出的对应的右通道采样以产生双耳输出信号的右通道。

[0195] 在图10的系统中,多通道音频输入信号(具有通道 X_i) 被引向两个并行处理路径并在其中经受处理:一个处理路径通过直接响应和早期反射处理子系统100;另一个处理路径通过晚期混响处理子系统200。图10系统被配置为向各通道 X_i 应用 $BRIR_i$ 。各 $BRIR_i$ 可分解成两个部分:直接响应和早期反射部分(通过子系统100被应用) 和晚期混响部分(通过子系统221被应用)。在操作中,直接响应和早期反射处理子系统100由此产生从虚拟化器输出的双耳音频信号的直接响应和早期反射部分,并且,晚期混响处理子系统(“晚期混响产生器”) 221由此产生从虚拟化器输出的双耳音频信号的晚期混响部分。子系统100和221的输出(通过子系统210) 被混合以产生典型地从子系统210向呈现系统(未示出) 断言的双耳音频信号,在该呈现系统中,该信号经受双耳呈现以供耳机回放。

[0196] (晚期混响处理子系统221的) 下混子系统201被配置为将多通道输入信号的通道下混为单声道下混(其是时域信号),并且 FDN 220 被配置为将晚期混响部分应用于该单声

道下混。

[0197] 参照图9,接下来描述可用作图10的虚拟化器的FDN 220的时域FDN的示例。图9的FDN包括输入滤波器400,该输入滤波器400被耦接以接收多通道音频输入信号的所有通道的单声道下混(例如,由图10系统的子系统201产生)。图9的FDN还包括耦接到滤波器400的输出的全通滤波器(APF) 401(对应于图4的APF 301),耦接到滤波器401的输出的输入增益元件401A,耦接到滤波器401的输出的加算元件402、403、404和405(对应于图4的加算元件302、303、304和305),以及四个混响箱。每个混响箱耦接到元件402、403、404和405中的不同的一个元件的输出,并且包括混响滤波器406和406A、407和407A、408和408A以及409和409A之一、与之耦接的延迟线410、411、412和413之一(对应于图4的延迟线307),以及耦接到延迟线之一的输出的增益元件417、418、419和420之一。

[0198] 酉矩阵415(对应于图4的酉矩阵308并且典型地实现为与酉矩阵308相同)被耦接至延迟线410、411、412和413的输出。矩阵415被配置为将反馈输出断言至元件402、403、404和405中的每一个的第二输入。

[0199] 当通过线410施加的延迟(n_1)短于通过线411施加的延迟(n_2),通过线411施加的延迟短于通过线412施加的延迟(n_3),以及通过线412施加的延迟短于通过线413施加的延迟(n_4)时,(第一和第三混响箱的)增益元件417和419的输出被断言至加算元件422的输入,并且(第二和第四混响箱的)增益元件418和420的输出被断言至加算元件423的输入。元件422的输出被断言至IACC和混合滤波器424的一个输入,并且元件423的输出被断言至IACC滤波和混合级424的另一个输入。

[0200] 将参照图4的元件310和311以及输出混合矩阵312的典型实现来描述图9的增益元件417~420以及元件422、423和424的实现的示例。图4的输出混合矩阵312(还被标识为矩阵 M_{out})是 2×2 矩阵,其被配置为对来自初始扫调的未混合双耳通道(分别是元件310和311的输出)进行混合,以产生具有希望的耳间相干性的左和右双耳输出通道(在矩阵312的输出处被断言的左耳“L”以及右耳“R”信号)。初始扫调由元件310和311实现,元件310和311中的每一个组合两个混响箱输出以产生未混合双耳通道之一,其中具有最短延迟的混响箱输出被断言至元件310的输入,并且具有次最短延迟的混响箱输出被断言至元件311的输入。图9实施例的元件422和423(对于被断言至它们的输入的时域信号)执行与图4实施例的(每一频带中的)元件310和311对被断言至它们的输入的(在相关频带中的)滤波器组域成分的流所执行的初始扫调相同类型的初始扫调。

[0201] (从图4的元件310和322或者图9的元件422和423输出的)未混合双耳通道(由于它们不包含任何公共的混响箱输出而接近于不相关)可(通过图4的矩阵312或者图9的级424)被混合,以实现获得左和右双耳输出通道的希望的耳间相干性的扫调图案。但是,由于混响箱延迟在各FDN(即,图9的FDN或者图4中的对于各不同频带实现的FDN)中不同,一个未混合双耳通道(元件310和311或者422和423之一的输出)总是领先于另一未混合双耳通道(元件310和311或者422和423中的另一个的输出)。

[0202] 因此,在图4的实施例中,如果混响箱延迟与扫调图案的组合对于所有频带而言都是相同,则将得到声音图像偏差(sound image bias)。如果扫调图案跨频带交替以使得混合的双耳输出通道在交替频带中相互领先和尾随,则此偏差被减轻。例如,如果希望的耳间相干性为 C_{oh} (其中, $|C_{oh}| \leq 1$),则在被奇数编号的频带中的输出混合矩阵312可被实现为将

向其断言的两个输入乘以具有以下形式的矩阵：

$$[0203] \quad M_{out} = \begin{bmatrix} \cos \beta & \sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}, \quad \text{其中 } \beta = \arcsin(\text{Coh}) / 2$$

[0204] 并且,在被偶数编号的频带中的输出混合矩阵312可被实现为将向其断言的两个输入乘以具有以下形式的矩阵：

$$[0205] \quad M_{out,alt} = \begin{bmatrix} \sin \beta & \cos \beta \\ \cos \beta & \sin \beta \end{bmatrix}$$

[0206] 其中 $\beta = \arcsin(\text{Coh}) / 2$.

[0207] 作为替代,在矩阵312输入的通道顺序对于交替频带被切换(例如,在奇数频带中,元件310的输出可被断言至矩阵312的第一输入并且元件311的输出可被断言至矩阵312的第二输入,而在偶数频带中,元件311的输出可被断言至矩阵312的第一输入并且元件310的输出可被断言至矩阵312的第二输入)的情况下,通过将矩阵312实现为在对于所有频带的FDN中相同,上文提及双耳输出通道中的声音图像偏差可被减轻。

[0208] 在图9的实施例(以及本发明的系统的FDN的其它时域实施例)中,有意义地是基于频率交替扫调以解决声音图像偏差,否则在从元件422输出的未混合双耳通道总是领先于(或者滞后于)从元件423输出的未混合双耳通道时会出现该声音图像偏差。此声音图像偏差在本发明的系统的FDN的典型时域实施例中以与典型地在本发明的系统的FDN的滤波器组域实施例中的解决方式不同的方式被解决。具体而言,在图9的实施例(以及本发明系统的FDN的一些其他时域实施例中),未混合双耳通道(例如,从图9的元件422和423输出的那些)的相对增益由增益元件(例如,图9的元件417、418、419和420)确定,以便补偿否则将由于显著的不平衡定时而导致的聲音图像偏差。通过实现用以衰减最早到达的信号(已例如通过元件422被扫调至一侧)的增益元件(例如,元件417)并且实现用以增强次最早到达的信号(已例如通过元件423被扫调至另一侧)的增益元件(例如,元件418),立体声信号被重新居中。因此,包含增益元件417的混响箱向元件417的输出应用第一增益,并且包含增益元件418的混响箱向元件418的输出应用第二增益(不同于第一增益),从而第一增益和第二增益使(从元件422输出的)第一未混合双耳通道相对于(从元件423输出的)第二未混合双耳通道衰减。

[0209] 更具体而言,在图9的FDN的典型实现中,四个延迟线410、411、412和413具有增加的长度,分别具有延迟值 n_1 、 n_2 、 n_3 和 n_4 。在此实现中,滤波器417再次应用增益 g_1 。由此,滤波器417的输出是已被应用了增益 g_1 的延迟线410的输入的延迟版本。类似地,滤波器418应用增益 g_2 ,滤波器419应用增益 g_3 ,并且滤波器420应用增益 g_4 。因此,滤波器418的输出是已被应用了增益 g_2 的延迟线411的输入的延迟版本,滤波器419的输出是已被应用了增益 g_3 的延迟线412的输入的延迟版本,并且滤波器420的输出是已被应用了增益 g_4 的延迟线413的输入的延迟版本。

[0210] 在此实现中,以下增益值的选择导致了(由从元件424输出的双耳通道指示的)输出声音图像到一侧(即,到左侧通道或右侧通道)的不希望的偏差: $g_1=0.5$, $g_2=0.5$, $g_3=0.5$,以及 $g_4=0.5$ 。根据本发明的实施例,(分别由元件417、418、419和420应用的)增益值 g_1 、 g_2 、 g_3 、 g_4 被如下地选择以便使声音图像居中: $g_1=0.38$, $g_2=0.6$, $g_3=0.5$,以及 $g_4=0.5$ 。因

此,根据本发明的实施例,通过使(在此示例中已通过元件422被扫调至一侧的)最早到达的信号相对于次最早到达的信号衰减(例如,通过选择 $g_1 < g_3$),并且通过使(在此示例中已通过元件423被扫调至另一侧的)次最早到达的信号相对于最新到达的信号增强(例如,通过选择 $g_4 < g_2$),输出立体声图像被重新居中。

[0211] 图9的时域FDN的典型实现与图4的滤波器组域(CQMF域)FDN具有以下差别和相似性:

[0212] 相同的酉反馈矩阵, A(图4的矩阵308和图9的矩阵415);

[0213] 相似的混响箱延迟, n_i (即,图4的CQMF实现中的延迟可以是 $n_1 = 17 * 64 T_s = 1088 * T_s$, $n_2 = 21 * 64 T_s = 1344 * T_s$, $n_3 = 26 * 64 T_s = 1664 * T_s$, 并且 $n_4 = 29 * 64 T_s = 1856 * T_s$, 这里 $1/T_s$ 是采样率($1/T_s$ 典型地等于48KHz),而在时域实现中的延迟可以是 $n_1 = 1089 * T_s$, $n_2 = 1345 * T_s$, $n_3 = 1663 * T_s$, 以及 $n_4 = 185 * T_s$ 。应指出,在典型CQMF实现中,存在如下实际约束:各延迟是64个采样的块的持续时间的某一整数倍(采样率典型地为48KHz),但是在时域中,对于各延迟的选择更加灵活,因此对于各混响箱的延迟的选择更加灵活);

[0214] 类似的全通滤波器实现(即,图4的滤波器301和图9的滤波器401的类似实现)。例如,全通滤波器可通过级联数个(例如,三个)全通滤波器来实现。例如,每一被级联的全通滤波器可具有形式

[0215] $\frac{g - Z^{-n_i}}{1 - g * Z^{-n_i}}$, 其中 $g = 0.6$ 。图4的全通滤波器301可由具有合适的采样块延迟(例如,

如, $n_1 = 64 * T_s$, $n_2 = 128 * T_s$, 以及 $n_3 = 196 * T_s$)的三个级联的全通滤波器实现,而图9的全通滤波器401(时域全通滤波器)可由具有相似延迟(例如, $n_1 = 61 * T_s$, $n_2 = 127 * T_s$, 以及 $n_3 = 191 * T_s$)的三个级联的全通滤波器实现。

[0216] 在图9的时域FDN的一些实现中,输入滤波器400被实现为使得其使得要由图9的系统应用的BRIR的直接与晚期比(DLR)(至少基本上)匹配目标DLR,并且使得要通过包含图9的系统的虚拟化器(例如,图10的虚拟化器)应用的BRIR的DLR可通过替换滤波器400(或者控制滤波器400的配置)而被改变。例如,在一些实施例中,滤波器400被实现为滤波器(例如,如图9A所示地耦接的第一滤波器400A和第二滤波器400B)的级联以实现目标DLR并且可选地还实现希望的DLR控制。例如,级联的滤波器是IIR滤波器(例如,滤波器400A是被配置为匹配目标低频特性的一阶ButterWorth高通滤波器(IIR滤波器),并且滤波器400B是被配置为匹配目标高频特性的二阶低架IIR滤波器)。对于另一示例,级联的滤波器是IIR和FIR滤波器(例如,滤波器400A是被配置为匹配目标低频特性的二阶ButterWorth高通滤波器(IIR滤波器),并且滤波器400B是被配置为匹配目标高频特性的十四阶FIR滤波器)。典型地,直接信号是固定的,并且滤波器400对晚期信号进行修正以实现目标DLR。全通滤波器(APF)401优选地被实现为执行如图4的APF 301所执行的功能相同的功能,即引入相位差异和增大的回声强度以产生更自然的发声FDN输出。APF 401典型地控制相位响应,而输入滤波器400控制振幅响应。

[0217] 在图9中,滤波器406和增益元件406A一起实现混响滤波器,滤波器407和增益元件407A一起实现另一个混响滤波器,滤波器408和增益元件408A一起实现另一混响滤波器,并且滤波器409和增益元件409A一起实现还另一混响滤波器。图9的滤波器406、407、408和409中的每一个优选地被实现为具有接近1(单位增益)的最大增益值的滤波器,并且增益元件

406A、407A、408A和409A中的每一个被配置为向滤波器406、407、408和409中对应的一个滤波器的输出应用衰变增益,其匹配希望的衰变(在相关的混响箱延迟 n_i 之后)。具体而言,增益元件406A被配置为向滤波器406的输出应用衰变增益(衰变增益 $_1$)以使得元件406A的输出具有使得(在混响箱延迟 n_1 之后的)延迟线410的输出具有第一目标衰变增益的增益,增益元件407A被配置为向滤波器407的输出应用衰变增益(衰变增益 $_2$)以使得元件407A的输出具有使得(在混响箱延迟 n_2 之后的)延迟线411的输出具有第二目标衰变增益的增益,增益元件408A被配置为向滤波器408的输出应用衰变增益(衰变增益 $_3$)以使得元件408A的输出具有使得(在混响箱延迟 n_3 之后的)延迟线412的输出具有第三目标衰变增益的增益,并且增益元件409A被配置为向滤波器409的输出应用衰变增益(衰变增益 $_4$)以使得元件409A的输出具有使得(在混响箱延迟 n_4 之后的)延迟线413的输出具有第四目标衰变增益的增益。

[0218] 图9的系统的滤波器406、407、408和409中的每一个以及元件406A、407A、408A和409A中的每一个优选地被实现为(其中,滤波器406、407、408和409中的每一个被实现为IIR滤波器,例如,架式型滤波器或者架式型滤波器的级联)实现要由包含图9的系统的虚拟化器(例如,图10的虚拟化器)应用的BRIR的目标T60特性,这里“T60”指示混响衰变时间(T_{60})。例如,在一些实施例中,滤波器406、407、408和409中的每一个被实现为架式型滤波器(例如,具有 $Q=0.3$ 以及500Hz的架频率(shelf frequency)的架式型滤波器,以实现图13中所示的T60特性,其中T60的单位为秒),或者两个IIR架式型滤波器的级联(例如,具有100Hz和1000Hz的架频率,以实现图14中所示的T60特性,其中T60的单位为秒)。各架式型滤波器的形状被确定为匹配希望的从低频到高频的改变曲线。当滤波器406被实现为架式型滤波器(或者架式型滤波器的级联)时,包含滤波器406和增益元件406A的混响滤波器也是架式型滤波器(或者架式型滤波器的级联)。同样,当滤波器407、408和409中的每一个被实现为架式型滤波器(或者架式型滤波器的级联)时,包含滤波器407(408或409)和对应的增益元件(407A、408A或409A)的各混响滤波器也是架式型滤波器(或者架式型滤波器的级联)。图9B是被实现为如图9B中所示地被耦接的第一架式型滤波器406B和第二架式型滤波器406C的级联的滤波器406的示例。滤波器407、408和409中的每一个可如滤波器406的图9实现那样被实现。

[0219] 在一些实施例中,元件406A、407A、408A和409A所应用的衰变延迟(衰变增益 n_i)如下地被确定:

[0220] 衰变增益 $_i = 10^{((-60 * (n_i / F_s) / T) / 20)}$

[0221] 这里, i 是混响箱索引(即,元件406A应用衰变增益 $_1$,元件407A应用衰变增益 $_2$,等等), n_i 是第 i 混响箱的延迟(例如 n_1 是通过延迟线410应用的延迟), F_s 是采样率, T 是在希望的低频的所希望的混响衰变时间(T_{60})。

[0222] 图11是图9的以下元件的实施例的框图:元件422和423以及IACC(耳间互相关系数)滤波和混合级424。元件422被耦接和配置为合计(图9的)滤波器417和419的输出并且将合计的信号断言至低架滤波器500的输入,并且元件423被耦接和配置为合计(图9的)滤波器418和420的输出并且将合计的信号断言至高通滤波器501的输入。滤波器500和501的输出被在元件502中合计(混合)以产生双耳左耳输出信号,并且滤波器500和501的输出被在元件502中混合(从滤波器501的输出减去滤波器500的输出)以产生双耳右耳输出信号。元

件502和503对滤波器500和501的经滤波输出进行混合(合计和相减)以产生双耳输出信号,该信号实现(在可接受的精度内的)目标IACC特性。在图11的实施例中,低架滤波器500和高通滤波器510中的每一个典型地被实现为一阶IIR滤波器。在滤波器500和501具有这样的实现的示例中,图11的实施例可实现在图12中被绘制为曲线“ I ”的示例性的IACC特性,其与在图12中被绘制为“ I_T ”的目标IACC特性良好匹配。

[0223] 图11A是图11的滤波器500的典型实现的频率响应(R1)、图11的滤波器501的典型实现的频率响应(R2)以及并行连接的滤波器500和501的响应的曲线图。从图11A中清楚可见,组合的响应希望地在范围100Hz~10,000Hz上是平坦的。

[0224] 因此,在一类实施例中,本发明是一种用于响应多通道音频输入信号的一组通道产生双耳信号(例如,图10的元件210的输出)的系统(例如图10的系统)和方法,包括向该组通道中的每一通道应用双耳房间脉冲响应(BRIR),由此产生经滤波的信号,包括使用单个反馈延迟网络(FDN)以向该组通道中的通道的下混应用公共晚期混响;并且组合经滤波器的信号以产生双耳信号。FDN在时域中实现。在一些这样的实施例中,时域FDN(例如,如图9中那样配置的图10的FDN 220)包括:

[0225] 输入滤波器(例如,图9的滤波器400),具有被耦接以接收该下混的输入,其中该输入滤波器被配置为响应该下混产生第一经滤波的下混;

[0226] 全通滤波器(例如,图9的全通滤波器401),被耦接并被配置为响应该第一经滤波的下混产生第二经滤波的下混;

[0227] 混响应用子系统(例如,图9的除元件400、401和424之外的所有元件),具有第一输出(例如,元件422的输出)和第二输出(例如,元件423的输出),其中,该混响应用子系统包括一组混响箱,每一混响箱具有不同的延迟,并且其中混响应用子系统被耦接并配置为响应第二经滤波的下混产生第一未混合双耳通道和第二未混合双耳通道,在第一输出处断言第一未混合双耳通道并且在第二输出处断言第二未混合双耳通道;以及

[0228] 耳间互相关系数(IACC)滤波和混合级(例如,图9的级424,可被实现为图11的元件500、501、502和503),被耦接到该混响应用子系统,并且被配置为响应第一未混合双耳通道和第二未混合双耳通道产生第一混合双耳通道和第二混合双耳通道。

[0229] 输入滤波器可被实现以产生(优选地,被实现为两个滤波器的级联,被配置为产生)第一经滤波的下混,使得每个BRIR具有至少基本上匹配目标直接与晚期比(DLR)的直接与晚期比(DLR)。

[0230] 每个混响箱可被配置为产生延迟信号,并且可包括混响滤波器(例如,被实现为架滤波器或架滤波器的级联),该混响滤波器被耦接并被配置为向在所述每个混响箱中传播的信号应用增益,使得该延迟信号具有至少基本上匹配用于所述延迟信号的目标衰变增益的增益,以致于实现每个BRIR的目标混响衰变时间特性(例如, T_{60} 特性)。

[0231] 在一些实施例中,第一未混合双耳通道领先于第二未混合双耳通道,混响箱包括被配置为产生具有最短延迟的第一延迟信号的第一混响箱(例如,图9的包括延迟线410的混响箱)和被配置为产生具有次最短延迟的第二延迟信号的第二混响箱(例如,图9的包括延迟线411的混响箱),其中第一混响箱被配置为向第一延迟信号应用第一增益,第二混响箱被配置为向第二延迟信号应用第二增益,第二增益与第一增益不同,并且第一增益和第二增益的应用导致第一未混合双耳通道相对于第二未混合双耳通道衰减。典型地,第一混

合双耳通道和第二混合双耳通道指示被重新居中的立体声图像。在一些实施例中, IACC滤波和混合级被配置为产生第一混合双耳通道和第二混合双耳通道, 使得所述第一混合双耳通道和第二混合双耳通道具有至少基本上匹配目标IACC特性的IACC特性。

[0232] 本发明的多个方面包括执行(或被配置为执行或支持执行) 音频信号(例如, 其音频内容包含扬声器通道的音频信号和/或基于对象的音频信号) 的双耳虚拟化的方法和系统(例如, 图2的系统20或者图3或图10的系统)。

[0233] 在一些实施例中, 本发明的虚拟化器为或者包含被耦接以接收或产生指示多通道音频输入信号的输入数据并且通过软件(或固件) 被编程并且/或者另外被配置为(例如, 响应控制数据) 对输入数据执行包括本发明的方法实施例的各种操作中的任一种的通用处理器。这种通用处理器典型地会与输入装置(例如, 鼠标和/或键盘)、存储器和显示装置耦接。例如, 可在通用处理器中实现图3系统(或图2的系统20或包含系统20的元件12、...、14、15、16和18的虚拟化器系统), 其中输入是指示音频输入信号的N个通道的音频数据, 输出是指示双耳音频信号的两个通道的音频数据。常规的数字模拟转换器(DAC) 可对输出数据操作, 以产生用于供扬声器(例如, 一对耳机) 再现的双耳信号通道的模拟版本。

[0234] 虽然这里描述了本发明的具体实施例和本发明的应用, 但本领域技术人员可以理解, 在不背离这里描述和要求权利的本发明的范围的情况下, 这里描述的实施例和应用的许多变化是可能的。应当理解, 虽然表示和描述了本发明的某些形式, 但本发明不限于描述和表示的特定实施例或描述的特定的方法。

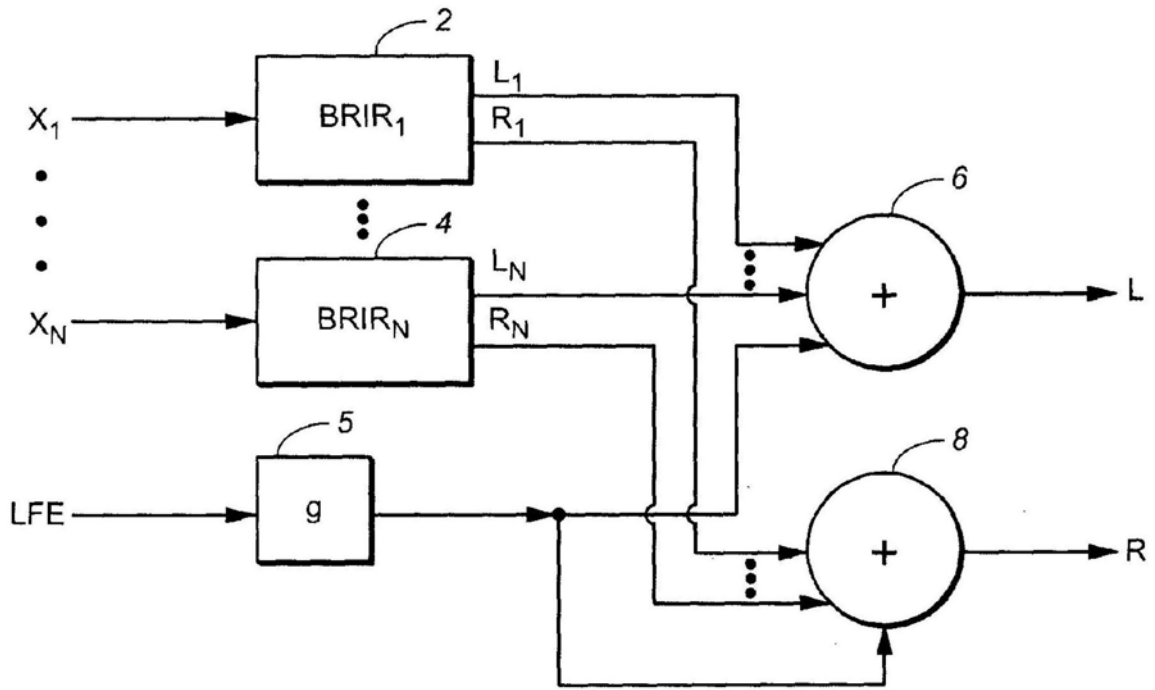


图1 (现有技术)

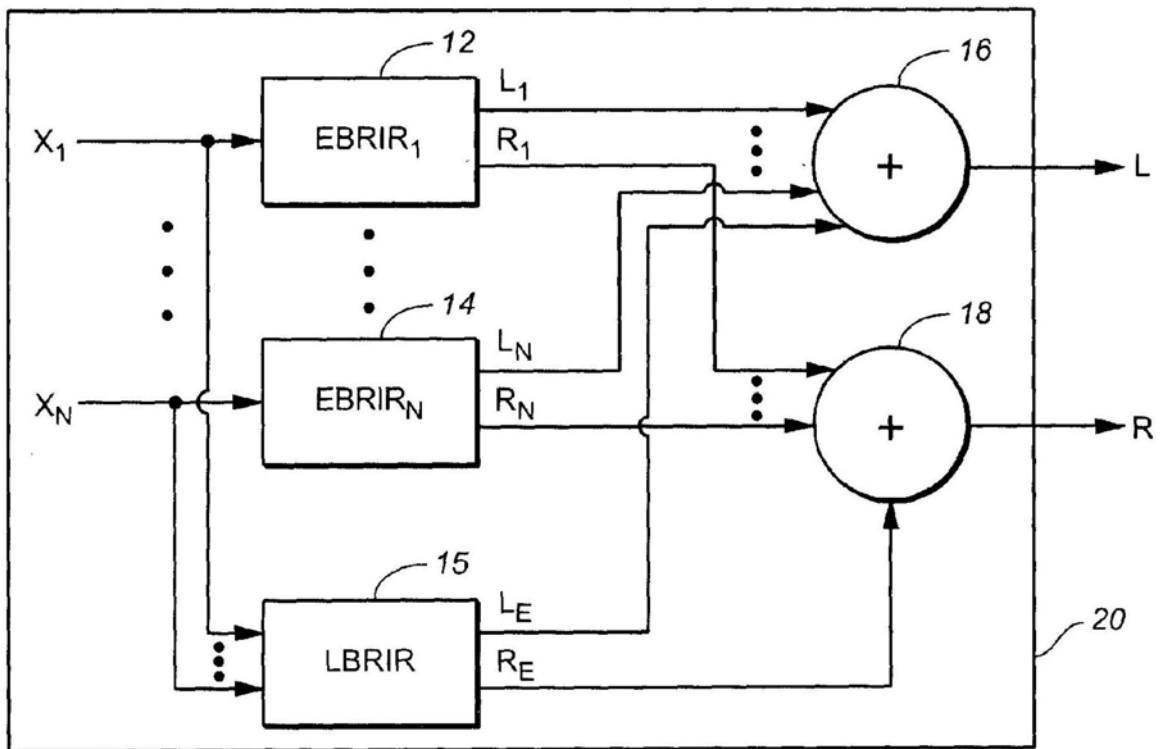


图2

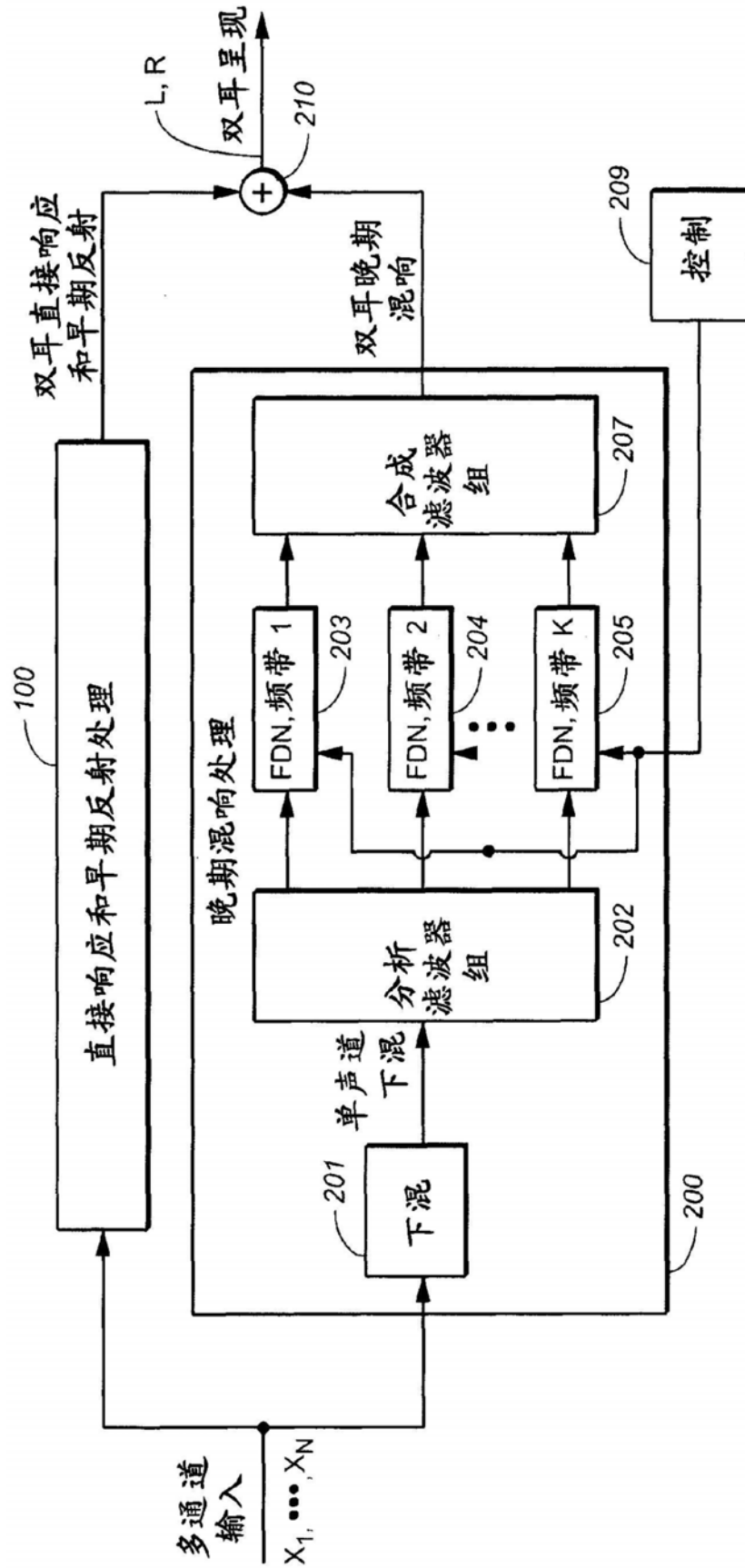


图3

对于各频带

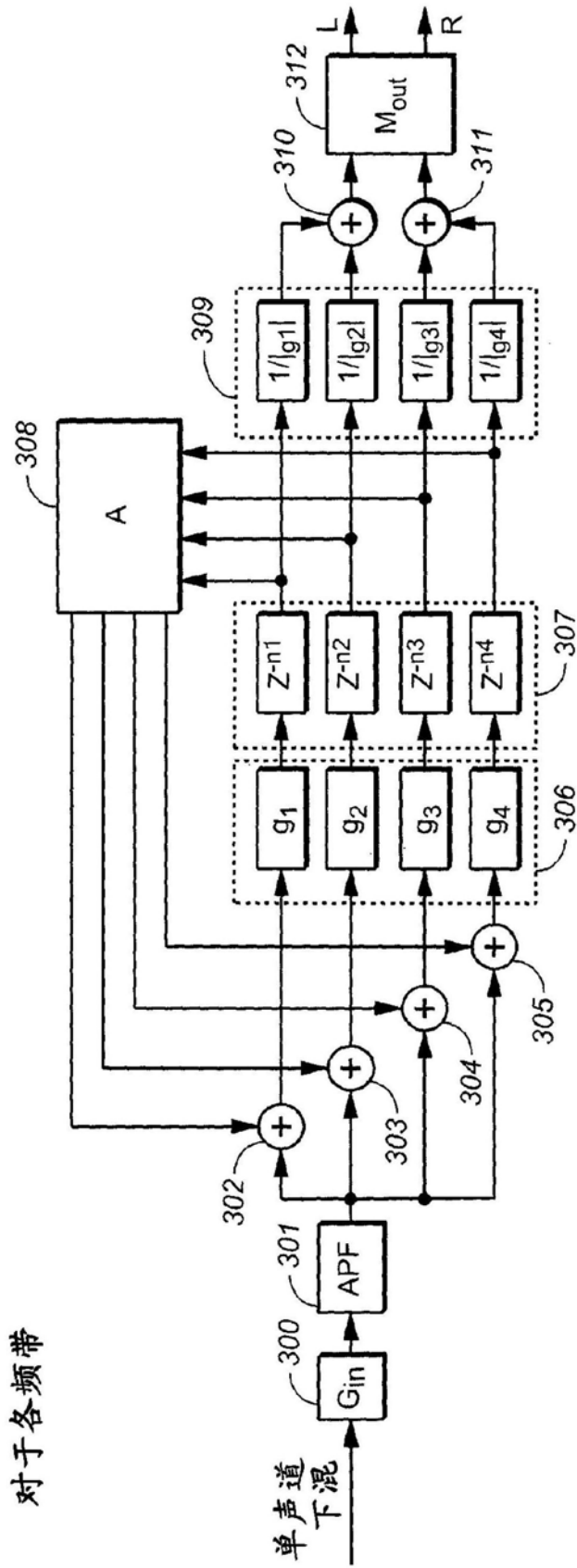


图4

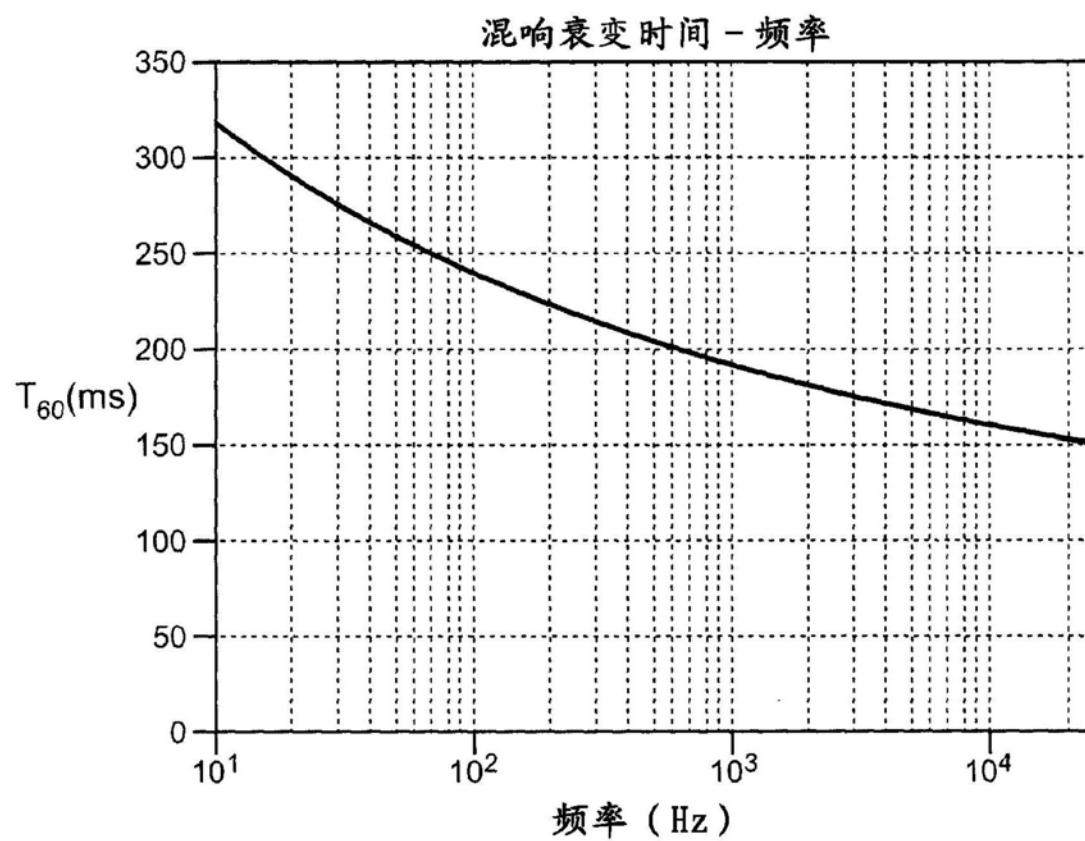


图5

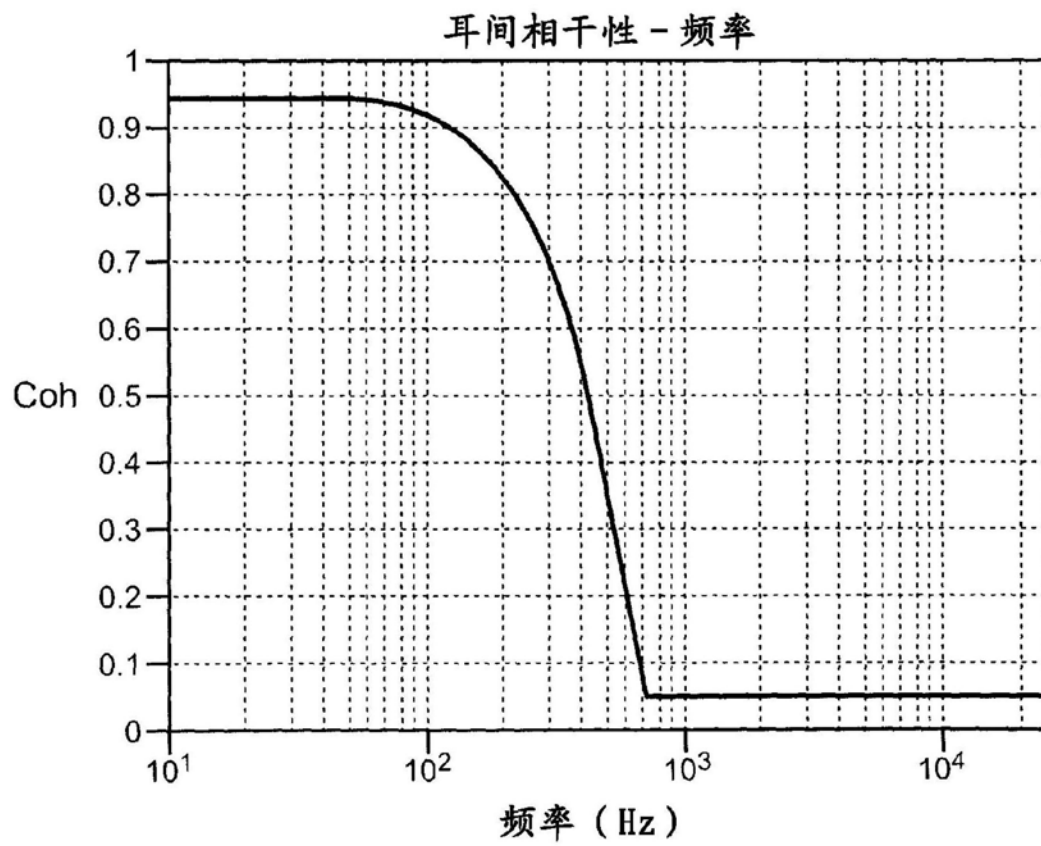


图6

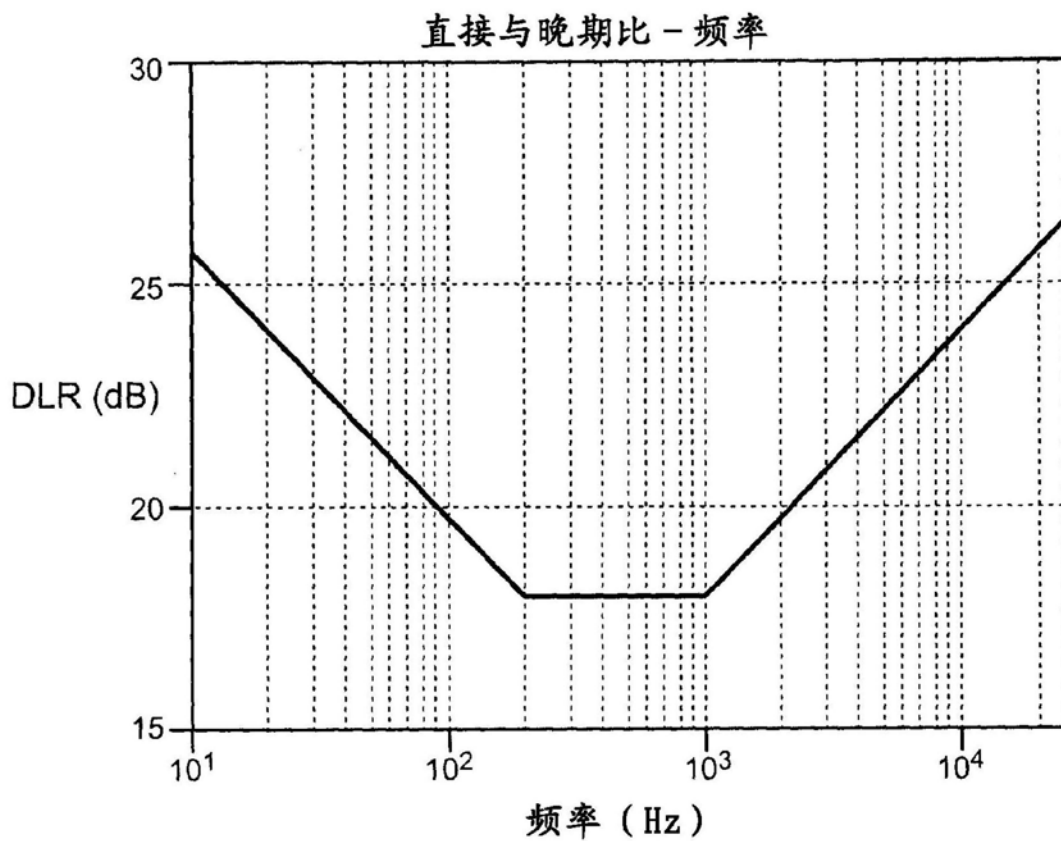


图7

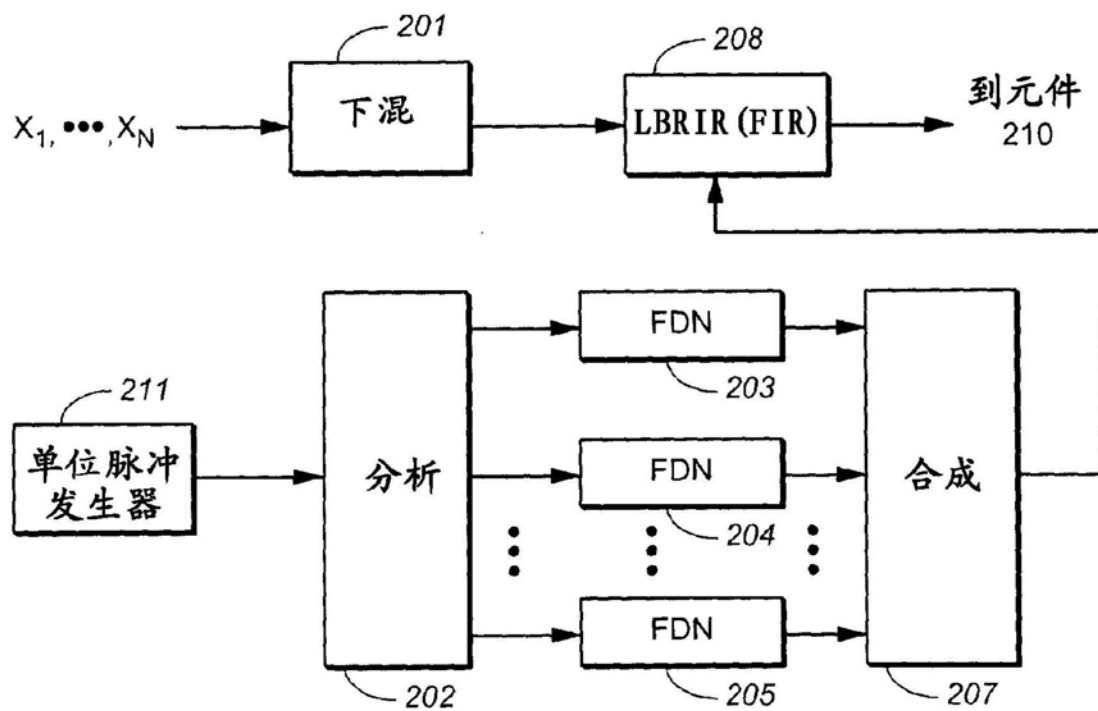


图8

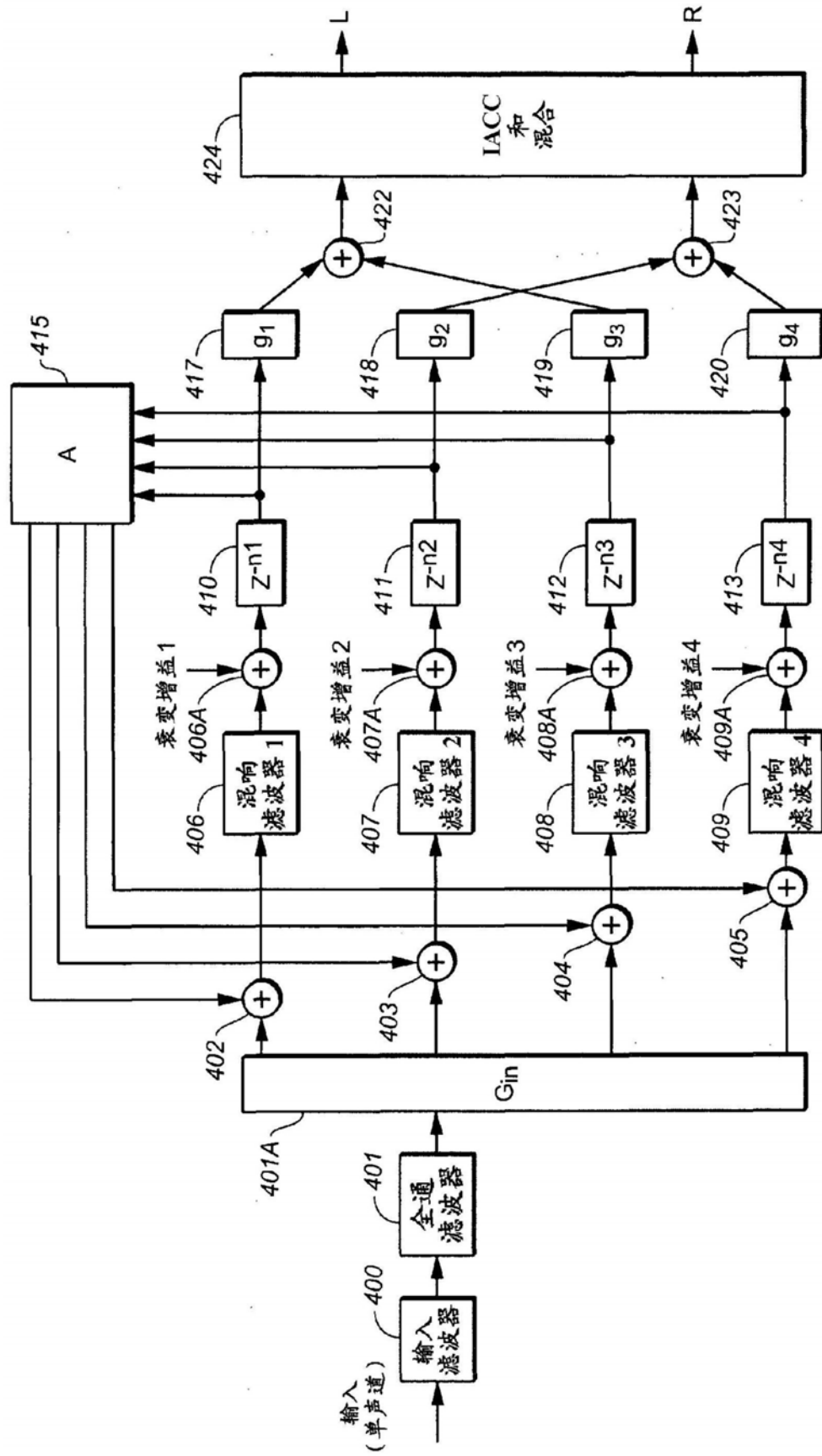


图9

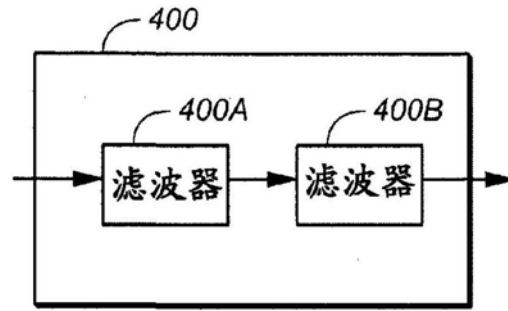


图9A

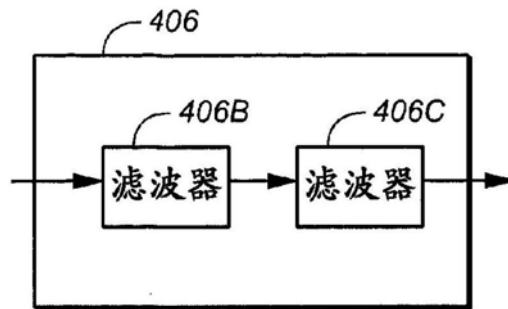


图9B

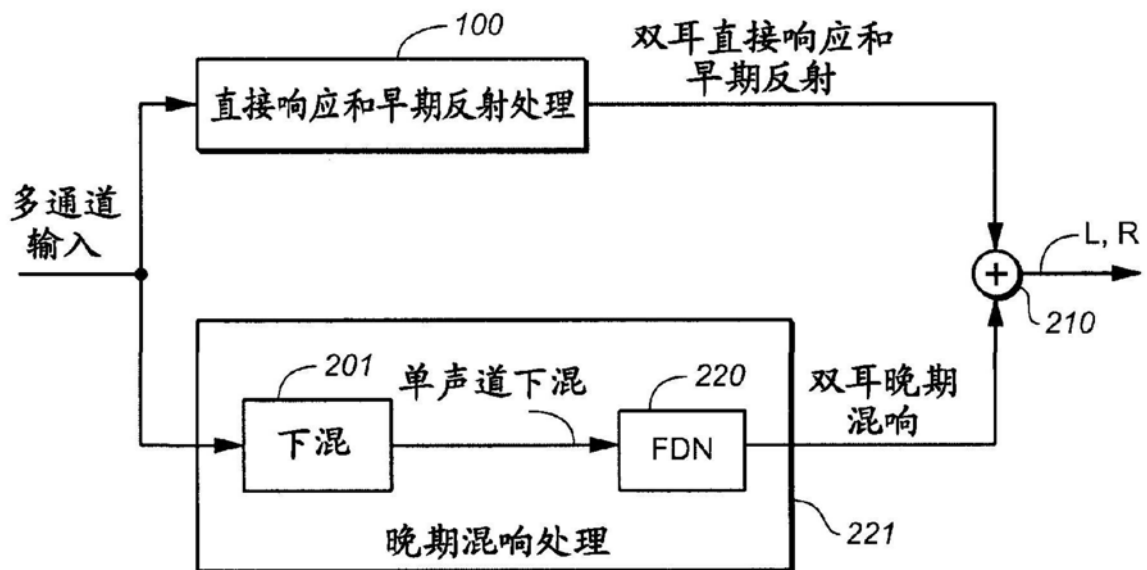


图10

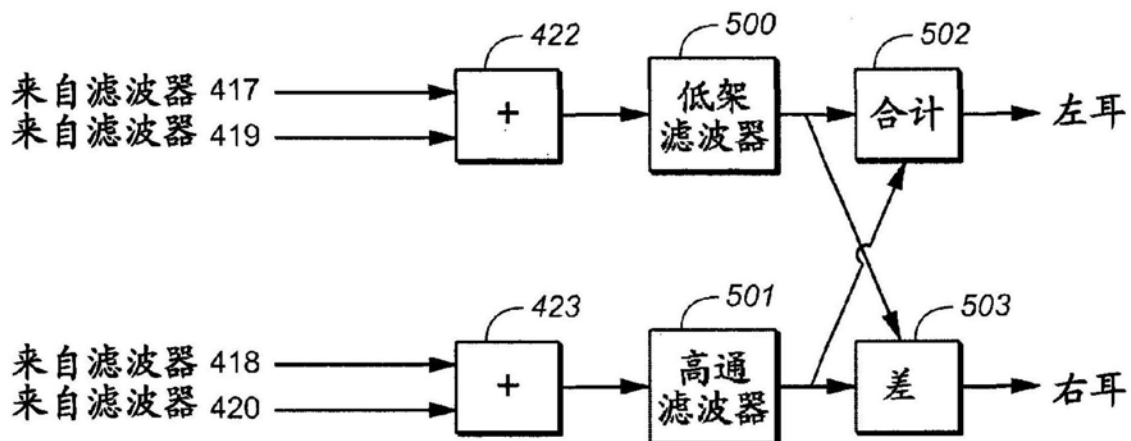


图11

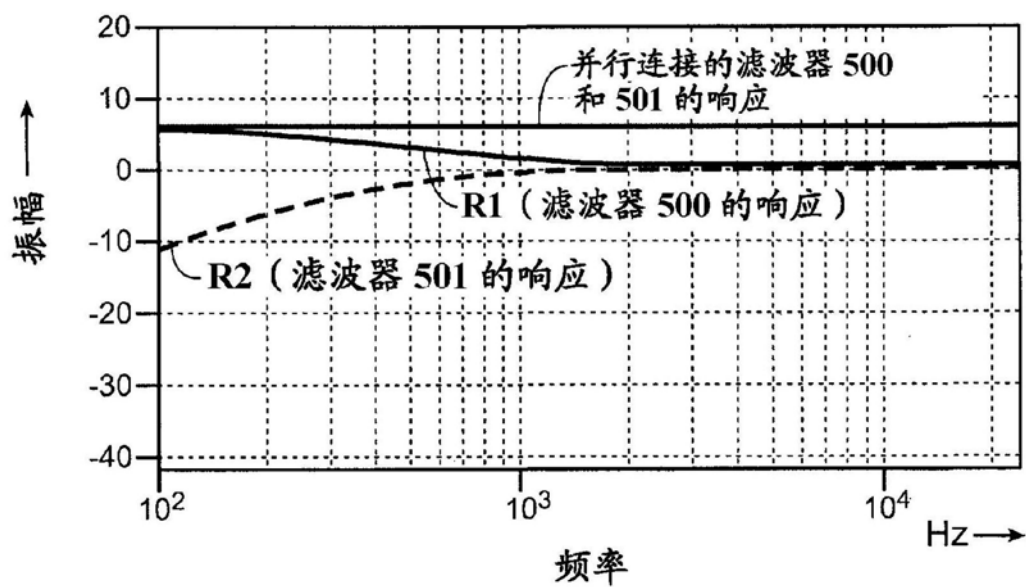


图11A

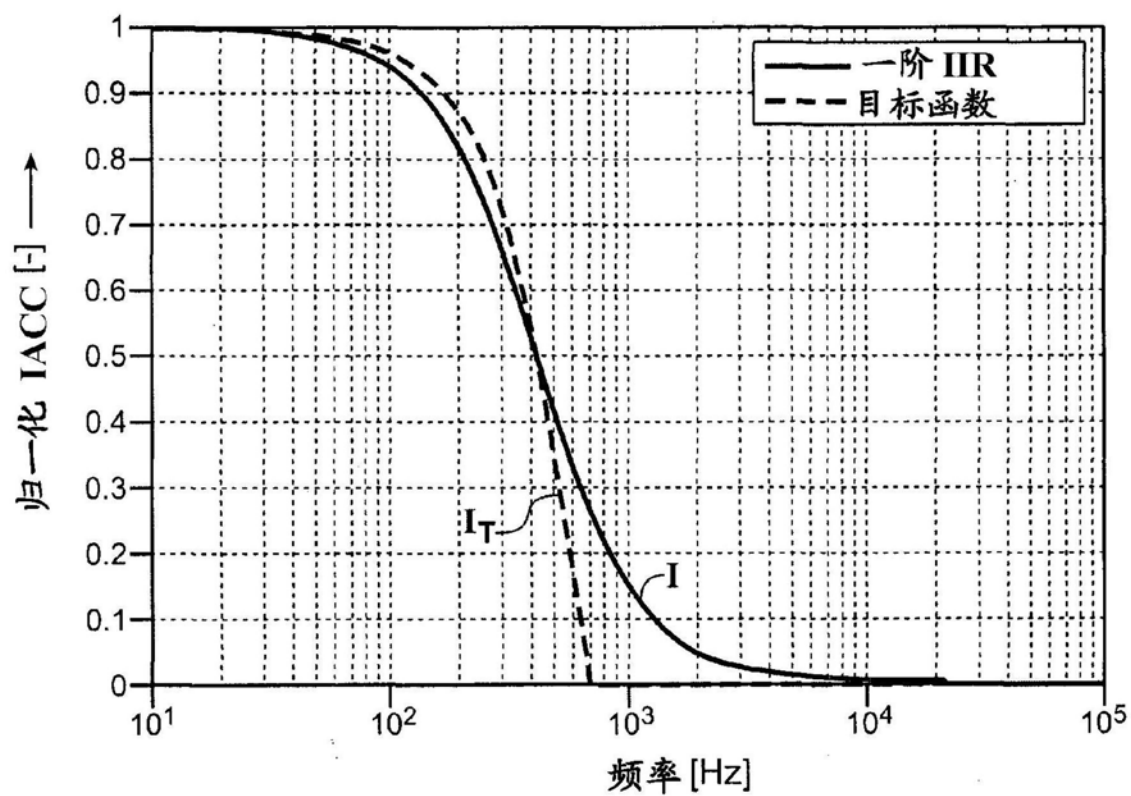


图12

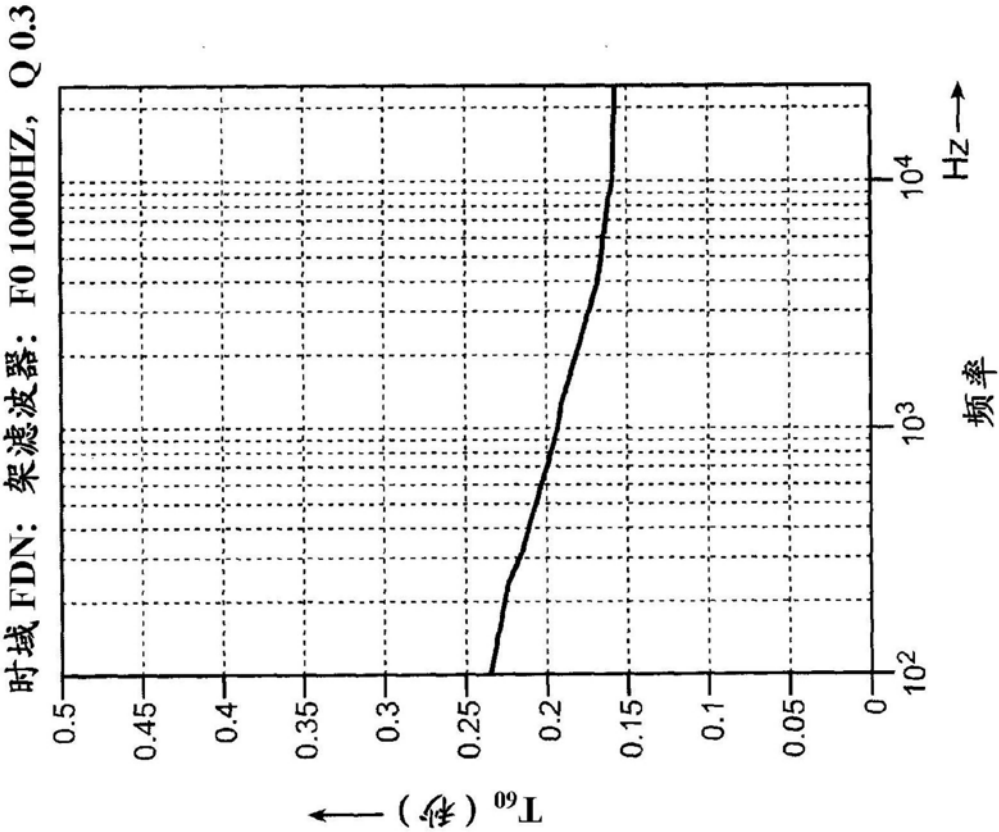


图13

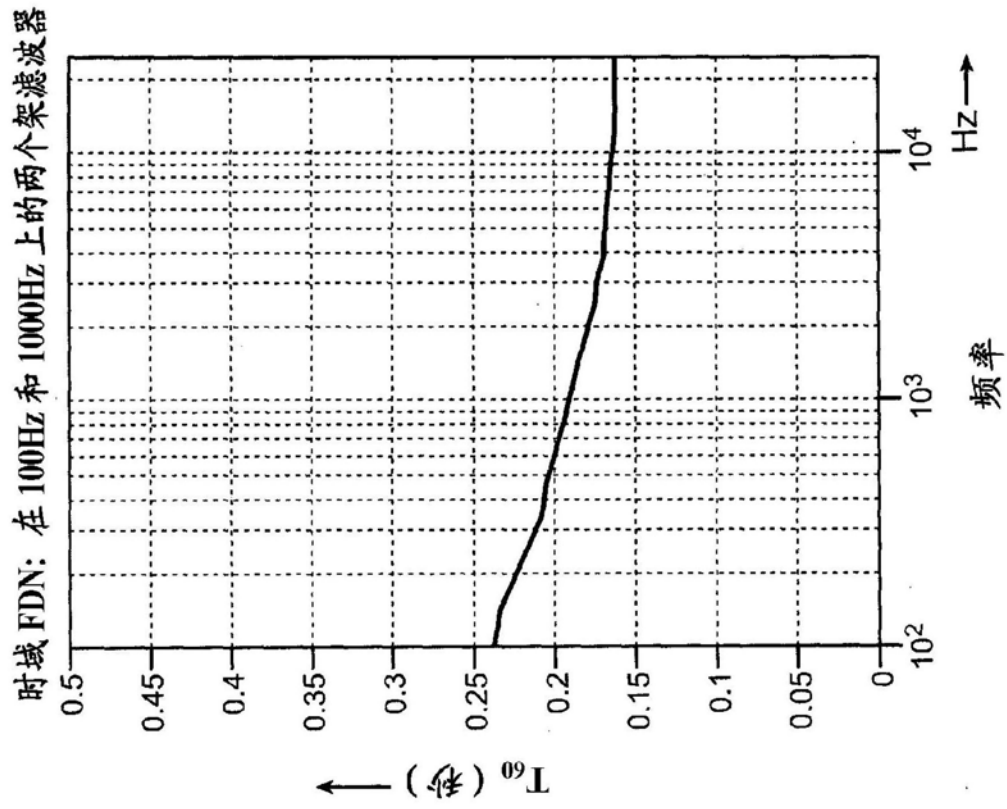


图14