



## (12)发明专利

(10)授权公告号 CN 105580391 B

(45)授权公告日 2019.04.12

(21)申请号 201480051924.2

(22)申请日 2014.07.14

(65)同一申请的已公布的文献号  
申请公布号 CN 105580391 A

(43)申请公布日 2016.05.11

(30)优先权数据  
13177368.1 2013.07.22 EP  
13189285.3 2013.10.18 EP

(85)PCT国际申请进入国家阶段日  
2016.03.21

(86)PCT国际申请的申请数据  
PCT/EP2014/065037 2014.07.14

(87)PCT国际申请的公布数据  
W02015/010937 EN 2015.01.29

(73)专利权人 弗朗霍夫应用科学研究促进协会  
地址 德国慕尼黑

(72)发明人 克里斯汀·卡特尔  
约翰内斯·希勒佩特  
安德烈·赫尔策 阿西姆·孔茨  
简·普洛格施蒂斯  
迈克尔·卡拉舒曼

(74)专利代理机构 北京清亦华知识产权代理事务  
所(普通合伙) 11201  
代理人 宋融冰

(51)Int.Cl.  
H04S 7/00(2006.01)  
H04S 5/00(2006.01)

(56)对比文件  
CN 101809654 A,2010.08.18,  
CN 102165797 A,2011.08.24,  
审查员 王亚辉

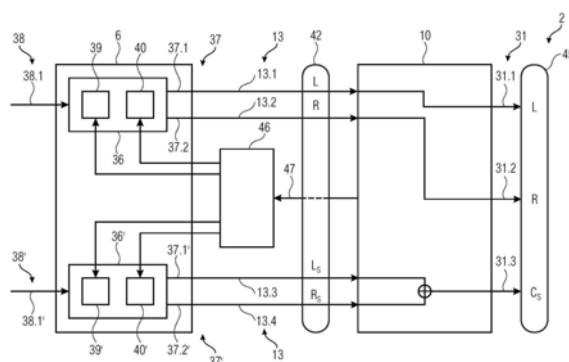
权利要求书2页 说明书16页 附图12页

(54)发明名称  
渲染器控制的空间升混

### (57)摘要

一种用于解码压缩输入音频信号的音频解  
码器装置,包括:至少一个核心解码器(6,24),具  
有用于根据处理器输入信号(38,38')产生处理  
器输出信号(37)的一个或多个处理器(36,36'),  
其中处理器输出信号(37,37')的输出声道  
(37.1,37.2,37.1',37.2')的数量高于处理器输  
入信号(38,38')的输入声道(38.1,38.1')的数  
量,其中一个或多个处理器(36,36')中的每一个  
包括解相关器(39,39')以及混合器(40,40'),其  
中具有多个声道(13.1,13.2,13.3,13.4)的核心  
解码器输出信号(13)包括处理器输出信号(37,  
37'),并且其中核心解码器输出信号(13)适用于  
参考扬声器设置(42);至少一个格式转换器装置  
(9,10),用于将核心解码器输出信号(13)转换成  
适用于目标扬声器设置(45)的输出音频信号  
(31);以及控制装置(46),用于以处理器(36,  
36')的解相关器(39,39')可以与处理器(36,

36')的混合器(40,40')独立地被控制的方式控  
制一个或多个处理器(36,36'),其中控制装置  
(46)用于根据目标扬声器设置(45)控制一个或  
多个处理器(36,36')的解相关器(39,39')中的  
至少一个。



1. 一种用于解码压缩输入音频信号的音频解码器装置,包括:

至少一个核心解码器(6,24),具有用于基于处理器输入信号(38,38')产生处理器输出信号(37)的一个或多个处理器(36,36'),其中所述处理器输出信号(37,37')的输出声道(37.1,37.2,37.1',37.2')的数量高于所述处理器输入信号(38,38')的输入声道(38.1,38.1')的数量,其中所述一个或多个处理器(36,36')中的每一个包括解相关器(39,39')以及混合器(40,40'),其中具有多个声道(13.1,13.2,13.3,13.4)的核心解码器输出信号(13)包括所述处理器输出信号(37,37'),并且其中所述核心解码器输出信号(13)适用于参考扬声器设置(42);

至少一个格式转换器装置(9,10),用于将所述核心解码器输出信号(13)转换成适用于目标扬声器设置(45)的输出音频信号(31);以及

控制装置(46),用于控制所述一个或多个处理器(36,36'),以便所述处理器(36,36')的所述解相关器(39,39')能够与所述处理器(36,36')的所述混合器(40,40')独立地被控制,其中所述控制装置(46)用于控制所述一个或多个处理器(36,36')的所述解相关器(39,39')中的至少一个,以便根据所述目标扬声器设置(45),当所述处理器(36,36')的所述解相关器(39,39')被关闭时,所述处理器(36,36')的所述混合器(40,40')是可操作的。

2. 如权利要求1所述的解码器装置,其中所述控制装置(46)用于去激活至少一个或多个处理器(36,36'),以便所述处理器输入信号(38,38')的输入声道(38.1,38.1')以未处理的形式提供至所述处理器输出信号(37,37')的输出声道(37.1,37.2,37.1',37.2')。

3. 如权利要求1或2所述的解码器装置,其中所述处理器(36,36')为一输入二输出的解码工具,其中所述解相关器(39,39')用于通过对所述处理器输入信号(38,38')的所述声道(38.1,38.1')中的至少一个进行解相关以生成解相关信号(48),其中所述混合器(40,40')基于声道电平差信号(49)和/或声道间相干性信号(50)混合所述处理器输入信号(38)以及所述解相关信号(46),以便所述处理器输出信号(37,37')由两个非相干输出声道(37.1,37.2,37.1',37.2')组成。

4. 如权利要求3所述的解码器装置,其中所述控制装置用于通过将所述解相关信号(48)设定为零或者通过阻止所述混合器(40,40')将所述解相关信号(46)混合至各个处理器(36,36')的所述处理器输出信号(37),关闭所述处理器(36,36')中的一个的所述解相关器(36,36')。

5. 如权利要求1或2所述的解码器装置,其中所述核心解码器(6)为用于音乐和语音的解码器,其中所述处理器(36,36')中的至少一个的所述处理器输入信号(38)包括声道对元素。

6. 如权利要求1或2所述的解码器装置,其中所述核心解码器(24)为参数化对象编码器。

7. 如权利要求1或2所述的解码器装置,其中所述参考扬声器设置(42)的扬声器数量高于所述目标扬声器设置(45)的扬声器数量。

8. 如权利要求1所述的解码器装置,其中所述控制装置(46)用于:关闭用于所述处理器输出信号(37')的所述输出声道的至少一个第一个(37.1')以及所述处理器输出信号(37')的所述输出声道的一个第二个(37.2')的所述解相关器(36'),如果根据所述目标扬声器设置将所述输出声道的所述第一个(37.1')和所述输出声道的所述第二个(37.2')混合至所

述输出音频信号(31)的共用声道(31.2),假设用于将所述输出声道的所述第一个(37.1')混合至所述共用声道(31.2)的第一比例因子超过第一阈值和/或将所述输出声道的所述第二个(37.2')混合至所述共用声道(31.2)的第二比例因子超过第二阈值。

9.如权利要求8所述的解码器装置,其中所述控制装置(46)用于从所述格式转换器装置(9,10)接收规则组(47),所述格式转换器装置(9,10)根据所述规则组(47)将所述核心解码器输出信号(13)的所述声道(13.1,13.2,13.3,13.4)按照所述目标扬声器设置(45)混合至所述输出音频信号(31)的所述声道(31.1,31.2,31.3),其中所述控制装置(46)用于根据接收到的所述规则组(47)控制所述处理器(36,36')中的至少一个。

10.如权利要求8或9所述的解码器装置,其中所述控制装置(46)用于控制所述处理器(36,36')的所述解相关器(39,39'),以便所述核心解码器输出信号(13)的非相干声道的数量等于所述输出音频信号(31)的所述声道(31.1,31.2,31.3)的数量。

11.如权利要求1或2所述的解码器装置,其中所述格式转换器装置(9,10)包括用于降混所述核心解码器输出信号(13)的降混器(10)。

12.如权利要求11所述的解码器装置,其中所述格式转换器装置(9,10)包括双耳渲染器(9)。

13.如权利要求12所述的解码器装置,其中所述核心解码器输出信号(13)作为双耳渲染器输入信号被提供至所述双耳渲染器(9)。

14.如权利要求12所述的解码器装置,其中所述降混器(10)的降混器输出信号作为双耳渲染器输入信号被提供至所述双耳渲染器(9)。

15.一种用于解码压缩输入音频信号的方法,所述方法包括以下步骤:

提供至少一个核心解码器(6,24),所述至少一个核心解码器(6,24)具有用于基于处理器输入信号(38,38')产生处理器输出信号(37)的一个或多个处理器(36,36'),其中所述处理器输出信号(37,37')的输出声道(37.1,37.2,37.1',37.2')的数量高于所述处理器输入信号(38,38')的输入声道(38.1,38.1')的数量,其中所述一个或多个处理器(36,36')中的每一个包括解相关器(39,39')以及混合器(40,40'),其中具有多个声道(13.1,13.2,13.3,13.4)的核心解码器输出信号(13)包括所述处理器输出信号(37,37'),并且其中所述核心解码器输出信号(13)适用于参考扬声器设置(42);

提供至少一个格式转换器装置(9,10),所述至少一个格式转换器装置(9,10)用于将所述核心解码器输出信号(13)转换成适用于目标扬声器设置(45)的输出音频信号(31);以及

提供控制装置(46),所述控制装置(46)用于控制所述一个或多个处理器(36,36')以便所述处理器(36,36')的所述解相关器(39,39')能够与所述处理器(36,36')的所述混合器(40,40')独立地被控制,其中所述控制装置(46)用于控制所述一个或多个处理器(36,36')的所述解相关器(39,39')中的至少一个,以便根据所述目标扬声器设置(45),当所述处理器(36,36')的所述解相关器(39,39')被关闭时,所述处理器(36,36')的所述混合器(40,40')是可操作的。

16.一种计算机可读存储介质,包括计算机程序,当所述计算机程序在计算机或信号处理器上运行时用于执行权利要求15所述的方法。

## 渲染器控制的空间升混

### 技术领域

[0001] 本发明涉及音频信号处理,特别地,本发明涉及多声道音频信号的格式转换。

### 背景技术

[0002] 格式转换描述将特定数量的音频声道映射至适用于通过不同数量的音频声道回放的另一呈现的过程。

[0003] 常见的格式转换的使用是将音频声道进行降混。在参考文献[1]中,给出示例,其中即使当无法获得完整的“家庭影院”5.1监控系统时,降混允许终端用户重放5.1源素材的版本。被设计为接受杜比数码素材,但只能提供单声道或立体声输出的设备(如便携式DVD播放器,机顶盒等),包含设施以降混原始5.1声道至标准的一个或两个输出声道。

[0004] 另一方面,格式转换也能够描述升混处理,例如升混立体声素材以形成5.1兼容的版本。再者,双耳渲染可视为格式转换。

[0005] 在下文中,将对压缩音频信号的解码处理的格式转换的影响进行讨论。在此,音频信号的压缩呈现(mp4文件)表示为通过固定扬声器设置进行回放准备的固定数量的音频声道。

[0006] 音频解码器以及随后的格式转换成所希望的播放格式之间的相互作用可分为三类:

[0007] 1. 此解码处理无关于最后的回放场景。因此,完整的音频呈现被重新取得并且转换处理随后被应用。

[0008] 2. 音频解码处理受限于其能力并且仅将输出固定格式。示例为接收立体声FM节目的单声道收音机,或者接收HE-AAC v2比特流的单声道HE-AAC解码器。

[0009] 3. 音频解码处理知道其最终回放设置并相应地调整其处理。示例如参考文献[2]中对MPEG环绕定义的“Scalable Channel Decoding for Reduced Speaker Configurations”。在此,解码器减少输出声道的数量。

[0010] 这些方法的缺点在于非必要的高复杂度和因解码素材的后续处理(用于降混的梳状滤波,用于升混的解除遮蔽)(1.)以及关于最终输出格式的受限灵活性(2.和3.)而导致的潜在伪象。

### 发明内容

[0011] 本发明的目的是提供一种改进的音频信号处理的概念。本发明的目的由权利要求1的解码器、权利要求14的方法以及权利要求15的计算机程序来实现。

[0012] 提供一种用于解码压缩输入音频信号的音频解码器装置,包括:具有用于根据处理器输入信号产生处理器输出信号的一个或多个处理器的至少一个核心解码器,其中处理器输出信号的输出声道的数量高于处理器输入信号的输入声道的数量,其中一个或多个处理器中的每一个包括解相关器以及混合器,其中具有多个声道的核心解码器输出信号包括处理器输出信号,并且其中核心解码器输出信号适用于参考扬声器设置;

[0013] 用于将核心解码器输出信号转换成适用于目标扬声器设置的输出音频信号的至少一个格式转换器装置;以及

[0014] 用于以处理器的解相关器可以与处理器的混合器独立地被控制的方式控制至少一个或多个处理器的控制装置,其中控制装置用于根据目标扬声器设置控制一个或多个处理器的解相关器中的至少一个。

[0015] 处理器的目的是建立具有多个非相干/非相关声道的处理器输出信号,且其声道的数量比处理器输入信号的输入声道的数量高。特别地,每个处理器生成具有多个非相干/非相关输出声道,例如具有两个输出声道的处理器输出信号,其中正确的空间线索来自具有较少数量的输入声道的处理器输入信号,例如来自单声道输入信号。

[0016] 这种处理器包括解相关器和混合器。解相关器用于从处理器输入信号的声道产生解相关器信号。典型的解相关器(解相关滤波器)由频率相依预延迟和其后的全通(IIR)部分所组成。

[0017] 解相关器信号和处理器输入信号的各个声道随后被送入混合器。混合器用于通过混合解相关器信号及处理器输入信号的各个声道以建立处理器输出信号,其中,使用边信息,以便合成正确的相干性/相关性以及处理器输出信号的输出声道的正确强度比。

[0018] 如果处理器输出信号的输出声道被送到不同位置处的不同扬声器,则处理器输出信号的输出声道为不相干/不相关的,以便处理器的输出声道被感知为独立声源。

[0019] 格式转换器可以转换核心解码器输出信号以适合在与参考扬声器设置不同的扬声器设置中播放。此设置被称为目标扬声器设置。

[0020] 在对于特定目标扬声器设置,随后的格式转换器不需要一个处理器的非相干/非相关形式的输出声道,正确相关的合成变得感知上不相关。因此,对于这些处理器,解相关器可以被省略。然而,当解相关器被关闭时,混合器通常仍然保持完全可操作。结果是,即使解相关器被关闭,处理器输出信号的输出声道仍被产生。

[0021] 必须注意的是,在这种情况下,处理器输出信号的声道是相干/相关但不相同。这意味着,在处理器的下游,可进一步彼此独立地对处理器输出信号的声道进行处理,其中,例如强度比和/或其他空间信息可用于格式转换器,以设置输出音频信号的声道的电平。

[0022] 由于解相关滤波需要大量的计算复杂度,整体解码的工作量可以由所提出的解码器装置大幅降低。

[0023] 虽然解相关器,尤其是它们的全通滤波器,被设计成对主观音质的影响降到最低,但它不能总是避免引入可听的伪象,例如由于相位失真或某些频率分量的“振铃(ringing)”导致的瞬变的模糊。因此,当解相关过程的副作用被避免时,可实现音频音质的改进。

[0024] 值得注意的是,此处理应仅被应用于其中应用解相关的频带。其中使用残余编码的频带不被影响。

[0025] 在优选实施例中,控制装置用于去激活至少一个或多个处理器,使得处理器输入信号的输入声道以未处理的形式被提供至处理器输出信号的输出声道。由此,不相同的声道的数量可以被减少。这可能是有益的,如果目标扬声器设置包括的扬声器的数量与参考扬声器设置中扬声器的数量相比非常小。

[0026] 在优选实施例中,处理器可为一输入二输出的解码工具(OTT),其中解相关器用于

通过对处理器输入信号的至少一个声道进行解相关,产生解相关信号,其中混合器基于声道电平差 (CLD) 信号和/或声道间相干 (ICC) 信号混合处理器输入音频信号以及解相关信号,使得处理器输出信号包括两个不相干输出声道。这种一输入二输出解码工具允许以简单的方式产生具有声道对的处理器输出信号,声道对具有关于彼此的正确振幅和相干性。

[0027] 在一些实施例中,控制装置用于通过将解相关信号设定为零或者通过阻止混合器将解相关信号混合至各个处理器的处理器输出信号,关闭一个处理器的解相关器。两种方式均允许以简单的方式关闭解相关器。

[0028] 在优选实施例中,核心解码器为用于音乐和语音的解码器,例如USAC解码器,其中处理器中的至少一个的处理器输入信号包括声道对元素,例如USAC声道对元素。在这种情况下,可以省略声道对元素的解码,如果其对于当前的目标扬声器设置不是必须的。这样,计算复杂度和由解相关处理以及降混合处理产生的伪象可以显著减小。

[0029] 在一些实施例中,核心解码器为参数化对象编码器,例如SAOC解码器。这样,计算复杂度和由解相关处理以及降混处理产生的伪象可以进一步减小。

[0030] 在一些实施例中,参考扬声器设置的扬声器数量高于目标扬声器设置的扬声器数量。这样,格式转换器可以降混核心解码器输出信号到音频的输出音频信号,其中输出声道的数量低于核心解码器输出信号的输出声道的数量。

[0031] 这里,降混描述了当参考扬声器设置中存在的扬声器数量高于目标扬声器设置中使用的扬声器数量的情况。在此情况下,一个或多个处理器的输出声道通常并不需要非相干信号的形式。如果处理器的解相关器被关闭,计算复杂度和由解相关处理以及降混处理所产生的伪象可以显著降低。

[0032] 在一些实施例中,控制装置用于关闭用于处理器输出信号的输出声道的至少一个第一个以及处理器输出信号的输出声道的一个第二个的解相关器,如果根据目标扬声器设置将输出声道的第一个和输出声道的第二个混合至输出音频信号的共用声道,假设用于将处理器输出信号的输出声道的第一个混合至共用声道的第一比例因子超过第一阈值和/或将处理器输出信号的输出声道的第二个混合至共用声道的第二比例因子超过第二阈值。

[0033] 在将输出声道的第一个和输出声道的第二个混合至输出音频信号的共用声道,的情况下,对于第一输出声道及第二输出声道,可以省略核心解码器处的解相关。这样,计算复杂度和由解相关处理以及降混处理所产生的伪象可以显著降低。以这种方式,可避免不必要的解相关。

[0034] 在更进一步的实施例中,可预测用于混合处理器输出信号的输出声道的第一个的第一比例因子。可以以同样的方式使用用于混合处理器输出信号的输出频道的第二个的第二比例因子。这里,比例因子是数值,其通常介于0和1之间,此比例因子描述原始声道的信号强度(处理器输出信号的输出声道)和混合声道(输出音频信号的共用声道)中的结果信号的信号强度间的比率。此比例因子可被包含在降混矩阵中。通过使用对于第一比例因子的第一阈值和/或通过使用对于第二比例因子的第二阈值,可以确保仅在第一输出声道的至少确定部分和/或第二输出声道的至少确定部分被混合至共用声道时,第一输出声道和第二输出声道的解相关才被关闭。举例来说,阈值可以被设定为0。

[0035] 在优选实施例中,控制装置用于从格式转换器接收规则组,格式转换器根据规则组将处理器输出信号的声道按照目标扬声器设置混合至输出音频信号的声道,其中控制装

置用于根据接收到的规则组控制处理器。这里,处理器的控制可包括解相关器和/或混合器的控制。由此,可以确保控制装置以精确的方式控制处理器。

[0036] 通过规则组,处理器的输出声道是否被随后的格式转换步骤所结合的信息可以被提供给控制装置。控制装置接收的规则通常为降混矩阵的形式,降混矩阵定义格式转换器使用的每个解码器输出声道至每个音频输出声道的比例因子。下一步中,通过控制装置可以从降混规则计算用于控制解相关器的控制规则。控制规则可以被包括在所谓的混合矩阵中,可以通过控制装置根据目标扬声器设置产生混合矩阵。然后,控制规则可以被用于控制解相关器和/或混合器。因此,控制装置可被适用于不同的目标扬声器设置且无需人工介入。

[0037] 在优选实施例中,控制装置用于以核心解码器输出信号的不相干声道的数量等于目标扬声器设置中扬声器的数量的方式控制核心解码器的解相关器。在这种情况下,计算复杂度和由解相关处理以及降混处理所产生的伪象可以显著降低。

[0038] 在一些实施例中,格式转换器包括用于降混核心解码器输出信号的降混器。降混器直接产生输出音频信号。然而,在一些实施例中,降混器可以被连接至格式转换器的另一元件,其然后产生输出音频信号。

[0039] 在一些实施例中,格式转换器包括双耳渲染器。双耳渲染器一般被用于将多声道信号转换至适用于立体声耳机的立体声信号。双耳渲染器产生被提供至双耳渲染器的信号的双耳降混,使得这个信号的每个声道由虚拟声源所表示。处理可以在正交镜像滤波器(QMF)域内被逐帧执行。双耳化是基于测量的双耳室脉冲响应,并且造成极高的计算复杂度,计算复杂度与被提供至双耳渲染器的信号的非相干/非相关声道的数量有关。

[0040] 在优选实施例中,核心解码器输出信号作为双耳渲染器输入信号被提供至双耳渲染器。在此情况下,控制装置通常用于控制核心解码器的处理器,以便核心解码器输出信号的声道数量比耳机的扬声器数量多。这可能是被需要的,例如为了产生三维音频效果,双耳渲染器可以使用包含在声道中的空间声音信息调整被提供至耳机的立体声信号的频率特性。

[0041] 在一些实施例中,降混器的降混器输出信号作为双耳渲染器输入信号被提供至双耳渲染器。在降混器的输出音频信号被提供至双耳渲染器的情况下,其输入信号的声道数量明显小于当核心解码器输出信号被提供至双耳渲染器的情况,由此降低计算复杂度。

[0042] 此外,提供一种用于解码压缩输入音频信号的方法,方法包括以下步骤:提供具有用于根据处理器输入信号产生处理器输出信号的一个或多个处理器的至少一个核心解码器,其中处理器输出信号的输出声道的数量高于处理器输入信号的输入声道的数量,其中一个或多个处理器中的每一个包括解相关器以及混合器,其中具有多个声道的核心解码器输出信号包括处理器输出信号,并且其中核心解码器输出信号适用于参考扬声器设置;提供至少一个格式转换器,至少一个格式转换器用于将核心解码器输出信号转换成适用于目标扬声器设置的输出音频信号;以及提供控制装置,控制装置用于以处理器的解相关器可以与处理器的混合器独立地被控制的方式控制一个或多个处理器,其中控装置用于根据目标扬声器设置控制一个或多个处理器的解相关器中的至少一个。

[0043] 此外,提供一种计算机程序,当计算机程序在计算机或信号处理器上运行时用于执行上述方法。

## 附图说明

- [0044] 下面,结合附图更详细地描述本发明的实施例,其中:
- [0045] 图1显示根据本发明的解码器的优选实施例的框图,
- [0046] 图2显示根据本发明的解码器的第二实施例的框图,
- [0047] 图3显示概念上的处理器的模型,其中解相关器被开启,
- [0048] 图4显示概念上的处理器的模型,其中解相关器被关闭,
- [0049] 图5示出格式转换和解码之间的交互作用,
- [0050] 图6显示根据本发明的解码器的实施例的细节的框图,其中产生5.1声道信号,
- [0051] 图7显示根据本发明的解码器的图6实施例的细节的框图,其中5.1声道被降混为2.0声道信号,
- [0052] 图8显示根据本发明的解码器的图6实施例的细节的框图,其中5.1声道被降混为4.0声道信号,
- [0053] 图9显示根据本发明的解码器的实施例的细节的框图,其中产生9.1声道信号,
- [0054] 图10显示根据本发明的解码器的图9实施例的细节的框图,其中9.1声道信号被降混为4.0声道信号,
- [0055] 图11示出3D音频编码器的概念描述的示意图,
- [0056] 图12示出3D音频解码器的概念描述的示意图,以及
- [0057] 图13示出格式转换器的概念描述的示意图。

## 具体实施方式

- [0058] 在描述本发明的实施例之前,提供更多本领域编码器解码器系统的背景知识。
- [0059] 图11示出3D音频编码器1的概念描述的示意图,图12示出3D音频解码器2的概念描述的示意图。
- [0060] 3D音频编解码器系统1,2可以基于用于声道信号4和对象信号5的编码的MPEG-D统一语音及音频编码(USAC)编码器3,以及基于用于编码器3的输出音频信号7的解码的MPEG-D统一语音及音频编码(USAC)解码器6。为了增加大量的对象5的编码效率,采用空间音频对象编码(SAOC)技术。三种渲染器8、9和10执行将对象11和12渲染至声道13以及将声道13渲染至耳机或将声道渲染至不同的扬声器设置的任务。
- [0061] 当对象信号被使用SAOC参数化编码或显式地传输时,相对应的对象元数据(OAM)14信息被压缩且被多路复用为3D音频比特流7。
- [0062] 在编码之前,可以选择性地使用预渲染器/混合器15以将声道和对象输入场景4,5转换成声道场景4,16。功能上,其与下面所描述的对象渲染器/混合器15相同。
- [0063] 对象5的预渲染确保编码器3的输入处的确定性信号熵基本与同时激活的对象信号5的数量无关。对于对象5的预渲染,不需要对象元数据14传输。
- [0064] 离散对象信号5被渲染至编码器3被配置使用的声道布局。对于每个声道16的对象5的权重从相关联的对象元数据14获得。
- [0065] 用于扬声器声道信号4、离散对象信号5、对象降混信号14和预渲染信号16的核心编解码器可以基于MPEG-D USAC技术。此MPEG-D USAC技术通过基于输入的声道和对象分配的几何及语义信息生成声道及对象映射信息,来处理大量的信号4、5和14的编码。这个映射

信息描述输入声道4和对象5如何被映射至USAC声道元素(即声道对元素(CPE)、单声道元素(SCE)、低频增强(LFE)),相应的信息被传输至解码器6。

[0066] 所有额外的有效负载例如SAOC数据17或对象元数据14可以作为扩展元素被传送,并且可以在编码器3的速率控制中被考虑。

[0067] 对象5的编码可以使用不同的方法,取决于用于渲染器的速率/失真需求和交互需求。下列对象编码变型是可能的:

[0068] -预渲染的对象16:在编码之前,对象信号5被预渲染和混合至声道信号4,例如22.2声道信号4。后续编码链看见22.2声道信号4。

[0069] -离散对象波形:对象5作为单声道波形被提供至编码器3。除了声道信号4以外,编码器3使用单声道元素(SCE)以传输对象5。解码对象18在接收器端被渲染并混合。压缩的对象元数据信息19,20被一起传输至接收器/渲染器21。

[0070] -参数化对象波形17:使用SAOC参数22及23的方式描述对象属性及其彼此之间的关系。对象信号17的降混使用USAC来编码。参数信息22被一起传输。降混声道17的数量根据对象5的数量及整体数据速率来选择。压缩的对象元数据信息23被传输至SAOC渲染器24。

[0071] 用于对象信号5的SAOC编码器25以及解码器24基于MPEG SAOC技术。系统可以基于较少量的传输声道7和额外的参数数据22及23,例如对象级差(OLD)、对象间相关性(IOC)及降混增益值(DMG),重新创建、修正及渲染多个音频对象5。额外的参数数据22及23显示的数据速率明显低于单独传输所有对象5所需要的数据速率,这使得编码效率非常高。

[0072] SAOC编码器25将单声道波形的对象/声道信号5作为输入,并且输出参数信息22(打包至3D音频比特流7)和SAOC传输声道17(使用单声道元素编码并且传输)。SAOC解码器24从解码的SAOC传输声道26和参数信息23重建对象/声道信号5,并且基于重现布局,解压缩的对象元数据信息20,以及可选地基于用户交互信息产生输出音频场景27。

[0073] 对于每个对象5,使用对象元数据编码器28通过对象属性在空间和时间上的量化,有效地编码指定对象在三维空间中的几何位置以及体积的相关联的对象元数据14。压缩对象元数据(cOAM)19作为边信息20被传输至接收器,边信息可以使用OAM解码器29来解码。

[0074] 对象渲染器21根据给定的重现格式,利用压缩的对象元数据20来产生对象波形12。每个对象5根据其对象元数据19及20被渲染至特定的输出声道12。块21的输出由部分结果的总和所产生。如果基于声道的两个内容11和30,以及离散/参数对象12和27被解码,那么基于声道的波形11,30和被渲染的对象波形12,27在输出结果波形13之前(或在将他们提供至后处理器模块9,10如双耳渲染器9或扬声器渲染器模块10之前)被混合器8混合。

[0075] 双耳渲染器模块9产生多声道音频材料13的双耳降混,使得每个输入声道13由虚拟声源所表示。处理在正交镜像滤波器(QMF)域内被逐帧执行。双耳化是基于测量的双耳室脉冲响应。

[0076] 图13中详细描述扬声器渲染器10在传输的声道配置13和期望的重现格式31之间转换。在下文中因此称之为“格式转换器”10。格式转换器10执行转换至较少数量的输出声道31,即通过降混器32产生降混。DMX配置器33自动地为给定的输入格式13和输出格式31的组合产生最佳降混矩阵,并且在降混处理32中应用这些矩阵,其中混合器输出布局34及重现布局35被使用。格式转换器10允许应用于标准扬声器配置以及非标准扬声器位置的随机配置。

[0077] 图1显示根据本发明的解码器2的优选实施例的框图。

[0078] 用于解码压缩的输入音频信号38,38' 的音频解码器装置2包括具有用于根据处理器输入信号38,38' 产生处理器输出信号37,37' 的一个或多个处理器36,36' 的至少一个核心解码器6,其中处理器输出信号37,37' 的输出声道37.1,37.2,37.1' 和37.2' 的数量高于处理器输入信号38,38' 的输入声道38.1,38.1' 的数量,其中一个或多个处理器36,36' 中的每一个包括解相关器39,39' 以及混合器40,40', 其中具有多个声道13.1,13.2,13.3和13.4的核心解码器输出信号13包括处理器输出信号37,37', 并且其中核心解码器输出信号13适用于参考扬声器设置42。

[0079] 进一步,音频解码器装置2包括至少一个格式转换器装置9,10,至少一个格式转换器装置9,10用于将核心解码器输出信号13转换成适用于目标扬声器设置45的输出音频信号31。

[0080] 此外,音频解码器装置2包括控制装置46,控制装置46用于以处理器36,36' 的解相关器39,39' 可以与处理器36,36' 的混合器40,40' 独立地被控制的方式控制一个或多个处理器36,36', 其中控制装置46用于根据目标扬声器设置45控制一个或多个处理器36,36' 的解相关器39,39' 的至少一个。

[0081] 处理器36,36' 的目的是生成具有多个非相干/非相关声道37.1,37.2,37.1' 和37.2' 且其数量比处理器输入信号38的输入声道38.1,38.1' 高的处理器输出信号37,37'。特别地,每个处理器36,36' 生成的处理器输出信号37可以具有多个非相干/非相关输出声道37.1,37.2,37.1' 和37.2', 具有来自具有较少数量的输入声道38.1,38.1' 的处理器输入信号38,38' 的正确空间线索。

[0082] 在图1所示的实施例中,第一处理器36具有从单声道输入信号38产生的两个输出声道37.1及37.2,第二处理器36' 具有从单声道输入信号38' 产生的两个输出声道37.1' 及37.2'。

[0083] 格式转换器装置9,10可以将核心解码器输出信号13转换成适用于在与参考扬声器设置42不同的扬声器设置45上播放。这个设置被称为目标扬声器设置45。

[0084] 在图1所示的实施例中,参考扬声器设置42包括左前方扬声器(L)、右前方扬声器(R)、左环绕扬声器(LS)及右环绕扬声器(RS)。进一步,目标扬声器设置45包括左前方扬声器(L)、右前方扬声器(R)及中心环绕扬声器(CS)。

[0085] 如果对于特定目标扬声器设置45,随后的格式转换器装置9,10不需要一个处理器36,36' 的非相干/非相关形式的输出声道37.1,37.2,37.1' 和37.2', 正确相关的合成将变得感知上无关。因此,对于这些处理器36,36', 解相关器39,39' 可以被省略。然而,当解相关器被关闭时,通常这些混合器40,40' 仍然保持完全可操作。由此,即使解相关器被关闭,仍产生处理器输出信号的输出声道37.1、37.2、37.1' 和37.2'。

[0086] 必须指出的地方在于,在这种情况下,处理器输出信号37,37' 的声道37.1、37.2、37.1' 和37.2' 是相干/相关但不相同的。这意味着,在处理器36,36' 的下游,可进一步彼此独立地处理处理器输出信号37,37' 的声道37.1、37.2、37.1' 和37.2', 其中例如,强度比和/或其它的空间信息可以被用于格式转换器装置9及10,以设置输出音频信号31的声道37.1、37.2、37.1' 和37.2' 的位准。

[0087] 由于解相关滤波需要大量的计算复杂度,整体解码的工作量可以被本发明所提出

的解码器装置2大幅降低。

[0088] 虽然解相关器39及39', 尤其是他们的全通滤波器, 被设计成对主观音质的影响最小, 但它无法总是避免引入可听的伪象, 例如由于相位失真导致或某些频率分量的“振铃”导致的瞬态的模糊。因此, 可实现音频音质的改进, 因为解相关过程的副作用被省略。

[0089] 值得注意的是, 此处理应仅被应用于其中应用解相关的频带。而使用残余编码的频带将不被影响。

[0090] 在优选实施例中, 控制装置46用于去激活至少一个或多个处理器36, 36', 使得处理器输入信号38的输入声道38.1, 38.1' 以未处理的形式被提供至处理器输出信号37, 37' 的输出声道37.1、37.2、37.1' 和37.2'。由此, 不相同的声道的数量可以被减少。这可能是有益的, 如果目标扬声器设置45包括的扬声器的数量与参考扬声器设置42中扬声器的数量相比非常小。

[0091] 在优选实施例中, 核心解码器6为用于音乐和语音的解码器6, 例如USAC解码器6, 其中处理器中的至少一个的处理器输入信号38, 38' 包括声道对元素, 例如USAC声道对元素。在这种情况下, 如果对于当前的目标扬声器设置45不是必须的, 则可以省略声道对元素的解码。这样, 计算复杂度和由解相关处理以及降混处理产生的伪象可以显著减小。

[0092] 在一些实施例中, 核心解码器为参数化对象编码器24, 例如SAOC解码器24。这样, 计算复杂度和由解相关处理以及降混处理产生的伪象可以进一步减小。

[0093] 在一些实施例中, 参考扬声器设置42的扬声器数量高于目标扬声器设置45的扬声器数量。这样, 格式转换器装置9, 10可以降低核心解码器输出信号13到音频的输出音频信号31, 其中输出声道31.1、31.2及31.3的数量低于核心解码器输出信号13的输出声道13.1、13.2、13.3及13.4的数量。

[0094] 这里, 降混描述当参考扬声器设置42中存在扬声器数量高于目标扬声器设置45中使用的扬声器数量的情况。在此情况下, 一个或多个处理器36及36' 的输出声道37.1、37.2、37.1' 及37.2' 通常并不需要非相干信号的形式。在图1中, 存在核心解码器输出信号13的四个解码器输出声道13.1、13.2、13.3及13.4, 但是音频输出信号31只存在三个输出声道31.1、31.2及31.3。如果处理器36及36' 的解相关器39及39' 被关闭, 计算复杂度和由解相关处理以及降混处理所产生的伪象可以显著降低。

[0095] 其原因解释如下, 在图1中, 解码器输出声道13.3及13.4不需要处于非相干信号的形式。因此, 解相关器39' 被控制装置46关闭, 而解相关器39和混合器40及40' 被开启。

[0096] 在一些实施例中, 控制装置46用于关闭用于处理器输出信号37及37' 的输出声道的至少一个第一个37.1' 以及处理器输出信号37及37' 的输出声道的一个第二个37.2及37.2' 的解相关器39', 如果根据目标扬声器设置(45)将输出声道的第一个37.1' 和输出声道的第二个37.2' 混合至输出音频信号31的共用声道31.3, 假设用于将处理器输出信号37' 的输出声道的第一个37.1' 混合至共用声道31.3的第一比例因子超过第一阈值和/或将处理器输出信号37' 的输出声道的第二个37.2' 混合至共用声道31.3的第二比例因子超过第二阈值。

[0097] 在图1中, 解码器输出声道13.3及13.4被混合至输出音频信号31的共用声道31.3。第一比例因子和第二比例因子可以是0.7071。当本实施例的第一阈值和第二阈值被设定为0, 其解相关器39' 被关闭。

[0098] 如果将输出声道的第一个37.1'与输出声道的第二个37.2'混合至输出音频信号31的共用声道31.3,核心解码器6处的针对第一和第二输出声道37.1'及37.2'的解相关可以省略。这样,计算复杂度和由解相关处理以及降混处理所产生的伪象可以显著降低。此方式可避免不必要的解相关。

[0099] 在更进一步的实施例中,可预测用于混合处理器输出信号37'的输出声道的第一个37.1'的第一比例因子。可以以同样的方式使用用于混合处理器输出信号37'的输出声道的第二个37.2'的第二比例因子。这里,比例因子是一个数值,其通常介于0和1之间,描述原始声道(处理器输出信号37'的输出声道37.1'及37.2')的信号强度和混合声道(输出音频信号31的共用声道31.1)中的结果信号的信号强度之间的比率。此比例因子可被包含在降混矩阵中。通过使用用于第一比例因子的第一阈值和/或通过使用用于第二比例因子的第二阈值,可以确保仅在第一输出声道37.1'的至少确定部分和/或第二输出声道37.2'的至少确定部分被混合至共用声道31.3时才关闭用于第一输出声道37.1'和第二输出声道37.2'的解相关。举例来说,阈值可以被设定为0。

[0100] 在图1的实施例中,解码器输出声道13.3及13.4被混合至输出音频信号31的共用声道31.3。第一比例因子和第二比例因子可以是0.7071。当本实施例的第一阈值和第二阈值被设定为0,其解相关器39'被关闭。

[0101] 在优选实施例中,控制装置46用于从格式转换器装置9及10接收规则组47,根据规则组47,格式转换器装置9及10将处理器输出信号37及37'的声道37.1、37.2、37.1'及37.2'按照目标扬声器设置45混合至输出音频信号31的声道31.1、31.2及31.3,其中控制装置46用于根据接收到的规则组47控制处理器36及36'。这里,处理器36,36'的控制可包括解相关器39,39'和/或混合器40,40'的控制。由此,可以确保控制装置46以精确的方式控制处理器36,36'。

[0102] 通过规则组47,处理器36,36'的输出声道是否被随后的格式转换步骤所结合的信息可以被提供给控制装置9,10。控制装置46接收的规则通常为降混矩阵的形式,降混矩阵表示由格式转换器装置9,10采用的每个核心解码器输出声道13.1、13.2、13.3及13.4至每个音频输出声道31.1、31.2及31.3的比例因子。下一步,可以由控制装置从降混规则计算控制解相关的控制规则。这个控制规则可以被包括在所谓的混合矩阵中,可以通过控制装置46根据目标扬声器设置45产生混合矩阵。然后,控制规则可以被用于控制解相关器39,39'和/或混合器40,40'。因此,控制装置46可被适用于不同的目标扬声器设置45且无需人工介入。

[0103] 在图1中,规则组47可以包含解码器输出声道13.3及13.4被混合至输出音频信号31的共用声道31.3的信息。这可以在图1的实施例中被执行,参考扬声器设置42的左环绕扬声器和右环绕扬声器被目标扬声器设置45中的中心环绕扬声器替换。

[0104] 在优选实施例中,控制装置46用于以核心解码器输出信号13的不相干声道的数量等于目标扬声器设置45中扬声器的数量的方式控制核心解码器6的解相关器39,39'。在这种情况下,计算复杂度和由解相关处理以及降混处理所产生的伪象可以显著降低。

[0105] 例如,在图1中存在三个不相干声道,第一个为解码器输出声道13.1、第二个为解码器输出声道13.2以及第三个为解码器输出声道13.3及13.4的每个,由于省略解相关器39',因此解码器输出声道13.3及13.4为相干的。

[0106] 在实施例中,例如在图1所示的实施例中,格式转换器装置9,10包括用于降混核心解码器输出信号13的降混器10。降混器10可直接产生输出音频信号31,如图1所示。然而,在一些实施例中,降混器10可以被连接至格式转换器10的另一元件,例如双耳渲染器9,其随后产生输出音频信号31。

[0107] 图2显示根据本发明的解码器的第二实施例的框图。在下文中仅讨论与第一实施例的差异。在图2中,格式转换器9,10包括双耳渲染器9。双耳渲染器9一般用于将多声道信号转换成适用于立体声耳机使用的立体声信号。双耳渲染器9产生被提供至双耳渲染器9的多声道信号的双耳降混LB和RB,使得信号的每个声道由虚拟声源所表示。多声道信号可以有高达32个声道或更多。然而,在图2中示出四声道信号以简化。处理可以在正交镜像滤波器(QMF)域内被逐帧执行。双耳化基于测量的双耳室脉冲响应,并且造成极高的计算复杂度,计算复杂度与被提供至双耳渲染器9的信号的非相干/非相关声道的数量相关。为了降低计算复杂度,可以关闭解相关器39,39'中的至少一个。

[0108] 在图2所示的实施例中,核心解码器输出信号13作为双耳渲染器输入信号13被提供至双耳渲染器9。在此情况下,控制装置46通常用于控制核心解码器6的处理器,以便核心解码器输出信号13的声道13.1、13.2、13.3及13.4的数量比耳机的扬声器数量多。这可能是被需要的,例如为了产生三维音频效果,双耳渲染器9可以使用包含在声道中的空间声音信息以调整被提供至耳机的立体声信号的频率特性。

[0109] 在未示出的实施例中,降混器10的降混器输出信号作为双耳渲染器输入信号被提供至双耳渲染器9。如果降混器10的输出音频信号被提供至双耳渲染器9,其输入信号的声道数量明显小于当核心解码器输出信号13被提供至双耳渲染器9的情况,由此降低计算复杂度。

[0110] 在有利的实施例中,处理器36是一输入二输出的解码工具(OTT)36,如图3及图4所示,。

[0111] 如图3所示,解相关器39用于通过对处理器输入信号38的至少一个声道38.1进行解相关,产生解相关信号48,其中混合器40基于声道电平差(CLD)信号49和/或声道间相干性(ICC)信号50混合处理器输入信号48以及解相关信号48,使得处理器输出信号37组成两个不相干输出声道37.1及37.2,以便处理器输出信号37包括两个不相干输出声道37.1及37.2。

[0112] 这种一输入二输出解码工具36允许以简单的方式生成具有声道对37.1及37.2的处理器输出信号37,声道对37.1及37.2具有关于彼此的正确幅度和相干性。典型的解相关器(解相关滤波器)由频率相依预延迟和其后的全通(IIR)部分所组成。

[0113] 在一些实施例中,控制装置用于通过将解相关信号48设定为零或者通过阻止混合器将解相关信号48混合至各个处理器36的处理器输出信号37,关闭一个处理器36的解相关器39。两种方式均可简单地关闭解相关器39。

[0114] 一些实施例可被定义用于基于“ISO/IEC IS 23003-3统一语音及音频编码”的多声道解码器2。

[0115] 对于多声道编码,USAC由不同声道元素所组成。下面给出了5.1音频声道的一个示例。

[0116] 简单比特流负荷的示例

[0117]

|            | 元素数量 | 元素索引 | usac 元素类别[元素索引] |
|------------|------|------|-----------------|
| 5.1 声道输出信号 | 4    | 1    | ID_USAC_SCE     |
|            |      | 2    | ID_USAC_CPE     |
|            |      | 3    | ID_USAC_CPE     |
|            |      | 4    | ID_USAC_LFE     |

[0118] 对于通过OTT 36进行的从单声道至立体声的升混,每个立体声元素ID\_USAC\_CPE可以使用MPEG环绕。如下所述,通过混合单声道输入信号和提供有这个单声道输入信号的解相关器39的输出,每个元件产生具有正确空间线索的两个输出声道37.1,37.2[2][3]。

[0119] 一个重要的建构块是解相关器39,其用于合成输出声道37.1及37.2的正确的相干性相关性。典型地,解相关滤波器由频率相依预延迟及其后的全通(IIR)部分所组成。

[0120] 如果一个OTT解码块36的输出声道37.1及37.2被随后的格式转换步骤降混,正确的相关性的合成将变得感知上无关。因此,对于这些升混块,解相关器39可以被省略。这可以如下被实现。

[0121] 如图5所示,格式转换9及10和解码之间的交互作用可被建立。可以产生是否通过随后的格式转换步骤降混OTT解码块36的输出声道37.1及37.2的信息。这个信息被包括在所谓的混合矩阵中,混合矩阵由矩阵计算器46所产生并且传送至USAC解码器6。矩阵计算器所处理的信息典型地为由格式转换模块9,10提供的降混矩阵。

[0122] 格式转换处理块9,10将音频数据转换成适合在与参考扬声器设置42不同的扬声器设置45上播放。这个设置被称为目标扬声器设置45。

[0123] 降混描述使用在目标扬声器设置45中使用的扬声器的数量小于参考扬声器设置42中存在的扬声器的数量的情况。

[0124] 图6中示出了核心解码器6,核心解码器6提供的核心解码器输出信号包括适用于5.1参考扬声器设置42的输出声道13.1至13.6,输出声道13.1至13.6包括左前方扬声器声道L、右前方扬声器声道R、左环绕扬声器声道LS、右环绕扬声器声道RS、中心前方扬声器声道C和低频增强扬声器声道LFE。当处理器36的解相关器39被开启时,处理器36基于提供至处理器的声道对元素(ID\_USAC\_CPE)产生输出声道13.1及13.2,作为解相关声道13.1及13.2。

[0125] 左前方扬声器声道L、右前方扬声器声道R、左环绕扬声器声道LS、右环绕扬声器声道RS及中心前方扬声器声道C为主要声道,而低频增强扬声器声道LFE为可选的。

[0126] 以同样的方式,当处理器36'的解相关器39'被开启时,通过处理器36'基于提供至处理器36'的声道对元素(ID\_USAC\_CPE)产生输出声道13.3及13.4,作为解相关声道13.3及13.4。

[0127] 输出声道13.5基于单声道元素(ID\_USAC\_SCE),而输出声道13.6基于低频增强元素ID\_USAC\_LFE。

[0128] 如果六个适合的扬声器是可用的,核心解码器输出信号13可以被用于不需要任何降混的播放。然而,如果只有立体声扬声器设置是可用的,核心解码器输出信号13可以被降混。

[0129] 典型地,降混过程可以通过定义每个源声道至每个目标声道的比例因子的降混矩阵而描述。

[0130] 例如,ITU BS775定义用于降混5.1主要声道至立体声的下述降混矩阵,其映射声道L、R、C、LS及RS至立体声声道L'及R'。

$$[0131] \quad M_{DMX} = \begin{pmatrix} 1,0 & 0,0 & 0,7071 & 0,701 & 0,0 \\ 0,0 & 1,0 & 0,7071 & 0,0 & 0,7071 \end{pmatrix}$$

[0132] 降混矩阵具有维度 $m \times n$ ,其中 $n$ 为源声道的数量且 $m$ 为目的声道的数量。

[0133] 在矩阵计算器处理块中从降混矩阵 $M_{DMX}$ 推导出所谓的混合矩阵 $M_{Mix}$ ,其描述源声道的哪一部分被组合,它具有维度 $n \times n$ 。

$$[0134] \quad M_{Mix}(i, j) = \begin{cases} 1, & \text{如果通过降混组合声道和声道} \\ 0, & \text{否则} \end{cases}$$

[0135] 请注意 $M_{Mix}$ 是对称矩阵。

[0136] 对于上述降混5声道至立体声的示例,混合矩阵 $M_{Mix}$ 如下:

$$[0137] \quad M_{Mix} = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

[0138] 下列伪码提供了一种用于获取混合矩阵的方法:

```

MMix = zero n x n Matrix
for i = 1 to m
    for j = 1 to n
        set_j = 0
        if MDmx(i, j) > thr
            set_j = 1
        end
        for k = 1 to n
            set_k = 0
[0139]    if MDmx(i, k) > thr
                set_k = 1
            end
            if set_j == 1 and set_k == 1
                MMix(j, k) = 1
            end
        end
    end
end
end
end

```

[0140] 举例而言, 阈值thr可以被设定成零。

[0141] 每个OTT解码块产生相对应于声道号码i及j的两个输出声道。如果混合矩阵 $M_{Mix}(i, j)$ 等于1, 对于这个解码块的解相关被关闭。

[0142] 对于省略解相关器39, 元素 $q^{l,m}$ 被设定为零。可选地, 解相关路径可以被省略, 如下所述。

[0143] 这导致升混矩阵 $H22_{OTT}^{l,m}$ 的元素 $H12_{OTT}^{l,m}$ 和 $H22_{OTT}^{l,m}$ 分别被设定为零或被省略。(详细内容参见参考文献[2]的“6.5.3.2 Derivation of arbitrary matrix element”)

[0144] 在另一个优选实施例中, 升混矩阵 $R_2^{l,m}$ 的元素 $H11_{OTT}^{l,m}$ 及 $H21_{OTT}^{l,m}$ 应当通过设定 $ICC^{l,m}=1$ 来计算。

[0145] 图7示出主要声道L、R、LS、LR及C至立体声声道L'及R'的降混。由于处理器36产生的声道L及R不被混合至输出音频信号31的共用声道, 处理器36的解相关器39保持开启。同样地, 当处理器36'产生的声道LS及RS不被混合至输出音频信号31的共用声道时, 处理器36'的解相关器39'保持开启。低频增强扬声器声道LFE可以被选择性地使用。

[0146] 图8示出图6所示的5.1参考扬声器设置42至4.0目标扬声器设置45的降混。由于处理器36产生的声道L及R不被混合至输出音频信号31的共用声道, 处理器36的解相关器39保持开启。然而, 处理器36'产生的声道13.3(图6中的LS)及13.4(图6中的RS)被混合至输出音

频信号31的共用声道31.3,以形成中心环绕扬声器声道CS。因此,处理器36'的解相关器39'被关闭,使得声道13.3是中心环绕扬声器声道CS',并且声道13.4是中心环绕扬声器声道CS"。通过这样做,产生修改的参考扬声器设置42'。值得注意的是,声道CS'及CS"为相关但不相同的。

[0147] 为了完整性,应当增加的是声道13.5(C)及13.6(LFE)被混合至输出音频信号31的共用声道31.4,以便形成中心前方扬声器声道C。

[0148] 图9示出核心解码器6,其提供包括适用于9.1参考扬声器设置42的输出声道13.1至13.10的核心解码器输出信号13,输出声道13.1至13.10包括左前方扬声器声道L、左前方中心扬声器声道LC、左环绕扬声器声道LS、左环绕后方垂直高度LVR、右前方扬声器声道R、右环绕扬声器声道RS、右前方中心扬声器声道RC、右环绕扬声器声道RS、右环绕后方垂直高度RVR、中心前方扬声器声道C及低频增强扬声器声道LFE。

[0149] 当处理器36的解相关器39被开启时,处理器36基于提供至处理器36的声道对元素(ID\_USAC\_CPE)产生输出声道13.1及13.2,作为解相关声道13.1及13.2。

[0150] 相似地,当处理器36'的解相关器39'被开启时,处理器36'基于提供至处理器36'的声道对元素(ID\_USAC\_CPE)产生输出声道13.3及13.4,作为解相关声道13.3及13.4。

[0151] 更进一步,当处理器36"的解相关器39"被开启时,处理器36"基于提供至处理器36"的声道对元素(ID\_USAC\_CPE)产生输出声道13.5及13.6,作为解相关声道13.5及13.6。

[0152] 此外,当处理器36"'的解相关器39"'被开启时,处理器36"'基于提供至处理器36"'的声道对元素(ID\_USAC\_CPE)产生输出声道13.7及13.8,作为解相关声道13.7及13.8。

[0153] 输出声道13.9基于单声道元素(ID\_USAC\_SCE),而输出声道13.10基于低频增强元素ID\_USAC\_LFE。

[0154] 图10示出图9中所示的9.1参考扬声器设置42至5.1目标扬声器设置45的降混。由于处理器36产生的声道13.1及13.2被混合至输出音频信号31的共用声道31.1以形成左前方扬声器声道L',处理器36的解相关器39被关闭,使得声道13.1为左前方扬声器声道L',并且声道13.2是左前方扬声器声道L"。

[0155] 更进一步,处理器36'产生的声道13.3及13.4被混合至输出音频信号31的共用声道31.2以形成左环绕扬声器声道LS。因此,处理器36'的解相关器39'被关闭,使得声道13.3为左环绕扬声器声道LS',并且声道13.4为左环绕扬声器声道LS"。

[0156] 处理器36"产生的声道13.5及13.6被混合至输出音频信号31的共用声道31.3以形成右前方扬声器声道R,处理器36"的解相关器39"被关闭,使得声道13.5是右前方扬声器声道R',并且声道13.6是右前方扬声器声道R"。

[0157] 此外,处理器36"'产生的声道13.7及13.8被混合至输出音频信号31的共用声道31.4以形成右环绕扬声器声道RS。因此,处理器36"'的解相关器39"'被关闭,使得声道13.7是右环绕扬声器声道RS',并且声道13.8是右环绕扬声器声道RS"。

[0158] 通过这样做,可以产生修改的参考扬声器设置42',其中核心解码器输出信号13的非相干声道的数量等于目标设置45的扬声器声道的数量。

[0159] 应当注意的是,此处理应仅被应用于其中应用解相关的频带。其中使用残差编码的频带不被影响。

[0160] 如之前所提到的,本发明适用于双耳渲染。双耳播放典型地出现在耳机和/或移动

设备上。因此,可能存在约束,其限制解码器和渲染复杂性。

[0161] 解相关器处理的减少/省略可以被执行。如果音频信号最终被处理用于双耳播放,则建议在所有或一些OTT解码块中省略或减少解相关。

[0162] 这将避免来自解码器中被解相关的音频信号的降混的伪象。

[0163] 用于双耳渲染的解码输出声道的数量可以被减少。除了省略解相关,可能需要解码成数量较少的非相干输出声道,然后使得用于双耳渲染的非相干输入声道的数量较少。例如,如果解码发生在移动设备上,原始的22.2声道材料,解码至5.1以及只有5个而不是22个声道的双耳渲染。

[0164] 为了降低解码器的整体复杂度,建议采用下列的处理:

[0165] A) 定义具有比原始声道配置的声道数量少的目标扬声器设置。目标声道的数量取决于质量及复杂度约束。

[0166] 为了达到目标扬声器设置,存在两种可能B1和B2,这两者也可以结合:

[0167] B1) 解码至数量较少的声道,即通过跳过解码器中的完整的OTT处理块。这需从双耳渲染器至(USAC)核心解码器的信息路径,以控制解码器处理。

[0168] B2) 应用从原始扬声器声道配置或中间声道配置应用至目标扬声器设置的格式转换(即,降混)步骤。这可以在(USAC)核心解码器之后的后处理步骤中完成,并且不需要改变的解码处理。

[0169] 最后执行步骤C):

[0170] C) 执行较少数量的声道的双耳渲染。

[0171] SAOC解码的应用

[0172] 上面描述的方法也可以被应用于参数化对象编码(SAOC)处理。

[0173] 具有解相关器处理的减少/省略的格式转换可以被执行。如果格式转换被应用在SAOC解码之后,则信息从格式转换器被传递至SAOC解码器。通过此种信息,控制SAOC解码器内部的相关性,以减少有伪象的解相关信号的数量。此信息可以为整个降混矩阵或导出的信息。

[0174] 更进一步,具有解相关处理器的减少/省略的双耳渲染可以被执行。在参数化对象编码(SAOC)的情况下,解相关被应用于解码处理。如果随后进行双耳渲染,SAOC解码器内部的解相关处理应当被省略或减少。

[0175] 此外,具有减少的声道数量的双耳渲染可能被执行。如果在SAOC解码之后应用双耳播放,SAOC解码器可以用于使用根据来自格式转换器的信息建构的降混矩阵,渲染至较少数量的声道。

[0176] 由于解相关滤波需要大量的计算复杂度,整体解码的工作量可以通过所提出的方法大幅降低。

[0177] 虽然全通滤波器被设计成对主观音质的影响最低,但它总无法避免引入可听的伪象,例如由于相位失真或某些频率分量的“振铃”导致的瞬态的模糊。因此,因为解相关过程的副作用被避免,可实现音频音质的改进。除此之外,通过随后的降混、升混或双耳处理的解相关器伪象的任何解除遮蔽(unmasking)都被避免。

[0178] 除此之外,双耳渲染与(USAC)核心解码器或SAOC解码器相结合的情况下的复杂度降低的方法也已经被讨论。

[0179] 关于解码器和编码器以及实施例的方法在下文被提到:

[0180] 虽然已经在装置的上下文中描述了一些方面,但显然,这些方面还表示对应的方法的描述,其中块或装置对应于方法步骤或方法步骤的特征。类似地,在方法步骤的上下文中描述的方面还表示对应装置的对应块或项目或特征的描述。

[0181] 根据某些实施要求,本发明的实施例可以以硬件或软件实施。可使用具有存储于其上的电子可读控制信号的数字存储介质,例如软盘、DVD、CD、ROM、PROM、EPROM、EEPROM或闪存,执行实施,电子可读控制信号与(或能够与)可编程计算机系统协作,从而执行各个方法。

[0182] 根据本发明的一些实施例包括具有电子可读控制信号的数据载体,电子可读控制信号能够与可编程计算机系统协作,从而执行本文中描述的方法中的一个。

[0183] 一般地,本发明的实施例可被实施为具有程序代码的计算机程序产品,程序代码可操作用于当计算机程序产品在计算机上运行时执行方法中的一个。程序代码可以,例如,存储于机器可读载体上。

[0184] 其他实施例包括存储于机器可读载体或者非易失性存储介质上的用于执行本文描述的方法中的一个的计算机程序。

[0185] 换言之,本发明的方法的实施例因此为具有程序代码的计算机程序,该程序代码用于当计算机程序在计算机上运行时执行本文描述的方法中的一个。

[0186] 本发明方法的进一步实施例因此为数据载体(如数字存储介质,或计算机可读介质),其包括记录于其上的用于执行本文描述的方法中的一个的计算机程序。

[0187] 本发明方法的进一步实施例因此为数据流或信号序列,其表示用于执行本文描述的方法中的一个的计算机程序。数据流或信号序列可以是,例如被配置为通过数据通信连接,例如,通过因特网,进行传送。

[0188] 进一步实施例包括处理装置,例如,计算机或可编程逻辑装置,其被配置为或适于执行本文描述的方法中的一个。

[0189] 进一步实施例包括计算机,其具有安装于其上用于执行本文描述的方法中的一个的计算机程序。

[0190] 在一些实施例中,可使用可编程逻辑设备(例如,现场可编程门阵列)执行本文描述的方法的一些或全部功能。在一些实施例中,现场可编程门阵列可与微处理器协作以执行本文描述的方法中的一个。通常,方法优选地被硬件装置执行。

[0191] 尽管已经根据一些实施例描述本发明,但是存在落在本发明的范围内的替换、变换和等同。还应注意的是,存在用于实现本发明的方法和组成的很多替换方式。因此,应理解的是,后面所附的权利要求被解释为包括落在本发明的真实精神和范围内的所有这些替换、变换和等同。

[0192] 参考文献:

[0193] [1] Surround Sound Explained-Part 5. Published in: soundonsound magazine, December 2001.

[0194] [2] ISO/IEC IS 23003-1, MPEG audio technologies-Part 1: MPEG Surround.

[0195] [3] ISO/IEC IS 23003-3, MPEG audio technologies-Part 3: Unified speech and audio coding.

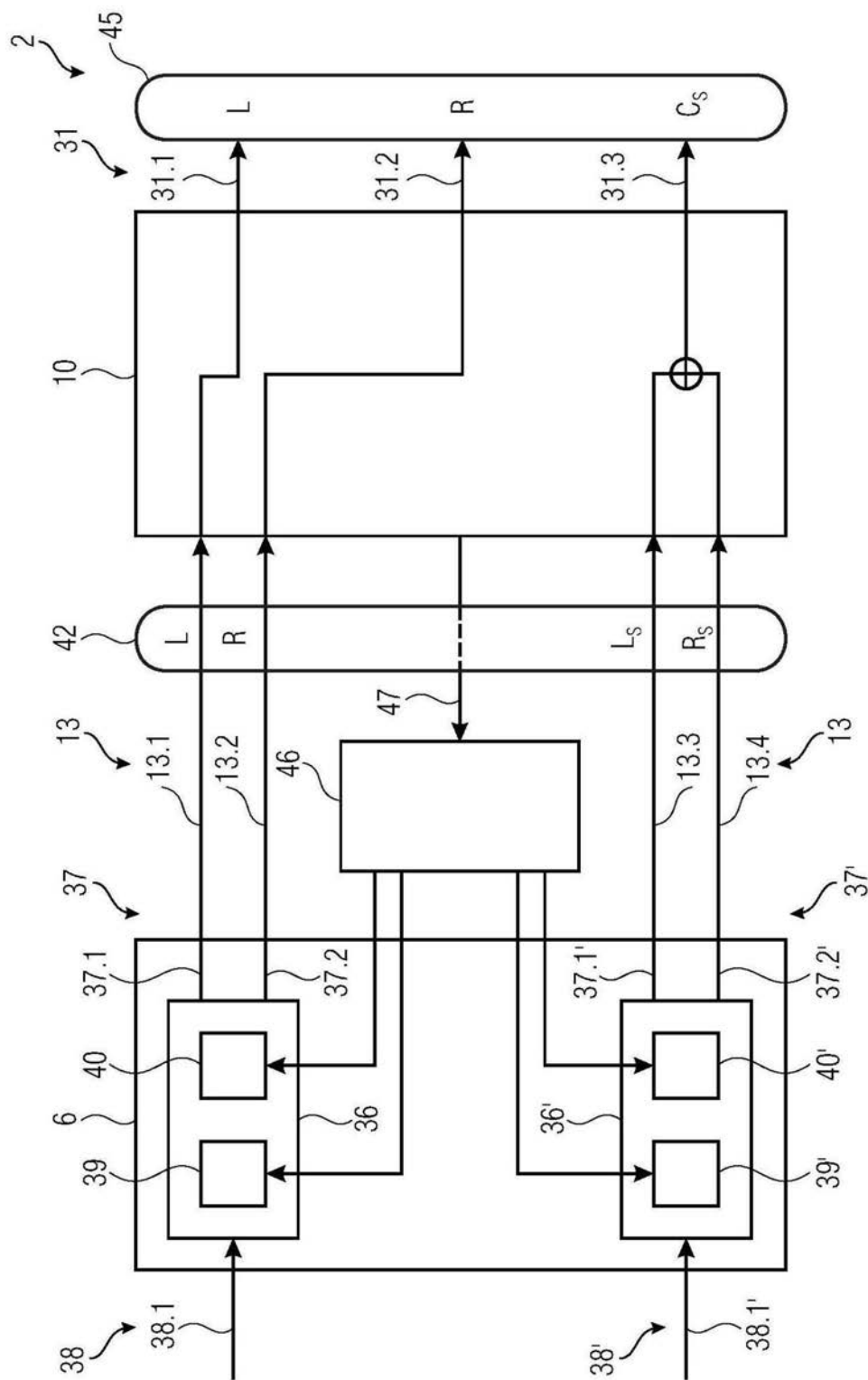


图1

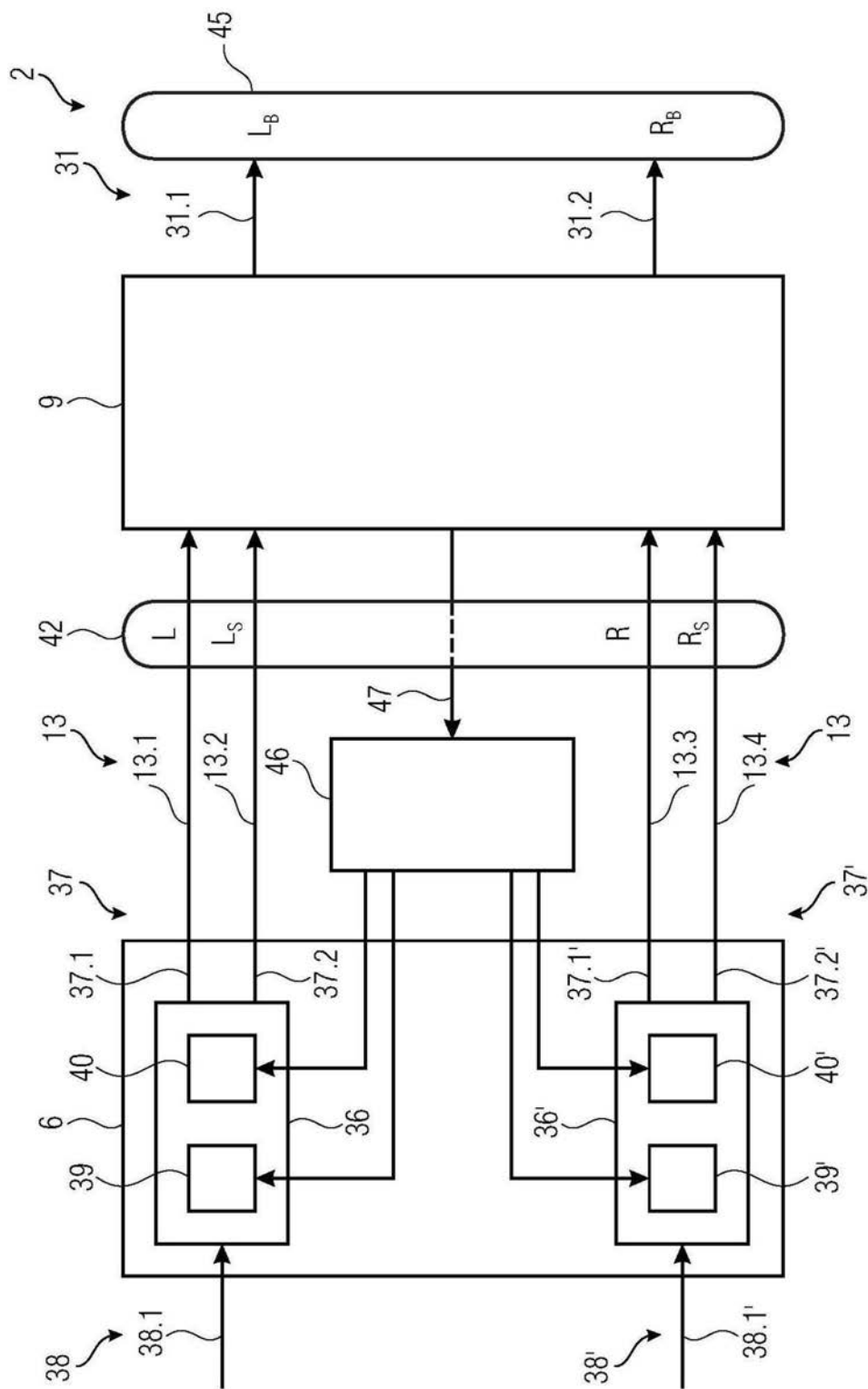


图2

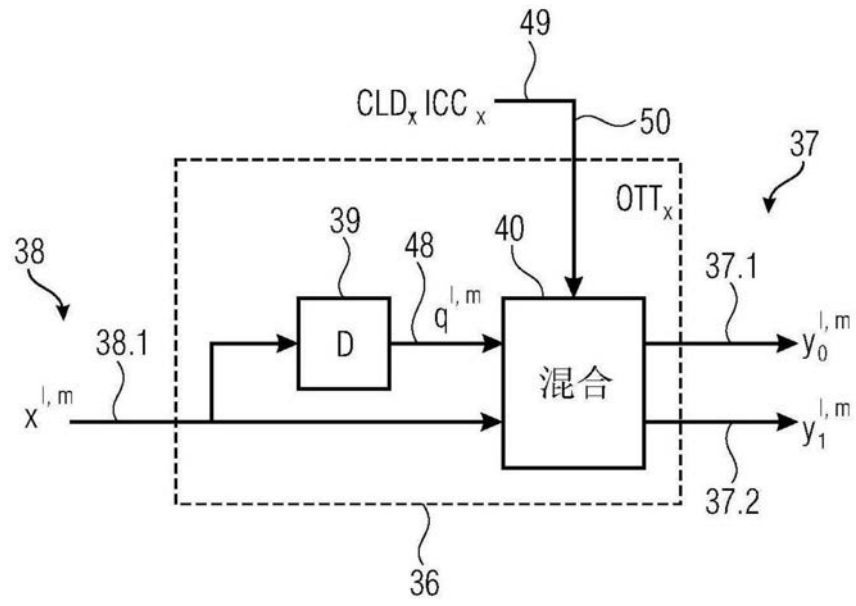


图3

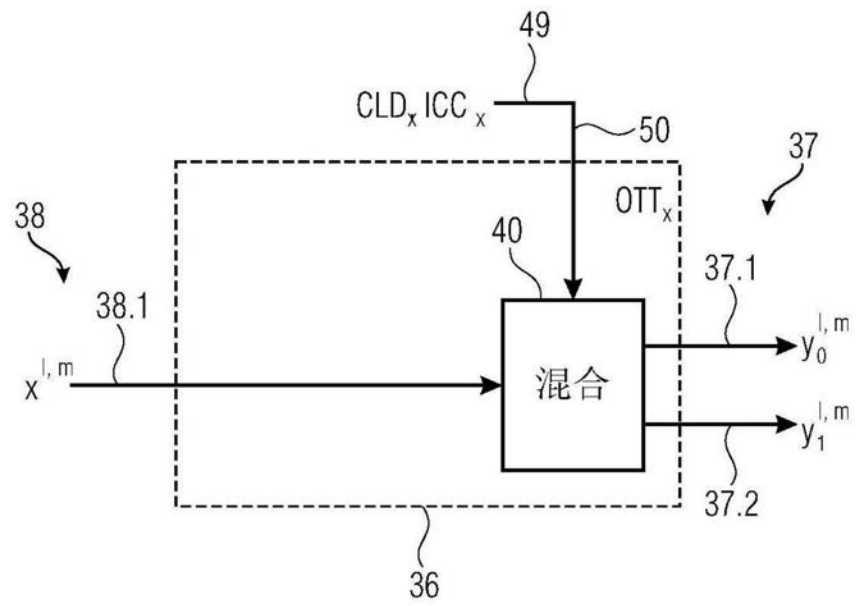


图4

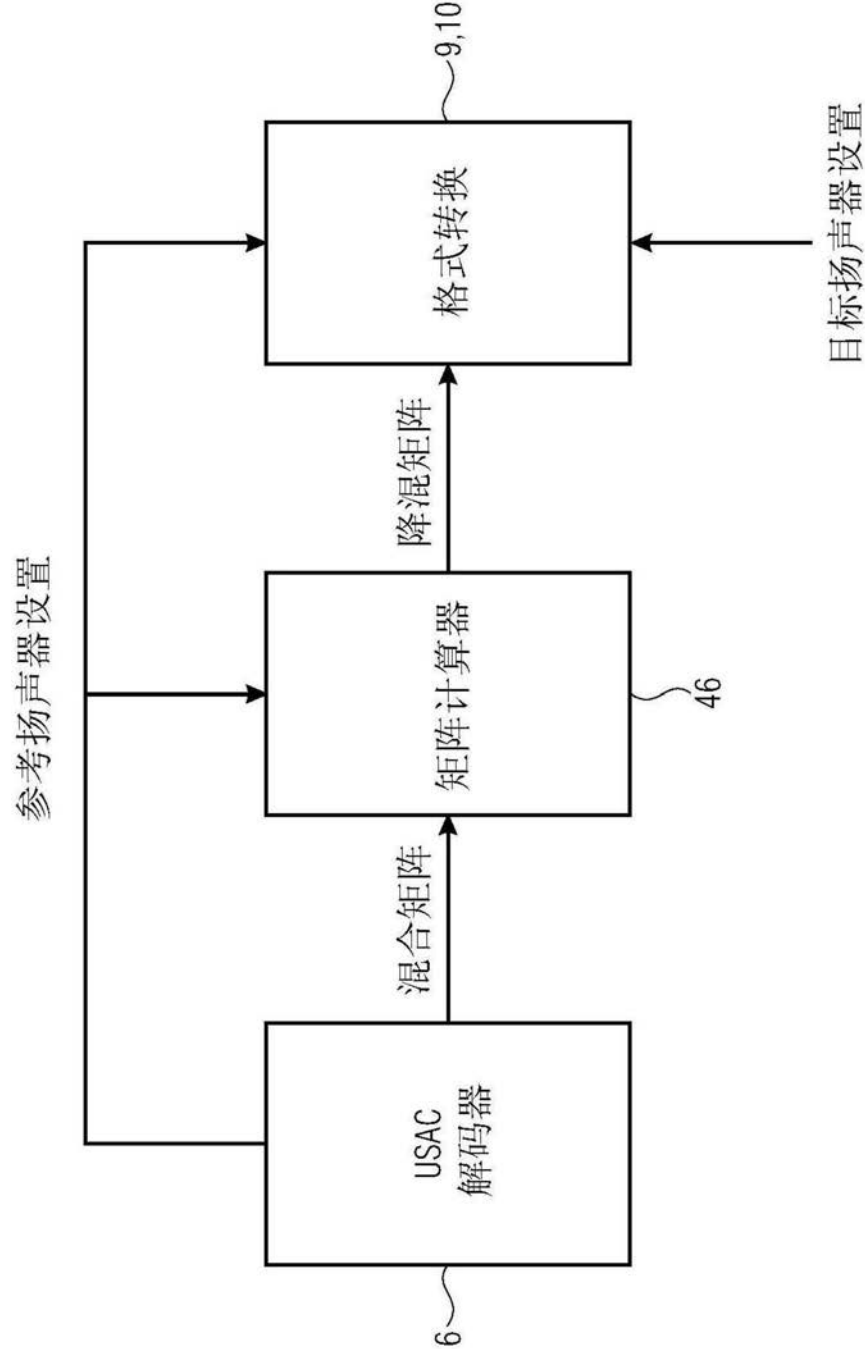


图5

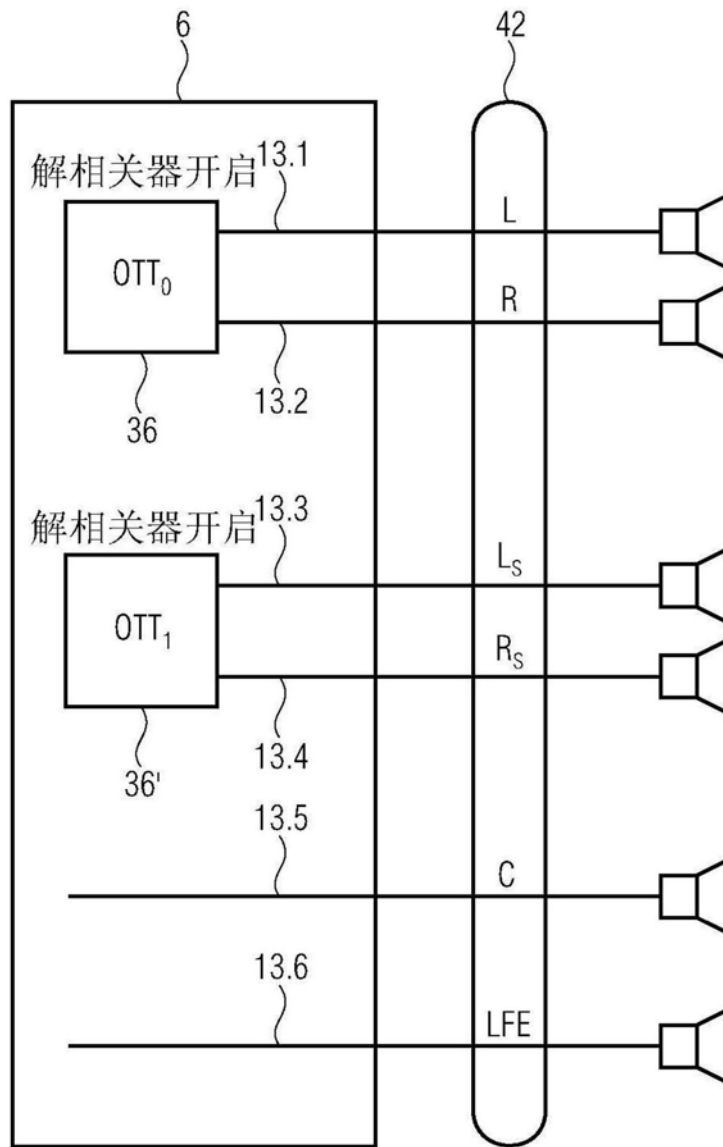


图6

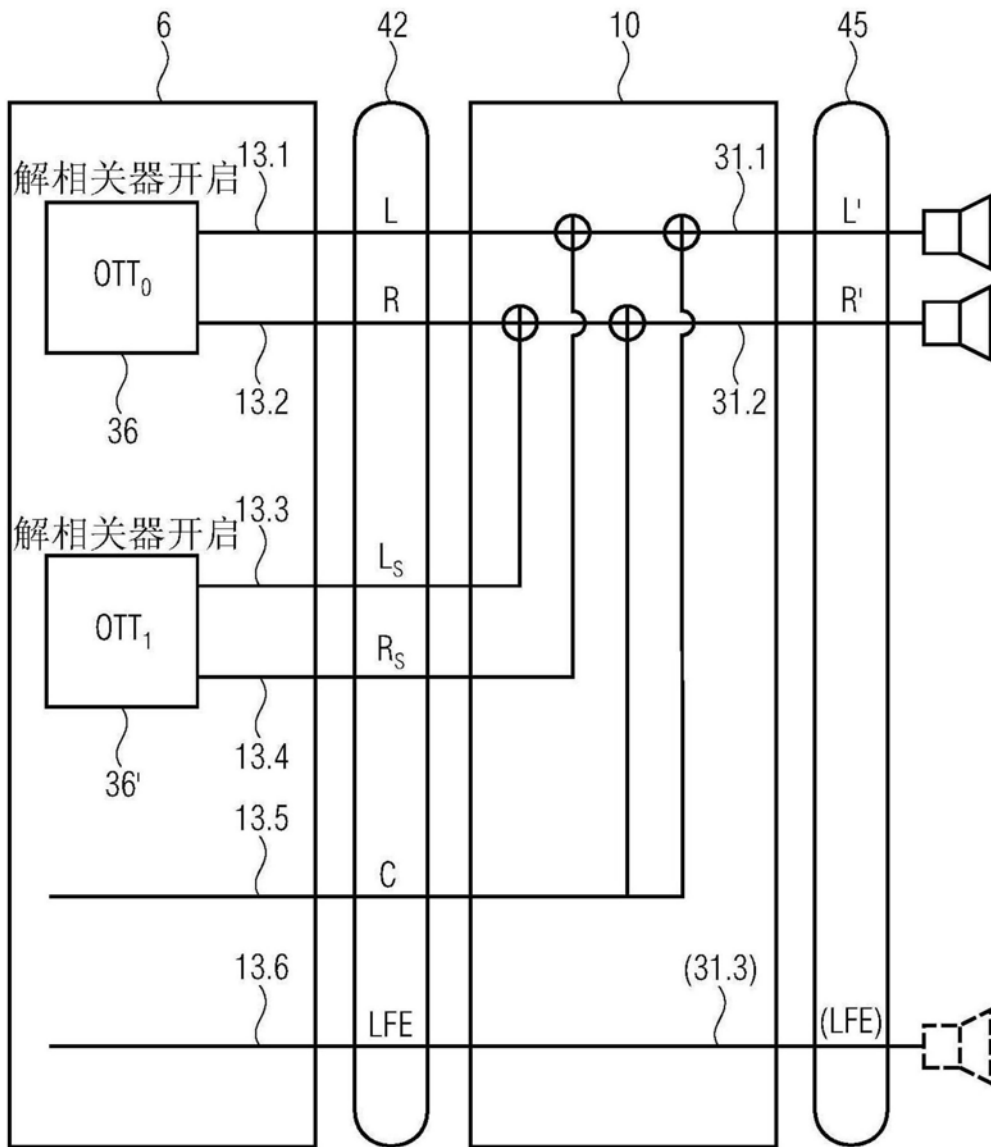


图7

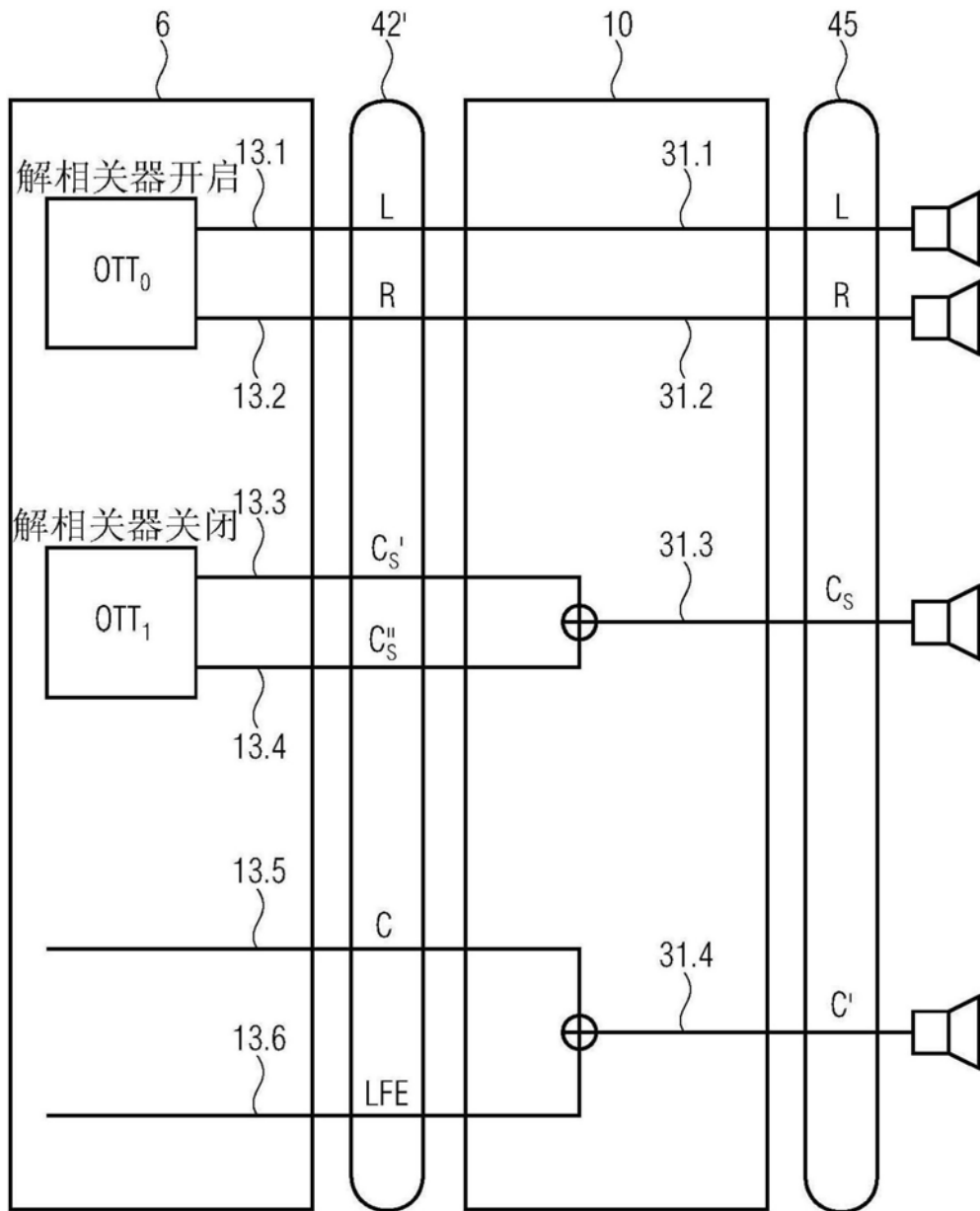


图8

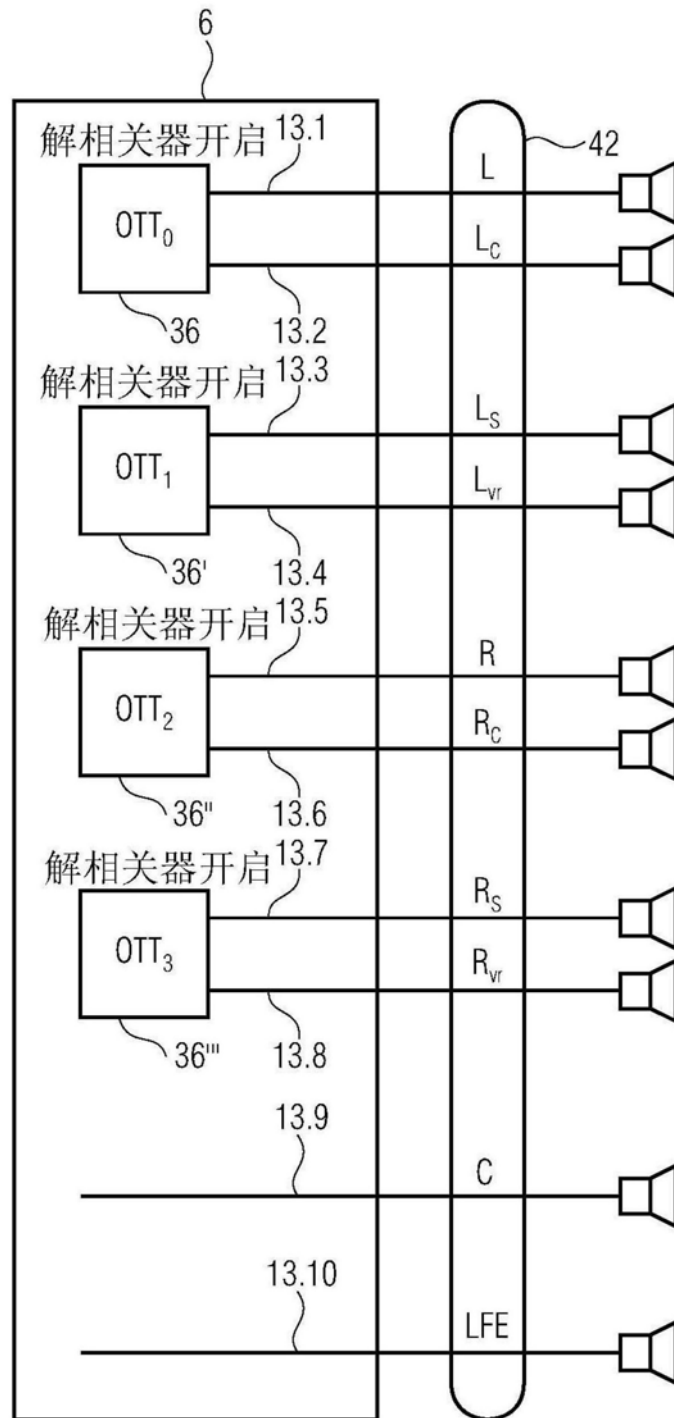


图9

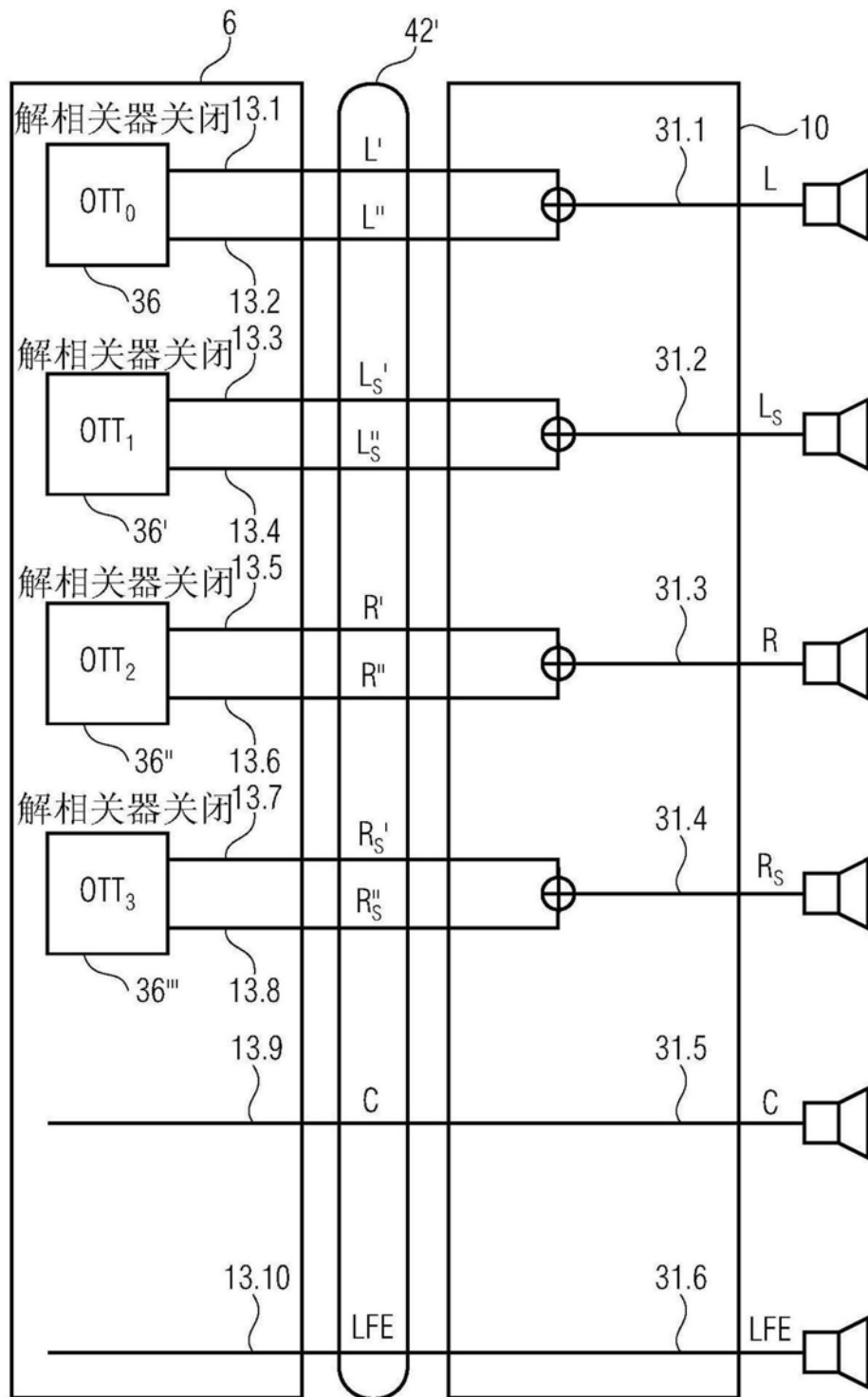


图10

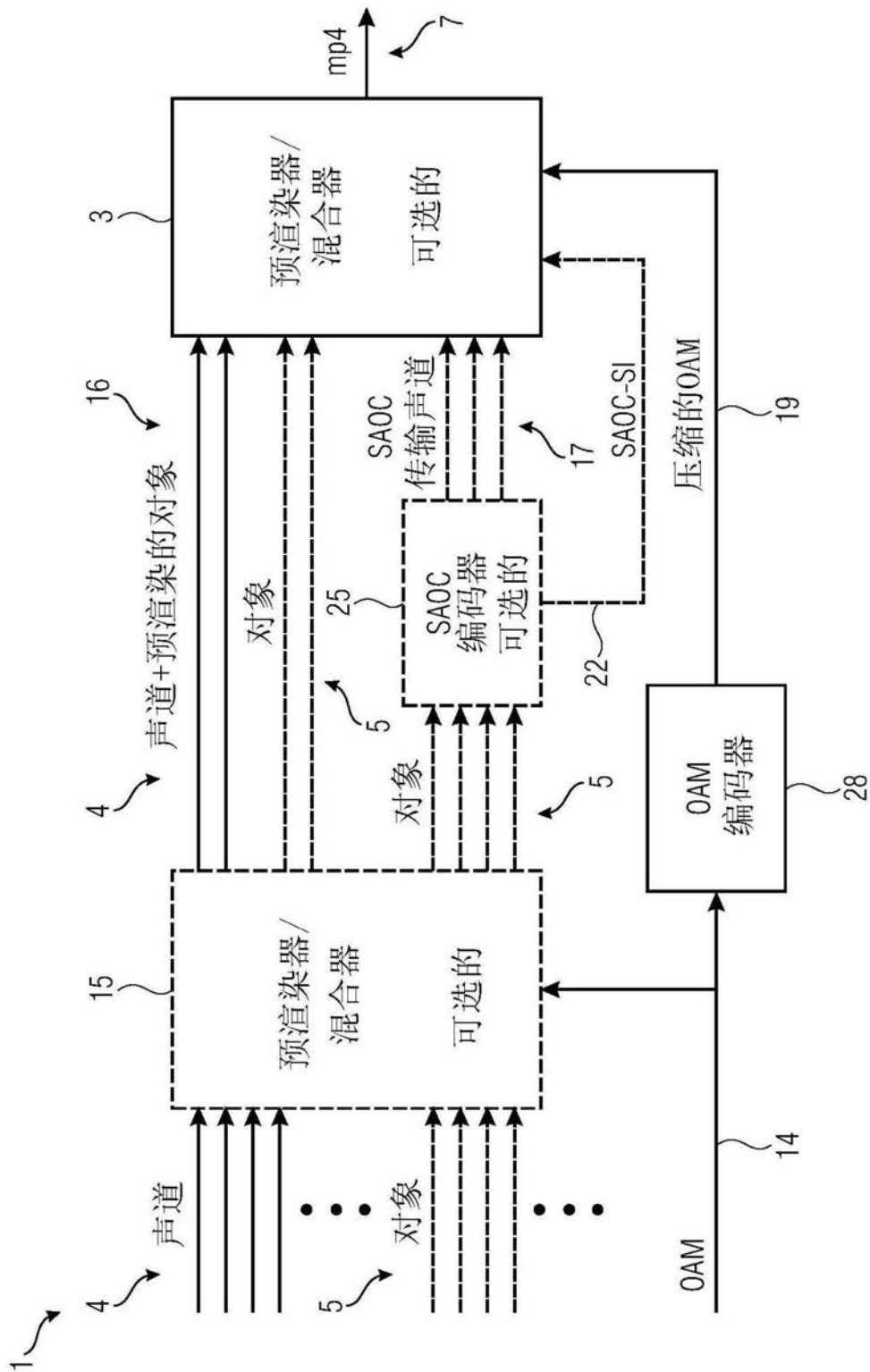


图11

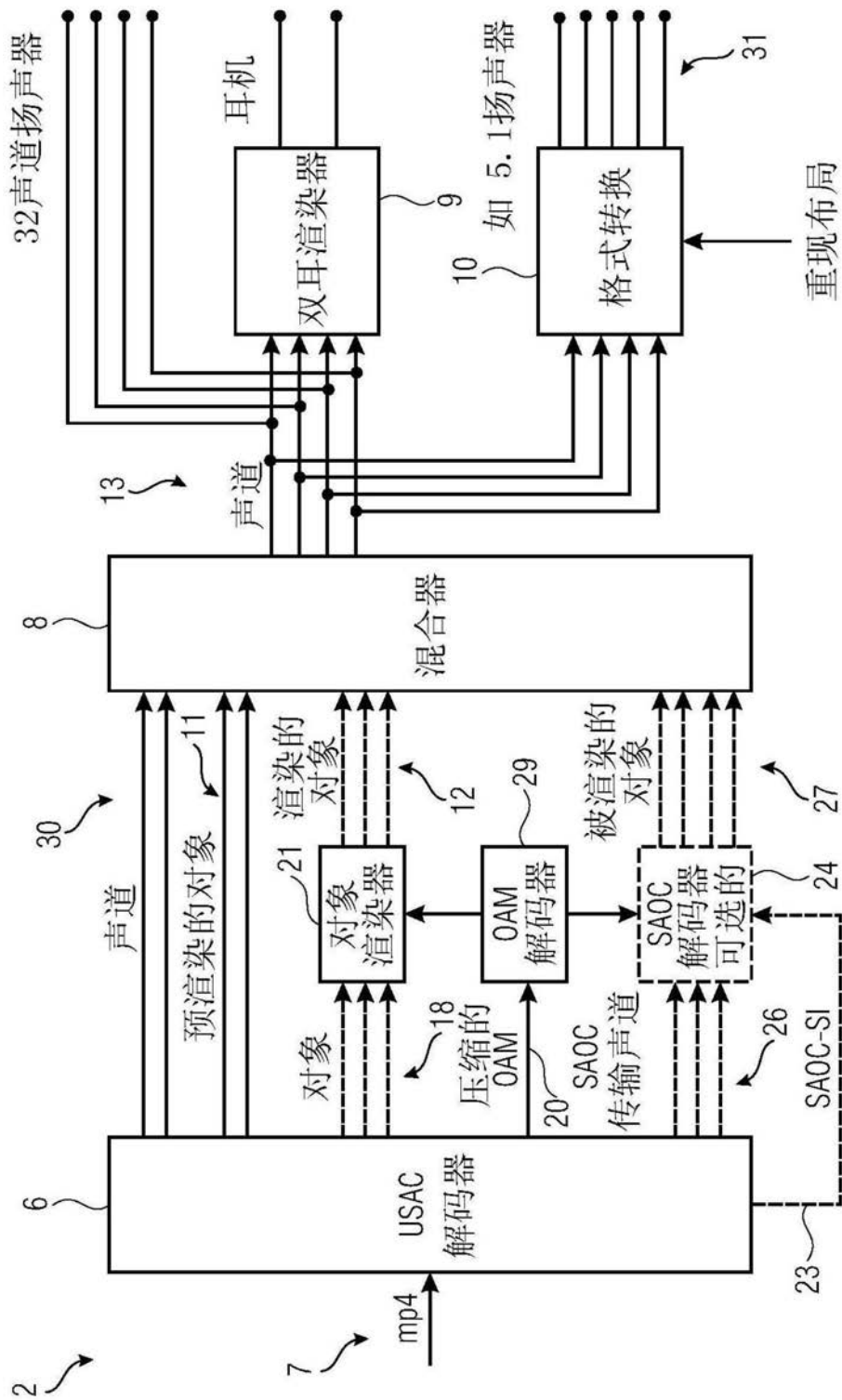


图12

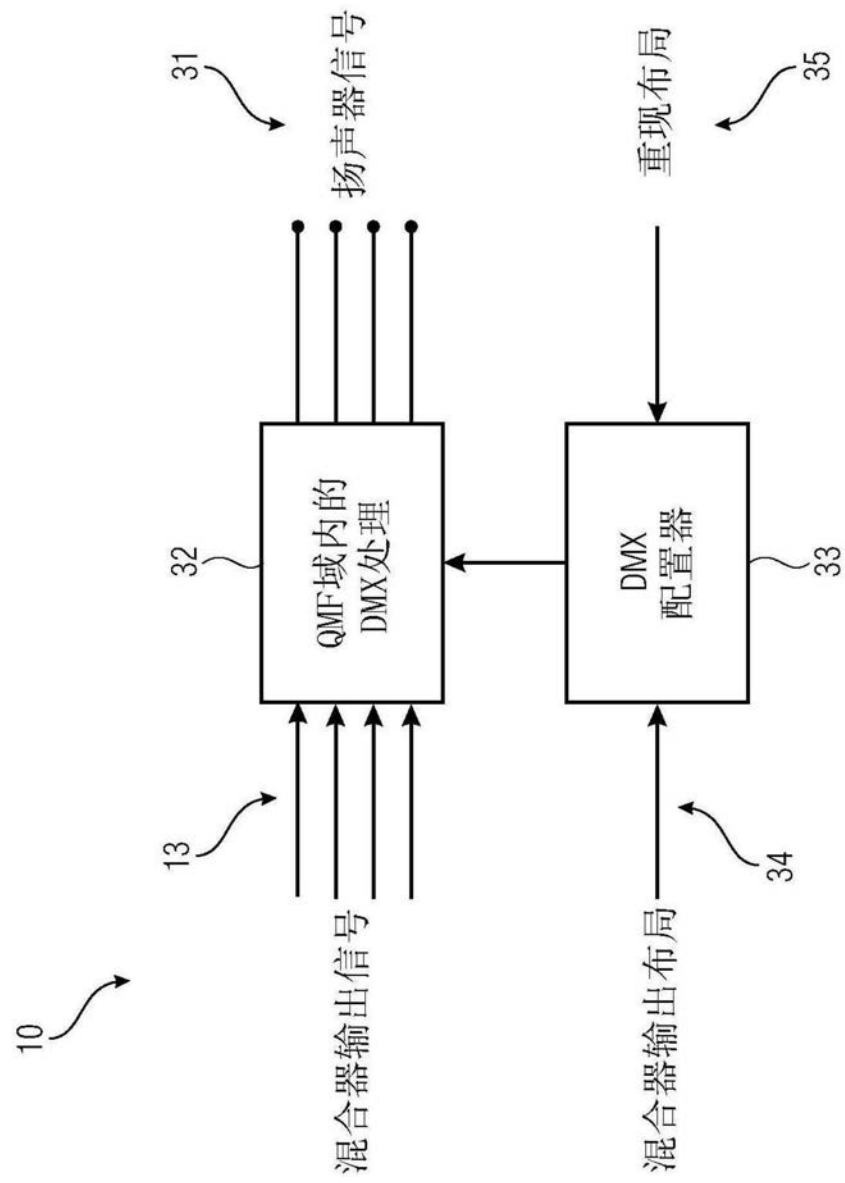


图13