



(12)发明专利

(10)授权公告号 CN 105706162 B

(45)授权公告日 2019.06.11

(21)申请号 201480060448.0

(22)申请日 2014.10.14

(65)同一申请的已公布的文献号
申请公布号 CN 105706162 A

(43)申请公布日 2016.06.22

(30)优先权数据
1360185 2013.10.18 FR

(85)PCT国际申请进入国家阶段日
2016.05.04

(86)PCT国际申请的申请数据
PCT/FR2014/052617 2014.10.14

(87)PCT国际申请的公布数据
W02015/055946 FR 2015.04.23

(73)专利权人 奥兰吉公司
地址 法国巴黎

(72)发明人 格雷戈里·帕洛内
马克·埃梅里特

(74)专利代理机构 上海天协和诚知识产权代理
事务所 31216

代理人 张轶

(51)Int.Cl.
G10K 15/12(2006.01)
H04S 1/00(2006.01)
H04S 3/00(2006.01)
H04S 7/00(2006.01)

(56)对比文件
CN 102165798 A,2011.08.24,
US 2008085008 A1,2008.04.10,
JP 2011083018 A,2011.04.21,
US 2012201389 A1,2012.08.09,
CN 102694517 A,2012.09.26,
徐宁等.基于状态空间模型的子频带语音转
换算法.《电子学报》.2010,第38卷(第3期),

审查员 薛彦杰

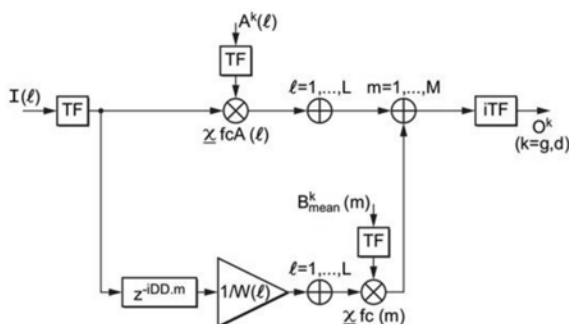
权利要求书3页 说明书10页 附图5页

(54)发明名称

具有复杂度优化的房间效果的声音空间化

(57)摘要

本发明涉及声音的空间化,它通过将具有房间效果的至少一个传递函数应用于至少一个声音信号。这一应用相当于在频谱范围内将声音信号的频谱分量与对应于该传递函数的滤波器的频谱分量相乘,且该滤波器的各个频谱分量在时间频率表示中具有时间演变。具体的说,对于上述分量的相乘,可以忽略该滤波器超出阈值频率($F_c^d(1)$ 、 $F_c^g(1)$ 、 $F_c^d(2)$ 、 $F_c^g(2)$)并在所述时间频率表示中在至少一个给定瞬间($m=1, m=2$)之后的频谱分量。



1. 适用于声音空间化的方法,包括将具有房间效果的至少一个传递函数应用于至少一个声音信号,所述应用相当于在频谱范围中将该声音信号的频谱分量与对应于所述传递函数的滤波器的频谱分量相乘,该滤波器的每一个频谱分量在时间频率表示中具有时间演变,

其中针对分量的所述相乘,忽略该滤波器超过阈值频率且在所述时间频率表示中在至少一个给定瞬间之后的所述频谱分量,并且其中,通过接收多个输入信号和提供至少两个输出信号的声音空间化模块的一种实施方式,将具有房间效果的传递函数应用于每个输入信号,以便提供各个输出信号,通过应用下述类型的公式给出所述各个输出信号:

$$O^k = \sum_{l=1}^L (I(l) *_{[0; \dots; f^k(l)]} A^k(l)) + \sum_{m=1}^M \left(z^{-iDDm} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) \right) *_{[0; \dots; f^k(m)]} B_{mean}^k(m)$$

O^k 是输出信号,并且k是与一个输出信号相关的索引,

$l \in [1; L]$ 是与所述输入信号中的一个输入信号相关的索引,L是输入信号的数量,并且I(l)是所述输入信号中的一个输入信号,

$A^k(l)$ 是具有房间效果的传递函数,专用于一个输入信号,

$B_{mean}^k(m)$ 是具有房间效果的通用的传递函数,为输入信号所共用,

$W^k(l)$ 是所选择的加权因子,并且G(I(l))是预先确定的功率补偿增益,

z^{-iDDm} 是延迟的应用,被认为采样的块的数量,对应于在与房间效果对应的在房间中的声音的发散与所述房间中混响场的出现开始之时之间的时间差,索引m对应于与这一延迟对应的持续时间中的多个采样的块,M为在时间频率表示中传递函数持续的块的总数量,

符号“.”表示乘积,

术语“ $*_{[0; \dots; f^k(l)]}$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至最高频率 $f^k(l)$,它是至少一个索引为1的输入信号的函数,以及,

术语“ $*_{[0; \dots; f^k(m)]}$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至频率 $f^k(m)$,它是索引为m的采样的块的函数。

2. 根据权利要求1所述的方法,其特征在于,所述阈值频率在时间频率表示中随着时间递减。

3. 根据权利要求1所述的方法,其特征在于,获取与声音信号中最高频率的频谱分量相关的信息并且其中所述阈值频率是预先确定的阈值频率和所述最高频率之间的最小值。

4. 根据权利要求3所述的方法,其特征在于,所述声音信号源自压缩解码器并且与最高频率的频谱分量相关的信息可由解码器提供。

5. 根据权利要求3所述的方法,其特征在于,以给定的采样频率采样声音信号并基于所述采样频率选择所述阈值频率。

6. 根据权利要求1所述的方法,其特征在于,在分别与第一和第二声道相关的至少第一和第二虚拟扬声器上对声音信号进行空间化,并且将具有房间效果的第一和第二传递函数

分别应用于所述第一和第二声道，

第一和第二传递函数中的其中一个应用同侧声路径效果而第一和第二传递函数中的另一个应用对侧声路径效果，同时消除超过给定筛选频率的声音信号的频谱分量，

并且其中针对应用对侧路径效果的传递函数的所述阈值频率是预先定义的阈值频率和所述筛选频率之间的最小值。

7. 根据权利要求1所述的方法，其特征在于，所述信号包括在信号间具有相同大小的连续的采样块，并且其中至少一个给定的瞬间在时间上位于与一系列块中的第一块不同的一个块的开始；这一给定的瞬间因此发生在直接传播之后并且位于声音反射的时刻或者混响场出现开始的时刻。

8. 一种非暂态的计算机可读存储介质，其上存储有可执行程序，其中该程序命令微处理器执行根据前述任一权利要求所述的方法。

9. 声音空间化模块，包括适用于将具有房间效果的至少一个传递函数应用于至少一个输入声音信号的处理器，所述应用相当于在频谱范围中将声音信号的频谱分量与对应于所述传递函数的滤波器的频谱分量相乘，该滤波器的每一个频谱分量具有在时间频率表示中的时间演变，

其特征在于，该处理器被配置为对于分量的所述相乘，忽略该滤波器超出阈值频率并且在所述时间频率表示中至少一个给定瞬间之后的所述频谱分量，并且所述声音空间化模块，接收多个输入信号，提供至少两个输出信号，该处理器被配置为将具有房间效果的传递函数应用于每一个输入信号，通过应用下述类型的公式给出所述各个输出信号：

$$O^k = \sum_{l=1}^L (I(l) *_{[0; \dots; f^k(l)]} A^k(l)) + \sum_{m=1}^M \left(z^{-iDDm} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) \right) *_{[0; \dots; f^k(m)]} B_{mean}^k(m)$$

O^k 是输出信号，并且k是与一个输出信号相关的索引，

$l \in [1; L]$ 是与所述输入信号中的一个输入信号相关的索引，L是输入信号的数量，并且I(l)是所述输入信号中的一个输入信号，

$A^k(l)$ 是具有房间效果的传递函数，专用于一个输入信号，

$B_{mean}^k(m)$ 是具有房间效果的通用的传递函数，为输入信号所共用，

$W^k(l)$ 是所选择的加权因子，并且G(I(l))是预先确定的功率补偿增益，

z^{-iDDm} 是延迟的应用，被认为采样的块的数量，对应于在与房间效果对应的在房间中的声音的发散与所述房间中混响场的出现开始之时之间的时间差，索引m对应于与这一延迟对应的持续时间中的多个采样的块，M为在时间频率表示中传递函数持续的块的总数量，

符号“.”表示乘积，

术语“ $*_{[0; \dots; f^k(l)]}$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至最高频率 $f^k(l)$ ，它是至少一个索引为l的输入信号的函数，以及，

术语“ $*_{[0; \dots; f^k(m)]}$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至频率

$f^k(m)$,它是索引为 m 的采样的块的函数。

具有复杂度优化的房间效果的声音空间化

技术领域

[0001] 本发明涉及具有房间效果的声音空间化。

背景技术

[0002] 本发明在分别从与虚拟扬声器 (例如在要呈现的声音的多声道表示或者环绕立体声表示中) 相关的L个声道发出的声音信号的处理中找到一种有益但非限制的应用,以便用于在真正扬声器 (例如,在双耳耳机中的两个听筒或者在反耳 (transaural) 中的两个分离的扬声器) 上进行空间化的渲染。

[0003] 举例来说,来自这些声道中的一个的信号可在双耳中处理为在左听筒具有第一贡献和在右听筒具有第二贡献,尤其是通过将具有房间效果的传递函数应用于这些贡献的每一个。这些具有房间效果的传递函数的应用有助于为收听者提供沉浸式的感觉,犹如将与那个声道相关的虚拟扬声器可相对于收听者所“放置”。

[0004] 在一个特定实施例中,具体如文档FR1357229所描述,可以一种BRIR (“双耳房间脉冲响应”) 类型的脉冲响应的形式,将在时域中具有房间效果的传递函数应用于相应声道的每一个声音信号。具体的说,在通过引用将其结合在本文中的那一文档中,BRIR传递函数被构建为下面两个函数的组合:

[0005] -第一传递函数,为每个信号所专用,以及,

[0006] -第二通用传递函数,为所有信号所共用并且具体以混响声场为特征,后者的出现通常在某段时间后发生在房间中,一般是在声波的第一反射之后。

[0007] 这一实施例有利地允许应用所有信号所共用的处理,它实际上物理地对应于当混响发生时声波的“混合”,因此在某段时间后 (特征化混响声场的出现的开始)。这一实施例减少了在多个初始声道上的具有房间效果的空间化处理的复杂度。

[0008] 然而,在具有发生在渲染之前的渲染的模块中,期望能进一步最小化空间化处理的复杂度。作为一个非限制性示例,由压缩解码器以编码的形式接收多声道的信号。一旦解码之后,此解码器就将声道的信号发送至用于在两个扬声器上渲染具有房间效果声音的空间化模块。然后,就希望能减少在此空间化步骤中的处理 (它在接收的信号解码之后) 的复杂度,使之在渲染之前接收信号时所有的解码和空间化步骤不会变慢。

发明内容

[0009] 本发明旨在改进该状况。

[0010] 为此,本发明提出减少具有房间效果的传递函数的应用的复杂度,尤其是通过在频谱范围内减少这一复杂度。在频谱范围内,通过表示传递函数的滤波器 (参考将在下文进一步详细描述图1),将传递函数的卷积变为信号的频谱分量的乘积。

[0011] 本发明基于有利的观察,即在直接传播后,声波在高频段趋于衰减,这是因为吸收声音的表面上的渐进反射 (特别是墙壁、收听者的脸等),尤其是高频。此外,在声音的传播中,空气本身也会吸收声音中的最高频的频谱分量。例如,对于混响声场来说,这一现象会

进一步增加,因此,对它来说,就不需要对非常高的频率(例如高于5至15kHz的频率范围)进行频率表示。

[0012] 于是,就有可能在应用具有房间效果的传递函数时减少处理的复杂度,在频谱范围中,当乘以前述频谱分量时,简化为不再考虑与高于预先确定的截止频率的频率相关的分量(例如,高于5至15kHz)。

[0013] 本发明因此关注一种适用于声音空间化的方法,包括将具有房间效果的至少一个传递函数应用于至少一个声音信号,所述应用相当于在频谱范围中将该声音信号的频谱分量与对应于所述传递函数的滤波器的频谱分量相乘。该滤波器的每一个频谱分量都具有在时间频率表示中的时间演变(将参考图3进一步描述)。

[0014] 具体的说,在所述时间频率表示中,对于上述分量的乘积来说,可以忽略超出阈值频率并且在至少一个给定的瞬间之后的该滤波器的这些频域分量。因此,在此给定的瞬间之后,该滤波器的频谱分量被考虑直至例如可被选择为5和15kHz之间的一截止频率(取决于要应用的房间效果和/或取决于要空间化的信号,如下文所述)。在该截止频率之外,甚至不再执行该乘法,它在数学上等同于该信号乘以零。

[0015] 这一给定的瞬间一般可表示为当声波开始经历混响的时刻(通过连续的反射,或者稍后,从混响声场的出现)。于是,一般来说,在其中该传递函数考虑房间效果中的混响(例如,考虑该混响声场)的实施例中,所述给定的瞬间可选择为这些混响的函数。举例来说,在房间效果混响中,所述给定的瞬间可在具有初始反射的直接声音传播之后,并因此对应于混响声场的出现开始之时。

[0016] 此外,可提供一种实施例,其中在所述时间频率表示中,上述阈值频率随着时间而降低。举例来说,如果信号是在数个连续时间块中进行采样的,对于第一块,在分量的乘积中,它可能被安排为诸如保留该信号中出现的频谱分量,然后对于在第一块之后的第二块忽略超出第一阈值频率的那些,然后对于在第二块之后的第三块忽略超出第二阈值频率的那些,等等;其中第二阈值频率低于第一阈值频率。

[0017] 因此,一般来说,在其中信号以多个连续块进行采样的实施例中,对于分量的乘积来说,可以忽略该滤波器的下述频谱分量:

[0018] -对于给定块,超出第一阈值频率的,

[0019] -然后,对于该给定块之后的块,超出第二阈值频率的,该第二阈值频率低于第一阈值频率。

[0020] 所述给定块可包括,举例来说,在时间上处于对应于当声波已经经历一个或多个反射时刻甚至在混响声场的出现开始之时的时间采样。在所述给定块之后(紧挨着或者相隔数个块)的块可包括,举例来说,在时间上处于混响声场的出现开始之后或者适于该开始的采样。

[0021] 这一实施例允许,举例来说,从在用于混响的高频声音衰减中减少有可能听得见的人工效果,可以数个块来逐渐实现这一实施例。还允许考虑具有混响声场特征的多种形式的传递函数(下文表示为 $B_{mean}^k(m)$,其中m为块的索引)。例如,有可能将传递函数 B_{mean}^k 应用于所述给定的块,并针对接下来的块将时间上逐渐截止的窗口(“淡出”类型的窗口)应用于这一传递函数 B_{mean}^k ,以便“结束”该混响声场的出现。

[0022] 在其中该方法由声音空间化模块所实施的实施例中,该声音空间化模块接收多个

输入信号并提供至少两个输出信号,将具有房间效果的传递函数应用于各个输入信号,以便提供各个输出信号,所述输出信号各自通过应用下述类型的公式给出:

$$\begin{aligned}
 [0023] \quad O^k &= \sum_{l=1}^L (I(l) *_{[0;\dots;f^k(l)]} A^k(l)) \\
 &+ \sum_{m=1}^M \left(z^{-iDDm} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) \right) *_{[0;\dots;f^k(m)]} B_{mean}^k(m)
 \end{aligned}$$

[0024] O^k 是输出信号,并且k是与一个输出信号相关的索引,

[0025] $l \in [1;L]$ 是与所述输入信号中的一个输入信号相关的索引,L是输入信号的数量,并且 $I(l)$ 是所述输入信号中的一个输入信号,

[0026] $A^k(l)$ 是具有房间效果的传递函数,专用于一个输入信号,

[0027] $B_{mean}^k(m)$ 是具有房间效果的通用的传递函数,为输入信号所共用,

[0028] $W^k(l)$ 是所选择的加权因子,并且 $G(I(l))$ 是预先确定的功率补偿增益,

[0029] z^{-iDDm} 是延迟的应用,被认为采样的块的数量,对应于在与房间效果对应的在房间中的声音的分散与所述房间中混响场的出现开始之时的时间差,索引m对应于与这一延迟对应的持续时间中的多个采样的块,M为在时间频率表示中传递函数持续的块的总数量,

[0030] 符号“.”表示乘积,

[0031] 术语“ $*_{[0;\dots;f^k(l)]}$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至最高频率 $f^k(l)$,它是至少一个索引为1的输入信号的函数,以及,

[0032] 术语“ $*_{[0;\dots;f^k(m)]}$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至频率 $f^k(m)$,它是索引为m的采样的块的函数。

[0033] 将在下文具体参考图2和5对这一实施例进行详细描述。

[0034] 还可将相乘计算限制在第一阈值频率之外,适于样本的第一块或多个块开始,基于信号的特征(例如,它的采样频率或者在信号的频谱分量中表示的最高频率)或者基于应用的空间特征(例如,针对如下文详述的对侧声路径限制高频分量)。

[0035] 在这种情况下,来自混响的信号(在反射后或在混响场中)通常并不包括高于初始信号的频率的频谱分量。上文提到的阈值频率因此并不高于此最高频率。

[0036] 更具体的说,在一个实施例中,获取与声音信号中最高频率的频谱分量相关的信息并且上述阈值频率被选择为预先确定的阈值频率(例如,在5和15kHz之间)和所述最高频率之间的最小值。

[0037] 典型的是,在一个实施例中,其中声音信号源自压缩解码器,与最高频率的频谱分量相关的信息可以由解码器提供。

[0038] 相似地,如果在能够支持不同信号格式的模块中执行该空间化,尤其是在这些信号的采样频率方面,所述最高频率不能高于采样频率的一半,因此还可基于该采样频率来选择用于实现本发明的阈值频率。

[0039] 在一个实施例中,其中在分别与第一和第二声音声道相关的至少第一和第二虚拟扬声器上对声音信号进行空间化,具有房间效果的第一和第二传递函数被分别应用于所述第一和第二声道,如上文在前言中所解释的那样(例如通过适配环绕立体声声道上的信号

以切换至双耳或反耳的渲染)。具体的说,在其中第一和第二传递函数中的其中一个应用于同侧声路径效果而第一和第二传递函数中的另一个应用于对侧声路径效果的情况下,可以提供对于超过给定筛选频率的声音信号的频谱分量的消除。这一“筛选”频率通过下述事实来解释,对于在虚拟扬声器和收听者的相关耳朵之间的对侧路径上,该收听者的头部位于该声音路径中并吸收声波的较高音质(因此消除了与该声波的较高频率相关的频谱分量)。因此,对于应用对侧路径效果的传递函数,所述阈值频率可被选择为预先定义的阈值频率(例如在5和15kHz之间选择)和所述筛选频率之间的最小值。这一实施例当甚至应用于采样的第一块时是有利的。然而,这并不排除对于下一块再次提高阈值频率的可能性,以模拟面对讨论中的耳朵的墙上的第一反射,这一第一反射通过同侧路径被那只耳朵所接收。

[0040] 不管怎样,应当理解该截止频率可能被选择为所有信号所共用,在一个可能的实施例中,在例如对应于混响场的出现开始所给定的瞬间之后。

[0041] 因此,在上文介绍的文档FR13 57299中描述的实施例在本发明的上下文中可为有利的,尤其是如果应用于信号的每一个传递函数包括:

[0042] -专用于此信号的传递函数,加上,

[0043] -通用的传递函数,为所有信号所共用并且表示混响场的出现,

[0044] 那么,所述给定的瞬间对于所有信号都是共同的并且例如对应于混响声场出现的开始。

[0045] 在一个实施例中,其中信号包括连续的采样块且在信号间具有相同的大小,可以提供至少一个给定的瞬间,以便限制频率分量的包含直到截止频率,所述给定的瞬间在时间上位于与一系列块中的第一块不同的一个块的开始。这一给定的瞬间因此发生在直接传播之后并且位于声音反射或者混响场出现的时刻。

[0046] 此实施例将在下文参考图5进行详述,在一个示例性的实施例中,还说明了实现本发明意义的方法包括由空间化模块的处理器所执行的计算机程序的可能算法。在这一方面,本发明一般还涉及一种包括当由处理器执行时用于实现上述方法的指令的计算机程序。

[0047] 本发明还关注一种声音空间化模块,包括用于将具有房间效果的至少一个传递函数应用于至少一个输入声音信号的处理器,所述应用相当于在频谱范围中,将声音信号的频谱分量与对应于所述传递函数的滤波器的频谱分量相乘,该滤波器的每一个频谱分量具有在时间频率表示中的时间演变。具体的说,此处理器被配置为对于分量的所述乘积,忽略该滤波器超出阈值频率并且在所述时间频率表示中至少一个给定瞬间之后的所述频谱分量。

[0048] 该声音空间化模块,接收多个输入信号,提供至少两个输出信号,该处理器被配置为将具有房间效果的传递函数应用于每一个输入信号,所述输出信号各自通过应用下述类型的公式给出:

$$O^k = \sum_{l=1}^L (I(l) *_{[0;\dots;f^k(l)]} A^k(l))$$

[0049]

$$+ \sum_{m=1}^M \left(z^{-iDDm} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) \right) *_{[0;\dots;f^k(m)]} B_{mean}^k(m)$$

- [0050] O^k 是输出信号,并且k是与一个输出信号相关的索引,
- [0051] $l \in [1;L]$ 是与所述输入信号中的一个输入信号相关的索引,L是输入信号的数量,并且I(l)是所述输入信号中的一个输入信号,
- [0052] $A^k(l)$ 是具有房间效果的传递函数,专用于一个输入信号,
- [0053] $B_{mean}^k(m)$ 是具有房间效果的通用的传递函数,为输入信号所共用,
- [0054] $W^k(l)$ 是所选择的加权因子,并且G(I(l))是预先确定的功率补偿增益,
- [0055] z^{-iDDm} 是延迟的应用,被认为采样的块的数量,对应于在与房间效果对应的在房间中的声音的发散与所述房间中混响场的出现开始之时的时间差,索引m对应于与这一延迟对应的持续时间中的多个采样的块,M为在时间频率表示中传递函数持续的块的总数量,
- [0056] 符号“.”表示乘积,
- [0057] 术语“ $*[0;...;f^k(l)]$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至最高频率 $f^k(l)$,它是至少一个索引为l的输入信号的函数,以及
- [0058] 术语“ $*[0;...;f^k(m)]$ ”表示在受限数量的频率上的卷积算子并且范围为最低频率至频率 $f^k(m)$,它是索引为m的采样的块的函数。
- [0059] 这一模块可集成于压缩解码器件中或者更一般的集成于渲染系统中。
- [0060] 图6表示了这一空间化模块SPAT和解码设备DECOD,在所示的示例中,它接收来自网络RES的压缩编码信号I'(l)(其中 $l=1, \dots, L$)并先于渲染解码它们,发送解码的信号I(l)(其中 $l=1, \dots, L$)至空间化模块。在所示的示例中,后一模块包括输入接口IN以接收解码的信号,以及计算装置诸如处理器PROC和与接口IN/OUT协作的工作存储器MEM以便空间化信号I(l)并通过输出接口OUT仅仅传递意欲提供至耳机CAS的各自听筒的两个信号 O^d 和 O^s 。

附图说明

- [0061] 根据下文的详细说明并参考附图,本发明的其它特征及优点将变得更为明显,附图包括:
- [0062] 图1说明了本发明方法的一般实施例;
- [0063] 图2说明了根据一个实施例的该方法的应用,其中传递函数是两个传递函数的组合形式,其中之一在延迟后应用于所要处理的信号;
- [0064] 图3示出了具有可变截止频率(或者上文所提到的“阈值频率”)的传递函数的时间频率表示的一个示例,具体作为时间函数的变化;
- [0065] 图4说明了对应于本发明背景的计算机程序的一种可能的一般算法的流程图;
- [0066] 图5示出了源于图2所示模式所获得的一种特定实施例,只不过针对多于两个连续时间块,其中传递函数 $B_{mean}^k(m)$ 表示随作为块m的函数而改变的混响场;
- [0067] 图6示出了对应于本发明背景的空间化模块的一个示例;
- [0068] 图7示意性说明了虚拟扬声器和当应用适当的传递函数时的房间效果,其中将所述传递函数的频率分量限制于适当的截止频率。
- [0069] 在描述图1和本发明的一般原理之前,我们将参考图7解释本发明的基础物理现象。

具体实施方式

[0070] 在所示出的示例中,多个虚拟扬声器围绕着收听者的头部TE。虚拟扬声器HPV各自初始被提供一个信号 $I(1)$,其中 $1 \in [1;L]$,举例来说,如上文参考图6所表示的那样,可以预先被解码。虚拟扬声器的安排可关注所要处理的信号 $I(1)$ 的一种多声道表示或者加上一种环绕立体声表示,以便在耳机CAS上一起以一种具有房间效果的空间化方式来渲染它们(图6)。为此,对于要提供的每个听筒信号 0^k 典型地为每个信号应用一个具有房间效果的传递函数,其中 $k=d$ (对于右侧的), g (对于左侧的)。于是,参考图7,对于每个虚拟扬声器HPV,我们考虑从扬声器HPV向着左耳OG的声音路径(在示出的示例中为同侧的TIL)和从扬声器HPV向着右耳OD的声音路径(在示出的示例中为对侧的TCL),还有在墙MUR上的反射(路径RIL),以及最终在多次反射之后的混响场。在每次反射时,都认为该声波的最高频率被衰减了。

[0071] 于是,参考图3,它关注适于图7所示的虚拟扬声器HPV的传递函数的时间频率表示,已经非常明显的是收听者的头自然地位于对侧路径中并且要为针对右耳OD的传递函数所考虑的最高频率低于要为针对左耳OG的传递函数所考虑的那些(它沿着一同侧路径面向虚拟扬声器HPV)。因此,考虑从0至 $N-1$ 的第一时间块,表示为 $m=0$,表示针对右耳的传递函数的滤波器的最大频率 $F_c^d(0)$ 可能低于表示针对左耳的传递函数的滤波器的最大频率 $F_c^g(0)$ 。这一滤波器的开发者因此可限制针对右耳的该滤波器的分量不超过截止频率 $F_c^d(0)$ (对应于头部筛选频率),即使所要处理的信号 $I(1)$ 可能具有不超过至少频率 $F_c^g(0)$ 的更高的频谱分量。

[0072] 然后,在反射后,该声波在高频趋于衰减,这确实发生在针对左耳以及右耳的传递函数的时间频率表示中,对于时刻 N 至 $2N-1$ 来说,对应于表示为 $m=1$ 的下一块。因此,表示这些传递函数的滤波器的开发者可限制针对右耳的滤波器的分量不超过截止频率 $F_c^d(1)$ 并且对于左耳不超过截止频率 $F_c^d(1)$ 。在参考图5具体说明的一个实施例中,我们可认为在块 $m=1$ 中,传递函数典型地特征化针对右耳和针对左耳的混响场,并因此它可(可能的并非限制性的)被创建为 $F_c^d(1) = F_c^g(1)$ 。

[0073] 然后,当存在具有声音的一般衰减(“淡出”)的混响场时,该声波趋于在高频更加衰减,这确实发生在如图3针对左耳和右耳的传递函数的时间频率表示中,对于 $2N$ 至 $3N-1$ 的瞬间,对应于表示为 $m=2$ 的块。因此,表示这些传递函数的滤波器的开发者可将针对右耳的滤波器的分量限制为不超过截止频率 $F_c^d(2)$ 并将针对左耳的限制为不超过截止频率 $F_c^g(2)$ 。

[0074] 应当注意的是,较短的块会允许要考虑的最高频率的更精确的变化,例如以便考虑第一反射RIL,对于该第一反射RIL来说在块 $m=0$ 的第一时刻中最高频率针对右耳增加(如图3所示 $F_c^d(0)$ 附近的虚线)。

[0075] 因此,我们看到可能不考虑表示传递函数的滤波器的所有频谱分量,具体是超出截止频率 F_c 的。于是,有利于在频谱范围内处理该传递函数的应用。信号 $I(1)$ 与传递函数的卷积在频谱范围内可以变为信号 $I(1)$ 的频谱分量与表示在频谱范围内的传递函数的滤波器的频谱分量的乘积,并且具体为,可执行这一相乘仅仅直到截止频率,它是诸如给定块和所要处理信号的函数。

[0076] 因此,参考图1, L 个输入信号 $I(1)$ 、 $I(2)$ 、...、 $I(L)$ 在各自的步骤TF11、TF12、...、TF1L中被变换至频域。作为选择的,这些输入信号有可能已经以频率形式(例如,在解码器

中)更为有效。

[0077] 在步骤BA11中,对应于来自声道1的信号 $I(1)$ 的时间形式的完整的空间化脉冲响应(典型的为BRIR-“双耳房间脉冲响应”)被存储在存储器中。在步骤TFA11中,将此脉冲响应转换为频率的形式,以便获得频谱范围中的相应滤波器。在一个有利的实施例中,该滤波器以它的频谱形式进行存储,以避免重复该变换计算。然后,此滤波器与来自声道1的频率形式的输入信号相乘(它等同于在时域中的卷积)。我们因此获得来自声道1的信号 $I(1)$ 的空间化信号。

[0078] 为 $L-1$ 个其它声道执行相同的操作。因此我们获得总共 L 个空间化的声道。这些声道然后被加和,以获得表示这 L 个声道的单个输出信号,并且在步骤ITF11中返回时域,以便输出提供给听筒的信号 O^k (其中 $k=d,g$)的其中一个。对于另一听筒执行相似的处理。在将参考图2和5在下文详细描述的一个实施例中,在加和之前不能单独访问这 L 个空间化的声道:通过逐渐加和每一个空间化的信道与之前的输出信号来构造该单个输出信号。

[0079] 针对所要构造的每一个输出信号 O^k 执行这些操作。在双耳的复制中,这些步骤一般可执行两次,一次针对所要提供给耳机的左侧听筒的输出信号,一次针对所要提供给该耳机的右侧听筒的输出信号。我们因此最后获得两个空间化的信号 O^d 和 O^g ,各自对应于一只耳朵。

[0080] 这 L 个输入信号可能典型地对应于意欲提供给(“虚拟”)扬声器的多声道音频内容的 L 个声道。这 L 个输入信号可能例如对应于在环绕立体声表示中的音频内容的 L 个环绕立体声信号。

[0081] 现在参考图2,它说明了本发明背景中的一个实施例,我们再次参见如图1所展示的 L 个声道的空间化的原理。然而,图2中的展示通过将 L 个输入信号组合为单条线 $I(1)$ 而被简化。因此, L 个输入信号 $I(1), I(2), \dots, I(L)$ 在步骤S21中被变换至频域。正如上文所表明的,这种输入信号可能作为选择地在频率形式中已经可用。在步骤S22中,对应于声道1的信号 $I(1)$ 的来自空间化(典型为BRIR类型)的脉冲响应 $A^k(1)$ 被变换至频谱范围,以便获得频率滤波器。这一脉冲响应 $A^k(1)$ 在图2的表示中是不完整的,因为它对应于采样 $m=0$ 的第一时间块。正如上文所表明的,此脉冲响应在频率形式中可能已经可用。此滤波器的分量然后与对应声道的频谱信号 $I(1)$ 相乘。在本发明背景中,这一相乘被配置(如下文参考图4所示的)可以忽略某些频率分量。典型的是,最高频率分量被忽略,以便减少计算的复杂度。在图2和5中,受限至截止频率的分量的相乘由符号 x 表示。

[0082] 定义截止频率 $f_{cA(I)}$,忽略超出该频率的频率分量(例如在声道的信号 $I(1)$ 中表示的最大频率或者它的采样频率的一半)。此外,此截止频率专用于每个滤波器并且专用于每个块(举例来说,它对于块 $m=1, m=2$ 来说降低)。由于这里的滤波器专用于每一个输入信号并专用于各个耳朵,截止频率专用于一个输入信号并专用于一只耳朵(并因此专用于一个输出信号)以及专用于一个时间块。

[0083] 然后,我们获得针对第一时间块的声道1的空间化信号。针对所有 L 个声道 $l=1, \dots, L$ 执行这些操作。这就提供了 L 个空间化的声道。然后,在步骤S23中,加和这些声道,以获得表示在第一时间块中的这 L 个声道的单个信号。

[0084] 实际上,该加和可以一种特定的方式来执行,以允许声道中的延迟实现特征化混响(反射和混响场),如下文所详述的。确实,在一个实施例中,在加和之前并不可独立访问

这L个空间化的声道:通过逐渐加和每一个空间化的声道和之前的输出信号来构建单个输出信号。为此,在步骤DBD中,输入信号 $I(1)$ 可延迟一延迟时间,通过 $z^{-iDD \cdot m}$ 给出,专用于每一个块 $m=1, \dots, M$ 。将注意到的是,针对第一块的延迟 m 是零。在频率表示的情况下,此延迟一般对应于针对第一块处理的信号帧的大小并且被解释为以其频率形式采用之前输入块的行为。

[0085] 在步骤S24中,来自对应于声道1的信号 $I(1)$ 空间化(典型为BRIR类型)的不完整脉冲响应 $B_m^k(1)$ 被转换至频谱范围,以便获得一个频率滤波器。这一脉冲响应 $B_m^k(1)$ 是不完整的,因为它对应于采样的第二时间块(然后对应于第三块,以此类推,对于 $m=1, \dots, M$)。如上文所表明的,作为一个变形,此脉冲响应在频率形式中可能已经可用。应用在文档FR13 57299中所描述的原理,可能的是通过假定 $B_m^k(1) = \dots = B_m^k(1) = \dots = B_m^k(L) = B_{\text{mean}}^k(m)$ 并且使得这一传递函数最终仅依赖于所关注的块 m (主要是混响场或者具有“淡出”衰减的次要混响场)以及耳朵 k 来降低处理的复杂度。相似地,该混响场并不依赖于声道并且可能的是为每个信道设置相同的截止频率 f_c (但是,它仍可从一个块下降至下一个,如之前参考图3所示的那样)。这一实施例如图5所示。

[0086] 再次参考图2,滤波器 $B_m^k(1)$ 然后与声道1的信号 $I(1)$ 相乘。该截止频率对于此第二时间块来说是不同的。如参考图3所讨论的,测量结果显示出在更远距离的时间块中(对应于混响声音和多次混响)高频更为衰减。针对这些更远距离块的截止频率因此与针对第一块的截止频率可更低。截止频率越低,操作的数量减少得越多。因此,这有利于降低计算的复杂度。

[0087] 对于这L个声道执行相同的操作,并且我们重复将滤波器与逐渐延迟的频谱信号相乘,针对每次延迟在步骤S25中加和该贡献直到我们获得表示覆盖所考虑的时间块 m 的集合 M 的L个声道的单个信号。通过逐渐加和每个空间化的声道与之前的输出信号来构造单个输出信号,正如现在将参考图4所讨论的。

[0088] 最后,在步骤S26中,我们返回时域,以便获得要提供给其中一个耳机听筒的输出信号。

[0089] 参考图4,我们现在描述针对给定时间块(例如值在时间间隔 $[0; N-1]$ 中的表示直接声场的块)以及针对对应于诸如右耳的信号的空间化方法。当然,针对对应于左耳的信号应用相同的方法。通过应用专用于各个耳朵的滤波器引入两只耳朵之间的区别。

[0090] 在步骤S40中,输出信号 S 可初始化为0。以频域方式表达这一输出信号。它具有受限的大小,长度大于截止频率 $f_c(1)$ 。举例来说,这一信号可定义为 $[0; f_s(1)/2]$, $f_s(1)$ 为此信号 $I(1)$ 的采样频率。第一计数变量 i 同样初始化为1。此第一计数变量标识为针对右耳的时间块 $[0; N-1]$ 中的声道信号 $I(1)$ 、 $I(2)$ 、 \dots 、 $I(1)$ 、 \dots 、 $I(L)$ 的其中一个。在步骤S41中,第二计数变量 j 初始化为0。此第二计数变量标识为针对右耳的时间块 $[0; N-1]$ 中的信号 $I(1)$ 的频率分量。

[0091] 在步骤S42中,系数 $c_{\text{BRIR}}(j; 1)$ 被存储在存储器中。这一系数对应于针对右耳的时间块 $[0; N-1]$ 中的滤波器BRIR(I)的频率分量 j 。相似地,系数 $c_I(j; 1)$ 被存储在存储器中。此系数对应于针对右耳的在时间块 $[0; N-1]$ 中的信号 $I(1)$ 的频率分量 j 。因此,系数 $c_{\text{BRIR}}(j; 1)$ 和 $c_{\text{BRIR}}(j; 1)$ 对应于相同的频率分量(由变量 j 标识)并因此可随后逐项地相乘(步骤S44)。

[0092] 在测试T47中,我们检查对应于变量 j 的频率是否小于(例如严格地小于)截止频率

$f_c(1)$ 。这一截止频率对应于针对右耳的时间块 $[0;N-1]$ 的信号 $I(1)$ 的截止频率。如果该频率小于截止频率 $I(1)$ ，则转到步骤S44。

[0093] 在步骤S44中，计算对应于系数 $C_{BRIR}(j;1)$ 和 $C_I(j;1)$ 的乘积的值 $MULT(j)$ 。这些系数被逐项相乘，因为它们对应于相同的频率分量 j （针对相同声道，以相同块以及针对同一耳朵）。

[0094] 在步骤S45中，此值 $MULT(j)$ 在频率 j 的位置处被递增地增加至信号 S 。

[0095] 因此，逐步地构造信号 S ，所述信号包括（在长度 $f_c(1)$ 的循环的末端处）不超过截止频率 $f_c(1)$ 的所有的频率分量（针对此信号 $I(1)$ ，在块 $[0;N-1]$ 中，并针对右耳）。因为在图4中当该循环开始时我们已经将所有分量初始化为0，在该循环的末端一个缓冲（初始为零）已经被填充直至该截止频率，连续地构造该信号 S 。系数的每次相乘 $MULT(j)$ 因此被逐步增加至正在构造的信号 S 。

[0096] 在步骤S46中，变量 j 被递增并且返回至步骤S42。如果变量 j 大于（例如，或者等于）截止频率 $f_c(1)$ ，则前进至测试T48。因此，针对间隔 $[0;f_c(1)]$ 填满了信号 S 。

[0097] 如上文所述，可针对比 $[0;f_c(1)]$ 更大的间隔来定义该信号（例如， $[0;f_s(1)/2]$ ）。此外，此信号的全部定义的间隔已经被初始化为0。因此，该间隔的未填充的剩余部分（例如 $[f_c(1);f_s(1)/2]$ ）仍为零。这就会提升复杂度，因为在信号 S 中填充的一些步骤还没有被执行，这些会减少所必需的的计算的数量。

[0098] 在测试T48中，检查对应于声道1的信号 $I(1)$ 的计数变量1是否小于（例如，严格小于）声道的数量 L 。如果变量1小于或等于 L ，变量1在步骤S49中递增并且该方法返回至步骤S41。如果变量1大于 L ，在步骤S50中对应于针对右耳的时间块 $[0;N-1]$ 的空间化信号的信号 S 是可用的。

[0099] 对应于时间块 $[0;N-1]$ 的这一信号 S 然后与针对其他时间块 $[N;2N-1]$ 、 $[2N;3N-1]$ 等的其它相似生成的信号相加（并且对它应用适当的延迟，例如依据上述参考图2所示步骤DBD）。

[0100] 典型的是，为了构造块 $[N;2N-1]$ ，可在频域应用对应于为所有输入信号 $I(1)$ 所共用的传递函数的滤波器，以便表示该混响场，在对应于下述两者的最小值的频域分量的乘积中使用截止频率 f_c ：

[0101] -如上述图3所示的混响场最大频率 F_c （混响）（举例来说，对于块 $m=1$ 在10至15kHz之间进行选择并且对于块 $m=2$ 在5至10kHz之间进行选择），以及，

[0102] -在每个输入信号中表示的最大频率 f_{max} （例如它的采样频率或者其频域分量不为零的最大频率，此值典型的是由压缩解码器给出的）。

[0103] 值得注意的是，该频率相乘，它终止于给定截止频率（它在数学上等价于超出那一点时乘以0），对于本领域技术人员来说，这是不重要的。确实，在滤波声音信号背景下，此类非常侵略性的低通滤波器一般会产生能听得见的混叠效应，这归因于从由循环卷积产生的时间混叠所造成的回音或前回音现象，一般令人满意的是避免它的出现。然而，在本发明的上下文中，低通滤波器并不被应用于声音信号而是应用于BRIR滤波器（它自身与该声音信号卷积），它已经由多次反射所组成；产生的人工效果因此在最坏的情况下被认为是原始BRIR滤波器的额外反射，并且在实践中是很难注意到的。尽管如此，但有可能通过轻微修改在截止频率之前的该滤波器的频率来减轻这些人工效果（例如，通过应用半汉宁窗口的轻

微衰减(淡出类型)。

[0104] 一般来说,参考图4,将会注意到,在同一循环实例(典型为一个时钟周期)中执行两个操作:相乘MULT(k)以及将它增加至输出信号S。这允许在具有在单个循环实例(典型为一个时钟周期)中执行多个操作的能力的处理器上执行这一方法,从而减少计算所需的时间。

[0105] 图5示出了根据上文所展示公式的该处理的一个完整的算法形式,它所获得的输出信号 O^k :

$$O^k = \sum_{l=1}^L (I(l) *_{[0;\dots;f^k(l)]} A^k(l))$$

[0106]

$$+ \sum_{m=1}^M \left(z^{-iDDm} \cdot G(I(l)) \cdot \sum_{l=1}^L \left(\frac{1}{W^k(l)} \cdot I(l) \right) \right) *_{[0;\dots;f^k(m)]} B_{mean}^k(m)$$

[0107] 正如上文所表明的,加权因子 $W^k(1)$ 和增益 $G(I(1))$ 可能固定为1。该增益 $G(I(1))$ 没有在图5中展示,因为此附图应当被读为在加权 $1/W^k(1)$ 处的增益的集成。此外,在该滤波器的设计期间,这两个参数被一劳永逸地一起确定、固定和相乘。

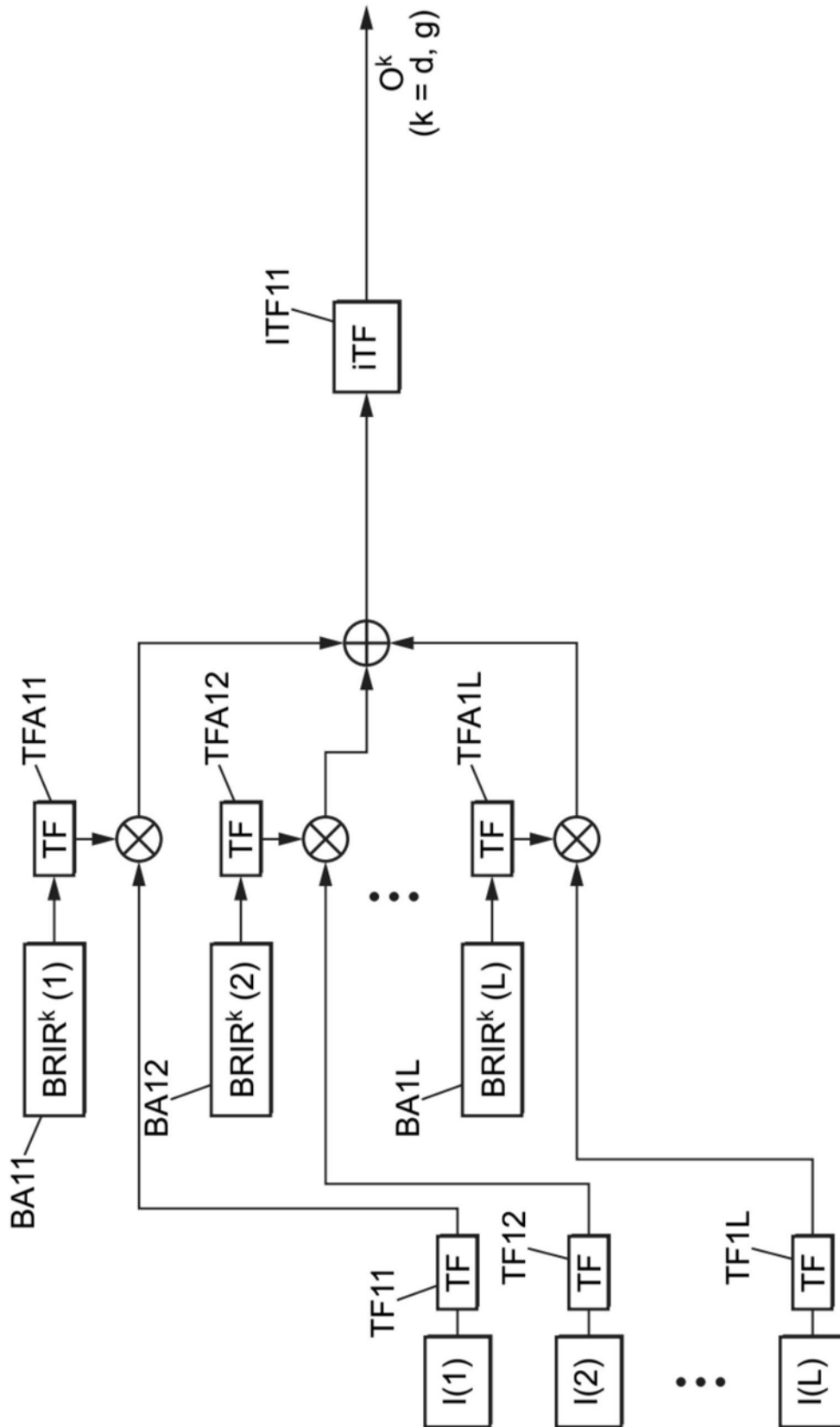


图1

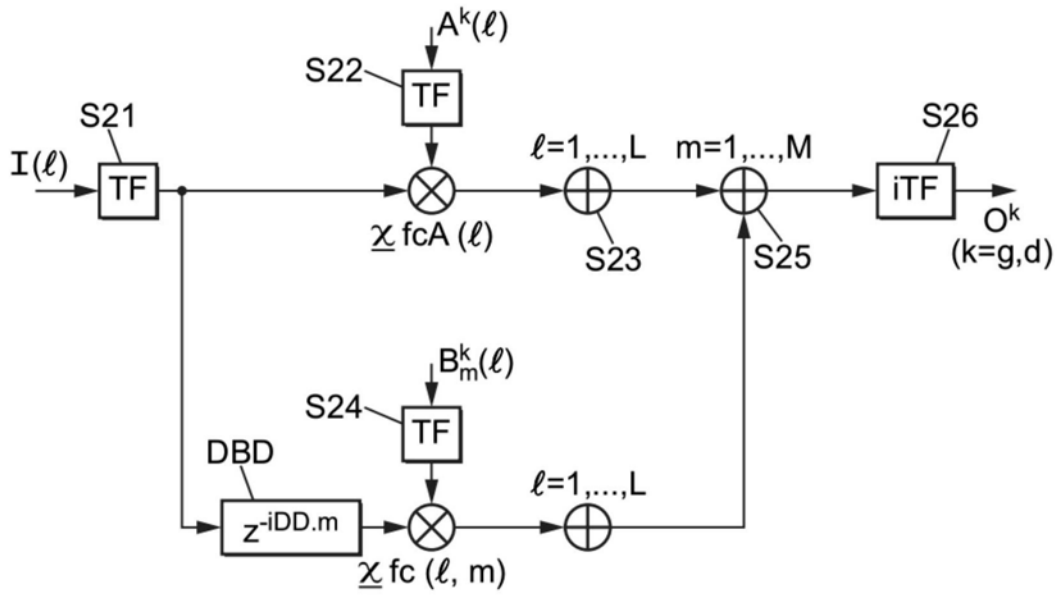


图2

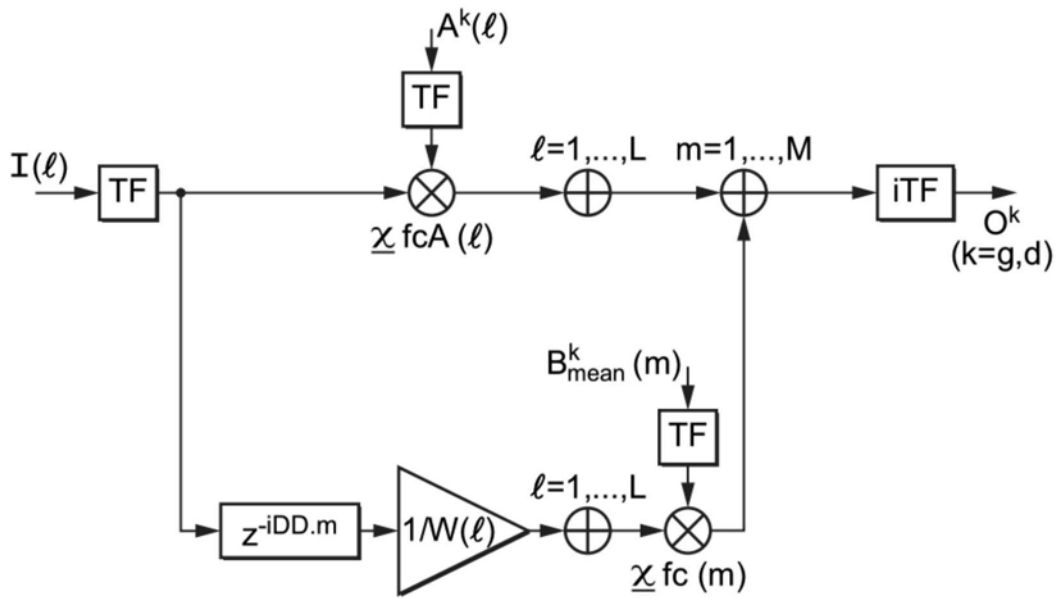


图5

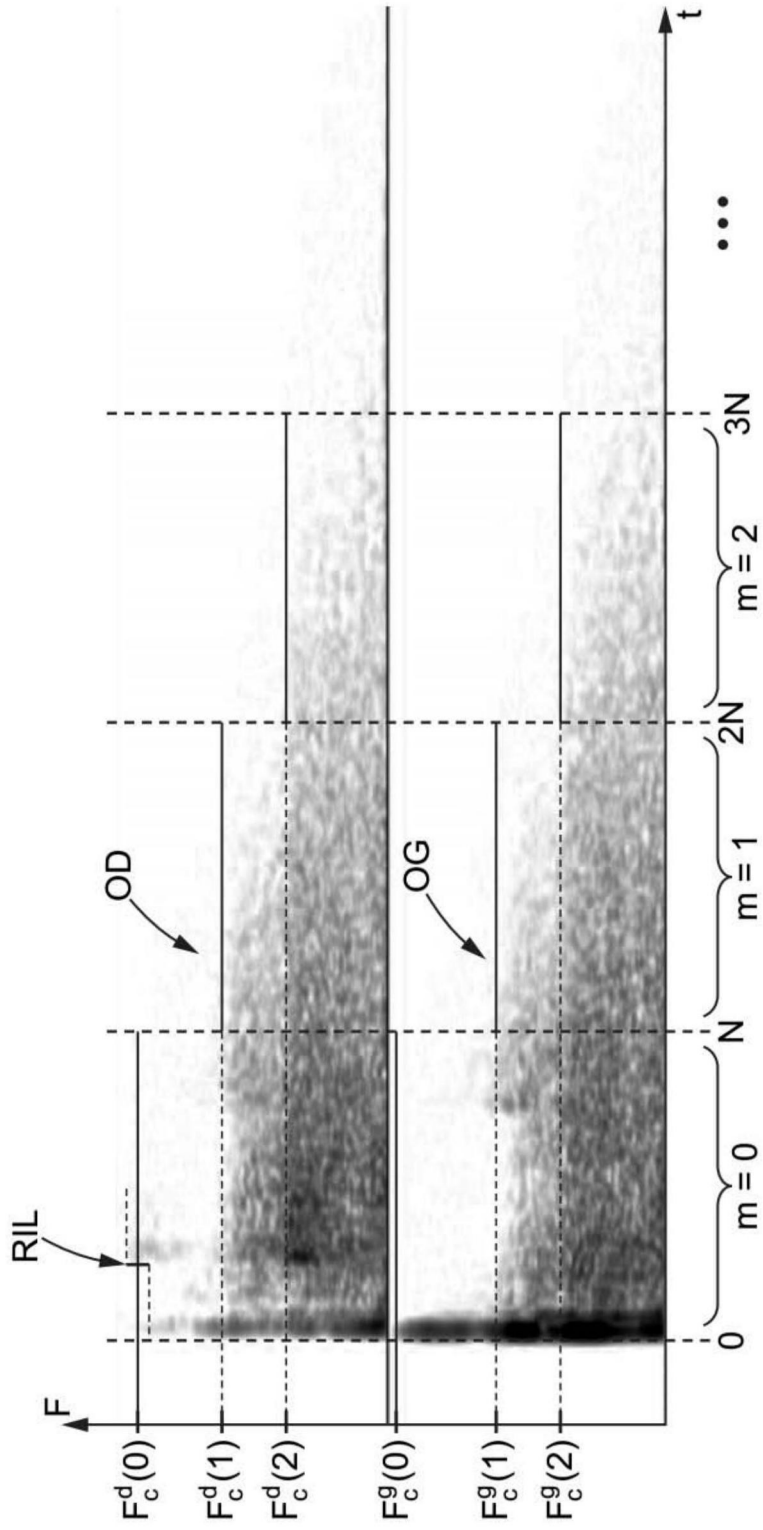


图3

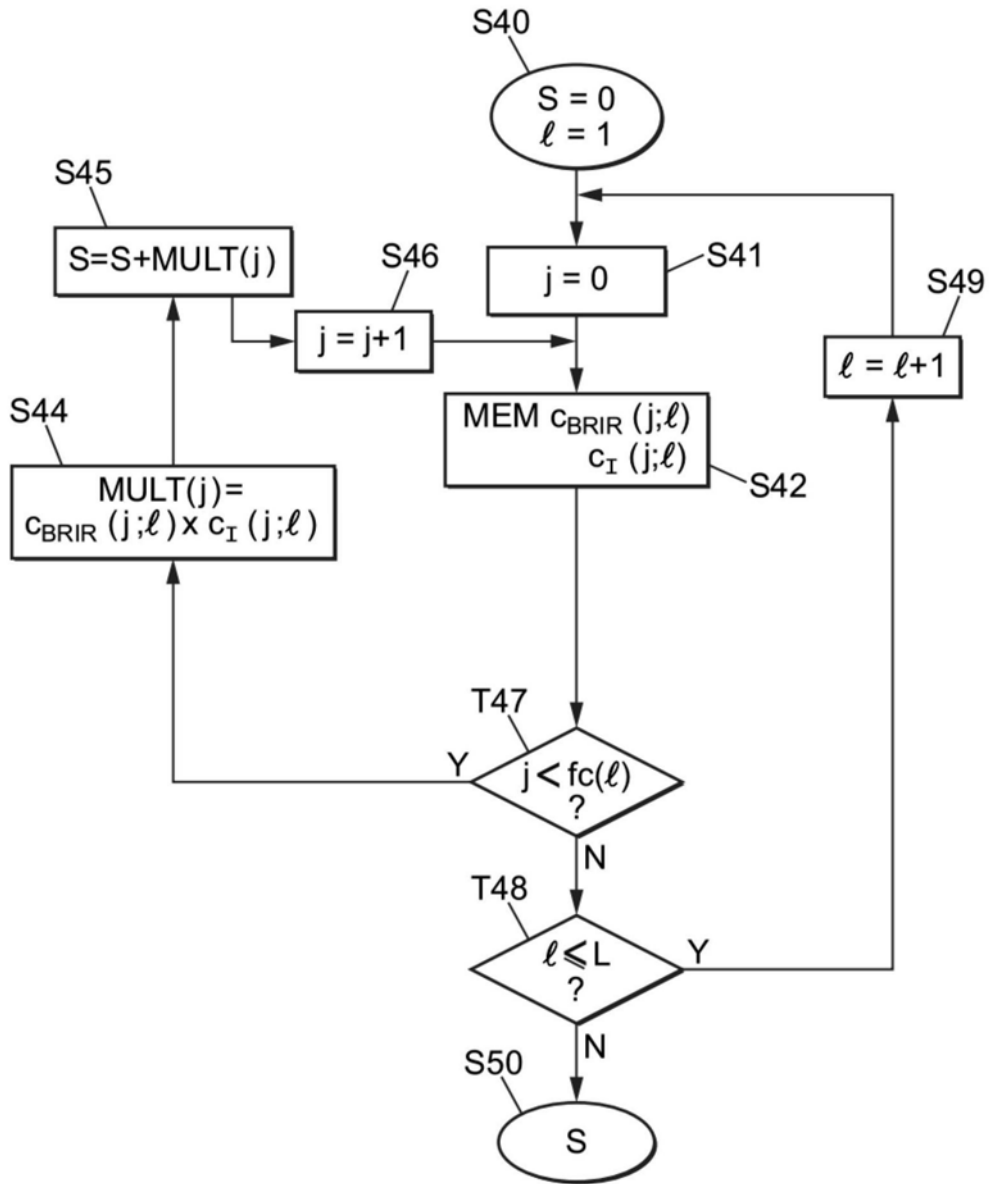


图4

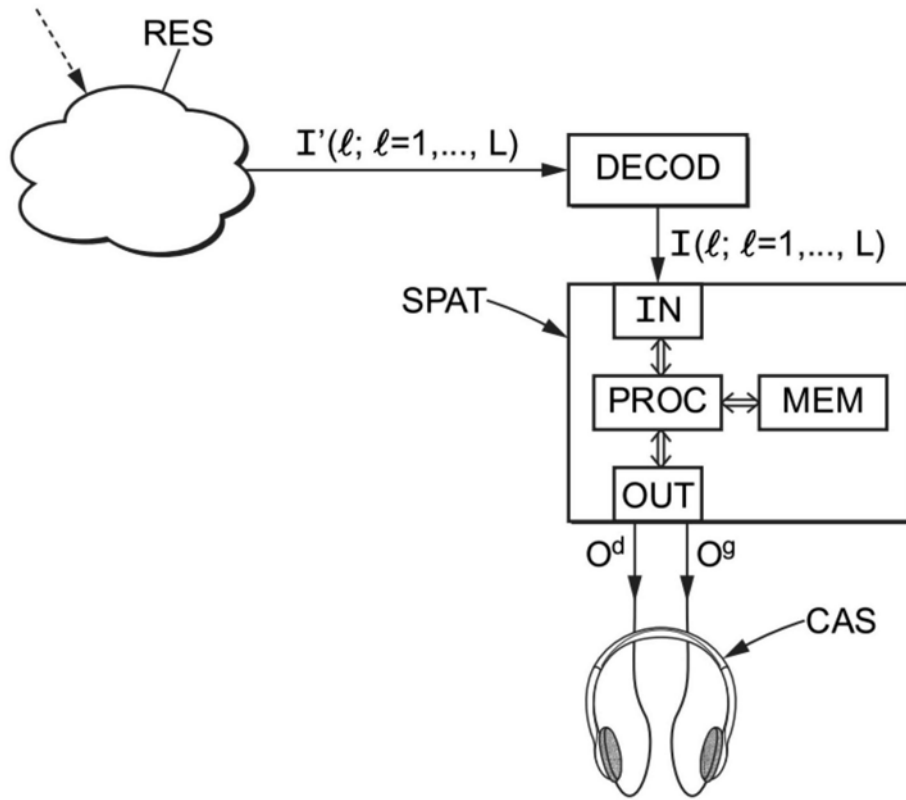


图6

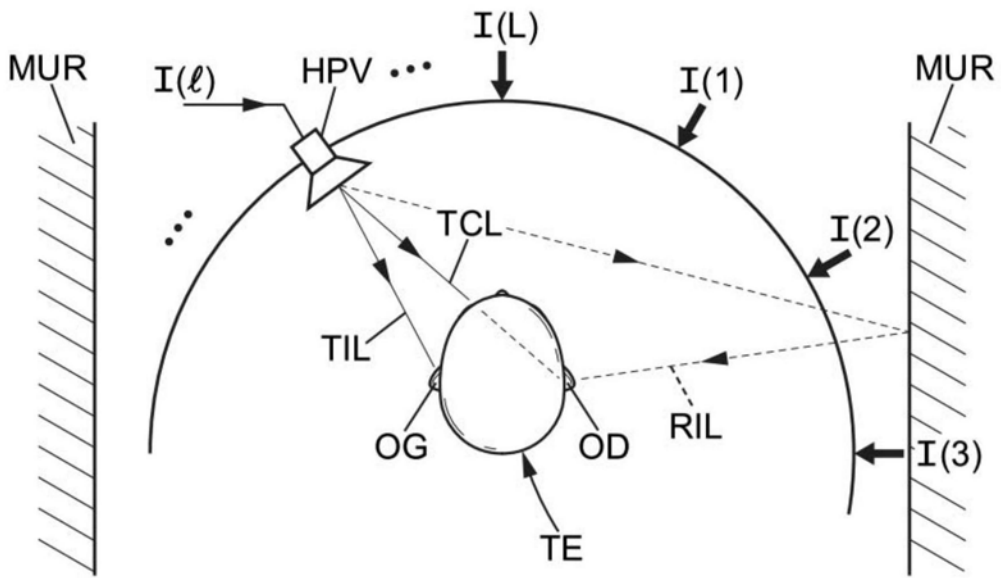


图7