

1. 一种传输媒体数据的方法,所述方法包括由源设备的基于文件的协议发送单元进行以下操作:

从所述源设备的形成区段的分段器接收包括媒体数据的区段的数据的流,所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件;

确定媒体传递事件 (MDE) 在媒体数据的所述流中的位置,其中,所述MDE包括针对所述区段中的一个区段的至少一部分的数据;

确定用于所述MDE的一个或多个传输时间要求,所述一个或多个传输时间要求表示要将所述MDE发送给客户端设备的时间;以及

根据所述源设备的物理层发送单元的可用传递时隙将所述MDE和表示所述MDE的所述一个或多个传输时间要求的数据提供给所述物理层发送单元。

2. 根据权利要求1所述的方法,其中,确定所述MDE的位置包括:使用模式匹配来确定所述MDE的位置。

3. 根据权利要求2所述的方法,其中,所述媒体数据包括视频数据,并且其中,使用模式匹配来确定所述MDE的位置包括:

获得定义所述流的视频帧的节奏的数据,其中,所定义的节奏表示发送所述视频帧的类型的顺序,所述视频帧的类型包括帧内预测视频帧和帧间预测视频帧;以及

基于所定义的节奏来确定所述MDE中的至少一些MDE,使得所述MDE在对应的视频帧边界之间。

4. 根据权利要求2所述的方法,其中,所述媒体数据包括音频数据,并且其中,使用模式匹配来确定所述MDE的位置包括:

确定所述流的固定数量的音频帧以包括在所述MDE中的每个MDE中;以及

确定所述MDE中的至少一些MDE均包括来自所述流的所述固定数量的音频帧。

5. 根据权利要求1所述的方法,其中,确定所述MDE的位置包括使用规则来确定所述MDE的位置,包括:

接收表示一个或多个规则的数据,所述一个或多个规则定义音频帧、视频帧以及定时文本实例中的一项或多项;

基于所述规则来确定至少一个音频帧、视频帧或者定时文本实例的位置;以及

基于所述至少一个音频帧、视频帧或者定时文本实例的位置来确定所述MDE。

6. 根据权利要求1所述的方法,其中,确定所述MDE的位置包括:基于用于所述物理层发送单元的物理层同步的定时信息来确定所述MDE的位置。

7. 根据权利要求1所述的方法,其中,确定所述MDE的位置包括:对提示轨道进行分析,所述提示轨道包括表示所述流内可播放数据的位置的元数据。

8. 根据权利要求1所述的方法,其中,确定所述MDE的所述一个或多个传输时间要求包括:基于所述源设备的系统配置来确定用于所述MDE中的一个MDE的最早传输时间。

9. 根据权利要求1所述的方法,其中,确定所述MDE的所述一个或多个传输时间要求包括:基于用于区段的区段可用时间来确定所述MDE中的一个MDE的最迟传输时间,在所述MDE中的所述一个MDE中包括针对所述区段的数据,所述方法还包括:根据数据的所述流的清单文件来确定区段可用时间。

10. 一种用于传输媒体数据的源设备,所述源设备包括:

分段器,其被配置为:形成包括媒体数据的区段的数据的流,所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件;

物理层发送单元,其被配置为:根据用于媒体传递事件 (MDE) 的一个或多个传输时间要求将所述MDE传递给客户端设备,其中,所述MDE包括针对所述区段中的一个区段的至少一部分的数据,其中,所述物理层发送单元被配置有用于接收要传递的数据的可用传递时隙;以及

基于文件的协议发送单元,其被配置为:

从所述分段器接收包括所述媒体数据的区段的数据的所述流;

确定所述MDE在媒体数据的所述流中的位置;

确定用于所述MDE的一个或多个传输时间要求,所述MDE的所述一个或多个传输时间要求表示要将所述MDE发送给所述客户端设备的时间;以及

根据所述物理层发送单元的所述可用传递时隙将所述MDE和表示所述MDE的所述一个或多个传输时间要求的数据提供给所述物理层发送单元。

11. 根据权利要求10所述的源设备,其中,所述基于文件的协议发送单元被配置为:使用模式匹配来确定所述MDE的位置。

12. 根据权利要求11所述的源设备,其中,所述媒体数据包括视频数据,并且其中,为了使用模式匹配来确定所述MDE的位置,所述基于文件的协议发送单元被配置为:

获得定义所述流的视频帧的节奏的数据,其中,所定义的节奏表示发送所述视频帧的类型的顺序,所述视频帧的类型包括帧内预测视频帧和帧间预测视频帧;以及

基于所定义的节奏来确定所述MDE中的至少一些MDE,使得所述MDE在对应的视频帧边界之间。

13. 根据权利要求11所述的源设备,其中,所述媒体数据包括音频数据,并且其中,为了使用模式匹配来确定所述MDE的位置,所述基于文件的协议发送单元被配置为:

确定所述流的固定数量的音频帧以包括在所述MDE中的每个MDE中;以及

确定所述MDE中的至少一些MDE均包括来自所述流的所述固定数量的音频帧。

14. 根据权利要求10所述的源设备,其中,所述基于文件的协议发送单元被配置为:使用规则来确定所述MDE的位置,并且其中,为了使用所述规则,所述基于文件的协议发送单元被配置为:

接收表示一个或多个规则的数据,所述一个或多个规则定义音频帧、视频帧以及定时文本实例中的一项或多项;

基于所述规则来确定至少一个音频帧、视频帧或者定时文本实例的位置;以及

基于所述至少一个音频帧、视频帧或者定时文本实例的位置来确定所述MDE。

15. 根据权利要求10所述的源设备,其中,所述基于文件的协议发送单元被配置为:基于用于所述物理层发送单元的物理层同步的定时信息来确定所述MDE的位置。

16. 根据权利要求10所述的源设备,其中,为了确定所述MDE的位置,所述基于文件的协议发送单元被配置为:对提示轨道进行分析,所述提示轨道包括表示所述流内可播放数据的位置的元数据。

17. 根据权利要求10所述的源设备,其中,为了确定所述MDE的所述一个或多个传输时间要求,所述基于文件的协议发送单元被配置为:基于所述源设备的系统配置来确定用于所述MDE中的一个MDE的最早传输时间。

18. 根据权利要求10所述的源设备,其中,为了确定所述MDE的所述一个或多个传输时间要求,所述基于文件的协议发送单元被配置为:基于用于区段的区段可用时间来确定用于所述MDE中的一个MDE的最迟传输时间,在所述MDE中的所述一个MDE中包括针对所述区段的数据,并且其中,所述基于文件的协议发送单元还被配置为:根据数据的所述流的清单文件来确定区段可用时间。

19. 一种用于传输媒体数据的源设备,所述源设备包括:

分段器,其被配置为:形成包括媒体数据的区段的数据的流,所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件;

物理层发送单元,其被配置为:根据用于媒体传递事件 (MDE) 的一个或多个传输时间要求将所述MDE传递给客户端设备,其中,所述MDE包括针对所述区段中的一个区段的至少一部分的数据,其中,所述物理层发送单元被配置有用于接收要传递的数据的可用传递时隙;

用于从所述分段器接收包括媒体数据的区段的数据的流的单元;

用于确定所述MDE在媒体数据的所述流中的位置的单元;

用于确定用于所述MDE的一个或多个传输时间要求的单元,所述一个或多个传输时间要求表示要将所述MDE发送给所述客户端设备的时间;以及

用于根据所述物理层发送单元的可用传递时隙将所述MDE和表示所述MDE的所述一个或多个传输时间要求的数据提供给所述物理层发送单元的单元。

20. 根据权利要求19所述的源设备,其中,所述用于确定所述MDE的位置的单元包括:用于使用模式匹配来确定所述MDE的位置的单元。

21. 根据权利要求20所述的源设备,其中,所述媒体数据包括视频数据,并且其中,所述用于使用模式匹配来确定所述MDE的位置的单元包括:

用于获得定义所述流的视频帧的节奏的数据的单元,其中,所定义的节奏表示发送所述视频帧的类型的顺序,所述视频帧的类型包括帧内预测视频帧和帧间预测视频帧;以及

用于基于所定义的节奏来确定所述MDE中的至少一些MDE,使得所述MDE在对应的视频帧边界之间的单元。

22. 根据权利要求20所述的源设备,其中,所述媒体数据包括音频数据,并且其中,所述用于使用模式匹配来确定所述MDE的位置的单元包括:

用于确定所述流的固定数量的音频帧以包括在所述MDE中的每个MDE中的单元;以及

用于确定所述MDE中的至少一些MDE均包括来自所述流的所述固定数量的音频帧的单元。

23. 根据权利要求19所述的源设备,其中,所述用于确定所述MDE的位置的单元包括:用于使用规则来确定所述MDE的位置的单元,包括:

用于接收表示一个或多个规则的数据的单元,所述一个或多个规则定义音频帧、视频帧以及定时文本实例中的一项或多项;

用于基于所述规则来确定至少一个音频帧、视频帧或者定时文本实例的位置的单元;

以及

用于基于所述至少一个音频帧、视频帧或者定时文本实例的位置来确定所述MDE的单元。

24. 根据权利要求19所述的源设备,其中,所述用于确定所述MDE的位置的单元包括:用于基于用于所述物理层发送单元的物理层同步的定时信息来确定所述MDE的位置的单元。

25. 根据权利要求19所述的源设备,其中,所述用于确定所述MDE的位置的单元包括:用于对提示轨道进行分析的单元,所述提示轨道包括表示所述流内可播放数据的位置的元数据。

26. 根据权利要求19所述的源设备,其中,所述用于确定所述MDE的所述一个或多个传输时间要求的单元包括:用于基于所述源设备的系统配置来确定用于所述MDE中的一个MDE的最早传输时间的单元。

27. 根据权利要求19所述的源设备,其中,所述用于确定所述MDE的所述一个或多个传输时间要求的单元包括:用于基于用于区段的区段可用时间来确定用于所述MDE中的一个MDE的最迟传输时间的单元,在所述MDE中的所述一个MDE中包括针对所述区段的数据,所述源设备还包括:用于根据数据的所述流的清单文件来确定区段可用时间的单元。

28. 一种其上存储有计算机程序的计算机可读存储介质,当由处理器执行时,所述计算机程序使得源设备的基于文件的协议发送单元进行操作,所述操作包括:

从所述源设备的形成区段的分段器接收包括媒体数据的区段的数据的流,所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件;

确定媒体传递事件 (MDE) 在媒体数据的所述流中的位置,其中,所述MDE包括针对所述区段中的一个区段的至少一部分的数据;

确定用于所述MDE的一个或多个传输时间要求,所述一个或多个传输时间要求表示要将所述MDE发送给客户端设备的时间;以及

根据所述源设备的物理层发送单元的可用传递时隙将所述MDE和表示所述MDE的所述一个或多个传输时间要求的数据提供给所述物理层发送单元。

29. 根据权利要求28所述的计算机可读存储介质,其中,确定所述MDE的位置包括:使用模式匹配来确定所述MDE的位置。

30. 根据权利要求29所述的计算机可读存储介质,其中,所述媒体数据包括视频数据,并且其中,使用模式匹配来确定所述MDE的位置包括:

获得定义所述流的视频帧的节奏的数据,其中,所定义的节奏表示发送所述视频帧的类型的顺序,所述视频帧的类型包括帧内预测视频帧和帧间预测视频帧;以及

基于所定义的节奏来确定所述MDE中的至少一些MDE,使得所述MDE在对应的视频帧边界之间。

31. 根据权利要求29所述的计算机可读存储介质,其中,所述媒体数据包括音频数据,并且其中,使用模式匹配来确定所述MDE的位置包括:

确定所述流的固定数量的音频帧以包括在所述MDE中的每个MDE中;以及

确定所述MDE中的至少一些MDE均包括来自所述流的所述固定数量的音频帧。

32. 根据权利要求28所述的计算机可读存储介质,其中,确定所述MDE的位置包括:使用

规则来确定所述MDE的位置,所述操作还包括:

接收表示一个或多个规则的数据,所述一个或多个规则定义音频帧、视频帧以及定时文本实例中的一项或多项;

基于所述规则来确定至少一个音频帧、视频帧或者定时文本实例的位置;以及

基于所述至少一个音频帧、视频帧或者定时文本实例的位置来确定所述MDE。

33.根据权利要求28所述的计算机可读存储介质,其中,确定所述MDE的位置包括:基于用于所述物理层发送单元的物理层同步的定时信息来确定所述MDE的位置。

34.根据权利要求28所述的计算机可读存储介质,其中,确定所述MDE的位置包括:对提示轨道进行分析,所述提示轨道包括表示所述流内可播放数据的位置的元数据。

35.根据权利要求28所述的计算机可读存储介质,其中,确定所述MDE的所述一个或多个传输时间要求包括:基于所述源设备的系统配置来确定用于所述MDE中的一个MDE的最早传输时间。

36.根据权利要求28所述的计算机可读存储介质,其中,确定所述MDE的所述一个或多个传输时间要求包括:基于用于区段的区段可用时间来确定用于所述MDE中的一个MDE的最迟传输时间,在所述MDE中的所述一个MDE中包括针对所述区段的数据,所述操作还包括:根据数据的所述流的清单文件来确定区段可用时间。

为媒体传输确定媒体传递事件位置的方法和装置

[0001] 本申请要求享有于2016年1月8日提交的美国临时申请No. 62/276,674的权益,其全部内容故此通过引用被并入本文。

技术领域

[0002] 本公开内容涉及媒体数据的传输。

背景技术

[0003] 数字音频和视频能力可以被并入到广泛的设备中,包括数字电视、数字直接广播系统、无线广播系统、个人数字助理 (PDA)、膝上型或台式计算机、数字相机、数字记录设备、数字媒体播放器、视频游戏设备、视频游戏控制台、蜂窝或卫星无线电话、视频电话会议设备等等。

[0004] 可以压缩媒体数据(例如,音频和视频数据)以用于传输。压缩通常包括对媒体数据的编码。在已对媒体数据进行编码之后,可以对视频数据进行打包以用于传输或存储。可以将媒体数据组装成符合各种标准(例如,国际标准化组织 (ISO) 基本媒体文件格式 (ISO BMFF) 及其扩展,例如AVC) 中的任何标准的媒体文件。

[0005] 可以使用各种技术经由计算机网络来传输媒体数据。例如,可以经由广播协议或单播协议来传递媒体数据。在广播协议中,服务器设备向多个订阅客户端设备发送媒体数据。在单播协议中,客户端设备从服务器设备请求媒体数据,并且服务器设备响应于该请求而传递媒体数据。

发明内容

[0006] 概括地说,本公开内容描述了与支持在例如经由ROUTE的媒体感知字节范围传递 (ATSC Working Draft:Signaling,Delivery,Synchronization, and Error Protection, 2015年11月20日) 中使用的媒体传递事件 (MDE) 和MMT (ISO/IEC:ISO/IEC 23008-1第2版,信息技术—High efficiency coding and media delivery in heterogeneous environments—第1部分:MPEG media transport (MMT) 或FLUTE (RFC 6726,其是ROUTE的适当子集) 相关的技术。这些传递技术还可以使用广播HTTP的动态自适应流式传输 (DASH)。共同媒体应用格式 (CMAF)、或者类似的区段或基于媒体感知字节范围的媒体传递方法。本公开内容还描述了用于确定检测到的MDE的时间标签的技术,以及用于在物理层处的自适应传递技术,这些技术考虑早期媒体传递相对于与媒体对齐的物理层成帧的影响。

[0007] 在一个例子中,一种用于传输媒体数据的方法包括由源设备的基于文件的协议发送单元进行以下操作:从所述源设备的形成区段的分段器接收包括媒体数据的区段的数据流,所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件;确定媒体传递事件 (MDE) 在所述媒体数据流中的位置,其中,所述MDE包括针对所述区段中的一个区段的至少一部分的数据;确定用于所述MDE的一个或多个传输时间要求,所述一个或多个传输时间要求表示要将所述MDE发送给客户端设

备的时间；以及根据所述源设备的物理层发送单元的可用传递时隙将所述MDE和表示所述传输时间要求的数据提供给所述物理层发送单元。

[0008] 在另一个例子中，一种用于传输媒体数据的设备包括分段器，其被配置为：形成包括媒体数据的区段的数据流，所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件，例如ISO BMFF格式。所述设备还包括物理层发送单元，其被配置为：根据用于所述MDE的传输时间要求将媒体传递事件 (MDE) 传递给客户端设备，其中，所述MDE包括针对所述区段中的一个区段的至少一部分的数据，其中，所述物理层发送单元被配置有用于接收要传递的数据的可用传递时隙。所述设备还包括基于文件的协议发送单元，其被配置为：从所述分段器接收包括所述媒体数据的区段的所述数据流；确定所述MDE在所述媒体数据流中的位置；确定用于所述MDE的所述传输时间要求中的一个或多个传输时间要求，所述一个或多个传输时间要求表示要将所述MDE发送给所述客户端设备的时间；以及根据所述物理层发送单元的所述可用传递时隙将所述MDE和表示所述传输时间要求的数据提供给所述物理层发送单元。

[0009] 在另一个例子中，一种用于传输媒体数据的设备包括分段器，其被配置为：形成包括媒体数据的区段的数据流，所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件。所述设备还包括物理层发送单元，其被配置为：根据用于所述MDE的传输时间要求将媒体传递事件 (MDE) 传递给客户端设备，其中，所述MDE包括针对所述区段中的一个区段的至少一部分的数据，其中，所述物理层发送单元被配置有用于接收要传递的数据的可用传递时隙。所述设备还包括：用于从所述分段器接收包括媒体数据的区段的数据流的单元；用于确定所述MDE在所述媒体数据流中的位置的单元；用于确定用于所述MDE的一个或多个传输时间要求的单元，所述一个或多个传输时间要求表示要将所述MDE发送给所述客户端设备的时间；以及用于根据所述物理层发送单元的可用传递时隙将所述MDE和表示所述传输时间要求的数据提供给所述物理层发送单元的单元。

[0010] 在另一个例子中，一种其上存储有指令的计算机可读存储介质，当所述指令被执行时使得源设备的基于文件的协议发送单元进行以下操作：从所述源设备的形成区段的分段器接收包括媒体数据的区段的数据流，所述区段中的每个区段包括与唯一的统一资源定位符 (URL) 或统一资源标识符 (URI) 相关联的对应可单独获取的文件；确定媒体传递事件 (MDE) 在所述媒体数据流中的位置，其中，所述MDE包括针对所述区段中的一个区段的至少一部分的数据；确定用于所述MDE的一个或多个传输时间要求，所述一个或多个传输时间要求表示要将所述MDE发送给客户端设备的时间；以及根据所述源设备的物理层发送单元的可用传递时隙将所述MDE和表示所述传输时间要求的数据提供给所述物理层发送单元。

[0011] 在附图和以下描述中阐述了一个或多个例子的细节。通过本描述和附图以及通过权利要求，其它特征、目标和优点将是显而易见的。

附图说明

[0012] 图1是示出了实现通过网络流式传输媒体数据的技术的示例性系统的框图。

[0013] 图2是更详细地示出了图1的获取单元52的示例性组件集合的框图。

[0014] 图3是示出了示例性多媒体内容的元素的概念图。

- [0015] 图4是示出了可以与表示的区段相对应的示例性视频文件的元素的框图。
- [0016] 图5是示出了示例性广播HTTP的动态自适应流式传输(DASH)基础设施系统的框图。
- [0017] 图6是示出了示例性MDE的高级组成的概念图。
- [0018] 图7是示出了示例性的帧类型节奏(cadence)和相应定义的MDE的概念图。
- [0019] 图8是表示系统时间的简单视图的概念图。
- [0020] 图9是表示发送基础设施中的时间关系的框图。
- [0021] 图10是示出了每个图片组(GOP)的具有交错的随机访问点(RAP)位置的多个区段的概念图。
- [0022] 图11是示出了用于媒体回放的解码时间的最小延迟的概念图。
- [0023] 图12是示出了可选的物理层同步和相关元数据的概念图。
- [0024] 图13是示出了获取事件的示例性序列的流程图。
- [0025] 图14是根据本公开内容的技术示出了另一个示例性方法的流程图。

具体实施方式

[0026] 概括地说,本公开内容描述了与传输媒体数据相关的技术。本公开内容的技术可以由服务器设备或将媒体数据流式传输到客户端设备的其它源设备来执行。具体而言,源设备可以包括分段器、基于文件的协议发送单元、以及物理层发送单元。分段器可以与形成媒体数据的区段的单元相对应,其中区段符合例如HTTP的动态自适应流式传输(DASH)区段。例如,分段器可以将网络抽象层(NAL)单元封装到对应的DASH区段中,其中区段均可以与唯一的统一资源定位符(URL)或统一资源标识符(URI)相关联。基于文件的协议发送单元可以被配置为:使用基于文件的协议(例如,单向传输的文件传递(FLUTE)或ROUTE)来发送数据。物理层发送单元可以在物理层例如使用以太网来发送媒体数据。

[0027] 根据本公开内容的技术,分段器可以以区段流的形式将媒体数据的两种区段提供给基于文件的协议发送单元。因此,基于文件的协议发送单元可以从分段器接收包括媒体数据的区段的数据流。基于文件的协议发送单元随后可以确定媒体传递事件(MDE)在媒体数据流中的位置。在一个例子中,基于文件的协议发送单元可以根据媒体数据流来形成MDE。例如,基于文件的协议发送单元可以确定流的哪些部分要包括在特定的MDE中,使得该MDE包括对于媒体解码器有用的数据(例如,能够被媒体解码器恰当地解码的数据,假设较早的数据正确地传递)。

[0028] 为了确定MDE的位置,基于文件的协议发送单元可以使用模式匹配技术,其中基于文件的协议发送单元被配置有用于流中所包括的媒体数据的一个或多个模式。这种模式可以表示例如以分层解码顺序对媒体数据的布置。例如,可以根据某一数量的帧内预测帧(I帧)、随后是某一数量的单向帧间预测帧(P帧)、随后是(或者间隔开)某一数量的双向帧间预测帧(B帧)来布置视频数据。基于文件的协议发送单元可以基于对I帧、P帧和B帧的检测来确定MDE的位置,并且例如确定MDE包括如下的视频序列:举一个例子,以第一I帧开始,并且以最后B帧结束,之后是后续的I帧。此外,基于文件的协议发送单元可以确定视频MDE与区段的特定字节范围相对应,其中该字节范围在帧边界处结束。举另一个例子,基于文件的协议发送单元可以确定音频MDE包括固定数量的音频帧。

[0029] 举另一个例子,为了确定MDE的位置(例如,以便形成MDE),可以根据规则来配置基于文件的协议发送单元,其中规则定义各种类型的媒体数据(例如音频数据、视频数据和/或定时文本数据)在流中的位置。规则可以定义例如各种类型的媒体数据在流中的布置,使得基于文件的协议发送单元可以自类型的媒体数据之间进行区分并且将各类型的媒体数据分配给对应的MDE。

[0030] 另外,基于文件的协议发送单元还可以确定用于MDE的一个或多个传递时间要求(还被称为传输时间要求),其中传输时间要求表示要将MDE发送给客户端设备的时间。例如,传输时间要求可以表示要将MDE发送给客户端设备的最早和/或最迟时间。MDE可以具有最早时间要求(可以以该最早时间要求将MDE发送给客户端设备)以便例如防止客户端设备处的缓冲器溢出。MDE也可以具有最迟时间要求(可以以该最迟时间要求将MDE发送给客户端设备)以便例如防止客户端设备处的缓冲器欠运行。

[0031] 基于文件的协议发送单元随后可以将MDE连同表示传输时间要求的数据发送给物理层发送单元。具体而言,基于文件的协议发送单元可以确定用于物理层发送单元的可用传递时隙,并且将MDE和传输时间要求数据发送给物理层发送单元,使得物理层发送单元能够根据传输时间要求来传递MDE。

[0032] 上面所讨论的源设备可以被配置成根据HTTP流式传输协议(例如,DASH)来流式传输数据。在HTTP流式传输中,频繁使用的操作包括HEAD、GET和部分GET。HEAD操作获取与给定的统一资源定位符(URL)、统一资源名称(URN)或统一资源标识符(URI)相关联的文件的报头,而不获取与URL、URN或URI相关联的有效载荷。GET操作获取与给定的URL、URN或URI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获取连续字节数量的文件,其中,字节数量对应于所接收的字节范围。因此,可以提供电影区段以用于HTTP流式传输,这是因为部分GET操作可以取得一个或多个单独的电影区段。在电影区段中,可以存在不同轨道的若干个轨道区段。在HTTP流式传输中,媒体呈现可以是客户端可访问的结构化数据集合。客户端可以请求并下载媒体数据信息,以向用户提供流式传输服务。

[0033] 在使用HTTP流式传输来流式传输3GPP数据的例子中,可能存在针对多媒体内容的视频和/或音频数据的多个表示(representation)。如下面说明的,不同的表示可以对应于不同的编码特性(例如,不同的视频编码标准简档或层级)、不同的编码标准或编码标准的扩展(例如,多视图和/或可缩放扩展)或者不同的比特速率。可以在媒体呈现描述(MPD)数据结构中定义这种表示的清单(manifest)。媒体呈现可以对应于HTTP流式传输客户端设备可访问的结构化数据集合。HTTP流式传输客户端设备可以请求并下载媒体数据信息,以向客户端设备的用户提供流式传输服务。可以在MPD数据结构(其可以包括MPD的更新)中描述媒体呈现。

[0034] 媒体呈现可以包含一个或多个时段的序列。可以由MPD中的Period(时段)元素来定义时段。每个时段可以具有MPD中的属性start(开始)。MPD可以包括针对每个时段的start属性和availableStartTime(可用开始时间)属性。对于实况服务,时段的start属性和MPD属性availableStartTime的和可以以UTC格式来指定时段的可用时间,特别是针对相应时段中的每个表示的第一媒体区段。对于按需服务,第一时段的start属性可以是0。对于任何其它时段,start属性可以指定相应时段的开始时间相对于第一时段的开始时间的时间偏移。每个时段可以扩展,直到下一时段的开始为止,或者在最后一个时段的情况下,直

到媒体呈现的结束为止。时段开始时间可以是精确的。时段开始时间可以反映由于播放所有先前时段的媒体而产生的实际定时。

[0035] 每个时段可以包含针对相同媒体内容的一个或多个表示。表示可以是音频或视频数据的多个替代的经编码版本中的一个版本。表示可以依据编码类型(例如,比特速率、分辨率、和/或针对视频数据和比特速率的编解码器、语言、和/或针对音频数据的编解码器)而不同。术语表示可以用于指代经编码的音频或视频数据的对应于多媒体内容的特定时段并以特定方式编码的部分。

[0036] 特定时段的表示可以分配给MPD中的属性(其指示表示所属的适配集)所指示的群组。相同适配集中的表示通常被视为彼此的替代,因为客户端设备可以动态地并且无缝地在这些表示之间切换,例如,以便执行带宽适配。例如,针对特定时段的视频数据的每个表示可以分配给相同的适配集,使得表示中的任何表示可以被选择用于解码,以呈现针对相应时段的多媒体内容的媒体数据(例如,视频数据或音频数据)。一个时段内的媒体内容可以由来自群组0的一个表示(如果存在的话)来表示,或者在一些例子中,由来自每个非零群组的最多一个表示的组合来表示。可以相对于时段的开始时间来表达针对时段的每个表示的定时数据。

[0037] 表示可以包括一个或多个区段。每个表示可以包括初始化区段,或者表示的每个区段可以是自初始化的。当存在初始化区段时,其可以包含用于访问表示的初始化信息。通常,初始化区段不包含媒体数据。可以用标识符(例如,统一资源定位符(URL)、统一资源名称(URN)或统一资源标识符(URI))来唯一地引用区段。MPD可以提供针对每个区段的标识符。在一些例子中,MPD还可以以range(范围)属性的形式来提供字节范围,其可以与针对可由URL、URN或URI访问的文件内的区段的数据相对应。

[0038] 可以选择不同的表示以用于大体上同时获取不同类型的媒体数据。例如,客户端设备可以选择从中获取区段的音频表示、视频表示以及定时文本表示。在一些例子中,客户端设备可以选择特定的适配集以用于执行带宽适配。即,客户端设备可以选择包括视频表示的适配集、包括音频表示的适配集、和/或包括定时文本的适配集。替代地,客户端设备可以选择针对某些类型的媒体(例如,视频)的适配集,并且直接地选择针对其它类型的媒体(例如,音频和/或定时文本)的表示。

[0039] 图1是示出了实现用于通过网络来流式传输媒体数据的技术的示例性系统10的框图。在该例子中,系统10包括内容准备设备20、服务器设备60和客户端设备40。客户端设备40和服务器设备60通过网络74(其可以包括互联网)通信地耦合。在一些例子中,内容准备设备20和服务器设备60也可以通过网络74或另一个网络耦合,或者可以直接通信地耦合。在一些例子中,内容准备设备20和服务器设备60可以包括相同的设备。

[0040] 在图1的例子中,内容准备设备20包括音频源22和视频源24。音频源22可以包括例如麦克风,其中麦克风产生电信号,所述电信号表示所捕获的要由音频编码器26编码的音频数据。替代地,音频源22可以包括:存储介质,其存储先前记录的音频数据;音频数据生成器,例如计算机化的合成器;或者任何其它的音频数据源。视频源24可以包括:视频相机,其产生要由视频编码器28进行编码的视频数据;编码有先前记录的视频数据的存储介质;视频数据生成单元,例如计算机图形源;或者任何其它的视频数据源。内容准备设备20不一定在所有的例子中都通信地耦合到服务器设备60,但是可以将多媒体内容存储到由服务器设

备60读取的单独介质。

[0041] 原始音频和视频数据可以包括模拟或数字数据。可以在由音频编码器26和/或视频编码器28对模拟数据进行编码之前对其进行数字化。当说话参与者正在说话时,音频源22可以从该说话参与者获得音频数据,并且视频源24可以同时获得该说话参与者的视频数据。在其它例子中,音频源22可以包括包含所存储的音频数据的计算机可读存储介质,并且视频源24可以包括包含所存储的视频数据计算机可读存储介质。以此方式,本公开内容中所描述的技术可以应用于实况的、流式传输的、实时的音频和视频数据,或者应用于经存档的、预先记录的音频和视频数据。

[0042] 与视频帧相对应的音频帧通常是包含音频数据的音频帧,其中该音频数据是由音频源22与包含在视频帧内的由视频源24捕获(或生成)的视频数据同时捕获(或生成)的。例如,当说话参与者通常通过说话来产生音频数据时,音频源22捕获音频数据,并且视频源24同时(即,当音频源22正在捕获音频数据时)捕获说话参与者的视频数据。因此,音频帧可以在时间上对应于一个或多个特定的视频帧。因此,与视频帧相对应的音频帧通常对应于以下情形:在该情形中,同时捕获音频数据和视频数据,并且对于该情形,音频帧和视频帧分别包括同时捕获的音频数据和视频数据。

[0043] 在一些例子中,音频编码器26可以将时间戳编码到每个经编码的音频帧中,其中该时间戳表示用于经编码的音频帧的音频数据被记录的时间,并且类似地,视频编码器28可以将时间戳编码到每个经编码的视频帧中,其中该时间戳表示用于经编码的视频帧的视频数据被记录的时间。在这些例子中,与视频帧相对应的音频帧可以包括:包括时间戳的音频帧,以及包括相同时间戳的视频帧。内容准备设备20可以包括内部时钟,其中音频编码器26和/或视频编码器28可以根据该内部时钟来生成时间戳,或者音频源22和视频源24可以使用该内部时钟来分别将音频和视频数据与时间戳相关联。

[0044] 在一些例子中,音频源22可以向音频编码器26发送与音频数据被记录的时间相对应的数据,并且视频源24可以向视频编码器28发送与视频数据被记录的时间相对应的数据。在一些例子中,音频编码器26可以将序列标识符编码到经编码的音频数据中,以指示经编码的音频数据的相对时间排序,而不必指示音频数据被记录的绝对时间,并且类似地,视频编码器28也可以使用序列标识符来指示经编码的视频数据的相对时间排序。类似地,在一些例子中,序列标识符可以被映射或者以其它方式与时间戳相关。

[0045] 音频编码器26通常产生经编码的音频数据流,而视频编码器28产生经编码的视频数据流。每个单独的数据流(无论是音频还是视频)可以被称为基本流。基本流是表示的单个的、经数字编码(可能经压缩)的分量。例如,表示的经编码的视频或音频部分可以是基本流。基本流可以在被封装到视频文件内之前转换为打包的基本流(PES)。在相同的表示内,可以使用流ID来将属于一个基本流的PES分组与属于其它基本流的PES分组进行区分。基本流的基本数据单元是打包的基本流(PES)分组。因此,经编码的视频数据通常对应于基本视频流。类似地,音频数据对应于一个或多个对应的基本流。

[0046] 在图1的例子中,内容准备设备20的封装单元30从视频编码器28接收包括经编码的视频数据的基本流,并且从音频编码器26接收包括经编码的音频数据的基本流。在一些例子中,视频编码器28和音频编码器26均可以包括打包器(packetizer)以用于根据经编码的数据来形成PES分组。在其它例子中,视频编码器28和音频编码器26均可以与用于根据经

编码的数据来形成PES分组的对应打包器对接。在其它例子中,封装单元30可以包括打包器,以用于根据经编码的音频和视频数据来形成PES分组。

[0047] 视频编码器28可以以各种方式来对多媒体内容的视频数据进行编码,以在各种比特速率下并且利用各种特性(例如,像素分辨率、帧速率、对各种编码标准的符合性、对用于各种编码标准的各种简档和/或简档层级的符合性、具有一个或多个视图的表示(例如,针对二维或三维回放),或其它此类特性)来产生多媒体内容的不同表示。如本公开内容中所使用的,表示可以包括以下各项中的一项:音频数据、视频数据、文本数据(例如,用于隐藏式字幕),或者其它此类数据。表示可以包括基本流,例如音频基本流或视频基本流。每个PES分组可以包括stream_id(流_id),其标识该PES分组所属的基本流。封装单元30负责将基本流组装成各种表示的视频文件(例如,区段)。

[0048] 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分组,并且根据PES分组来形成相应的网络抽象层(NAL)单元。在H.264/AVC(高级视频编码)的例子中,将经编码的视频区段组织成NAL单元,其中NAL单元提供处理例如视频电话、存储、广播或流式传输等应用的“网络友好的”视频表示。NAL单元可以分类为视频编码层(VCL)NAL单元和非VCL NAL单元。VCL单元可以包含核心压缩引擎并且可以包括块、宏块和/或切片层级数据。其它NAL单元可以是非VCL NAL单元。在一些例子中,一个时间实例中的经编码图片(通常呈现为主要经编码图片)可以包含在访问单元中,其中访问单元可以包括一个或多个NAL单元。

[0049] 非VCL NAL单元可以包括参数集NAL单元和SEI NAL单元等等。参数集可以包含序列层级报头信息(在序列参数集(SPS)中)以及很少改变的图片层级报头信息(在图片参数集(PPS)中)。利用参数集(例如,PPS和SPS),不需要针对每个序列或图片重复很少改变的信息,因此可以提高编码效率。此外,参数集的使用可以实现重要报头信息的带外传输,从而避免了需要针对错误复原而进行冗余传输。在带外传输的例子中,可以在与其它NAL单元(例如,SEI NAL单元)不同的频道上发送参数集NAL单元。

[0050] 补充增强信息(SEI)可以包含对于解码来自VCL NAL单元的经编码图片样本来说不是必要的但是可以辅助与解码、显示、错误恢复以及其它目的相关的过程的信息。SEI消息可以包含在非VCL NAL单元中。SEI消息是一些标准规范的规范性部分,并且因此对于符合标准的解码器实现方式来说并非总是强制的。SEI消息可以是序列层级SEI消息或图片层级SEI消息。一些序列层级信息可以包含在SEI消息中,例如,在SVC的例子中的可缩放信息SEI消息,以及在MVC中的视图可缩放信息SEI消息。这些示例性的SEI消息可以传送关于例如操作点的提取和操作点的特性的信息。另外,封装单元30可以形成清单文件,例如,对表示的特性进行描述的媒体呈现描述符(MPD)。封装单元30可以根据可扩展标记语言(XML)来格式化MPD。

[0051] 封装单元30可以将用于多媒体内容的一个或多个表示的数据连同清单文件(例如,MPD)提供给输出接口32。输出接口32可以包括网络接口或者用于向存储介质写入的接口,例如,通用串行总线(USB)接口、CD或DVD写入器或烧录器、到磁性或闪速存储介质的接口、或者用于存储或发送媒体数据的其它接口。封装单元30可以将多媒体内容的表示中的每个表示的数据提供给输出接口32,其中输出接口32可以经由网络传输或存储介质来将数据发送给服务器设备60。在图1的例子中,服务器设备60包括存储各种多媒体内容64的存储

介质62,其中每个多媒体内容64包括对应的清单文件66以及一个或多个表示68A-68N(表示68)。在一些例子中,输出接口32还可以将数据直接发送给网络74。

[0052] 在一些例子中,表示68可以被分离成适配集。即,表示68的各个子集可以包括对应的共同特性集合,例如,编解码器、简档和层级、分辨率、视图数量、区段的文件格式、文本类型信息(其可以对要利用表示来显示的文本和/或要由例如说话者解码并呈现的音频数据的语言或其它特性进行标识)、相机角度信息(其可以描述针对适配集中的表示的场景的相机角度或真实世界相机视角)、描述针对特定观众的内容合适性的分级信息等。

[0053] 清单文件66可以包括对与特定适配集相对应的表示68的子集以及针对适配集的共同特性进行指示的数据。清单文件66还可以包括对针对适配集的单独表示的单独特性(例如,比特速率)进行表示的数据。以此方式,适配集可以提供简化的网络带宽适配。可以使用清单文件66的适配集元素的子元素来指示适配集中的表示。

[0054] 服务器设备60包括请求处理单元70和网络接口72。在一些例子中,服务器设备60可以包括多个网络接口。此外,可以在内容传递网络的其它设备(例如,路由器、桥接器、代理设备、交换机或其它设备)上实现服务器设备60的特征中的任何或所有特征。在一些例子中,内容传递网络的中间设备可以高速缓存多媒体内容64的数据,并且包括大体上符合服务器设备60的组件的组件。通常,网络接口72被配置为经由网络74来发送和接收数据。

[0055] 请求处理单元70被配置为从客户端设备(例如,客户端设备40)接收针对存储介质62的数据的网络请求。例如,请求处理单元70可以实现如RFC 2616,1999年6月,IETF,网络工作组,R.Fielding等人的“超文本传输协议-HTTP/1.1”中所描述的超文本传输协议(HTTP)版本1.1。即,请求处理单元70可以被配置为接收HTTP GET或部分GET请求并且响应于该请求而提供多媒体内容64的数据。请求可以指定表示68中的一个表示的区段(例如,使用该区段的URL或URI)。在一些例子中,请求还可以指定区段的一个或多个字节范围,因此包括部分GET请求。请求处理单元70还可以被配置为对HTTP HEAD请求进行服务,以提供表示68中的一个表示的区段的报头数据。在任何情况下,请求处理单元70可以被配置为对请求进行处理,以向请求设备(例如,客户端设备40)提供所请求的数据。

[0056] 另外地或替代地,请求处理单元70可以被配置为经由广播或多播协议(例如,eMBMS或DVB-T/T2)来传递媒体数据。内容准备设备20可以以与所描述的方式大体上相同的方式来创建DASH区段和/或子区段,但是服务器设备60可以使用eMBMS、DVB-T/T2或另一个广播或多播网络传输协议来传递这些区段或子区段。例如,请求处理单元70可以被配置为从客户端设备40接收多播群组加入请求。即,服务器设备60可以向与特定媒体内容(例如,实况事件的广播)相关联的客户端设备(包括客户端设备40)通告与多播群组相关联的互联网协议(IP)地址。客户端设备40转而可以提交加入多播群组的请求。可以在整个网络74(例如,构成网络74的路由器)上传播该请求,使得促使路由器将以关联于多播群组的IP地址为目的地的业务引导到订阅客户端设备(例如,客户端设备40)。

[0057] 此外,服务器设备60的网络接口72可以被配置为执行本公开内容的技术。另外地或替代地,请求处理单元70和网络接口72可以被配置为执行本公开内容的技术。出于说明的目的,假设网络接口72包括图1中未显式示出的子组件,例如下面更详细讨论的各个例子。例如,网络接口72可以包括发送机、调度器、成帧器、以及激励器放大器。另外,在该例子中,封装单元30表示分段器的例子(即,形成媒体数据的区段的单元,其中区段表示与对应

的URL/URI相关联的单独文件)。具体而言,封装单元30可以将MDE封装在区段中,其中每个区段可以包括一个或多个MDE。

[0058] 根据本公开内容的技术,封装单元30可以向服务器设备60提供媒体数据的区段流,其中服务器设备60最终将区段流引导到网络接口72以用于传输至客户端设备40。区段流可以对应于表示,例如表示68中的一个表示。此外,封装单元30可以向服务器设备60提供表示针对区段中所包括的MDE的传输时间要求的数据,其中服务器设备60再次最终将传输时间要求引导到网络接口72。例如,可以在清单文件66内指定传输时间要求。传输时间要求通常指示要将MDE传递给客户端设备40的最早和/或最迟时间。因此,网络接口72可以根据传输时间要求向客户端设备40发送MDE(例如,区段或区段的字节范围)。例如,网络接口72可以传递MDE以使得MDE不早于最早传输时间和/或不迟于最迟传输时间抵达客户端设备40处。

[0059] 如图1的例子中所示出的,多媒体内容64包括清单文件66,其中清单文件66可以对应于媒体呈现描述(MPD)。清单文件66可以包含对不同的替代表示68(例如,具有不同质量的视频服务)的描述,并且该描述可以包括例如编解码器信息、简档值、层级值、比特速率以及表示68的其它描述性特性。客户端设备40可以获取媒体呈现的MPD,以确定如何访问表示68的区段。

[0060] 具体而言,获取单元52可以获取客户端设备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的渲染能力。配置数据还可以包括以下各项中的任何或所有项:客户端设备40的用户所选择的语言偏好、与客户端设备40的用户所设置的深度偏好相对应的一个或多个相机视角、和/或客户端设备40的用户所选择的分级偏好。获取单元52可以包括例如被配置为提交HTTP GET和部分GET请求的web浏览器或媒体客户端。获取单元52可以与客户端设备40的一个或多个处理器或处理单元(未示出)所执行的软件指令相对应。在一些例子中,可以在硬件、或者硬件、软件、和/或固件的组合中实现关于获取单元52所描述的功能中的全部或部分功能,其中,可以提供必要的硬件来执行针对软件或固件的指令。

[0061] 获取单元52可以将客户端设备40的解码和渲染能力与清单文件66的信息所指示的表示68的特性进行比较。获取单元52可以初始地获取清单文件66的至少一部分以确定表示68的特性。例如,获取单元52可以请求清单文件66的描述一个或多个适配集的特性的部分。获取单元52可以选择表示68的子集(例如,适配集),该子集具有客户端设备40的编码和渲染能力能够满足的特性。获取单元52然后可以确定针对适配集中的表示的比特速率,确定当前可用的网络带宽量,以及从表示(其具有网络带宽能够满足的比特速率)中的一个表示中获取区段。

[0062] 通常,较高比特速率的表示可以产生较高质量的视频回放,而当可用的网络带宽减小时,较低比特速率的表示可以提供足够质量的视频回放。因此,当可用的网络带宽相对高时,获取单元52可以从相对高的比特速率的表示中获取数据,而当可用的网络带宽低时,获取单元52可以从相对低的比特速率的表示中获取数据。以此方式,客户端设备40可以通过网络74来流式传输多媒体数据,同时也适应网络74的变化的网络带宽可用性。

[0063] 另外地或替代地,获取单元52可以被配置为根据广播或多播网络协议(例如,eMBMS或IP多播)来接收数据。在这些例子中,获取单元52可以提交加入与特定媒体内容相

关联的多播网络群组的请求。在加入多播群组之后,获取单元52可以接收多播群组的数据而无需向服务器设备60或内容准备设备20发出进一步的请求。当不再需要多播群组的数据时,获取单元52可以提交离开多播群组的请求,例如,以便停止回放或将频道改变到不同的多播群组。

[0064] 网络接口54可以接收所选择的表示的区段的数据并将其提供给获取单元52,获取单元52转而可以将区段提供给解封装单元50。解封装单元50可以将视频文件的元素解封装为组成PES流,对PES流进行解包以获取经编码的数据,以及将经编码的数据发送给音频解码器46或视频解码器48,这取决于经编码的数据是音频流还是视频流的一部分(例如,如流的PES分组报头所指示的)。音频解码器46对经编码的音频数据进行解码并将经解码的音频数据发送给音频输出42,而视频解码器48对经编码的视频数据进行解码并将经解码的视频数据(其可以包括流的多个视图)发送给视频输出44。

[0065] 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、获取单元52以及解封装单元50均可以视适用情况实现为各种适当的处理电路中的任何处理电路,例如,一个或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、分立逻辑电路、软件、硬件、固件或者其任何组合。视频编码器28和视频解码器48均可以包括在一个或多个编码器或解码器中,其中任一项可以集成为组合的视频编码器/解码器(CODEC)的一部分。同样地,音频编码器26和音频解码器46均可以包括在一个或多个编码器或解码器中,其中任一项可以集成为组合的CODEC的一部分。包括视频编码器28、视频解码器48、音频编码器26、音频解码器46、封装单元30、获取单元52和/或解封装单元50的装置可以包括集成电路、微处理器和/或无线通信设备(例如,蜂窝电话)。

[0066] 客户端设备40、服务器设备60和/或内容准备设备20可以被配置为根据本公开内容的技术来操作。出于举例的目的,本公开内容关于客户端设备40和服务器设备60来描述这些技术。但是,应当理解的是,作为服务器设备60的替代(或者除服务器设备60之外),内容准备设备20也可以被配置为执行这些技术。

[0067] 封装单元30可以形成NAL单元,其包括标识NAL单元所属的节目的报头以及有效载荷,例如音频数据、视频数据或者描述NAL单元所对应的传输或节目流的数据。例如,在H.264/AVC中,NAL单元包括1个字节的报头和变化大小的有效载荷。在其有效载荷中包括视频数据的NAL单元可以包括各种粒度水平的视频数据。例如,NAL单元可以包括视频数据块、多个块、视频数据切片或者视频数据的整个图片。封装单元30可以从视频编码器28接收具有基本流的PES分组形式的经编码的视频数据。封装单元30可以将每个基本流与相应的节目相关联。

[0068] 封装单元30还可以根据多个NAL单元来组装访问单元。通常,访问单元可以包括用于表示视频数据帧的一个或多个NAL单元,以及对应于该帧的音频数据(当该音频数据可用时)。访问单元通常包括针对一个输出时间实例的所有NAL单元,例如,针对一个时间实例的所有音频和视频数据。例如,如果每个视图具有每秒20帧(fps)的帧速率,则每个时间实例可以对应于0.05秒的时间间隔。在该时间间隔期间,可以同时渲染针对相同访问单元(相同时间实例)的所有视图的特定帧。在一个例子中,访问单元可以包括一个时间实例中的经编码图片,其可以被呈现为主要经编码图片。

[0069] 因此,访问单元可以包括共同时间实例的所有音频帧和视频帧,例如对应于时间X

的所有视图。本公开内容还将特定视图的经编码图片称为“视图分量”。即，视图分量可以包括针对特定时间处的特定视图的经编码图片(或帧)。因此，访问单元可以定义为包括共同时间实例的所有视图分量。访问单元的解码顺序不必与输出顺序或显示顺序相同。

[0070] 媒体呈现可以包括媒体呈现描述(MPD)，其可以包含对不同的替代表示(例如，具有不同质量的视频服务)的描述，并且该描述可以包括例如编解码器信息、简档值和层级值。MPD是清单文件(例如，清单文件66)的一个例子。客户端设备40可以获取媒体呈现的MPD，以确定如何访问各种表示的电影区段。电影区段可以位于视频文件的电影区段盒(moof盒)中。

[0071] 清单文件66(其可以包括例如MPD)可以对表示68的区段的可用性进行通告。即，MPD可以包括对表示68中的一个表示的第一区段变为可用的挂钟时间进行指示的信息，以及对表示68内的区段的持续时间进行指示的信息。以此方式，客户端设备40的获取单元52可以基于在特定区段之前的区段的开始时间以及持续时间来确定每个区段何时可用。

[0072] 在封装单元30已经基于接收到的数据将NAL单元和/或访问单元组装成视频文件之后，封装单元30将视频文件传递给输出接口32以用于输出。在一些例子中，封装单元30可以本地地存储视频文件或者经由输出接口32将视频文件发送给远程服务器，而不是直接地将视频文件发送给客户端设备40。输出接口32可以包括例如发射机、收发机、用于向计算机可读介质(例如，光学驱动器、磁性介质驱动器(例如，软盘驱动器))写入数据的设备、通用串行总线(USB)端口、网络接口或者其它输出接口。输出接口32将视频文件输出到计算机可读介质，例如传输信号、磁性介质、光学介质、存储器、闪速驱动器或者其它计算机可读介质。

[0073] 网络接口54可以经由网络74来接收NAL单元或访问单元，并且经由获取单元52将NAL单元或访问单元提供给解封装单元50。解封装单元50可以将视频文件的元素解封装为组成PES流(consituent PES stream)，对PES流进行解包以获取经编码的数据，并且将经编码的数据发送给音频解码器46或视频解码器48，这取决于经编码的数据是音频流还是视频流的一部分(例如，如流的PES分组报头所指示的)。音频解码器46对经编码的音频数据进行解码并将经解码的音频数据发送给音频输出42，而视频解码器48对经编码的视频数据进行解码并将经解码的视频数据(其可以包括流的多个视图)发送给视频输出44。

[0074] 图1的例子的某些元件可以被配置为执行本公开内容的技术。例如，网络接口72和/或请求处理单元70可以被配置为独立地或相结合地操作以执行本公开内容的技术。

[0075] 图2是更详细地示出了图1的获取单元52的示例性组件集合的框图。在该例子中，获取单元52包括eMBMS中间件单元100、DASH客户端110、以及媒体应用112。

[0076] 在该例子中，eMBMS中间件单元100还包括eMBMS接收单元106、高速缓存104、以及服务器单元102。在该例子中，eMBMS接收单元106被配置为例如根据单向传输文件传递(FLUTE)(在可从<http://tools.ietf.org/html/rfc6726>获得的网络工作组，RFC 6726，2012年11月，T.Paila等人的“FLUTE-File Delivery over Unidirectional Transport”中描述的)、经由eMBMS来接收数据。即，eMBMS接收单元106可以从例如服务器设备60(其可以充当BM-SC)经由广播接收文件。

[0077] eMBMS中间件单元100接收用于文件的数据，eMBMS中间件单元可以将接收到的数据存储在高速缓存104中。高速缓存104可以包括计算机可读存储介质，例如闪存、硬盘、

RAM、或者任何其它适当的存储介质。

[0078] 本地服务器单元102可以充当用于DASH客户端110的服务器。例如，本地服务器单元102可以向DASH客户端110提供MPD文件或其它清单文件。本地服务器单元102可以通告MPD文件中区段的可用时间、以及能够从中获得区段的超链接。这些超链接可以包括与客户端设备40相对应的本地主机地址前缀(例如，针对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请求区段。例如，对于可从链接http://127.0.0.1/rep1/seg3获得的区段，DASH客户端110可以构造包括针对http://127.0.0.1/rep1/seg3的请求的HTTP GET请求，并将该请求提交给本地服务器单元102。响应于这些请求，本地服务器单元102可以从高速缓存104获得所请求的数据并将数据提供给DASH客户端110。

[0079] 图3是示出了示例性多媒体内容120的元素的概念图。多媒体内容120可以与多媒体内容64(图1)或存储在存储器介质62中的另一个多媒体内容相对应。在图3的例子中，多媒体内容120包括媒体呈现描述(MPD)122以及多个表示124A-124N(表示124)。表示124A包括可选的报头数据126和区段128A-128N(区段128)，而表示124N包括可选的报头数据130和区段132A-132N(区段132)。为了方便起见，使用字母N来标示表示124中的每个表示中的最后一个电影区段。在一些例子中，在表示124之间可能存在不同数量的电影区段。

[0080] MPD 122可以包括与表示124分离的数据结构。MPD 122可以对应于图1的清单文件66。同样地，表示124可以对应于图2的表示68。通常，MPD 122可以包括通常描述表示124的特性(例如，编码和渲染特性、适配集、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，对包括时间子序列的表示进行指示的信息)和/或用于获取远程时段的信息(例如，用于在回放期间将目标广告插入到媒体内容中))的数据。

[0081] 报头数据126(当存在时)可以描述区段128的特性，例如，随机访问点(RAP，还被称为流访问点(SAP))的时间位置，区段128中的哪个区段包括随机访问点，到区段128内的随机访问点的字节偏移，区段128的统一资源定位符(URL)或URI，或者区段128的其它方面。报头数据130(当存在时)可以描述区段132的类似特性。另外地或替代地，这些特性可以完全地包括在MPD 122内。

[0082] 区段128、124包括一个或多个经编码的视频样本，其中每个视频样本可以包括视频数据帧或切片。区段128的经编码的视频样本中的每个视频样本可以具有类似的特性，例如，高度、宽度以及带宽要求。可以由MPD 122的数据来描述这些特性，尽管图3的例子中没有示出这种数据。MPD 122可以包括如由3GPP规范描述的特性，外加本公开内容中所描述的用信号发送的信息中的任何或所有信息。

[0083] 区段128、132中的每个区段可以与唯一的统一资源定位符(URL)或URI相关联。因此，可以使用流式传输网络协议(例如，DASH)来独立地获取区段128、132中的每个区段。以此方式，目的地设备(例如，客户端设备40)可以使用HTTP GET请求来获取区段128或132。在一些例子中，客户端设备40可以使用HTTP部分GET请求来获取区段128或132的特定字节范围。

[0084] 图4是示出了可以与表示的区段(例如，图3的区段114、124中的一个区段)相对应的示例性视频文件150的元素的框图。虽然图4的例子示出了视频文件，但应当理解，音频文件或其它文件可以包括与视频文件150的数据类似的数据。区段128、132中的每个区段可以

包括大体上符合图4的例子中所示出的数据的布置的数据。视频文件150可以被认为对区段进行封装。如上面所描述的,根据ISO基本媒体文件格式及其扩展的视频文件将数据存储在一系列对象中,被称为“盒(box)”。在图4的例子中,视频文件150包括文件类型(FTYP)盒152、电影(MOOV)盒154、区段索引(sidx)盒162、电影区段(MOOF)盒164、以及电影区段随机访问(MFRA)盒166。虽然图4表示视频文件的例子,但应当理解,其它媒体文件可以包括其它类型的媒体数据(例如,音频数据、定时本文数据等等),该媒体数据根据ISO基本媒体文件格式及其扩展、与视频文件150的数据类似地构造。

[0085] 文件类型(FTYP)盒152通常描述视频文件150的文件类型。文件类型盒152可以包括标识描述针对视频文件105的最佳使用的规范的数据。文件类型盒152可以替代地位于MOOV盒154、电影区段盒164和/或MFRA盒166之前。

[0086] 在一些例子中,区段(例如,视频文件150)可以在FTYP盒152之前包括MPD更新盒(未示出)。MPD更新盒可以包括指示与包括要更新的视频文件150的表示相对应的MPD的信息,连同用于对MPD进行更新的信息。例如,MPD更新盒可以提供要用于对MPD进行更新的资源的URI或URL。举另一个例子,MPD更新盒可以包括用于对MPD进行更新的数据。在一些例子中,MPD更新盒可以紧接在视频文件150的区段类型(STYP)盒(未示出)之后,其中STYP盒可以定义视频文件150的区段类型。下面更详细讨论的图7提供了关于MPD更新盒的另外信息。

[0087] 在图4的例子中,MOOV盒154包括电影报头(MVHD)盒156、轨道(TRAK)盒158、以及一个或多个电影扩展(MVEX)盒160。通常,MVHD盒156可以描述视频文件150的一般特性。例如,MVHD盒156可以包括描述以下各项的数据:视频文件150何时被初始地创建、视频文件150何时被最后修改、视频文件150的时间刻度、视频文件150的回放持续时间、或者一般性描述视频文件150的其它数据。

[0088] TRAK盒158可以包括用于视频文件150的轨道的数据。TRAK盒158可以包括轨道报头(TKHD)盒,其描述与TRAK盒158相对应的轨道的特性。在一些例子中,TRAK盒158可以包括经编码的视频图片,而在其它例子中,轨道的经编码的视频图片可以包括在电影区段164中,其中电影区段164可以由TRAK盒158和/或sidx盒162的数据引用。

[0089] 在一些例子中,视频文件150可以包括一个以上轨道。因此,MOOV盒154可以包括与视频文件150中的轨道数量相等数量的TRAK盒。TRAK盒158可以描述视频文件150的相应轨道的特性。例如,TRAK盒158可以描述相应轨道的时间和/或空间信息。当封装单元30(图3)包括视频文件(例如,视频文件150)中的参数集轨道时,与MOOV盒154的TRAK盒158类似的TRAK盒可以描述参数集轨道的特性。封装单元30可以在描述参数集轨道的TRAK盒内的参数集轨道中用信号通知序列级SEI消息的存在。

[0090] MVEX盒160可以描述相应的电影区段164的特性,以例如用信号通知除了包括在MOOV盒154内所包括的视频数据(如果有的话)之外,视频文件150还包括电影区段164。在流式传输视频数据的上下文中,经编码的视频图片可以包括在电影区段164中而不是MOOV盒154中。因此,所有经编码的视频样本可以包括在电影区段164中而不是MOOV盒154中。

[0091] MOOV盒154可以包括与视频文件150中的电影区段164的数量相等数量的MVEX盒160。MVEX盒160中的每个MVEX盒可以描述电影区段164中的相应一个电影区段的特性。例如,每个MVEX盒可以包括电影扩展报头盒(MEHD)盒,该MEHD盒描述电影区段164中的相应一个电影区段在时间上的持续时间。

[0092] 如上面提到的,封装单元30可以在不包括实际经编码的视频数据的视频样本中存储序列数据集。视频样本通常可以对应于访问单元,该访问单元是对在特定时间实例处经编码的图片的表示。在AVC的上下文中,经编码的图片包括包含用于构造访问单元的所有像素的信息的一个或多个VCL NAL单元,以及其它相关联的非VCL NAL单元,例如SEI消息。因此,封装单元30可以在电影区段164中的一个电影区段中包括序列数据集,其可以包括序列级SEI消息。封装单元30还可以用信号通知序列数据集和/或序列级SEI消息的存在,其存在于与一个电影区段164相对应的一个MVEX盒160内的一个电影区段164中。

[0093] SIDX盒162是视频文件150的可选元素。即,符合3GPP文件格式或其它此类文件格式的视频文件不必包括SIDX盒162。根据3GPP文件格式的例子,SIDX盒可以用于标识区段(例如,包含在视频文件150内的区段)的子区段。3GPP文件格式将子区段定义为“具有相应媒体数据盒的一个或多个连续电影区段盒的自包含集合,并且包含由电影区段盒引用的数据的媒体数据盒必须在该电影区段盒之后并且在包含关于相同轨道的信息的下一个电影区段盒之前”。3GPP文件格式还指示SIDX盒“包含对由该盒记录的(子)区段的子区段的引用序列。被引用的子区段在呈现时间上是连续的。类似地,由区段索引盒所引用的字节在区段内总是连续的。被引用的大小给出了所引用的材料中的字节数量的计数。”

[0094] SIDX盒162通常提供表示视频文件150中所包括的区段的一个或多个子区段的信息。例如,此类信息可以包括:子区段开始和/或结束的回放时间、子区段的字节偏移、子区段是否包括(例如,开始于)流访问点(SAP)、SAP的类型(例如,SAP是否是瞬时解码器刷新(IDR)图片、干净随机访问(CRA)图片、断链访问(BLA)图片等)、SAP在子区段中的位置(在回放时间和/或字节偏移方面)等等。

[0095] 电影区段164可以包括一个或多个经编码的视频图片。在一些例子中,电影区段164可以包括一个或多个图片组(GOP),其中每个GOP可以包括多个经编码的视频图片,例如,帧或图片。另外,如上面所描述的,在一些例子中电影区段164可以包括序列数据集。电影区段164中的每个电影区段可以包括电影区段报头盒(MFHD,图4中未示出)。MFHD盒可以描述相应的电影区段的特性,例如电影区段的序列号。电影区段164可以按序列号顺序包括在视频文件150中。

[0096] MFRA盒166可以描述视频文件150的电影区段164内的随机访问点。这可以协助执行特技模式,例如执行寻找由视频文件150所封装的区段内的特定时间位置(即,回放时间)。MFRA盒166通常是可选的并且在一些例子中不需要包括在视频文件中。类似地,客户端设备(例如,客户端设备40)不一定需要参考MFRA盒166来正确地解码和显示视频文件150的视频数据。MFRA盒166可以包括与视频文件150的轨道数量、或者在一些例子中与视频文件150的媒体轨道(例如,无提示轨道)数量相等数量的轨道区段随机访问(TFRA)盒(未示出)。

[0097] 在一些例子中,电影区段164可以包括一个或多个流访问点(SAP),例如IDR图片。类似地,MFRA盒166可以提供对SAP在视频文件150内的位置的指示。因此,可以根据视频文件150的SAP来形成视频文件150的时间子序列。时间子序列还可以包括其它图片,例如依赖于SAP的P帧和/或B帧。时间子序列的帧和/或切片可以布置在区段内,使得依赖于子序列的其它帧/切片的时间子序列的帧/切片可以被恰当地解码。例如,在数据的分层布置中,用于其它数据的预测的数据也可以包括在时间子序列中。

[0098] 图5是示出了示例性的广播DASH基础设施系统180的框图。系统180包括系统管理

单元198、GOP持续时间媒体缓冲器182、媒体编码器184、分段器186、发送机188、调度器190、成帧器192、以及激励器/放大器194。系统180以及因此系统180的这些组件中的每个组件可以包括在源设备中,例如图1的服务器设备60或图1的内容准备设备20。再次,服务器设备和内容准备设备可以在功能上集成到单个设备中。

[0099] 在系统180的例子中,GOP持续时间媒体缓冲器182接收并缓冲要编码的媒体数据。媒体编码器184(其可以实际上表示多个媒体编码器,例如图1的音频编码器26和图1的视频编码器28)对媒体数据(例如,音频数据、视频数据和/或定时文本数据)进行编码,将媒体数据封装成网络抽象层(NAL)单元,并将处于NAL单元形式中经编码的媒体数据202提供给分段器186。另外,媒体编码器184将表示显式MDE标识符200的数据发送给分段器186。

[0100] 分段器186将接收到的NAL单元封装成对应的区段。因此,通常,分段器186不提供指示MDE在区段内的位置的显式信息。然而,有效地,分段器186以区段内的字节范围的形式将MDE提供给发送机188(206)。此外,分段器186还确定MDE的传输时间要求204,并将表示传输时间要求204的数据发送给发送机188。

[0101] 发送机188表示根据本公开内容的技术的基于文件的协议发送单元的例子。在该例子中,发送机188根据ROUTE来发送接收到的区段的数据。在其它例子中,发送机188可以根据FLUTE或其它此类基于文件的协议来发送接收到的区段的数据。根据本公开内容的技术,发送机188将MDE数据(封装在基于网络的数据单元内,例如IP、UDP、ROUTE、ALP等等)210和传输时间要求208二者发送给调度器190和成帧器192。成帧器192形成网络帧,该网络帧根据由调度器190确定的调度数据将网络数据封装到激励器/放大器194,其中网络帧还包括对物理层数据传递事件(DDE)的基带描述212。激励器/放大器194例如经由物理网络层、例如经由ATSC OTA广播、以太网等等来主动地发送数据。

[0102] 广播系统(例如,高级电视系统委员会(ATSC)系统)可以支持在ROUTE或包括媒体感知字节范围的其它对象传递协议内使用利用MDE实现的媒体感知字节范围。对系统的单独MDE的标识(定义)可能是要求的行为,使得系统能够传递MDE。即,系统必须能够确定什么媒体感知字节范围表示MDE(或类似的数据对象)以便对MDE打包并以适当的时间标签将其发送给调度器。

[0103] 对MDE的示例性定义是对于接收媒体编解码器而言有意义的媒体的媒体感知字节范围。由于编解码器的帧结构,MDE可以是或者可以不是术语样本的ISO BMMF使用中的多个样本。包含多个音频样本的高效率高级音频编码(HE-AAC)(ISO/IEC JTC1/SC29/WG11/N7016(2005年1月11日))中的音频帧本身是ISO BMFF上下文中的单个样本。HEVC上下文中的ISO BMFF样本可以是视频数据的单个帧。MDE可以包括单个或多个ISO BMFF样本。由于在连续视频帧之间的相互依赖性而存在这种可能性。存在最小可能MDE,但是为了方便或其他目的,这些原子MDE(即,最小可能MDE)可以被聚合成较大的MDE。MDE通常不重叠。

[0104] 图5中系统180的可以首先感知MDE的元件是媒体编码器184、分段器186以及发送机188。图5中的功能实体是概念性的。级联的这种逻辑构造是为了方便考虑系统而不是必需的。各功能可以在功能上集成到单个实体中,或者可能以不同方式在另外的或替代的单元之间进行细分。MDE形成的逻辑位置在媒体编码器184、分段器186或发送机188的输入部分的范围内。

[0105] MDE的范围和定义是根据被处理的媒体类型来定义的。例如,音频可以被打包成单

独的区段或者被复用在区段内,即,ISO BMFF文件/对象。例如,MPEG DASH(国际标准ISO/IEC 23009-1第二版2014-05-01信息技术—HTTP的动态自适应流式传输(DASH)第1部分:媒体呈现描述和区段格式)定义了经复用和未经复用的区段二者。这仅是配置的一个方面。如上面图5中所示出的,可以存在向各种功能提供配置的系统配置功能。

- [0106] 一些潜在地经配置方面的例子包括:
 - [0107] ●对于DASH
 - [0108] ○区段持续时间
 - [0109] ○区段组成即,特定的媒体对象中所包含的媒体类型
 - [0110] ○区段类型节奏例如,特定的媒体段是否包含流访问点(SAP)
 - [0111] ●对于视频编码器
 - [0112] ○GoP持续时间(特定的编解码器命名针对此类结构不同)
 - [0113] ○GoP类型例如,开启还是关闭
 - [0114] ○帧类型节奏例如,I、B、B、P、B、B、P...
 - [0115] ○帧速率、分辨率、渐进式相对于交织等等。
 - [0116] ○经由模式匹配或根据允许的媒体结构的MDE边界,例如
 - [0117] ■音频帧是原子MDE和ISO BMFF样本
 - [0118] ■单独的I(IDR)帧是原子MDE
 - [0119] ■视频帧的群组是原子MDE,如图3中所描绘的
 - [0120] ○这种确定可以在编码器中做出或已知或者由分段器来确定
 - [0121] ●对于音频编码器
 - [0122] ○样本速率
 - [0123] ○编码工具的选择
 - [0124] ○根据特定的编解码器的音频帧构造的细节
- [0125] 就这些方面是静态配置而言,可以将相关联的MDE定义可以提供作为静态定义,即,分段器和或发送机已知编码器的配置,因为它是静态配置的设置集合。假设视频运行视频帧类型的单个定义的节奏,则MDE边界可以由配置来定义。关于特定视频节奏及其原子MDE的例子,参见下面的图7。应当注意,该图描绘了视频帧的显式顺序。
- [0126] 图6是示出了MDE的示例性高级组成的概念图。
- [0127] 图7是示出了示例性帧类型节奏和相应定义的MDE的概念图。具体而言,图7示出了视频帧250-296。各个视频帧被标记为I、P或B,以表示这些帧是帧内预测(I),单向帧间预测(P)还是双向帧间预测(B)。
- [0128] 图7中所描绘的视频帧按照显式顺序来示出。假如按照传递顺序来描绘图,则帧群组将在来自视频编码器的源流中并且在ISO BMFF容器内显示为连续的字节范围,即,DASH的媒体区段。
- [0129] 对于这种经定义的节奏的MDE边界的标识可以经由表来完成,其中该表例如定义GoP的前N个视频帧为MDE 1,接下去的M个帧为MDE 2等等,或者在编解码器语法中识别的特定的帧类型节奏,例如,模式匹配。如果帧类型节奏在每个区段基础上是静态的,则使用表的方法是简单和方便的。对于可能每个MDE以及每个编解码器具有固定数量的音频编解码器帧的音频媒体类型而言,这潜在地是方便的方法。(这也是平凡规则的例子。)可能包含对

模式或规则的显式描述的这些表可以被表达为XML。图7示出了一些原子MDE可以被聚合以形成较大的MDE。这可以经由规则或模式匹配或者其组合来确定。这些方法可以实现在媒体编码器内并作为显式描述向前传递给分段器和发送机,例如ROUTE或MMT。

[0130] 如上面简要提到的,可以根据规则来形成配置表。此类规则的逻辑可以允许MDE(媒体感知字节范围)标识功能的操作是自适应的。例如如果区段(ISO BMFF对象)是经复用的(即,包含多个媒体类型),则这是方便的。分段器的MDE标识功能可以解析例如隐藏式字幕、音频和视频。此外,例如视频帧类型的节奏可以是自适应的。

[0131] 从MDE的角度而言,视频的自适应方面的例子包括:

[0132] ●视频帧类型节奏以启用区段序列中开启和关闭的GoP

[0133] ●每个区段或时段基础上的视频帧计数和类型节奏

[0134] ○例如,24fps广告和30fps内容交替

[0135] ○帧速率的这种改变会导致视频帧类型节奏改变,并因此导致MDE节奏改变。

[0136] 描述了用于支持区段内的MDE标识的技术,即,与支持字节范围感知协议(例如ROUTE或MMT)的相关联的媒体编解码器上下文相关的媒体感知字节范围。

[0137] 用于支持这种MDE定义(即,媒体编解码器感知字节范围)的技术的例子可以包括:

[0138] ●描述特定序列的表格方法,例如:

[0139] ○视频帧类型

[0140] ○视频帧类型的序列

[0141] ○例如视频或音频帧的数量

[0142] ○所描述的结构的组合

[0143] ○例如视频帧类型的特定序列的模式匹配

[0144] ○这些方法实现在媒体编码器、分段器或发送机中

[0145] ●基于规则的基于MDE的检测

[0146] ○例如,HEVC编解码器的特定配置可以允许帧类型序列的这些逻辑构造:

[0147] ■帧序列(MDE)的开始满足该第一条件集合。

[0148] ■帧序列(MDE)的结束满足第二条件集合。

[0149] ■帧类型的序列满足第三条件集合。

[0150] ●所描述的模式和规则方法可以组合,例如:

[0151] ○发现此模式

[0152] ○则应当应用此规则

[0153] ●媒体编码器可以具有带有特定MDE构造的经定义的操作模式库。

[0154] ○媒体编码器实时地传递所述操作模式,其中经编码媒体传递给分段器

[0155] ●有关基于XML或基于文件规则的逻辑和基于模式的逻辑的例子,请参见附录A。

[0156] 每个MDE可以定义有最早和/或最迟传输时间。这是为了调度器的益处,调度器的任务是将媒体分配给传递时隙,例如,在物理层上的ATSC 3.0中的特定FEC帧。这些最早和最迟属性可以由不同的方面来约束,例如:

[0157] ●最早传输时间:这在字面上是可以辐射给定MDE(或区段)的第一比特的时刻。该最早辐射的比特包括所有相关的封装协议。

[0158] ●最迟传输时间:这在字面上是必须辐射MDE(或区段)的最后比特的最后时刻。该

最后比特包括所有相关的封装协议。

[0159] 最早传输时间可以由系统配置来约束。例如,物理层帧节奏可以具有在ATSC物理层时间(48b TAI秒,10b毫秒或可能的其它分数秒)的整个第二边界上开始的帧。该物理层帧节奏可以或者可以不与媒体区段持续时间直接相关。例如,在作为替换广告的广告之前的整秒的结束之前,可能必须完整地传递节目的给定时段。例如,拖尾替换广告直到至少下一整秒开始之前不应开始传递。对于远离时段边界的时间,最早传输时间可以是例如在先前的整秒中。

[0160] 最迟传输时间可以与包含在标记的MDE内的最后解码的媒体的解码时间相关。应用于媒体区段级传递的标记可以由额定DASH/区段可用时间线来确定。最迟传输时间可以与满足区段的传递期限一致。区段的最后MDE(媒体感知字节范围)的最迟传输时间要求可以与完整的区段相同。

[0161] 图8是表示系统时间的简单视图的概念图。即,图8描绘了系统时间的顶级视图。对于区段级回放或基于字节范围的回放均是这种情况。对于MDE回放而言的一个不同之处在于静态接收机延迟可以比区段回放显著更短。站挂钟时间可以经由物理层和支持协议传递给接收机。在同步之后的任何时刻的接收机挂钟是当前站挂钟时间减去当时从主导发射机到接收机的当前信号飞行时间。

[0162] 图9是表示系统300的框图,其表示被配置有时间关系的发送基础设施。即,图9提供了时间延迟可以如何影响系统300的发送过程的概念表示。在该例子中,系统300包括GOP持续时间媒体缓冲器306、媒体编码器308、分段器310、速率控制304、系统管理单元302、发送机312、调度器314、以及激励器/放大器316。

[0163] 在该例子中,GoP持续时间媒体缓冲器306存储(即,缓冲)要被编码的原始媒体数据。媒体编码器308从GoP持续时间媒体缓冲器306获取要被编码的媒体数据,对媒体数据进行编码,以NAL单元形式来封装经编码的媒体数据,并将NAL单元320传递给分段器310。应当理解,经编码的媒体数据还包括MDE,例如,如图7中所示出的。此外,如上面说明的,每个MDE可以具有传输时间要求,例如最早传输时间和/或最迟传输时间要求。

[0164] 分段器310从媒体编码器308接收NAL单元320,并将一个或多个MDE(即,NAL单元中的一个或多个)封装在对应的区段中。再次,区段可以与关联于对应URL的可独立传递的文件相对应,其中URL均可以唯一地标识相应的区段。MDE可以与区段的对应字节范围相对应。另外,分段器310可以确定每个MDE的传输时间要求,并将传输时间要求322和MDE(以区段的字节范围的形式)324二者传递给发送机312。

[0165] 发送机312可以根据基于文件的传递协议(例如,ROUTE)来操作。因此,发送机312可以用对应的基于网络的传递单元328(例如,分组或帧)将区段的对应部分(例如,MDE)封装到调度器314以供传输。另外,发送机312将传输时间要求326提供给调度器314,使得调度器314能够根据传输时间要求将基于网络的传递单元传递给客户端设备(图9中未示出)。更具体而言,调度器314形成对物理层数据传递事件(DDE)的基带描述330,并将对物理层DDE的基带描述330传递给激励器/放大器316。随后,激励器/放大器316在物理层处将数据例如发送给客户端设备。

[0166] 图9还表示由系统300在向客户端设备发送媒体数据时所引入的各种延迟。例如,媒体数据被缓冲至媒体数据被调度用于传输之间的时间被称为分析持续时间延迟332。分

析持续时间延迟特别包括针对特定样本的单次延迟334,其是媒体编码器308对样本进行编码并将样本封装在NAL单元中以及将NAL单元提供给分段器310之间的时间,如图9中所示出的。此外,具有用于传输的经调度数据与数据实际被发送之间的时间被称为发射机和STL延迟336。

[0167] 定义媒体数据的时间是考虑系统300的发送基础设施的各种功能块在辐射时间之前多久必须进行操作以及在媒体在接收之后多久将被切换到媒体解码器的事情。该讨论因此被划分成辐射之前和接收之后。为什么实际辐射时间是必要的?为了实现针对诸如个性化广告插入的应用在源流之间的潜在同步切换。限定媒体的传输的实际时间可以受约束。传递给接收机的时间被同步到物理层的自引导特征的辐射时间。必须理解和控制特定的时段存在于物理层上的时间跨度。MDE或媒体区段被标记有允许辐射的时间范围,因此在实际发射之前的所有过程在时间上通常偏移与特定的块或功能相关联的恒定延迟。图9描绘了发射基础设施内的时域的合理组织。该划分不是要求,仅是从概念的角度而言方便。以下是一些可能的时域及其合理的持续时间的列表。

[0168] 发射机和演播室发射机链路(STL)延迟336:跨越这些块定义的延迟足够长,使得刚开始传递给STL的给定物理层帧可以在该时段过去时开始由发射机输出。该延迟名义上包括至少物理层帧持续时间、交织器持续时间和SFN分配网络中的最长发送延迟。可以考虑使该延迟更短,例如,在子帧基础上。支持子帧潜在地可能导致演播室发射机链路上的演播室上的增加的峰值速率,即,演播室发射机链路必须在至少子帧持续时间内保持物理层的最大允许吞吐量。对全帧实现方式的约束是在全帧基础上测量的峰值允许速率分配,其可以低于每个子帧,但是如果所有活动PLP具有相同的容量,则可以是相同的。

[0169] 分析持续时间延迟332:调度器314的任务之一是以满足其定义的时间时限、并且满足每个当前物理层PLP定义的容量约束的形式向物理层分配媒体。为了实现这一点,调度器314应当分析至少物理层帧的媒体的最长持续时间。如果调度器314每个经过的物理层帧时间没有成功地调度至少一个物理层帧时间的经编码媒体,则调度器314可能最终失败。这是累积任务。具有小于其媒体时间的持续时间的物理层帧可以与表示更多媒体时间的物理层帧偏移。

[0170] 物理层帧的持续时间可以短于最长媒体区段的持续时间或GoP持续时间。对于这种情况,限定性界限与至少最长媒体区段持续时间或GoP持续时间相关,GoP由多个媒体区段组成。调度器314可以确保最长持续时间媒体区段/GoP被成功调度。调度器314还可以每个媒体秒生成至少一秒的经编码媒体区段/GoP。如此处描述的,分析持续时间延迟332不是恒定的,因为调度器314在实际辐射时间上具有最迟辐射时间—最早辐射时间灵活性。

[0171] 单次延迟334:假如输入媒体直接在媒体编码器308和分段器310上流式传输,则将存在对于给定配置是静态的、从GoP的第一帧的输入到第一经解码帧的解码时间的解码延迟,如在得到的第一区段中所描述的。这是概念性的构造。

[0172] 可能的情况是,媒体编码器308(其可以表示多个编码器)很可能每个媒体区段时间产生一个以上媒体时间的经编码内容。实际约束很可能由编码方案使用的“轮次”的数量来限定。在该上下文中,“轮次”是媒体编码的一个实例。如果媒体编码器308能够在一个编码轮次上满足所有目标,则要求将是每个经过的媒体区段时间至少一个媒体区段持续时间的媒体。在第一轮次取得期望的数据大小、使得调度器可以尝试将轮次限制为二次编码是

相当具有挑战性的。在二次编码中,第一轮次可以提供对要编码的媒体数据的复杂性的校准,并且第二轮次可以用于调整在第一轮次期间作出的决定以满足物理层的容量。如果初始的编码轮次得到大体上与最终期望/要求的数据大小不同的数据大小,则可以执行第三轮次或对数据速率的其它调整。应当注意,调度器必须对要在包括所有封装的物理层中传递的实际数据大小进行操作。

[0173] 假设媒体实时地流式传输到媒体编码器308,并且这是多次编码方案,则潜在地需要全GoP延迟以在媒体编码器之前捕获实时的流式传输媒体。在接收机处向经编码的媒体提供的解码时间必须反映基础设施和接收机的完整延迟。预期这些是基于媒体区段回放的。可以根据相对于区段级回放原本会发生的MDE RAP启动点来计算MDE时间。

[0174] 图10是示出了具有交错RAP位置的每个GOP的多个区段的概念图。如图10中所示出的交错RAP区段时间(即,每个GoP的多个分析持续时间)可以减少在调度媒体之前所需的最小捕获时间,并且因此降低总体延时。这具有有限的统计复用益处,因为如所示出的,给定物理层帧的数据速率由单个RAP区段数据大小控制。媒体的时间标签与媒体一起捕获。

[0175] 如上面所描述的,发射延迟、或至辐射的延迟可以被表达为:

[0176] 分析延迟持续时间-针对特定样本的单次延迟+发射机和STL延迟(帧时间+交织器深度+SFN分配)

[0177] 图11是示出了用于媒体回放的解码时间的最小延迟的概念图。即,图11描绘了影响媒体区段和MDE回放的媒体解码时间的最小延迟。媒体区段和MDE回放之间的一个区别在于,在区段级延迟的MDE回放的开始时间必须考虑最长堆栈延迟加上建议的呈现延迟,而MDE级回放是单独设备静态的实际延迟加上某个延迟以确保ROUTE缓冲器足够满。总体MDE延迟值可以小于区段时间。

[0178] 总体的端到端区段解码延迟可以被表达为:

[0179] 分析延迟持续时间-针对特定样本的单次延迟+发射机和STL延迟(帧时间+交织器深度+SFN分配)+飞行时间+至可用开始时间的延迟+MPD@建议的呈现延迟+针对特定样本的区段内延迟总体的端到端MDE解码延迟可以被表达为:

[0180] 分析延迟持续时间-针对特定样本的单次延迟+发射机和STL延迟(帧时间+交织器深度+SFN分配)+飞行时间+ROUTE或类似的MMT接收机处的延迟抵达+流式传输字节范围推迟(hold-off)时间+

[0181] 针对特定样本的区段内延迟

[0182] 上面已讨论了用于调度的时间元数据方法的某些细节。此处讨论用于调度的时间元数据方法的进一步细节。用于MDE的时间元数据的构造可以基于由于区段开始而由所要求的最迟传输时间定义的锚点。MDE所要求的最迟时间按照在区段开始之后每个对应MDE的第一解码时间的顺序可以是线性的。最迟时间可以由当前要由在MDE中解码的样本相对于区段的第一经解码样本的解码时间来约束,该第一经解码样本也是区段的第一经解码MDE。这种方法在解码时间上是线性的,这仅是用于定义最迟传输时间元数据的一种方法。

[0183] 存在要考虑的另外方面,例如有可能前端加载的媒体组件,例如音频和字幕。由于这些媒体组件相对于视频是小的,但在没有这些媒体组件的情况下完整呈现是不可能实现的,因此有可能最简单的是在视频RAP之前但在可能需要的元数据对象(例如,MPD和IS)之后发送这些媒体组件。即,通常,可以用于确保所有需要的小的数据对象首先传递的方法。

[0184] 图12是示出了可选的物理层同步和相关元数据的概念图。物理层帧与T-RAP的关系是关于最早传输的相关方面。例如,考虑图12的例子,示出了最早时间在额定物理帧开始之前移动的自由度。这会影响频道改变时间,除非在所有层中存在开始元数据的完整序列。

[0185] 物理层的设计也许能或也许不能如上所述在GoP边界上精确地渲染同步。同步位置的定义可以由物理层的数学方案(numerology)约束到其它位置。一种可行的方法是使用靠近GoP持续时间的位置。靠近可以是例如下一可用时间,即,在承载变换的先前业务的额定边界之前或之后。最小变换的时间刻度可以是1毫秒的刻度,而最大值可以是10毫秒的范围。

[0186] 调度功能可以名义上考虑固定的时间块,可能与如前所述的媒体GoP持续时间或区段持续时间相关。对于更高效的GoP持续时间方案,可能可以调度所有媒体,并且可以使一些显著容量在(N)中可用以供下一个分析间隔(N+1)潜在使用。在这种状况下,调度器可以尝试用来自下一分析间隔(N+1)的媒体来填充(N)中的该容量。这可以通过时间元数据来实现,从而允许经由最早参数在较早的间隔中传输。一旦知道了进行中的间隔的已填充部分,物理层组件就可以插入可选的帧同步元素,其被提供用于允许比利用准固定周期性成帧可能的时间更早地获取媒体服务,例如ATSC 3.0的引导程序和前导码。

[0187] 图13描绘了包括与潜在系统事件相对应的步骤序列以实现对多层媒体数据的服务获取的方法的一个例子。即,图13是示出了获取事件的示例性序列的流程图。可以由本文所示出的各种设备和系统(例如,图1的内容准备设备20和源设备60、图5的系统180、或者图9的系统300)来执行图13的方法。出于举例和说明的目的,关于图5的系统180来说明图13的方法。

[0188] 系统180传递使得能够开始线性或基于应用的线性服务(350)的数据序列。图13描绘了接收机处的事件的典型序列以及系统(180)传递相关的元数据、其它对象、以及与服务接收的典型开始相关的媒体文件的顺序。服务的这种开始通常由频道号输入、频道向上/向下键输入、或者服务的引导选择来发起(350)。物理层的获取由接收引导程序或系统同步模式(352)来开始。与物理层相关的系统元数据和另外的系统元数据的获取经由前导码或类似的物理层描述元数据接收(354)来完成。在标识了期望的PLP或其它物理层分组流在波形中的位置时,就接收到期望的分组流,其包含ALP或其它类似封装协议,包括分组流ID(PLP ID)到所包含的IP或其它类似分组流ID或地址的映射(交叉引用)(356)。前导码或类似的物(354)包含标识系统元数据(LLS)的存在的标志,其标识诸如系统时间、服务组成(SLT)、可能的告警信息(AEAT)(358)等服务细节。在LLS(358)中可用的SLT包含SLS中的每个服务细节的位置(360)。携带服务的PLP也可以包含可以包括运行时应用(362)的应用,其可以允许电视服务在基于浏览器的接口中操作。如果存在,则应用可以在服务开始之前或之后启动。一旦由相关的数据接收驱动的这些步骤完成,就可以根据对IS(364)和相关的媒体区段(366)的接收和解码来发起媒体回放。随后可以开始线性或基于应用的线性服务(368)。

[0189] MDE的端到端的解码延迟是比区段级解码早的静态时间。可以经由MPD中现有的元数据或者对MPD的重写(其改变元数据的值以实现MDE的MPD回放)来实现该静态偏移。可以根据经由MDE相对于从区段级将播放第一RAP的时间的第一RAP播放来确定偏移的值。

[0190] 与MDE的调度相关的特征可以包括:

[0191] ●使获取时间最小化的媒体类型和元数据的时间组织。

[0192] ○提供具有较早最迟时间的较小的媒体数据类型,以确保完整传递并且作为相对于尝试经由许多较小的经调度对象与其它媒体交织的简化。该方法得到在调度上的更大灵活性,即,具有更多时间维度的一些较大对象而不是具有严格要求的许多较小对象。增加的调度灵活性通常得到在总体效率方面的优越成果。

[0193] ○假如针对音频和必要的元数据期望更稳健的性能,则这些可以在更稳健的物理层管道中一起调度。

[0194] ○在公共传递管道中期望媒体类型的特定交织的情况下,使用经复用的区段可以更简单地实现。按照传递/解码顺序来提供经复用的媒体对象内部的样本。

[0195] ●用于通过可选的物理层同步位置的手段来保留快速获取的方法。

[0196] ○在额定位置之前识别T-RAP传递,即,拖尾额定物理层同步或相关的额定分析间隔。

[0197] ○基于已满足所有传递请求的当前调度周期进行适配,即,如果已满足在额定帧物理层周期中最新的所有数据请求。

[0198] ●利用发射机必须在已知的整秒和整毫秒上开始的规则,适应物理层帧到与媒体GoP对准的分析间隔的非精确持续时间的手段。(同步模式按照定义将总是在偶数秒和毫秒上运行。)符号模式将始终保持从ATSC时间的秒和毫秒开始的整数个符号经过时间。(符号速率可以在配置信息中指定的或从配置信息中确定的预定义值之间改变。)

[0199] ○将帧同步分配到时间拖尾额定位置相对于分析间隔的最接近的可用符号位置。

[0200] ○将帧同步分配到符号主导的额定分析间隔位置中最接近的可用位置。

[0201] ○相对于定义的分析间隔的任何其它基于规则的基于物理层同步的方案。

[0202] ○链接到媒体GoP或类似结构的定义的分析间隔。

[0203] ●基于分析间隔、基于区段持续时间而不是GoP的可能的低延时方法。

[0204] ○基于RAP区段的主速率管理。

[0205] ○调整拖尾非RAP区段以补偿失败的PSNR或其它质量度量。

[0206] 概括地说,本公开内容描述了可以单独或以任意组合用于在文件传输协议(例如ROUTE或MMT)内标识MDE的位置、媒体感知字节范围的各种技术。这些技术可以基于对有效模式的显式描述和匹配或者通过规则来确定有效模式。可以组合这些技术。对于组合的顺序没有限制。类似地,本公开内容还描述了时间方面作为用于组织物理层同步和其它传递方面的技术的讨论的背景。这些技术可以实现在具有相对于与媒体相关的分析间隔的可变或精确同步位置的系统中的快速获取。

[0207] 图14是根据本公开内容的技术示出了另一个示例性方法的流程图。图14的方法被说明为由基于文件的协议发送单元(例如,图5的发送机188或者图9的发送机312)来执行。出于说明的目的,关于图9的发送机312来说明图14的方法,但应当要理解,其它单元可以被配置为执行该方法或类似方法。如上面所讨论的,发送机312表示基于文件的协议发送单元的例子,例如,根据ROUTE协议或FLUTE协议或者其它基于文件的传输协议来发送数据的单元。

[0208] 初始地在图14的例子中,发送机312接收包括一个或多个MDE的一个或多个区段(380)。例如,MDE可以与区段的可传递字节范围相对应。发送机312随后确定MDE位置(382)。通常,发送机312不接收用信号通知MDE的位置的显式信息,而是替代地使用本公开内容的

技术中的一种或多种技术来对区段的数据进行处理以确定MDE的位置。例如,发送机312可以执行模式匹配技术或者定义可以如何标识MDE的规则,如上面所讨论的。发送机312因此可以使用模式匹配技术、规则或其它此类技术来标识区段内的MDE。

[0209] 发送机312还可以确定MDE的传输时间要求(384)。例如,发送机312可以分析与区段相对应的清单文件的数据,并根据清单文件来确定传输时间要求。清单文件可以包括例如根据DASH的MPD。传输时间要求可以表示针对一个或多个MDE的最早传输时间或最迟传输时间中的一者或二者,如上面所讨论的。

[0210] 最终,发送机312可以向物理层发送单元(例如,调度器314和激励器/放大器316)提供MDE和表示MDE的传输时间要求的数据(386)。以此方式,调度器314可以根据传输时间要求来调度MDE的传输,使得激励器/放大器316不早于最早传输时间和/或不迟于最迟传输时间来发送MDE的数据。

[0211] 以此方式,图14的方法表示包括由源设备的基于文件的协议发送单元进行以下操作的方法的例子:从源设备的形成区段的分段器接收包括媒体数据的区段的数据流,每个区段包括与唯一的统一资源定位符(URL)或统一资源标识符(URI)相关联的对应可单独获取的文件;确定媒体传递事件(MDE)在媒体数据流中的位置,其中MDE包括针对区段中的一个区段的至少一部分的数据;确定MDE的一个或多个传输时间要求,该一个或多个传输时间要求表示要将MDE发送给客户端设备的时间;以及根据源设备的物理层发送单元的可用传递时隙将MDE和表示传输时间要求的数据供给该物理层发送单元。

[0212] 在一个或多个例子中,可以在硬件、软件、固件、或者其任意组合中实现所描述的功能。如果用软件来实现,则所述功能可以作为一个或多个指令或代码存储在计算机可读介质上或者通过其传输,并且由基于硬件的处理单元来执行。计算机可读介质可以包括计算机可读存储介质,其对应于有形介质(例如数据存储介质)或者通信介质,其中通信介质包括有助于例如根据通信协议将计算机程序从一个地方向另一个地方转移的任何介质。以此方式,计算机可读介质通常可以对应于(1)非暂时性的有形计算机可读存储介质或者(2)通信介质,例如信号或载波波形。数据存储介质可以是能够由一个或多个计算机或一个或多个处理器存取以获取用于实现本公开内容中所描述的技术的指令、代码和/或数据结构的任何可用介质。计算机程序产品可以包括计算机可读介质。

[0213] 通过举例而非限制性的方式,这种计算机可读介质可以包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储、磁盘存储或其它磁存储设备,闪存、或者可用于存储具有指令或数据结构形式的期望程序代码并且可以由计算机存取的任何其它介质。此外,任何连接可以适当地称为计算机可读介质。例如,如果使用同轴电缆、光纤光缆、双绞线、数字用户线(DSL)或诸如红外线、无线电和微波之类的无线技术从网站、服务器或者其它远程源传输指令,则同轴电缆、光纤光缆、双绞线、DSL或诸如红外线、无线电和微波之类的无线技术包括在介质的定义中。然而,应当要理解,计算机可读存储介质和数据存储介质不包括连接、载波、信号或者其它暂时性介质,而是替代地涉及非暂时性有形存储介质。如本文所使用的,磁盘(disk)和光盘(disc)包括压缩光盘(CD)、激光光盘、光盘、数字多功能光盘(DVD)、软盘和蓝光光盘,其中磁盘通常磁性地复制数据,而光盘利用激光来光学地复制数据。上面各项的组合也应当包括在计算机可读介质的范围内。

[0214] 指令可以由一个或多个处理器执行,例如一个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)、或者其它等效的集成或分立逻辑电路。因此,如本文使用的,术语“处理器”可以指代前述结构或任何适合用于实现本文所描述的技术的任何其它结构。另外,在一些方面中,本文所描述的功能可以在被配置用于编码和解码的专用硬件和/或软件模块内提供,或者被并入经组合的编解码器中。此外,所述技术可以完全实现在一个或多个电路或逻辑元件中。

[0215] 可以用各种各样的设备或装置(包括无线手持装置、集成电路 (IC) 或一组 IC(例如,芯片组))来实现本公开内容的技术。在本公开内容中描述的各种组件、模块或单元是为了对被配置为执行所公开的技术的设备的功能性方面进行强调,而不一定要求由不同的硬件单元来实现。更确切地说,如上面所描述的,各种单元可以组合到编解码器硬件单元中,或者由可互操作的硬件单元的集合(包括如上面所描述的一个或多个处理器)结合适当的软件和/或固件来提供。

[0216] 已经描述了各种例子。这些例子和其它例子落在所附权利要求的范围内。

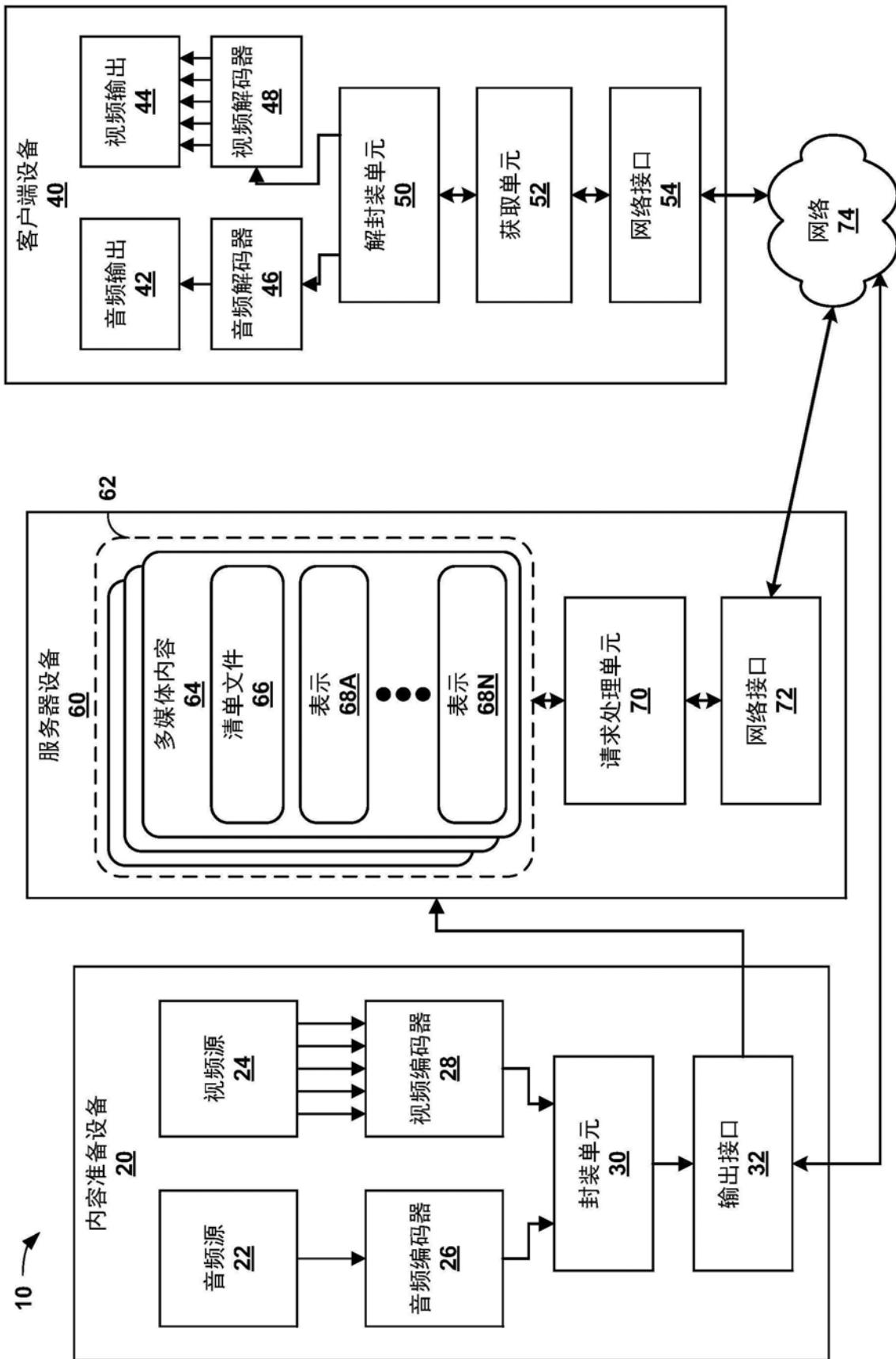


图1

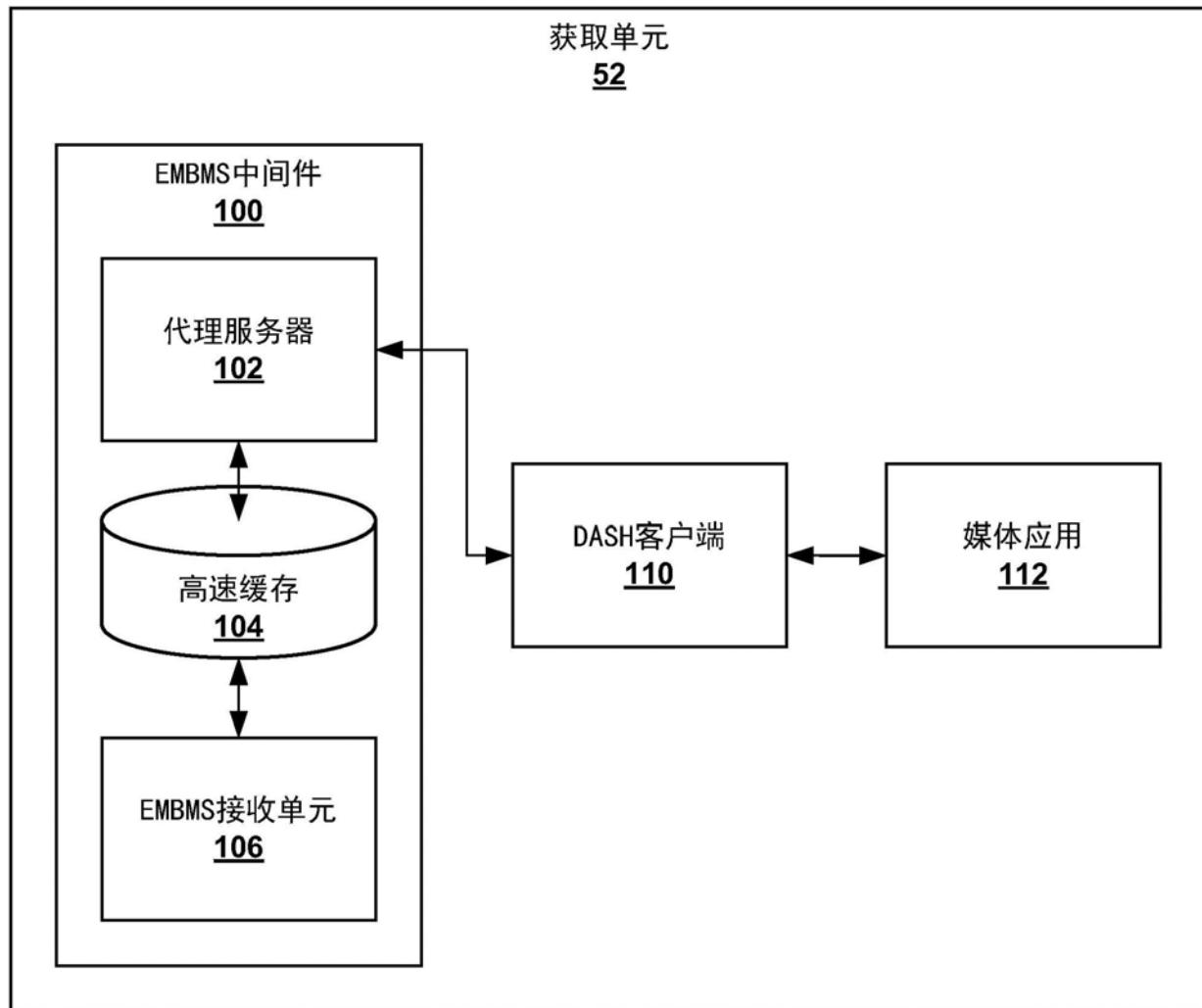


图2

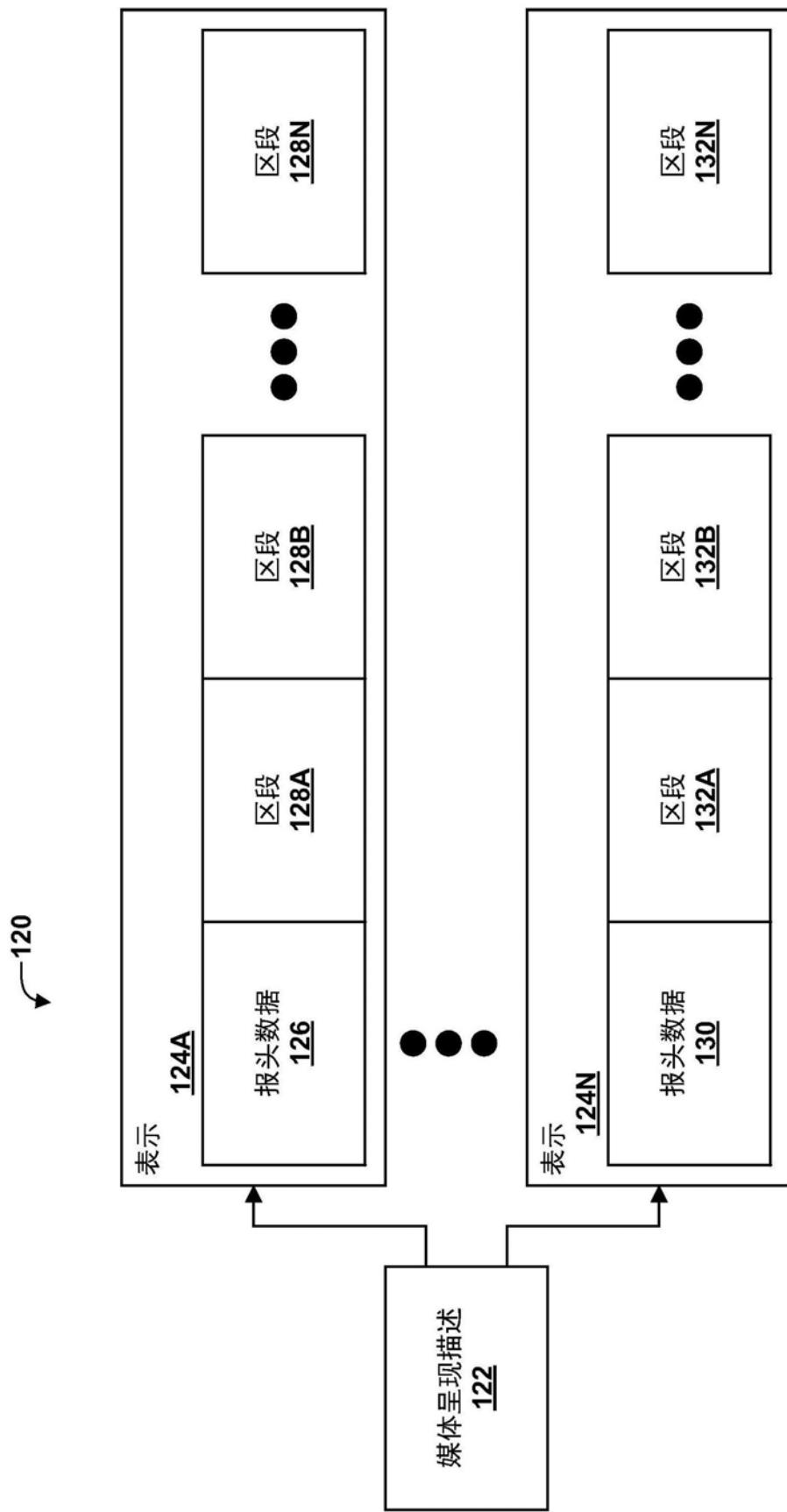


图3

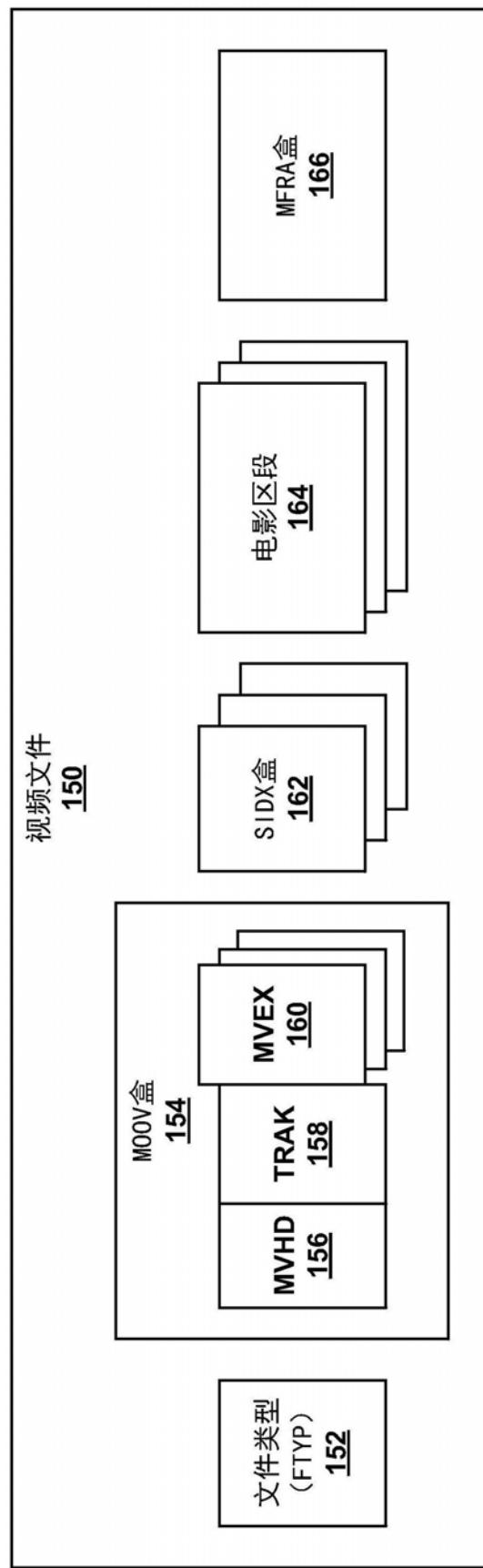


图4

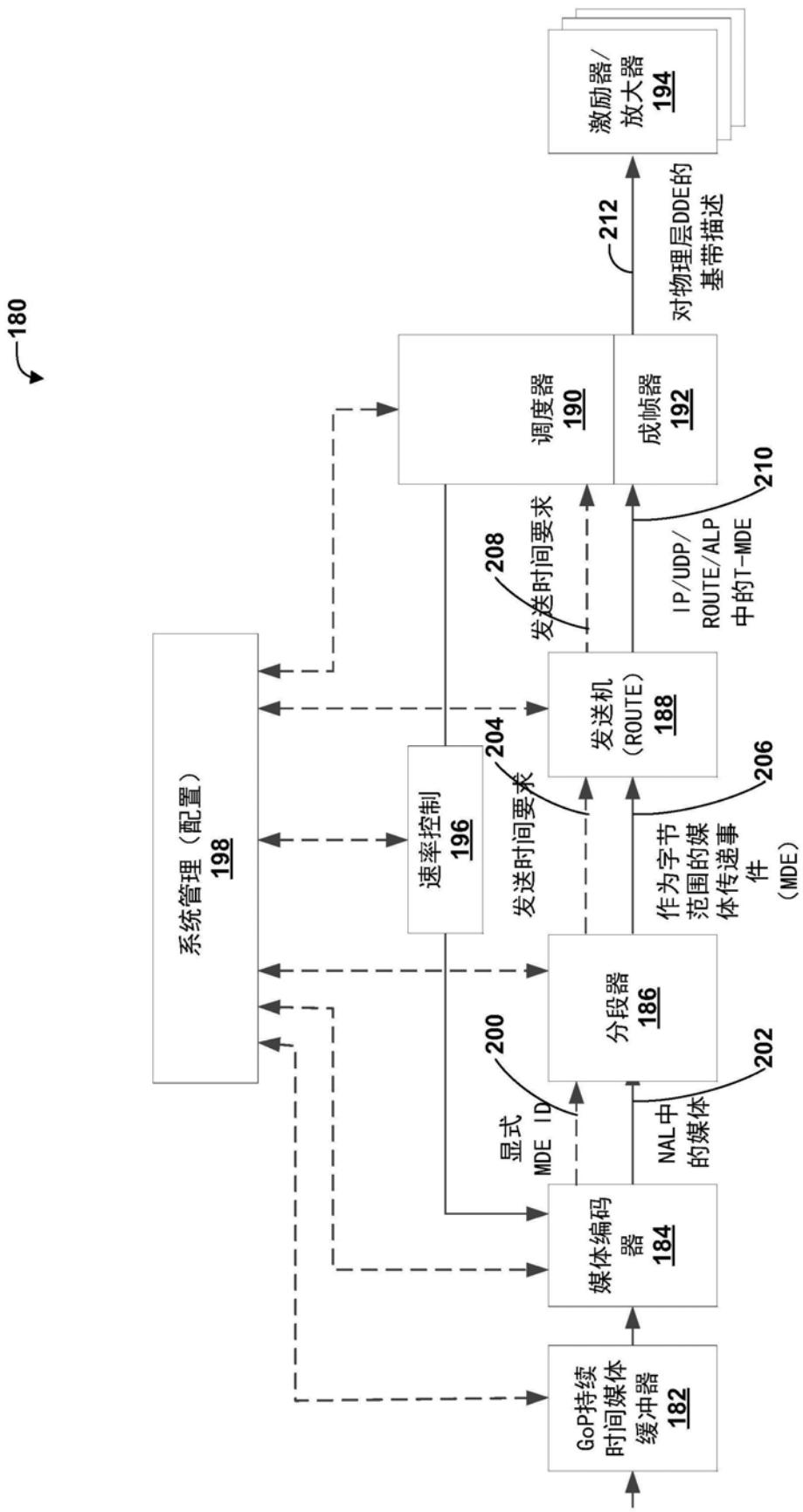


图5

音频

| 音频帧N个音频样本 | 音频帧N个音频样本 | 音频帧N个音频样本 | 音频帧N个音频样本 | 音频帧N个音频样本 | 音频帧N个音频样本 |
|---------------|---------------|---------------|---------------|---------------|---------------|
| ISO BMFF样本 | ISO BMFF样本 | ISO BMFF样本 | ISO BMFF样本 | ISO BMFF样本 | ISO BMFF样本 |
| 原子MDE | 原子MDE | 原子MDE | 原子MDE | 原子MDE | 原子MDE |
| 经聚合的MDE | | | 经聚合的MD | | |

视频

| N=1 视频帧 | J 视频帧 | I 视频帧 | M 视频帧 | K 视频帧 | L 视频帧 |
|---------------|-----------------|-----------------|-----------------|-----------------|---------------|
| ISO BMFF样本 | 多个ISO BMFF样本 | 多个ISO BMFF样本 | 多个ISO BMFF样本 | 多个ISO BMFF样本 | ISO BMFF样本 |
| MDE | MDE | MDE | MDE | MDE | MDE |
| 经聚合的MDE | | | 经聚合的MDE | | |

图6

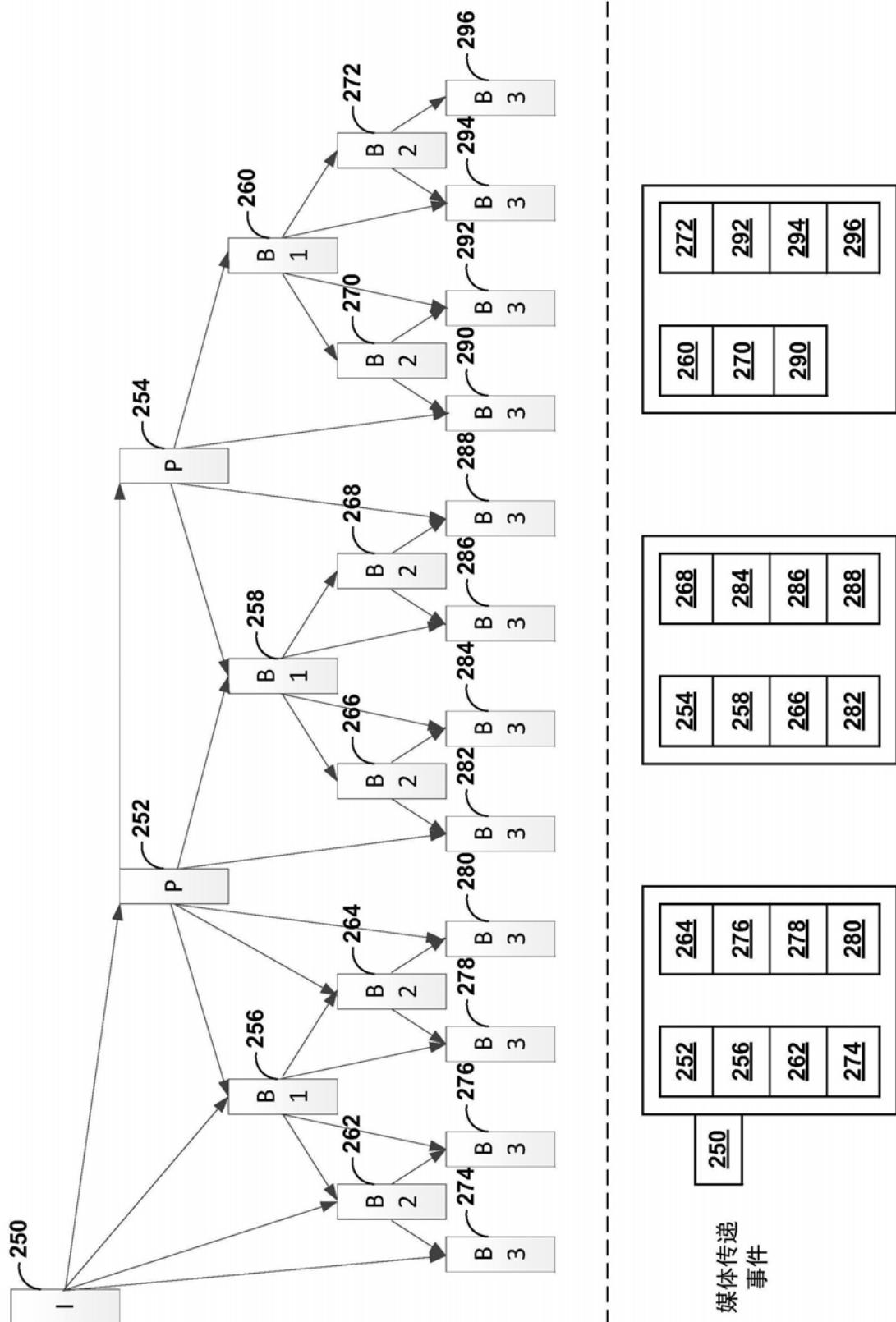


图7

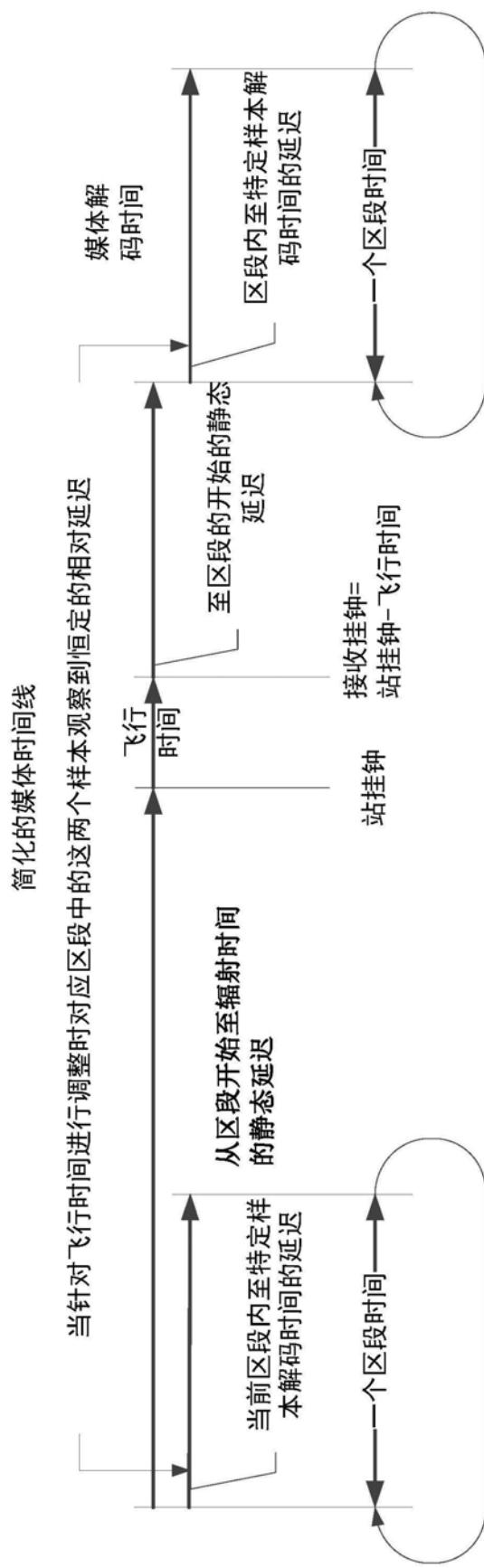


图8

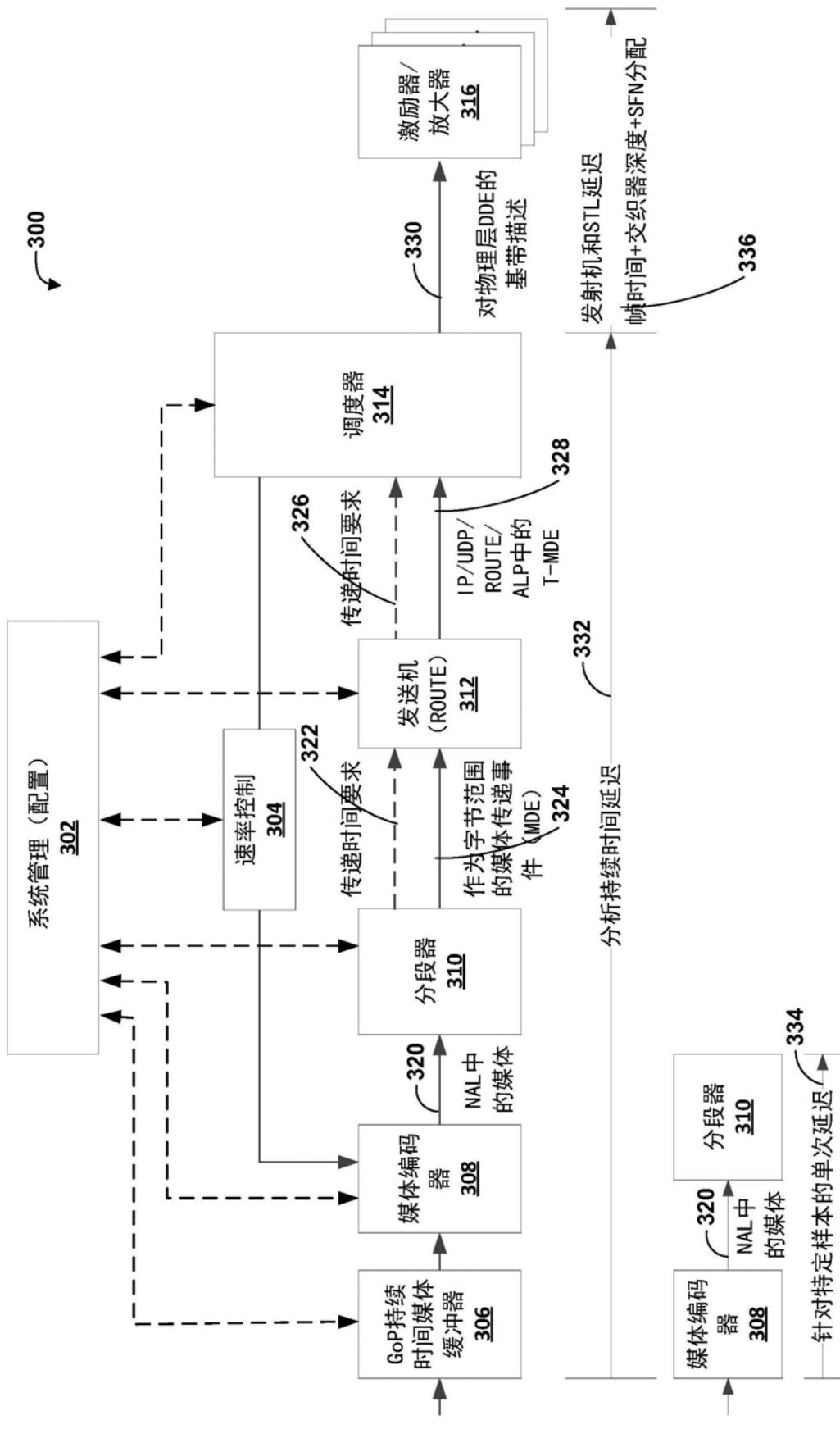




图 10

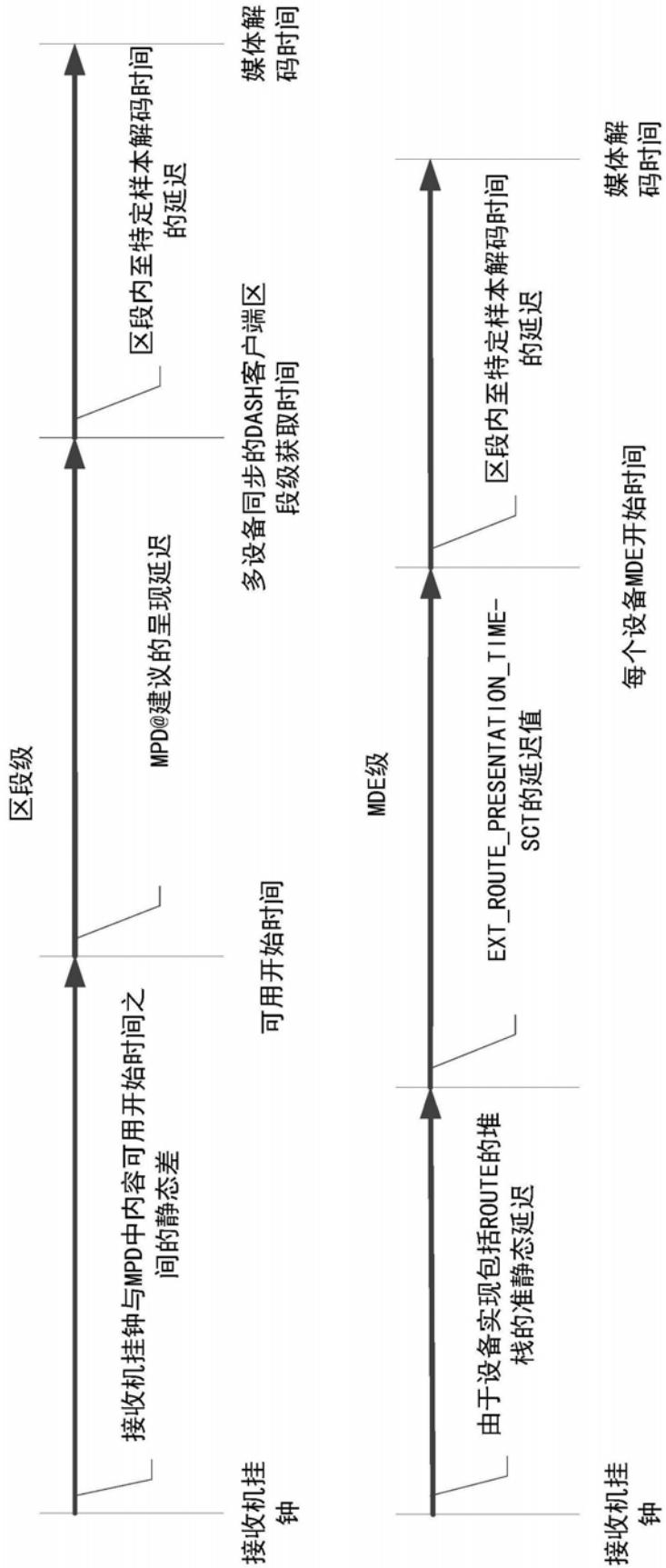


图 11

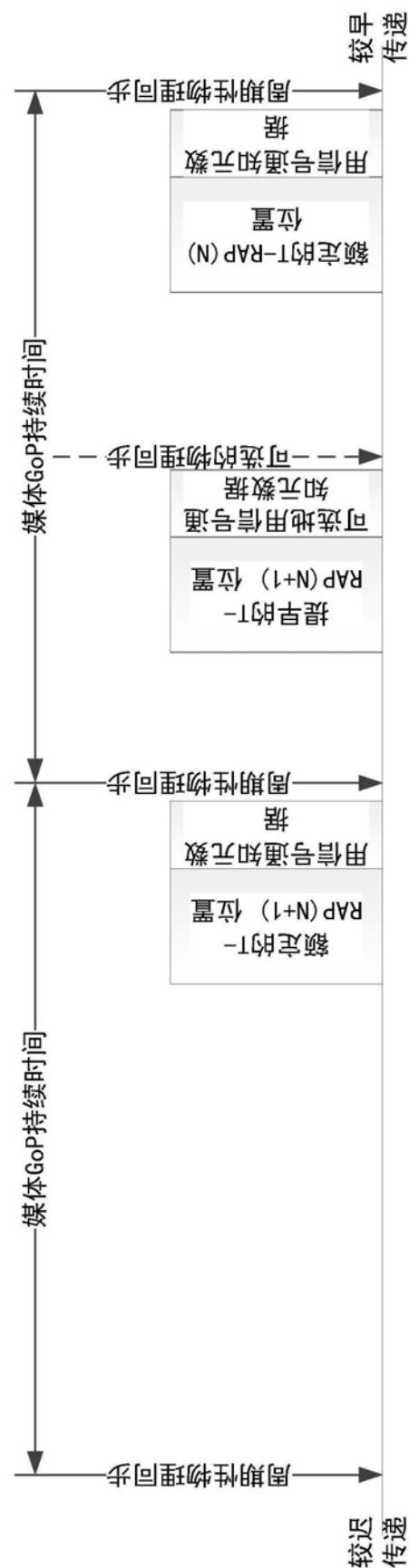


图12

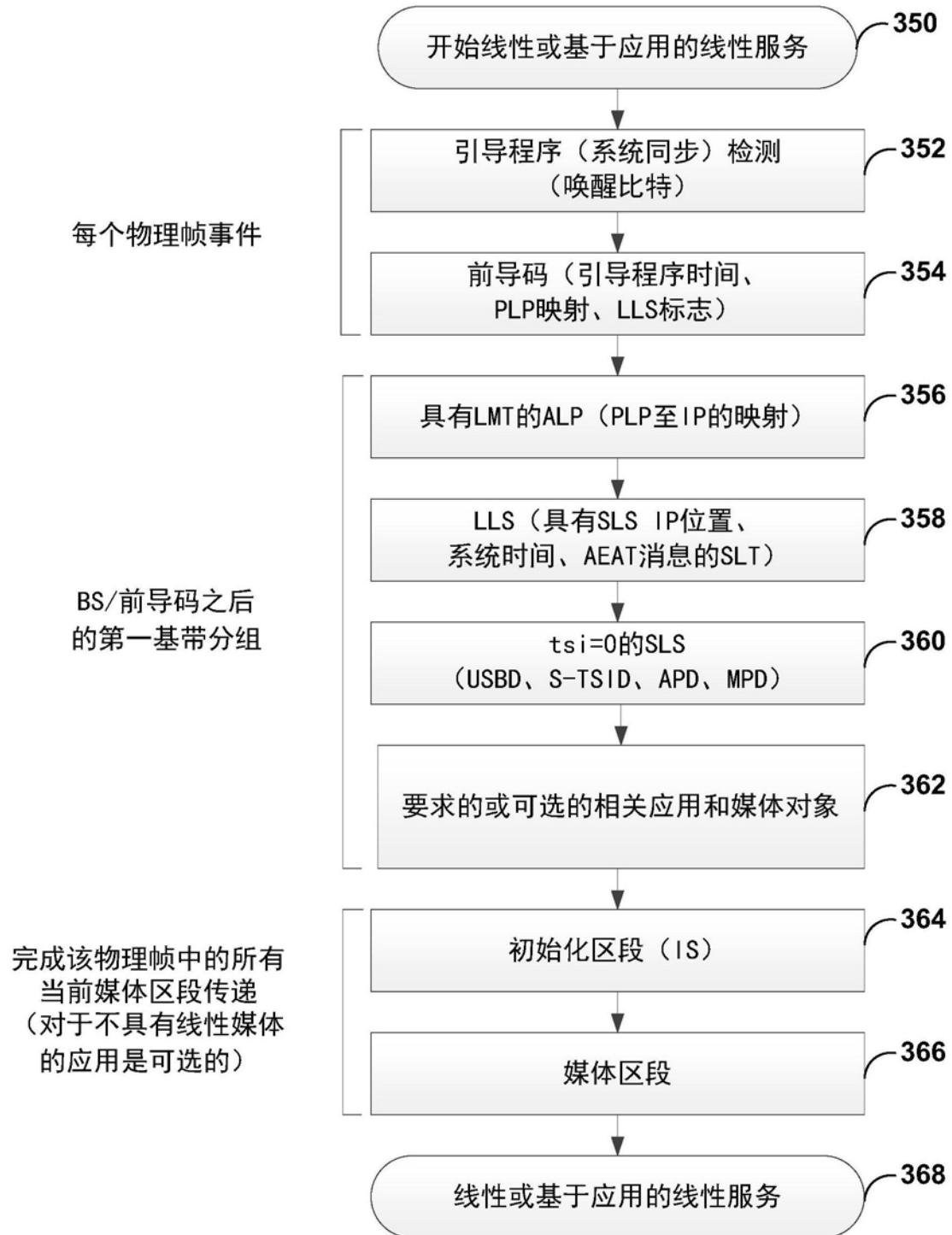


图13

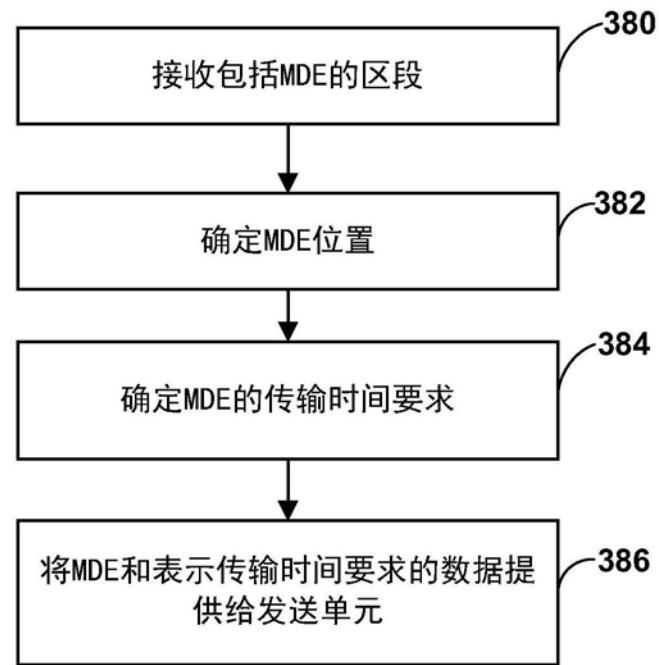


图14