(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 120077434 A
(43)申请公布日 2025.05.30

(21)申请号 202380071307.8

(22)申请日 2023.09.15

(30)优先权数据
63/378,498 2022.10.05 US
63/578,537 2023.08.24 US

(85)PCT国际申请进入国家阶段日
2025.04.07

(86)PCT国际申请的申请数据
PCT/US2023/074317 2023.09.15

(87)PCT国际申请的公布数据
WO2024/076829 EN 2024.04.11

(71)申请人 杜比实验室特许公司
地址 美国
申请人 杜比国际公司

(72)发明人 K·孔灵 H·普恩哈根 D·古那万 B·索斯威尔 L·萨米尔森

(74)专利代理机构 中国贸促会专利商标事务所有限公司 11038
专利代理师 宿小猛

(51)Int.Cl.
G10L 19/008(2006.01)
G10L 19/16(2006.01)
H04S 3/00(2006.01)

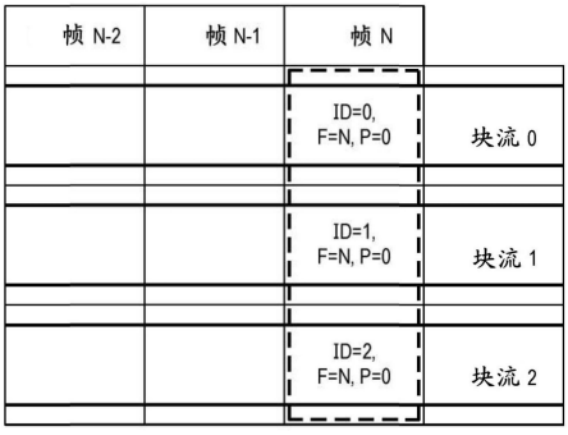
权利要求书2页 说明书24页 附图18页

(54)发明名称

用于音频比特流和关联回声参考信号的编码和解码的方法、装置和介质

(57)摘要

一种用于生成包括多个音频信号的音频节目的编码比特流的帧的方法,其中所述帧包括两个或更多个独立的编码数据块,所述方法包括对于所述多个音频信号中的一个或多个音频信号,接收指示与所述一个或多个音频信号相关联的回放设备的信息;对于所指示的回放设备,接收指示一个或多个附加的关联回放设备的信息;接收与所指示的一个或多个附加的关联回放设备相关联的一个或多个音频信号;对与所述回放设备相关联的所述一个或多个音频信号进行编码;对与所指示的一个或多个附加的关联回放设备相关联的所述一个或多个音频信号进行编码;将与所述回放设备相关联的所述一个或多个编码音频信号和指示所述一个或多个附加的关联回放设备的信令信息组合成第一独立块;将与所述一个或多个附加的关联回放设备相关联的所述一个或多个编码音频信号组合成一个或多个附加的独立块;以及将所述第一独立块和所述一个或多个附加的独立块组合成所述编码比特流的所述帧。



CN 120077434 A

1.一种用于生成包括多个音频信号的音频节目的编码比特流的帧的方法,其中所述帧包括两个或更多个独立的编码数据块,所述方法包括:

对于所述多个音频信号中的一个或多个音频信号,接收指示与所述一个或多个音频信号相关联的回放设备的信息;

对于所指示的回放设备,接收指示一个或多个附加的关联回放设备的信息;

接收与所指示的一个或多个附加的关联回放设备相关联的一个或多个音频信号;

对与所述回放设备相关联的所述一个或多个音频信号进行编码;

对与所指示的一个或多个附加的关联回放设备相关联的所述一个或多个音频信号进行编码;

将与所述回放设备相关联的所述一个或多个编码音频信号和指示所述一个或多个附加的关联回放设备的信令信息组合成第一独立块;

将与所述一个或多个附加的关联回放设备相关联的所述一个或多个编码音频信号组合成一个或多个附加的独立块;以及

将所述第一独立块和所述一个或多个附加的独立块组合成所述编码比特流的所述帧。

2.如权利要求1所述的方法,其中,所述多个音频信号包括与所述回放设备或所述一个或多个附加的关联回放设备不相关联的一组或多组音频信号,所述方法还包括:

将与所述回放设备或所述一个或多个附加的关联回放设备不相关联的一组或多组音频信号中的每一组音频信号编码成相应的独立块;以及

将所述一个或多个组中的每一组的相应独立块组合成所述编码比特流的所述帧。

3.如权利要求1或2所述的方法,其中与所指示的一个或多个附加的关联回放设备相关联的一个或多个音频信号专门旨在用作用于对回放设备执行回声管理的回声参考。

4.如权利要求3所述的方法,其中,与和所述回放设备相关联的所述一个或多个音频信号相比,使用更少的数据来传送旨在用作回声参考的所述一个或多个音频信号。

5.如权利要求3或4所述的方法,其中使用参数编码工具对旨在用作回声参考的所述一个或多个音频信号进行编码。

6.如权利要求1或2所述的方法,其中与所述一个或多个其他回放设备相关联的所述一个或多个音频信号适合于从所述一个或多个其他回放设备回放。

7.一种用于从编码比特流的帧中解码与回放设备相关联的一个或多个音频信号的方法,其中所述帧包括两个或更多个独立的编码数据块,其中所述回放设备包括一个或多个麦克风,所述方法包括:

从所述编码比特流中识别对应于与所述回放设备相关联的所述一个或多个音频信号的独立的编码数据块;

从所述编码比特流中提取所识别的独立的编码数据块;

从所识别的独立的编码数据块中提取与所述回放设备相关联的所述一个或多个音频信号;

从所述编码比特流识别对应于与一个或多个其它回放设备相关联的一个或多个音频信号的一个或多个其它独立的编码数据块;

从所述一个或多个其他独立的编码数据块中提取与所述一个或多个其他回放设备相关联的所述一个或多个音频信号;

使用所述回放设备的所述一个或多个麦克风捕获一个或多个音频信号;以及

使用与所述一个或多个其他回放设备相关联的所提取的一个或多个音频信号作为回声参考,用于响应于所捕获的一个或多个音频信号对所述回放设备执行回声管理。

8.如权利要求7所述的方法,所述方法还包括:

确定所述编码比特流包括一个或多个附加的独立的编码数据块;以及

忽略所述一个或多个附加的独立的编码数据块。

9.如权利要求8所述的方法,其中忽略所述一个或多个附加的独立的编码数据块包括跳过所述一个或多个附加的独立的编码数据块而不提取所述一个或多个附加的独立的编码数据块。

10.如权利要求7至9中任一项所述的方法,其中,与所述一个或多个其他回放设备相关联的所述一个或多个音频信号专门旨在用作用于对回放设备执行回声管理的回声参考。

11.如权利要求10所述的方法,其中,与和所述回放设备相关联的所述一个或多个音频信号相比,使用更少的数据来传送专门旨在用作回声参考的所述一个或多个音频信号。

12.如权利要求10或11所述的方法,其中,根据所述一个或多个音频信号的参数表示来重构专门旨在用作回声参考的所述一个或多个音频信号。

13.如权利要求7至9中任一项所述的方法,其中,与所述一个或多个其他回放设备相关联的所述一个或多个音频信号适合于从所述一个或多个其他回放设备回放。

14.如权利要求7所述的方法,其中,所述编码信号包括指示所述一个或多个其他回放设备用作所述回放设备的回声参考的信令信息。

15.如权利要求14所述的方法,其中,对于当前帧的由所述信令信息指示的所述一个或多个其他回放设备不同于对于前一帧的用作回声参考的一个或多个其他回放设备。

16.一种装置,被配置为执行如权利要求1-15中任一项所述的方法。

17.一种非暂态计算机可读存储介质,包括指令序列,所述指令序列在被执行时使得一个或多个设备执行如权利要求1-15中任一项所述的方法。

用于音频比特流和关联回声参考信号的编码和解码的方法、装置和介质

[0001] 相关申请的交叉引用

[0002] 本申请要求2022年10月5日提交的美国临时申请第63/378,498号和2023年8月24日提交的美国临时申请第63/578,537号的优先权,这些申请均通过引用整体并入本文。

技术领域

[0003] 本公开总体上涉及音频信号处理,并且更具体地涉及用于设备之间沉浸式音频节目的音频信号的低延迟交换的音频源编码和解码。

背景技术

[0004] 音频流式传输在当今社会中是常见的。随着用户对质量的期望提高,音频流式传输的要求变得越来越高,而且用户设置也变得更加复杂,具有大量扬声器以及不同的扬声器类型的,可能在相同设置内。流式传输通常至少部分地在无线链路上进行,这然后要求无线链路具有良好质量,然而如许多人可能已经经历的那样,情况并不总是如此。

[0005] 因此,需要定义用于用例的交换格式,其中某一格式被从云/服务器流式传输,并且随后在设备上转码为更合适的低延迟格式以便于在无线(或者,在一些情况下有线)链路上分发。示例性用例是家庭内连接,以及电话到汽车连接,然而该格式在期望音频信号从单个设备低延迟分发到一个或多个连接设备的任何场景中可能是有益的。

[0006] 除了音频信息被无线地发送和流式传输之外,还可以有其他类型的信息被并入流中。这样的其他类型的信息也会受无线链路质量影响,并且可能具有与音频类似的缺点。

[0007] 因此,克服与用于与其他类型的信息或信号组合的不同类型的流音频的无线流式传输相关联的问题将是有利的。

发明内容

[0008] 本公开的一个目的是克服至少部分地关于与其他类型的信息组合的音频的无线流式传输的上述问题。

[0009] 根据本公开的第一方面,提供了一种用于生成包括多个音频信号的音频节目的编码比特流的帧的方法,其中所述帧包括两个或更多个独立的编码数据块,所述方法包括对于所述多个音频信号中的一个或多个音频信号,接收指示与所述一个或多个音频信号相关联的回放设备的信息;对于所指示的回放设备,接收指示一个或多个附加的关联回放设备的信息;接收与所指示的一个或多个附加的关联回放设备相关联的一个或多个音频信号;对与所述回放设备相关联的所述一个或多个音频信号进行编码;对与所指示的一个或多个附加的关联回放设备相关联的所述一个或多个音频信号进行编码;将与所述回放设备相关联的所述一个或多个编码音频信号和指示所述一个或多个附加的关联回放设备的信令信息组合成第一独立块;将与所述一个或多个附加的关联回放设备相关联的所述一个或多个编码音频信号组合成一个或多个附加的独立块;以及将所述第一独立块和所述一个或多个

附加的独立块组合成所述编码比特流的所述帧。

[0010] 根据本公开的第二方面,提供了一种用于从编码比特流的帧中解码与回放设备相关联的一个或多个音频信号的方法,其中所述帧包括两个或更多个独立的编码数据块,其中所述回放设备包括一个或多个麦克风,所述方法包括:从所述编码比特流中识别对应于与所述回放设备相关联的所述一个或多个音频信号的独立的编码数据块;从所述编码比特流中提取所识别的独立的编码数据块;从所识别的独立的编码数据块中提取与所述回放设备相关联的所述一个或多个音频信号;从所述编码比特流识别对应于与一个或多个其它回放设备相关联的一个或多个音频信号的一个或多个其它独立的编码数据块;从所述一个或多个其他独立的编码数据块中提取与所述一个或多个其他回放设备相关联的所述一个或多个音频信号;使用所述回放设备的所述一个或多个麦克风捕获一个或多个音频信号;以及使用与所述一个或多个其他回放设备相关联的所提取的一个或多个音频信号作为回声参考,用于响应于所捕获的一个或多个音频信号对所述回放设备执行回声管理。

[0011] 根据本公开的第三方面,提供了一种被配置为执行第一和/或第二方面中任一者的装置。

[0012] 根据本公开的第四方面,提供了一种包括指令序列的非暂时性计算机可读存储介质,所述指令序列在被执行时使一个或多个设备执行第一和/或第二方面中任一者的方法。

[0013] 本公开的进一步示例在从属权利要求中被限定。

[0014] 在一些示例中,与关联回放设备相关联的一个或多个音频信号专门旨在用作用于对回放设备执行回声管理的回声参考。

[0015] 在一些示例中,与和所述回放设备相关联的一个或多个音频信号相比,使用更少的数据来传送旨在用作回声参考的一个或多个音频信号。

[0016] 在本公开中,帧表示所有信号的整体的时间片。块流表示会话持续时间内的信号集合。块表示块流的一个帧。对于具有给定采样频率的数字音频,帧大小等同于任何音频信号的帧中音频样本的数量。帧大小通常在会话的持续时间内保持恒定。

[0017] 唤醒词可以包括一个词,或者包含处于固定顺序的两个或更多个词的短语。

[0018] 贯穿包括在权利要求中的本公开,表述“系统”在广义上用于表示设备、系统或子系统。例如,实现解码器的子系统可以被称为解码器系统,并且包括这样的子系统的系统(例如,响应于多个输入生成X个输出信号的系统,其中子系统生成M个输入,而其他X-M个输入被从外部源接收)也可以被称为解码器系统。

附图说明

[0019] 将参照附图更详细地描述本公开的示例。在以下附图中,相同的附图标记用于指代相同的元件。尽管以下附图描绘了各种示例,但是一个或多个实施方式并不限于附图中描绘的示例。

[0020] 图1示出了家庭内连接低延迟转码的示例。

[0021] 图2示出了汽车连接音频流式传输的示例。

[0022] 图3示出了使用无线设备的增强TV音频流式传输的示例。

[0023] 图4示出了具有简单无线扬声器的音频流式传输的示例。

[0024] 图5示出了从比特流元素到设备的元数据映射的示例。

[0025] 图6示出了如何部署音频流式传输的简单灵活渲染的示例。

[0026] 图7示出了灵活渲染数据到比特流元素的映射的示例。

[0027] 图8示出了当用多个设备播放音频和收听命令时回声参考的信令的示例。

[0028] 图9示出了帧、块和分组如何彼此相关的示例。

[0029] 图10示出了当前MPEG-4结构中的另外的编解码器信息的示例。

[0030] 图11示出了当前MPEG-4结构中的还另外的编解码器信息的示例。

[0031] 图12示出了集成收听能力和语音识别的示例。

[0032] 图13示出了包括多个块的帧的示例。

[0033] 图14示出了包括多个块的比特流的示例。

[0034] 图15示出了具有不同优先级的块的示例。

[0035] 图16示出了包括具有不同优先级的多个块的比特流的示例。

[0036] 图17示出了具有不同优先级的帧的示例。

[0037] 图18示出了包括具有不同优先级的帧的比特流的示例。

具体实施方式

[0038] 现在将参考附图中示出的各种示例来描述本发明的原理。应当认识到,对这些实施例的描绘仅仅是为了使本领域技术人员能够更好地理解并进一步实现本发明,而不旨在以任何方式限制本发明的范围。

[0039] 在图1中,沉浸式音频流从云或服务器10流式传输,并且在TV或集线器装置20上解码。沉浸式音频流可被以任何现有格式编码,包括例如Dolby Digital Plus、AC-4等。输出随后被转码成低延迟交换格式,以用于进一步传输到连接的设备30,该设备优选地通过本地无线连接(例如WiFi软接入点或蓝牙连接)连接。低延迟通常依赖于各种因素,例如帧大小、采样率、硬件和/或软件计算资源等,但低延迟通常小于40ms、20ms或10ms。

[0040] 在另一个典型用例中,电话从云或服务器获取沉浸式音频流,将之转码为交换格式,随后传输到连接的汽车。在图2所示的示例中,移动设备(例如,电话或平板电脑)20连接到服务器10并接收沉浸式音频流,执行到低延迟交换格式的低延迟转码,并将转码后的信号传输到支持沉浸式音频回放的汽车30。沉浸式音频流的示例是包括Dolby Atmos格式的音频的流,并且支持沉浸式音频回放的汽车30的示例是被配置为回放Dolby Atmos沉浸式格式的汽车。

[0041] 通常,交换格式优选地具有低延迟、低编码和解码复杂度、可缩放到高质量的能力、以及合理的编码效率。该格式优选地还支持可配置的延迟,使得延迟可以与效率以及容错率进行折中,以便在变化的连接状况下可操作。

[0042] 具有带有或不带有收听能力的无线扬声器的集线器或增强显示器

[0043] 图3中示出了如下示例,驱动一组无线扬声器30的集线器20,或可能具有内置扬声器30的显示器(例如,电视机或TV)20,该显示器20可被扩充有若干无线扬声器30。扩充(augmentation)表明,在显示器20包括扬声器30的示例中,显示器20也是音频再现的一部分。无线扬声器/设备30可以接收相同的完整信号(广播模式,图3中的左侧示出)或为特定设备定制的单独流(单播-多点模式,图3中的右侧示出)。

[0044] 每个扬声器可以包括多个驱动器,覆盖同一声道的不同频率范围,或者在规范映

射中对应于不同声道。例如,扬声器可以具有两个驱动器,其中一个驱动器可以是向上发射驱动器,其输出对应于高度声道的信号以仿真高架扬声器。无线设备可以具有收听能力(例如,“智能扬声器”),并且因此可能需要回声管理,这可能需要扬声器接收一个或多个回声参考。回声参考可以是本地的(例如,对于相同的扬声器/设备而言),或者可替代地可表示来自附近的其他扬声器/设备的相关信号。

[0045] 扬声器可以被放置在任意位置,在这种情况下,可以执行所谓的“灵活渲染”,由此该渲染考虑了扬声器的实际位置(例如,与规范假设相反,其中扬声器被假设位于固定的、预定义的位置)。灵活渲染可以在集线器或TV中发生,随后,渲染后的信号或者以广播模式传送到扬声器/设备,其中将每个渲染后的信号传送到每个设备,并且每个相应的设备提取并输出合适的信号,或者作为单独的流传送到单独的设备。作为替代,灵活渲染可以在每个设备上本地发生,由此每个设备接收完整沉浸式节目的表示,例如,基于7.1.4声道的沉浸式表示,并且根据该表示渲染适合于相应设备的输出信号。

[0046] 无线设备可以通过设备提供的软接入点或者通过家庭中的本地接入点连接到集线器或TV。这可能对比特率和延迟施加不同的要求。

[0047] 汽车应用中的移动设备投影

[0048] 在另一个用例中,移动设备(例如,电话或平板电脑)20从云或服务器10获取沉浸式音频流,将沉浸式音频流转码成交换格式,并且随后将转码后的流传送到连接的汽车30。在图2的示例中,支持Dolby Atmos的电话20连接到服务器10并接收Dolby Atmos流,用于低延迟转码和传输到支持Dolby Atmos的汽车3。在该用例中,可利用比起居室用例中更高的比特率,并且无线信道还可以存在不同特性。汽车用例中较高比特率可能归因于噪声较少的无线环境,因为与起居室相比,汽车充当一种法拉第笼并且屏蔽车内环境免受外部无线干扰,而在起居室中,诸如来自邻居、其他房间等的所有无线设备都向无线环境添加噪声。同时,与家庭中的所有无线设备(包括起居室无线设备)通常都竞争可用带宽的起居室用例相比,汽车中通常有非常少的无线设备来竞争可用的无线带宽。

[0049] 对于此用例,预想到要由移动装置转码且传送到汽车的信号可为基于信道的沉浸式表示、基于对象的表示、基于场景的表示(例如,Ambisonics表示),或甚至不同表示的组合。对于该示例,不同的渲染架构以及广播对比多点可能不相关,因为完整的呈现通常将从移动设备传递到单个端点(例如,汽车)。

[0050] 交换格式描述

[0051] 沉浸式交换格式建立在具有感知激励量化和编码的修正离散余弦变换(MDCT)上。它具有可配置的延迟,例如支持给定采样速率下的不同变换大小。示例性帧大小是在48kHz和44.1kHz的采样率下的128、256、512、1024和120、240、480、960和192、384、768个样本。

[0052] 该格式可以支持单声道、立体声、5.1和其他声道配置,包括沉浸式声道配置(例如,包括但不限于5.1.2、5.1.4、7.1.2、7.1.4、9.1.6和22.2,或如ISO/IEC 23091-3:2018,表-2中规定的独特声道组成的任何其他声道配置)。该格式还可以支持基于对象的音频和基于场景的表示,诸如Ambisonics(例如,一阶或更高阶)。该格式还可以使用适于与现有格式(例如,诸如也可被称为MPEG-4音频标准的ISO/IEC 14496-3、也可被称为MPEG-4系统标准的ISO14496-1、也可被称为ISO基本媒体文件格式标准的ISO14496-12、和/或也可被称为MP4文件格式标准的ISO14496-14中描述的那些格式)集成的信令方案。

[0053] 此外,系统可以支持跳过一部分语法元素而仅解码给定扬声器的相关部分的能力,支持诸如延迟对齐、电平调整和均衡的元数据控制的灵活渲染方面,支持使用具有收听能力的智能扬声器(包括传送一组信号的场景,允许该组信号独立地馈送智能扬声器中的每一驱动器/扬声器),以及相关联地支持通过回声参考的信号传递进行回声管理,支持具有低重叠窗口以及50%重叠窗口两者的MDCT域中的量化和编码,和/或支持MDCT域中沿着频率轴进行滤波以进行量化噪声的时间整形,例如,TNS(时间噪声整形)。因此,在各种示例中,重叠窗口是对称的或不对称的。

[0054] 在一些示例中,所公开的交换格式可以提供改进的联合信道编码,利用在灵活渲染用例中信号之间的增加的相关性、通过跨信道元素共享的比例因子而改进的编码效率、在MDCT域中包括高频重构和噪声添加技术、相对于传统编码结构的各种改进,以允许更好的效率和对手边用例的适用性。

[0055] 在广播模式下控制各个回放设备的一些示例中,可以使用可跳过块和元数据。在广播模式设置中,每个无线设备将需要为对应于特定设备的特定驱动器/信道播放完整音频的相关部分。这意味着设备需要知道流的哪些部分与该设备相关,并从正在广播的完整音频流中提取那些部分。为了实现低复杂度的解码操作,优选地是流被构造成使得特定设备的解码器可以有效地越过与解码无关的元素而跳到与该设备以及该设备上的给定驱动器相关的元素。

[0056] 然而,应指出,仍然允许目的地为不同设备的信号(例如,在最简单的情况下是立体声呈现的左声道和右声道)之间的联合编码可能是有益的(从压缩角度看)。

[0057] 因此,该格式可以在比特流中启用“可跳过块”以使得能够对仅与特定设备相关的部分进行有效解码,包括如下元数据,其使得能够将一个或多个可跳过块灵活地映射到特定设备,同时保留在对应于不同扬声器/设备的信号之间应用联合编码技术的能力。

[0058] 在图4中,示出了任意设置的示例。存在三个连接的无线设备31、32、33,其中两个是单信道扬声器32、33(意味着音频的一个信道),并且一个扬声器31是具有三个不同驱动器并由此对三个不同信号进行操作的更高级的扬声器31。在该示例中,第一扬声器31对3个单独的信号进行操作,而第二和第三扬声器32、33对立体声表示(例如,左和右)进行操作。如此,对立体声表示的信号进行联合编码可能是有益的。

[0059] 给定上述场景,该格式规定比特流包含可跳过块以使得第一设备可以仅提取流的相关部分以用于解码该扬声器的信号,而解码器复杂度对比效率可以针对立体声对通过构造信号而折中,其中每个扬声器需要解码两个信号以输出单个信号,好的一点是可以执行联合编码。

[0060] 该格式规定了能够实现通用且灵活的表示的元数据格式,其将多个可跳过块中的特定一个块映射到一个或多个设备。这在图5中示出,其中映射可以表示为将每个设备31、32、33关联到一个或多个比特流元素的矩阵,使得用于给定设备的解码器将知道要输出和解码什么比特流元素。

[0061] 例如,在图5的示例中,第一块或跳过块Blk1包括3个单信道元素(设备1 31的每个驱动器各一个),而第二块或跳过块Blk2包含信道对元素,其可以包括要由设备2 32和3 33输出的信号的联合编码版本。设备1 31提取映射元数据,并确定其所需的信号在跳过块1Blk1中。因此,它提取跳过块1Blk1,解码其中的三个单信道元素,并将它们分别提供给驱

动器1、2a和2b。

[0062] 此外,设备1 31忽略跳过块2,Blk2。类似地,设备2 32提取映射元数据,并确定其所需的信号在跳过块2Blk2中。因此,设备2 32跳过跳过块1Blk1,并提取跳过块2Blk2。设备2 32解码信道对元素,并将CPE的左信道输出提供给其驱动器。类似地,设备3 33提取映射元数据并且确定其需要的信号在跳过块2Blk2中,因此设备3 33跳过跳过块1Blk 1并且还提取跳过块2Blk2。设备3 33解码信道对元素,并将CPE的右信道输出提供给其驱动器。

[0063] 在一些示例中,设备2 32可以确定其仅需要来自跳过块2Blk2的信号的子集。在这样的示例中,当可能时,设备2 32可以仅执行完全解码跳过块2Blk2中的信号所需的操作的子集。具体地,在图5的示例中,设备2 32可以仅执行提取CPE的左信道所需的那些处理操作,因此使得能够降低计算复杂度。类似地,设备3 33可以仅执行提取CPE的右信道所需的那些处理操作。

[0064] 例如,在图5的情况下,可能存在使用联合信道编码对CPE进行编码的时间,以及使用独立信道编码对其进行编码的其他时间。在使用独立信道编码来对CPE进行编码的时间期间,装置2 32可仅提取CPE的第一(例如,左)信道,而装置3 33可仅提取CPE的第二(例如,右)信道。

[0065] 在另一示例中,可以使用联合信道编码来对CPE的信道进行编码,在这种情况下,设备2 32和3 33必须提取CPE的两个中间信道。然而,设备2 32仍然能够通过仅执行从中间信道提取左信道所需的那些操作来以降低的计算复杂度进行操作。类似地,设备3 33仍然能够通过仅执行从中间解码信道提取右信道所需的那些操作来以降低的计算复杂度进行操作。

[0066] 依赖于联合信道编码的具体情况,其他优化可能是可能的。可以在系统初始化或设置阶段期间定义不同设备/解码器中的每一者的身份。这样的设置是常见的,并且通常涉及测量房间的声学、扬声器到最佳听音点(sweet-spot)的距离等。

[0067] 灵活渲染(预编码和后解码)的分布和延时等后解码器的应用

[0068] 对于在将相关信号发送到特定设备/扬声器之前在集线器/TV中应用灵活渲染的用例,渲染可能创建从编码角度看更难以编码的信号,例如当考虑此类信号的联合编码时更难以编码的信号。一个原因是灵活渲染可对不同设备应用不同的延时(delay)、均衡和/或增益调整(例如,取决于扬声器相对于其他扬声器和收听者的放置)。还可以在初始设置时预先设置例如增益和延时的信息,并且灵活地仅渲染均衡。在其他示例中,预先设置信息和灵活渲染信息的其他变体也是可能的。应指出在本文中,术语“增益”应被解释为表示任何水平调整(例如,衰减、放大或直通),而不是仅限于某些水平调整(例如,放大)。

[0069] 在图6的示例中,右信道33和左信道32扬声器被给予不同的延迟,反映了相对于收听者的不同放置(例如,由于扬声器32、33可能不是与收听者等距的,因此不同的延迟可以被应用于从扬声器32、33输出的信号,使得来自不同扬声器32、33的相干声音同时到达收听者)。对旨在不同扬声器31、32、33上回放的相干信号引入这样的不同延迟使得这种信号的联合编码具有挑战性。

[0070] 为了解决这个挑战,灵活渲染过程的各方面可以被参数化并在信号的解码之后应用于端点设备中。

[0071] 在图7的示例中,每个设备的延时和增益值被参数化并且被包括在传送到相应扬

声器31、32、33的编码信号中。各个信号可以由相应设备31、32、33解码，该设备然后可以将参数化的增益和延时值引入到相应的解码信号。

[0072] 在一个示例，例如图7所示的示例中，用于不同设备31、32、33的编码信号可在可分离块(例如，可跳过块)中传送，参数(例如，延时和增益)也可在可分离块中传送，使得设备31、32、33可以仅提取该设备31、32、33所需的参数的子集，而忽略(并且跳过)该设备31、32、33不需要的那些参数。在这种情况下，可以向每个设备提供指示哪些参数被包括在哪些块中的映射元数据。

[0073] 还应指出，虽然图7仅指示延时和增益参数，但是也可以包括其他参数，诸如均衡参数。均衡参数可以例如包括将被应用于不同频率区域的多个增益、将由回放设备31、32、33应用的预定均衡曲线的指示、一组或多组无限脉冲响应(IIR)或有限脉冲响应(FIR)滤波器系数、一组双二阶滤波器系数、指定参数均衡器的特性的参数、以及用于规定本领域技术人员已知的均衡的其他参数。

[0074] 此外，灵活渲染方面的参数化不需要是静态的，而是可以是动态的(例如，在音频节目回放期间收听者移动的情况)。因此，允许参数动态地改变可能是优选的。在音频节目期间一个或多个参数改变的情况下，装置可在先前延时及/或增益参数与更新后延时及/或增益参数之间内插以提供平滑转变。这在系统动态地跟踪收听者的位置并且相应地更新动态渲染的最佳听音点的情境下可能尤其有用。

[0075] 还应指出并且将随后描述，当应用灵活渲染时，可以发生信道之间的相关性水平增加，并且这可被更灵活的联合编码利用。

[0076] 回声参考编码和信令

[0077] 对于如图8概述的用例，在该用例中有多个设备/扬声器30一起操作以在图8左侧示出的广播模式下或在图8右侧示出的单播-多点模式下接收信号，同时在设备30上有麦克风40以启用“收听”能力，出现了对回声管理的需要。

[0078] 当对于多个扬声器/设备执行回声管理时，使用来自不仅仅是本地扬声器设备的回声参考可能是有益的。作为示例，一个设备可以靠近另一个设备放置，因此来自附近设备的信号将影响具有有源麦克风的设备的回声管理。在广播模式用例中，每个设备接收所有设备的信号。当设备具有其他设备的信号时，使用那些信号作为回声参考可能是有益的。为此，有必要向特定设备发信号告知哪些信号可以用作其他设备的回声参考。

[0079] 在一个示例中，这可以通过提供如下元数据来完成，所述元数据不仅映射要由特定扬声器/设备播放的(来自整个集合)的信道/信号，而且映射要用作特定扬声器/设备的回声参考的(来自整个集合)的信道/信号。这样的元数据或信令可以是动态的，使得优选回声参考的指示能够随时间变化。

[0080] 对于每个设备/扬声器仅接收其要播放的特定信号的用例，为了提供合适的回声参考，可能需要向每个设备/扬声器传送附加信号(例如，回声参考信号)。此外，为了这样做，需要提供设备特定的信令，以便每个设备可以选择用于播放的合适信号和用于回声管理的合适信号。

[0081] 由于用于回声管理的信号仅由该设备使用以用于执行回声管理，而不由该设备向收听者播放，因此回声参考信号可以与旨在由设备回放的信号被不同地编码或表示。具体地，因为成功的回声管理可以通过以比通常用于向收听者播放的速率更低的速率编码的信

号来实现,所以额外的压缩工具(例如信号的参数表示,其可能根本不适合于向收听者播放,但是捕获了音频信号的必要特征)可以以显著降低的传输成本提供良好的回声管理。

[0082] 分类语法元素

[0083] 在一些示例中,块用于优化音频的传输。在所描述的格式中,每个帧可以被分成块,如上面关于可跳过块所描述的。块可由其所属的帧编号、可用于将具有相同ID的不同帧的连续块关联到块流的块ID、以及重传优先级来标识。上述的一个示例是具有多个帧N-2、N-1和N的流,帧N包括被标识为ID1、ID2、ID3的多个块,其在图13中示出。在图14中示出了此流的示例,其为比特流格式。在一些示例中,对于沉浸式音频节目的音频信号,块的块ID可以指示整个沉浸式音频节目的哪组信号由该块承载。

[0084] 所描述格式的一个用例是音频在诸如Wifi的无线网络上的低延迟的可靠传输。例如,Wifi使用基于分组的网络协议。分组大小通常是有限的。IP网络中典型的最大分组大小是1500字节。流的基于块的架构允许在组装用于传输的分组时的灵活性。例如,具有较小帧的分组可被来自其他帧的重传块填充。大的帧可以在打包之前在块边界上分割,以减少网络协议层上分组之间的依赖性。

[0085] 图9示出了帧、块和分组之间的关系。帧承载音频数据,优选地所有音频数据,其表示具有开始时间、结束时间以及作为结束时间和开始时间之差的持续时间的音频信号的连续段。连续段可包括根据ISO/IEC 14496-3,第4子部分,第4.5.2.1.1节的时间段。第4.5.2.1.1节描述了raw_data_block()的内容。帧还可以承载该段的冗余表示,例如,以较低的数据率编码。在编码之后,该帧可以被分割成块。这些块可以被组合成用于在基于分组的网络上传输的分组。来自不同帧的块可以组合在单个分组中,和/或可以无序地传送。

[0086] 在示例中,块用于寻址单独的设备。数据分组由单独的设备或相关的设备组接收。可跳过块的概念可以用于寻址单独的设备或相关的设备组。即使网络在向不同设备发送分组时以广播模式操作,音频的处理(例如,解码、渲染等)也可被减少到寻址到该设备的块。所有其他块,即使在同一分组内被接收,也可被简单地跳过。在一些示例中,块可以基于它们的解码或呈现时间被以正确的顺序引入。如果还已经接收到具有较高优先级的相同块,则可以移除具有较低优先级的重传块。然后可以将块流馈送到解码器。

[0087] 在一些示例中,流和设备的配置在带外传送。编解码器允许建立如下连接,其中音频流以相当高的速率传输但具有低延迟。这种连接的配置可以在这种连接的持续时间内保持稳定。在那种情况下,不是使配置成为音频流的一部分,而是可以在带外传送它。对于这种带外传输,甚至可以使用不同的网络或网络协议。例如,音频流可以使用用户数据报简档(UDP)用于低延迟传输,而配置可以使用传输控制协议(TCP)来确保配置的可靠传输。

[0088] 在MPEG-4音频情境中启用编解码器

[0089] 本技术的一个具体应用是在MPEG-4音频内使用。在MPEG-4音频中,为不同的编解码器技术定义了不同的AOT(音频对象类型)。对于本文档中描述的格式,可以定义新的AOT,其允许特定于该格式的信令和数据。此外,MPEG-4中解码器的配置是在DecoderSpecificInfo()有效载荷中完成的,该有效载荷继而携带AudioSpecificConfig()有效载荷。在后者中,定义了不知晓该特定格式的某些通用信令,例如采样率和信道配置,以及用于特定AOT的特定信息。对于传统格式,其中整个流由单个设备解码,这可能是有意义的。然而,在广播模式中,其中单个流被传送到若干解码器,其中每个相应解码器仅解

码流的一部分,前置信道配置信令(作为设置设备的输出功能的手段)可能是次优的。

[0090] 图10示出了传统的MPEG-4高级结构(黑色),其中修改为灰色。为了支持广播用例,定义了codecSpecificConfig()(其中“codec”可以是通用占位符名称),其中针对特定用例重新定义了信令,使得可以将特定信道元素映射到特定设备,以及包括其他相关静态参数。具有值“0”的MPEG-4元素channelConfiguration被定义为在codecSpecificConfig中定义的信道配置。因此,此值可用于使得能够进行编解码器特定配置内的信道配置的信令的修订。

[0091] 进一步地,在MPEG-4的精神下,假定codecSpecificConfig()是可解码的,则为手边(at hand)的特定解码器规定了原始有效载荷。然而,该格式确保动态元数据是原始有效载荷的一部分,并且确保长度信息可用于所有原始有效载荷,使得解码器可以容易地跳过与特定设备无关的元素。

[0092] 在图11中,在右侧给出如在MPEG-4中定义的示例raw_data_block(原始数据块)的一部分。原始数据块包含处于给定顺序的信道元素(单信道元素(SCE)或信道对元素(CPE))。然而,在常规MPEG-4音频语法中,希望跳过可能与手边的输出设备不相关的这些信道元素的部分的解码器将不得不解析(并且在某种程度上还解码)所有信道元素以能够提取相关部分。在图11左侧所示的新原始数据块中,内容将由可跳过的块组成,使得解码器可以跳过不相关的部分,并且仅解码由元数据指示为与手边设备相关的信道元素。在一个示例中,可跳过块包括raw_data_block和相关信息。

[0093] 使用块进行重传也将是可能的。每个帧可以被分成块,如上面的可跳过块部分中所述。块可由其所属的帧编号、可用于将具有相同ID的不同帧的连续块关联到块流的块ID以及重传优先级来标识,如图15和16中所说明。例如,高重传优先级(例如,在图15和图16中以优先级0指示)发信号告知一个块在接收器处将优先于具有相同块ID和帧计数但具有较低重传优先级的另一个块(例如,在图15和图16中以优先级1指示)。应指出,如在图15和图16的示例中,优先级可以随着优先级索引的增加而降低(例如,优先级1可以是比优先级0更低的优先级),而在其他示例中,优先级可以随着优先级索引的增加而增加(例如,优先级0是比优先级1更低的优先级)。还有其他示例对于本领域技术人员来说将是显而易见的。

[0094] 语法可以支持音频元素的重传。对于重传,可以支持各种质量等级。因此,重传的块可以携带‘优先级’标志,以指示对于解码器应该优先考虑具有相同块ID的哪些块,这是因为具有相同帧计数和块ID的所接收的块是冗余的,并且因此对于解码器是互斥的,如图17和18中所示。

[0095] 块的重传可被以降低的数据率进行。这种降低的数据率可以通过降低音频信号的信噪比、降低音频信号的带宽、降低音频信号的信道计数(例如,如美国专利11,289,103中所述,其通过引用而全文并入此)或其任何组合来实现。为了使解码器挑选提供最佳可能质量的音频块,提供最高质量信号的块可具有最高解码优先级,具有次高质量信号的块可具有次高优先级,以此类推。

[0096] 块也可被以相同的质量等级重传。在这种情况下,优先级可以反映重传块的延迟。

[0097] 提高核心编码效率的工具

[0098] 另外的示例可以允许在信道元素上共享MDCT量化比例因子。在一些情况下,可在某些信道上共享MDCT比例因子以降低相关联侧比特率。此外,比例因子共享可以扩展到不

同的信道元素,例如对于7.1.4输入。被共享比例因子的使用可以由允许3种活动配置的2个信令比特来指示。一种可能的配置将是在左水平信道、左顶部信道、右水平信道和右顶部信道中共享比例因子。特定语法可以与可跳过块概念相结合,以确保比例因子仅在跳过块内共享。

[0099] 还可能具有多于两个信道的联合编码。在图4、5、6和7中图示的示例中,块Blk 1携带三个SCE,对于在该场景中考虑的智能扬声器中的驱动器中的每一个各一个SCE。在这样的情况下,可能有益的是使得能够进行不只是两个信道的联合编码(如由CPE提供的,其可以被扩展以包括立体声预测,例如,如ISO/IEC 23003-3中所描述的基于MDCT的复杂预测立体声工具,被称为MPEG-D USAC),而是多于两个信道的联合编码,例如如ETSI TS103 190中引入的SAP(立体声音频处理)工具中所描述的。

[0100] 类似地,应指出,在灵活渲染的情况下,跨设备的信号之间可能存在高相关性。对于这样的信号,构造覆盖多个设备并且允许使用上文概述的工具在多于两个信道上应用联合信道编码的跳过块可能是有益的。

[0101] 对于某些变换长度,可以使用的另一种联合信道编码工具是信道耦合,其中对于中高频传送复合信道和比例因子信息。这可以为良好质量范围中的回放提供比特率降低。例如,针对变换长度为256的帧,对应于用于低延迟编码的约256个样本的帧长度,使用信道耦合工具可能是有益的。

[0102] 本公开还将允许对带限信号进行有效编码。在发送馈送智能扬声器的不同驱动器的分离信号的场景中,这些信号中的一些可能是带限的,诸如例如具有低音扬声器、中音和高音扬声器的三路驱动器配置的情况。因此,期望对这种带限信号进行有效编码,这可以转化为被特别调谐的心理声学模型和比特分配策略以及语法的潜在修改,以便以提高的编码效率和/或降低的计算复杂度来处理这种情况(例如,使得能够将带限IMDCT用于低音扬声器馈送)。

[0103] 在一些示例中,MDCT域中高频重构和噪声添加的组合将是可能的。现代音频编解码器通常被设计为支持参数编码技术,且类似地,可设想在MDCT域中包含例如高频重构方法以在MDCT域中保留低延迟且执行噪声添加方案以管理重构高频带中的音调噪声比。这种参数编码技术对于较低的操作点,并且特别在重传作为FEC(前向纠错)方案的一部分而发生的场景中,是尤其有用的,在该场景中通常传送主信号,然后以延时方式以较低比特率重传相同的信号。

[0104] 这些示例将进一步允许编码器中的峰值复杂度降低。这在用于恒定比特率传输信道的缓冲模型限制不到位的情况下将是适用的。在具有缓冲模式的恒定比特率时,在量化和计数之后可能发生用于待编码帧的所得比特数高于被允许比特的限制。需要进行至少一个新的、更粗略的量化和比特计数步骤来满足比特要求。如果该限制更宽松,则编码器可以保持第一量化结果,并且将以稍高的瞬时比特率完成。仍然可以通过使用相同的比特储存控制机制以及根据遵守缓冲要求的编码器而更新的虚拟缓冲满度来实现与具有缓冲模型的恒定比特率编码器类似的编码器行为。这将节省额外的量化和比特计数步骤,并且与具有总比特率略微增加的缺点的恒定比特率情况相比,所得到的音频质量将相同或更好。

[0105] 回传信道概念

[0106] 对于播放和收听用例,智能扬声器30可以具有一个或多个麦克风40(例如,麦克风

或麦克风阵列),其被配置为捕获单声道或空间声场(例如,单声道、立体声、A格式、B格式、或任何其他各向同性或各向异性声道格式),并且需要编解码器能够以低延迟方式进行此类格式的高效编码,因此相同的编解码器用于向智能扬声器设备30的广播/传输,如同用于回传信道,例如图12所示。

[0107] 还应指出,虽然在智能扬声器设备上执行唤醒词检测,但是语音识别通常在云中完成,其中被记录语音的合适片段由唤醒词检测触发。

[0108] 在此上下文中,通过在例如编解码器内的中间格式/表示上创建语音分析系统来总体上节省系统复杂性可能是有趣的。对于最简单的用例,其中没有进行人与人的对话,可以定义特定于语音识别任务的合适表示,因为它不会被另一个人收听。这可以是频带能量、梅尔频率倒谱系数(MFCC)等,或者MDCT编码频谱的低比特率版本。

[0109] 在存在正在进行的人与人对话并且存在交织的唤醒词和所需语音识别的用例中,人们可能也希望能够为此提取流的相关部分。这种提取可能需要简单附加数据与主语音信号并行传送的分层编码结构。还设想了从现有解码的MDCT谱到相关表示的定义的代码转换,并且这样做简化了结构。实质上,解码器将解码并输出人类可听音频,直到用信号告知它也应该(并行地)解码语音识别表示。该表示可以是“剥离”的分层流,简单地是相同完整数据的替代解码,或者简单地是流中附加表示的解码和输出。在该用例中,信令是指示接收侧的解码器应当输出语音识别相关表示的使能片段。

[0110] 列举示例

[0111] 在下文中,并非权利要求的七组列举的示例性示例(EEE-A、EEE-B、EEE-C、EEE-D、EEE-E、EEE-F和EEE-G)描述了本文公开的示例的方面。

[0112] EEE-A1.一种音频信号解码方法,所述方法包括:

[0113] 接收包括至少一个帧的比特流,其中所述至少一个帧中的每个帧包括多个块;

[0114] 基于输出设备的设备信息,从信令数据确定用于标识在解码时要跳过的所述多个块中的一个或多个块的部分的信息;以及

[0115] 在跳过所标识的所述一个或多个块的所述部分的情况下解码所述比特流。

[0116] EEE-A2.根据EEE-A1所述的方法,其中,用于标识在解码时要跳过的所述多个块中的一个或多个块的部分的信息包括将多个输出设备中的每个输出设备与一个或多个比特流元素相关联的矩阵。

[0117] EEE-A3.根据EEE-A2所述的方法,其中,所述一个或多个比特流元素是对于相应的关联输出设备解码所述比特流所需的。

[0118] EEE-A4.根据EEE-A1至EEE-A3中任一项所述的方法,其中,所述输出设备可以包括无线设备、移动设备、平板电脑、单声道扬声器和/或多声道扬声器中的至少一者。

[0119] EEE-A5.根据EEE-A1至EEE-A4中任一项所述的方法,其中所标识的部分包括至少一个块。

[0120] EEE-A6.根据EEE-A1至EEE-A5中任一项所述的方法,其中输出设备是第一输出设备,还包括在到第二输出设备和第三输出设备的比特流的一个或多个信号之间应用联合编码技术。

[0121] EEE-A7.根据EEE-A1至EEE-A6中任一项所述的方法,其中,在系统初始化阶段期间定义每个输出设备和/或解码器的身份。

[0122] EEE-A8.根据EEE-A1至EEE-A7中任一项所述的方法,其中,由所述比特流的元数据确定所述信令数据。

[0123] EEE-A9.一种装置,被配置为执行根据EEE-A1至EEE-A8中任一项所述的方法。

[0124] EEE-A10。一种非暂时性计算机可读存储介质,包括指令序列,所述指令序列在被执行时使一个或多个设备执行根据EEE-A1至EEE-A8中任一项所述的方法。

[0125] EEE-B1.一种用于从包括多个音频信号的音频节目生成编码比特流的方法,所述方法包括:

[0126] 对于所述多个音频信号中的每一个,接收指示与相应音频信号相关联的回放设备的信息;

[0127] 对于每个回放设备,接收指示与相应回放设备相关联的延时、增益和均衡曲线中的至少一者的信息;

[0128] 从所述多个音频信号确定两个或更多个相关音频信号的组;

[0129] 对所述组中的所述两个或更多个相关音频信号应用一个或多个联合编码工具以获得经联合编码的音频信号;

[0130] 将所述经联合编码的音频信号、与所述经联合编码的音频信号相关联的回放设备的指示、以及和与所述经联合编码的音频信号相关联的相应回放设备相关联的延时和增益的指示组合成编码比特流的独立块。

[0131] EEE-B2.根据EEE-B1所述的方法,其中,与所述相应回放设备相关联的所述延时、增益和/或均衡曲线取决于所述相应回放设备相对于收听者位置的位置。

[0132] EEE-B3.根据EEE-B1或EEE-B2所述的方法,其中,与所述相应回放设备相关联的所述延时、增益和/或均衡曲线取决于所述相应回放设备相对于其他回放设备的位置的位置。

[0133] EEE-B4.根据EEE-B1至EEE-B3中任一项所述的方法,其中,所述延时、增益和/或均衡曲线是动态可变的。

[0134] EEE-B5.根据EEE-B4所述的方法,其中,所述延时、增益和/或均衡曲线响应于所述收听者的位置的变化被调整。

[0135] EEE-B6.根据EEE-B4或EEE-B5所述的方法,其中,所述延时、增益和/或均衡曲线响应于所述回放设备的位置的变化被调整。

[0136] EEE-B7.根据EEE-B4至EEE-B6中任一项所述的方法,其中,所述延时、增益和/或均衡曲线响应于所述其他回放设备中的一者或多者的位置的变化被调整。

[0137] EEE-B8.根据EEE-B1至EEE-B7中任一项所述的方法,还包括从所述多个音频信号确定不是所述两个或更多个相关音频信号的组的一部分的音频信号。

[0138] EEE-B9.根据EEE-B8所述的方法,还包括,对于不是所述两个或更多个相关音频信号的组的一部分的音频信号,应用和与该音频信号相关联的回放设备相关联的延时、增益和/或均衡曲线。

[0139] EEE-B10.根据EEE-B9所述的方法,还包括独立编码不是所述两个或更多个相关音频信号的组的一部分的音频信号,并且将经独立编码的音频信号和与经独立编码的音频信号相关联的回放设备的指示组合成编码比特流的单独的可独立解码的子集。

[0140] EEE-B11.根据EEE-B8所述的方法,还包括独立编码不是所述两个或更多个相关音频信号的组的一部分的音频信号,并且将经独立编码的音频信号、与经独立编码的音频信

号相关联的回放设备的指示、以及和与经独立编码的音频信号相关联的回放设备相关联的延时、增益和/或均衡曲线的指示组合成编码比特流的单独的可独立解码的子集。

[0141] EEE-B12.一种用于从编码比特流的帧解码与回放设备相关联的一个或多个音频信号的方法,其中所述帧包括一个或多个独立的编码数据块,所述方法包括:

[0142] 从所述编码比特流识别对应于与所述回放设备相关联的所述一个或多个音频信号的独立的编码数据块;

[0143] 从所述编码比特流中提取所识别的独立的编码数据块;

[0144] 确定所提取的独立的编码数据块包括两个或更多个联合编码音频信号;

[0145] 将一个或多个联合解码工具应用于所述两个或更多个联合编码音频信号,以获得与所述回放设备相关联的所述一个或多个音频信号;

[0146] 从所提取的独立的编码数据块确定与所述回放设备相关联的延时、增益和均衡曲线中的至少一者;

[0147] 将与所述回放设备相关联的所述延时、增益和/或均衡曲线应用于与所述回放设备相关联的所述一个或多个音频信号。

[0148] EEE-B13.根据EEE-B12所述的方法,其中,所确定的与所述回放设备相关联的延时、增益和/或均衡曲线取决于所述回放设备相对于收听者的位置的位置。

[0149] EEE-B14.根据EEE-B12或EEE-B13所述的方法,其中,所确定的与所述回放设备相关联的延时、增益和/或均衡曲线取决于所述回放设备相对于其他回放设备的位置。

[0150] EEE-B15.根据EEE-B12至EEE-B14中任一项所述的方法,其中,所确定的所述回放设备的延时、增益和/或均衡曲线是动态可变的。

[0151] EEE-B16.根据EEE-B15所述的方法,其中,当所确定的与所述回放设备相关联的延时、增益和/或均衡曲线不同于先前确定的与所述回放设备相关联的延时、增益和/或均衡曲线时,所述方法还包括在先前确定的与所述回放设备相关联的延时、增益和/或均衡曲线和所确定的与所述回放设备相关联的延时、增益和/或均衡曲线之间进行内插。

[0152] EEE-B17.根据EEE-B16所述的方法,其中所确定的延时、增益和/或均衡曲线由于收听者的位置变化而不同于先前确定的延时、增益和/或均衡曲线。

[0153] EEE-B18.根据EEE-B16或EEE-B17所述的方法,其中所确定的延时、增益和/或均衡曲线由于回放设备的位置变化而不同于先前确定的延时、增益和/或均衡曲线。

[0154] EEE-B19.根据EEE-B16至EEE-B18中任一项所述的方法,其中,所确定的延时、增益和/或均衡曲线由于一个或多个其他回放设备的位置变化而不同于先前确定的延时、增益或均衡曲线。

[0155] EEE-B20.根据EEE-B12至EEE-B19中任一项所述的方法,其中,所述编码比特流的帧包括两个或更多个独立的编码数据块,所述方法还包括:

[0156] 确定所述独立块中的一个或多个独立块包含与所述回放设备不相关联的音频信号;以及

[0157] 忽略包含与所述回放设备不相关联的音频信号的所述一个或多个独立块。

[0158] EEE-B21.根据EEE-B12至EEE-B20中任一项所述的方法,其中,应用一个或多个联合解码工具包括:识别与所述回放设备相关联的联合编码音频信号的子集,以及仅重构所述联合编码音频信号的所述子集以获得与所述回放设备相关联的所述一个或多个音频信

号。

[0159] EEE-B22.根据EEE-B12至EEE-B20中任一项所述的方法,其中,应用一个或多个联合解码工具包括:重构所述联合编码音频信号中的每一者,识别与所述回放设备相关联的重构的联合编码音频信号的子集,以及从与所述回放设备相关联的重构的联合编码音频信号的所述子集中获得与所述回放设备相关联的所述一个或多个音频信号。

[0160] EEE-B23.一种装置,被配置为执行根据EEE-B1至EEE-B22中任一项所述的方法。

[0161] EEE-B24.一种非暂时性计算机可读存储介质,包括指令序列,所述指令序列在被执行时使一个或多个设备执行根据EEE-B1至EEE-B22中任一项所述的方法。

[0162] EEE-C1.一种用于生成包括多个音频信号的音频节目的编码比特流的帧的方法,其中所述帧包括两个或更多个独立的编码数据块,所述方法包括:

[0163] 对于所述多个音频信号中的一个或多个音频信号,接收指示与所述一个或多个音频信号相关联的回放设备的信息;

[0164] 对于所指示的回放设备,接收指示一个或多个附加的关联回放设备的信息;

[0165] 接收与所指示的一个或多个附加的关联回放设备相关联的一个或多个音频信号;

[0166] 对与所述回放设备相关联的所述一个或多个音频信号进行编码;

[0167] 对与所指示的一个或多个附加的关联回放设备相关联的所述一个或多个音频信号进行编码;

[0168] 将与所述回放设备相关联的所述一个或多个编码音频信号和指示所述一个或多个附加的关联回放设备的信令信息组合成第一独立块;

[0169] 将与所述一个或多个附加的关联回放设备相关联的所述一个或多个编码音频信号组合成一个或多个附加的独立块;以及

[0170] 将所述第一独立块和所述一个或多个附加的独立块组合成所述编码比特流的所述帧。

[0171] EEE-C2.根据EEE-C1所述的方法,其中,所述多个音频信号包括与所述回放设备或所述一个或多个附加的关联回放设备不相关联的一组或多组音频信号,所述方法还包括:

[0172] 将与所述回放设备或所述一个或多个附加的关联回放设备不相关联的一组或多组音频信号中的每一者编码成相应独立块;以及

[0173] 将所述一个或多个组中的每一者的所述相应独立块组合成所述编码比特流的所述帧。

[0174] EEE-C3.根据EEE-C1或EEE-C2所述的方法,其中与所指示的一个或多个附加的关联回放设备相关联的一个或多个音频信号专门旨在用作用于对回放设备执行回声管理的回声参考。

[0175] EEE-C4.根据EEE-C3所述的方法,其中,相比于与所述回放设备相关联的所述一个或多个音频信号,旨在用作回声参考的所述一个或多个音频信号被使用更少的数据来传送。

[0176] EEE-C5.根据EEE-C3或EEE-C4所述的方法,其中使用参数编码工具对旨在用作回声参考的所述一个或多个音频信号进行编码。

[0177] EEE-C6.根据EEE-C1或EEE-C2所述的方法,其中与所述一个或多个其他回放设备相关联的所述一个或多个音频信号适合于从所述一个或多个其他回放设备回放。

[0178] EEE-C7.一种用于从编码比特流的帧中解码与回放设备相关联的一个或多个音频信号的方法,其中所述帧包括两个或多个独立的编码数据块,其中所述回放设备包括一个或多个麦克风,所述方法包括:

[0179] 从所述编码比特流中识别对应于与所述回放设备相关联的所述一个或多个音频信号的独立的编码数据块;

[0180] 从所述编码比特流中提取所识别的独立的编码数据块;

[0181] 从所识别的独立的编码数据块中提取与所述回放设备相关联的所述一个或多个音频信号;

[0182] 从所述编码比特流识别对应于与一个或多个其它回放设备相关联的一个或多个音频信号的一个或多个其它独立的编码数据块;

[0183] 从所述一个或多个其他独立的编码数据块中提取与所述一个或多个其他回放设备相关联的所述一个或多个音频信号;

[0184] 使用所述回放设备的所述一个或多个麦克风捕获一个或多个音频信号;以及

[0185] 使用与所述一个或多个其他回放设备相关联的所提取的一个或多个音频信号作为回声参考,用于响应于所捕获的一个或多个音频信号对所述回放设备执行回声管理。

[0186] EEE-C8.根据EEE-C7所述的方法,所述方法还包括:

[0187] 确定所述编码比特流包括一个或多个附加的独立的编码数据块;以及

[0188] 忽略所述一个或多个附加的独立的编码数据块。

[0189] EEE-C9.根据EEE-C8所述的方法,其中忽略所述一个或多个附加的独立的编码数据块包括跳过所述一个或多个附加的独立的编码数据块而不提取所述一个或多个附加的独立的编码数据块。

[0190] EEE-C10.根据EEE-C7至EEE-C9中任一项所述的方法,其中,与所述一个或多个其他回放设备相关联的所述一个或多个音频信号专门旨在用作用于对回放设备执行回声管理的回声参考。

[0191] EEE-C11.根据EEE-C10所述的方法,其中,相比于与所述回放设备相关联的所述一个或多个音频信号,专门旨在用作回声参考的所述一个或多个音频信号被使用更少的数据来传送。

[0192] EEE-C12.根据EEE-C10或EEE-C11所述的方法,其中,根据所述一个或多个音频信号的参数表示来重构专门旨在用作回声参考的所述一个或多个音频信号。

[0193] EEE-C13.根据EEE-C7至EEE-C9中任一项所述的方法,其中,与所述一个或多个其他回放设备相关联的所述一个或多个音频信号适合于从所述一个或多个其他回放设备回放。

[0194] EEE-C14.根据EEE-C7所述的方法,其中,所述编码信号包括指示所述一个或多个其他回放设备用作所述回放设备的回声参考的信令信息。

[0195] EEE-C15.根据EEE-C14所述的方法,其中,对于当前帧的由所述信令信息指示的所述一个或多个其他回放设备不同于对于前一帧的用作回声参考的一个或多个其他回放设备。

[0196] EEE-C16.一种装置,被配置为执行根据EEE-C1至EEE-C15中任一项所述的方法。

[0197] EEE-C17.一种非暂时性计算机可读存储介质,包括指令序列,所述指令序列在被

执行时使一个或多个设备执行根据EEE-C1至EEE-C15中任一项所述的方法。

[0198] EEE-D1.一种用于传送音频信号的方法,所述方法包括:

[0199] 生成包括比特流的部分的数据分组,其中所述比特流包括多个帧,其中所述多个帧中的每个帧包括多个块,其中所述生成包括:

[0200] 将数据分组与所述多个块中的一个或多个块进行整合,其中来自不同帧的块被组合成单个分组和/或被无序地传送;以及

[0201] 经由基于分组的网络传送所述数据分组。

[0202] EEE-D2.根据EEE-D1所述的方法,其中,所述多个块中的每个块包括标识信息。

[0203] EEE-D3.根据EEE-D2所述的方法,其中,所述标识信息包括块ID、与所述块相关联的相应帧号、和/或重传优先级中的至少一者。

[0204] EEE-D4.根据EEE-D1至EEE-D3中任一项所述的方法,其中,所述多个帧中的每一帧承载表示具有开始时间、结束时间和持续时间的音频信号的连续段的所有音频数据。

[0205] EEE-D5.一种音频信号解码方法,所述方法包括:

[0206] 接收包括比特流的部分的数据分组,其中所述比特流包括多个帧,其中所述多个帧中的每个帧包括多个块;

[0207] 确定所述多个块中的寻址到设备的块集合;以及

[0208] 对寻址到所述设备的所述块集合进行解码,并且跳过对未寻址到所述设备的所述多个块中的块的解码。

[0209] EEE-D6.一种用于传送音频流的方法,所述方法包括:

[0210] 传送所述音频流,其中所述音频流包括多个帧,其中所述多个帧中的每个帧包括多个块,其中所述传送包括在带外传送所述音频流的配置信息。

[0211] EEE-D7.根据EEE-D6所述的方法,其中,在带外传送所述音频流的配置信息包括:

[0212] 经由第一网络和/或第一网络协议传送所述音频流;以及

[0213] 经由第二网络和/或第二网络协议传送所述配置信息。

[0214] EEE-D8.根据EEE-D7所述的方法,其中,所述第一网络协议是用户数据报协议(UDP),并且所述第二网络协议是传输控制协议(TCP)。

[0215] EEE-D9.一种音频信号解码方法,所述方法包括:

[0216] 接收比特流,所述比特流包括:与静态配置方面的信令相对应的信息;静态元数据;以及

[0217] 基于所述信息和/或静态元数据将一个或多个信道元素映射到一个或多个设备。

[0218] EEE-D10.根据EEE-D9所述的方法,其中所述比特流由被配置为解码所述比特流的多个解码器接收,其中所述多个解码器中的每个解码器被配置为解码所述比特流的一部分。

[0219] EEE-D11.根据EEE-D9或EEE-D10所述的方法,其中,所述比特流还包括动态元数据。

[0220] EEE-D12.根据EEE-D9至EEE-D11中任一项所述的方法,其中,所述比特流包括多个块,其中,所述多个块中的每个块包括:使得能够在解码期间跳过所述块的一部分的信息,其中,所述部分不是设备所需要的;以及动态元数据。

[0221] EEE-D13.一种用于重传音频信号的块的方法,所述方法包括:

[0222] 传送比特流的一个或多个块,其中所述比特流包括多个块,其中所述比特流的一个或多个块中的每一个块先前已经被传送;以及其中所述一个或多个块中的每一个块包括解码优先级指示符。

[0223] EEE-D14.根据EEE-D13所述的方法,其中,所述解码优先级指示符向解码器指示用于解码所述比特流的所述一个或多个块的优先级顺序。

[0224] EEE-D15.根据EEE-D13或EEE-D14所述的方法,其中所述一个或多个块中的每个块包括相同的块ID。

[0225] EEE-D16.根据EEE-D13至EEE-D15中任一项所述的方法,其中所述比特流的所述一个或多个块的传输是通过与先前传输相比降低数据速率来传输的。

[0226] EEE-D17.根据EEE-D16所述的方法,其中,降低所述数据速率包括降低音频信号的信噪比、降低音频信号的带宽和/或降低音频信号的信道计数中的至少一者。

[0227] EEE-E1.一种用于生成包括多个音频信号的音频节目的编码比特流的帧的方法,其中所述帧包括一个或多个独立的编码数据块,所述方法包括:

[0228] 对于所述多个音频信号中的每一个,接收指示与该相应音频信号相关联的回放设备的信息;

[0229] 对与相应回放设备相关联的一个或多个音频信号进行编码,以获得一个或多个编码音频信号;

[0230] 将与所述相应回放设备相关联的所述一个或多个编码音频信号组合成所述帧的第一独立块;

[0231] 将所述多个音频信号中的一个或多个其他音频信号编码成一个或多个附加独立块;以及

[0232] 将所述第一独立块和所述一个或多个附加独立块组合成所述编码比特流的所述帧。

[0233] EEE-E2.根据EEE-E1所述的方法,其中两个或更多个音频信号与所述回放设备相关联,并且所述两个或更多个音频信号中的每一个是旨在由所述回放设备的相应驱动器回放的带限信号,并且其中对于每一个带限信号使用不同的编码技术。

[0234] EEE-E3.根据EEE-E2所述的方法,其中对于每个带限信号使用不同的心理声学模型和/或不同的比特分配技术。

[0235] EEE-E4.根据EEE-E1至EEE-E3中任一项所述的方法,其中,所述编码信号的瞬时帧速率是可变的,并且由缓冲满度模型约束。

[0236] EEE-E5.根据EEE-E1所述的方法,其中,对与相应回放设备相关联的一个或多个音频信号进行编码包括将与相应回放设备相关联的一个或多个音频信号和与一个或多个附加回放设备相关联的一个或多个附加音频信号联合编码到帧的第一独立块中。

[0237] EEE-E6.根据EEE-E5所述的方法,其中,对所述一个或多个音频信号和一个或多个附加音频信号进行联合编码包括跨两个或更多个音频信号共享一个或多个比例因子。

[0238] EEE-E7.根据EEE-E6所述的方法,其中,所述两个或更多个音频信号在空间上相关。

[0239] EEE-E8.根据EEE-E7所述的方法,其中,所述两个或更多个空间相关音频信号包括左水平信道、左顶部信道、右水平信道或右顶部信道。

[0240] EEE-E9.根据EEE-E5所述的方法,其中对所述一个或多个音频信号和一个或多个附加音频信号进行联合编码包括应用耦合工具,包括:

[0241] 将两个或更多个音频信号组合成指定频率以上的复合信号;以及

[0242] 对于所述两个或更多音频信号中的每一者,确定使所述复合信号的能量与每个相应信号的能量相关的比例因子。

[0243] EEE-E10.根据EEE-E5所述的方法,其中,对所述一个或多个音频信号和一个或多个附加音频信号进行联合编码包括将联合编码工具应用于两个以上的信号。

[0244] EEE-E11.一种用于从编码比特流的帧解码与回放设备相关联的一个或多个音频信号的方法,其中所述帧包括一个或多个独立的编码数据块,所述方法包括:

[0245] 从所述编码比特流识别对应于与所述回放设备相关联的所述一个或多个音频信号的独立的编码数据块;

[0246] 从所述编码比特流中提取所识别的独立的编码数据块;

[0247] 从所述独立的编码数据块解码与所述回放设备相关联的所述一个或多个音频信号,以获得一个或多个解码音频信号;

[0248] 从所述编码比特流识别对应于一个或多个附加音频信号的一个或多个附加的独立的编码数据块;以及

[0249] 解码或跳过所述一个或多个附加的独立的编码数据块。

[0250] EEE-E12.根据EEE-E11所述的方法,其中,两个或更多个音频信号与所述回放设备相关联,并且所述两个或更多个音频信号中的每一个是旨在由所述回放设备的相应驱动器回放的带限信号,并且其中使用不同的解码技术来解码所述两个或更多个音频信号。

[0251] EEE-E13.根据EEE-E12所述的方法,其中,对于每个带限信号使用不同的心理声学模型和/或不同的比特分配技术来编码。

[0252] EEE-E14.根据EEE-E11或EEE-E13所述的方法,其中,所述编码比特流的瞬时帧速率是可变的,并且由缓冲满度模型约束。

[0253] EEE-E15.根据EEE-E11所述的方法,其中解码与回放设备相关联的一个或多个音频信号包括从独立的编码数据块联合解码与相应回放设备相关联的一个或多个音频信号和与一个或多个附加回放设备相关联的一个或多个附加音频信号。

[0254] EEE-E16.根据EEE-E15所述的方法,其中,联合解码所述一个或多个音频信号和一个或多个附加音频信号包括提取跨两个或更多个音频信号共享的一个或多个比例因子。

[0255] EEE-E17.根据EEE-E16所述的方法,其中,所述两个或更多个音频信号在空间上相关。

[0256] EEE-E18.根据EEE-E17所述的方法,其中,所述两个或更多个空间相关音频信号包括左水平信道、左顶部信道、右水平信道或右顶部信道。

[0257] EEE-E19.根据EEE-E15所述的方法,其中联合解码所述一个或多个音频信号和一个或多个附加音频信号包括应用解耦工具。

[0258] EEE-E20.根据EEE-E19所述的方法,其中,所述解耦工具包括:

[0259] 提取指定频率以下的独立解码信号;

[0260] 提取所述指定频率以上的复合信号;

[0261] 由所述复合信号以及将所述复合信号的能量与各个信号的能量相关的比例因子

确定所述指定频率以上的各个解耦信号;以及

[0262] 将每个独立解码信号与相应的解耦信号组合以获得联合解码的信号。

[0263] EEE-E21.根据EEE-E15所述的方法,其中联合解码所述一个或多个音频信号和一个或多个附加音频信号包括应用联合解码工具来提取两个以上的音频信号。

[0264] EEE-E22.根据EEE-E11所述的方法,其中,对与所述回放设备相关联的所述一个或多个音频信号进行解码包括在与所述音频信号被编码相同的域中对所述音频信号应用带宽扩展。

[0265] EEE-E23.根据EEE-E22所述的方法,其中所述域是修正离散余弦变换(MDCT)域。

[0266] EEE-E24.根据EEE-E22或EEE-E23所述的方法,其中带宽扩展包括自适应噪声添加。

[0267] EEE-E25.一种装置,被配置为执行根据EEE-E1至EEE-E24中任一项所述的方法。

[0268] EEE-E26.一种非暂时性计算机可读存储介质,包括指令序列,所述指令序列在被执行时使一个或多个设备执行根据EEE-E1至EEE-E24中任一项所述的方法。

[0269] EEE-F1.一种由具有一个或多个麦克风的设备执行的用于生成编码比特流的方法,所述方法包括:

[0270] 由所述一个或多个麦克风捕获一个或多个音频信号;

[0271] 分析所捕获的音频信号以确定唤醒词的存在;

[0272] 在检测到唤醒词存在时:

[0273] 设置标志以指示将对所捕获的音频信号执行语音识别任务;

[0274] 对所捕获的音频信号进行编码;

[0275] 将所编码的音频信号和所述标志整合到所述编码比特流中。

[0276] EEE-F2.根据EEE-F1所述的方法,其中,所述一个或多个麦克风被配置为捕获单声道或空间声场。

[0277] EEE-F3.根据EEE-F2所述的方法,其中,所述空间声场是A格式或B格式。

[0278] EEE-F4.根据EEE-F1至EEE-F3中任一项所述的方法,其中,所捕获的音频信号旨在仅用于执行所述语音识别任务。

[0279] EEE-F5.根据EEE-F4所述的方法,其中,所捕获的音频信号被编码为使得当所捕获的音频信号被解码时,所解码的音频信号的质量足以供执行语音识别任务,但不足以供人收听。

[0280] EEE-F6.根据EEE-F4或EEE-F5所述的方法,其中在对所捕获的音频信号进行编码之前,将所捕获的音频信号转换为包括频带能量、梅尔频率倒谱系数或修正离散余弦变换(MDCT)频谱系数中的一者或多者的表示。

[0281] EEE-F7.根据EEE-F1至EEE-F3中任一项所述的方法,其中,所捕获的音频信号旨在用于供人收听和用于供执行所述语音识别任务。

[0282] EEE-F8.根据EEE-F7所述的方法,其中,所捕获的音频信号被编码为使得当所捕获的音频信号被解码时,所解码的音频信号的质量足以供人收听。

[0283] EEE-F9.根据EEE-F7所述的方法,其中,编码所捕获的音频信号包括生成所捕获的音频信号的第一编码表示和所捕获的音频信号的第二编码表示,其中所述第一编码表示被生成为使得当从所述第一编码表示解码所捕获的音频信号时,所解码的音频信号的质量足

以供人收听,并且其中所述第二编码表示被生成为使得当从所述第二编码表示解码所捕获的音频信号时,所解码的音频信号的质量足以供执行语音识别任务,但不足以供人收听。

[0284] EEE-F10.根据EEE-F9所述的方法,其中生成所捕获的音频信号的第二编码表示包括:在对所捕获的音频信号进行编码之前,将所捕获的音频信号转换为参数表示、粗略波形表示或包括频带能量、梅尔频率倒谱系数或修正离散余弦变换(MDCT)谱系数中的一者或多者的表示中的一者或多者。

[0285] EEE-F11.根据EEE-F9或EEE-F10所述的方法,其中将编码音频信号整合成比特流包括将第一编码表示插入编码比特流的第一独立块,以及将第二编码表示插入编码比特流的第二独立块。

[0286] EEE-F12.根据EEE-F9或EEE-F10所述的方法,其中第一编码表示被包括在编码比特流的第一层中,第二编码表示被包括在编码比特流的第二层中,并且第一层和第二层被包括在编码比特流的单个块中。

[0287] EEE-F13.根据EEE-F1至EEE-F12中任一项所述的方法,还包括:当未检测到存在所述唤醒词时:

[0288] 设置所述标志以指示不对所捕获的音频信号执行语音识别任务;

[0289] 对所捕获的音频信号进行编码;

[0290] 将所编码的音频信号和所述标志整合到所述编码比特流中。

[0291] EEE-F14.一种音频信号解码方法,包括:

[0292] 接收编码比特流,所述编码比特流包括编码音频信号和指示是否要执行语音识别任务的标志;

[0293] 解码所述编码音频信号以获得解码音频信号;以及

[0294] 当所述标志指示要执行所述语音识别任务时,对解码音频信号执行所述语音识别任务。

[0295] EEE-F15.根据EEE-F14所述的方法,其中,所述解码音频信号旨在仅用于执行所述语音识别任务。

[0296] EEE-F16.根据EEE-F15所述的方法,其中,所述解码音频信号的质量足以供执行所述语音识别任务,但不足以供人收听。

[0297] EEE-F17.根据EEE-F15或EEE-F16所述的方法,其中,在对所捕获的音频信号进行编码之前,所述解码音频信号的表示包括频带能量、梅尔频率倒谱系数或修正离散余弦变换(MDCT)频谱系数中的一者或多者。

[0298] EEE-F18.根据EEE-F14所述的方法,其中,所捕获的音频信号被编码为使得当对所捕获的音频信号进行解码时,所解码的音频信号的质量足以供人收听。

[0299] EEE-F19.根据EEE-F18所述的方法,其中,所述编码音频信号包括一个或多个音频信号的第一编码表示和所述一个或多个音频信号的第二编码表示。

[0300] EEE-F20.根据EEE-F18所述的方法,其中从第一表示解码的音频信号的质量足以供人收听,并且其中从第二表示解码的音频信号的质量足以供执行语音识别任务,但是不足以供人收听。

[0301] EEE-F21.根据EEE-F19或EEE-F20所述的方法,其中第一表示在编码比特流的第一独立块中,并且第二表示在编码比特流的第二独立块中。

[0302] EEE-F22.根据EEE-F19或EEE-F20所述的方法,其中第一表示在编码比特流的第一层中,第二编码表示被包括在编码比特流的第二层中,并且第一层和第二层被包括在编码比特流的单个块中。

[0303] EEE-F23.根据EEE-F18至EEE-F22中任一项所述的方法,其中,解码所述编码音频信号包括仅解码所述第二表示,而忽略所述第一表示。

[0304] EEE-F24.根据EEE-F18至EEE-F23中任一项所述的方法,其中,从所述第二编码表示解码的音频信号为参数表示、波形表示、或包括频带能量、梅尔频率倒谱系数或修正离散余弦变换(MDCT)频谱系数中的一者或多者的表示。

[0305] EEE-F25.一种装置,被配置为执行根据EEE-F1至EEE-F24中任一项所述的方法。

[0306] EEE-F26.一种非暂时性计算机可读存储介质,包括指令序列,所述指令序列在被执行时使一个或多个设备执行根据EEE-F1至EEE-F24中任一项所述的方法。

[0307] EEE-G1.一种用于低延迟传输到一个或多个回放设备的沉浸式音频节目的音频信号的编码方法,所述方法包括:

[0308] 接收所述沉浸式音频节目的多个时域音频信号;

[0309] 选择帧大小;

[0310] 响应于所述帧大小提取时域音频信号的帧,其中时域音频信号的所述帧与时域音频信号的前一帧重叠;

[0311] 将音频信号分割成重叠帧;

[0312] 将所述时域音频信号的帧变换为频域信号;

[0313] 对所述频域信号进行编码;

[0314] 使用感知激励量化工具对编码的频域信号进行量化;

[0315] 将量化和编码的频域信号整合成帧内的一个或多个独立块;以及

[0316] 将所述一个或多个独立块整合成编码帧。

[0317] EEE-G2.根据EEE-G1所述的方法,其中,所述多个音频信号包括具有定义的声道配置的基于声道的信号。

[0318] EEE-G3.根据EEE-G2所述的方法,其中,所述声道配置是单声道、立体声、5.1、5.1.2、5.1.4、7.1.2、7.1.4、9.1.6或22.2中的一个。

[0319] EEE-G4.根据EEE-G1至EEE-G3中任一项所述的方法,其中,所述多个音频信号包括一个或多个基于对象的信号。

[0320] EEE-G5.根据EEE-G1至EEE-G4中任一项所述的方法,其中,所述多个音频信号包括所述沉浸式音频节目的基于场景的表示。

[0321] EEE-G6.根据EEE-G1至EEE-G5中任一项所述的方法,其中,所选择的帧大小是128、256、512、1024、120、240、480或960个样本中的一个。

[0322] EEE-G7.根据EEE-G1至EEE-G6中任一项所述的方法,其中,时域音频信号的所述帧与时域音频信号的所述前一帧之间的重叠为50%或更低。

[0323] EEE-G8.根据EEE-G1至EEE-G7中任一项所述的方法,其中,所述变换是修正离散余弦变换(MDCT)。

[0324] EEE-G9.根据EEE-G1至EEE-G8中任一项所述的方法,其中,所述多个音频信号中的两个或更多个被联合编码。

[0325] EEE-G10.根据EEE-G1至EEE-G9中任一项所述的方法,其中,每个独立块包含用于一个或多个回放设备的编码信号。

[0326] EEE-G11.根据EEE-G1至EEE-G10中任一项所述的方法,其中,至少一个独立块包含用于两个或更多个回放设备的编码信号,并且所述编码信号包括联合编码音频信号。

[0327] EEE-G12.根据EEE-G1至EEE-G11中任一项所述的方法,其中,至少一个独立块包含旨在从回放设备的不同驱动器回放的覆盖不同带宽的多个编码信号。

[0328] EEE-G13.根据EEE-G1至EEE-G12中任一项所述的方法,其中,至少一个独立块包含用于由回放设备执行的回声管理的编码回声参考信号。

[0329] EEE-G14.根据EEE-G1至EEE-G13中任一项所述的方法,其中,对量化的频域信号进行编码包括应用以下工具中的一者或多者:时间噪声整形(TNS)、联合信道编码、跨信号共享比例因子、确定用于高频重构的控制参数、以及确定用于噪声替代的控制参数。

[0330] EEE-G15.根据EEE-G1至EEE-G14中任一项所述的方法,其中,一个或多个独立块包括用于控制回放设备的延时、增益和均衡中的一者或多者的参数。

[0331] EEE-G16.一种用于从编码信号解码沉浸式音频节目的音频信号的低延迟方法,所述方法包括:

[0332] 接收包括一个或多个独立块的编码帧;

[0333] 从所述一个或多个独立块中提取量化和编码的频域信号;

[0334] 对量化和编码的频域信号进行去量化;

[0335] 解码去量化的频域信号;

[0336] 对解码的频域信号进行逆变换,得到时域信号;以及

[0337] 将所述时域信号与来自前一帧的时域信号重叠并相加,以提供所述沉浸式音频节目的多个音频信号。

[0338] EEE-G17.根据EEE-G16所述的方法,其中,所述多个音频信号包括具有定义的声道配置的基于声道的信号。

[0339] EEE-G18.根据EEE-G17所述的方法,其中,所述声道配置是单声道、立体声、5.1、5.1.2、5.1.4、7.1.2、7.1.4、9.1.6或22.2中的一个。

[0340] EEE-G19.根据EEE-G16至EEE-G18中任一项所述的方法,其中,所述多个音频信号包括一个或多个基于对象的信号。

[0341] EEE-G20.根据EEE-G16至EEE-G19中任一项所述的方法,其中,所述多个音频信号包括所述沉浸式音频节目的基于场景的表示。

[0342] EEE-G21.根据EEE-G16至EEE-G20中任一项所述的方法,其中,时域样本的帧包括128、256、512、1024、120、240、480或960个样本中的一者。

[0343] EEE-G22.根据EEE-G16至EEE-G21中任一项所述的方法,其中,与前一帧的重叠为50%或更低。

[0344] EEE-G23.根据EEE-G16至EEE-G22中任一项所述的方法,其中,所述逆变换是逆修正离散余弦变换(IMDCT)。

[0345] EEE-G24.根据EEE-G16至EEE-G23中任一项所述的方法,其中,每个独立块包含用于一个或多个回放设备的量化和编码的频域信号。

[0346] EEE-G25.根据EEE-G16至EEE-G24中任一项所述的方法,其中,至少一个独立块包

含用于两个或更多个回放设备的量化和编码的频域信号,并且所述量化和编码的频域信号是联合编码的音频信号。

[0347] EEE-G26.根据EEE-G16至EEE-G25中任一项所述的方法,其中,至少一个独立块包含旨在从回放设备的不同驱动器回放的覆盖不同带宽的多个量化和编码的频域信号。

[0348] EEE-G27.根据EEE-G16至EEE-G26中任一项所述的方法,其中,至少一个独立块包含用于由回放设备执行的回声管理的编码的回声参考信号。

[0349] EEE-G28.根据EEE-G16至EEE-G27中任一项所述的方法,其中,解码所述去量化的频域信号包括应用以下解码工具中的一者或多者:时间噪声整形(TNS)、联合信道解码、跨信号共享比例因子、高频重构和噪声替代。

[0350] EEE-G29.根据EEE-G16至EEE-G28中任一项所述的方法,其中,一个或多个独立块包括用于控制回放设备的延时、增益和均衡中的一者或多者的参数。

[0351] EEE-G30.根据EEE-G16至EEE-G29中任一项所述的方法,其中,所述方法由回放设备执行,并且其中,从一个或多个独立块中提取量化和编码的信号包括仅选择包含用于由所述回放设备回放的量化和编码的频域信号的那些块,而忽略包含用于由其他回放设备回放的量化和编码的频域信号的独立块。

[0352] EEE-G31.一种装置,被配置为执行根据EEE-G1至EEE-G30中任一项所述的方法。

[0353] EEE-G32.一种非暂时性计算机可读存储介质,包括指令序列,所述指令序列在被执行时使一个或多个设备执行根据EEE-G1至EEE-G30中任一项所述的方法。

[0354] 在下面的权利要求和本文的描述中,术语“包括”、“由……组成”或“其包括”中的任一者是开放性术语,其意味着至少包括随后的要素/特征,但不排除其他要素/特征。因此,当在权利要求书中使用时,术语“包括”不应被解释为限于其后列出的装置或元件或步骤。例如,表述“设备包括A和B”的范围不应限于设备仅由元件A和B组成。如本文所使用的术语“包含”或“其包含”中的任一者也是开放术语,其也意味着至少包含该术语之后的要素/特征,但不排除其他要素/特征。因此,包括与包含同义并且意味着包含。

[0355] 应当认识到,在本发明的示例的以上描述中,出于精简本公开并帮助理解各个发明方面中的一个或多个方面的目的,各种特征有时被一起分组在单个示例、图或其描述中。然而,本公开的方法不应被解释为反映需要比每个权利要求中明确记载的特征更多的特征的意图。相反,如所附权利要求所反映的,创造性方面在于少于单个前述公开的示例的所有特征。因此,具体实施方式之后的权利要求书特此明确地并入本具体实施方式中,其中每个权利要求独立地作为本发明的单独示例。

[0356] 此外,虽然本文描述的一些示例包括其他示例中包括的一些特征,但不包括其他特征,但如本领域技术人员将理解的,不同示例的特征的组合意在被涵盖,并且形成不同示例。例如,在以下权利要求中,所要求保护的示例中的任一个可被以任何组合使用。

[0357] 此外,示例中的一些在本文中被描述为可以由计算机系统的处理器或由执行功能的其他手段实现的方法或方法的元素的组合。因此,具有用于执行这种方法或方法的要素的必要指令的处理器形成用于执行该方法或方法的要素的手段。此外,装置的本文描述的元件是用于实行由该元件执行的功能的手段的示例。

[0358] 此外,本文描述的一些示例公开了连接的解决方案,其应当被解释为可能在诸如有线和/或无线系统的分发和/或传输系统中实现。例如通过使用任何电、光和/或移动系

统,诸如3G、4G以及5G。

[0359] 因此,虽然已经描述了本发明的具体示例,但是本领域那些技术人员将认识到,可以进行其他和进一步的修改,并且旨在要求保护所有这样的改变和修改。例如,上面给出的任何式子仅仅是可以使用的过程的代表。可以从框图中添加或删除功能,并且可以在功能块之间互换操作。可以对于所描述的方法添加或删除步骤。

[0360] 上文公开的系统、设备和方法可以被实现为软件、固件、硬件或其组合。例如,本申请的方面可以至少部分地体现在设备、包括多于一个设备的系统、方法、计算机程序产品等中。

[0361] 在硬件实现方式中,上述描述中所提及的功能单元之间的任务划分并不一定与物理单元的划分相对应;相反,一个物理组件可以具有多个功能,并且一个任务可以由若干物理组件协作执行。

[0362] 某些组件或所有组件可以被实现为由数字信号处理器或微处理器执行的软件或者被实现为硬件或专用集成电路。这样的软件可以分布在计算机可读介质上,所述计算机可读介质可以包括计算机存储介质(或非暂时性介质)和通信介质(或暂时性介质)。

[0363] 如本领域技术人员所公知的,术语“计算机存储介质”包括用于存储诸如计算机可读指令、数据结构、程序模块或其他数据的信息的、以任何方法或技术实现的易失性和非易失性、可拆装和不可拆装介质两者。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其它存储器技术、CD-ROM、数字通用盘(DVD)或其它光盘存储、盒式磁带、磁带、磁盘存储或其它磁存储设备、或可用于存储所需信息并可由计算机访问的任何其它介质。

[0364] 进一步地,本领域技术人员公知的是,通信介质通常在诸如载波或其他传输机制的调制数据信号中体现计算机可读指令、数据结构、程序模块或其他数据,并且包括任何信息递送介质。

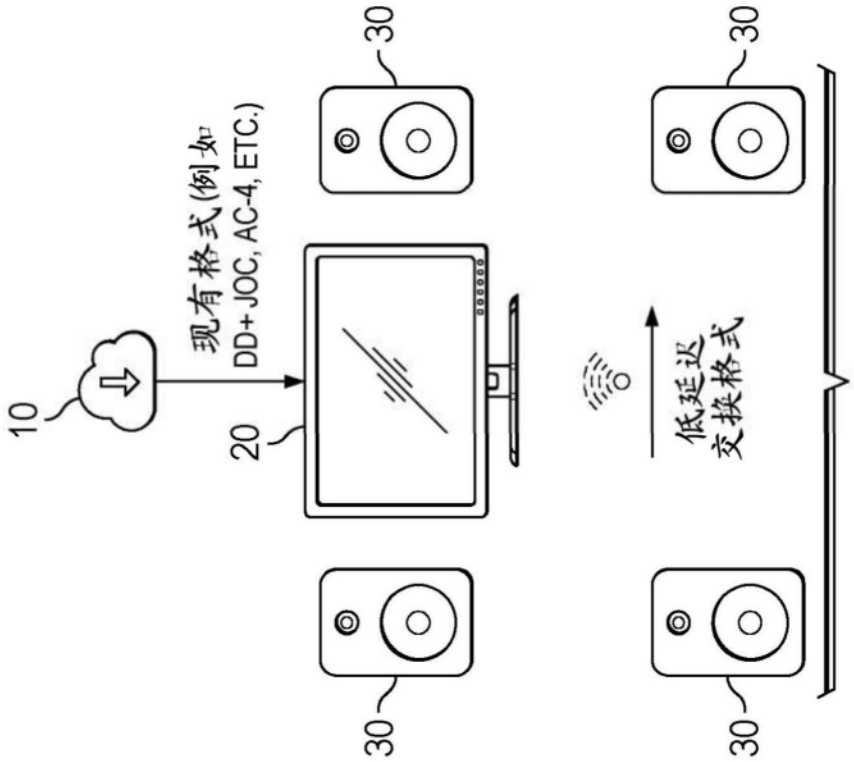


图1

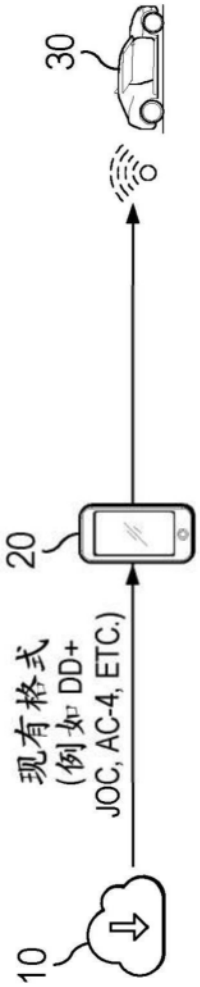


图2

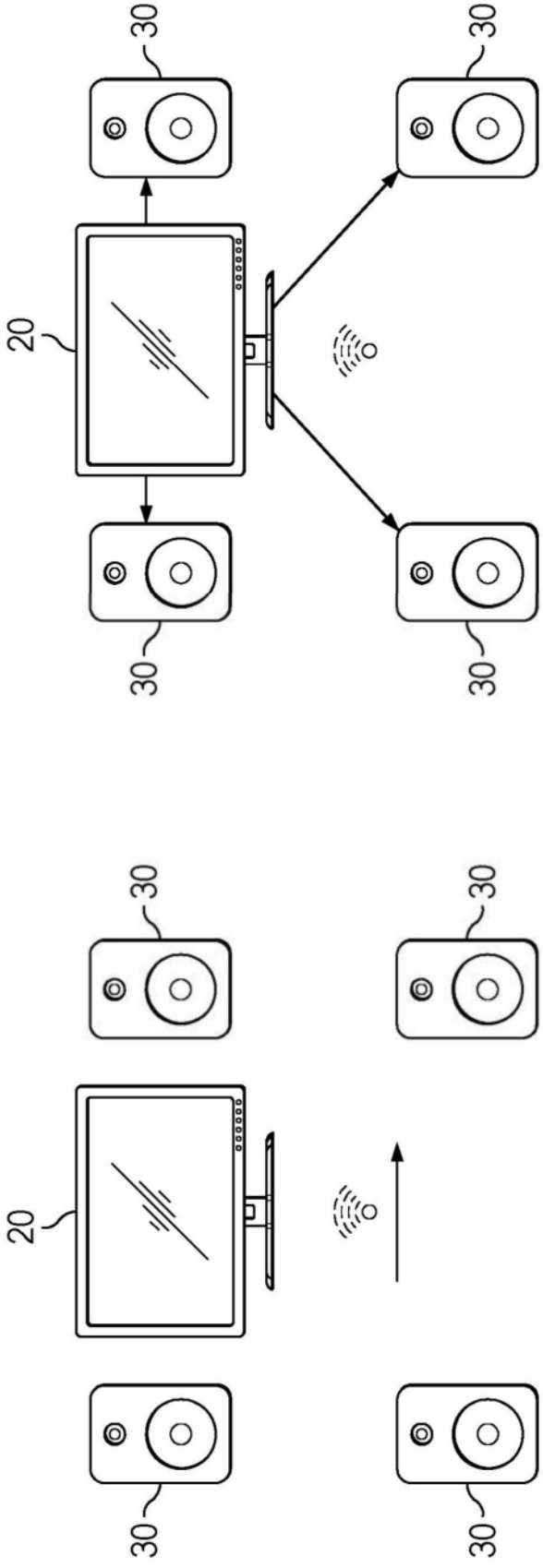


图3

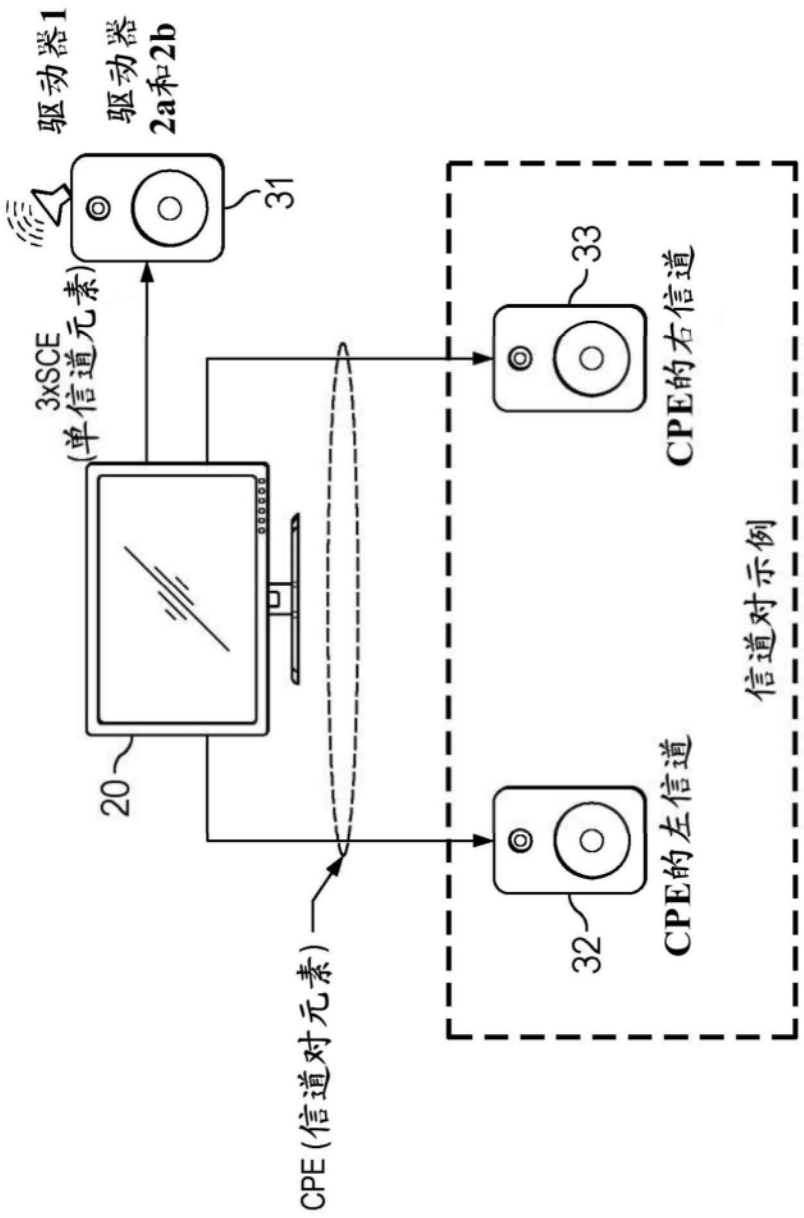

驱动器1
驱动器2a和2b
31
3xSCE
(单信道元素)
20
33
CPE的右信道
信道对示例
CPE的左信道
32
CPE (信道对元素)

图4

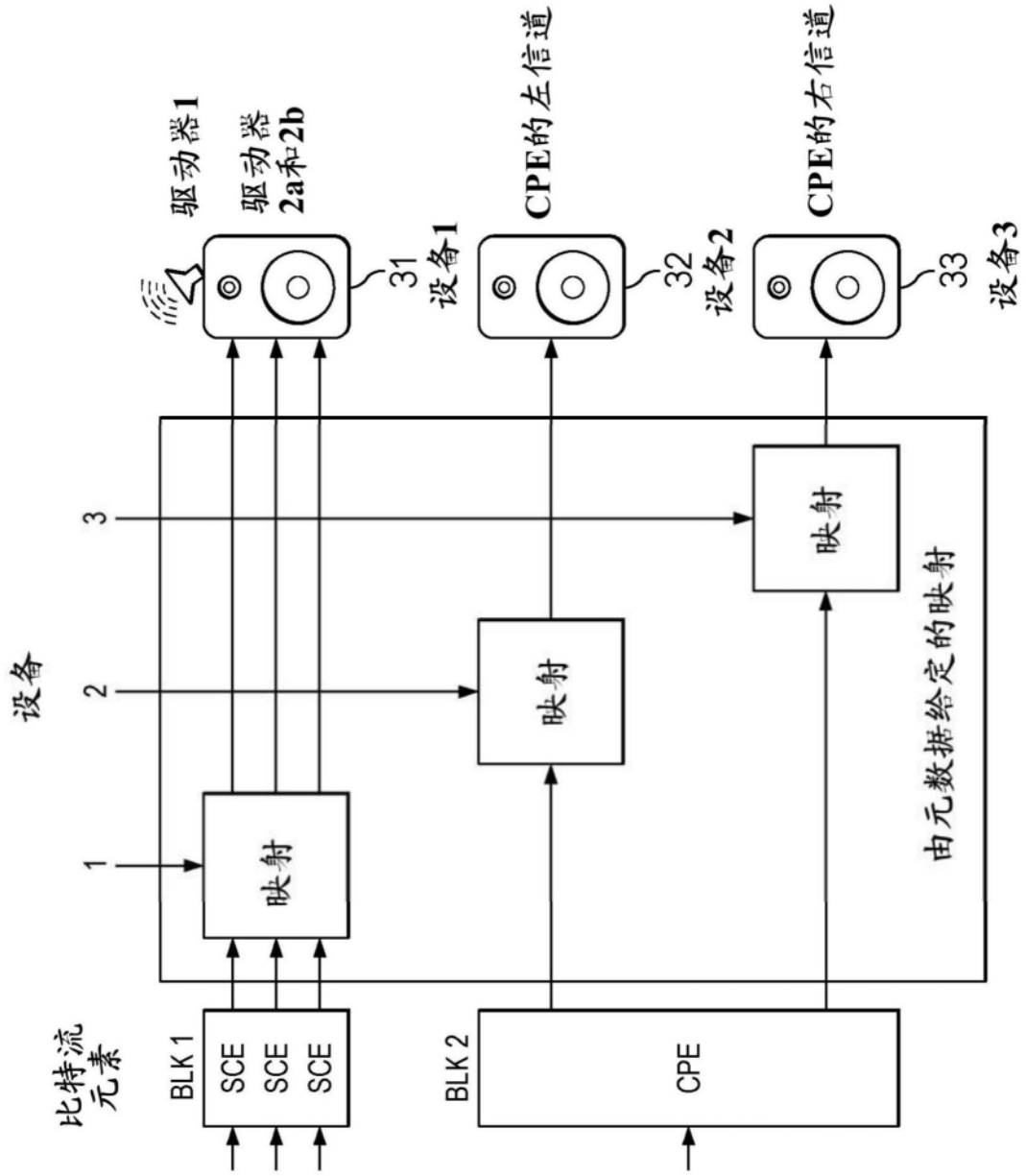

比特流元素
BLK 1
SCE
SCE
SCE
BLK 2
CPE
设备
1
2
3
映射
映射
映射
由元数据给定的映射
驱动器1
驱动器2a和2b
CPE的左信道
CPE的右信道
31
设备1
32
设备2
33
设备3

图5

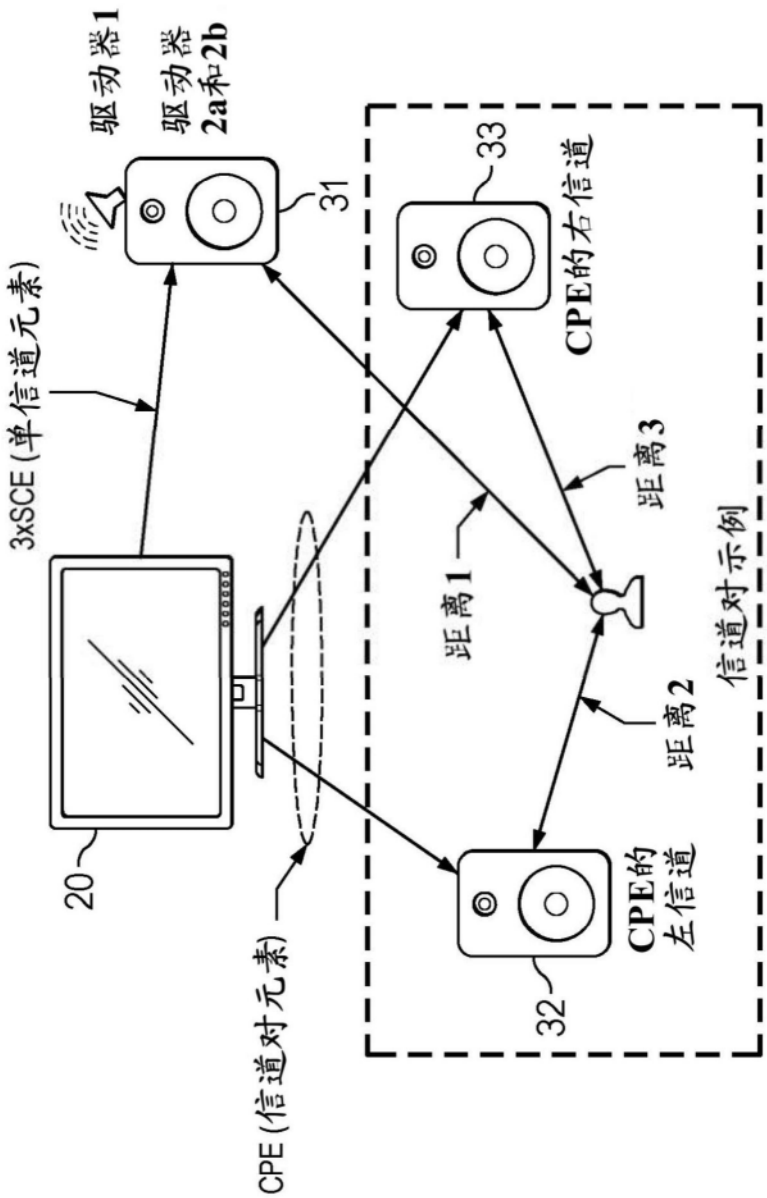

驱动器1
驱动器2a和2b
31
33
CPE的右信道
3xSCE (单信道元素)
距离3
距离1
信道对示例
距离2
20
CPE的左信道
32
CPE (信道对元素)

图6

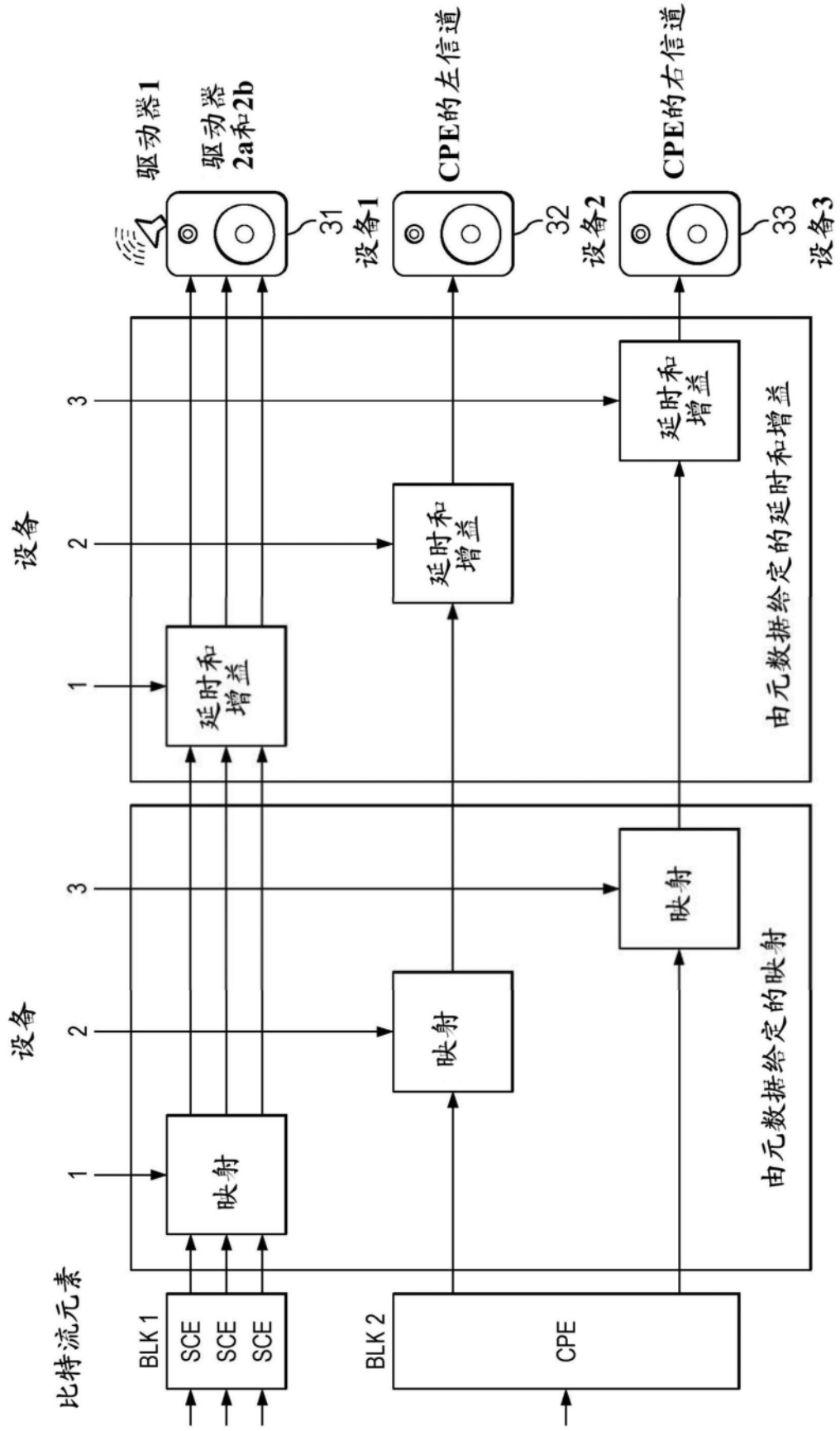

比特流元素
BLK 1
SCE
SCE
SCE
BLK 2
CPE
设备
1
2
3
映射
映射
映射
由元数据给定的映射
延时和增益
延时和增益
延时和增益
由元数据给定的延时和增益
驱动器1
驱动器2a和2b
31
设备1
CPE的左信道
32
设备2
CPE的右信道
33
设备3

图7

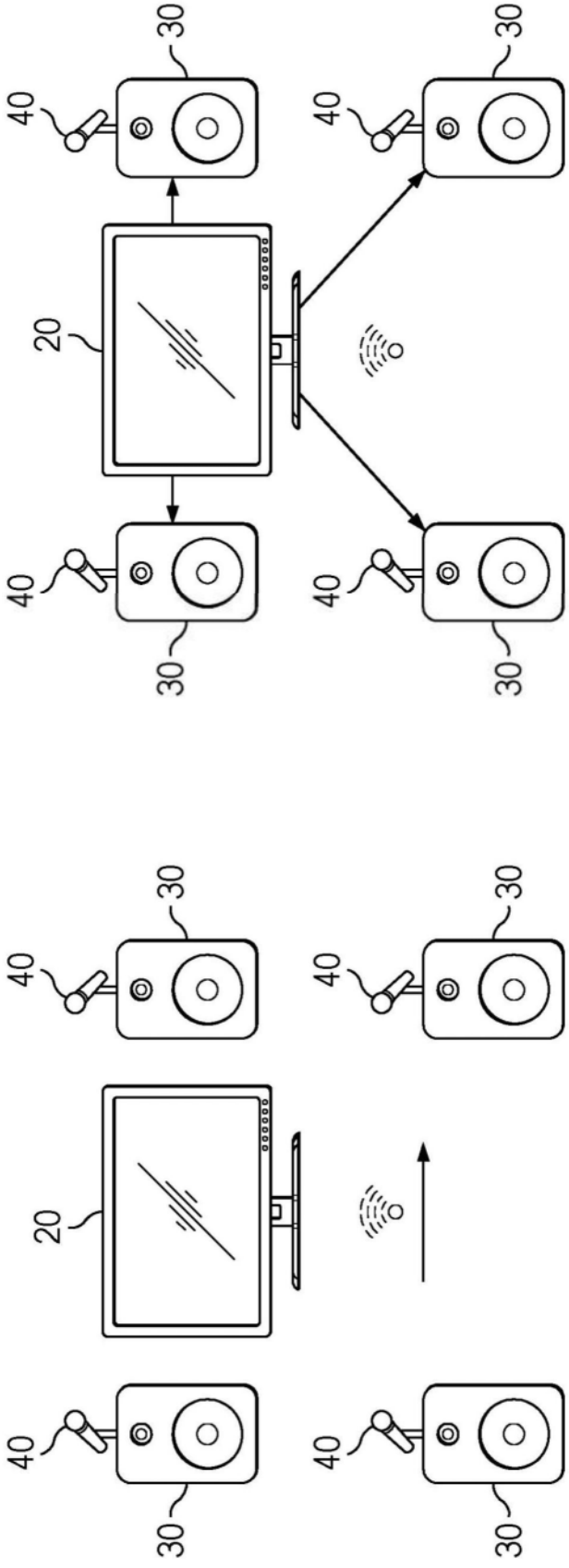


图8

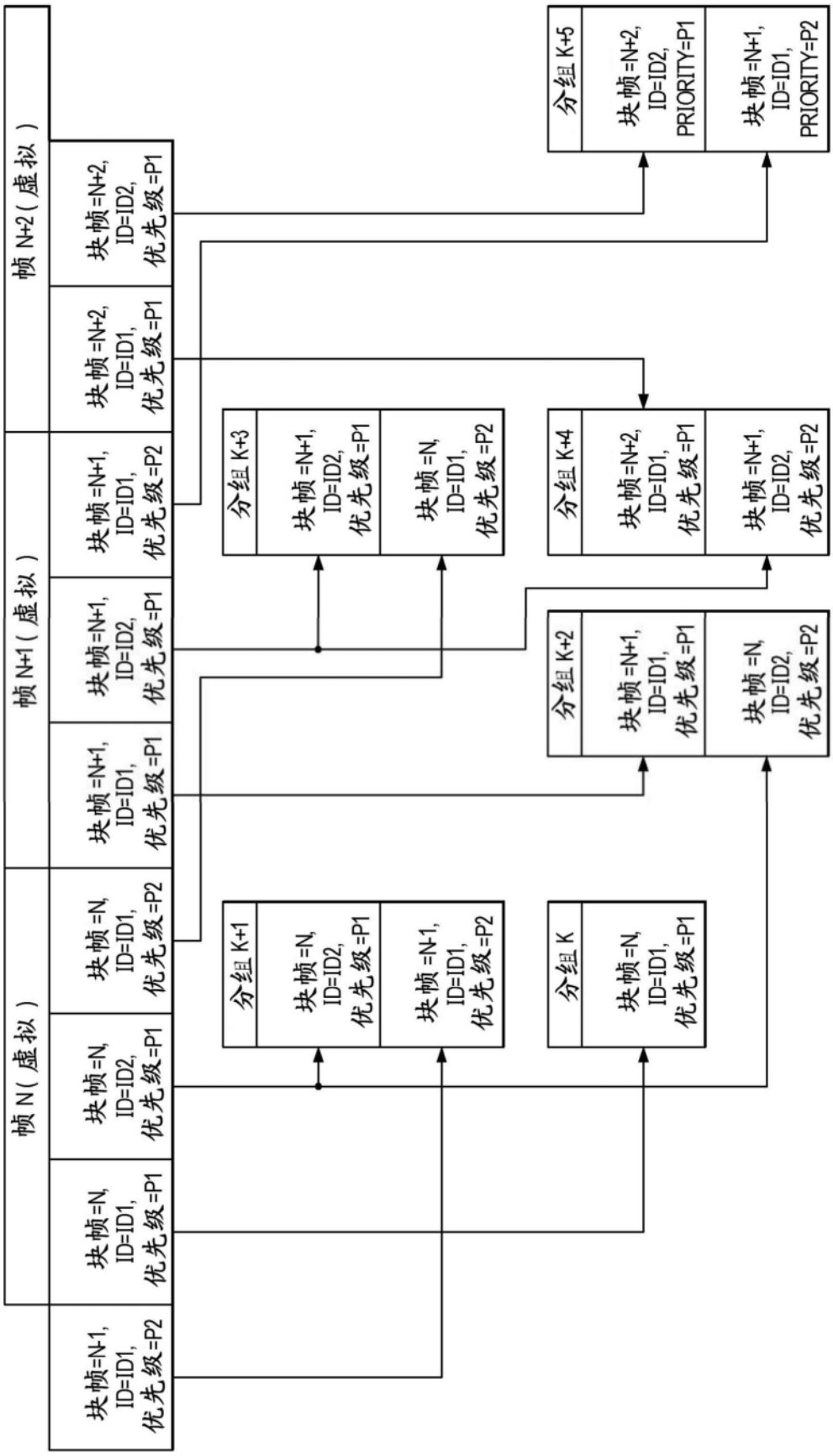

帧 N(虚拟)
块帧=N-1, ID=ID1, 优先级=P2
块帧 =N, ID=ID1, 优先级=P1
块帧 =N, ID=ID2, 优先级=P1
块帧 =N, ID=ID1, 优先级=P2
帧 N+1(虚拟)
块帧=N+1, ID=ID1, 优先级=P1
块帧=N+1, ID=ID2, 优先级=P1
块帧=N+1, ID=ID1, 优先级=P2
帧 N+2(虚拟)
块帧=N+2, ID=ID1, 优先级=P1
块帧=N+2, ID=ID2, 优先级=P1
分组 K
块帧 =N, ID=ID1, 优先级=P1
分组 K+1
块帧 =N, ID=ID2, 优先级=P1
块帧 =N-1, ID=ID1, 优先级=P2
分组 K+2
块帧 =N+1, ID=ID1, 优先级=P1
块帧 =N, ID=ID2, 优先级=P2
分组 K+3
块帧 =N+1, ID=ID2, 优先级=P1
块帧 =N, ID=ID1, 优先级=P2
分组 K+4
块帧 =N+2, ID=ID1, 优先级=P1
块帧 =N+1, ID=ID2, 优先级=P2
分组 K+5
块帧 =N+2, ID=ID2, PRIORITY=P1
块帧 =N+1, ID=ID1, PRIORITY=P2

图9

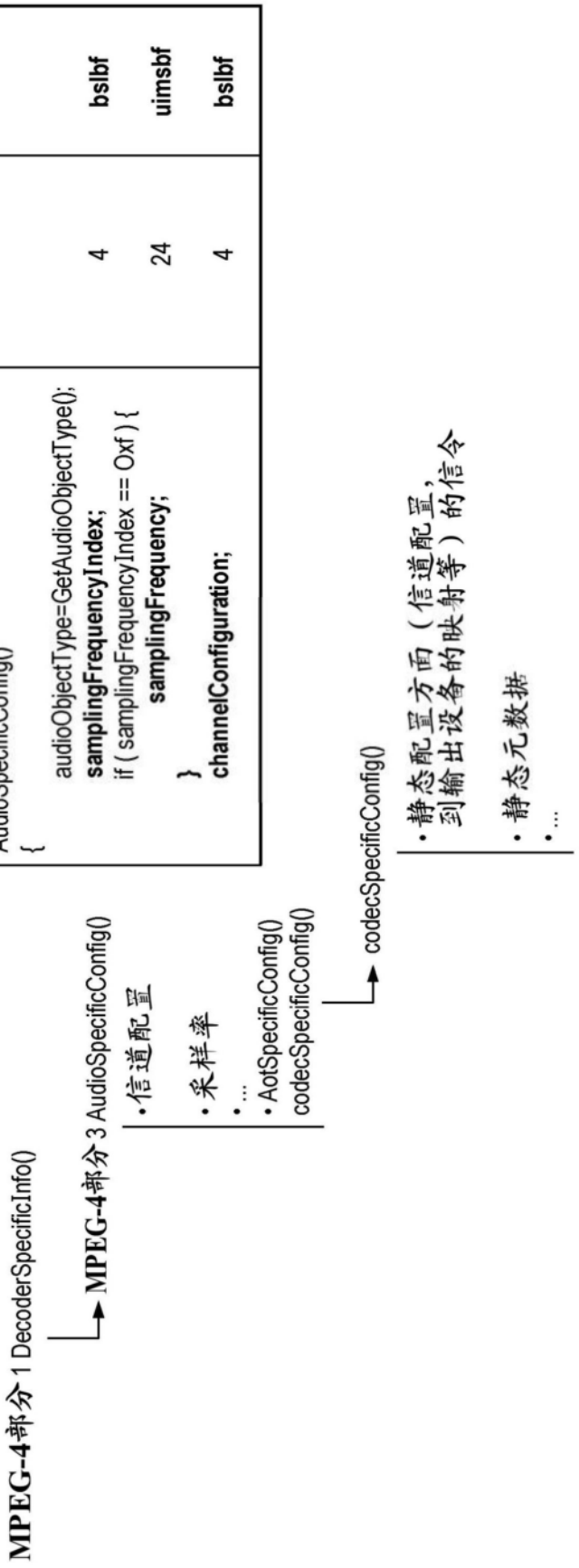

语法
比特数
助记符
AudioSpecificConfig()
{
audioObjectType=GetAudioObjectType();
samplingFrequencyIndex;
4
bslbf
if (samplingFrequencyIndex == 0xf) {
samplingFrequency;
24
uimsbf
}
channelConfiguration;
4
bslbf
MPEG-4部分1 DecoderSpecificInfo()
MPEG-4部分3 AudioSpecificConfig()
•信道配置
•采样率
•...
•AotSpecificConfig()
codecSpecificConfig()
codecSpecificConfig()
•静态配置方面（信道配置，到输出设备的映射等）的信令
•静态元数据
•...

图10

语法	比特数	助记符

raw_data_block()
{
while((id= **id_syn_ele**) != ID_END){
switch (id) {
case ID_SCE: single_channel_element();
break;
case ID_CPE: channel_pair_element();
break;
case ID_CCE: coupling_channel_element();
break;
case ID_LFE: life_channel_element();
break;
case ID_DSE: data_stream_element();
break;
case ID_PCE: program_config_element();
break;
case ID_FIL: fill_element();
	3	**uimsbf**

raw_data_block_codec(configData)

- 类似于**MPEG-4**中已有原始数据块
- 具有附加信息以使得能够跳过给定设备不需要的数据块的部分
- 动态元数据

图11

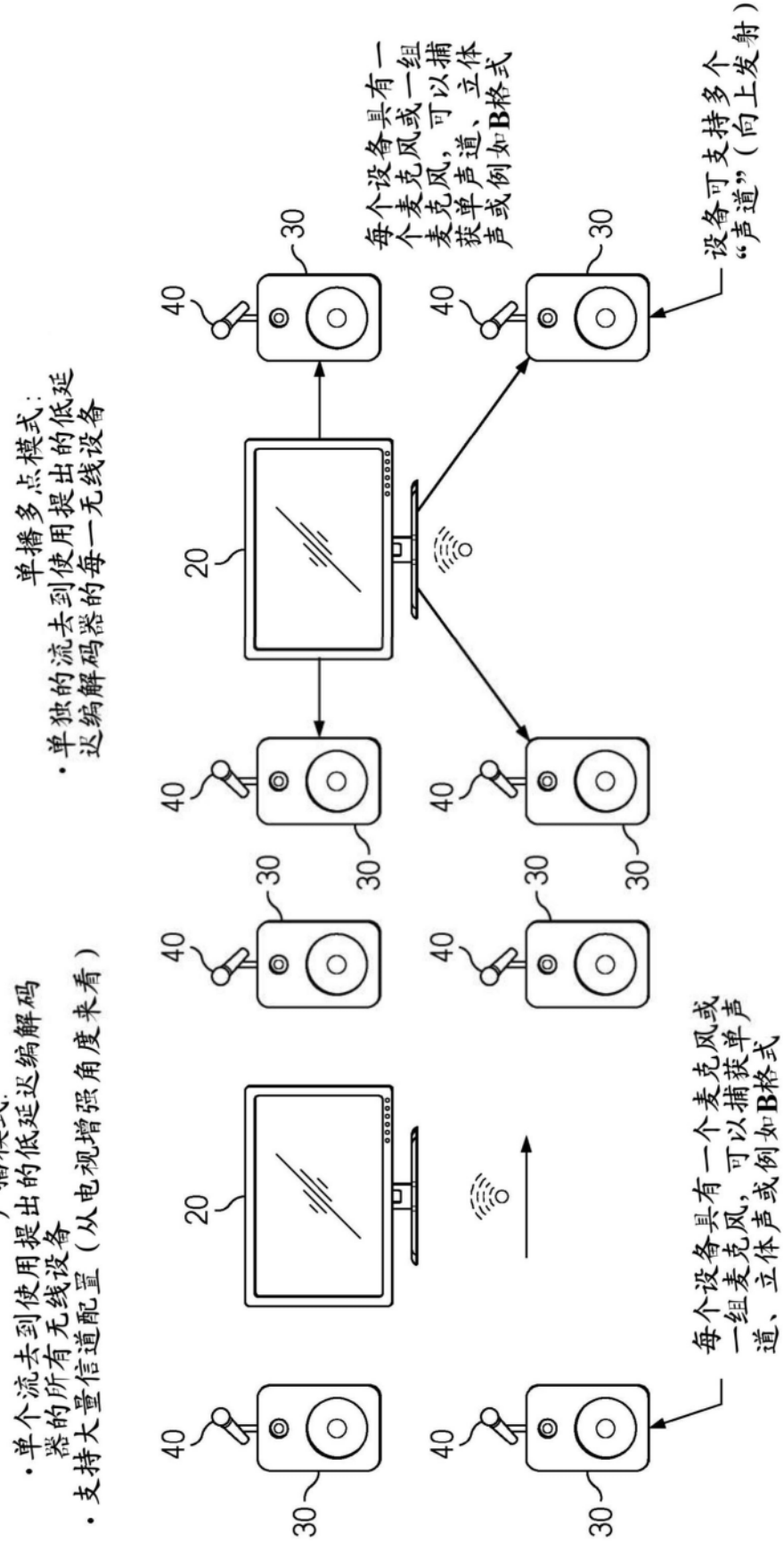

广播模式:
•单个流去到使用提出的低延迟编解码器的所有无线设备
•支持大量信道配置(从电视增强角度来看)
单播多点模式:
•单独的流去到使用提出的低延迟编解码器的每一无线设备
每个设备具有一个麦克风或一组麦克风,可以捕获单声道、立体声或例如B格式
设备可支持多个"声道"(向上发射)
每个设备具有一个麦克风或一组麦克风,可以捕获单声道、立体声或例如B格式
20
30
40

图12



图13

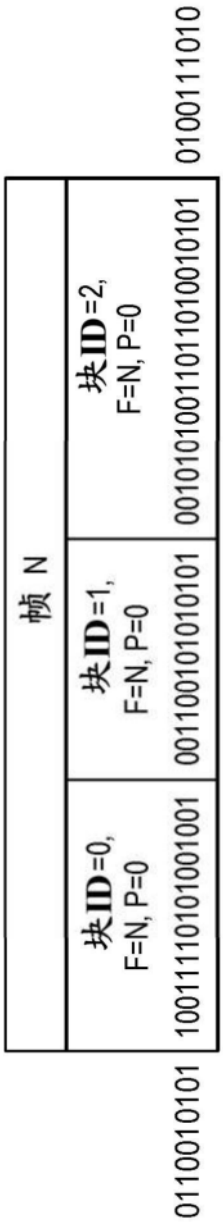

帧 N
块ID=0,
F=N, P=0
块ID=1,
F=N, P=0
块ID=2,
F=N, P=0
0110010101
1001111010100100 1
0011001010101010 1
0010101001101101 0010010101
0100111010

图14

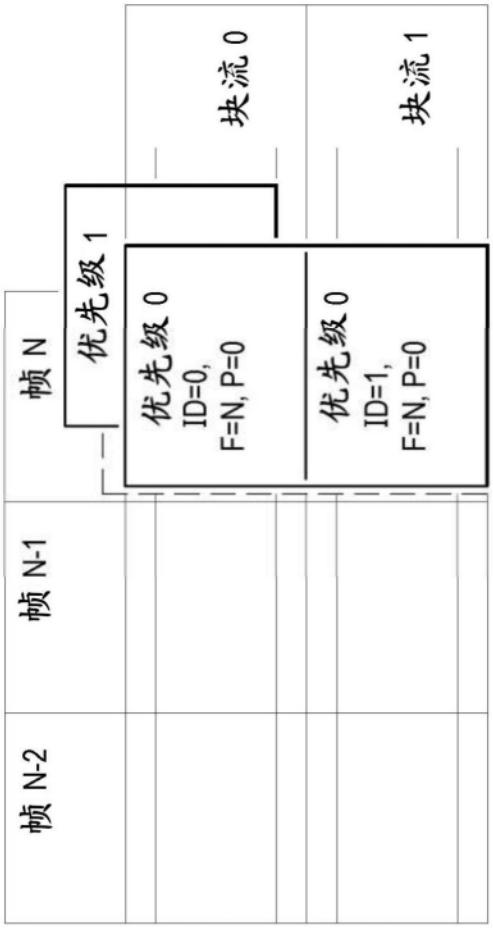


图15

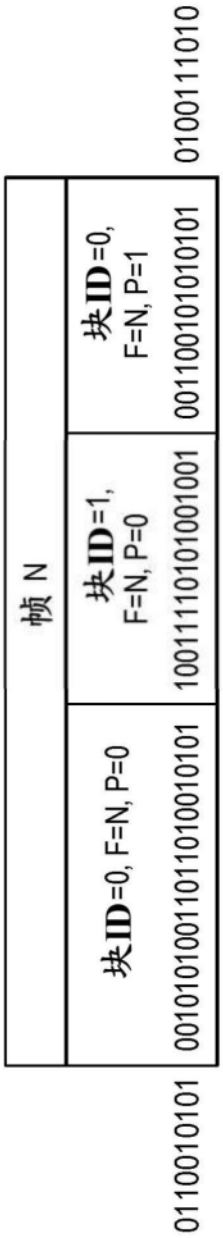

帧 N
块ID=0, F=N, P=0
0010101001101101010010101
块ID=1, F=N, P=0
1001111010100100 1001
块ID=0, F=N, P=1
0011001010101010 1
0110010101
0100111010

图16

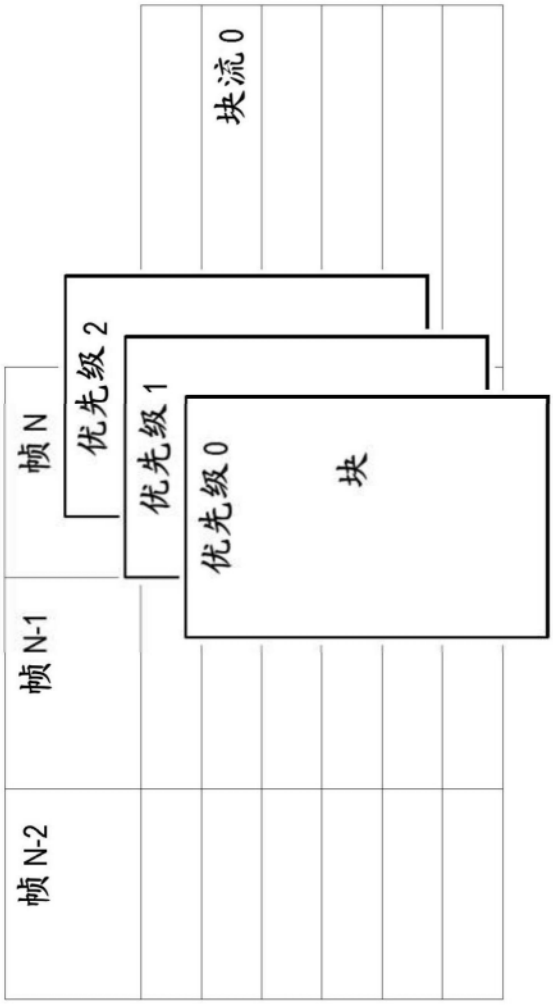


图17

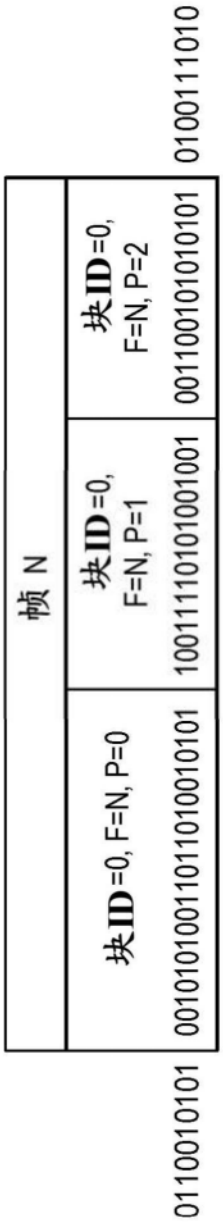

帧 N
块ID=0, F=N, P=0
0010101001101101010010101
0110010101
块ID=0, F=N, P=1
1001111010101001001
块ID=0, F=N, P=2
0011001010101010101
0100111010

图18