



(10) 授权公告号 CN 109076238 B

(45) 授权公告日 2022. 08. 05

(21) 申请号 201780028785.5

(22) 申请日 2017.05.24

(65) 同一申请的已公布的文献号  
申请公布号 CN 109076238 A

(43) 申请公布日 2018.12.21

(30) 优先权数据  
62/341,012 2016.05.24 US  
15/602,887 2017.05.23 US

(85) PCT国际申请进入国家阶段日  
2018.11.09

(86) PCT国际申请的申请数据  
PCT/US2017/034261 2017.05.24

(87) PCT国际申请的公布数据  
W02017/205504 EN 2017.11.30

(73) 专利权人 高通股份有限公司  
地址 美国加利福尼亚州

(72) 发明人 王业奎

(74) 专利代理机构 永新专利商标代理有限公司  
72002

专利代理师 赵腾飞

(51) Int.Cl.  
H04N 19/597 (2006.01)  
H04N 21/218 (2006.01)  
H04N 21/2343 (2006.01)  
H04N 21/81 (2006.01)  
H04N 21/845 (2006.01)  
H04N 21/8543 (2006.01)  
H04N 19/70 (2006.01)

(56) 对比文件  
W0 2015197818 A1,2015.12.30  
CN 104335590 A,2015.02.04  
CN 105144728 A,2015.12.09  
CN 104854835 A,2015.08.19  
W0 2015012605 A1,2015.01.29  
W0 2014022017 A1,2014.02.06

审查员 胡西

权利要求书4页 说明书24页 附图11页

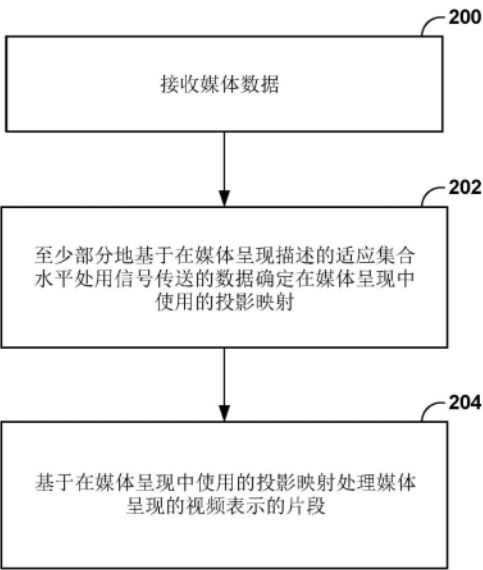
(54) 发明名称

通过HTTP在动态自适应流式传输中用信号  
传送虚拟现实视频

(57) 摘要

一种用于处理媒体数据的装置经配置以进行以下操作:接收包含虚拟现实VR视频数据的媒体数据;至少部分地基于在用于媒体呈现的媒体呈现描述的适应集合水平处用信号传送的数据确定在所述媒体呈现中使用的投影映射;基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的视频表示的片段。一种用于处理媒体数据的装置经配置以进行以下操作:产生包含VR视频数据的媒体数据;在所述媒体数据中包含在识别在包含于所述媒体数据中的媒体呈现中使用的投影映射的媒体呈现描述的适应集合水平处用信号传送的数据;以及基于在所述媒体呈现中使用的所述投影映射发送所述媒体呈现的视

频表示的片段。



1. 一种处理媒体数据的方法,所述方法包括:

接收媒体呈现的虚拟现实VR视频表示,其中,所述媒体呈现包括一个或多个周期,所述一个或多个周期包括多个视频表示,所述多个视频表示包括虚拟现实VR视频表示和非VR视频表示,并且所述视频表示包括一个或多个片段;

至少部分地基于在用于描述所述媒体呈现的媒体呈现描述中用信号传送的数据确定在所述媒体呈现中使用的投影映射;以及

基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的所述VR视频表示的所述片段,

其中,在所述媒体呈现描述中用信号传送的所述数据是在适应集合水平处用信号传送的,并且不是在表示水平处用信号传送的。

2. 根据权利要求1所述的方法,其中基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的所述VR视频表示的所述片段包括:

至少部分地基于在所述媒体呈现中使用的所述投影映射确定是否检索所述VR视频表示的所述片段;以及

检索所述VR视频表示的所述片段。

3. 根据权利要求1所述的方法,其中基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的所述VR视频表示的所述片段包括:

至少部分地基于在所述媒体呈现中使用的所述投影映射确定如何解码所述VR视频表示的所述片段。

4. 根据权利要求1所述的方法,其进一步包括:

确定在所述媒体呈现中使用的所述投影映射的指示是否存在于所述媒体呈现的周期水平处,并且其中至少部分地基于在媒体呈现描述的所述适应集合水平处用信号传送的所述数据确定在所述媒体呈现中使用的所述投影映射是响应于确定在所述媒体呈现中使用的所述投影映射的所述指示并非存在于所述媒体呈现的所述周期水平处执行的。

5. 根据权利要求1所述的方法,其进一步包括:

确定在所述媒体呈现中使用的所述投影映射的指示是否存在于所述媒体呈现的周期水平处,并且其中至少部分地基于在媒体呈现描述的所述适应集合水平处用信号传送的所述数据确定在所述媒体呈现中使用的所述投影映射是响应于确定在所述媒体呈现中使用的所述投影映射的所述指示存在于所述媒体呈现的所述周期水平处执行的。

6. 根据权利要求1所述的方法,其进一步包括:

至少部分地基于在第二媒体呈现的周期水平处用信号传送的数据且无需在所述第二媒体呈现的表示水平处的额外的用信号传送确定在所述第二媒体呈现中使用的投影映射;

基于在所述第二媒体呈现中使用的所述投影映射处理所述第二媒体呈现的视频表示的片段。

7. 根据权利要求1所述的方法,其中在所述适应集合水平处用信号传送的所述数据包括在适应集合中的元素,其中所述VR视频表示属于所述适应集合。

8. 根据权利要求1所述的方法,其进一步包括:

基于在所述媒体呈现中使用的所述投影映射解码所述媒体呈现的所述VR视频表示的所述片段;以及

输出所述媒体呈现的所述VR视频表示的所述经解码片段。

9. 根据权利要求1所述的方法,其进一步包括:

将所述媒体呈现的所述VR视频表示的所述片段发送到客户端装置。

10. 一种用于处理媒体数据的装置,所述装置包括:

存储器,其经配置以存储所述媒体数据;以及

一或多个处理器,其经配置以进行以下操作:

接收媒体呈现的虚拟现实VR视频表示,其中,所述媒体呈现包括一个或多个周期,所述一个或多个周期包括多个视频表示,所述多个视频表示包括虚拟现实VR视频表示和非VR视频表示,并且所述视频表示包括一个或多个片段;

至少部分地基于在用于描述所述媒体呈现的媒体呈现描述中用信号传送的数据确定在所述媒体呈现中使用的投影映射;以及

基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的所述VR视频表示的所述片段,

其中,在所述媒体呈现描述中用信号传送的所述数据是在适应集合水平处用信号传送的,并且不是在表示水平处用信号传送的。

11. 根据权利要求10所述的装置,其中为了基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的所述VR视频表示的所述片段,所述一或多个处理器经配置以至少部分地基于在所述媒体呈现中使用的所述投影映射确定是否检索所述VR视频表示的所述片段以及检索所述VR视频表示的所述片段。

12. 根据权利要求10所述的装置,其中为了基于在所述媒体呈现中使用的所述投影映射处理所述媒体呈现的所述VR视频表示的所述片段,所述一或多个处理器经配置以至少部分地基于在所述媒体呈现中使用的所述投影映射确定如何解码所述VR视频表示的所述片段。

13. 根据权利要求10所述的装置,其中所述一或多个处理器进一步经配置以进行以下操作:

确定在所述媒体呈现中使用的所述投影映射的指示是否存在于所述媒体呈现的周期水平处,并且其中所述一或多个处理器经配置以至少部分地基于在媒体呈现描述的所述适应集合水平处用信号传送的所述数据响应于确定在所述媒体呈现中使用的所述投影映射的所述指示并非存在于所述媒体呈现的所述周期水平处确定在所述媒体呈现中使用的所述投影映射。

14. 根据权利要求10所述的装置,其中所述一或多个处理器进一步经配置以进行以下操作:

确定在所述媒体呈现中使用的所述投影映射的指示是否存在于所述媒体呈现的周期水平处,并且其中所述一或多个处理器经配置以至少部分地基于在媒体呈现描述的所述适应集合水平处用信号传送的所述数据响应于确定在所述媒体呈现中使用的所述投影映射的所述指示存在于所述媒体呈现的所述周期水平处确定在所述媒体呈现中使用的所述投影映射。

15. 根据权利要求10所述的装置,其中所述一或多个处理器进一步经配置以进行以下操作:

至少部分地基于在第二媒体呈现的周期水平处用信号传送的数据且无需在所述第二媒体呈现的表示水平处的额外的用信号传送确定在所述第二媒体呈现中使用的投影映射；

基于在所述第二媒体呈现中使用的所述投影映射处理所述第二媒体呈现的视频表示的片段。

16. 根据权利要求10所述的装置，其中在所述适应集合水平处用信号传送的所述数据包括在适应集合中的元素，其中所述VR视频表示属于所述适应集合。

17. 根据权利要求10所述的装置，其中所述一或多个处理器进一步经配置以进行以下操作：

基于在所述媒体呈现中使用的所述投影映射解码所述媒体呈现的所述VR视频表示的所述片段；以及

输出所述媒体呈现的所述VR视频表示的所述经解码片段。

18. 根据权利要求10所述的装置，其中所述一或多个处理器进一步经配置以将所述媒体呈现的所述VR视频表示的所述片段发送到客户端装置。

19. 根据权利要求10所述的装置，其进一步包括经配置以接收所述媒体数据的接收器。

20. 根据权利要求19所述的装置，其中所述装置包括无线通信装置，并且其中所述接收器经配置以根据无线通信标准来解调包括所述媒体数据的信号。

21. 一种处理媒体数据的方法，所述方法包括：

产生媒体呈现，其中，所述媒体呈现包括一个或多个周期，所述一个或多个周期包括多个视频表示，所述多个视频表示包括虚拟现实VR视频表示和非VR视频表示，并且所述视频表示包括一个或多个片段；

在用于描述所述媒体呈现的媒体呈现描述中包括用于识别在所述媒体呈现中使用的投影映射的数据；以及

基于在所述媒体呈现中使用的所述投影映射发送所述媒体呈现的VR视频表示的片段，

其中，用于识别所述投影映射的所述数据是在适应集合水平处被包括在所述媒体呈现描述中的，并且不是在表示水平处被包括在所述媒体呈现描述中的。

22. 根据权利要求21所述的方法，其进一步包括：

确定是否在所述媒体呈现的周期水平处包含在所述媒体呈现中使用的所述投影映射的指示，并且其中响应于确定在所述媒体呈现的所述周期水平处不包含在所述媒体呈现中使用的所述投影映射的所述指示执行在所述媒体数据中包含在识别在所述媒体呈现中使用的所述投影映射的媒体呈现描述的所述适应集合水平处用信号传送的数据。

23. 根据权利要求21所述的方法，其进一步包括：

确定是否在所述媒体呈现的周期水平处包含在所述媒体呈现中使用的所述投影映射的指示，并且其中响应于确定在所述媒体呈现的所述周期水平处包含在所述媒体呈现中使用的所述投影映射的所述指示执行在所述媒体数据中包含在识别在所述媒体呈现中使用的所述投影映射的媒体呈现描述的所述适应集合水平处用信号传送的数据。

24. 根据权利要求21所述的方法，其进一步包括：

在所述媒体数据中包含在识别在第二媒体呈现中使用的投影映射的第二媒体呈现的周期水平处用信号传送的数据；以及

基于在所述第二媒体呈现中使用的所述投影映射发送所述第二媒体呈现的视频表示

的片段。

25. 根据权利要求21所述的方法, 其中在所述适应集合水平处用信号传送的所述数据包括在适应集合中的元素, 其中所述VR视频表示属于所述适应集合。

26. 一种用于处理媒体数据的装置, 所述装置包括:

存储器, 其经配置以存储所述媒体数据; 以及

一或多个处理器, 其经配置以进行以下操作:

产生媒体呈现, 所述媒体呈现包括一个或多个周期, 所述一个或多个周期包括多个视频表示, 所述多个视频表示包括虚拟现实VR视频表示和非VR视频表示, 并且所述视频表示包括一个或多个片段;

在用于描述所述媒体呈现的媒体呈现描述中包括用于识别在所述媒体呈现中使用的投影映射的数据; 以及

基于在所述媒体呈现中使用的所述投影映射发送所述媒体呈现的VR视频表示的所述片段,

其中, 用于识别所述投影映射的所述数据是在适应集合水平处被包括在所述媒体呈现描述中的, 并且不是在表示水平处被包括在所述媒体呈现描述中的。

27. 根据权利要求26所述的装置, 其中所述一或多个处理器进一步经配置以进行以下操作:

确定是否在所述媒体呈现的周期水平处包含在所述媒体呈现中使用的所述投影映射的指示, 并且其中所述一或多个处理器经配置以响应于确定在所述媒体呈现的所述周期水平处不包含在所述媒体呈现中使用的所述投影映射的所述指示在所述媒体数据中包含在识别在所述媒体呈现中使用的所述投影映射的媒体呈现描述的所述适应集合水平处用信号传送的数据。

28. 根据权利要求26所述的装置, 其中所述一或多个处理器进一步经配置以进行以下操作:

确定是否在所述媒体呈现的周期水平处包含在所述媒体呈现中使用的所述投影映射的指示, 并且其中所述一或多个处理器经配置以响应于确定在所述媒体呈现的所述周期水平处包含在所述媒体呈现中使用的所述投影映射的所述指示在所述媒体数据中包含在识别在所述媒体呈现中使用的所述投影映射的媒体呈现描述的所述适应集合水平处用信号传送的数据。

29. 根据权利要求26所述的装置, 其中在所述适应集合水平处用信号传送的所述数据包括在适应集合中的元素, 其中所述VR视频表示属于所述适应集合。

30. 根据权利要求26所述的装置, 其中所述装置包括无线通信装置, 所述无线通信装置包括发射器, 其中所述发射器经配置以根据无线通信标准调制包括所述媒体数据的信号。

## 通过HTTP在动态自适应流式传输中用信号传送虚拟现实视频

[0001] 本申请主张2016年5月24日递交的第62/341,012号美国临时专利申请的权益,所述临时专利申请的全部内容以引用的方式并入本文中。

### 技术领域

[0002] 本发明涉及经编码视频数据的存储和传送。

### 背景技术

[0003] 数字视频能力可并入到广泛范围的装置中,包含数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型计算机或桌上型计算机、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电电话、视频电话会议装置及类似者。数字视频装置实施视频压缩技术,例如在由MPEG-2、MPEG-4、ITU-T H.263或ITU-T H.264/MPEG-4,第10部分,高级视频译码(AVC)、ITU-T H.265-高效视频译码(HEVC)所定义的标准和此类标准的扩展中所描述的那些,以更有效地发射和接收数字视频信息。

[0004] 视频压缩技术执行空间预测和/或时间预测以减少或移除视频序列中固有的冗余。对于基于块的视频译码来说,可将视频帧或切片分割为多个宏块。可进一步分割每个宏块。使用相对于相邻宏块的空间预测来编码帧内译码(I)帧或切片中的宏块。经帧间译码(P或B)帧或切片中的宏块可使用相对于同一帧或切片中的相邻宏块的空间预测或相对于其它参考帧的时间预测。

[0005] 在视频数据已经编码之后,可将视频数据封包化以用于发射或存储。视频数据可以汇编到符合例如国际标准化组织(ISO)基础媒体文件格式及其扩展(例如,HEVC文件格式)的多种标准中的任一个的视频文件中。

### 发明内容

[0006] 一般来说,本发明描述与虚拟现实(VR)相关的技术。更确切地说,本发明描述用于通过HTTP(DASH)的动态自适应流式传输中的VR视频的用信号传送的技术。

[0007] 根据本发明的一个实例,用于处理媒体数据的方法包含:接收媒体数据,其中所述媒体数据包括虚拟现实(VR)视频数据;至少部分地基于在用于媒体呈现的媒体呈现描述的适应集合水平处用信号传送的数据确定在媒体呈现中使用的投影映射;以及基于在媒体呈现中使用的投影映射处理媒体呈现的视频表示的片段。

[0008] 根据本发明的另一实例,用于处理媒体数据的装置包含:存储器,其经配置以存储媒体数据;以及一或多个处理器,其经配置以进行以下操作:接收媒体数据,其中所述媒体数据包括虚拟现实(VR)视频数据;至少部分地基于在用于媒体呈现的媒体呈现描述的适应集合水平处用信号传送的数据确定在媒体呈现中使用的投影映射;以及基于在媒体呈现中使用的投影映射处理媒体呈现的视频表示的片段。

[0009] 根据本发明的另一实例,处理媒体数据的方法包含:产生媒体数据,其中所述媒体数据包含虚拟现实(VR)视频数据;在所述媒体数据中包含在识别在包含于媒体数据中的媒

体呈现中使用的投影映射的媒体呈现描述的适应集合水平处用信号传送的数据;以及基于在媒体呈现中使用的投影映射处理媒体呈现的视频表示的片段。

[0010] 根据本发明的另一实例,用于处理媒体数据的装置包含:存储器,其经配置以存储媒体数据;以及一或多个处理器,其经配置以进行以下操作:产生媒体数据,其中所述媒体数据包含虚拟现实 (VR) 视频数据;在所述媒体数据中包含在识别在包含于媒体数据中的媒体呈现中使用的投影映射的媒体呈现描述的适应集合水平处用信号传送的数据;以及基于在媒体呈现中使用的投影映射发送媒体呈现的视频表示的片段。

[0011] 在附图以及以下描述中阐述一或多个实例的细节。其它特征、目标和优点将从所述描述、图式以及权利要求书中显而易见。

## 附图说明

[0012] 图1是说明实施用于经由网络的流式传输媒体数据的技术的实例系统的框图。

[0013] 图2是说明基于图块的虚拟现实 (VR) 视频部分解码的实例的概念图。

[0014] 图3是说明第一可缩放译码VR视频部分解码的概念图。

[0015] 图4是说明第二可缩放译码VR视频部分解码的概念图。

[0016] 图5是说明基于第一同时联播译码的VR视频部分解码的概念图。

[0017] 图6是说明基于第二同时联播译码的VR视频部分解码的概念图。

[0018] 图7是更详细地说明检索单元的组件的实例集合的框图。

[0019] 图8是说明实例多媒体内容的元素的概念图。

[0020] 图9是说明实例视频文件的元素的框图。

[0021] 图10是根据本发明的技术用于处理媒体数据的流程图。

[0022] 图11是根据本发明的技术用于处理媒体数据的流程图。

## 具体实施方式

[0023] 一般来说,本发明描述与虚拟现实 (VR) 相关的技术。更确切地说,本发明描述用于通过HTTP (DASH) 的动态自适应流式传输中的VR视频的用信号传送的技术。在ISO/IEC 23009-1中指定DASH的版本。本发明的技术可以应用于符合根据ISO基础媒体文件格式、可缩放视频译码 (SVC) 文件格式、高级视频译码 (AVC) 文件格式、第三代合作伙伴计划 (3GPP) 文件格式和/或多视图视频译码 (MVC) 文件格式或其它类似视频文件格式中的任一者封装的视频数据的视频文件。

[0024] 在HTTP流式传输中,经常使用的操作包含HEAD、GET和部分GET。HEAD操作检索与给定的统一资源定位符 (URL) 或统一资源名 (URN) 相关联的文件的标头,而不检索与URL或URN相关联的有效负载。GET操作检索与给定URL或URN相关联的整个文件。部分GET操作接收字节范围作为输入参数,并且检索文件的连续数目的字节,其中所述字节数目对应于接收到的字节范围。因而,可以提供电影片段以用于HTTP流式传输,因为部分GET操作可以获得一或多个单独的电影片段。在一个电影片段中,可能存在不同轨迹的若干个轨迹片段。在HTTP流式传输中,媒体呈现可以是客户端可存取的数据的经构造集合。ISO/IEC 23009-1将媒体呈现定义为建立媒体内容的有界的或无界的呈现的数据的集合。客户端可以请求和下载媒体数据信息以向用户呈现流式传输服务。

[0025] 在使用HTTP流式传输来流式传输3GPP数据的实例中,可能存在用于多媒体内容的视频和/或音频数据的多个表示。如下文所解释,不同表示可对应于不同译码特性(例如,视频译码标准的不同简档或水平)、不同译码标准或译码标准的扩展(例如,多视图及/或可缩放扩展)或不同位速率。此类表示的清单可以在媒体呈现描述(MPD)数据结构中定义。媒体呈现可以对应于HTTP流式传输客户端装置可存取的数据的经构造集合。HTTP流式传输客户端装置可以请求和下载媒体数据信息以向客户端装置的用户呈现流式传输服务。媒体呈现可在MPD数据结构中描述,MPD数据结构可包含MPD的更新。

[0026] 媒体呈现可含有一或多个周期的序列。ISO/IEC 23009-1定义媒体呈现的间隔,其中所有周期的连续序列构成媒体呈现。每个周期可延伸直到下一周期的开始为止,或者在最后一个周期的情况下直到媒体呈现的结束为止。每个周期可以含有相同媒体内容的一或多个表示。表示可以是音频、视频、定时文本或其它此类数据的多个替代的编码版本中的一个。表示可按编码类型而不同,例如,按用于视频数据的位速率、分辨率和/或编解码器和用于音频数据的位速率、语言和/或编解码器而不同。术语表示可用于指代经编码音频或视频数据的对应于多媒体内容的特定周期并且用特定方式编码的区段。

[0027] 特定周期的表示可指派给由MPD中的属性(所述属性指示表示所属的适应集合)指示的群组。ISO/IEC 23009-1将表示定义为传送格式的且与描述性元数据相关联的一或多个媒体流的集合和封装并且定义适应集合为一个或若干个媒体内容分量的可互换的经编码版本的集合。同一适应集合中的表示通常被视为彼此的替代,因为客户端装置可在这些表示之间动态地且无缝地切换,例如,以执行带宽适应。举例来说,特定周期的视频数据的每个表示可指派给同一适应集合,使得可选择这些表示中的任一者进行解码以呈现对应周期的多媒体内容的媒体数据(例如,视频数据或音频数据)。在一些实例中,一个周期内的媒体内容可以由来自群组0(如果存在的话)的一个表示来表示,或者由来自每个非零群组的至多一个表示的组合来表示。可以相对于周期的开始时间来表达所述周期的每个表示的定时数据。

[0028] 一个表示可包含一或多个片段。每个表示可包含初始化片段,或表示的每个片段可自初始化。当存在时,初始化片段可含有用于存取所述表示的初始化信息。一般来说,初始化片段不含媒体数据。片段可由识别符唯一地参考,例如,统一资源定位符(URL)、统一资源名(URN)或统一资源识别符(URI)。MPD可以提供每个片段的识别符。在一些实例中,MPD还可提供呈范围属性的形式的字节范围,所述范围属性可对应于可通过URL、URN或URI存取的文件内的片段的数据。ISO/IEC 23009-1定义片段为与HTTP-URL相关联的数据的单元且任选地通过MPD指定的字节范围并且定义初始化片段为含有呈现封装在媒体片段中的媒体流所必需的元数据的片段。ISO/IEC 23009-1定义媒体片段为符合在使用中的媒体格式的片段并且当与零或多个先前片段组合时启用回放,以及初始化片段(如果存在)。

[0029] 不同表示可选择用于不同类型的媒体数据的基本上同时的检索。举例来说,客户端装置可选择音频表示、视频表示及定时文本表示,从这些表示中检索片段。在一些实例中,客户端装置可选择特定适应集合以用于执行带宽适应。也就是说,客户端装置可选择包含视频表示的适应集合、包含音频表示的适应集合和/或包含定时文本的适应集合。替代地,客户端装置可以选择用于特定类型的媒体(例如,视频)的适应集合,并且直接地选择用于其它类型的媒体(例如,音频和/或定时文本)的表示。

[0030] 图1是说明实施用于经由网络的流式传输媒体数据的技术的实例系统10的框图。在这个实例中,系统10包含内容准备装置20、服务器装置60和客户端装置40。客户端装置40和服务器装置60通过网络74以通信方式耦合,所述网络可包括因特网。在一些实例中,内容准备装置20和服务器装置60还可通过网络74或另一网络耦合,或者可直接以通信方式耦合。在一些实例中,内容准备装置20和服务器装置60可包括相同装置。

[0031] 在图1的实例中,内容准备装置20包括音频源22和视频源24。举例来说,音频源22可包括麦克风,所述麦克风产生表示待通过音频编码器26编码的所俘获的音频数据的电信号。替代地,音频源22可包括存储媒体(其存储先前记录的音频数据)、音频数据产生器(例如,计算机化的合成器)或音频数据的任何其它来源。视频源24可包括:摄像机,其产生待通过视频编码器28编码的视频数据;存储媒体,其通过先前记录的视频数据编码;视频数据产生单元,例如,计算机图形源;或视频数据的任何其它来源。内容准备装置20未必在所有实例中都以通信方式耦合到服务器装置60,而是可以将多媒体内容存储到由服务器装置60读取的分开的媒体。

[0032] 原始音频和视频数据可包括模拟或数字数据。模拟数据在通过音频编码器26和/或视频编码器28编码之前可以被数字化。音频源22可以在说话参与者正在说话时从说话参与者获得音频数据,并且视频源24可以同时获得说话参与者的视频数据。在其它实例中,音频源22可包含包括所存储的音频数据的计算机可读存储媒体,且视频源24可包含包括所存储的视频数据的计算机可读存储媒体。以此方式,本发明中所描述的技术可应用于实况、流式传输、实时音频和视频数据或所存档的、预先记录的音频和视频数据。

[0033] 对应于视频帧的音频帧通常为含有通过音频源22俘获(或产生)的音频数据的音频帧,所述音频数据同时伴随包含于视频帧内的通过视频源24俘获(或产生)的视频数据。举例来说,当说话参与者通常通过说话而产生音频数据时,音频源22俘获音频数据,且视频源24同时(即,在音频源22正俘获音频数据的同时)俘获说话参与者的视频数据。因此,音频帧可在时间上对应于一或多个特定视频帧。相应地,对应于视频帧的音频帧通常对应于同时俘获到音频数据和视频数据且音频帧和视频帧分别包括同时俘获到的音频数据和视频数据的情况。

[0034] 在一些实例中,音频编码器26可以对每个经编码音频帧中的表示记录经编码音频帧的音频数据的时间的时戳进行编码,并且类似地,视频编码器28可以对每个经编码视频帧中的表示记录经编码视频帧的视频数据的时间的时戳进行编码。在此类实例中,对应于视频帧的音频帧可包括包含时戳的音频帧和包含相同时戳的视频帧。内容准备装置20可包含内部时钟,音频编码器26和/或视频编码器28可从所述内部时钟产生时戳,或者音频源22和视频源24可使用所述内部时钟分别使音频和视频数据与时戳相关联。

[0035] 在一些实例中,音频源22可向音频编码器26发送对应于记录音频数据的时间的数据,且视频源24可向视频编码器28发送对应于记录视频数据的时间的数据。在一些实例中,音频编码器26可对经编码音频数据中的序列识别符进行编码以指示经编码音频数据的相对时间排序,但未必指示记录音频数据的绝对时间,且类似地,视频编码器28也可使用序列识别符来指示经编码视频数据的相对时间排序。类似地,在一些实例中,序列识别符可以映射或以其它方式与时戳相关。

[0036] 音频编码器26通常产生经编码音频数据流,而视频编码器28产生经编码视频数据

流。每个单独的数据流(不论是音频还是视频)可被称为基本流。基本流是一个表示的单个经数字译码的(可能经压缩的)分量。举例来说,所述表示的经译码视频或音频部分可以是基本流。基本流可以在被封装在视频文件内之前被转换成包化基本流(PES)。在相同表示内,可以使用流ID来区分属于一个基本流的PES包与属于其它基本流的PES包。基本流的数据的基本单元是包化基本流(PES)包。因此,经译码视频数据通常对应于基本视频流。类似地,音频数据对应于一或多个相应的基本流。

[0037] 例如ITU-T H.264/AVC和高效视频译码(HEVC)标准等许多视频译码标准定义无误差位流的语法、语义和解码过程,所述无误差位流中的任一者符合特定简档或水平。视频译码标准通常并不指定编码器,但编码器具有保证所产生的位流对于解码器来说是符合标准的任务。在视频译码标准的情形下,“简档”对应于算法、特征或工具及适用于其的约束的子集。举例来说,如通过H.264标准所定义,“简档”是通过H.264标准指定的整个位流语法的子集。“水平”对应于解码器资源消耗的限制,例如,举例来说,解码器存储器和计算,其与图片分辨率、位速率和块处理速率相关。可以使用profile\_idc(简档指示符)值用信号表示简档,而可以使用level\_idc(水平指示符)值用信号表示水平。

[0038] 举例来说,H.264标准承认,在给定简档的语法所施加的界限内,仍然可能需要编码器和解码器的性能有较大变化,这取决于位流中的例如经解码图片的指定大小等语法元素所取的值。H.264标准进一步承认,在许多应用中,实施能够处理特定简档内的语法的假设使用的解码器既不实际又不经济。相应地,H.264标准将“水平”定义为强加于位流中的语法元素值的指定约束集合。这些约束可以是对值的简单的限制。替代地,这些约束可以采用值的算术组合(例如,图片宽度乘以图片高度乘以每秒解码的图片数目)的约束的形式。H.264标准进一步规定单独的实施方案对于每个所支持的简档可以支持一种不同水平。

[0039] 符合一个简档的解码器一般支持简档中定义的所有特征。举例来说,作为一个译码特征,B图片译码在H.264/AVC的基础简档中不被支持,但是在H.264/AVC的其它简档中被支持。符合一个水平的解码器应能够对任何不需要超出所述水平中定义的限制的资源的位流进行解码。简档和水平的定义可对可解释性有帮助。举例来说,在视频发射期间,可针对整个发射会话协商和同意一对简档和水平定义。更确切地说,在H.264/AVC中,水平可定义对于需要处理的宏块数目、经解码图片缓冲器(DPB)大小、经译码图片缓冲器(CPB)大小、垂直运动向量范围、每两个连续MB的运动向量的最大数目和B块是否可具有小于8x8像素的子宏块分区的限制。以此方式,解码器可确定解码器是否能够恰当地对位流进行解码。

[0040] 在图1的实例中,内容准备装置20的封装单元30从视频编码器28接收包括经译码视频数据的基本流,且从音频编码器26接收包括经译码音频数据的基本流。在一些实例中,视频编码器28和音频编码器26可以各自包含用于从经编码数据形成PES包的包化器。在其它实例中,视频编码器28和音频编码器26可各自与用于从经编码数据形成PES包的相应的包化器介接。在另外其它实例中,封装单元30可包含用于从经编码音频和视频数据形成PES包的包化器。

[0041] 视频编码器28可以用多种方式对多媒体内容的视频数据进行编码,以便以各种位速率并且用各种特性产生多媒体内容的不同表示,所述特性例如是像素分辨率、帧率、对于各种译码标准的符合性、对于各种译码标准的各种简档和/或简档水平的符合性、具有一或多个视图的表示(例如,对于二维或三维回放)或其它此类特性。如本发明中所使用,表示可

包括音频数据、视频数据、文本数据(例如,用于隐藏字幕)或其它此类数据中的一者。表示可包含例如音频基本流或视频基本流的基本流。每个PES包可包含stream\_id,其识别PES包所属于的基本流。封装单元30负责将基本流汇编成各种表示的视频文件(例如,片段)。

[0042] 封装单元30从音频编码器26和视频编码器28接收用于表示的基本流的PES包且从所述PES包形成对应的网络抽象层(NAL)单元。在H.264/AVC(高级视频译码)的实例中,经译码视频片段被组织成NAL单元,所述单元提供“网络友好”的视频表示,其处理例如视频电话、存储、广播或流式传输等的应用。NAL单元可分类为视频译码层(VCL)NAL单元和非VCL NAL单元。VCL单元可含有核心压缩引擎,且可包含块、宏块和/或切片水平数据。其它NAL单元可以是非VCL的NAL单元。在一些实例中,一个时间例子中的经译码图片(通常呈现为初级经译码图片)可包含在存取单元中,所述存取单元可包含一或多个NAL单元。

[0043] 非VCL的NAL单元可包含参数集NAL单元和SEI NAL单元等等。参数集可以含有序列水平标头信息(在序列参数集(SPS)中)和不频繁改变的图片水平标头信息(在图片参数集(PPS)中)。对于参数集(例如,PPS和SPS),不频繁改变的信息不需要对于每个序列或图片重复,因此可改进译码效率。另外,使用参数集可以实现重要标头信息的带外发射,避免了对于用于错误恢复的冗余发射的需要。在带外发射实例中,参数集NAL单元可以在与其它NAL单元(例如,SEI NAL单元)不同的信道上发射。

[0044] 此外,封装单元30可形成清单文件,例如描述表示的特性的媒体呈现描述符(MPD)。封装单元30可根据可扩展标记语言(XML)将MPD格式化。封装单元30可向输出接口32提供用于多媒体内容的一或多个表示的数据以及清单文件(例如,MPD)。输出接口32可包括网络接口或用于向存储媒体进行写入的接口,例如,通用串行总线(USB)接口、CD或DVD写入器或烧录器、到磁性或快闪存储媒体的接口,或用于存储或发射媒体数据的其它接口。封装单元30可以向输出接口32提供多媒体内容的表示中的每一个的数据,所述输出接口可以经由网络发射或存储媒体向服务器装置60发送所述数据。在图1的实例中,服务器装置60包含存储媒体62,其存储各种多媒体内容64,每个多媒体内容包含一个相应的清单文件66和一或多个表示68A-68N(表示68)。在一些实例中,输出接口32还可向网络74直接发送数据。

[0045] 在一些实例中,表示68可以分成适应集合。也就是说,表示68的各种子集可以包含特性的相应的共同集合,例如,编解码器、简档和水平、分辨率、视图数目、片段的文件格式、可以通过待解码和呈现的表示和/或音频数据识别待显示的文本的语言或其它特性的文本类型信息,例如,通过扬声器,描述用于适应集合中的表示的场景的相机角度或现实世界相机视角的相机角度信息、描述用于特定观众的内容适用性的分级信息,或类似者。

[0046] 清单文件66可包含指示对应于特定适应集合的表示68的子集以及所述适应集合的共同特性的数据。清单文件66还可包含表示适应集合的单独的表示的单独的特性(例如,位速率)的数据。以此方式,适应集合可提供用于简化的网络带宽适应。适应集合中的表示可使用清单文件66的适应集合元素的子代元素来指示。

[0047] 服务器装置60包含请求处理单元70和网络接口72。在一些实例中,服务器装置60可包含多个网络接口。另外,服务器装置60的所述特征中的任一个或全部可以在内容传递网络的其它装置(例如,路由器、桥接器、代理装置、交换器或其它装置)上实施。在一些实例中,内容传递网络的中间装置可高速缓冲存储多媒体内容64的数据,且包含基本上与服务器装置60的组件相符的组件。一般来说,网络接口72经配置以经由网络74发送及接收数据。

[0048] 请求处理单元70经配置以从客户端装置(例如,客户端装置40)接收对于存储媒体62的数据的网络请求。举例来说,请求处理单元70可以实施超文本传送协议(HTTP)版本1.1,如RFC 2616中R.Fielding等人(网络工作组,IETF,1999年6月)的“超文本传送协议-HTTP/1.1(Hypertext Transfer Protocol-HTTP/1.1,)”中所描述。也就是说,请求处理单元70可经配置以接收HTTP GET或部分GET请求,且响应于所述请求而提供多媒体内容64的数据。所述请求可例如使用表示68中的一者的片段的URL来指定所述片段。在一些实例中,所述请求还可指定所述片段的一或多个字节范围,因而包括部分GET请求。请求处理单元70可进一步经配置以服务于HTTP HEAD请求以提供表示68中的一者的片段的标头数据。在任何情况下,请求处理单元70可经配置以处理所述请求以向请求装置(例如,客户端装置40)提供所请求的数据。

[0049] 另外或替代地,请求处理单元70可经配置以经由广播或多播协议(例如,eMBMS)传递媒体数据。内容准备装置20可以与所描述的基本上相同的方式产生DASH片段和/或子片段,但服务器装置60可使用eMBMS或另一广播或多播网络传送协议传递这些片段或子片段。举例来说,请求处理单元70可经配置以从客户端装置40接收多播群组加入请求。也就是说,服务器装置60可以将与多播群组相关联的因特网协议(IP)地址播发到与特定媒体内容(例如,实况事件的广播)相关联的客户端装置,包含客户端装置40。客户端装置40继而可提交加入所述多播群组的请求。此请求可遍及网络74(例如,构成网络74的路由器)传播,使得致使所述路由器将去往与多播群组相关联的IP地址的业务导向预订的客户端装置(例如,客户端装置40)。

[0050] 如图1的实例中所说明,多媒体内容64包含清单文件66,其可以对应于媒体呈现描述(MPD)。清单文件66可以含有不同替代表示68(例如,具有不同品质的视频服务)的描述,并且所述描述可包含例如编解码器信息、简档值、水平值、位速率和表示68的其它描述性特性。客户端装置40可检索媒体呈现的MPD以确定如何存取表示68的片段。

[0051] 具体地说,检索单元52可以检索客户端装置40的配置数据(未示出)以确定视频解码器48的解码能力和视频输出44的再现能力。配置数据还可包含客户端装置40的用户选择的语言偏好中的任一个或全部、对应于客户端装置40的用户设置的深度偏好的一或多个相机视角和/或客户端装置40的用户选择的分级偏好。举例来说,检索单元52可包括网络浏览器或媒体客户端,其经配置以提交HTTP GET和部分GET请求。检索单元52可对应于客户端装置40的一或多个处理器或处理单元(未示出)执行的软件指令。在一些实例中,相对于检索单元52描述的功能性中的全部或部分可在硬件或硬件、软件和/或固件的组合中实施,其中可提供必需的硬件以执行针对软件或固件的指令。

[0052] 检索单元52可将客户端装置40的解码和再现能力与清单文件66的信息所指示的表示68的特性进行比较。检索单元52可初始地检索清单文件66的至少一部分以确定表示68的特性。举例来说,检索单元52可请求描述一或多个适应集合的特性的清单文件66的一部分。检索单元52可以选择具有可以由客户端装置40的译码和再现性能满足的特性的表示68的子集(例如,适应集合)。检索单元52可以随和确定适应集合中的表示的位速率、确定网络带宽的当前可用量和从具有可以由网络带宽满足的位速率的表示中的一个中检索片段。

[0053] 一般来说,较高位速率表示可以产生较高品质的视频回放,而在可用的网络带宽减少时较低位速率表示可以提供充足品质的视频回放。因此,当可用的网络带宽相对高时,

检索单元52可以从相对高的位速率表示中检索数据,而当可用的网络带宽较低时,检索单元52可以从相对低的位速率表示中检索数据。以此方式,客户端装置40可以经由网络74流式传输多媒体数据,同时还适应于网络74的变化的网络带宽可用性。

[0054] 另外或替代地,检索单元52可经配置以根据例如eMBMS或IP多播等广播或多播网络协议接收数据。在此类实例中,检索单元52可提交加入与特定媒体内容相关联的多播网络群组请求。在加入多播群组之后,检索单元52可在无发布到服务器装置60或内容准备装置20的进一步请求的情况下接收多播群组的数据。检索单元52可提交当不再需要多播群组的数据时离开多播群组的请求,例如,停止回放或将信道改变到不同多播群组。

[0055] 网络接口54可以接收所选表示的片段的数据并且向检索单元52提供所述数据,所述检索单元继而向解封装单元50提供所述片段。解封装单元50可将视频文件的元素解封装成组成PES流,将PES流解包化以检索经编码数据,并且取决于经编码数据是音频流还是视频流的一部分(例如,如通过流的PES包标头所指示)而向音频解码器46或视频解码器48发送经编码数据。音频解码器46对经编码音频数据进行解码,并且向音频输出42发送经解码音频数据,而视频解码器48对经编码视频数据进行解码,并且向视频输出44发送经解码视频数据,其可包含流的多个视图。

[0056] 视频编码器28、视频解码器48、音频编码器26、音频解码器46、封装单元30、检索单元52和解封装单元50各自可以实施为适用的多种合适的处理电路中的任一个,所述处理电路例如是一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑电路、软件、硬件、固件或其任何组合。视频编码器28和视频解码器48中的每一个可以包含在一或多个编码器或解码器中,所述编码器或解码器中的任一个可以集成为组合视频编码器/解码器(编解码器)的一部分。类似地,音频编码器26和音频解码器46中的每一个可包含在一或多个编码器或解码器中,所述编码器或解码器中的任一个可集成为组合编解码器的一部分。包含视频编码器28、视频解码器48、音频编码器26、音频解码器46、封装单元30、检索单元52和/或解封装单元50的设备可包括集成电路、微处理器和/或无线通信装置,例如,蜂窝式电话。

[0057] 客户端装置40、服务器装置60和/或内容准备装置20可经配置以根据本发明的技术操作。出于实例的目的,本发明相对于客户端装置40和服务器装置60描述这些技术。然而,应理解,代替服务器装置60(或除此之外),内容准备装置20可经配置以执行这些技术。

[0058] 封装单元30可形成NAL单元,所述NAL单元包括识别NAL单元所属的程序的标头,以及有效负载,例如,音频数据、视频数据或描述NAL单元对应于的传送或程序流的数据。举例来说,在H.264/AVC中,NAL单元包含1-字节标头和变化大小的有效负载。在其有效负载中包含视频数据的NAL单元可包括各种粒度水平的视频数据。举例来说,NAL单元可包括视频数据块、多个块、视频数据的切片,或视频数据的整个图片。封装单元30可以从视频编码器28接收呈基本流的PES包的形式的经编码视频数据。封装单元30可使每个基本流与对应的程序相关联。

[0059] 封装单元30还可汇编来自多个NAL单元的存取单元。一般来说,存取单元可包括用于表示视频数据帧以及对应于所述帧的音频数据(当此类音频数据可用时)的一或多个NAL单元。存取单元通常包含用于一个输出时间例子的所有NAL单元,例如,用于一个时间例子的所有音频和视频数据。举例来说,如果每个视图具有20帧每秒(fps)的帧率,那么每个时

间例子可以对应于0.05秒的时间间隔。在此时间间隔期间,用于相同存取单元的所有视图的特定帧(相同时间例子)可同时再现。在一个实例中,存取单元可包括一个时间例子中的经译码图片,其可呈现为初级经译码图片。

[0060] 相应地,存取单元可包括共同时间例子的所有音频和视频帧,例如,对应于时间X的所有视图。本发明还将特定视图的经编码图片称为“视图分量”。也就是说,视图分量可包括在特定时间针对特定视图的经编码图片(或帧)。相应地,存取单元可被定义为包括共同时间例子的所有视图分量。存取单元的解码次序不需要一定与输出或显示次序相同。

[0061] 媒体呈现可包含媒体呈现描述(MPD),其可含有不同替代表示(例如,具有不同品质的视频服务)的描述,且所述描述可包含例如编解码器信息、简档值和水平值。MPD为例如清单文件66等清单文件的一个实例。客户端装置40可检索媒体呈现的MPD以确定如何存取各种呈现的电影片段。电影片段可位于视频文件的电影片段框(moof框)中。

[0062] 清单文件66(其可包括例如MPD)可播发表示68的片段的可用性。也就是说,MPD可包含指示表示68中的一个的第一片段变得可用时的挂钟时间的信息,以及指示表示68内的片段的持续时间的信息。以此方式,客户端装置40的检索单元52可基于开始时间以及在特定片段之前的片段的持续时间而确定何时每一片段可用。

[0063] 在封装单元30已基于所接收数据将NAL单元和/或存取单元汇编到视频文件中之后,封装单元30将视频文件传递到输出接口32以用于输出。在一些实例中,封装单元30可将视频文件存储在本地,或经由输出接口32将视频文件发送到远程服务器,而非将视频文件直接发送到客户端装置40。输出接口32可包括例如发射器、收发器、用于将数据写入到计算机可读媒体的装置,例如,光盘驱动器、磁性媒体驱动器(例如,软盘驱动器)、通用串行总线(USB)端口、网络接口或其它输出接口。输出接口32将视频文件输出到计算机可读媒体,例如,发射信号、磁性媒体、光学媒体、存储器、快闪驱动器,或其它计算机可读媒体。

[0064] 网络接口54可经由网络74接收NAL单元或存取单元,且经由检索单元52将NAL单元或存取单元提供到解封装单元50。解封装单元50可将视频文件的元素解封装成组成PES流,将PES流解包化以检索经编码数据,且取决于经编码数据是音频流还是视频流的一部分(例如,如通过流的PES包标头所指示)而向音频解码器46或视频解码器48发送经编码数据。音频解码器46对经编码音频数据进行解码,并且向音频输出42发送经解码音频数据,而视频解码器48对经编码视频数据进行解码,并且向视频输出44发送经解码视频数据,其可包含流的多个视图。

[0065] 视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1 Visual、ITU-T H.262或ISO/IEC MPEG-2 Visual、ITU-T H.263、ISO/IEC MPEG-4 Visual、ITU-T H.264或ISO/IEC MPEG-4 AVC,包含其可缩放视频译码(SVC)和多视图视频译码(MVC)扩展,以及高效视频译码(HEVC),也被称为ITU-T H.265和ISO/IEC 23008-2,包含其可缩放译码扩展(即,可缩放高效视频译码,SHVC)、多视图扩展(即,多视图高效视频译码,MV-HEVC),以及3D扩展(即,3D高效视频译码,3D-HEVC)。

[0066] 现在将简要地论述文件格式和文件格式标准。文件格式标准包含ISO基础媒体文件格式(ISOBMFF,ISO/IEC 14496-12,下文中,“ISO/IEC 14996-12”),以及从ISOBMFF中导出的其它文件格式标准,包含MPEG-4文件格式(ISO/IEC 14496-15)、3GPP文件格式(3GPP TS 26.244)和含有用于AVC的文件格式的ISO/IEC 14496-15(ISO/IEC 14496-15,下文中

“ISO/IEC 14496-15”)及其扩展以及用于HEVC的文件格式及其扩展。因此,ISO/IEC 14496-12指定ISO基础媒体文件格式。其它文档扩展ISO基础媒体文件格式以用于特定应用程序。举例来说,ISO/IEC 14496-15描述ISO基础媒体文件格式中的NAL单元构造视频的运载。H.264/AVC和HEVC以及它们的扩展是NAL单元构造视频的实例。ISO/IEC 14496-15包含描述H.264/AVC NAL单元的运载的章节。另外,ISO/IEC 14496-15的章节8描述HEVC NAL单元的运载。因此,ISO/IEC 14496-15的章节8被称为描述HEVC文件格式。在第114次MPEG会议之后,基于从若干国家主体接收到的评论,准备了含有将适用于ISO/IEC 14496-15草案规范的新版本的ISO/IEC 14496-15的一些变化的处置文档。此处置文档被称作“MPEG输出文档N15297”。

[0067] ISOBMFF被用作许多编解码器封装格式(例如,AVC文件格式)以及许多多媒体容器格式(例如,MPEG-4文件格式、3GPP文件格式(3GP)和DVB文件格式)的基础。除连续媒体(例如,音频和视频)、静态媒体(例如,图像)之外,元数据同样可以存储在符合ISOBMFF的文件中。根据ISOBMFF构造的文件可用于许多目的,其包含本地媒体文件回放、远程文件的即看式下载、通过HTTP的动态自适应流式传输(DASH)的片段、待流式传输的内容的容器及其包化指令,以及记录接收到的实时媒体流。因此,尽管ISOBMFF最初设计用于存储,但是已经证实ISOBMFF对于流式传输(例如,即看式下载或DASH)而言是有价值的。出于流式传输的目的,可以使用在ISOBMFF中定义的电影片段。除连续媒体(例如,音频和视频)、静态媒体(例如,图像)之外,元数据同样可以存储在符合ISOBMFF的文件中。

[0068] 符合HEVC文件格式的文件可以包括一系列对象,被称作框。“框”可以是唯一类型识别符和长度定义的面向对象的构建块。框是在ISOBMFF中的基础语法结构,其包含四字符经译码的框类型、框的字节数和有效负载。换句话说,框可以是包括经译码框类型、框的字节数和有效负载的语法结构。在一些例子中,符合HEVC文件格式的文件中的所有数据可以含在框内并且在不在框中的文件中可能不存在数据。因此,ISOBMFF文件由一系列框构成,并且框可以含有其它框。举例来说,框的有效负载可以包含一或多个额外的框。在本发明中的其它处详细地描述的图9根据本发明的一或多种技术示出了文件内的实例框。

[0069] 符合ISOBMFF的文件可以包含各种类型的框。举例来说,符合ISOBMFF的文件可以包含文件类型框、媒体数据框、电影框、电影片段框等等。在此实例中,文件类型框包含文件类型和兼容性信息。媒体数据框可含有样本(例如,经译码图片)。电影框(“moov”)含有用于存在于文件中的连续媒体流的元数据。连续媒体流中的每一个可以在文件中表示为轨迹。举例来说,电影框可含有关于电影的元数据(例如,样本之间的逻辑和定时关系,以及到样本的位置的指针)。电影框可以包含若干类型的子框。电影框中的子框可以包含一或多个轨迹框。轨迹框可以包含关于电影的单独的轨迹的信息。轨迹框可以包含指定单个轨迹的整体信息的轨迹标头框。另外,轨迹框可以包含含有媒体信息框的媒体框。媒体信息框可以包含含有轨迹中的数据编索引媒体样本的样本表框。样本表框中的信息可用于及时定位样本,并且用于将轨迹、类型、大小、容器和偏移的样本中的每一个定位到样本的容器中。因此,用于轨迹的元数据封闭在轨迹框(“trak”)中,而轨迹的媒体内容或者封闭在媒体数据框(“mdat”)中或者直接地封闭在单独的文件中。轨迹的媒体内容包括或其组成为一系列样本,例如,音频或视频存取单元。

[0070] ISOBMFF指定以下类型的轨迹:媒体轨迹(含有基础媒体流)、提示轨迹(包含媒体

发射指令或者表示接收到的包流)以及定时元数据轨迹(包括时间同步的元数据)。用于每个轨迹的元数据包含样本描述条目的列表,每一个提供译码或封装格式,所述格式在轨迹和用于处理所述格式的初始化数据中使用。每个样本与轨迹的样本描述条目中的一个相关联。

[0071] ISOBMFF实现了以各种机制指定样本特定的元数据。已经标准化在样本表框(“stbl”)内的特定框以响应于共同需要。举例来说,同步样本框(“stss”)用于列出轨迹的随机存取样本。样本分组机制实现根据四字符分组类型将样本映射到共享相同性质的样本的群组中,所述性质被指定为文件中的样本群组描述条目。在ISOBMFF中已经指定了若干分组类型。样本表框含有样本表,所述样本表含有轨迹中的媒体样本的所有时间和数据编索引。使用样本表框中的表,可有可能及时定位样本、确定它们的类型(例如,I帧或不是),并且确定它们的大小、容器和到容器中的偏移。

[0072] 举例来说,同步样本框(“stss”)是样本表框内的框。同步样本框用于列出轨迹的随机存取样本。本发明可将通过同步样本框列出的样本称为同步样本。在另一实例中,样本分组机制实现根据四字符分组类型将样本映射到共享相同性质的样本的群组中,所述性质被指定为文件中的样本群组描述条目。在ISOBMFF中已经指定了若干分组类型。

[0073] ISOBMFF规范指定与DASH一起使用的流存取点(SAP)的六种类型。前两种SAP类型(类型1和2)对应于在H.264/AVC和HEVC中的IDR图片。第三SAP类型(类型3)对应于开启GOP随机存取点,因此对应于在HEVC中的BLA或CRA图片。第四SAP类型(类型4)对应于GDR随机存取点。

[0074] 电影片段框是顶部水平框。每个电影片段框提供先前应已经在电影框中的信息。电影片段框可含有一或多个轨迹片段(“traf”)框。在电影片段内存在轨迹片段(每轨迹零或多个)的集合。轨迹片段继而含有零或多个轨迹运行,其中的每一个用文档记录用于该轨迹的样本的连续运行。举例来说,每个轨迹运行可含有按特定次序(例如,解码次序)连续的图片的样本。轨迹片段框在14996-12规范中定义并且包括用于一或多个轨迹片段的元数据。举例来说,轨迹片段框可以包含指示轨迹ID、基础数据偏移、样本描述索引、默认样本持续时间、默认样本大小和默认样本旗标的轨迹片段标头框。轨迹片段框可以包含一或多个轨迹片段运行框,每一个记录用于轨迹的样本的连续集合。举例来说,轨迹片段框可以包含指示样本计数、数据偏移、样本旗标、样本持续时间、样本大小、样本组成时间偏移等等的语法元素。在这些结构内,许多字段是任选的并且可以是默认的。

[0075] 样本表框可以包含一或多个SampleToGroup框和一或多个样本群组描述框(即,SampleGroupDescription框)。SampleToGroup框可用于确定样本所属的样本群组,连同所述样本群组的相关联的描述。换句话说,SampleToGroup框可以指示样本所属的群组。SampleToGroup框可能具有“sbgp”的框类型。SampleToGroup框可以包含分组类型元素(例如,grouping\_type)。在一些例子中,在本发明中,框的元素也可被称作语法元素。分组类型元素可以是识别样本分组的类型(即,用于形成样本群组的标准)的整数。此外,SampleToGroup框可以包含一或多个条目(即,样本群组条目)。SampleToGroup框中的每个样本群组条目可以与轨迹中的连续样本的不同的非重叠的系列相关联。每个样本群组条目可以指示样本计数元素(例如,sample\_count)和群组描述索引元素(例如,group\_description\_index)。样本群组条目的样本计数元素可以指示与样本群组条目相关联的样

本的数目。换句话说,样本群组条目的样本计数元素可以是产生具有相同样本群组描述符的连续样本的数目的整数。在SampleGroupDescription框内,群组描述索引元素可以识别含有与样本群组条目相关联的样本的描述的群组描述条目。多个样本群组条目的群组描述索引元素可以识别相同SampleGroupDescription框。

[0076] 在ISO/IEC 23009-1中指定的通过HTTP (DASH) 的动态自适应流式传输是HTTP (自适应) 流式传输应用的标准。DASH主要指定媒体呈现描述 (MPD) (也被称为清单) 的格式和媒体片段格式。MPD描述在服务器上可用的媒体并且让DASH客户端在它感兴趣的媒体时间处自主地下载媒体版本。

[0077] DASH是基于层级数据模型的。呈现是通过MPD文档描述的,所述MPD文档及时描述组成媒体呈现的周期的序列。周期通常表示媒体内容周期,在此期间媒体内容的经编码版本的恒定的集合是可用的,例如,可用的位速率、语言、字幕、副标题等的集合在周期期间并不改变。

[0078] 在周期内,材料被布置到适应集合中。适应集合表示一个或若干个媒体内容分量的可互换的经编码版本的集合。举例来说,可能存在用于主视频分量的一个适应集合和用于主音频分量的单独的适应集合。其它可用的材料(例如,字幕或音频描述)可以各自具有单独的适应集合。材料也可以提供于多路复用形式中,在此情况下多路复用的可互换的版本可以被描述为单个适应集合,例如,在一周期含有主音频和主视频两者的适应集合。多路复用分量中的每一个可以通过媒体内容分量描述单独地描述。

[0079] 适应集合含有表示的集合。表示描述一个或若干个媒体内容分量的可传递经编码版本。表示包含一或多个媒体流(用于多路复用中的每个媒体内容分量的一个)。适应集合内的任何单个表示足以再现包含的媒体内容分量。通过收集一个适应集合中的不同表示,媒体呈现作者表达表示在感知上表示等效内容。通常,这意味着客户端可以在适应集合内的表示之间动态地切换以便适应于网络状况或其它因数。切换是指直到特定时间 $t$ 的解码数据的呈现,以及从时间 $t$ 向前的另一表示的解码数据的呈现。如果表示包含在一个适应集合中,并且客户端恰当地切换,那么媒体呈现预期为跨越切换顺畅地被察觉。客户端可以忽略依赖于编解码器或它们并不支持或以其它方式不合适的其它再现技术的表示。在表示内,内容可以被及时分割成片段以用于恰当地可存取性和传递。为了存取片段,提供URL用于每个片段。因此,片段是可以通过单个HTTP请求检索的数据的最大单元。

[0080] 用于基于DASH的HTTP流式传输的典型过程包含以下步骤:

[0081] 1) 客户端获得流式传输内容(例如,电影)的MPD。MPD包含关于流式传输内容的不同的替代表示的信息,例如,位速率、视频分辨率、帧率、音频语言,以及HTTP资源(初始化片段和媒体片段)的URL。

[0082] 2) 基于MPD中的信息和客户端的本地信息,例如,网络带宽、解码/显示能力和用户偏好,客户端每次向所期望的表示请求一个片段(或其部分)。

[0083] 3) 当客户端检测到网络带宽改变时,它向不同表示的片段请求更好地匹配的位速率,理想地从开始于随机存取点的片段开始。

[0084] 在HTTP流式传输“会话”以响应于用户请求以向后搜寻到过去位置或向前搜寻到将来位置期间,客户端请求从接近于理想位置且理想地开始于随机存取点的片段开始的过去或将来片段。用户也可以请求快进内容,这可以通过请求足以用于解码仅经帧内译码视

频图片或仅视频流的时间子集的数据实现。

[0085] 虚拟现实 (VR) 是实际上存在于非物质世界的能力,所述非物质世界由通过允许与该世界交互的沉浸用户的动作相关的天然和/或合成图像和声音的再现产生。通过在再现装置中作出的最新进展,例如,头戴式显示器 (HMD) 和VR视频 (通常也被称作360度视频) 创建,可以提供显著品质的体验。VR应用包含游戏、训练、教育、体育视频、在线购物、成人娱乐等等。

[0086] 典型的VR系统可以包含以下组件和步骤:

[0087] a. 相机集合,其通常由指向不同方向且理想地一起覆盖围绕相机集合的所有视点的多个单独的相机组成。

[0088] b. 图像拼接,其中通过多个单独的相机拍摄的视频图片在时域中同步并且在空间中经拼接,以成为球面视频,但是映射到矩形格式,例如,相等矩形 (如同世界地图) 或立方体地图。

[0089] c. 映射的矩形格式的视频是使用视频编解码器经编码/压缩的,例如,H.265/HEVC 或H.264/AVC。

[0090] d. 经压缩视频位流可以存储和/或封装在媒体格式中并且通过网络发射 (可能仅覆盖仅用户看到的区域的子集) 到接收器。

[0091] e. 接收器接收视频位流或其部分,可能封装在一种格式中,并且将经解码视频信号或其部分发送到再现装置。

[0092] f. 再现装置可以是例如HMD,其可以追踪头部移动并且甚至眼睛移动时刻并且再现视频的对应的部分使得沉浸经历传递到用户。

[0093] VR视频与普通视频相比的不同特征在于VR通常仅是所显示的由视频图片表示的整个视频区的子集,对应于当前视野 (FOV),即,用户当前看到的区域,而在普通视频应用程序中通常显示整个视频区。FOV有时也被称作视口。可以利用此特征以改进VR视频系统的性能,例如,通过使用视口依赖性投影映射或视口依赖性视频译码。在呈现给用户的视频部分的相同分辨率/品质下,与常规的VR视频系统相比性能改进可以是较低发射带宽和较低解码复杂度中的任一个或两个。

[0094] 视口依赖性投影映射也可被称作不对称投影映射。一个实例是对立方体地图进行二次采样。典型的立方体地图由六个相等大小的表面组成。在对立方体地图进行二次采样的一个实例中,表面中的一个可以保持不变,而在相对侧上的表面可以进行二次采样或向下缩放到位于原始表面形状的中心区域的较小大小,并且其它表面相应地在几何学上按比例缩放 (仍然保持六个表面)。极端情况是向下按比例缩放相对侧上的表面为单个点,并且因此立方体变为角锥。在对立方体地图进行二次采样的另一实例中,一些表面是向下按比例地设定大小的,例如,通过2x2比率 (即,在平行于表面的边缘的每个方向上是2:1)。

[0095] 此类向下设定大小按比例缩放也可以适用于其它投影映射的不同区,例如,相等矩形。一个实例是缩小上部和下部区 (即,极点)。

[0096] 视口依赖性视频译码也可被称作视口依赖性部分视频解码,因为关键是实现解码仅部分地整个经编码视频区以提供足够信息以供显示当前FOV或视口。

[0097] 常规地,VR视频通过相等矩形或立方体地图保护映射来表示。视频图片是使用时间帧间预测 (TIP) 经编码为单层位流的,整个经译码位流存储在服务器处,如果所需的话,

那么被发射到接收器侧,完全地通过解码器来解码,并且对应于当前视口的经解码图片的区被再现给用户。

[0098] VR视频图片可以使用运动受约束图块经译码,使得覆盖视口的每个潜在区可以跨越时间独立于其它区解码。对于特定当前视口,覆盖视口的图块的最小集合被发送到客户端,经解码并且经再现。此方法被称作基于简单图块的部分解码(STPD),并且通过图2来描绘。

[0099] 此方法的一个问题在于当用户快速地转动他或她的头部到未由当前正被发送的图块覆盖的(完全地或部分地)新视口时,在覆盖新视口的图块到达之前在新的区域(由新视口而非旧视口覆盖)中什么也看不到(并且数据是根据缓冲时间表充分地缓冲的)。因此,此方法可以仅在网络往返时间极其低的情况下起作用,例如,在10ms的大小,这并非是可行的或者至少是在当今或在近期内的巨大挑战。

[0100] 在本章节中呈现两个基于可缩放译码的部分解码(SLPD)方案,它们被称作SLPD#1和SLPD#2。在SLPD#1中,如通过图3所描绘,VR视频是使用具有多分辨率的SHVC(或另一类型的可缩放视频译码)空间可缩放性经可缩放译码的。最低分辨率视频的位流,即,基底层(BL)总是完全地发送,使得对于任何视口在任何时间至少最低分辨率视频是可供用于再现的。最低分辨率视频根本不需要使用图块进行译码,虽然如果它是使用图块或运动受约束图块经译码的,那么它同样起作用。

[0101] 增强层(EL)是使用运动受约束图块经译码的,使得覆盖视口的每个潜在区可以跨越时间与其它区独立地解码,具有启用的层间预测(ILP)。对于特定当前视口,覆盖视口的图块的最小集合被发送到客户端,经解码并且经再现。从存储的观点来看,需要存储所有分辨率/层的完整流。

[0102] 当两个以上层经译码时,覆盖不同视口的图块可以选自不同分辨率。对于当前视口,图块是选自最高分辨率的;对于与当前视口相邻的视口,图块是选自第二最高分辨率的;并且以此类推。

[0103] 在SLPD#2中,如通过图4所描绘,VR视频同样是使用具有多分辨率的SHVC空间可缩放性经可缩放译码的。BL是与在SLPD#1中相同的经译码的,而EL是与在SLPD#1中类似地经译码的但是时间帧间预测(TIP)被停用。

[0104] 在本章节中呈现两个基于同时联播译码的部分解码(SMPD)方案,它们被称作SMPD#1和SMPD#2。SMPD#1是通过图5描绘的。此方法与SLPD#1相同,不同之处在于并未使用层间预测。换句话说,除了最低分辨率之外,所有其它分辨率是按照与STPD中的相同方式经译码的。SMPD#2是通过图6描绘的。在SMPD#2与SMPD#1之间存在两个差异:在SMPD#2中,最低分辨率同样使用与其它分辨率相同的运动受约束图块经译码的。当使用SMPD#2时,最低分辨率同样未被完全地发送,而是仅覆盖几何学上距离当前视口最远的视口的图块以及未由其它分辨率发送图块覆盖的任何其它视口。

[0105] 截至今天,仍然缺乏支持DASH中的VR信令的设计。举例来说,为了使基于DASH的VR流式传输系统很好地起作用,本发明提出了信令以指示所有视频表示或一些视频表示含有VR视频而不是普通视频(即,传统地非VR视频)。在一个实例中,在并不支持VR视频再现的传统DASH客户端并不尝试消耗VR视频表示的意义上信令机构是后向兼容的。对于另一实例,作为VR视频的基础部分的所使用的投影映射可以有效地用信号传送的。此外,当基于运

动受约束图块的视口依赖性VR视频译码方案在使用中时,在DASH处的一些信令可用于使得DASH客户端能够最初(基于客户端能力和用户偏好)以及在用于流适应(于改变的网络状况等)的流式传输会话期间两者处选择特定表示。

[0106] 本发明描述了用于DASH中的VR视频的用信号传送的若干技术。这些技术中的一或多个可以独立地应用或与其它技术组合应用。所述技术可以相对于在不同水平处用信号传送的各种数据来描述。如上文所介绍,从最高到最低,此类水平包含文档水平、MPD水平、周期水平、适应集合水平和表示水平。每个水平包含具有唯一语义(例如,元素和属性的唯一组合)的数据结构。在此情形下,属性可指代单个值,而元素指代一或多个属性的集合。元素还可包含额外的元素。

[0107] 提出MPD水平(例如,通过MPD水平元素或属性)上的信令以指示以下各项中的一或多个:

[0108] a. 媒体呈现的所有视频表示含有VR视频。对于以后向兼容方式起作用的此情况,定义专门地用于VR视频的新DASH简档。对于符合此简档的媒体呈现,媒体呈现的所有视频表示将含有VR视频。因此,传统DASH客户端可经配置以根本不尝试存取媒体呈现。

[0109] b. 存在媒体呈现的视频表示中的至少一个含有VR视频并且存在媒体呈现的视频表示中的至少一个含有普通视频。在此情况下,可使用允许普通视频表示的现有DASH简档。然而,对于含有VR视频的所有表示或含有VR视频的每个表示的每个视频适应集合应该随后通过新DASH简档或某些其它指示用信号传送以防止传统DASH客户端尝试使用含有VR视频的任何视频表示。

[0110] 举例来说,例如内容准备装置20的装置可以在媒体呈现的MPD数据结构中用信号传送指示媒体呈现中的至少一个视频表示含有VR视频的数据,并且所述装置可以发送媒体呈现中的至少一个视频表示的片段。例如客户端装置40或服务器装置60的装置可以至少部分地基于在媒体呈现的MPD数据结构中用信号传送的数据确定媒体呈现中的至少一个视频表示含有VR视频,并且检索媒体呈现中的至少一个视频表示的片段。为了检索片段,所述装置可以至少部分地基于经配置以解码VR视频的装置检索媒体呈现中的至少一个视频表示的片段。例如客户端装置40的装置可以解码媒体呈现中的至少一个视频表示。例如服务器装置60的装置可以将至少一个视频表示的片段发射到另一装置,例如,客户端装置40。在一些实例中,为了确定媒体呈现中的至少一个视频表示含有VR视频,所述装置可以基于在MPD数据结构中用信号传送的数据确定媒体呈现中的所有视频表示含有VR视频。

[0111] 举例来说,数据可以是MPD水平元素或MPD水平属性。媒体呈现可以包含含有媒体呈现的一或多个视频表示的视频适应集合。视频适应集合可以符合用于VR视频的特定DASH简档,其中媒体呈现中的所有视频表示符合含有特定DASH简档的VR视频。媒体呈现中的至少一个视频表示也可以含有非VR视频。

[0112] 所使用的投影映射可以是在周期水平(例如,通过周期水平元素或属性)和适应集合水平(例如,通过适应集合水平元素或属性)两者上用信号传送的,但是不在表示水平上以阻止用于VR视频表示与非VR视频表示之间的流适应的表示切换。

[0113] a. 当周期水平信令存在时,适应集合水平上的信令是任选的,并且在此情况下如果信令同样存在用于适应集合,那么它将不与周期水平信令相抵触。

[0114] b. 当周期水平信令不存在时,那么适应集合水平信令是必选的(即,必须被呈现)。

[0115] 举例来说,例如内容准备装置20的装置可以在媒体呈现描述的周期水平或媒体呈现描述的适应集合水平中的至少一个中发送数据,并且所述数据可以指示在媒体呈现中使用的投影映射。所述装置可以发送媒体呈现的视频表示的片段。例如客户端装置40或服务器装置60的装置可以检索媒体数据,方法是至少部分地基于在媒体呈现描述的周期水平或媒体呈现描述的适应集合水平中的至少一个中用信号传送的数据确定在媒体呈现中使用的投影映射并且基于所述确定检索媒体呈现的视频表示的片段。所述装置还可以至少部分地基于在媒体呈现中使用的投影映射确定是否检索视频表示的片段。所述装置还可以至少部分地基于在媒体呈现中使用的投影映射确定如何解码视频表示的片段。用信号传送投影映射的数据并非在媒体呈现的表示水平处用信号传送。

[0116] 在一些实例中,当用信号传送投影映射的数据呈现于媒体呈现的周期水平处时,在媒体呈现的适应集合水平处的投影映射的用信号传送可以是任选的。在一些实例中,当用信号传送投影映射的数据呈现于媒体呈现的适应集合水平处时,在媒体呈现的周期水平处的投影映射的用信号传送可以是必选的。用信号传送投影映射的数据可能并非在媒体呈现的表示水平处用信号传送。

[0117] 适应集合水平信令(例如,通过适应集合水平元素或属性)用于指示基于运动受约束的图块的视口依赖性VR视频译码方案是否在使用中,并且如果是的话,那么是哪一个。举例来说,可以指定可使用如在章节2.5.2中描述的三个方法STPD、SLPD#1和SMPD#2中的一个,并且属性用于指示这三个中的哪一个在使用中,例如,值1、2或3分别指示STPD、SLPD#1或SMPD#2在使用。

[0118] 举例来说,例如内容准备装置20在装置可以在用于媒体呈现的媒体呈现描述的适应集合水平处发送指示基于运动受约束图块的视口依赖性VR视频译码方案在媒体呈现中在使用中的数据并且发送媒体呈现的片段。数据可以是第一数据,所述装置还可以发送在媒体呈现的适应集合水平处用信号传送的指示哪个基于运动受约束图块的视口依赖性VR译码方案在媒体呈现中在使用中的第二数据。例如客户端装置40或服务器装置60的装置可以基于在用于媒体呈现的媒体呈现描述的适应集合水平处用信号传送的第一数据确定基于运动受约束图块的视口依赖性VR视频译码方案是否在媒体呈现中在使用中,并且所述装置可以检索媒体呈现的片段。响应于确定基于运动受约束图块的视口依赖性VR译码方案在媒体呈现中在使用中,所述装置可以基于在媒体呈现的适应集合水平处用信号传送的第二数据确定哪个基于运动受约束图块的视口依赖性VR译码方案在媒体呈现中在使用中。

[0119] 根据基于第一运动受约束图块的视口依赖性VR译码方案,所述装置可以使用运动受约束图块编码或解码VR视频图片,使得覆盖视口的每个潜在区跨越时间可独立于其它区解码。根据基于第二运动受约束图块的视口依赖性VR译码方案,(1) VR视频可以是使用具有多分辨率的可缩放视频译码空间可缩放性经可缩放译码的;(2) 基底层可以是完全地发送的,使得在任何时间对于任何视口至少最低分辨率视频是可供用于再现的;(3) 增强层(EL)是使用运动受约束图块经译码的使得覆盖视口的每个潜在区可以跨越时间独立于其它区解码,其中启用层间预测;以及(4) 时间帧间预测在基底层和E1中启用。根据基于第三运动受约束图块的视口依赖性VR译码方案,(1) VR视频可以是使用具有多分辨率的可缩放视频译码空间可缩放性经可缩放译码的;(2) 基底层可以是完全地发送的,使得在任何时间对于任何视口至少最低分辨率视频是可供用于再现的;(3) EL可以是使用运动受约束图块经译

码的使得覆盖视口的每个潜在区可以跨越时间独立于其它区解码,其中启用层间预测;以及(4)时间帧间预测可以在基底层中启用但是在EL中停用。在媒体呈现的适应集合水平处用信号传送的第二数据可以指示基于第一、第二或第三运动受约束图块的视口依赖性VR译码方案中的哪一个在媒体呈现中在使用中。

[0120] 在文件格式水平上的基于图块的视口依赖性部分VR视频编码和解码方案的用信号传送可以使用ISO/IEC 14496-15的条款10中的HEVC和分层HEVC(L-HEVC)的图块化存储。在一些实例中,使用图块轨迹,每个运载一个运动受约束图块或图块区,以避免当使用图块区样本群组映射时使用许多字节范围用于来自特定DASH表示(对应于一个轨迹)的覆盖视口的图块的请求。

[0121] 举例来说,例如内容准备装置20的装置可以发送存储在含有媒体呈现的文件的文件格式水平处的指示基于图块的视口依赖性VR视频编码和解码方案的数据并且发送媒体呈现的片段。例如客户端装置40或服务器装置60的装置可以基于存储在含有媒体呈现的文件的文件格式水平处的数据确定基于图块的视口依赖性VR视频编码和解码方案,并且检索媒体呈现的片段。所述文件可以例如包含多个轨迹,其中所述多个轨迹的每个相应的轨迹含有运动受约束图块或图块区。

[0122] ISO/IEC 23008-2视频提供对被称作图块的矩形区的译码的支持。HEVC图块并不具有与相同经译码图片中的其它HEVC图块的译码依赖性,但是可能具有与来自先前经译码图片的其它HEVC图块的译码依赖性或可独立地解码。在ISO/IEC 14496-15的条款10中,矩形图块区被定义为形成在不含有其它HEVC图块的且可以不必是按解码次序连续的一或多个切片中经编码的矩形区的任何整数数目的HEVC图块。在ISO/IEC14496-15的条款10中,不受约束的图块区被定义为由一或多个完整HEVC图块组成的任何数目的完整切片并且可以不必是按解码次序连续的。在ISO/IEC 14496-15的条款10中,不受约束的图块区也可以用于表示跨越若干HEVC图块的关注区。

[0123] 此外,如在ISO/IEC 14496-15的子条款10.6中所描述,

[0124] 存在其中在不同轨迹中存储可独立解码的HEVC(相应地L-HEVC)图块可以是可用于快速空间和时间存取视频内容的情况。对于此类情况,轨迹可以是使用HEVCTileSampleEntry(相应地LHEVCTileSampleEntry)样本描述格式创建的。

[0125] HEVC(相应地L-HEVC)图块轨迹是其中存在对运载图块所属于的相关联的HEVC层的NALU的HEVC(相应地L-HEVC)轨迹的“tbas”参考的视频轨迹。HEVC图块轨迹的样本描述类型将为“hvt1”。L-HEVC图块轨迹的样本描述类型将为“lht1”。

[0126] 图块轨迹中的样本或样本描述框将都不含有VPS、SPS或PPS NAL单元,这些NAL单元将在含有相关联的层的轨迹的样本中或样本描述框中,如由“tbas”轨迹参考所识别。如由“tbas”轨迹参考所指示的HEVC/L-HEVC图块轨迹和含有相关联的层的轨迹两者均可使用如在A.7中所定义的提取器来指示如何重构原始位流;在一些应用域中可约束这些轨迹中的提取器的存在。

[0127] 存储在图块轨迹中的HEVC或L-HEVC样本是用于一或多个图块的切片的完整集合,如在ISO/IEC 23008-2中所定义的。通常,如果轨迹由单个HEVC图块组成,那么仅用于译码此图块的切片将在样本中被发现。图块轨迹通常包含一个RectTileRegGroupEntry(单图块轨迹),或一个UnconstrTileRegGroupEntry和一或多个依赖性RectTileRegGroupEntry,这

个不受约束的图块区由(多图块轨迹)制成。

[0128] 存储在图块轨迹中的HEVC样本被视为同步样本,前提是样本中的VCL NAL单元是解码及时刷新(IDR)NAL单元、清洁随机存取(CRA)NAL单元或断链存取(BLA)NAL单元。

[0129] 存储在图块轨迹中的L-HEVC样本被视为同步样本,前提是样本中的VCL NAL单元是IRAP NAL单元并且对应的存取单元不具有RASL图片。

[0130] 定义用于常规HEVC(相应地L-HEVC)样本的子样本和样本分组具有用于HEVC(相应地L-HEVC)图块样本的相同定义。

[0131] 注意1:实施方案可以决定仅解码HEVC序列的完整图块的子集。在此情况下,可以使用RectTileRegGroupEntry和UnconstrTileRegGroupEntry样本群组描述中的图块依赖性信息以在解码HEVC序列时抛弃不需要的轨迹或忽略一些提取器。

[0132] 注意2:当仅需要由特定HEVC或L-HEVC图块轨迹表示的视频的一部分时,文件解析器可以使用以下实例过程以产生输出位流以馈送到视频解码器:[版本(YK):这应该是针对需要HEVC或L-HEVC图块轨迹的特定集合的情况一般化的]。

[0133] -首先,包含于样本条目中的参数集和SEI NAL单元按照所列出的次序通过以下项被输出:必须在VCL NAL单元(例如,参数集NAL单元、前缀SEI NAL单元)之前呈现的基础轨迹(如由“tbas”轨迹参考所指示)中的第一样本firstSampleInBaseTrack中的NAL单元、在图块轨迹中的firstSampleInBaseTrack的对应的样本中的NAL单元、在必须在VLC NAL单元(例如,EOS NAL单元、EOB NAL单元)之后呈现的firstSampleInBaseTrack中的NAL单元、在必须在VLC NAL单元之前呈现的基础轨迹中的第二样本secondSampleInBaseTrack中的NAL单元、在图块轨迹中的secondSampleInBaseTrack的对应的样本中的NAL单元、在必须在VLC NAL单元之后呈现的secondSampleInBaseTrack中的NAL单元等等。

[0134] -在基础轨迹和图块轨迹中的样本是通过解码时间同步的。换句话说,在基础轨迹中的样本和在图块轨迹中的其对应的样本具有相同解码时间。

[0135] 当基于运动受约束图块的视口依赖性VR视频译码方案在使用中时,并且如在14496-15的条款10中所指定的图块轨迹被使用,例如,每个运动受约束图块或图块区专门地运载在轨迹或DASH表示中,适应集合水平元素用于在每个运动受约束图块或图块区与运载它的表示之间用信号传送所述映射。所述元素可含有条目计数,随后是{表示ID,区位置,和区大小}的值的循环。当相同VR视频通过多空间分辨率(每个空间分辨率应对应于多个图块轨迹)表示并且所有表示包含于一个适应集合中时,每个区位置和区大小可以相对于最高分辨率指定,或替代地相对于特定分辨率的整个图片区指定,其中特定分辨率通过等于与特定分辨率相关联的表示ID的属性指示。表示ID据称与作为完全地或部分地运载在具有该表示ID的表示中的图片的分辨率的分辨率相关联。

[0136] 举例来说,例如内容准备装置20的装置可以发送包含在运动受约束图块或图块区与运载运动受约束图块或图块区的表示之间的映射的媒体呈现的适应集合水平元素,并且所述装置可以发送媒体呈现的片段。例如客户端装置40或服务器装置60的装置可以基于媒体呈现的适应集合水平元素确定在运动受约束图块或图块区与运载运动受约束图块或图块区的表示之间的映射,并且所述装置可以检索媒体呈现的片段。适应集合水平元素可以包含条目计数和条目的集合,其中条目的集合中的每个条目包含表示识别符值、区位置值和区大小值。条目计数可以指示条目的集合中的条目的数目。

[0137] 当基于运动受约束图块的视口依赖性VR视频译码方案在使用中时,并且如在14496-15的条款10中指定的图块轨迹不使用,即,每个轨迹或DASH表示运载整个经译码图片,那么存取单元水平或样本水平字节范围需要用于DASH客户端以请求特定运动受约束图块或图块区,并且通过字节范围参考的数据可以专门地用于特定运动受约束图块或图块区,或用于所有运动受约束图块或图块区。对于MPD来说包含此类字节范围是过于繁重的。一个可能的可行方式是可能地以相对方式定义新框以含有此类字节范围,例如,第0个字节意味着紧跟在新框之后的字节,并且此类新框的例子可以包含在片段索引框中,并且在片段中的所有样本或存取单元的字节范围被记录在新框中。

[0138] 举例来说,例如内容准备装置20的装置可以在存储使用基于运动补偿图块的视口依赖性VR视频译码方案的媒体呈现的文件中的框中发送特定运动受约束图块或图块区的字节范围,并且所述装置可以基于特定运动受约束图块或图块区的字节范围发送特定运动受约束图块或图块区的数据。例如客户端装置40或服务器装置60的装置可以基于存储使用基于运动补偿图块的视口依赖性VR视频译码方案的媒体呈现的文件中的框确定特定运动受约束图块或图块区的字节范围,并且所述装置可以基于特定运动受约束图块或图块区的字节范围获得特定运动受约束图块或图块区的数据。所述框可以例如是片段索引框。所述框可以包含记录在媒体呈现的片段中的所有样本或存取单元的字节范围的数据。

[0139] 图7是更详细地说明图1的检索单元52的组件的实例集合的框图。在此实例中,检索单元52包含eMBMS中间件单元100、DASH客户端110和媒体应用程序112。

[0140] 在此实例中,eMBMS中间件单元100进一步包含eMBMS接收单元106、高速缓冲存储器104和服务器单元102。在此实例中,eMBMS接收单元106经配置以经由eMBMS接收数据,例如,根据通过单向传送的文件传递(FLUTE),在T.Paila等人的“FLUTE—通过单向传送的文件传递(FLUTE-File Delivery over Unidirectional Transport)”,网络工作组,RFC 6726,2012年11月中描述,可在<http://tools.ietf.org/html/rfc6726>获得。也就是说,eMBMS接收单元106可以经由广播从例如可以充当BM-SC的服务器装置60接收文件。

[0141] 因为eMBMS中间件单元100接收用于文件的数据,所以eMBMS中间件单元可以在高速缓冲存储器104中存储接收到的数据。高速缓冲存储器104可以包括计算机可读存储媒体,例如,快闪存储器、硬盘、RAM或任何其它合适的存储媒体。

[0142] 本地服务器单元102可以充当用于DASH客户端110的服务器。举例来说,本地服务器单元102可以将MPD文件或其它清单文件提供到DASH客户端110。本地服务器单元102可以播发可用性时间用于MPD文件中的片段,以及可以从中检索片段的超链接。这些超链接可以包含对应于客户端装置40的本地主机地址前缀(例如,用于IPv4的127.0.0.1)。以此方式,DASH客户端110可以使用HTTP GET或部分GET请求从本地服务器单元102请求片段。举例来说,对于可从链接<http://127.0.0.1/rep1/seg3>获得的片段,DASH客户端110可以构造包含用于<http://127.0.0.1/rep1/seg3>的请求的HTTP GET请求,并且将请求提交到本地服务器单元102。本地服务器单元102可以从高速缓冲存储器104检索所请求的数据并且响应于此类请求将所述数据提供到DASH客户端110。

[0143] 图8是说明实例多媒体内容120的元素的框图。多媒体内容120可对应于多媒体内容64(图1),或存储在存储媒体62中的另一多媒体内容。在图8的实例中,多媒体内容120包含媒体呈现描述(MPD)122和多个表示124A到124N(表示124)。表示124A包含任选的标头

数据126和片段128A到128N(片段128),而表示124N包含任选的标头数据130和片段132A到132N(片段132)。为了方便起见,使用字母N来标明表示124中的每一者中的最后一个电影片段。在一些实例中,表示124之间可能存在不同数目的电影片段。

[0144] MPD 122可包括与表示124分开的数据结构。MPD 122可对应于图1的清单文件66。类似地,表示124可对应于图1的表示68。一般来说,MPD 122可包含通常描述表示124的特性的数据,例如,译码及再现特性、适应集合、MPD 122所对应的简档、文本类型信息、相机角度信息、分级信息、特技模式信息(例如,指示包含时间子序列的表示的信息)和/或用于检索遥远周期(例如,用于在回放期间向媒体内容中插入定向广告)的信息。

[0145] 标头数据126当存在时可描述片段128的特性,例如,随机存取点(RAP,也称为流存取点(SAP))的时间位置、片段128中的哪一个包含随机存取点、到片段128内的随机存取点的字节偏移、片段128的统一资源定位符(URL),或片段128的其它方面。标头数据130当存在时可描述片段132的类似特性。另外或替代地,此类特性可完全包含在MPD 122内。

[0146] 片段128、132包含一或多个经译码视频样本,其中的每一个可包含视频数据的帧或切片。片段128的经译码视频样本中的每一个可具有类似特性,例如,高度、宽度和带宽要求。可通过MPD 122的数据描述此类特性,但在图8的实例中未说明此类数据。MPD 122可包含如3GPP规范所描述的特性,其中添加了本发明中所描述的用信号传送信息中的任一个或全部。

[0147] 片段128、132的每一个可与唯一统一资源定位符(URL)相关联。因此,片段128、132中的每一个可使用流式传输网络协议(例如,DASH)可独立地检索。以此方式,例如客户端装置40的目的地装置可以使用HTTP GET请求来检索片段128或132。在一些实例中,客户端装置40可使用HTTP部分GET请求来检索片段128或132的特定字节范围。

[0148] 图9是说明可以对应于例如图8的片段128、132中的一个的表示的片段的实例视频文件150的元素的框图。片段128、132中的每一个可以包含基本上符合在图9的实例中说明的数据的布置的数据。视频文件150可以被称为封装一个片段。如上文所描述,根据ISO基础媒体文件格式及其扩展的视频文件在被称作“框”的一系列对象中存储数据。在图9的实例中,视频文件150包含文件类型(FTYP)框152、电影(MO0V)框154、片段索引(sidx)框162、电影片段(MOOF)框164和电影片段随机存取(MFRA)框166。虽然图9表示视频文件的实例,但是应理解根据ISO基础媒体文件格式及其扩展,其它媒体文件可以包含类似于视频文件150的数据构造的其它类型的媒体数据(例如,音频数据、定时文本数据或类似者)。

[0149] 文件类型(FTYP)框152通常描述视频文件150的文件类型。文件类型框152可以包含识别描述视频文件150的最佳用途的规范的数据。文件类型框152可以替代地放置在MO0V框154、电影片段框164和/或MFRA框166之前。

[0150] 在一些实例中,例如视频文件150的片段可以包含在FTYP框152之前的MPD更新框(未示出)。MPD更新框可以包含指示对应于包含待更新的视频文件150的表示的MPD的信息,连同用于更新MPD的信息。举例来说,MPD更新框可以提供用于待用于更新MPD的资源的URI或URL。作为另一实例,MPD更新框可以包含用于更新MPD的数据。在一些实例中,MPD更新框可以紧跟视频文件150的片段类型(STYP)框(未示出),其中STYP框可以定义用于视频文件150的片段类型。

[0151] 在图9的实例中,M00V框154包含电影标头(MVHD)框156、轨迹(TRAK)框158和一或

多个电影扩展 (MVEX) 框160。一般来说, MVHD框156可以描述视频文件150的一般特性。举例来说, MVHD框156可以包含描述何时最初创建视频文件150、何时最后修改视频文件150、视频文件150的时间标度、视频文件150的回放的持续时间的数据, 或通常描述视频文件150的其它数据。

[0152] TRAK框158可以包含用于视频文件150的轨迹的数据。TRAK框158可以包含描述对应于TRAK框158的轨迹的特性的轨迹标头 (TKHD) 框。在一些实例中, TRAK框158可以包含经译码视频图片, 而在其它实例中, 轨迹的经译码视频图片可以包含于电影片段164中, 所述电影片段可以通过TRAK框158和/或sidx框162的数据参考。

[0153] 在一些实例中, 视频文件150可以包含一个以上轨迹。相应地, MOOV框154可以包含与视频文件150中的轨迹的数目相等的数目的TRAK框。TRAK框158可以描述视频文件150的对应的轨迹的特性。举例来说, TRAK框158可以描述用于对应的轨迹的时间和/或空间信息。当封装单元30 (图1) 包含在视频文件 (例如, 视频文件150) 中的参数集轨迹时, 类似于MOOV框154的TRAK框158的TRAK框可以描述参数集轨迹的特性。封装单元30可以在描述参数集轨迹的TRAK框内的参数集轨迹中用信号传送序列水平SEI消息的存在。

[0154] MVEX框160可以描述对应的电影片段164的特性, 例如, 以用信号传送除包含在MOOV框154内的视频数据 (如果存在的话) 之外, 视频文件150还包含电影片段164。在流式传输视频数据的情形下, 经译码视频图片可以包含于电影片段164中而非包含于MOOV框154中。相应地, 所有经译码视频样本可以包含于电影片段164中, 而非包含于MOOV框154中。

[0155] MOOV框154可以包含与视频文件150中的电影片段164的数目相等的数目的MVEX框160。MVEX框160中的每一个可以描述电影片段164中的对应的一个的特性。举例来说, 每个MVEX框可以包含描述电影片段164中的对应的一个的持续时间的电影扩展标头框 (MEHD) 框。

[0156] 如上文所指出, 封装单元30可以在并不包含实际经译码视频数据的视频样本中存储序列数据集。视频样本可以通常对应于存取单元, 所述存取单元是在特定时间例子处经译码图片的表示。在AVC和HEVC的情形下, 经译码图片包含含有信息的一或多个VCL NAL单元以构造存取单元和其它相关联的非VCL NAL单元的所有像素, 例如, SEI消息。相应地, 封装单元30可以包含序列数据集, 所述序列数据集可以包含在电影片段164中的一个中的序列水平SEI消息。封装单元30可以进一步用信号传送序列数据集和/或序列水平SEI消息的存在作为存在于在对应于电影片段164中的一个的MVEX框160中的一个内的电影片段164中的一个中。

[0157] SIDX框162是视频文件150的任选的元素。也就是说, 符合3GPP文件格式或其它此类文件格式的视频文件不一定包含SIDX框162。根据3GPP文件格式的实例, SIDX框可用于识别片段 (例如, 包含在视频文件150内的片段) 的子片段。3GPP文件格式将子片段定义为“具有对应的媒体数据框的一或多个连续的电影片段框的独立的集合并且包含电影片段框所参考的数据的媒体数据框必须在电影片段框之后并且在包含关于相同轨迹的信息的下一个电影片段框之前”。3GPP文件格式还指示SIDX框“包含通过框记录的 (子) 片段的子片段的参考的序列”。所参考的子片段在呈现时间中是连续的。类似地, 由片段索引框所指代的字节在片段内始终是连续的。所参考的大小给出在所参考的材料中的字节的数目的计数。

[0158] SIDX框162通常提供表示包含于视频文件150中的片段的一或多个子片段的信息。

举例来说,此类信息可以包含子片段开始和/或结束的回放时间、用于子片段的字节偏移、子片段是否包含(例如,起始于)流存取点(SAP)、用于SAP的类型(例如,SAP是否是即时解码器刷新(IDR)图片、清洁随机存取(CRA)图片、断链存取(BLA)图片,或类似者)、在子片段中的SAP的位置(就回放时间和/或字节偏移而言),及类似者。

[0159] 电影片段164可以包含一或多个经译码视频图片。在一些实例中,电影片段164可以包含图片的一或多个群组(GOP),其中的每一个可以包含多个经译码视频图片,例如,帧或图片。另外,如上文所述,在一些实例中,电影片段164可以包含序列数据集。电影片段164中的每一个可以包含电影片段标头框(MFHD,在图9中未示出)。MFHD框可以描述对应的电影片段的特性,例如,电影片段的序列号。可以按视频文件150中的序列号的次序包含电影片段164。

[0160] MFRA框166可以描述视频文件150的电影片段164内的随机存取点。这可以辅助于执行特技模式,例如,执行对通过视频文件150封装的片段内的特定时间位置(即,回放时间)的寻找。在一些实例中,MFRA框166通常是任选的并且无需包含于视频文件中。类似地,客户端装置(例如,客户端装置40)不一定需要参考MFRA框166以正确地解码和显示视频文件150的视频数据。MFRA框166可以包含与视频文件150的轨迹的数目相等的数目的轨迹片段随机存取(TFRA)框(未示出),或在一些实例中,包含与视频文件150的媒体轨迹(例如,非提示轨迹)的数目相等的数目的轨迹片段随机存取(TFRA)框(未示出)。

[0161] 在一些实例中,电影片段164可以包含一或多个流存取点(SAP),例如,IDR图片。类似地,MFRA框166可以提供SAP的视频文件150内的位置的指示。相应地,视频文件150的时间子序列可以由视频文件150的SAP形成。时间子序列还可包含其它图片,例如,取决于SAP的P帧和/或B帧。时间子序列的帧和/或切片可以布置在片段内,使得取决于子序列的其它帧/切片的时间子序列的帧/切片可以恰当地得到解码。举例来说,在数据的分级布置中,用于预测其它数据的数据也可以包含于时间子序列中。

[0162] 图10是根据本发明的技术用于处理媒体数据的流程图。将相对于通用装置描述图10的技术。所述装置可以例如对应于例如服务器装置60的服务器装置、例如客户端装置40的客户端装置,或经配置以处理媒体数据的一些其它类型的装置。所述装置接收媒体数据(200)。媒体数据可以包含VR视频数据和非VR视频数据两者。所述装置至少部分地基于在用于媒体呈现的媒体呈现描述的适应集合水平处用信号传送的数据确定在媒体呈现中使用的投影映射(202)。在适应集合水平处用信号传送的数据可以例如是适应集合中的元素,并且视频表示可以属于适应集合。

[0163] 所述装置基于在媒体呈现中使用的投影映射处理媒体呈现的视频表示的片段(204)。服务器装置可以例如基于在媒体呈现中使用的投影映射处理媒体呈现的视频表示的片段,方法是至少部分地基于在媒体呈现中使用的投影映射确定是否检索视频表示的片段。服务器装置可以将媒体呈现的视频表示的片段发送到客户端装置。客户端装置可以例如基于在媒体呈现中使用的投影映射处理媒体呈现的视频表示的片段,方法是至少部分地基于在媒体呈现中使用的投影映射确定如何解码视频表示的片段。客户端装置可以例如基于在媒体呈现中使用的投影映射解码媒体呈现的视频表示的片段并且输出媒体呈现的视频表示的经解码片段。

[0164] 在一些实施方案中,取决于周期水平信令的存在,适应水平信令可以是任选的。在

一些实例中,所述装置还可以确定在媒体呈现中使用的投影映射的指示是否存在于媒体呈现的周期水平处,并且响应于确定在媒体呈现中使用的投影映射的指示并非存在于媒体呈现的周期水平处确定在媒体呈现中使用的投影映射。在其它实例中,所述装置还可以确定在媒体呈现中使用的投影映射的指示是否存在于媒体呈现的周期水平处,并且响应于确定在媒体呈现中使用的投影映射的指示并非存在于媒体呈现的周期水平处确定在媒体呈现中使用的投影映射。

[0165] 所述装置还可以至少部分地基于在第二媒体呈现的周期水平处用信号传送的数据且无需在第二媒体呈现的表示水平处的额外的信令确定在第二媒体呈现中使用的投影映射并且基于在第二媒体呈现中使用的投影映射处理第二媒体呈现的视频表示的片段。

[0166] 图11是根据本发明的技术用于处理媒体数据的流程图。将相对于通用装置描述图11的技术。举例来说,所述装置可以对应于例如内容准备装置20的内容准备装置或经配置以发送媒体数据的一些其它类型的装置。所述装置产生媒体数据(210)。举例来说,媒体数据可以包含VR视频数据和非VR视频数据两者。所述装置在媒体数据中包含在识别在媒体呈现中使用的投影映射的媒体呈现描述的适应集合水平处用信号传送的数据(212)。在适应集合水平处用信号传送的数据可以例如是适应集合中的元素,并且视频表示可以属于适应集合。所述装置基于在媒体呈现中使用的投影映射发送媒体呈现的视频表示的片段(214)。

[0167] 在一些实施方案中,取决于周期水平信令的存在,适应水平信令可以是任选的。在一些实例中,所述装置还可以确定是否在媒体呈现的周期水平处包含在媒体呈现中使用的投影映射的指示,并且响应于确定在媒体呈现的周期水平处不包含在媒体呈现中使用的投影映射的指示包含在识别在媒体呈现中使用的投影映射的媒体呈现描述的适应集合水平处用信号传送的数据。在其它实例中,所述装置还可以确定是否在媒体呈现的周期水平处包含在媒体呈现中使用的投影映射的指示,并且响应于确定在媒体呈现的周期水平处包含在媒体呈现中使用的投影映射的指示包含在识别在媒体呈现中使用的投影映射的媒体呈现描述的适应集合水平处用信号传送的数据。

[0168] 所述装置还可以在媒体数据中包含在识别在第二媒体呈现中使用的投影映射的第二媒体呈现的周期水平处用信号传送的数据并且基于在第二媒体呈现中使用的投影映射发送第二媒体呈现的视频表示的片段。

[0169] 在一或多个实例中,所描述的功能可以在硬件、软件、固件或其任何组合中实施。如果在软件中实施,那么所述功能可作为一或多个指令或代码在计算机可读媒体上存储或发射,并且由基于硬件的处理单元执行。计算机可读媒体可以包含计算机可读存储媒体,其对应于例如数据存储媒体或通信媒体的有形媒体,通信媒体(例如)根据通信协议包含有助于将计算机程序从一处传送到另一处的任何媒体。以此方式,计算机可读媒体通常可对应于(1)非暂时性的有形计算机可读存储媒体,或(2)通信媒体,例如,信号或载波。数据存储媒体可以是可由一或多个计算机或一或多个处理器存取以检索用于实施本发明中所描述的技术的指令、代码和/或数据结构的任何可用媒体。计算机程序产品可以包含计算机可读媒体。

[0170] 借助于实例而非限制,此类计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器,或可用于存储呈指令或数据结构形式的所期望的程序代码且可由计算机存取的任何其它媒体。并且,适当地

将任何连接称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或例如红外线、无线电和微波等无线技术从网站、服务器或其它远程源发射指令,那么同轴电缆、光纤电缆、双绞线、DSL或例如红外线、无线电和微波等无线技术包含在媒体的定义中。然而,应理解,所述计算机可读存储媒体和数据存储媒体并不包含连接、载波、信号或其它暂时性媒体,而是实际上涉及非暂时性的有形存储媒体。如本文中所使用的磁盘和光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘和蓝光光盘,其中磁盘通常是以磁性方式再现数据,而光盘是用激光以光学方式再现数据。以上各项的组合也应包含于计算机可读媒体的范围内。

[0171] 指令可以由一或多个处理器执行,所述一或多个处理器例如是一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效的集成或离散逻辑电路。相应地,如本文中所使用的术语“处理器”可指代上述结构或适用于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面中,本文中所描述的功能性可在经配置以用于编码和解码的专用硬件和/或软件模块内提供,或并入在组合编解码器中。并且,所述技术可完全实施于一或多个电路或逻辑元件中。

[0172] 本发明的技术可以在多种多样的装置或设备中实施,包含无线手持机、集成电路(IC)或IC的集合(例如,芯片集合)。本发明中所描述的各种组件、模块或单元是为了强调经配置以执行所公开的技术的装置的功能方面,但未必需要由不同硬件单元实现。实际上,如上文所描述,各种单元可结合合适的软件和/或固件组合在编解码器硬件单元中,或通过互操作硬件单元的集合来提供,所述硬件单元包含如上文所描述的一或多个处理器。

[0173] 已经描述了各种实例。这些和其它实例在所附权利要求书的范围内。

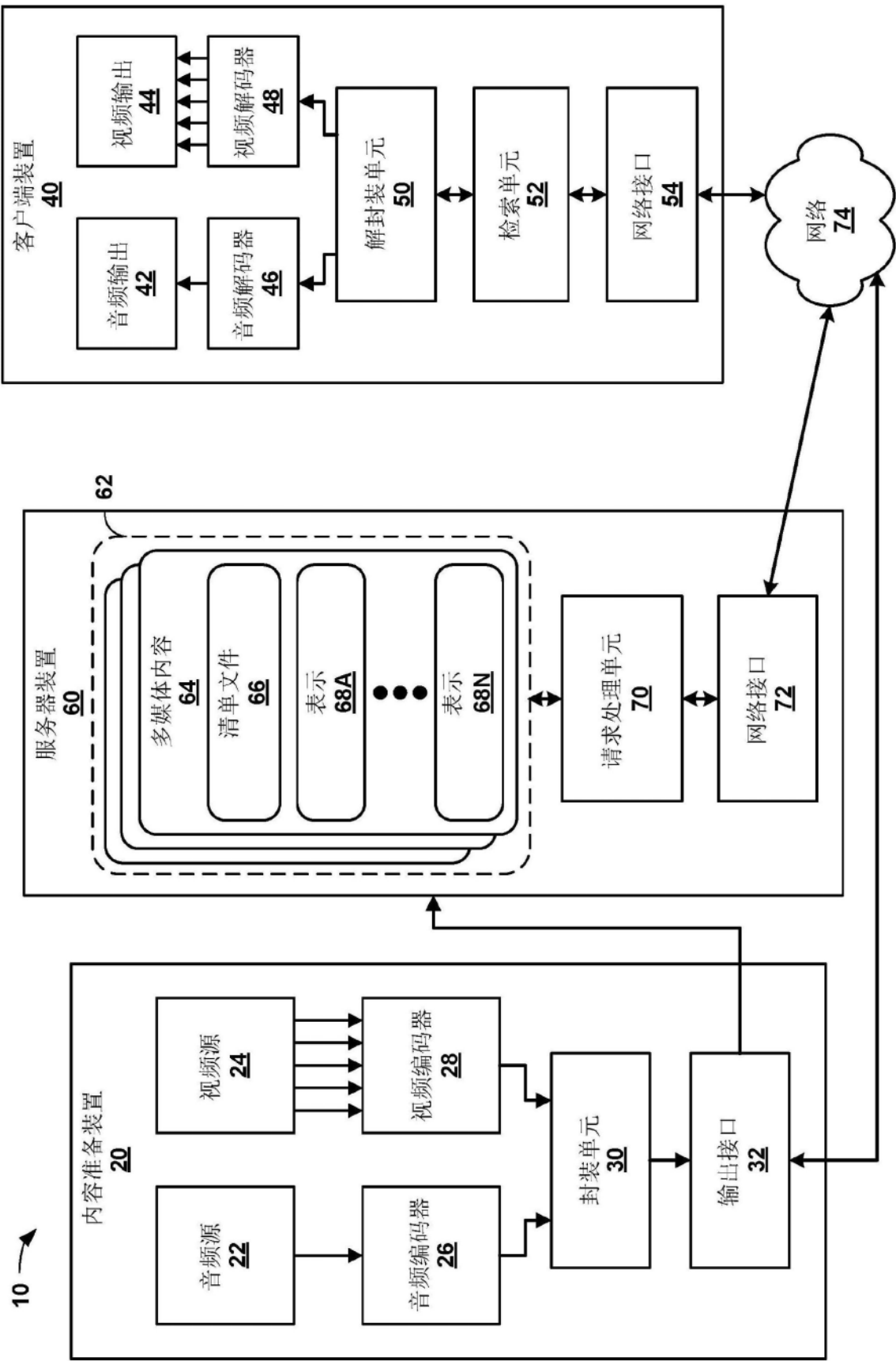


图1

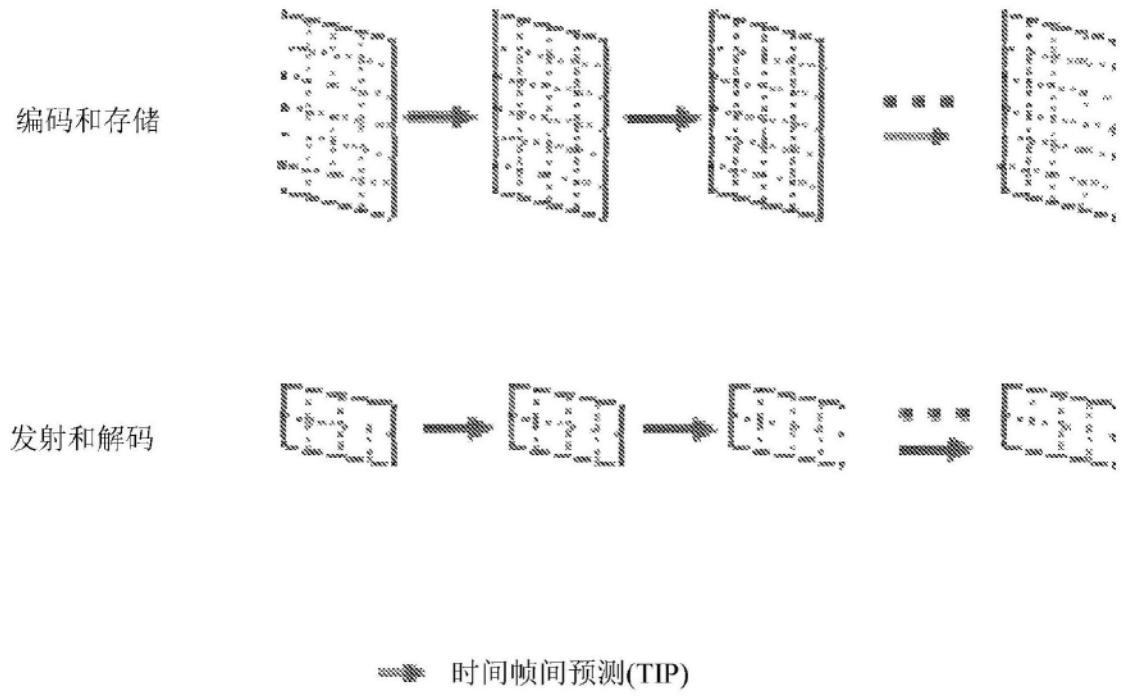


图2

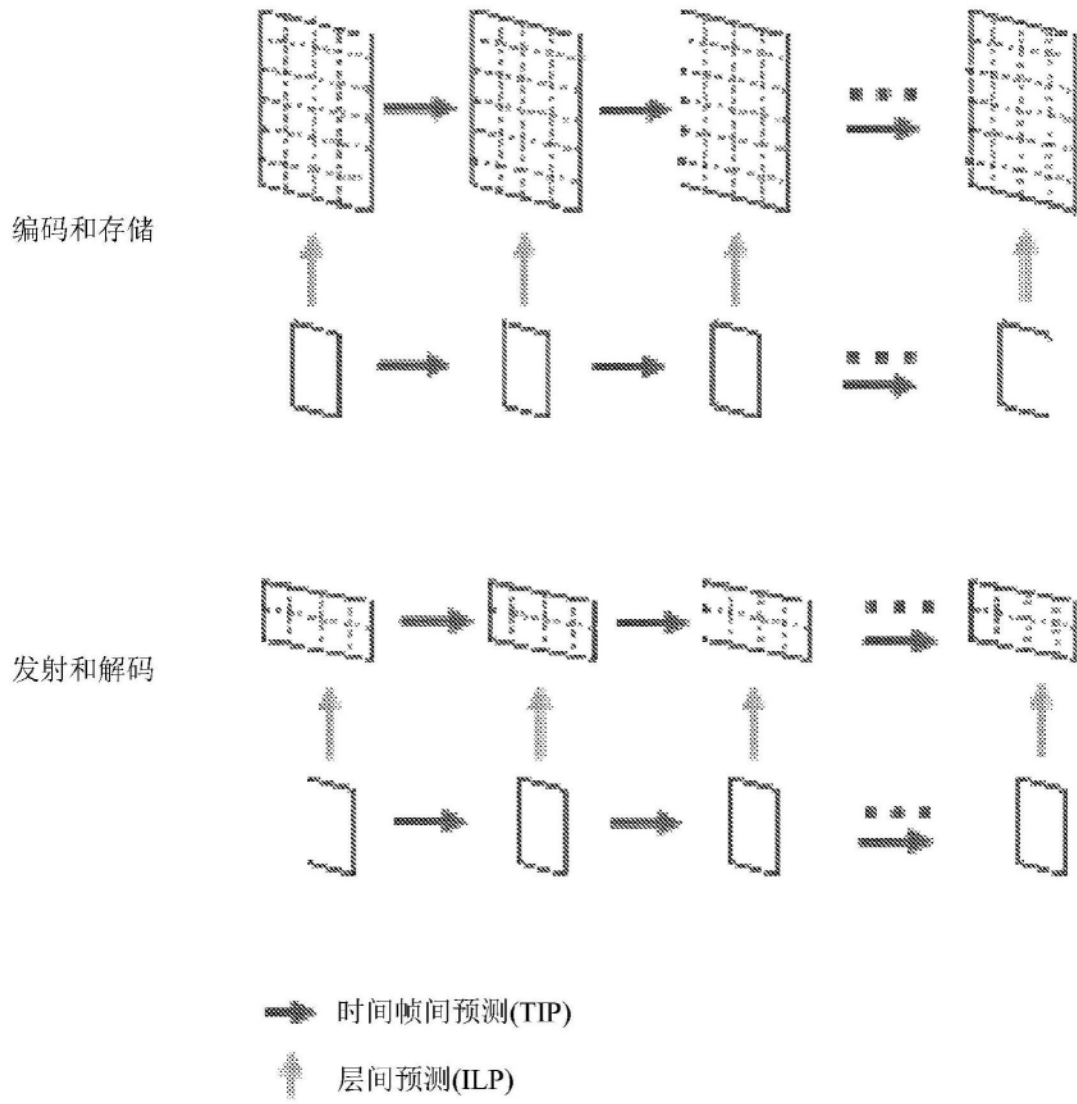
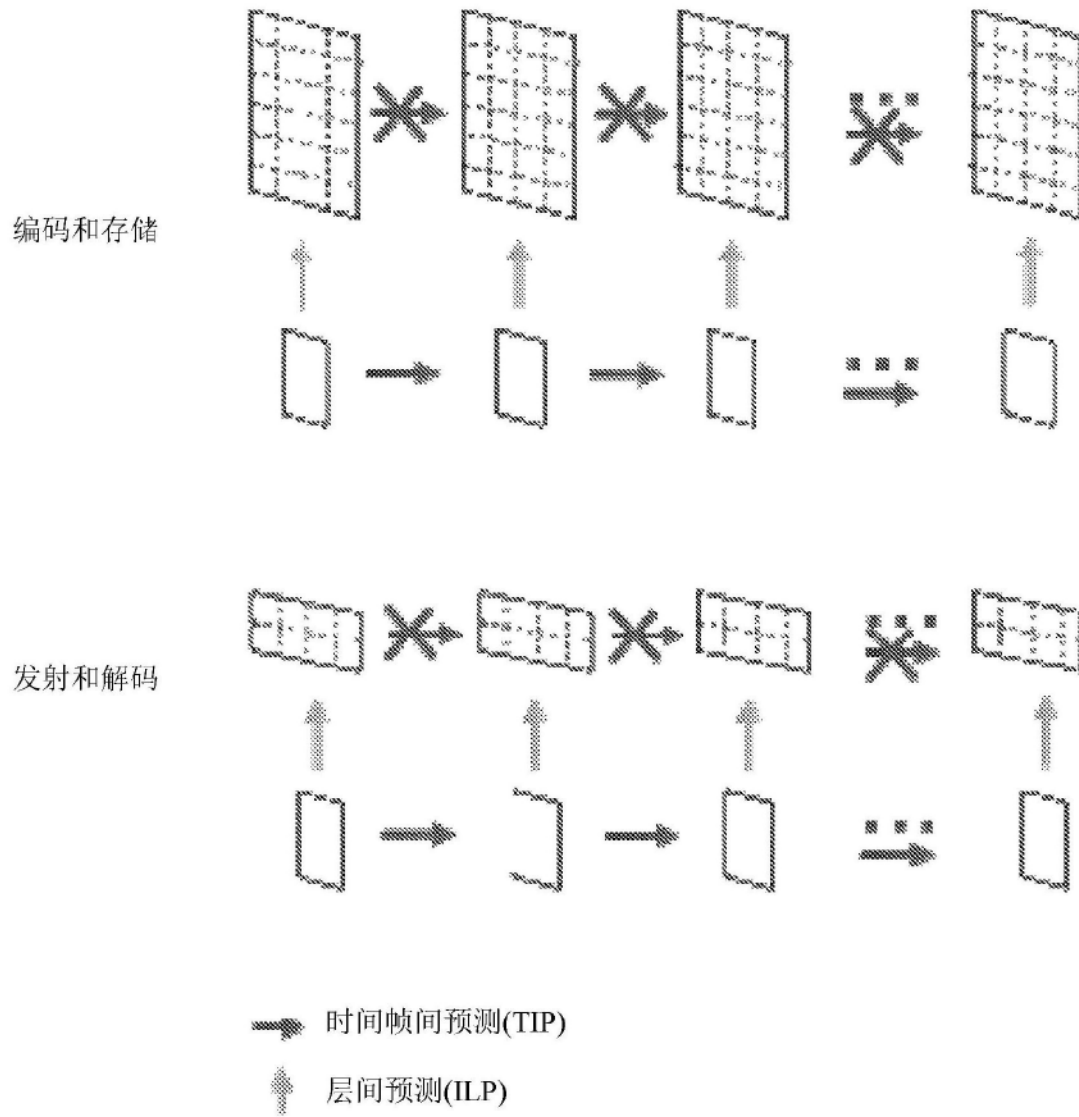


图3



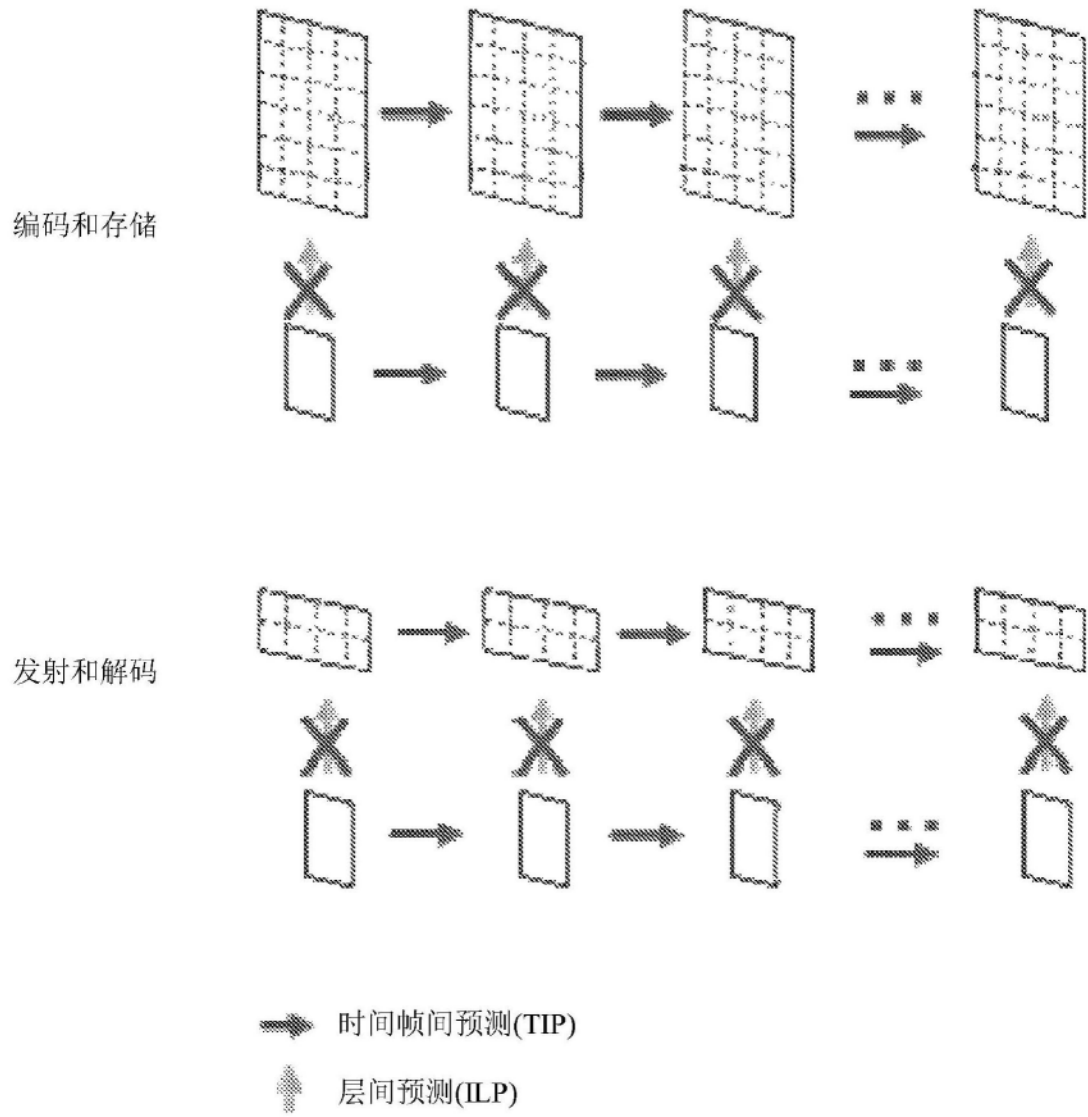


图5

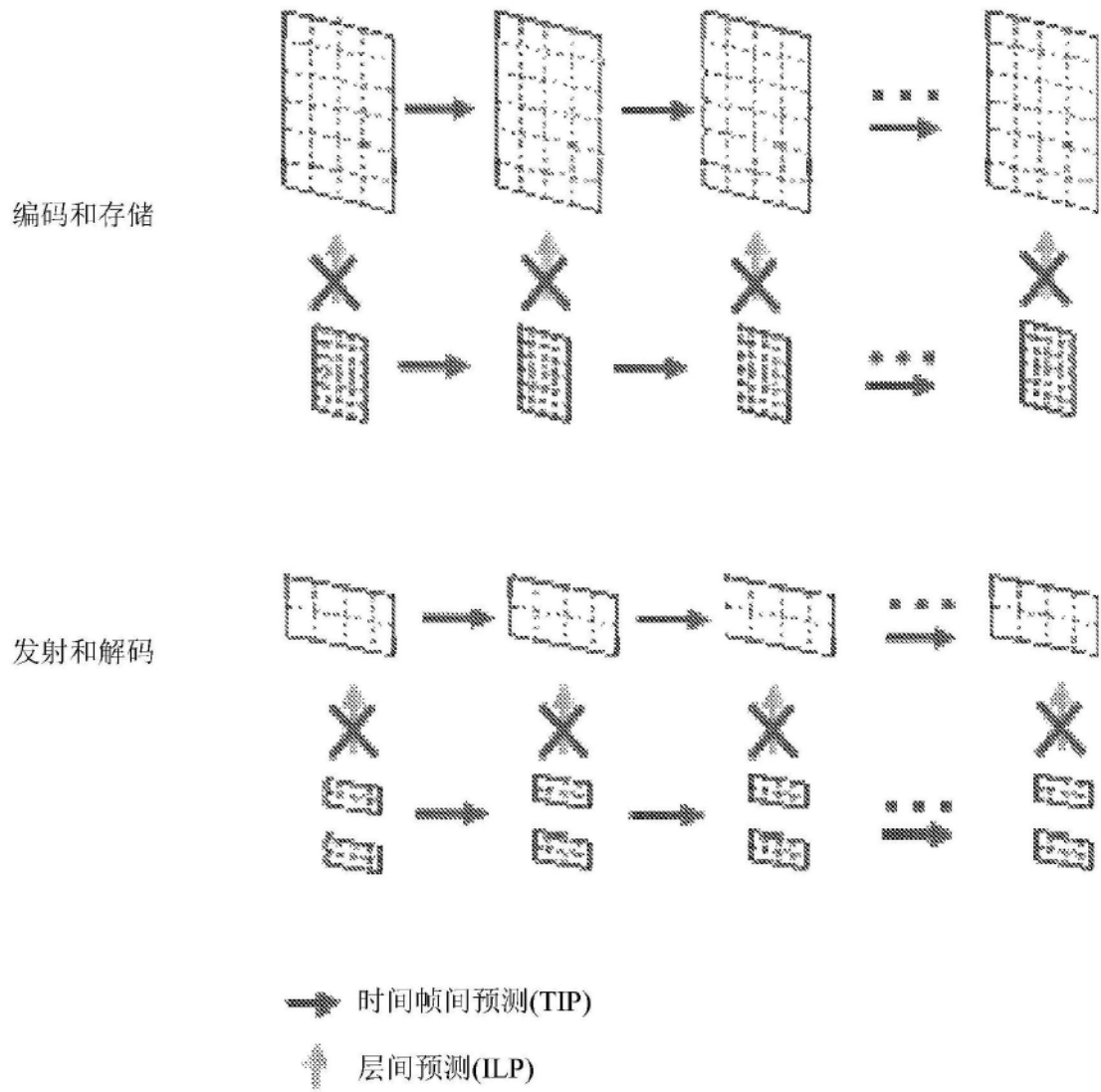


图6

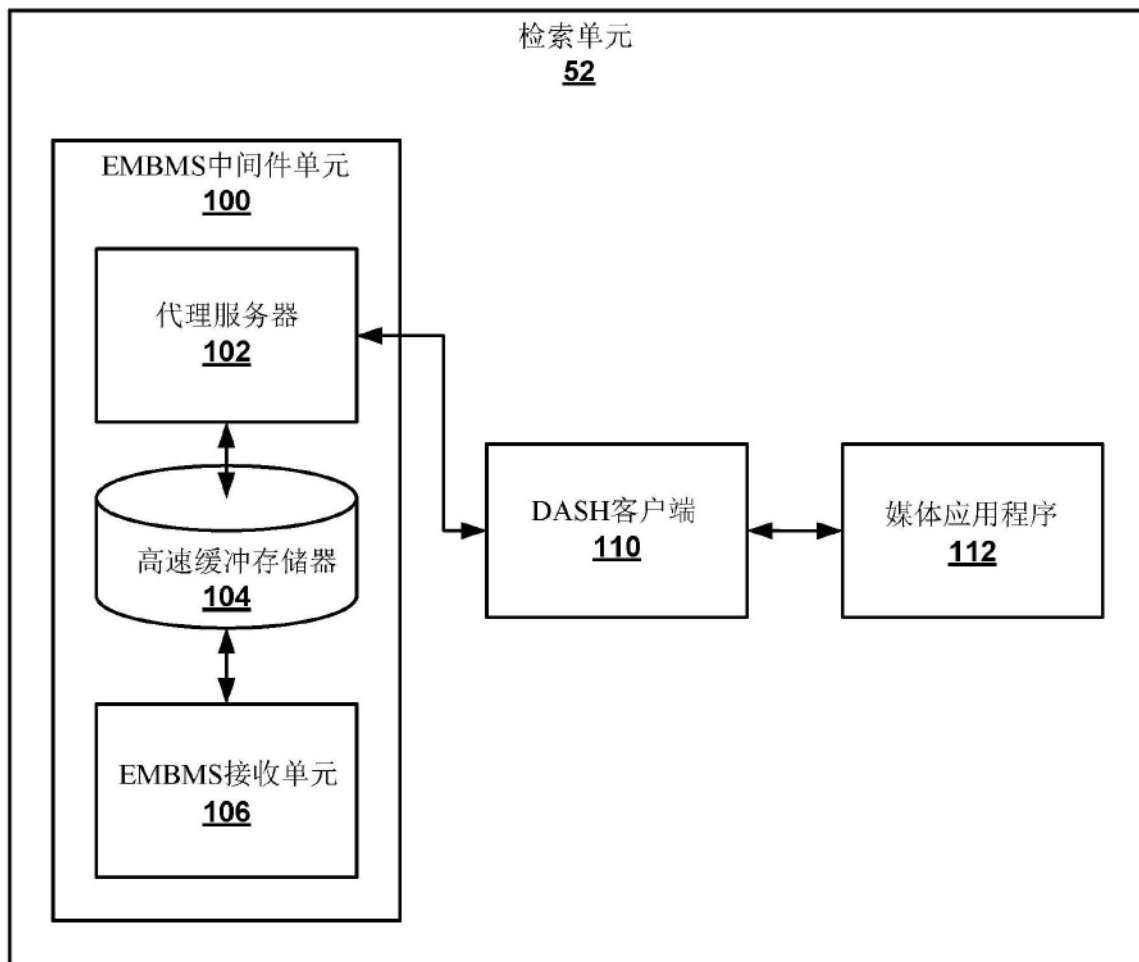


图7

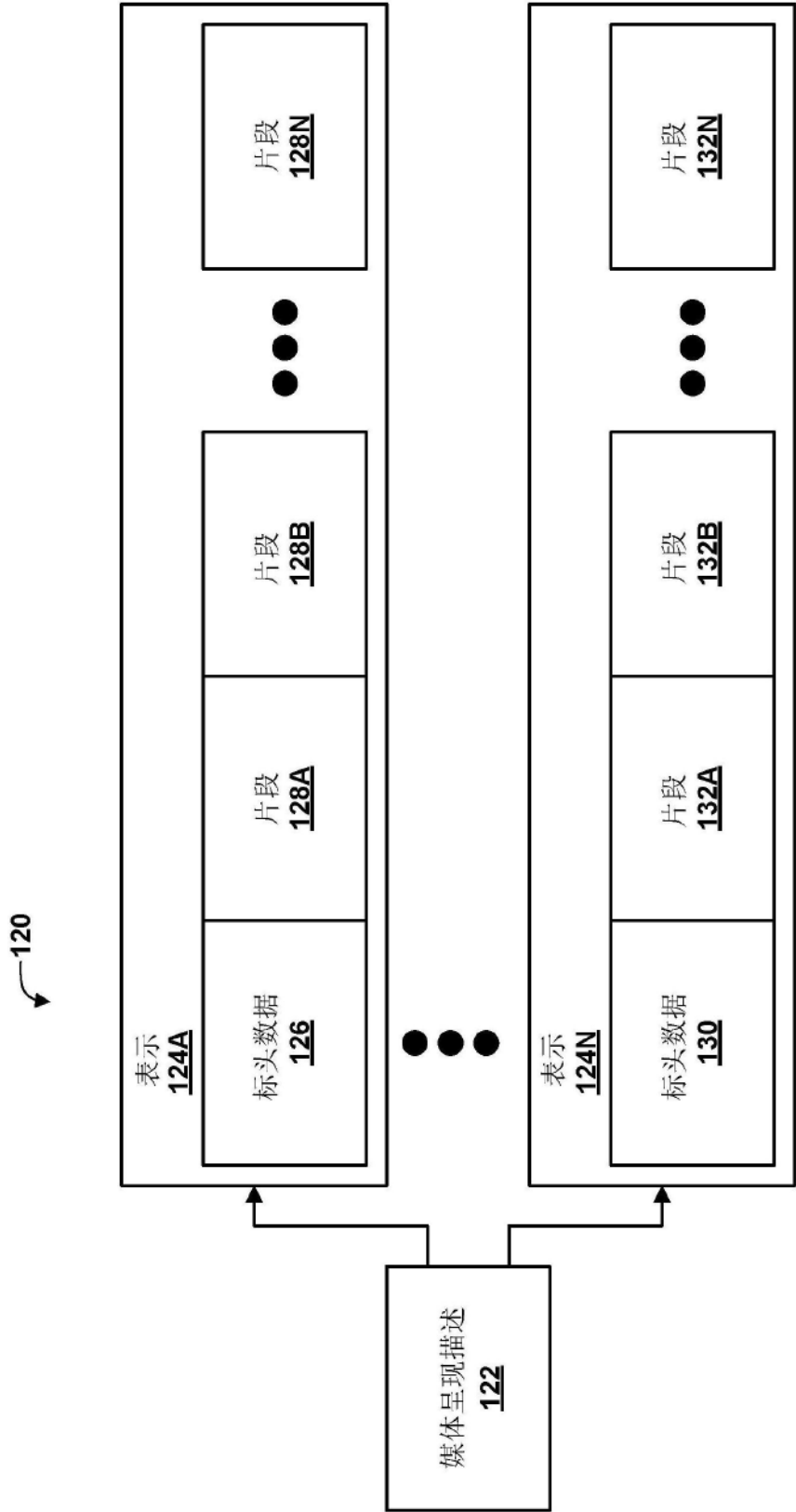


图8

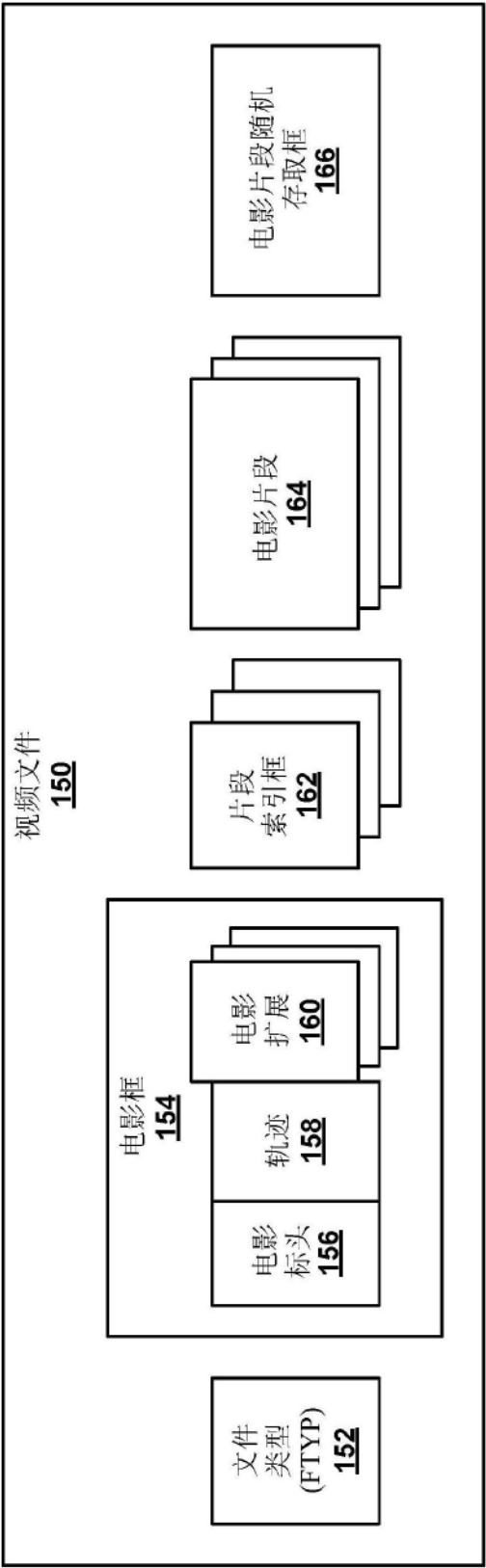


图9

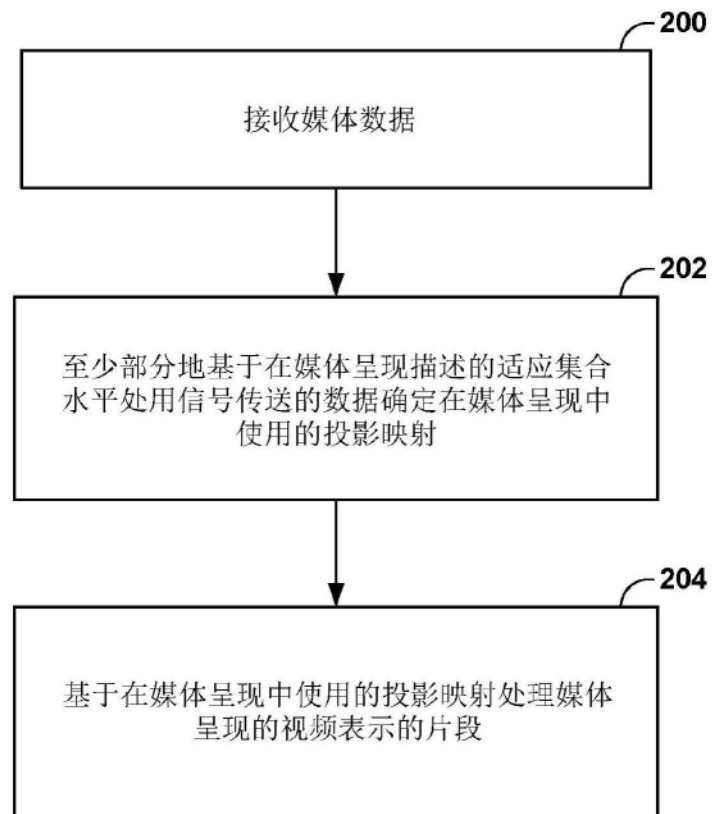


图10

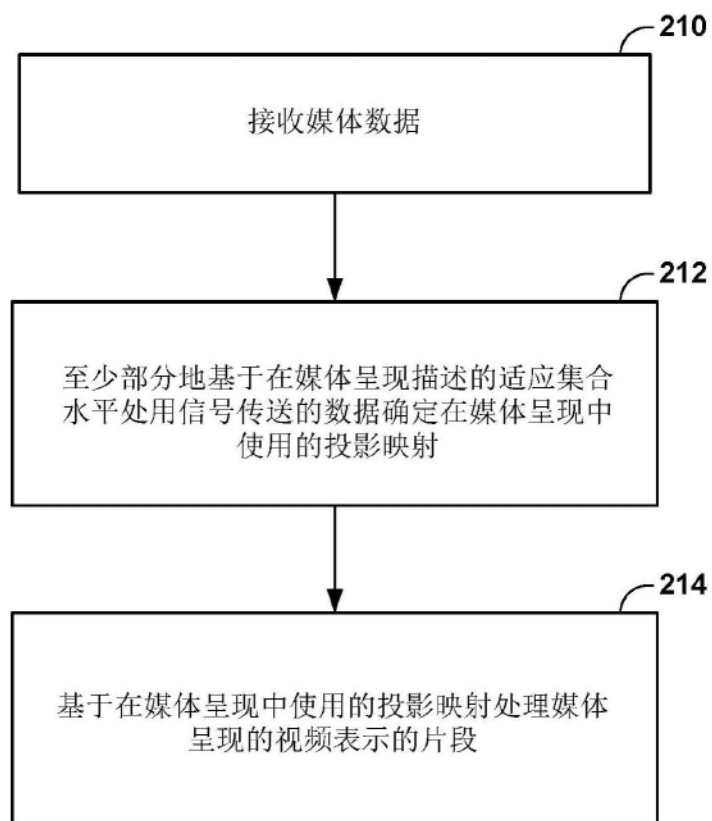


图11