



(12) 发明专利

(10) 授权公告号 CN 112740325 B

(45) 授权公告日 2024.04.16

(21) 申请号 201980061575.5
 (22) 申请日 2019.08.20
 (65) 同一申请的已公布的文献号
 申请公布号 CN 112740325 A
 (43) 申请公布日 2021.04.30
 (30) 优先权数据
 62/720,680 2018.08.21 US
 (85) PCT国际申请进入国家阶段日
 2021.03.19
 (86) PCT国际申请的申请数据
 PCT/EP2019/072258 2019.08.20
 (87) PCT国际申请的公布数据
 W02020/038938 EN 2020.02.27
 (73) 专利权人 杜比国际公司
 地址 荷兰阿姆斯特丹
 (72) 发明人 克里斯托弗·费尔施

丹尼尔·费希尔

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287

专利代理师 刘锋

(51) Int.Cl.
G10L 19/16 (2006.01)

(56) 对比文件
 CN 101103393 A, 2008.01.09
 CN 105745704 A, 2016.07.06
 CN 107430867 A, 2017.12.01
 US 2016232910 A1, 2016.08.11
 US 2018197556 A1, 2018.07.12
 WO 2014124377 A2, 2014.08.14
 WO 2018130577 A1, 2018.07.19
 IN 201617011930 A, 2016.08.19

审查员 张彩云

权利要求书5页 说明书16页 附图6页

(54) 发明名称

即时播放帧 (IPF) 的生成、传输及处理的方法、设备及系统

(57) 摘要

本文描述一种用于对编码音频数据的位流进行解码的音频解码器,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值,所述音频解码器包括:确定器,所述确定器经配置以确定所述编码音频数据的位流的帧是否是包括与当前帧相关联的编码音频样本值及额外信息的即时播放帧;及初始化器,所述初始化器经配置以在所述确定器确定所述帧是即时播放帧时初始化所述解码器,其中初始化所述解码器包括在对与所述当前帧相关联的所述编码音频样本值进行解码之前对由所述额外信息包括的所述编码音频样本值进行解码。进一步描述一种用于对所述编码音频数据的位流进行解码的方法,以及一种音频编码器、一种音频编码器的系统及一种

用于生成具有即时播放帧的所述编码音频数据的位流的方法。此外,还描述一种用于在编码音频数据的位流中生成即时播放帧或用于从编码音频数据的位流中移除即时播放帧的设备及相应的非暂时性数字存储媒体。



1. 一种用于对编码音频数据的位流进行解码的音频解码器,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值,所述音频解码器包括:

确定器,所述确定器经配置以确定所述编码音频数据的位流的帧是否是包括与当前帧相关联的编码音频样本值及额外信息的即时播放帧,其中所述编码音频数据的位流为MPEG-4音频位流,其中所述额外信息是经由MPEG-4音频位流扩展机制来传输,所述MPEG-4音频位流扩展机制是新扩展有效负载元素类型EXT_AUDIO_PRE_ROLL的extension_payload元素,其中所述extension_payload元素位于所述MPEG-4音频位流中的第一位置,且其中所述extension_payload元素含在填充元素(ID_FIL)内,

其中所述额外信息包括在所述即时播放帧先前的数个帧的编码音频样本值,其中先前帧的所述编码音频样本值是使用与所述当前帧相同的编解码器配置进行编码,

其中对应于预滚帧的先前帧的数量对应于所述解码器建立完整信号以便每当即时播放帧被解码时就处于输出与所述当前帧相关联的有效音频样本值的位置所需的帧的数量;及

初始化器,所述初始化器经配置以在所述确定器确定所述帧是即时播放帧时初始化所述解码器,

其中初始化所述解码器包括在对与所述当前帧相关联的所述编码音频样本值进行解码之前对由所述额外信息包括的所述编码音频样本值进行解码,

其中如果所述确定器确定所述帧是即时播放帧,且如果所述当前帧的所述音频样本值已使用不同编解码器配置进行编码,那么所述初始化器经配置以将所述音频解码器从当前编解码器配置切换到所述不同编解码器配置,且

其中如果所述确定器确定所述帧是即时播放帧,且如果所述当前帧的所述音频样本值已使用所述当前编解码器配置进行编码,那么所述解码器经配置以使用所述当前编解码器配置对所述当前帧进行解码并丢弃所述额外信息。

2. 根据权利要求1所述的音频解码器,其中所述额外信息进一步包括关于用于对与所述当前帧相关联的所述音频样本值进行编码的所述编解码器配置的信息,且其中所述确定器经进一步配置以确定所述额外信息的所述编解码器配置是否与所述当前编解码器配置不同。

3. 根据权利要求1或2所述的音频解码器,其中所述即时播放帧包括作为扩展有效负载的所述额外信息,且其中所述确定器经配置以评估所述即时播放帧的所述扩展有效负载。

4. 根据权利要求1或2所述的音频解码器,其中所述extension_payload元素具有特定实例标签,所述特定实例标签发信号通知所述extension_payload元素的有效负载是所述额外信息。

5. 根据权利要求1或2所述的音频解码器,其中所述额外信息进一步包括唯一标识符,且其中任选地,所述唯一标识符用于检测所述不同编解码器配置。

6. 根据权利要求1或2所述的音频解码器,其进一步包括交叉渐变器,所述交叉渐变器经配置以对通过刷新先前编解码器配置中的所述解码器而获得的输出样本值及通过对与所述当前帧相关联的所述编码音频样本值进行解码而获得的输出样本值执行交叉渐变。

7. 根据权利要求1或2所述的音频解码器,其中包括在所述额外信息中的所述数个帧的

最早帧相对于所述最早帧先前的任何帧不进行时差编码或熵编码,且其中所述即时播放帧相对于所述即时播放帧先前的所述数个帧的所述最早帧先前的任何帧或相对于所述即时播放帧先前的任何帧不进行时差编码或熵编码。

8. 一种用于对编码音频数据的位流进行解码的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值,所述方法包括:

确定所述编码音频数据的位流的帧是否是包括与当前帧相关联的编码音频样本值及额外信息的即时播放帧,其中所述编码音频数据的位流为MPEG-4音频位流,其中所述额外信息是经由MPEG-4音频位流扩展机制来传输,所述MPEG-4音频位流扩展机制是新扩展有效负载元素类型EXT_AUDIO_PRE_ROLL的extension_payload元素,其中所述extension_payload元素位于所述MPEG-4音频位流中的第一位置,且其中所述extension_payload元素含在填充元素(ID_FIL)内,

其中所述额外信息包括在所述即时播放帧先前的数个帧的编码音频样本值,

其中所述先前帧的所述编码音频样本值是使用与所述即时播放帧相同的编解码器配置进行编码,

其中对应于预滚帧的先前帧的数量对应于解码器建立完整信号以便每当即时播放帧被解码时就处于输出与所述当前帧相关联的有效音频样本值的位置所需的帧的数量;

如果确定所述帧是即时播放帧,那么初始化所述解码器,其中所述初始化包括在对与所述当前帧相关联的所述编码音频样本值进行解码之前对由所述额外信息包括的所述编码音频样本值进行解码;

如果确定所述帧是即时播放帧,且如果所述即时播放帧的所述音频样本值已使用不同编解码器配置进行编码,那么将所述音频解码器从当前编解码器配置切换到所述不同编解码器配置;及

如果确定所述帧是即时播放帧,且如果所述即时播放帧的所述音频样本值已使用所述当前编解码器配置进行编码,那么使用所述当前编解码器配置对所述即时播放帧进行解码并丢弃所述额外信息。

9. 根据权利要求8所述的方法,其中所述额外信息进一步包括关于用于对与所述当前帧相关联的所述音频样本值进行编码的所述编解码器配置的信息,所述方法进一步包括确定所述额外信息的所述编解码器配置是否与用于对与所述即时播放帧先前的所述位流中的帧相关联的音频样本值进行编码的所述当前编解码器配置不同。

10. 根据权利要求8或9所述的方法,其中所述extension_payload元素具有特定实例标签,所述特定实例标签发信号通知所述extension_payload元素的有效负载是所述额外信息。

11. 根据权利要求8或9所述的方法,其中所述额外信息进一步包括唯一标识符,且其中任选地,所述唯一标识符用于检测所述不同编解码器配置。

12. 根据权利要求8或9所述的方法,其中所述编码音频数据的位流包括使用第一编解码器配置编码的第一数量的帧及在所述第一数量的帧之后并使用第二编解码器配置编码的第二数量的帧,其中所述第二数量的帧的第一帧为所述即时播放帧。

13. 一种用于生成具有即时播放帧的编码音频数据的位流的音频编码器,其中所述编

码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频数据音频样本值,其中所述音频编码器包括:

核心编码器,所述核心编码器经配置以使用预定义编解码器配置对与所述多个帧相关联的未压缩音频样本值进行编码;

缓冲器,所述缓冲器经配置以存储使用所述预定义编解码器配置编码的所述多个帧的当前帧的数个先前帧的编码音频样本值;及

嵌入器,所述嵌入器经配置以在所述多个帧的所述当前帧中写入即时播放帧,其中所述即时播放帧包括与所述当前帧相关联的编码音频样本值及与所述当前帧的所述数个先前帧的所述编码音频样本值相对应的额外信息,其中所述生成的编码音频数据的位流为MPEG-4音频位流,其中所述嵌入器经进一步配置以:经由MPEG-4音频位流扩展机制将所述额外信息嵌入在所述位流中,所述MPEG-4音频位流扩展机制是新扩展有效负载元素类型EXT_AUDIO_PRE_ROLL的extension_payload元素;将所述extension_payload元素定位在所述MPEG-4音频位流中的第一位置;并将所述extension_payload元素嵌入在填充元素(ID_FIL)内。

14. 根据权利要求13所述的音频编码器,其中所述嵌入器经进一步配置以在所述额外信息中包含关于所述预定义编解码器配置的信息。

15. 根据权利要求13或14所述的音频编码器,其中所述嵌入器经进一步配置以在所述即时播放帧中包含所述额外信息。

16. 根据权利要求13或14所述的音频编码器,其中所述嵌入器经进一步配置以分配特定实例标签,所述特定实例标签发信号通知所述extension_payload元素的有效负载是所述额外信息。

17. 根据权利要求13或14所述的音频编码器,其中所述嵌入器经进一步配置以将唯一标识符包含到所述额外信息中,且其中任选地,所述唯一标识符发信号通知所述预定义编解码器配置。

18. 根据权利要求13或14所述的音频编码器,其中所述音频编码器经进一步配置以使包括在所述额外信息中的所述数个帧的最早帧相对于所述最早帧先前的任何帧不进行时差编码或熵编码,且其中所述音频编码器经进一步配置以使所述即时播放帧相对于所述即时播放帧先前的所述数个帧的所述最早帧先前的任何帧或相对于所述即时播放帧先前的任何帧不进行时差编码或熵编码。

19. 一种系统,所述系统包括两个或多于两个根据权利要求13到18中任一权利要求所述的音频编码器,以用于生成编码音频数据的多个位流,每一位流具有即时播放帧,其中编码音频数据的每一位流表示音频样本值的序列并包括多个帧,且其中每一帧包括相关联的编码音频样本值。

20. 根据权利要求19所述的系统,其中对于所述两个或多于两个音频编码器的所述核心编码器中的每一者,预定取样率相同。

21. 根据权利要求19或20所述的系统,其中所述系统进一步包括用于延迟对准所述多个位流的延迟对准单元。

22. 一种由音频编码器生成具有即时播放帧的编码音频数据的位流的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码

音频样本值,所述方法包括以下步骤:

由核心编码器使用预定义编解码器配置对与所述多个帧相关联的未压缩音频样本值进行编码;

由缓冲器存储使用所述预定义编解码器配置编码的所述多个帧的当前帧的数个先前帧的编码音频样本值;及

由嵌入器在所述多个帧的所述当前帧中写入即时播放帧,其中所述即时播放帧包括与所述当前帧相关联的编码音频样本值及与所述当前帧的所述数个先前帧的所述编码音频样本值相对应的额外信息,其中所述所生成的编码音频数据的位流为MPEG-4音频位流,其中所述额外信息由所述嵌入器经由MPEG-4音频位流扩展机制嵌入在所述位流中,所述MPEG-4音频位流扩展机制是新扩展有效负载元素类型EXT_AUDIO_PRE_ROLL的extension_payload元素,其中所述extension_payload元素由所述嵌入器定位在所述MPEG-4音频位流中的第一位置,且其中所述extension_payload元素由所述嵌入器嵌入在填充元素(ID_FIL)内。

23. 根据权利要求22所述的方法,其中所述额外信息进一步包括关于所述预定义编解码器配置的信息。

24. 根据权利要求22或23所述的方法,其中所述即时播放帧进一步包括所述额外信息。

25. 根据权利要求22或23所述的方法,其中所述extension_payload元素由所述嵌入器分配特定实例标签,所述特定实例标签发信号通知所述extension_payload元素的有效负载是所述额外信息。

26. 根据权利要求22或23所述的方法,其中所述额外信息进一步包括唯一标识符,且其中任选地,所述唯一标识符发信号通知所述预定义编解码器配置。

27. 根据权利要求22或23所述的方法,其中由所述音频编码器使包括在所述额外信息中的所述数个帧的最早帧相对于所述最早帧先前的任何帧不进行时差编码或熵编码,且其中由所述音频编码器使所述即时播放帧相对于所述即时播放帧先前的所述数个帧的所述最早帧先前的任何帧或相对于所述即时播放帧先前的任何帧不进行时差编码或熵编码。

28. 一种用于在编码音频数据的位流中生成即时播放帧的设备,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值,其中所述设备包括:

接收器,所述接收器经配置以接收编码音频数据的MPEG-4音频位流,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值;

缓冲器,所述缓冲器经配置以存储所述多个帧的当前帧的数个先前帧的编码音频样本值,及

嵌入器,所述嵌入器经配置以在所述多个帧的所述当前帧中写入即时播放帧,其中所述即时播放帧包括与所述当前帧相关联的编码音频样本值及与所述当前帧的所述数个先前帧的编码音频样本值相对应的额外信息,其中所述嵌入器经进一步配置以:经由MPEG-4音频位流扩展机制将所述额外信息嵌入在所述位流中,所述MPEG-4音频位流扩展机制是新扩展有效负载元素类型EXT_AUDIO_PRE_ROLL的extension_payload元素;将所述extension_payload元素定位在所述MPEG-4音频位流中的第一位置;并将所述extension_

payload元素嵌入在填充元素(ID_FIL)内。

29.一种用于从编码音频数据的位流中移除即时播放帧的设备,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值,其中所述设备包括:

接收器,所述接收器经配置以接收编码音频数据的位流,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值;及

嵌入器,所述嵌入器经配置以通过从即时播放帧中移除与所述即时播放帧已被写入的所述多个帧的当前帧的数个先前帧的编码音频样本值相对应的额外信息,将所述即时播放帧转换为正常帧。

30.一种非暂时性数字存储媒体,其具有存储在其上的计算机程序,以在所述计算机程序由计算机或处理器运行时执行根据权利要求8到12中任一权利要求所述的方法。

31.一种非暂时性数字存储媒体,其具有存储在其上的计算机程序,以在所述计算机程序由计算机或处理器运行时执行根据权利要求22到27中任一权利要求所述的方法。

即时播放帧(IPF)的生成、传输及处理的方法、设备及系统

[0001] 相关申请案的交叉引用

[0002] 本申请案要求以下优先权申请案的优先权:2018年8月21日申请的第62/720,680号美国临时申请案(参考文献:D18080USP1),其通过引用的方式并入本文中。

技术领域

[0003] 本公开大体上涉及音频编码器、编码方法、音频解码器及解码方法,包含:一种用于对编码音频数据的位流进行解码的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值;及一种用于生成具有即时播放帧的编码音频数据的位流的方法。本公开进一步涉及一种用于在编码音频数据的位流中生成即时播放帧或用于从编码音频数据的位流中移除即时播放帧的设备。

[0004] 虽然本文将特别参考所述公开内容来描述一些实施例,但应理解,本公开不限于此类使用领域,且适用于更广泛的上下文。

背景技术

[0005] 目前,在ISO/IEC 14496-3《视听对象编码第3部分:用于生成、传输及处理即时播放帧(IPF)的音频(Coding of audio-visual objects-Part 3:Audio,for generating,transporting and processing Immediate Playout Frames(IPF))》中标准化的MPEG-4音频中还存在不足。IPF向特殊帧提供信息,所述特殊帧允许即时初始化解码器,从而在切换到包括所述特殊帧的数据流时即时播放。以另一种方式陈述,IPF是解码器在其接收时能够即时从编码到此IPF中的第一样本产生正确样本的帧,因为它含有要这样做的所有信息。因此,IPF表示只能使用来自其自身的信息来解码的独立可解码帧。

[0006] 编码音频通常以数据帧或数据块的形式出现。在MPEG-4中标准化的音频上下文中,帧/块可被称为颗粒,编码块/帧被称为接入单元(AU),及解码块被称为合成单元(CU)。在传输系统中,音频信号只能在这些编码块(接入单元)的粒度内进行接入及寻址。

[0007] 在自适应流的上下文中,当音频切换到不同配置(例如,不同位率,例如在MPEG-DASH中的自适应集内配置的位率)时,为了从一开始就准确地再现音频样本,解码器需要提供表示音频节目的对应时间段的 AU_n ,及 AU_n 先前的额外 AU_{n-1} 、 AU_{n-2} 、 \dots AUs及配置数据。否则,由于不同编码配置(例如,窗口数据、SBR相关数据、PS相关数据),不能保证解码器在仅解码 AU_n 时产生正确的输出。因此,要用新配置解码的第一 AU_n 必须携带新配置数据及用新配置初始化解码器所需的所有预滚数据(以 AU_{n-x} 的形式,表示 AU_n 之前的时间段)。这可通过MPEG-H 3D音频标准或MPEG-D USAC标准中界定的即时播放帧(IPF)来实现。

[0008] 鉴于上文,因此本发明的目的是提供能够处理MPEG-4音频中的IPF的音频解码器与解码方法以及音频编码器、音频编码器的系统、设备及编码方法。

发明内容

[0009] 根据本公开的第一方面,提供一种用于对编码音频数据的位流进行解码的音频解

码器,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0010] 解码器可包括确定器,所述确定器经配置以确定所述编码音频数据的位流的帧是否是包括与当前帧相关联的编码音频样本值及额外信息的即时播放帧,其中所述额外信息可包括在所述即时播放帧先前的数个帧的编码音频样本值,其中所述先前帧的所述编码音频样本值可使用与所述当前帧相同的编解码器配置进行编码,其中对应于预滚帧的先前帧的数量可对应于所述解码器建立完整信号以便每当即时播放帧被解码时就处于输出与所述当前帧相关联的有效音频样本值的位置所需的帧的数量。

[0011] 且解码器可包括初始化器,所述初始化器经配置以在所述确定器确定所述帧是即时播放帧时初始化所述解码器,其中初始化所述解码器可包括在对与所述当前帧相关联的所述编码音频样本值进行解码之前对由所述额外信息包括的所述编码音频样本值进行解码,其中如果所述确定器确定所述帧是即时播放帧,且如果所述当前帧的所述音频样本值已使用不同编解码器配置进行编码,那么所述初始化器可经配置以将所述音频解码器从当前编解码器配置切换到所述不同编解码器配置,且其中如果所述确定器确定所述帧是即时播放帧,且如果所述当前帧的所述音频样本值已使用所述当前编解码器配置进行编码,那么所述解码器可经配置以使用所述当前编解码器配置对所述当前帧进行解码并丢弃所述额外信息。

[0012] 在一些实施例中,所述额外信息可进一步包括关于用于对与所述当前帧相关联的所述音频样本值进行编码的所述编解码器配置的信息,且所述确定器可经进一步配置以确定所述额外信息的所述编解码器配置是否与所述当前编解码器配置不同。

[0013] 在一些实施例中,所述即时播放帧可包括作为扩展有效负载的所述额外信息,且所述确定器可经配置以评估所述即时播放帧的所述扩展有效负载。

[0014] 在一些实施例中,所述编码音频数据的位流可为MPEG-4音频位流。

[0015] 在一些实施例中,所述额外信息可经由MPEG-4音频位流扩展机制来传输,所述MPEG-4音频位流扩展机制是数据流元素(DSE)或extension_payload元素。

[0016] 在一些实施例中,所述数据流元素(DSE)或所述extension_payload元素可位于所述MPEG-4音频位流中的预定义位置及/或可具有特定实例标签,所述特定实例标签发信号通知所述数据流元素(DSE)或所述extension_payload元素的有效负载是所述额外信息。

[0017] 所述extension_payload元素可例如含在所述MPEG-4音频位流语法的不同位置。因此,这允许也在MPEG-4音频中使用即时播放帧功能。

[0018] 在一些实施例中,所述extension_payload元素可含在填充元素(ID_FIL)内。

[0019] 在一些实施例中,所述额外信息可进一步包括唯一标识符,且任选地,所述唯一标识符可用于检测所述不同编解码器配置。

[0020] 在一些实施例中,所述解码器可进一步包括交叉渐变器,所述交叉渐变器经配置以对通过刷新先前编解码器配置中的所述解码器而获得的输出样本值及通过对与所述当前帧相关联的所述编码音频样本值进行解码而获得的输出样本值执行交叉渐变。

[0021] 在一些实施例中,包括在所述额外信息中的所述数个帧的最早帧可相对于所述最早帧先前的任何帧不进行时差编码或熵编码,且所述即时播放帧可相对于所述即时播放帧先前的所述数个帧的所述最早帧先前的任何帧或相对于所述即时播放帧先前的任何帧不

进行时差编码或熵编码。

[0022] 根据本公开的第二方面,提供一种用于对编码音频数据的位流进行解码的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0023] 所述方法可包括确定所述编码音频数据的位流的帧是否是包括与当前帧相关联的编码音频样本值及额外信息的即时播放帧,其中所述额外信息可包括在所述即时播放帧先前的数个帧的编码音频样本值,其中所述先前帧的所述编码音频样本值可使用与所述即时播放帧相同的编解码器配置进行编码,其中对应于预滚帧的先前帧的数量可对应于解码器建立完整信号以便每当即时播放帧被解码时就处于输出与所述当前帧相关联的有效音频样本值的位置所需的帧的数量。

[0024] 所述方法可进一步包括:如果确定所述帧是即时播放帧,那么初始化所述解码器,其中所述初始化可包括在对与所述当前帧相关联的所述编码音频样本值进行解码之前对由所述额外信息包括的所述编码音频样本值进行解码。

[0025] 所述方法可进一步包括:如果确定所述帧是即时播放帧,且如果所述即时播放帧的所述音频样本值已使用不同编解码器配置进行编码,那么将所述音频解码器从当前编解码器配置切换到所述不同编解码器配置。

[0026] 且所述方法可包括:如果确定所述帧是即时播放帧,且如果所述即时播放帧的所述音频样本值已使用所述当前编解码器配置进行编码,那么使用所述当前编解码器配置对所述即时播放帧进行解码并丢弃所述额外信息。

[0027] 如所提议的那样配置,所述方法允许例如如ISO/IEC 14496-3中所界定的音频对象类型(AOT)的切换,结合连续地产生正确的输出样本并在所述音频输出中不引入静默间隙。

[0028] 在一些实施例中,所述额外信息可进一步包括关于用于对与所述当前帧相关联的所述音频样本值进行编码的所述编解码器配置的信息,其中所述方法可进一步包括确定所述额外信息的所述编解码器配置是否与用于对与所述即时播放帧先前的所述位流中的帧相关联的音频样本值进行编码的所述当前编解码器配置不同。

[0029] 在一些实施例中,所述编码音频数据的位流可为MPEG-4音频位流。

[0030] 在一些实施例中,所述额外信息可经由MPEG-4音频位流扩展机制来传输,所述MPEG-4音频位流扩展机制是数据流元素(ID_DSE)或extension_payload元素。

[0031] 在一些实施例中,所述数据流元素(ID_DSE)或所述extension_payload元素可位于MPEG-4音频位流中的预定义位置及/或可具有特定实例标签,所述特定实例标签发信号通知所述数据流元素(ID_DSE)或所述extension_payload元素的有效负载是所述额外信息。

[0032] 在一些实施例中,所述extension_payload元素可含在填充元素(ID_FIL)内。

[0033] 在一些实施例中,所述额外信息可进一步包括唯一标识符,且任选地,所述唯一标识符可用于检测所述不同编解码器配置。

[0034] 在一些实施例中,所述编码音频数据的位流可包括使用第一编解码器配置编码的第一数量的帧及在所述第一数量的帧之后并使用第二编解码器配置编码的第二数量的帧,其中所述第二数量的帧的第一帧可为所述即时播放帧。

[0035] 根据本公开的第三方面,提供一种用于生成具有即时播放帧的编码音频数据的位流的音频编码器,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频数据音频样本值。

[0036] 所述音频编码器可包括核心编码器,所述核心编码器经配置以使用预定义编解码器配置对与所述多个帧相关联的未压缩音频样本值进行编码。

[0037] 所述音频编码器可进一步包括缓冲器,所述缓冲器经配置以存储使用所述预定义编解码器配置编码的所述多个帧的当前帧的数个先前帧的编码音频样本值。

[0038] 且所述音频编码器可包括嵌入器,所述嵌入器经配置以在所述多个帧的所述当前帧中写入即时播放帧,其中所述即时播放帧可包括与所述当前帧相关联的编码音频样本值及与所述当前帧的所述数个先前帧的所述编码音频样本值相对应的额外信息。

[0039] 在一些实施例中,所述嵌入器可经进一步配置以在所述额外信息中包含关于所述预定义编解码器配置的信息。

[0040] 在一些实施例中,所述嵌入器可经进一步配置以在所述即时播放帧中包含所述额外信息。

[0041] 在一些实施例中,所述所生成的编码音频数据的位流可为MPEG-4音频位流。

[0042] 在一些实施例中,所述嵌入器可经进一步配置以经由MPEG-4音频位流扩展机制将所述额外信息嵌入在所述位流中,所述MPEG-4音频位流扩展机制是数据流元素(ID_DSE)或extension_payload元素。

[0043] 在一些实施例中,所述嵌入器可经进一步配置以将所述数据流元素(ID_DSE)或所述extension_payload元素定位在所述MPEG-4音频位流中的预定义位置及/或分配特定实例标签,所述特定实例标签发信号通知所述数据流元素(ID_DSE)或所述extension_payload元素的有效负载是所述额外信息。

[0044] 在一些实施例中,所述嵌入器可经进一步配置以将所述extension_payload元素嵌入在填充元素(ID_FIL)内。

[0045] 在一些实施例中,所述嵌入器可经进一步配置以将唯一标识符包含到所述额外信息中,且任选地,所述唯一标识符可发信号通知所述预定义编解码器配置。

[0046] 在一些实施例中,所述音频编码器可经进一步配置以使包括在所述额外信息中的所述数个帧的最早帧相对于所述最早帧之前的任何帧不进行时差编码或熵编码,且所述音频编码器可经进一步配置以使所述即时播放帧相对于所述即时播放帧之前的所述数个帧的所述最早帧之前的任何帧或相对于所述即时播放帧之前的任何帧不进行时差编码或熵编码。

[0047] 根据本公开的第四方面,提供一种系统,所述系统包括两个或多于两个音频编码器,以用于生成编码音频数据的多个位流,每一位流具有即时播放帧,其中编码音频数据的每一位流表示音频样本值的序列并包括多个帧,且其中每一帧包括相关联的编码音频样本值。

[0048] 在一些实施例中,对于所述两个或多于两个音频编码器的核心编码器中的每一者,预定取样率可相同。因此,可避免解码器处的重取样及额外延迟处理。

[0049] 在一些实施例中,所述系统可进一步包括用于延迟对准所述多个位流的延迟对准单元。因此,这允许通过补偿不同编码器延迟在所述解码器处无缝切换。

[0050] 根据本公开的第五方面,提供一种由音频编码器生成具有即时播放帧的编码音频数据的位流的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0051] 所述方法可包括由核心编码器使用预定义编解码器配置对与所述多个帧相关联的未压缩音频样本值进行编码的步骤。

[0052] 所述方法可进一步包括由缓冲器存储使用所述预定义编解码器配置编码的所述多个帧的当前帧的数个先前帧的编码音频样本值的步骤。

[0053] 且所述方法可进一步包括由嵌入器在所述多个帧的所述当前帧中写入即时播放帧的步骤,其中所述即时播放帧可包括与所述当前帧相关联的编码音频样本值及与所述当前帧的所述数个先前帧的所述编码音频样本值相对应的额外信息。

[0054] 在一些实施例中,所述额外信息可进一步包括关于所述预定义编解码器配置的信息。

[0055] 在一些实施例中,所述即时播放帧可进一步包括所述额外信息。

[0056] 在一些实施例中,所述所生成的编码音频数据的位流可为MPEG-4音频位流。

[0057] 在一些实施例中,所述额外信息可由所述嵌入器经由MPEG-4音频位流扩展机制嵌入在所述位流中,所述MPEG-4音频位流扩展机制可为数据流元素(ID_DSE)或extension_payload元素。

[0058] 在一些实施例中,所述数据流元素(ID_DSE)或所述extension_payload元素可由所述嵌入器定位在所述MPEG-4音频位流中的预定义位置处及/或可被分配特定实例标签,所述特定实例标签发信号通知所述数据流元素(ID_DSE)或所述extension_payload元素的有效负载是所述额外信息。

[0059] 在一些实施例中,所述extension_payload元素可由所述嵌入器嵌入在填充元素(ID_FIL)内。

[0060] 在一些实施例中,所述额外信息可进一步包括唯一标识符,且任选地,所述唯一标识符可发信号通知所述预定义编解码器配置。

[0061] 在一些实施例中,由所述音频编码器使包括在所述额外信息中的所述数个帧的最早帧可相对于最早帧先前的任何帧不进行时差编码或熵编码,且由所述音频编码器使所述即时播放帧可相对于所述即时播放帧先前的所述数个帧中的所述最早帧先前的任何帧或相对于所述即时播放帧先前的任何帧不进行时差编码或熵编码。

[0062] 根据本公开的第六方面,提供一种用于在编码音频数据的位流中生成即时播放帧或用于从编码音频数据的位流中移除即时播放帧的设备,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0063] 所述设备可包括接收器,所述接收器经配置以接收编码音频数据的位流,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0064] 且所述设备可包括嵌入器,所述嵌入器经配置以在所述多个帧的当前帧中写入即时播放帧,其中所述即时播放帧可包括与所述当前帧相关联的编码音频样本值及与所述当前帧的数个先前帧的编码音频样本值相对应的额外信息。

[0065] 如所提议的那样配置,所述设备允许在编码音频数据的任何已存在的位流中单独

地生成即时播放帧,即在分发之前(如果需要)。

[0066] 在一些实施例中,所述设备可进一步包括缓冲器,所述缓冲器经配置以存储所述多个帧的所述当前帧的所述数个先前帧的编码音频样本值。

[0067] 在一些实施例中,所述嵌入器可经进一步配置以从所述即时播放帧中移除与所述当前帧的所述数个先前帧的所述编码音频样本值相对应的所述额外信息。

[0068] 因此,这允许例如在音频编码器仅生成即时播放帧的情况下,从编码音频数据的位流中单独地移除即时播放帧。

[0069] 根据本公开的第七方面,提供一种非暂时性数字存储媒体,其具有存储在其上的计算机程序,以在所述计算机程序由计算机或处理器运行时执行用于对编码音频数据的位流进行解码的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0070] 根据本公开的第八方面,提供一种非暂时性数字存储媒体,其具有存储在其上的计算机程序,以在所述计算机程序由计算机或处理器运行时执行由音频编码器生成具有即时播放帧的编码音频数据的位流的方法,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

附图说明

[0071] 现在将仅以示例的方式,参考附图来描述本公开的示例实施例,其中:

[0072] 图1说明编码音频数据的MPEG-4音频位流中的即时播放帧的示例。

[0073] 图2说明用于对编码音频数据的位流进行解码的方法的示例,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0074] 图3说明用于对编码音频数据的位流进行解码的方法的另一示例,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0075] 图4说明用于对编码音频数据的位流进行解码的音频解码器的示例,其中所述编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0076] 图5说明用于生成具有随机接入点(即时播放帧,IPF)的编码音频数据的位流的音频编码器的示例。

[0077] 图6说明用于生成具有随机接入点(即时播放帧,IPF)的编码音频数据的位流的方法的示例。

[0078] 图7说明用于生成编码音频数据的多个位流的音频编码器的系统的示例,每一位流具有时间对准的随机接入点(即时播放帧,IPF)。

[0079] 图8说明用于在编码音频数据的位流中生成随机接入点(即时播放帧,IPF)或用于从编码音频数据的位流中移除随机接入点(即时播放帧,IPF)的设备的示例。

[0080] 图9说明具有用于运行存储在非暂时性数字存储媒体上的计算机程序的处理器装置的示例。

具体实施方式

[0081] 本发明涉及MPEG-4音频中IPF的创建(编码)、传输(位流)及处理(解码),例如在包含根据另一标准(例如MPEG-H 3D音频标准)标准化的音频的数据流的上下文中。在这里及下文中,MPEG-4音频位流可指与ISO/IEC 14496-3《视听对象编码第3部分:音频》及其所有未来版本、修订及修正案(以下简称“MPEG-4音频”)中规定的标准兼容的位流。为了在MPEG-4音频中启用IPF功能,有几个选项用于在AU_n之前及时生成并传输AUs及配置数据,作为用于AU_n的相同有效负载包的一部分。这样做是为了能够从解码AU_n得到的第一个样本正确输出音频样本。

[0082] 编码音频数据的位流可含有音频样本值的序列,例如有效负载分组。编码音频数据的位流可进一步包括多个帧。每一帧可包括相关联的编码音频样本值。换句话说,每一有效负载分组可属于相应的帧或AU。

[0083] 在一个示例中,音频有效负载分组可符合如ISO/IEC 14496-3中所界定的以下语法:

```
raw_data_block() {
    while( (id = id_syn_ele) != ID_END ){ 3 uimbsf
        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: lfe_channel_element(); break;
[0084] case ID_DSE: data_stream_element(); break;
            case ID_PCE: program_config_element(); break;
            case ID_FIL: fill_element();
        }
    }
    byte_align()
}
```

[0085] 上述音频有效负载分组可与MPEG标准的当前及未来版本(例如MPEG-4音频标准)兼容。在一个实施例中,编码音频数据的位流可为MPEG-4音频位流(即MPEG-4音频标准兼容位流)。

[0086] 编码音频数据的位流的帧可为即时播放帧(随机接入点、特殊帧),包括与当前帧相关联的编码音频样本值及额外信息。额外信息可包括在即时播放帧先前的数个帧的编码音频样本值,其中先前帧的编码音频样本值可使用与当前帧相同的编解码器配置进行编码。对应于预滚帧的先前帧的数量可对应于解码器建立完整信号以便每当即时播放帧被解码时就处于输出与当前帧相关联的有效音频样本值的位置所需的帧的数量。例如,可在解码器的启动或重新启动期间建立完整信号。即时播放帧可为例如解码器启动时的第一帧。

[0087] 在一个实施例中,额外信息可经由MPEG-4音频位流扩展机制来传输,所述MPEG-4

音频位流扩展机制可为数据流元素 (ID_DSE) 或 extension_payload 元素。extension_payload 元素可例如含在 MPEG-4 音频位流语法的不同位置, 例如在不同层级。在一个实施例中, extension_payload 元素可含在填充元素 (ID_FIL) 内。

[0088] 因此, 可经由 MPEG-4 音频位流扩展机制来传输额外信息, 例如, 基于以下选项之一:

[0089] 选项1:

[0090] raw_data_block() -> case ID_DSE -> data_stream_element() ->

[0091] data_stream_byte[element_instance_tag][i];

[0092] 选项2:

[0093] raw_data_block() -> case ID_FIL -> fill_element() -> extension_payload (cnt) ->

[0094] switch(extension_type) -> extension_type == EXT_DATA_ELEMENT (+ convention

[0095] how to identify) or EXT_AUDIO_PRE_ROLL

[0096] 选项3:

[0097] er_raw_data_block() or er_raw_data_block_eld() -> extension_payload (cnt) ->

[0098] switch(extension_type) -> extension_type == EXT_DATA_ELEMENT (+ convention

[0099] how to identify) or EXT_AUDIO_PRE_ROLL

[0100] ISO/IEC 14496-3 及/或未来标准中界定的数据流元素 (经由等于 ID_DSE 的 id_syn_ele 发信号通知) 或填充元素 (经由等于 ID_FIL 的 id_syn_ele 发信号通知) 或其等效者可用于承载扩展有效负载, 其可用于进一步扩展在此类有效负载分组中传输的信息而不破坏与遗留解码器的兼容性。

[0101] 因此, 在 MPEG-4 音频上下文中, 可使用数据流元素 (ID_DSE) 或可包含在填充元素 (ID_FIL) 内的 extension_payload 元素, 以便在与 AU_n (即当前帧、即时播放帧) 相同的有效负载分组内传输表示 AU_n (即额外信息) 之前的时间段的 AUs 及配置信息。这可进一步用于应用处理, 其允许在 MPEG-4 音频中也使用 IPF 功能。类似地, 在 MPEG-D USAC 中, 扩展机制 (usacExtElement) 可用来传递 AudioPreRoll() 有效负载, 但有一些不同。

[0102] 在一个实施例中, 可将扩展元素 (例如 extension_payload 元素) 插入到 AU 中, 其中元素类型在位流中针对每一元素及 AU 发信号通知。在一个示例中, 扩展元素可为有效负载分组中的第一元素, 在第一音频元素先前。

[0103] 下面界定的每一音频预滚元素可由通用唯一标识符 (UUID) 来识别。UUID 字段可用于由音频编码器发信号通知并由音频解码器检测流配置的切换。如果 UUID 字段相对于先前帧或初始状态已更改 (例如, 在第一解码器启动时), 那么流配置可能已更改, 且必须评估预滚有效负载以确保正确解码。如果 UUID 相对于先前帧没有更改, 解码器可跳过 audio_preroll_element() 有效负载并继续进行常规解码。

[0104] 在一个示例中, 如果不存在 UUID, 解码器可将 audio_preroll_element() 的 AudioSpecificConfig 与当前解码器配置进行比较, 以检测流配置开关。

[0105] 字段“flags”用作8位位集,可用于向解码器发送额外信息。这可能是关于是否应用交叉渐变或交叉渐变类型(例如,线性、对数)的信息。在下面的示例中,一位用于表示在位流中是否存在预滚有效负载或UUID。

[0106] 在处理任何AU之前,解码器必须知道流配置。在MPEG-4音频中,解码器配置在AudioSpecificConfig元素中输送。解码器配置及UUID是预滚负载的一部分。此外,预滚有效负载包括MPEG-4音频AU(raw_data_block)的依赖于配置的数量。

[0107] MPEG-4中的即时播放帧(IPF)可包括如下所述的正好一个audio_preroll_element()有效负载及ISO/IEC 14496-3中界定的一或多个音频元素流(例如,single_channel_element())。音频基本元素流与当前时间戳相关联。预滚有效负载可在MPEG-4音频的扩展有效负载机制的一者中携带。

[0108] 图1展示即时播放帧(AU_n)1,包括两个预滚帧(AU_{n-1}、AU_{n-2})2、3与对应的流配置4(AudioSpecificConfig)及流标识符5(UUID)。

[0109] 可基于以下内容界定音频预滚元素:

```
audio_preroll_element() {
    flags; 8 uimbsf
    if((flags & 0x01) == 1)
        uuid; 128 uimbsf
    if((flags & 0x02) == 0) return; //不存在有效负载
    asc_size = bs_asc_size; 8 uimbsf
    if(asc_size == 255)
        asc_size += esc; 8 uimbsf
[0110] AudioSpecificConfig(); asc_size * 8
    n_preroll_frames; 8 uimbsf
    for(f = 0; f < n_preroll_frames; ++f) {
        au_size = bs_au_size; 8 uimbsf
        if(au_size == 255)
            au_size += esc; 8 uimbsf
        raw_data_block(); au_size * 8
    }
}
```

[0111]

flags	旗标控制,例如交叉渐变等。旗标&0x02用于指示有效负载是否存在
uuid	通用唯一标识符,也称为MPEG-D中的流ID。用于发信号通知配置更改(流切换)。如果uuid与先前帧不同,那么应解码有效负载
asc_size	AudioSpecificConfig()的大小(字节)。
n_preroll_frames	音频预滚帧所述数个

[0112]	au_size	要遵循的 AU 的大小(字节)
--------	----------------	-----------------

[0113] 在一个示例中,音频预滚元素(例如,audio_preroll_element())是字节对准的,因此可由extension_payload元素(例如,填充元素内部)及数据流元素来传输而无需进一步的字节对准。

[0114] 在一个示例中,IPF如下涉及预滚元素:IPF包含正确解码所需的当前AU及额外AU(即先前帧的数量)。额外AU被封装为预滚元素的一部分,而预滚元素又被封装到raw_data_block()中(经由ID_DSE或ID_FIL)。此类raw_data_block可为IPF。

[0115] IPF可经由各种方法编码。在一个示例中,预滚帧必须是独立可解码的,例如,如果使用SBR,那么必须存在SBR标头。在一个实施例中,audio_preroll_element()可封装在数据流元素中。例如,audio_preroll_element()可基于以下语法封装在数据流元素中:

```

data_stream_element() {
    element_instance_tag; 4 uimsbf
    data_byte_align_flag; 1 uimsbf
    cnt = count; 8 uimsbf
    if (cnt == 255)
[0116]         cnt += esc_count; 8 uimsbf
    if (data_byte_align_flag)
        byte_alignment();
    for (i = 0; i < cnt; i++)
        data_stream_byte[element_instance_tag][i]; 8 uimsbf
}
    
```

[0117]	element_instance_tag	识别数据流元素的数字。如果存在多个 DSE, 必须小心处理。
	cnt	音频 audio_preroll_element()的大小(字节)(如果没有 IPF, 那么为 0)
	data_stream_byte	音频预滚负载, 即 audio_preroll_element()

[0118] 可使用惯例来识别携带audio_preroll_element()的数据流元素(ID_DSE)。在一个实施例中,ID_DSE可位于流中的预定义位置及/或可具有特定实例标签,所述特定实例标签发信号通知有效负载是audio_preroll_element()。

[0119] 在另一实施例中,audio_preroll_element()可封装在填充元素内的extension_payload元素中。例如,audio_preroll_element()可基于以下语法进行封装:

```

[0120] fill_element() {
    
```

```

        cnt = count; 4 uimbsf
        if (cnt == 15)
            cnt += esc_count - 1; 8 uimbsf
[0121]    while (cnt > 0) {
            cnt -= extension_payload(cnt);
        }
    }
}

```

	count	音频预滚元素的大小(字节)(如果没有 IPF, 那么为 0)
[0122]	extension_payload	扩展有效负载元素, 携带 audio_preroll_element()。当前用于传输例如 SBR 数据, 填充字节。扩展类型可为“EXT_DATA_ELEMENT”或新界定的元素类型, 例如“EXT_AUDIO_PRE_ROLL”。

[0123] 如在先前示例中, 可使用惯例来识别 extension_payload 元素的类型, 例如, 如果在有效负载分组中的某个预定义位置传送 EXT_DATA_ELEMENT, 那么有效负载是 audio_preroll_element()。因此, 在一个实施例中, extension_payload 元素可位于流中的预定义位置及/或可具有特定实例标签, 所述特定实例标签表示有效负载是 audio_preroll_element()。

[0124] 在一个示例中, 根据本发明的扩展有效负载可使用新扩展有效负载类型来发信号通知, 例如 extension_type = EXT_AUDIO_PRE_ROLL = 1010b。

```

        extension_payload(cnt) {
            extension_type; 4 uimbsf
            align = 4;
            switch( extension_type ) {
[0125]                case EXT_AUDIO_PRE_ROLL:
                    //总是字节对准, 不修改对准
                    audio_preroll_element();
                    break;
                [...]
            }
        }

```

[0126] 在一个示例中, 携带 audio_preroll_element() 的数据流元素 (ID_DSE) 或 extension_payload 元素 (例如, 在填充元素 (ID_FIL) 内) 可在相同有效负载分组中的任何音频元素之前嵌入位流中。以这种方式嵌入 IPF 有效负载的一些示例包含:

[0127] Mono: <ID_(DSE|FIL)><ID_SCE>...<ID_END>

[0128] 5.1: <ID_(DSE|FIL)><ID_SCE><ID_CPE><ID_CPE><ID_LFE>...<ID_END>

[0129] 本发明的一个方面涉及 IPF 的解码。参照图 2 的示例, 解码过程可包含接收编码音频数据的位流的步骤 S101。编码音频数据的位流可表示音频样本值的序列, 并可包括多个帧, 其中每一帧可包括相关联的编码音频样本值。

[0130] 所述方法可进一步包括确定编码音频数据的位流的帧是否是即时播放帧的步骤

S102。即时播放帧可包括与当前帧相关联的编码音频样本值及额外信息。额外信息可包括在即时播放帧先前的数个帧的编码音频样本值,先前帧的编码音频样本值是使用与当前帧相同的编解码器配置进行编码。对应于预滚帧的先前帧的数量可对应于解码器建立完整信号以便每当即时播放帧被解码时就处于输出与当前帧相关联的有效音频样本值的位置所需的帧的数量。例如,可在解码器的启动或重新启动期间建立完整信号。即时播放帧可为例如解码器启动时的第一帧。

[0131] 且所述方法可包含在确定帧是即时播放帧时初始化解码器的步骤S103。所述初始化可包括在对与当前帧相关联的编码音频样本值进行解码之前对由额外信息包括的编码音频样本值进行解码。

[0132] 现在参考图3的示例,如果确定帧是即时播放帧,且如果即时播放帧的音频样本值已使用不同编解码器配置进行编码,那么音频解码器可从当前编解码器配置切换到不同编解码器配置。可使用当前编解码器配置对即时播放帧进行解码,且如果确定所述帧是即时播放帧且如果已使用当前编解码器配置对即时播放帧的音频样本值进行编码,那么可丢弃额外信息。

[0133] 在一个示例中,解码过程可利用来自MPEG-D USAC中的IPF解码的元素。再次参考图3的示例,解码过程可如下详细地进行。

[0134] ●如果有效负载存在且如果uuid从先前帧更改(在框S104处的决策)

[0135] 或

[0136] ●如果有效负载存在且uuid不存在(在框S105处的决策)

[0137] 1.从audio_preroll_element()读取新的流配置S106,即AudioSpecificConfig()

[0138] 2.例如,使用“NULL”访问单元,刷新解码器状态并将结果存储在缓冲器S107中

[0139] ■将结果存储在缓冲器(交叉渐变缓冲器)

[0140] 3.重新配置(重新初始化)解码器S108

[0141] 4.解码audio_preroll_element S109中的n_preroll_frames并丢弃输出

[0142] 5.解码位流中的下一音频元素(例如SCE/CPE/LFE),并将结果存储在缓冲器(交叉渐变缓冲器B)中

[0143] ■如果由flags发信号通知,那么在交叉渐变缓冲器A与解码器输出之间应用交叉渐变S110,并将结果写入输出PCM缓冲器。交叉渐变的结果构成此帧的合成单元

[0144] ■将解码器输出直接写入PCM缓冲器

[0145] 6.继续下一帧

[0146] ●其它

[0147] 1.跳过音频audio_preroll_element()并解码帧S111

[0148] 2.将结果写入PCM缓冲器S112

[0149] 解码过程可进一步包括:

[0150] ●如果有效负载不存在,那么解码相应帧并继续下一帧

[0151] 在一个示例中,此处理可允许如ISO/IEC 14496-3中所界定的AudioObjectTypes (AOT)的切换,结合连续产生正确的输出样本并在音频输出中不引入静默间隙。在一个示例中,在两者之间切换的AOT可包含AOT 2(AAC)、AOT 5(SBR)、AOT 29(PS)及其它兼容的AOT。

[0152] 现在参考图4的示例,本发明的一个方面涉及用于对IPF进行解码的音频解码器。音频解码器100可包括确定器101。确定器101可经配置以确定编码音频数据的位流的帧是否是包括与当前帧相关联的编码音频样本值及额外信息的即时播放帧。额外信息可包括在即时播放帧先前的数个帧的编码音频样本值。先前帧的经编码音频样本值可使用与当前帧相同的编解码器配置进行编码。对应于预滚帧的先前帧的数量可对应于解码器100建立完整信号以便每当即时播放帧被解码时就处于输出与当前帧相关联的有效音频样本值的位置所需的帧的数量。例如,可在解码器100的启动或重新启动期间建立完整信号。例如,即时播放帧可为解码器100启动时的第一帧。

[0153] 且音频解码器100可包括初始化器102。如果确定器101确定帧是即时播放帧,那么初始化器102可经配置以初始化解码器100。初始化解码器100可包括在对与当前帧相关联的编码音频样本值进行解码之前对由额外信息包括的编码音频样本值进行解码。如果确定器101确定帧是即时播放帧且如果当前帧的音频样本值已使用不同编解码器配置进行编码,那么初始化器102还可经配置以将音频解码器100从当前编解码器配置切换到不同编解码器配置。且如果确定器101确定所述帧是即时播放帧且如果当前帧的音频样本值已使用当前编解码器配置进行编码,那么解码器100可经配置以使用当前编解码器配置对当前帧进行解码并丢弃额外信息。

[0154] 现在参考图5的示例,本发明的一个方面涉及用于生成具有即时播放帧(随机接入点)的编码音频数据的位流的音频编码器,其中编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0155] 音频编码器200可包括核心编码器202,核心编码器202经配置以使用预定义编解码器配置对与多个帧相关联的未压缩音频样本值进行编码。例如,使用预定义编解码器配置可包含使用预定取样率。在一个示例中,核心编码器202可将未压缩音频样本编码为符合MPEG-4音频标准的解码。

[0156] 音频编码器200可进一步包括缓冲器203,缓冲器203经配置以存储使用预定义编解码器配置(如虚线所指示)编码的多个帧的当前帧的数个先前帧的编码音频样本值。

[0157] 当编码例如帧N时,相关的先前帧N-1、N-2……可总是被缓冲/存储。当被指示在帧N中写入IPF时(例如,每2秒以便允许动态切换),然后可将存储的相关先前帧N-1、N-2……取出并封装到当前帧N中。

[0158] 且音频编码器200可包括嵌入器204,嵌入器204经配置以在多个帧的当前帧中写入即时播放帧,其中即时播放帧可包括与所述当前帧相关联的编码音频样本值及与所述当前帧的数个先前帧的编码音频样本值相对应的额外信息。

[0159] 虽然在图5的示例中,嵌入器204经界定为音频编码器200的一部分,但注意,替代地或另外,嵌入器204还可单独实施,以在编码音频数据的位流的任何当前帧中写入即时播放帧,或通过从即时播放帧中移除额外信息,将编码音频数据的位流中的即时播放帧转换为“正常”帧。在这种情况下,嵌入器204可为编码器链的一部分,但不一定必须是。

[0160] 在一个实施例中,嵌入器204可经进一步配置以在额外信息中包含关于预定义编解码器配置的信息。在这种情况下,额外信息可向解码器提供关于预定义编解码器配置的信息。

[0161] 在一个实施例中,嵌入器204可经进一步配置以在即时播放帧中包含额外信息。以

此方式,额外信息可在位流中传输到解码器。

[0162] 在一个实施例中,所生成的编码音频数据的位流可为MPEG-4音频位流。

[0163] 在一个实施例中,嵌入器204可经进一步配置以经由MPEG-4音频位流扩展机制将额外信息嵌入在位流中(例如,用于输送),MPEG-4音频位流扩展机制可为数据流元素(ID_DSE)或extension_payload元素。

[0164] 在一个实施例中,嵌入器204可经进一步配置以将数据流元素(ID_DSE)或extension_payload元素定位在MPEG-4音频位流中的预定义位置及/及分配特定实例标签,所述特定实例标签发信号通知数据流元素(ID_DSE)或extension_payload元素的有效负载是额外信息。预定义位置可对应于MPEG-4音频位流中的第一位置,即总是在帧中的第一位置,因为它可携带解码帧n所需的解码器配置(图1,raw_data_block()[n])。解码器因此可假设,如果帧中的第一元素是ID_DSE或extension_payload元素(可含在ID_FIL元素内),那么此元素携带预滚数据(前帧、预滚帧)。

[0165] 在一个实施例中,嵌入器204可经进一步配置以在额外信息中包含唯一标识符。任选地,唯一标识符可发信号通知预定义编解码器配置。然后,解码器可使用预定义编解码器配置对如上所述的帧n进行解码。基于唯一标识符,解码器能够识别位流中的额外信息并相应地解析位流。

[0166] 在一个实施例中,音频编码器200可经进一步配置以使包括在额外信息中的数个帧的最早帧相对于最早帧之前的任何帧对不进行时差编码或熵编码,且音频编码器200可经进一步配置以使即时播放帧相对于即时播放帧之前的数个帧的最早帧之前的任何帧或相对于即时播放帧之前的任何帧不进行时差编码或熵编码。

[0167] 现在参考图6的示例,本发明的一个方面涉及由音频编码器生成具有即时播放帧(随机接入点)的编码音频数据的位流的方法,其中编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0168] 所述方法可包括由核心编码器使用预定义编解码器配置对与多个帧相关联的未压缩音频样本值进行编码的步骤S201。例如,使用预定义编解码器配置可包含使用预定取样率。所述方法可进一步包括由缓冲器存储使用预定义编解码器配置编码的多个帧的当前帧的数个先前帧的编码音频样本值的步骤S202。

[0169] 且所述方法可包括由嵌入器在多个帧的当前帧中写入即时播放帧的步骤S203,其中即时播放帧包括与所述当前帧相关联的编码音频样本值及与所述当前帧的数个先前帧的编码音频样本值相对应的额外信息。

[0170] 在一个实施例中,额外信息可进一步包括关于预定义编解码器配置的信息。如上所述,解码器可在解码过程中使用预定义编解码器配置。

[0171] 在一个实施例中,即时播放帧可进一步包括额外信息。以此方式,额外信息可在位流中传输。

[0172] 在一个实施例中,所生成的编码音频数据的位流可为MPEG-4音频位流。

[0173] 在一个实施例中,额外信息可由嵌入器经由MPEG-4音频位流扩展机制嵌入在位流中(例如,用于输送),MPEG-4音频位流扩展机制是数据流元素(ID_DSE)或extension_payload元素。extension_payload元素可例如含在MPEG-4音频位流语法的不同位置。在一个实施例中,extension_payload元素可由嵌入器嵌入(例如,用于输送)在填充元素(ID_

FIL) 内。

[0174] 如上所详细描述,在一个实施例中,数据流元素(ID_DSE)或extension_payload元素可由嵌入器定位在MPEG-4音频位流中的预定义位置处及/或可被分配特定实例标签,所述特定实例标签发信号通知数据流元素(ID_DSE)或extension_payload元素的有效负载是额外信息。预定义位置可总是帧中的第一位置,因为它可携带解码当前帧可能需要的解码器配置。

[0175] 在一个实施例中,额外信息可进一步包括唯一标识符。任选地,所述唯一标识符可发信号通知预定义编解码器配置。

[0176] 在一个实施例中,可由音频编码器使包括在额外信息中的数个帧的最早帧相对于最早帧先前的任何帧不进行时差编码或熵编码,且可由音频编码器使即时播放帧相对于即时播放帧先前的数个帧的最早帧先前的任何帧或相对于即时播放帧先前的任何帧不进行时差编码或熵编码。

[0177] 现在参考图7的示例,本发明的一个方面涉及一种系统,所述系统包括两个或多于两个音频编码器,以用于生成编码音频数据的多个位流,每一位流具有即时播放帧(时间对准的随机接入点),其中编码音频数据的每一位流表示音频样本值的序列并包括多个帧,且其中每一帧包括相关联的编码音频样本值。虽然系统中音频编码器的数量不受限制,但在图7的示例中,说明包括两个音频编码器的系统。系统可在相同的内容上并行运行两个音频编码器,以相同的节奏插入即时播放帧,但具有不同配置(例如位率)。每一编码链可输出自己的位流。以这种方式创建的两个位流可例如存储在web服务器上。客户端可开始回放流A(例如,以高位率)。在某个时间点,客户端可决定切换到较低位率的流B,因此可请求相同的内容,但位率不同。当流B的第一区段到达解码器时,此类区段可总是以IPF开始(这可例如经由根据MPEG-DASH的清单文件来发信号通知),这使得解码器能够从一开始就输出正确的音频。

[0178] 再次参考图7的示例,并行地说明两个音频编码器,每一音频编码器包括核心编码器202(编码器核心#1、编码器核心#2)、缓冲器(未展示)及嵌入器204(IPF插入)。

[0179] 在一个实施例中,对于每一核心编码器202,预定取样率可相同。否则,在解码器侧可能需要重取样及额外延迟处理。然而,核心编码器202可经配置成以不同帧速率运行(例如AAC-LC 1024;HE-AAC 2048)。进一步来说,核心编码器配置可能需要不同数量的预滚帧p。可能需要帧n-p及n都是独立可解码的,即可能不必依赖于来自先前帧的信息(对于HE-AAC,它们可能含有SBR标头)。在核心编码之后,IPFs的解码时间可跨不同流对准。

[0180] 在一个实施例中,系统可进一步包括延迟对准单元201(延迟#1,延迟#2),用于延迟对准多个位流。对于解码器处的无缝切换,可能必须对准即时播放帧(IPF)的解码时间。延迟对准级可延迟输入PCM样本(未压缩音频样本值、输入音频样本值)以补偿不同编码器/解码器延迟。

[0181] 现在参考图8的示例,本发明的一个方面涉及用于在编码音频数据的位流中生成即时播放帧(随机接入点)或用于从编码音频数据的位流中移除即时播放帧(随机接入点)的设备,其中编码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0182] 设备300可包括接收器301,接收器301经配置以接收编码音频数据的位流,其中编

码音频数据的位流表示音频样本值的序列并包括多个帧,其中每一帧包括相关联的编码音频样本值。

[0183] 且设备300可包括嵌入器302,嵌入器302经配置以在多个帧的当前帧中写入即时播放帧,其中即时播放帧包括与所述当前帧相关联的编码音频样本值及与所述当前帧的数个先前帧的编码音频样本值相对应的额外信息。在生成IPF时,嵌入器302可按照上面结合图5到7描述的原理操作。这可以说此操作对应于将“正常”(非IPF)帧转换为IPF。

[0184] IPF插入可能会考虑实际的复制及封装。通过分割核心编码及IPF插入,与上述设备一样,可在没有IPF的情况下存储位流,且如果需要,仅在分发之前嵌入IPF。如果经配置为无缝切换,每一IPF可能必须包含流配置ASC。

[0185] 在一个实施例中,设备300可进一步包括缓冲器,所述缓冲器经配置以存储多个帧中当前帧的数个先前帧的编码音频样本值。

[0186] 在一个实施例中,嵌入器302可经进一步配置以从即时播放帧中移除与所述当前帧的数个先前帧的编码音频样本值相对应的额外信息。这可以说对应于将IPF转换为“正常”帧。例如,音频编码器可仅生成ipf,其随后可由嵌入器根据相应信道/服务可能具有的约束来移除。

[0187] 现在参考图9的示例,本发明的方面可包含非暂时性数字存储媒体,所述非暂时性数字存储媒体具有存储在其上的计算机程序,以在所述计算机程序由计算机或处理器运行时执行本文所述的方法。图9说明具有可运行所述计算机程序的处理器401的装置400。或者,装置400可表示相应的计算机。

[0188] 本文所描述的方法及系统可经实施为软件、固件及/或硬件。某些组件可例如经实施为在数字信号处理器或微处理器上运行的软件。其它组件可例如经实施为硬件及/或特定于应用的集成电路。在所描述的方法及系统中遇到的信号可存储在例如随机存取存储器或光存储媒体的媒体上。它们可经由网络传输,例如无线网络、卫星网络、无线网络或有线网络,例如因特网。使用本文所描述的方法、设备及系统的典型装置是用于存储及/或呈现音频信号的便携式电子装置或其它消费型装备。

[0189] 应注意,说明书及附图仅说明所提议的方法、系统及设备的原理。所属领域的技术人员将能够实施各种布置,这些布置虽然在本文中并没有明确描述或展示,但体现本发明的原理,且包含在本发明的精神及范围内。此外,本文档中概述的所有示例及实施例主要希望明确地仅用于解释目的,以帮助读者理解所提出方法的原理。此外,本文中提供本发明的原理、方面及实施例的所有陈述及其特定示例希望包含其等效者。

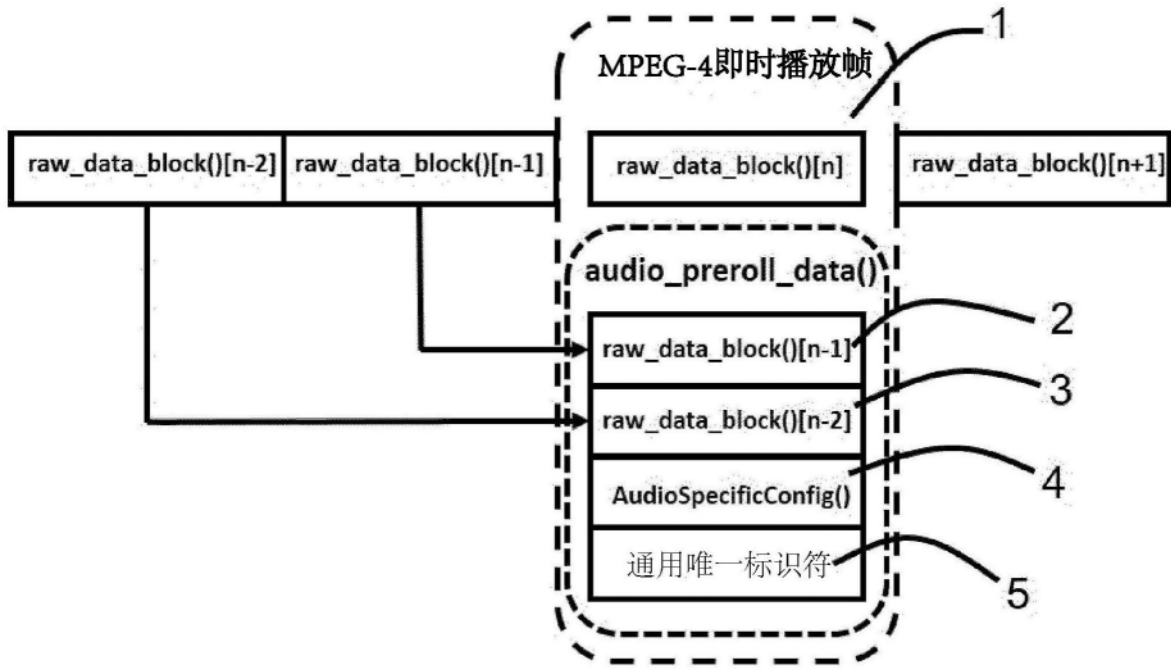


图1



图2

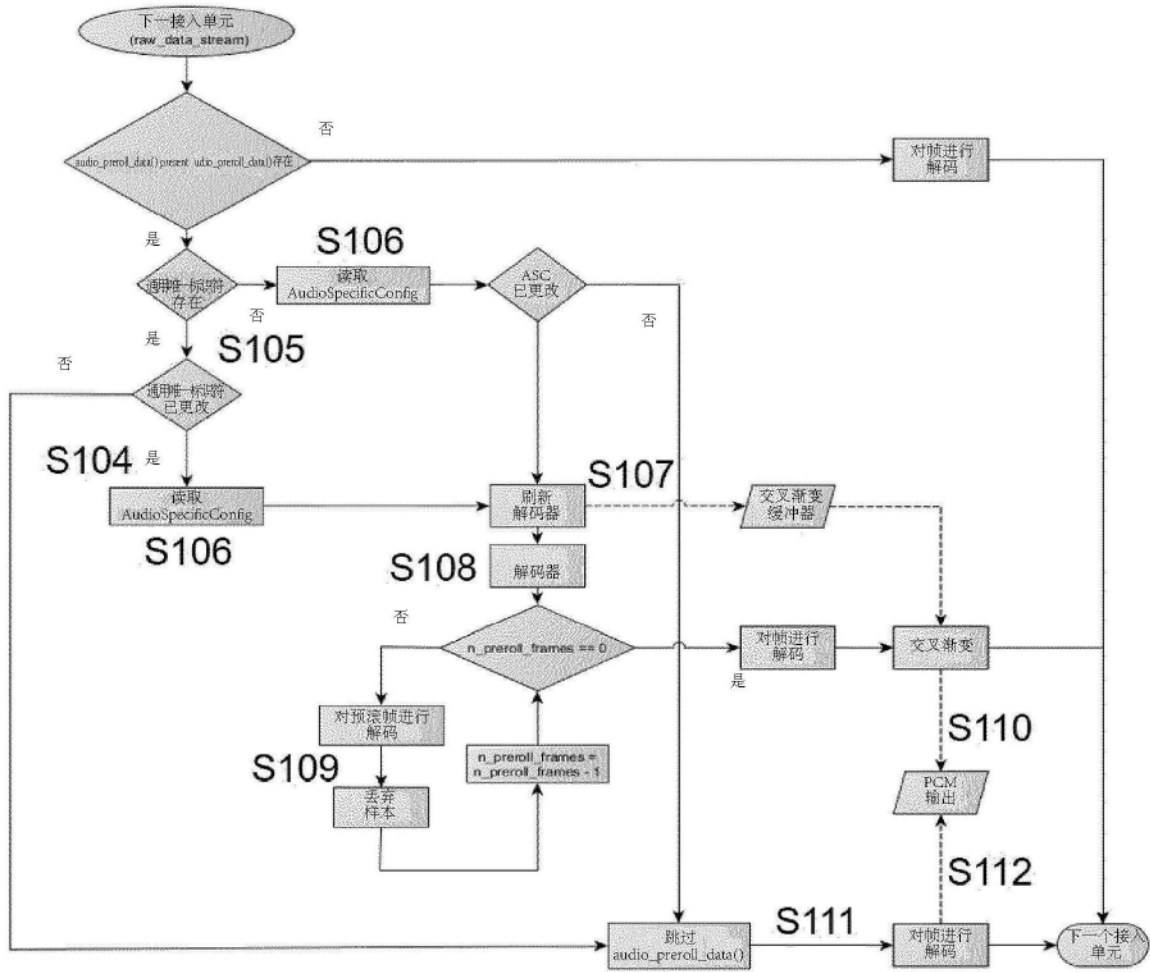


图3

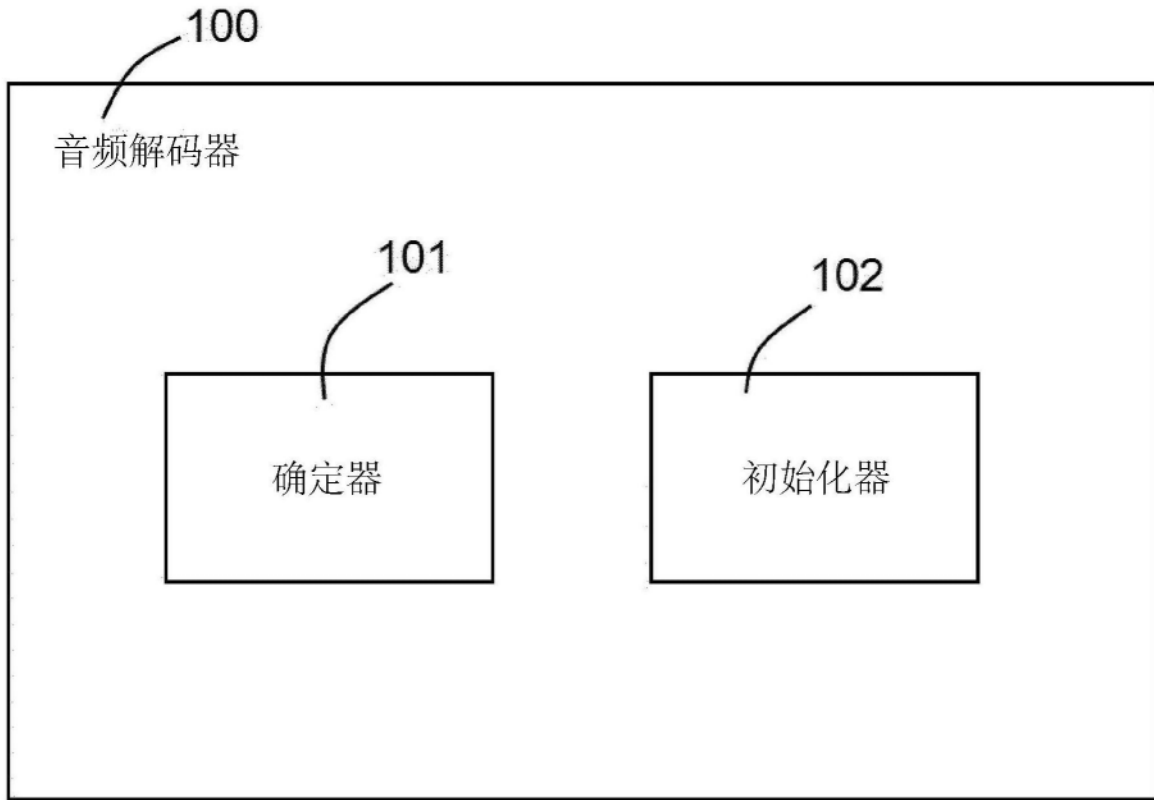


图4

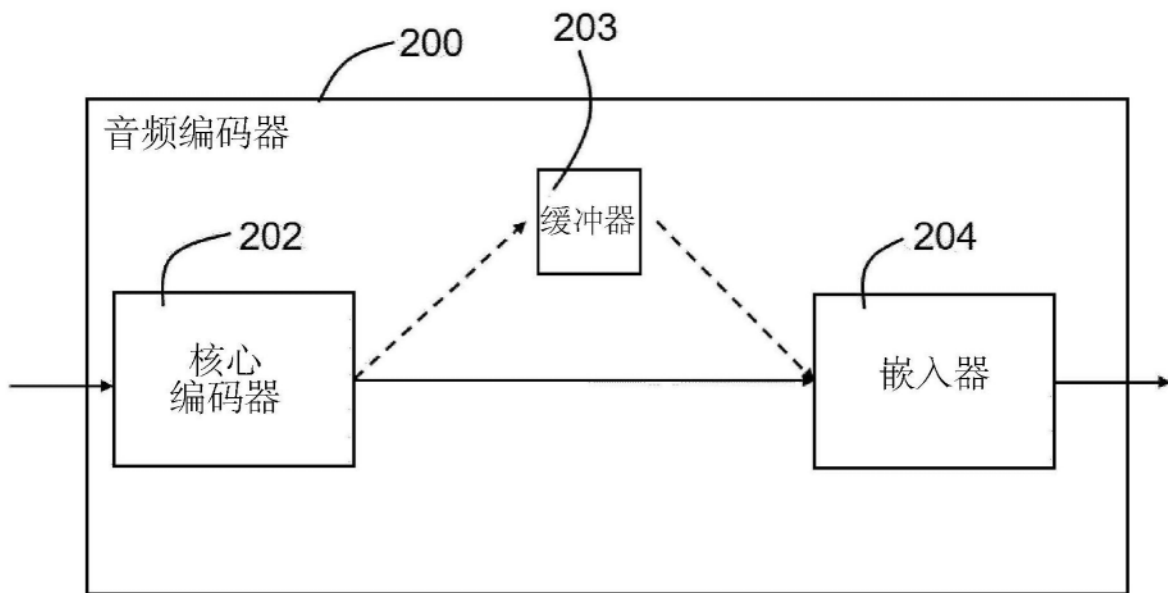


图5



图6

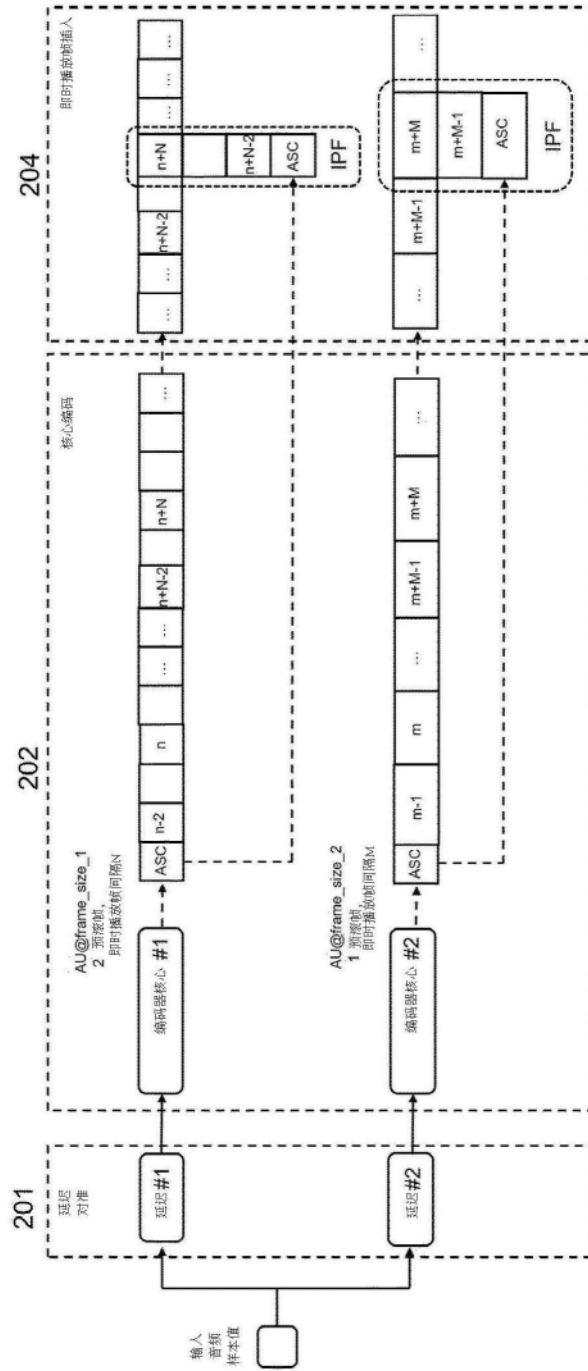


图7

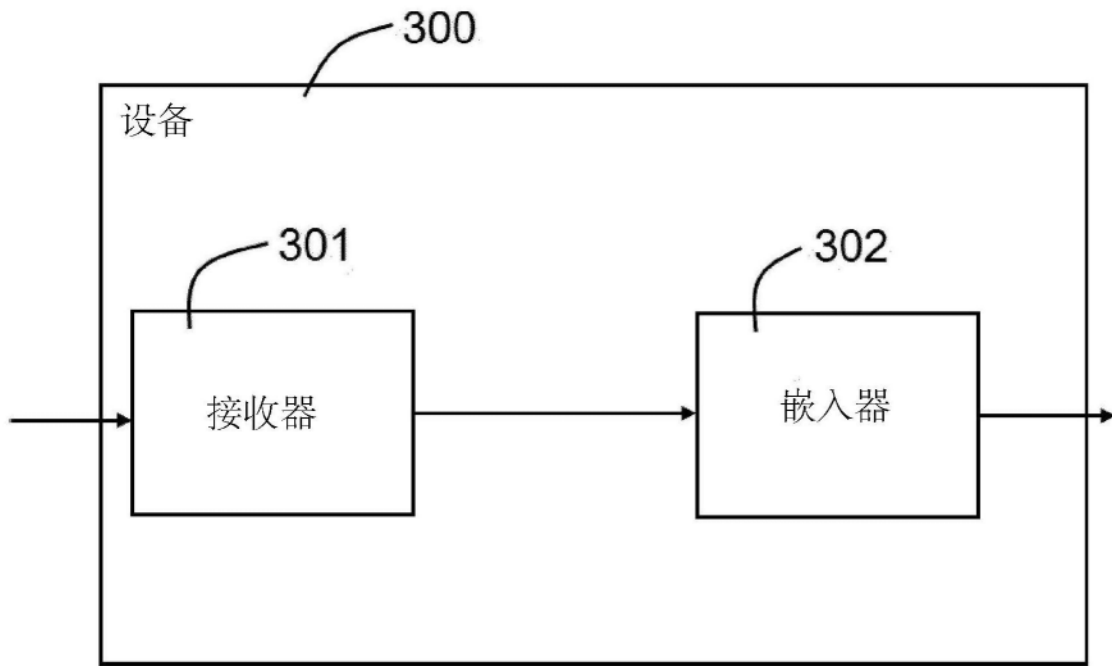


图8

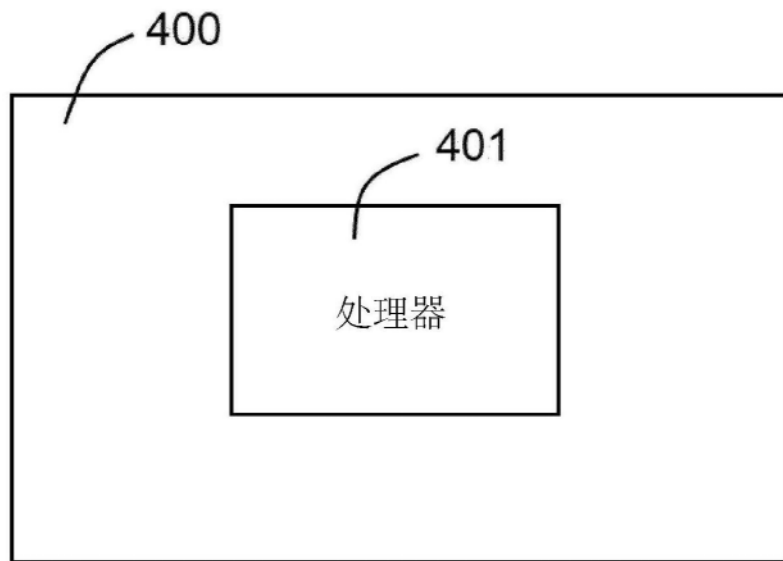


图9