



(12) 发明专利申请

(10) 申请公布号 CN 105453581 A

(43) 申请公布日 2016. 03. 30

(21) 申请号 201480024731. 8

H04N 21/854(2006. 01)

(22) 申请日 2014. 04. 08

H04N 21/43(2006. 01)

(30) 优先权数据

G11B 27/10(2006. 01)

61/817, 471 2013. 04. 30 US

(85) PCT国际申请进入国家阶段日

2015. 10. 29

(86) PCT国际申请的申请数据

PCT/US2014/033349 2014. 04. 08

(87) PCT国际申请的公布数据

W02014/179003 EN 2014. 11. 06

(71) 申请人 杜比实验室特许公司

地址 美国加利福尼亚

(72) 发明人 H·A·德圣萨尔维 G·J·麦加里

(74) 专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 欧阳帆

(51) Int. Cl.

H04N 21/81(2006. 01)

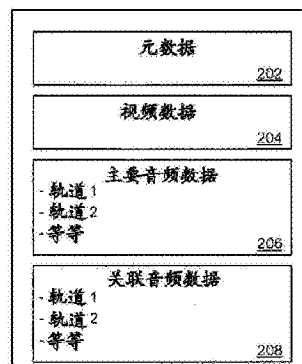
权利要求书3页 说明书12页 附图4页

(54) 发明名称

从单个容器输出多种语言的音频和关联音频的系统和方法

(57) 摘要

一种存储和输出关联音频的方法。关联音频和主要音频被存储在单个多媒体容器中并且从该单个多媒体容器输出。以这种方式,当改变主要语言时,在最少音频假象的情况下自动地改变关联音频。



容器 200 (逻辑图)

1. 一种输出音频的方法,包括:

存储多媒体容器,其中所述多媒体容器包括多个主要音频数据集以及多个关联音频数据集,并且在所述多媒体容器中所述多个主要音频数据集的数据元素与所述多个关联音频数据集的数据元素交错;

接收第一选择信息,其中所述第一选择信息对应于多种语言中的第一语言;以及

从所述多媒体容器输出第一音频数据流,其中所述第一音频数据流对应于所述多个主要音频数据集中的具有第一语言的第一主要音频数据集,

其中所述多个主要音频数据集和多个关联音频数据集具有多种语言。

2. 根据权利要求1所述的方法,还包括:

接收第二选择信息,其中第二选择信息与所述多个关联音频数据集有关;

从所述多媒体容器输出第二音频数据流,其中第二音频数据流对应于所述多个关联音频数据集中的具有第一语言的第一关联音频数据集;

接收第三选择信息,其中所述第三选择信息对应于所述多种语言中的第二所选语言;

从所述多媒体容器输出第三音频数据流代替第一音频数据流,其中所述第三音频数据流对应于所述多个主要音频数据集中的具有第二所选语言的第二所选主要音频数据集;以及

从所述多媒体容器自动地输出第四音频数据流代替第二音频数据流,其中所述第四音频数据流对应于所述多个关联音频数据集中的具有第二所选语言的第二所选关联数据集。

3. 根据权利要求2所述的方法,其中输出第二音频数据流包括:

与输出所述第一音频数据流同时地从所述多媒体容器输出第二音频数据流。

4. 根据权利要求2或3所述的方法,其中自动地输出第四音频数据流包括:

与输出所述第三音频数据流同时地从所述多媒体容器自动地输出第四音频数据流。

5. 根据权利要求1-4中任何一个所述的方法,还包括:

同时读取所述第一主要音频数据集的数据元素和所述第一关联音频数据集的数据元素。

6. 根据权利要求2-5中任何一个所述的方法,还包括:

同时读取第二所选主要音频数据集的数据元素和第二所选关联音频数据集的数据元素。

7. 根据权利要求2-6中任何一个所述的方法,其中平滑的切换源于由于所述多个主要音频数据集的数据元素与所述多个关联音频数据集的数据元素交错而自动地输出第四音频数据流来代替第二音频数据流。

8. 根据权利要求1-7中任何一个所述的方法,其中所述多媒体容器还包括使所述多个主要音频数据集中的一个是所述多个关联音频数据集中的一个是同步的同步信息。

9. 根据权利要求2-8中任何一个所述的方法,其中所述多媒体容器还包括使所述多个主要音频数据集中的一个是所述多个关联音频数据集中的一个是同步的同步信息,其中输出第二音频数据流包括:

根据所述同步信息,与输出所述第一音频数据流同时地从所述多媒体容器输出所述第二音频数据流。

10. 根据权利要求2-9中任何一个所述的方法,其中所述多媒体容器还包括使所述多个

主要音频数据集中的—个与所述多个关联音频数据集中的—个同步的同步信息,其中自动地输出第四音频数据流包括:

根据所述同步信息,与输出所述第三音频数据流同时地自动地从所述多媒体容器输出所述第四音频数据流。

11.根据权利要求1-10中任何—个所述的方法,其中所述多媒体容器还包括使所述多个主要音频数据集中的—个与所述多个关联音频数据集中的—个同步的同步信息,还包括:

根据所述同步信息,同时读取所述第—主要音频数据集的数据元素和所述第—关联音频数据集的数据元素。

12.根据权利要求2-11中任何—个所述的方法,其中所述多媒体容器还包括使所述多个主要音频数据集中的—个与所述多个关联音频数据集中的—个同步的同步信息,还包括:

根据所述同步信息,同时读取第二所选主要音频数据集的数据元素和第二所选关联音频数据集的数据元素。

13.根据权利要求1-12中任何—个所述的方法,其中所述多媒体容器还包括使所述多个主要音频数据集中的—个与所述多个关联音频数据集中的—个同步的同步信息。

14.根据权利要求1-13中任何—个所述的方法,其中所述多个关联音频数据集中的每个与所述多个主要音频数据集中的对应—个相关联。

15.根据权利要求1-14中任何—个所述的方法,其中所述多个关联音频数据集中的—个与所述多个主要音频数据集中的—个相关联,并且其中所述多个主要音频数据集中的所述—个的声道的数量不大于所述多个主要音频数据集中的所述—个的声道的数量。

16.根据权利要求1-15中任何—个所述的方法,其中所述多媒体容器包括与所述多个主要音频数据集和所述多个关联音频数据集有关的元数据。

17.根据权利要求1-16中任何—个所述的方法,其中所述多媒体容器包括与所述多个主要音频数据集和所述多个关联音频数据集有关的元数据,其中所述元数据包括使所述多个主要音频数据集中的—个与所述多个关联音频数据集中的—个同步的同步信息。

18.根据权利要求2-17中任何—个所述的方法,其中所述多媒体容器包括与所述多个主要音频数据集和所述多个关联音频数据集有关的元数据,还包括:

根据所述元数据,与输出所述第—音频数据流同时地从所述多媒体容器输出所述第二音频数据流。

19.根据权利要求1-18中任何—个所述的方法,其中所述多媒体容器包括与所述多个主要音频数据集和所述多个关联音频数据集有关的元数据,还包括:

根据所述元数据,在所述多媒体容器中使所述多个主要音频数据集的数据元素与所述多个关联音频数据集的数据元素交错。

20.根据权利要求1-19中任何—个所述的方法,其中所述多个关联音频数据集包括多个音频描述数据集。

21.根据权利要求2-20中任何—个所述的方法,其中所述多个关联音频数据集包括导演的评论数据集,还包括:

接收第四选择信息,其中所述第四选择信息对应于导演的评论数据集;

从所述多媒体容器输出第五音频数据流代替所述第四音频数据流,其中所述第五音频数据流对应于所述导演的评论数据集。

22.根据权利要求1-21中任何一个所述的方法,还包括:

渐进地下载所述多媒体容器,其中输出所述第一音频数据流包括在已经渐进地下载所述多媒体容器的规定量之后输出所述第一音频数据流。

23.根据权利要求1-22中任何一个所述的方法,还包括:

渐进地下载所述多媒体容器,其中输出所述第一音频数据流包括在已经渐进地下载所述多媒体容器的规定量之后输出所述第一音频数据流,并且其中在正在渐进地下载所述多媒体容器时继续输出所述第一音频数据流。

24.根据权利要求1-23中任何一个所述的方法,还包括:

渐进地下载所述多媒体容器,其中输出所述第一音频数据流包括在已经渐进地下载所述多媒体容器的规定量之后输出所述第一音频数据流,并且其中在已经渐进地下载了所述多媒体容器之后继续输出所述第一音频数据流。

25.一种用于处理多媒体容器的装置,所述装置包括:

处理器;

存储设备,耦接到所述处理器而且被配置为存储所述多媒体容器,其中所述多媒体容器包括多个主要音频数据集和多个关联音频数据集,其中所述多个主要音频数据集和所述多个关联音频数据集具有多种语言,并且在所述多媒体容器中所述多个主要音频数据集的数据元素与所述多个关联音频数据集的数据元素交错;以及

音频解码器,从所述多媒体容器输出第一音频数据流和第二音频数据流,其中所述第一音频数据流对应于所述多个主要音频数据集中的具有第一语言的主要音频数据集,并且所述第二音频数据流对应于所述多个关联音频数据集中的具有第一语言的关联音频数据集。

26.一种计算机程序,存储在非暂时的计算机可读介质中,用于控制用于输出多种语言的音频和关联音频的包括处理器和存储设备的装置执行包括如下的处理:

存储多媒体容器,其中所述多媒体容器包括多个主要音频数据集以及多个关联音频数据集,并且在所述多媒体容器中所述多个主要音频数据集的数据元素与所述多个关联音频数据集的数据元素交错;

接收第一选择信息,其中所述第一选择信息对应于多种语言中的第一语言;以及

从所述多媒体容器输出第一音频数据流,其中所述第一音频数据流对应于所述多个主要音频数据集中的具有第一语言的第一主要音频数据集,

其中所述多个主要音频数据集和所述多个关联音频数据集具有多种语言。

## 从单个容器输出多种语言的音频和关联音频的系统和方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2013年4月30日提交的美国临时专利申请No.61/817,471的优先权,其整体通过参考被并入于此。

### 技术领域

[0003] 本发明涉及视听回放,特别应用于从计算机网络下载的内容。

### 背景技术

[0004] 除非在此另有指明,本节中描述的方法并非对于本申请中的权利要求的现有技术并且没有由于包括在本节中而被承认为现有技术。

[0005] 消费者越来越多通过下载或流式传输(streaming)来在线获得内容,而不是购买物理介质。提供内容的示例在线服务包括Netflix Watch Instantly™服务、Apple iTunes™服务、以及Amazon Instant Video™服务。虽然从最终用户视角来看流式传输和下载可能是类似的,但是技术差别是,下载的内容被存储在本地并且随后从本地存储器回放。从最终用户视角来看,其中在下载仍然在进行中的同时可以开始内容的回放的渐进式(progressive)下载进一步对与流式传输的相似度有贡献。

### 发明内容

[0006] 由本发明实施例解决的问题涉及与内容有关的关联音频。关联音频可以是用于视觉受损的用户的音频描述轨道、导演评论轨道等。关联音频区别于主要音频,该主要音频是内容的主要音轨,包括对话轨道、音乐音轨和效果音轨。例如,AC3数字音频压缩标准限定两个主要的和六个关联的服务。主要音频服务是完整的主要(CM)以及音乐和效果(ME);关联服务是视觉受损(VI)、听力受损(HI)、对话(D)、评论(C)、紧急(E)以及画外音(VO)。

[0007] 对于诸如数字视频盘(DVD)的物理介质,关联内容(诸如音频描述)可以被包括在物理介质上。然而,这种关联内容可以在与主要音频分离的容器中,以及在其中关联内容可以已经与较低质量(例如,立体声)版本的主要音频预先混合的容器中。

[0008] 可替代地,对于DVD,关联内容可以不被设置在物理介质本身上,而是可以经由流式传输或下载在线获得。因此,当用户切换到另一个语言轨道时,DVD播放器设备必须协调用于如下两个组件的新设定:用于来自DVD的新的语言轨道的DVD读取器组件,以及用于对应的关联内容的网络组件(对于流式传输内容)或硬盘驱动组件(对于下载的内容)。除物理组件自身的协调之外,协调还必须还考虑切换中涉及的各个软件层(设备驱动器、操作系统、DVD播放器设备控制程序等)。如果这种协调是不完善的,可能得到诸如爆裂声(pops)或毛刺(glitches)之类的听觉假象,对用户体验有负面影响。

[0009] 当一个服务器提供主要音频而另一个服务器提供关联音频时,对于在线获得的内容,情形是类似的。对于流式传输内容,客户播放器设备必须将新的流式传输设定传递给服务器设备,其可能以起始于不同时间的它们的新的流响应。对于下载的内容,客户播放器设

备必须协调用于本地存储设备的新的读取设定,其在本地存储设备从新的位置读取时可能引入听觉假象。

[0010] 最终,可能有用户界面缺陷,其要求与改变用于主要内容的设定相分离地改变关联内容的设定。例如,用户将主要音频从英语切换到法语,但是关联内容继续是英语的,直到它被分离地改变。另一个用户界面缺陷涉及下载内容的管理。当与主要内容分离地提供关联内容时,用户必须管理从它们的分离的来源获得主要内容和关联内容。

[0011] 响应于以上提到的缺点,本发明的一个实施例涉及播放来自单个媒体容器文件的主要内容和关联内容。单个容器有助于确保关联内容与主要内容的同步以及在添加关联内容的回放时的无缝转换。一个实施例还在切换主要内容(例如,以另一种语言输出)时自动地切换关联内容。通过使用单个媒体容器文件,与许多现有的系统相比,媒体播放器设备的听觉假象的发生减少。

[0012] 根据一个实施例,一种方法输出多种语言的音频和关联音频。该方法包括存储多媒体容器。多媒体容器包括多个主要音频数据集和多个关联音频数据集。多个主要音频数据集和多个关联音频数据集具有多种语言。该方法还包括接收第一选择信息。第一选择信息对应于该多种语言中的第一所选语言。该方法还包括从多媒体容器输出第一音频数据流。第一音频数据流对应于该多个主要音频数据集中的具有第一所选语言的第一所选主要音频数据集。该方法还包括接收第二选择信息。第二选择信息与多个关联音频数据集有关。该方法还包括从多媒体容器输出第二音频数据流。第二音频数据流对应于该多个关联音频数据集中的具有第一所选语言的第一所选关联音频数据集。

[0013] 该方法还包括接收第三选择信息。第三选择信息对应于该多种语言中的第二所选语言。该方法还包括从多媒体容器输出第三音频数据流代替第一音频数据流。第三音频数据流对应于该多个主要音频数据集中的具有第二所选语言的第二所选主要音频数据集。该方法还包括从多媒体容器自动地输出第四音频数据流代替第二音频数据流。第四音频数据流对应于该多个关联音频数据集中的具有第二所选语言的第二所选关联数据集。

[0014] 一种装置输出多种语言的音频和关联音频。该装置包括处理器和存储设备。存储设备被配置为存储多媒体容器。处理器被配置为控制该装置执行与上面描述的方法对应的处理。

[0015] 一种计算机程序控制用于输出多种语言的音频和关联音频的装置(包括处理器和存储设备)。计算机程序根据上面描述的方法控制该装置。计算机程序被存储在非暂时的计算机可读介质中。

[0016] 以下详细描述和附图提供对本发明的性质和优点的进一步理解。

## 附图说明

[0017] 图1是示出媒体播放器设备和有关的组件的图。

[0018] 图2是示出由媒体播放器设备(参见图1)用来存储内容的容器的逻辑布置的图。

[0019] 图3A是示出容器的物理布置的图。

[0020] 图3B示出对于图3A中示出的视频数据、主要音频数据和关联音频数据的更多细节。

[0021] 图4是输出多种语言的音频和关联音频的方法的流程图。

[0022] 图5是对于媒体播放器设备(参见图1)的控制程序的框图。

### 具体实施方式

[0023] 在此描述用于媒体回放的技术。在下面的描述中,出于说明的目的,阐述了许多的示例和具体细节以便提供对本发明的彻底的了解。然而,本领域技术人员将明白,如由权利要求所限定的本发明可以单独或者与下面描述的其它特征结合地包括这些示例中的特征中的一些或全部,以及可以进一步包括在此描述的概念和特征的等同物和变型。

[0024] 在下面的描述中,各种方法、处理和过程被详述。虽然可以以某一顺序描述特定步骤,但是这种顺序主要为了方便和清楚起见。特定步骤可以被重复多于一次,可以发生在其它步骤之前或之后(即使以另一种顺序描述了那些步骤),以及可以与其它步骤并行地发生。只有当第一步骤必须在第二步骤开始之前完成时,才要求第二步骤跟随第一步骤。当从上下文不清楚时这种情形将被具体地指出。

[0025] 在本文档中,术语“和”、“或”和“和/或”被使用。这种术语要被读作具有相同的意思;也就是说,包括地。例如,“A和B”可以意指至少以下:“A和B两者”,“仅A”,“仅B”,“至少A和B两者”。作为另一个示例,“A或B”可以意指至少以下:“仅A”,“仅B”,“A和B两者”,“至少A和B两者”。当意图“互斥或”时,这将被具体地提出(例如,“或者A或者B”,“A和B中的至多一个”)。

### [0026] 概述和一般说明

[0027] 图1是示出媒体播放器设备100和有关的组件的图,包括计算机网络102(例如,因特网)、音频/视频接收器(AVR)104、电视机106(例如,高清1080p电视机)以及一个或多个扬声器108。图1的组件可以被部署在用户的家庭中作为家庭影院系统。

[0028] 媒体播放器设备100一般如下与其它组件交互。媒体播放器设备100访问计算机网络102以用于获得内容或与内容有关的其它信息。媒体播放器设备100将内容作为视频和音频信号152输出到AVR 104。媒体播放器设备100与AVR 104之间的连接可以是高清晰度媒体接口(HDMI)连接。信号152中的音频数据可以具有脉冲编码调制(PCM)格式。AVR 104在各种内容源(诸如媒体播放器设备100、DVD播放器设备、蓝光盘设备、紧致盘(CD)播放器设备、无线电接收机设备等)与各种输出设备(诸如电视机106和扬声器108)之间作为接口。AVR 104例如经由HDMI连接将视频流154发送到电视机106。AVR 104将音频流156发送到扬声器108,以及执行用于扬声器108的一般的放大和驱动功能。

[0029] 音频流156的组成将基于扬声器108的数量和配置而改变。例如,对于单个扬声器,音频流156可以是单个单音频声道。对于两个扬声器,音频流156可以是两个立体声音频声道。AVR 104可以将具有环绕声5.1格式的六个声道发送到六个扬声器(例如,左、中心、右、右环绕、左环绕以及亚低音扬声器(subwoofer)),或者将具有环绕声7.1格式的八个声道发送到八个扬声器。

[0030] 媒体播放器设备100包括处理器、内存、网络连接以及存储系统。通常,存储系统存储媒体播放器设备100经由网络连接(例如从网络102)获得的内容。媒体播放器设备100然后根据媒体播放器控制程序控制的那样输出内容。控制程序可以以软件、固件、硬件或其组合实现。用户可以经由用户界面而与控制程序交互。来自苹果公司的Mac Mini™计算机可以被用来实现媒体播放器设备100。控制程序可以被实现为在OS X操作系统(苹果公司)上

运行的应用程序。

[0031] 可替代地,通用计算机可以实现媒体播放器设备100;用于计算机的监视器可以用作电视机106并且计算机扬声器可以用作扬声器108。类似地,诸如移动电话或平板计算机之类的移动设备可以实现媒体播放器设备100;屏幕可以用作电视机106并且内置的扬声器或头戴耳机(headphones)可以用作扬声器108并且输出虚拟化的环绕声信号。

[0032] 媒体播放器设备100可以与各种服务(经由网络102)交互以便获得内容,所述各种服务包括Netflix Watch Instantly™服务、Apple iTunes™服务以及Amazon Instant Video™服务。

[0033] 图2是示出由媒体播放器设备100(参见图1)用来存储内容的容器200的逻辑布置的图。通常,一个容器存储一个电影或剧集(episode)。容器200包括元数据202、视频数据204、主要音频数据206以及关联音频数据208。容器200还可以包括其它数据,为了简便起见省略对其它数据的讨论。元数据202限定并且描述容器200以及容器200中的内容。更具体地,元数据202限定容器的数据(例如,视频数据204、主要音频数据206以及关联音频数据208)的位置以及该数据之间的关系。元数据202可以包括媒体播放器设备100使用来同步容器的数据的读取、处理或输出的同步信息。元数据202可以还包括关于容器200或内容的其它数据,诸如电影名字、持续时间、菜单、轨道列表、场景索引、封面图像等。

[0034] 视频数据204一般对应于内容的视觉成分。例如,视频数据204可以是视频的帧。视频数据204的准确的格式可以根据各种编码和压缩函数而变化。例如,视频数据204可以是高清(HD)视频,诸如H.264、具有23.976fps的1080p。

[0035] 主要音频数据206一般对应于内容的主要音轨。主要音轨一般包括各种对话轨道,其中每个对话轨道具有不同的语言。例如,对于美国制作的电影,默认的主要音轨具有英语对话轨道;其它主要音轨将具有西班牙语、法语、德语、葡萄牙语、日语、朝鲜语等的对话轨道。另外,对于单语言,可以有多个轨道,例如具有不同的格式(例如,7.1声道环绕格式、立体声格式等)、比特率(192kbps、384kbps等)等的多个轨道。

[0036] 关联音频数据208一般对应于超出主要音频轨道的额外的音频轨道。关联音频的一个示例是音频描述轨道(还被称为“视频描述”或“描述视频服务(DVS)”)。音频描述轨道为视觉受损用户来描述视频。可以有与主要音频轨道的每种语言对应的音频描述轨道。关联音频的另一个示例是导演的评论轨道。一般导演的评论轨道为单语言,而不管选择哪个主要音频轨道。通常,当输出音频描述轨道时,导演的评论轨道不被输出。

[0037] 主要音频数据206和关联音频数据208可以以各种格式编码。例如,主要音频数据206可以被编码作为具有192kbps的5.1环绕声声道Dolby Digital Plus™(DD+)格式,或者作为具有384kbps的7.1环绕声声道DD+格式等。关联音频数据208可以被编码作为具有64kbps的单声道DD+,作为具有96kbps的立体声DD+等。

[0038] 容器200可以根据标准格式(例如,动态图像专家组版本4(MPEG-4)部分14格式,也被称为MP4格式)布置。

[0039] 通常,关联音频数据208具有与主要音频数据206一样多的声道,或者比主要音频数据206更少的声道。例如,如果主要音频数据206具有5.1声道格式,则关联音频数据208可以具有单声道、立体声、5.1声道格式等。如果主要音频数据206具有立体声格式,则关联音频数据208可以具有单声道、立体声等格式,但不具有5.1声道格式。



[0040] 通常,视频数据204对应于与对话轨道中的一个同时输出的单个视频轨道。类似地,当启用时,音频描述轨道中的一个与视频轨道同时地输出。此外,视频和音频轨道被同步地输出,如下面进一步讨论的。

[0041] 图3A是示出容器300的物理布置的图。尽管图2示出了逻辑结构,但是图3A示出了物理结构——如何将数据物理地存储或者布置在容器300中(例如,容器200是容器300的逻辑图)。元数据302处于容器300的“开头”(或者“顶部”或者“前面”)。然后在元数据302之后视频数据304(304a、304b等)、主要音频数据306(306a、306b等)以及关联音频数据308(308a、308b等)依次交错(interleave)。通常,对于视频数据的给定部分(例如,304a),要被同步(例如,同时、并行等)输出的对应的主要音频数据(例如,306a)和关联音频数据(例如,308a)在视频数据的下一部分(例如,304b)之前被交错。

[0042] 这种布置帮助媒体播放器设备100(参见图1)执行渐进式下载。在渐进式下载中,媒体播放器设备100首先下载元数据302,然后渐进地下载交错的视频数据304、主要音频数据306以及关联音频308;起始于内容的开头,并且依次进行到末尾。这使得媒体播放器设备100一旦已经下载了规定量的数据就能够开始播放内容。例如,如果规定量对应于30秒的内容,则媒体播放器设备100可以在它已经下载了与第一个30秒对应的交错的数据部分之后开始播放内容;在内容正在输出时的第一个30秒中,媒体播放器设备100执行后续的交错的数据部分的下载;等等。规定量可以根据各种因素(诸如与网络102的连接速度、内容的大小等)调节。

[0043] 由于每个视频数据部分与它的对应主要音频部分和对应关联音频部分物理邻近,因此这三个部分中的每一个可以一起被并行地读取,处理,并且同步地输出。

[0044] 图3B示出了图3A中示出的视频数据、主要音频数据和关联音频数据(例如,304x、306x和308x)的更多细节。如图3A中,交错的视频数据部分304x与交错的主要音频数据部分306x和交错的关联音频数据部分308x关联。每个音频数据部分内是轨道部分。通常,图3B通过示出每个数据部分内的多个关联轨道和多个主要音频轨道而给图3A增加更多细节。

[0045] 更具体地,主要音频数据部分306x包括m个主要音频轨道部分316a、316b、316m等。m个主要音频轨道部分对应于内容的m个主要音轨。例如,内容可以具有用于对话的三种不同语言(英语、法语和德语)的三个主要音轨。如果轨道部分316a对应于英语,则来自全部交错的主要音频数据部分306的具有英语的主要音频轨道部分的每一个的集合体(aggregation)对应于英语的整个主要对话轨道。

[0046] 类似地,关联音频数据部分308x包括对应于内容的m个音频描述轨道的m个关联音频轨道部分318a、318b、318m等。关联音频数据部分308x可以还包括其它关联音频轨道部分320x。其它关联音频轨道部分320x可以对应于导演的评论轨道。如同主要音频数据部分306x一样,给定语言的整个音频描述轨道对应于来自全部交错的关联音频数据部分308的用于该语言的关联音频轨道部分(例如,318a)的每一个的集合体。

[0047] 图4是输出多种语言的音频和关联音频的方法400的流程图。方法400可以通过媒体播放器设备100(参见图1)例如如一个或多个计算机程序控制的那样实现。

[0048] 在402处,存储多媒体容器。多媒体容器包括主要音频数据集和关联音频数据集。主要音频数据集和关联音频数据集具有多种语言。多媒体容器可以类似于容器200和300(参见图2-3)。多媒体容器可以根据渐进式下载来存储。

[0049] 在404处,接收第一选择信息。第一选择信息对应于该多种语言中的第一所选语言。例如,对于第一选择信息的默认设定可以是英语(或者另一种默认语言),如容器中的元数据指定的。作为另一个示例,用户可以通过从许多可用的语言中选择语言来提供第一选择信息,如通过媒体播放器设备100的用户界面显示的。作为另一个例子,第一选择信息可以根据对于先前容器中的内容的先前用户选择来设定。

[0050] 在406处,从多媒体容器输出第一音频数据流。第一音频数据流对应于主要音频数据集中的具有第一所选语言的第一所选主要音频数据集。例如,媒体播放器设备100可以将作为视频和音频流152中的一个的第一音频数据流输出到AVR 104(参见图1)。

[0051] 在408处,接收第二选择信息。第二选择信息与关联音频数据集有关。例如,用户可以通过与媒体播放器设备100的用户界面交互来指示要“开启”音频描述,从而提供第二选择信息。

[0052] 在410处,从多媒体容器输出第二音频数据流。第二音频数据流对应于关联音频数据集中的具有第一所选语言的第一所选关联音频数据集。例如,如果用户在404中选择法语,也以法语输出音频描述。

[0053] 在412处,接收第三选择信息。第三选择信息对应于该多种语言中的第二所选语言。例如,如果404中默认为英语,用户可以通过经由媒体播放器设备100的用户界面选择德语来提供第三选择信息。

[0054] 在414处,从多媒体容器输出第三音频数据流代替第一音频数据流。第三音频数据流对应于主要音频数据集中的具有第二所选语言的第二所选主要音频数据集。例如,如果用户在412中选择德语,主要音频切换到德语。

[0055] 在416处,从多媒体容器自动地输出第四音频数据流来代替第二音频数据流。第四音频数据流对应于关联音频数据集中的具有第二所选语言的第二所选关联数据集。注意,切换是“自动的”,因为用户不需要作为额外动作地重新启用音频描述(408中已做)或者为音频描述选择新的语言。例如,如果用户在412中选择德语,则音频描述自动地也切换到德语。

[0056] 参照回图3B,由于关联音频部分与主要音频部分交错,当用户改变语言选择(参见412)时,因为这些部分是物理邻近的,所以对于对话轨道和音频描述轨道两者的语言得到平滑的切换。另外,由于主要音频数据和关联音频数据在单个容器(参见图2)中,因此在发生语言切换时对于两者涉及相同的软件和硬件层(如通过媒体播放器控制程序管理的),进一步对平滑的切换有贡献。

[0057] 媒体播放器细节

[0058] 本节提供关于媒体播放器设备100(参见图1)的额外细节。

[0059] 图5是对于媒体播放器设备100(参见图1)的控制程序500(也被称为媒体播放器应用程序、媒体播放器或者简单地“应用程序”)的框图。控制程序500控制媒体播放器设备100操作作为全功能的启用Dolby Digital Plus的媒体播放器。如上面所讨论的(参见图1),控制程序500可以由与7.1声道AVR和1080p TV显示器连接的Mac Mini计算机(例如,型号标识符:Macmini5,1)执行。应用程序可以被写入以便在OS X 10.8.x(Mountain Lion)上进行操作。该应用程序可以是64位应用程序。

[0060] 应用程序500包括解多路复用器(demuxer)组件502、用户界面组件504、视频解码

器506、音频解码器508以及HDMI发送器组件510。解多路复用器组件502读取容器200中的视频和音频数据并且如用户界面组件504所指引地解多路复用各个音频流。解多路复用器组件502将所选的音频流提供给音频解码器508,并且将视频流提供给视频解码器组件506。

[0061] 用户界面组件504接收各种用户选择,诸如用于选择可用的容器中的一个的文件选择、用于选择对话轨道的轨道选择器、音频描述是否被启用、以及导演的评论是否被启用。用户界面组件504还实现用户选择的其它播放器功能(诸如播放、暂停、拖放等),并且将对控制信号发送给其它组件。

[0062] 视频解码器组件506读取容器200中的视频数据,对视频数据进行解码,并且将解码后的视频数据提供给HDMI发送器组件510。

[0063] 音频解码器组件508可以是DD+音频解码器,并且包括解码器520和混合器522。解码器520可以是DD+双声道解码器。解码器520接收所选的主要数据轨道和所选的关联数据轨道,并且将它们解码成脉冲编码调制(PCM)数据和混合元数据。混合器522从解码器520接收PCM数据和混合元数据,根据混合元数据来混合PCM数据,并且将混合后的音频数据提供给HDMI发送器组件510。

[0064] HDMI发送器组件510接收解码后的视频数据和混合后的音频数据,并且将这个数据封装作为HDMI输出信号以从媒体播放器设备100输出。

[0065] 应用程序控制媒体播放器设备100操作作为具有以下基本回放能力的全功能音频/视频播放器:播放、暂停和定位器(scrubber)(拖放功能)。媒体播放器支持解码DD+音频(5.1和7.1声道)。应用程序经由HDMI输出(在连接多声道的AVR时)输出多声道的PCM音频。

[0066] 媒体播放器解析源容器(例如,MP4文件)并且识别各个主要轨道和它们的语言以及各个关联音频轨道。媒体播放器支持可用的主要音频轨道之间的切换。媒体播放器支持两个DD+音频轨道(例如,5.1/7.1中的主要音频以及单声道中的关联音频)的双解码。在双解码的情况下,媒体播放器混合主要音频和关联音频以便创建混合的音频(例如,5.1或者7.1,取决于主要音频)。

[0067] 当对于名称存在多个音频描述轨道(例如,对于主要音频音轨中的每种语言存在匹配的音频描述轨道)时,应用程序深入地自动地挑选合适的音频描述轨道,即应该自动地挑选具有与当前有效的(播放的)主要音频音轨相同的媒体语言标签的音频描述轨道。如果只存在一个音频描述轨道,则应用程序不管主要音频的语言如何都启用这个音频描述轨道(根据用户选择)。

[0068] 应用程序支持内容类型如下。一种支持的类型是具有仅一个音频描述轨道(尽管对于主要音频有多种语言)的内容。另一种支持的类型是具有匹配的数量的音频描述轨道和主要音频轨道以及匹配的语言(即对于每种主要音频语言,存在该相同语言的对应音频描述轨道)的内容。在任一种情况下,还可以存在导演的评论轨道;上面描述的内容支持特别涉及音频描述轨道的数量,不涉及次要音频轨道的数量。

[0069] 应用程序能够识别容器中的每个音频轨道的语言(例如通过依赖媒体语言元数据标签)。应用程序能够在“音频描述”关联音频轨道之间以及“导演评论”关联音频轨道之间、还以及其它类型的关联音频内容之间识别。例如,元数据可以包括应用程序用来识别关联音频内容的类型的与DD+轨道格式有关的标签,诸如“bsmod”标签。

[0070] 媒体播放器确保混合的音频输出没有修剪或者过载。

[0071] 应用程序能够将源内容下混合至5.1声道输出或者至立体声输出。应用程序能够认识到下游AVR仅与5.1声道扬声器系统连接并且能够将任何7.1声道音频内容下混合至5.1。应用程序能够认识到仅立体声输出可用,并且能够将任何5.1声道或者7.1声道音频内容下混合至立体声。

[0072] 音频解码器508默认在具有0%比例因子(即最大动态范围)的线路模式(Line mode)动态范围控制(DRC)中操作。应用程序能够切换到射频(RF)模式DRC或者便携模式DRC。

[0073] 应用程序响应于来自以下源的输入事件:鼠标移动和点击,键盘映射,以及红外信号(例如,诸如Apple Remote™设备之类的遥控)。可替代地,当通过移动设备实现应用程序时,应用程序响应于触摸输入。

[0074] 应用程序在起动时自动地支持显示器分辨率(以便匹配显示设备(例如,经由HDMI连接的TV)的固有分辨率)。应用程序在刚起动时以全屏启动并且以全屏操作;可替代地应用程序可以在窗口中启动和操作。在起动时,应用程序缩放视频回放以便匹配屏幕使得整个视频被显示以及使得视频的高宽比被保持。不要求在运行时刻动态地适应显示器的分辨率的变化。

[0075] 在回放期间,当到达内容的末尾时,应用程序自动地返回到内容选择屏幕。

[0076] 当鼠标光标被保持在屏幕顶部处多于1s时应用程序显示OS X系统菜单栏;在该情况下,应用程序将能通过点击:DD+电影>退出DD+电影而终止。

[0077] 不需要用于应用程序的安装程序,然而期望的是应用程序的安装要求将关联资源放置在文件系统上的特定位置中(与资源被捆在应用程序包内部相反)。

[0078] 应用程序一直对用户交互进行响应。导航应用程序用户界面UI(例如在可用的电影之间循环,或者导航回放菜单)是平滑并且流畅的,而不是迟钝的。用户界面控制的状态(例如,高亮对比非高亮)被立即更新(例如,小于200ms)以反映用户动作,而没有滞后。

[0079] 即使在多次切换音频轨道或者多次启用和禁用关联音频轨道之后,音频与视频也一直保持同步。

[0080] 选择新的语言轨道快速且没有爆裂声或毛刺地从当前音频轨道切换到新的音频轨道。当切换到新的语言轨道时,对新的轨道的音频回放立即或在小于1秒钟内恢复。如果切换到新的音频轨道需要多于200毫秒,则视频流暂停直到它能够与音频一起恢复,使得音频内容不被丢失。

[0081] 选择关联音频轨道自动地在该轨道中与主要音频混合,而没有爆裂声和毛刺。当启用或禁用关联音频轨道时,启用这个轨道、回放混合的音频的操作立即或在至多1秒钟之后恢复。

[0082] 以平滑和流畅的方式动画化用户界面控制条的逐渐淡入和淡出。在可应用的情况下,用户界面还使用优美的动画用于屏幕转变,例如在上下循环通过可用的电影时,在从内容选择屏幕转变到回放屏幕并且反之亦然时(水平滑动动画),以及在显示或隐藏UI控制条中的上拉菜单时(向上滑动/向下滑动动画)。

[0083] 在左右移动拖放柄(handle)的同时,视频显示器定期地更新,例如示出视频中的I-帧,使得用户具有剪辑的哪一部分对应于拖放的视觉反馈。

[0084] 用户界面细节

[0085] 应用程序的用户界面可以处于各种状态。这些状态包括启动时的初始屏幕(状态1)和回放屏幕(状态2)。

[0086] 状态1:启动时的初始屏幕

[0087] 当应用程序启动时,第一视图呈现可用于观看的电影资源,与关于电影的相关技术信息一起,特别地是音频内容。在任何给定时间在屏幕上仅呈现一个电影名称,以便使屏幕占地面积(real estate)最大化,但是相邻的电影名称海报在当前可见的电影海报上方和/或下方稍微可见(暗淡)。用户能够用鼠标指针点击那些在当前名称上方或下方的变暗淡的电影海报以便浏览电影。除了鼠标之外,用户还能够如下地从各个电影名称中进行选择:通过按压键盘上的向上或向下箭头键,或者通过按压遥控器上的向上或向下按钮。

[0088] 应用程序如下地从一个电影转变到另一个电影。当用户点击当前电影上方或下方的暗淡的海报时,首先文字消失。其次,居中的电影海报逐渐地变得暗淡。第三,所有电影海报向上滑动或者向下滑动直到点击的新电影的海报在屏幕的中心。滑动动画使用缓入/缓出定时曲线,即动画速度在中间最快并且在开始和结束时渐小。第四,新的点击的电影海报逐渐地变为全亮度。第五,新的对应文字出现。所有这些步骤能够同时发生,或者至少有一些交迭(例如,文字能够在电影变得暗淡的同时消失)。可替代地,在电影海报从中心向上滑动时,其在搁板(shelf)上的倒影成比例地向下滑动。并且同样地,在新的电影海报朝向搁板向下滑动时,它的倒影向上滑动直到它碰到海报的底部。

[0089] 用户能够通过用鼠标点击电影海报、按压键盘上的回车键和/或按压遥控器上的“进入”或者“播放”来选择电影。这将立即触发电影的回放,使用语言和其它选项(Extras)(例如,音频描述等)的默认设定。回放可以从电影的开头开始。可替代地,回放根据“从我之前离开处恢复的功能”而开始。

[0090] 状态2:回放屏幕

[0091] 当回放开始时,控制条被显示在屏幕上,覆盖在视频帧之上。控制条在3秒钟之后逐渐消失,除非存在新的事件(鼠标移动/点击,键盘输入,遥控信号等)。新的输入事件或者维持控制条在屏幕上或者在控制条已经逐渐消失的情况下将其带回。控制条上的控制包括播放/暂停轮换控制、语言轨道选择菜单按钮、其它选项选择菜单按钮、拖动条、以及退出按钮。

[0092] 用户界面一般如下地操作。用户能够使用鼠标指针点击控制条中的按钮。用户还能够使用键盘箭头以便在不同控制之间导航并且能够使用回车键以激活控制。用户还能够按压键盘上的空格键。这用作对播放/暂停功能的直接控制,而不管用户处于控制菜单中的何处。如果控制条已经逐渐消失,按压空格键不应该使控制条回到屏幕上;它应该简单地暂停或者恢复回放。

[0093] 用户还能够使用遥控器上的箭头按钮在不同控制之间导航并且使用选择按钮以激活控制。用户还能够按压遥控器上的播放/暂停按钮以用作对播放/暂停功能的直接控制,而不管用户处于控制菜单中的何处。如果控制条已经逐渐消失,按压遥控器上的播放/暂停按钮不应该使控制条回到屏幕上;它应该简单地暂停或者恢复回放。

[0094] 用户能够使用鼠标指针来选择拖动条中的位置按钮并且将它向左或向右拖曳以便移动到电影中的不同位置。屏幕上的图像周期性地跟踪这个按钮的移动位置,即使鼠标点击没有被释放。一旦释放,同步的音频与视频的回放从新的位置恢复。用户还应该能够通

过按压并保持遥控器上的“向右”或者“向左”按钮来实行这个拖动功能。“向右”按钮应该将位置按钮在电影中向前移动,而“向左”按钮应该向后移动该位置。位置按钮应该逐渐地移动预定义的增量(例如,按钮被维持按压每半秒,就有10秒的增量)。增量可以是要跳过的时间单位或者视频的预定义数量的I帧。拖动条不需要是使用键盘输入方法可控的。

[0095] 当用于语言或者其它选项的菜单按钮中的一个被激活(例如,点击)时,对于这个控制的菜单从菜单按钮向上滑动,而在覆盖的控制之下继续视频回放。一旦菜单已经完成向上滑动转变,这个菜单上的区域变得可通过使用键盘或者遥控器上的向上/向下箭头或者通过使用鼠标指针来导航。

[0096] 子菜单项目能够通过按压键盘上的回车键、遥控器上的“选择”或者通过用鼠标点击来选择。当项目被点击/选择/确认时,这个菜单子项目中的文字的颜色在返回到默认文字颜色之前应该短暂地以蓝色高亮以便确认用户选择。子菜单面板还在1秒之后自动地向下滑动回来。

[0097] 当用户界面控制条上的菜单按钮有效(例如,鼠标停留于按钮上,或者用户已经按压遥控器或者键盘上的箭头键使得给定按钮准备好被选择/激活)时,该按钮表现为高亮状态。子菜单中的当前所选/有效项目也呈现高亮,直到进行新的选择。

[0098] 通常,控制的导航不中断视频和音频回放(除非激活退出按钮)。

[0099] 默认语言是第一语言轨道。点击新的语言立即切换音轨到这个声道。其它选项的默认值是“关闭”(例如,导演的评论或者音频描述轨道不应该有效)。

[0100] 根据一个实施例,在任何给定时间仅一个其它选项有效。点击导演的评论子项目立即开始这个轨道与主要音频轨道的混合。点击音频描述子项目立即停止混合导演的评论轨道并且开始混入适当的音频描述轨道作为替代。点击“关闭”关掉任何关联音频轨道。一次应该至多一个关联音频轨道有效。

[0101] 当已经显示菜单面板时,点击键盘上的“退出”或者遥控器上的“菜单”取消菜单选择并且关闭菜单面板。

[0102] 点击退出按钮中断回放并且将用户带回到初始电影选择屏幕,显示选择的最近的电影。退出控制还能够通过按压键盘上的退出(Esc)键或者通过按压遥控器上的“菜单”按钮激活。

[0103] 用户界面图形资源特征

[0104] 用户通过浏览用于电影海报以及它的有关描述的图形选择观看的内容。

[0105] 偏好(preferences)窗口

[0106] 应用程序具有可使用菜单栏(点击应用程序名并且选择“偏好...”菜单项)或者使用快捷键访问的“偏好”窗口。应用程序的偏好窗口不是模式的(即它允许用户保持观看和与应用程序的其余部分交互,以便理解各个设定的影响)。通常,应用程序的偏好窗口不是可调整大小的,并且它的缩放和最小化按钮被禁用。当应用程序的偏好窗口关闭并且再打开时,至少在应用程序保持运行时,它应该再打开到用户在用户关闭它时选择的相同的窗格。在用户进行偏好窗口中的设定的改变时,改变立即生效,而不要求用户点击OK或者应用按钮并且不等待用户关闭窗口。用户设定的偏好/设定是持续不变的(即它们对于应用程序的整个会话以及在应用程序被终止和重新启动之后仍保持有效)。

[0107] 应用程序的偏好窗口应该具有遵循以下的对于“通用的”偏好窗格的描述的用户

界面。通用的偏好窗格包括用于选择优选的语言的下拉式菜单、用于选择动态范围压缩(线路模式、RF模式以及便携模式)的单选按钮以及用于选择输出声道的数量(7.1、5.1、立体声、单声道等)的下拉式菜单。应用程序的偏好窗口的通用的窗格提供改变用于DD+解码器的DRC简档、选择输出声道配置和选择优选的(默认的)语言的能力。

[0108] 对于默认/优选的语言轨道的控制默认为英语,并且可以至少提供以下选择:英语、法语、德语、西班牙语。应用程序通过在首次开始回放电影时自动地选择默认语言轨道(如果这个语言可用的话)来遵从对于默认/优选的语言的控制。在回放剪辑/电影期间,如果用户选择与默认值不同的语言,则这个新的语言在会话的持续时间内(即直到应用程序被终止)被记住。

[0109] 对于动态范围压缩的控制默认为具有0%比例因子的“线路模式”DRC(即最大动态范围)。可使用滑动器以1的增量在0与100之间调节比例因子。该控制提供将DRC简档改为RF模式或便携模式的能力。当选择“便携模式”时,下拉式菜单提供在-8、-11与-14LKFS之间选择便携模式输出水平的能力(其中-11LKFS为默认值)。当未选择便携模式时,下拉式菜单被禁用。当未选择线路模式时,用于比例因子的滑动器被禁用。应用程序遵从用户设定的DRC简档。

[0110] 应用程序的偏好窗口提供在立体声(2声道)输出、5.1声道输出与7.1声道输出之间改变用于应用程序的输出声道的数量的能力。当选择5.1或7.1时,应用程序仅经由HDMI输出端输出音频。当选择5.1时,任何主要或次要的音频轨道在到达混合器之前在解码器中被下混合到5.1(即仅独立的子流被解码)。当选择立体声时,任何主要或次要的音频轨道在到达混合器之前在DD+解码器中被下混合到立体声。在可应用时,应用程序经由HDMI链接以及头戴耳机端口、Sony/Philips数字互连格式(S/PDIF)输出以及内置的扬声器输出立体声音频。

[0111] 流式传输

[0112] 虽然上面已经描述了媒体播放器设备100的渐进式下载特征,但是它还可以被配置为对流式传输内容(例如,来自苹果公司的超文本传输协议(HTTP)实时流式传输(HLS))进行操作。

[0113] 实施方式细节

[0114] 应当理解在此描述的各个技术可以与硬件或软件、或适当时与两者的组合结合地实现。因此,本发明的方法与装置或其特定方面或部分可以采取程序代码(即,指令)的形式,该程序代码具体实现在有形的介质(诸如软盘、CD-ROM、硬盘驱动、固态驱动或任何其它机器可读的存储介质)中,其中在程序代码被载入机器(诸如计算机)中并且由该机器执行时,该机器变为用于实践本发明的装置。在可编程计算机上执行程序代码的情况下,计算设备一般包括处理器、可由处理器读取的存储介质(包括易失性的和非易失性的存储器和/或存储元件)、至少一个输入设备以及至少一个输出设备。一个或更多个程序可以实现或利用与本发明结合描述的处理,例如,通过使用API、可重复使用的控制等。这种程序优选地以高级程序的或面向对象的编程语言实现以便与计算机系统通信。然而,如果需要,一个或更多个程序能够以汇编或机器语言实现。在任何情况下,该语言可以是编译的或者解释的语言,并且与硬件实现相组合。

[0115] 虽然示例性实施例可能在一个或更多个独立计算机系统的上下文中提及利用本

发明的方面,但是本发明不限于此,而是可以与任何计算环境(诸如网络或分布计算环境)结合地实现。更进一步,本发明的方面可以被实现在多个处理芯片或设备中或横跨多个处理芯片或设备,并且存储可以类似地横跨多个设备地实现。这种设备可以包括个人计算机、网络服务器、手持设备、超级计算机、或集成到其它系统(诸如汽车和飞机)中的计算机。

[0116] 作为具体示例,媒体播放器设备100(参见图1)可以运行控制它的操作的软件或固件,以便例如如上所述地接收、处理和输出视频和音频数据。设备100可以包括经由因特网连接到一个或多个其它计算机的计算机,或者连接到该计算机,或者被该计算机控制。

[0117] 以上描述示出本发明的各个实施例以及可以如何实现本发明的方面的示例。以上示例和实施例不应该被认为是唯一的实施例,并且被提出以便示出如由以下权利要求所限定的本发明的优点和灵活性。基于上述公开以及下面的权利要求,本领域技术人员将明白其它布置、实施例、实现方式和等同物并且可以在不脱离如由权利要求所限定的本发明精神和范围的情况下采用。



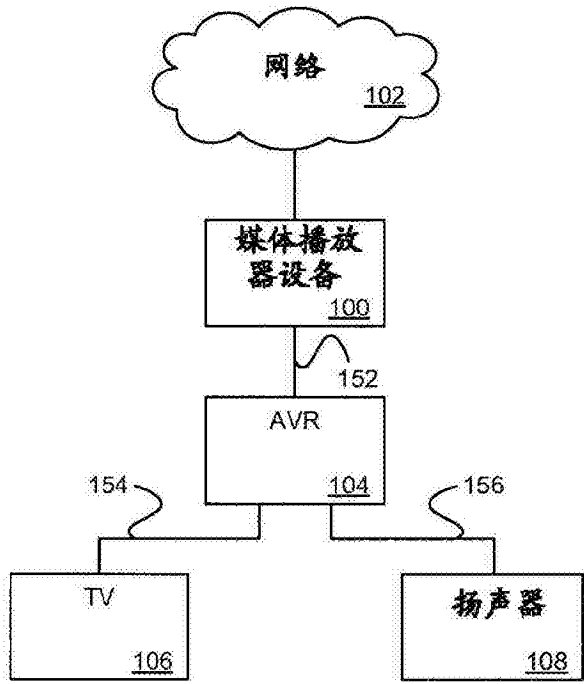


图1

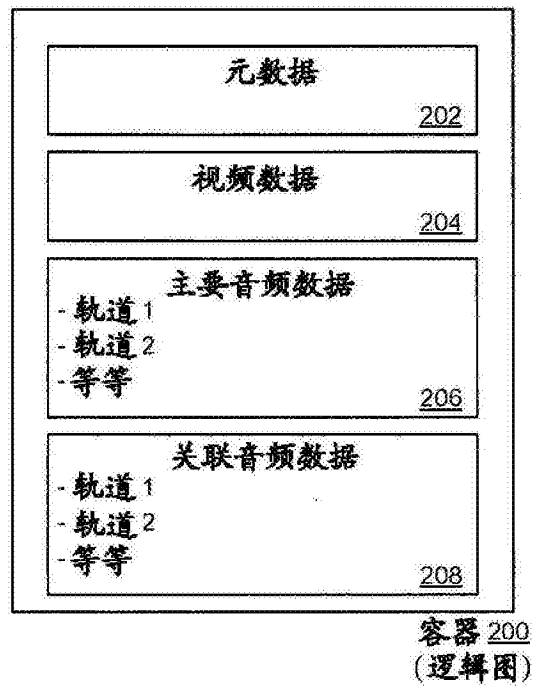
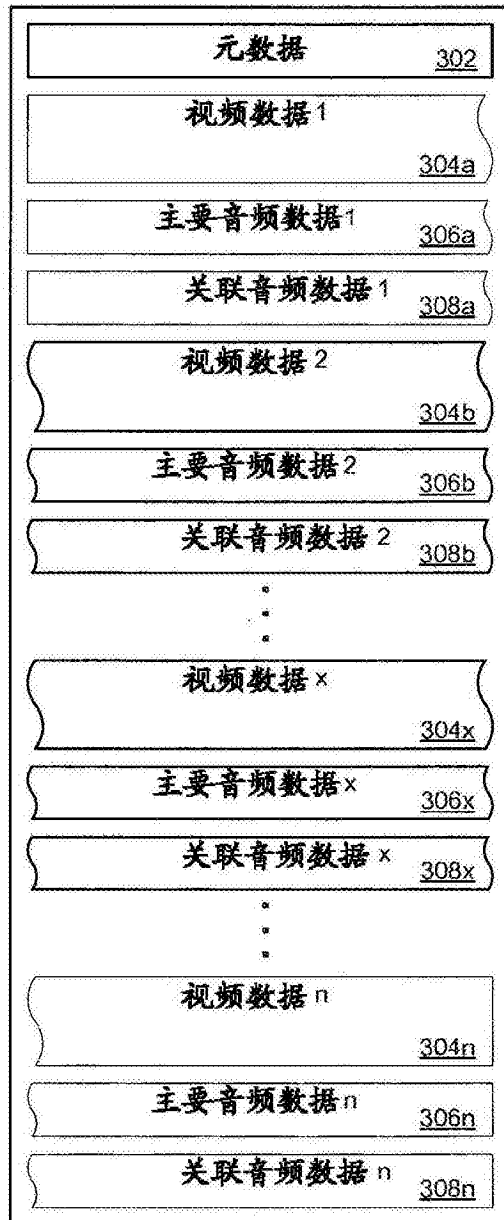
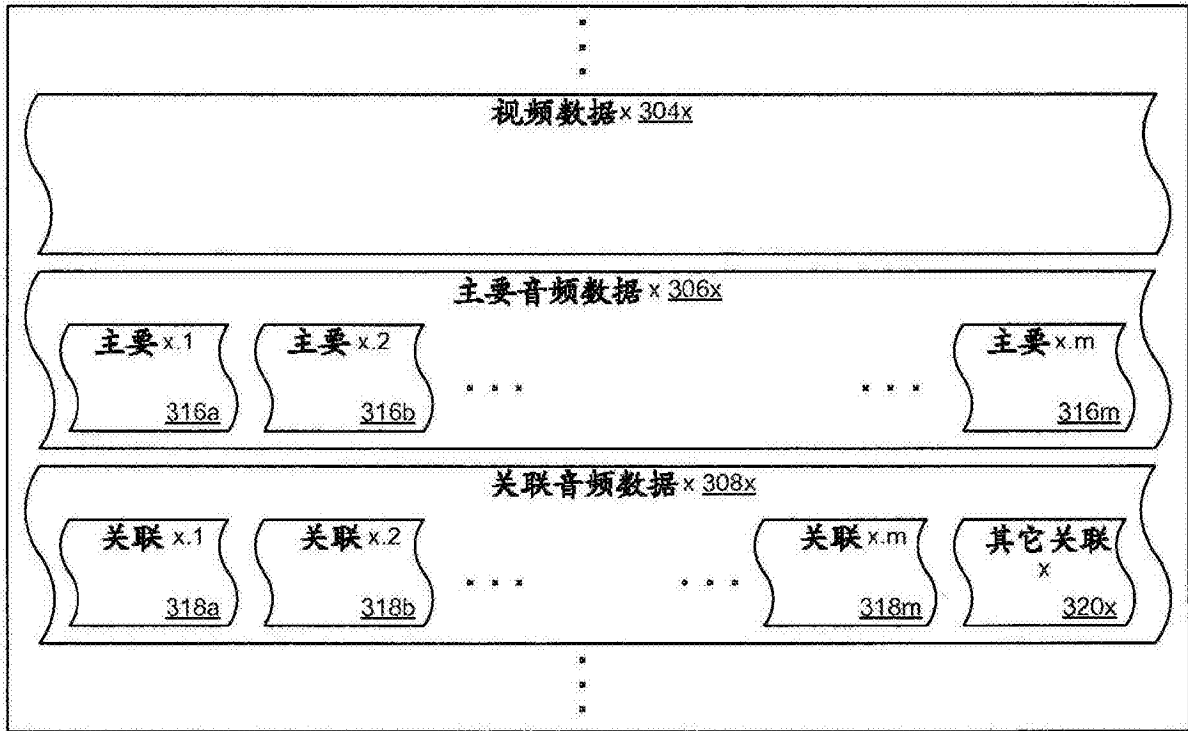


图2



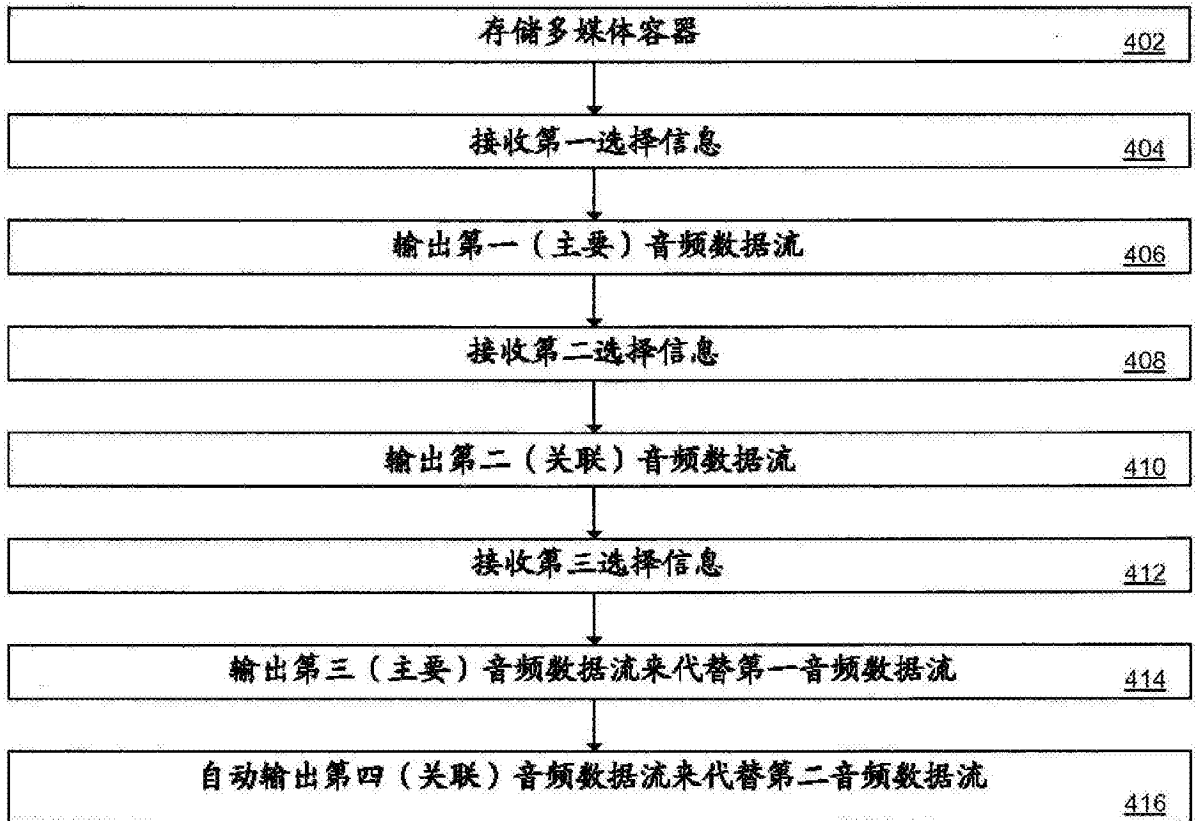
容器 300  
(物理图)

图3A



容器 300  
(物理图)

图3B



400

图4

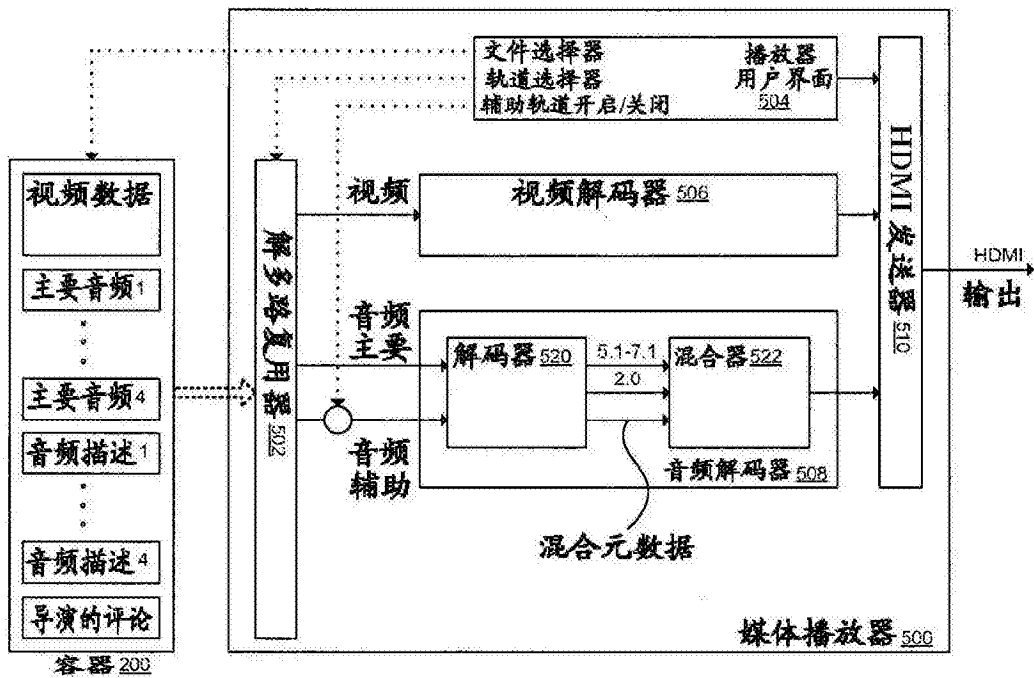


图5