



# (12) 发明专利申请

(10) 申请公布号 CN 119856499 A

(43) 申请公布日 2025. 04. 18

(21) 申请号 202380065418.8

(22) 申请日 2023.09.14

(30) 优先权数据

17/933,791 2022.09.20 US

(85) PCT国际申请进入国家阶段日

2025.03.12

(86) PCT国际申请的申请数据

PCT/US2023/032704 2023.09.14

(87) PCT国际申请的公布数据

W02024/064005 EN 2024.03.28

(71) 申请人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 L·马 T·斯托克哈默

A·里科阿尔瓦里诺 N·K·梁

I·布阿兹兹

(74) 专利代理机构 北京市柳沈律师事务所

11105

专利代理师 安之斐

(51) Int.Cl.

H04N 21/24 (2006.01)

H04L 1/00 (2006.01)

H04N 21/63 (2006.01)

H04N 21/6332 (2006.01)

H04N 21/6437 (2006.01)

H04N 21/81 (2006.01)

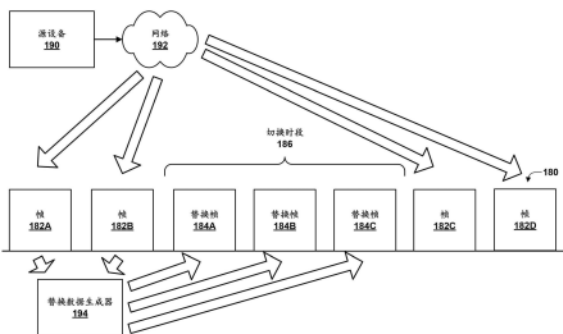
权利要求书5页 说明书27页 附图10页

## (54) 发明名称

响应于网络中断的视频内容的自动生成

## (57) 摘要

一种用于呈现媒体数据的示例设备包括：存储器，被配置为存储媒体数据；以及一个或多个处理器，实施于电路中并且被配置为：接收媒体比特流的第一媒体数据集合；响应于确定在一段时间段内将不接收媒体比特流的在第一媒体数据集合之后的第二媒体数据集合，使用第一媒体数据集合的至少一个子集来预测性地生成用于第二媒体数据集合的替换媒体数据；以及呈现第一媒体数据集合和替换媒体数据。



1. 一种呈现媒体数据的方法,所述方法包括:  
接收媒体比特流的第一媒体数据集合;  
响应于确定在一时间段内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合的至少一个子集来预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及  
呈现所述第一媒体数据集合和所述替换媒体数据。
2. 根据权利要求1所述的方法,其中接收所述第一媒体数据集合包括经由网络接收所述第一媒体数据集合,并且其中确定将不接收所述第二媒体数据集合包括确定网络中断将阻止接收所述第二媒体数据集合。
3. 根据权利要求2所述的方法,还包括确定所述网络中断的开始时间和所述网络中断的结束时间,其中预测性地生成所述替换数据包括预测性地生成在所述网络中断的开始时间和所述网络中断的结束时间之间的时间段内的所述替换媒体数据。
4. 根据权利要求1所述的方法,其中接收所述第一媒体数据集合包括经由网络从源设备接收所述第一媒体数据集合,并且其中确定将不接收所述第二媒体数据集合包括从所述源设备接收指示将不接收所述第二媒体数据集合的数据。
5. 根据权利要求4所述的方法,其中从所述源设备接收指示将不接收所述第二媒体数据集合的数据包括从所述源设备接收补充增强信息 (SEI) 消息或实时传输协议 (RTP) 消息中的至少一者。
6. 根据权利要求4所述的方法,其中,接收所述第一媒体数据集合包括经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合,并且其中,从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据包括经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据,所述信令信道与所述媒体数据信道分离。
7. 根据权利要求1所述的方法,还包括:  
在将不接收媒体数据的情况下,向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及  
从所述用户接收输入数据,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。
8. 根据权利要求1所述的方法,其中,所述媒体数据包括视频数据或扩展现实 (XR) 数据中的至少一个,所述XR数据包括增强现实 (AR) 数据、虚拟现实 (VR) 数据或混合现实 (MR) 数据中的至少一个。
9. 根据权利要求1所述的方法,还包括:  
接收所述媒体比特流的第三媒体数据集合;  
将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及  
在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。
10. 根据权利要求1所述的方法,其中,预测性地生成替换媒体数据包括将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程。
11. 一种用于呈现媒体数据的设备,所述设备包括:

存储器,被配置为存储媒体数据;以及

一个或多个处理器,在电路中实现并且被配置为:

接收媒体比特流的第一媒体数据集合;

响应于确定在一时间段内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合的至少一个子集来预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及

呈现所述第一媒体数据集合和所述替换媒体数据。

12. 根据权利要求11所述的设备,其中所述一个或多个处理器被配置为经由网络接收所述第一媒体数据集合,并且其中为了确定将不接收所述第二媒体数据集合,所述一个或多个网络设备被配置为确定网络中断将阻止接收所述第二媒体数据集合。

13. 根据权利要求12所述的设备,其中所述一个或多个处理器还被配置为确定所述网络中断的开始时间和所述网络中断的结束时间,其中预测性地生成所述替换数据包括预测性地生成在所述网络中断的开始时间和所述网络中断的结束时间之间的时间段内的所述替换媒体数据。

14. 根据权利要求11所述的设备,其中所述一个或多个处理器被配置为经由网络从源设备接收所述第一媒体数据集合,并且其中为了确定将不接收所述第二媒体数据集合,所述一个或多个处理器被配置为从所述源设备接收指示将不接收所述第二媒体数据集合的数据。

15. 根据权利要求14所述的设备,其中为了从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据,所述一个或多个处理器被配置为从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者。

16. 根据权利要求14所述的设备,其中所述一个或多个处理器被配置为经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合,并且其中所述一个或多个处理器被配置为经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据,所述信令信道与所述媒体数据信道分离。

17. 根据权利要求11所述的设备,还包括显示器,其中所述一个或多个处理器还被配置为:

在将不接收媒体数据的情况下,经由所述显示器向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及

从所述用户接收输入数据,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

18. 根据权利要求11所述的设备,其中所述媒体数据包括视频数据或扩展现实(XR)数据中的至少一个,所述XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一个。

19. 根据权利要求11所述的设备,其中所述一个或多个处理器还被配置为:

接收所述媒体比特流的第三媒体数据集合;

将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及

在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

20. 根据权利要求11所述的设备,其中,为了预测性地生成所述替换媒体数据,所述一

个或多个处理器被配置为将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能(AI)过程或机器学习(ML)过程。

21.一种用于呈现媒体数据的设备,所述设备包括:

用于接收媒体比特流的第一媒体数据集合的装置;

用于响应于确定在一段时间内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合的至少一个子集来预测性地生成用于所述第二媒体数据集合的替换媒体数据的装置;以及

用于呈现所述第一媒体数据集合和所述替换媒体数据的装置。

22.根据权利要求21所述的设备,其中用于接收所述第一媒体数据集合的装置包括用于经由网络接收所述第一媒体数据集合的装置,还包括用于确定网络中断将阻止接收所述第二媒体数据集合的装置。

23.根据权利要求22所述的设备,还包括用于确定所述网络中断的开始时间和所述网络中断的结束时间的装置,其中用于预测性地生成所述替换数据的装置包括用于预测性地生成在所述网络中断的开始时间和所述网络中断的结束时间之间的时间段内的所述替换媒体数据的装置。

24.根据权利要求21所述的设备,其中用于接收所述第一媒体数据集合的装置包括用于经由网络从源设备接收所述第一媒体数据集合的装置,并且其中用于确定将不接收所述第二媒体数据集合的装置包括用于从所述源设备接收指示将不接收所述第二媒体数据集合的数据的装置。

25.根据权利要求24所述的设备,其中用于从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的装置包括用于从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者的装置。

26.根据权利要求24所述的设备,其中用于接收所述第一媒体数据集合的装置包括用于经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合的装置,并且其中用于从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的装置包括用于经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的装置,所述信令信道与所述媒体数据信道分离。

27.根据权利要求21所述的设备,还包括:

用于在将不接收媒体数据的情况下向用户呈现请求用户同意预测性地生成替换媒体数据的提示的装置;以及

用于从所述用户接收输入数据的装置,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

28.根据权利要求21所述的设备,其中所述媒体数据包括视频数据或扩展现实(XR)数据中的至少一个,所述XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一个。

29.根据权利要求21所述的设备,还包括:

用于接收所述媒体比特流的第三媒体数据集合的装置;

用于将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合的装置;以及

用于在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分的装

置。

30. 根据权利要求21所述的设备,其中,用于预测性地生成所述替换媒体数据的装置包括:用于将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能(AI)过程或机器学习(ML)过程的装置。

31. 一种其上存储有指令的计算机可读存储介质,所述指令在被执行时使得处理器:  
接收媒体比特流的第一媒体数据集合;

响应于确定在一段时间内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合的至少一个子集来预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及

呈现所述第一媒体数据集合和所述替换媒体数据。

32. 根据权利要求31所述的计算机可读存储介质,其中使得所述处理器接收所述第一媒体数据集合的所述指令包括使得所述处理器经由网络接收所述第一媒体数据集合的指令,还包括使得所述处理器确定网络中断将阻止接收所述第二媒体数据集合的指令。

33. 根据权利要求32所述的计算机可读存储介质,还包括使得所述处理器确定所述网络中断的开始时间和所述网络中断的结束时间的指令,其中使得所述处理器预测性地生成所述替换数据的指令包括使得所述处理器预测性地生成在所述网络中断的开始时间和所述网络中断的结束时间之间的时间段内的所述替换媒体数据的指令。

34. 根据权利要求31所述的计算机可读存储介质,其中使得所述处理器接收所述第一媒体数据集合的所述指令包括使得所述处理器经由网络从源设备接收所述第一媒体数据集合的指令,还包括使得所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的数据的指令。

35. 根据权利要求34所述的计算机可读存储介质,其中使得所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的所述指令包括使得所述处理器从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者的指令。

36. 根据权利要求34所述的计算机可读存储介质,其中使得所述处理器接收所述第一媒体数据集合的所述指令包括使得所述处理器经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合的指令,以及其中使得所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的所述指令包括使得所述处理器经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的指令,所述信令信道与所述媒体数据信道分离。

37. 根据权利要求31所述的计算机可读存储介质,还包括使得所述处理器进行以下操作的指令:

在将不接收媒体数据的情况下,向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及

从所述用户接收输入数据,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

38. 根据权利要求31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或扩展现实(XR)数据中的至少一个,所述XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一个。

39. 根据权利要求31所述的计算机可读存储介质,还包括使得所述处理器进行以下操作的指令:

接收所述媒体比特流的第三媒体数据集合;

将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及

在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

40. 根据权利要求31所述的计算机可读存储介质,其中使得所述处理器预测性地生成替换媒体数据的所述指令包括使得所述处理器将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程的指令。

## 响应于网络中断的视频内容的自动生成

[0001] 本申请要求于2022年9月20日提交的第17/933,791号美国专利申请的优先权,其全部内容通过引用并入本文。

### 技术领域

[0002] 本公开涉及媒体数据的传输。

### 背景技术

[0003] 数字视频能力可并入到广泛范围的设备中,所述设备包括数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型或台式计算机、数码相机、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、蜂窝式或卫星无线电电话、视频电话会议设备等。数字视频设备实施视频压缩技术,例如由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(Advanced Video Coding,AVC)、ITU-T H.265(也被称作高效视频译码(High Efficiency Video Coding,HEVC))定义的标准及此类标准的扩展中所描述的那些技术,以更高效地发送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,可将视频帧或切片分割成宏块。每个宏块可以被进一步分割。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。帧间译码(P或B)帧或切片中的宏块可使用相对于相同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] 在视频数据已被编码之后,视频数据可被打包以供发送或存储。视频数据可组装成符合多种标准(诸如国际标准化组织(ISO)基本媒体文件格式及其扩展,例如AVC)中的任一者的视频文件。

### 发明内容

[0006] 通常,本公开描述了用于响应于确定在特定时间段(例如,可用性的时间段)内将不接收所请求的媒体内容而自动生成媒体内容(例如,图像内容、视频内容和/或音频内容)的技术。例如,媒体数据可以经由基于计算机的网络从源设备流式传输到目的地设备。网络中断(例如,源设备和目的地设备之间的连接中断)可以阻止媒体数据的一部分及时从源设备发送到目的地设备,以允许后续媒体数据的无缝回放。网络中断可以由诸如分组丢失的网络错误和/或在无线电接入网络(RAN)的小区之间的切换期间引起。

[0007] 响应于这样的网络中断,目的地设备可以生成用于将不被递送到目的地设备的媒体数据的替换媒体数据。例如,目的地设备可以应用人工智能和/或机器学习过程来生成替换媒体数据。媒体数据可以是两个用户装备之间的实况媒体会话的一部分,诸如多媒体呼叫(例如,视频电话会议或扩展现实(XR)会话,诸如增强现实(AR)、混合现实(MR)或虚拟现实(VR)会话)。以这种方式,目的地设备可以避免请求将不递送的媒体数据的重传(这可以避免增加媒体通信会话的延迟),同时还通过避免黑屏或复制/拷贝的帧来改善用户体验。

[0008] 在一个示例中,一种呈现媒体数据的方法包括:接收媒体比特流的第一媒体数据集合;响应于确定在一时间段内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0009] 在另一示例中,一种用于呈现媒体数据的设备包括:存储器,被配置为存储媒体数据;以及一个或多个处理器,实施于电路中并且被配置为:接收媒体比特流的第一媒体数据集合;响应于确定在一时间段内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0010] 在另一示例中,一种用于呈现媒体数据的设备包括:用于接收媒体比特流的第一媒体数据集合的装置;用于响应于确定在一时间段内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合针对所述第二媒体数据集合预测性地生成替换媒体数据的装置;以及用于呈现所述第一媒体数据集合和所述替换媒体数据的装置。

[0011] 在另一示例中,一种计算机可读存储介质,其上存储有指令,所述指令在被执行时使得一个或多个处理器:接收媒体比特流的第一媒体数据集合;响应于确定在一时间段内将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0012] 在附图和以下描述中阐述了一个或多个示例的细节。其他特征、目的和优点从描述和附图以及权利要求书中将是清晰的。

## 附图说明

[0013] 图1是示出实施用于通过网络流式传输媒体数据的技术的示例系统的框图。

[0014] 图2是示出检索(retrieval)单元的组件的示例集合的框图。

[0015] 图3是示出示例多媒体内容的元素的概念图。

[0016] 图4是示出示例视频文件的元素的框图,示例视频文件可对应于表示的区段(segment)。

[0017] 图5是示出根据本公开的技术的生成替换帧的示例的概念图。

[0018] 图6是示出根据本公开的技术的生成和呈现替换媒体数据的示例方法的流程图。

[0019] 图7是示出根据本公开的技术的生成和呈现替换媒体数据的另一示例方法的流程图。

[0020] 图8是示出根据本公开的技术的生成和呈现媒体数据的示例方法的流程图。

[0021] 图9是示出包括可被配置为执行本公开的技术的设备的示例系统的框图。

[0022] 图10是示出包括用户装备(UE)设备和伴随设备的示例系统的框图。

## 具体实施方式

[0023] 通常,本公开描述了用于生成用于两个或更多个媒体设备之间的媒体流会话的替换媒体数据的技术。例如,两个用户装备(UE)设备可以参与多媒体电话会议或多媒体会



话,诸如视频电话会议或扩展现实(XR)会话,例如,增强现实(AR)、混合现实(MR)或虚拟现实(VR)会话。两个UE设备可经由基于计算机的网络(例如,蜂窝网络、局域网或短程数据传输网络)交换媒体数据。

[0024] 例如,一个或两个UE设备可由在无线电接入网络(RAN)的小区之间移动的用户携带。当在小区之间移动时,UE设备可以执行从一个小区的基站到不同小区的基站的切换过程。

[0025] 在这样的切换期间,UE设备和基站可以交换无线电资源控制(RRC)消息并且遵循随机接入信道(RACH)过程。RACH过程可花费大约50ms或更长时间来执行。当执行从一个小区到另一个小区的切换时,作为正在进行的通信会话(诸如媒体电话会议)的一部分的数据传输被中断,使得多个视频帧(或其他媒体数据,诸如XR数据或图像数据)将不满足递送期限,这可能显著影响用户的体验。

[0026] 根据本公开的技术的UE设备可以检测该网络中断,并且作为响应,生成用于由于中断而在一段时间内将不接收的媒体数据的替换媒体数据,而不是简单地呈现黑屏或拷贝的(例如,复制的)视频帧。将不接收的媒体数据可以是完全丢失、毁坏和/或延迟以至延迟的媒体数据晚于UE设备处的预期呈现时间到达的程度的媒体数据。UE设备可以使用人工智能(AI)和/或机器学习(ML)过程来生成替换媒体数据。UE设备可以将最近接收的媒体数据提供给AI/ML过程,并且AI/ML过程可以预测性地生成用于由于网络中断而在一段时间内将不接收的媒体数据的替换媒体数据。该时间段可以是媒体数据否则将可用(例如,可解码和可呈现)的时间。

[0027] 安装在UE上的执行媒体通信会话的某些应用(例如,Apple FaceTime、Microsoft Teams、Zoom Cloud Meeting、Google Meet、Facebook Messenger等)可以被配置为允许启用或禁用根据本公开的技术的替换媒体数据的生成。例如,应用可以被配置为启用或禁用替换媒体数据生成。

[0028] UE设备可以请求用户同意在发起(或建立)媒体会话之前、在发起(或建立)媒体会话时和/或响应于检测到网络中断,生成替换媒体数据,例如,作为媒体会议应用的设置。在一些情况下,用户可能不想要具有替换媒体数据,例如,如果媒体会话涉及有新闻价值的事件,诸如直播体育事件、政治事件或其他这样的事件。同样,UE设备可以被配置为不生成某些类型的媒体数据(诸如预先录制的电影或电视节目)的替换数据,以避免知识产权(例如,版权)权利侵犯和/或某些类型的媒体会话(诸如个人视频电话会议(相对于工作相关会话))。换句话说,在一些实现中,在UE设备处是否允许(或启用)替换媒体数据的生成可以进一步(或替代地)基于媒体数据的一个或多个特性、媒体数据的类型和/或媒体会话的类型。

[0029] 通过应用本公开的技术,UE设备可以响应于确定在一段时间(例如,期间媒体数据可用(到指定的输出时间可解码和可呈现)的一段时间)内将不接收某些媒体数据而生成替换媒体数据。通过执行这些技术,UE设备可以避免请求将不接收的媒体数据的重传,这将向媒体流送会话引入等待时间并消耗附加网络资源(例如,带宽)。同样地,UE设备还可以呈现替换媒体数据,这是对呈现拷贝(例如,复制)的帧或空白屏幕的用户体验的改进。以这种方式,本公开的技术提供了对媒体传输、生成和呈现领域的技术改进,以及对UE设备和媒体传输和呈现中涉及的其他网络设备的操作的技术改进。

[0030] 本公开的技术可应用于符合根据ISO基本媒体文件格式、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。

[0031] 在HTTP流式传输中,频繁使用的操作包括HEAD、GET及部分GET。HEAD操作检索与给定统一资源定位符 (URL) 或统一资源名称 (URN) 相关联的文件的标头,而不检索与URL或URN相关联的有效载荷。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个字节,其中字节数目对应于所接收的字节范围。因此,可提供电影片段用于HTTP流式传输,这是因为部分GET操作可获得一个或多个个别电影片段。在电影片段中,可以存在不同轨道 (track) 的若干轨道片段。在HTTP流式传输中,媒体呈现可以是客户端可访问的数据的结构化集合。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。

[0032] 在使用HTTP流式传输流送3GPP数据的示例中,可存在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级别 (level))、不同译码标准或译码标准的扩展 (例如多视图和/或可缩放扩展) 或不同比特速率。此类表示的清单可在媒体呈现描述 (MPD) 数据结构中定义。媒体呈现可以对应于HTTP流式传输客户端设备可访问的数据的结构化集合。HTTP流式传输客户端设备可以请求和下载媒体数据信息以向客户端设备的用户呈现流式传输服务。可以在MPD数据结构中描述媒体呈现,所述MPD数据结构可以包括MPD的更新。

[0033] 媒体呈现可含有一个或多个时段 (Period) 的序列。每个时段可以延展直到下一时段的开始,或者在最后时段的情况下直到媒体呈现的结束。每一时段可含有用于相同媒体内容的一个或多个表示。表示可为音频、视频、计时文本或其它此类数据的数个替代编码版本中的一者。表示可因编码类型而不同,例如,因用于视频数据的比特速率、分辨率和/或编解码器,以及用于音频数据的比特速率、语言和/或编解码器而不同。术语表示可用于指代对应于多媒体内容的特定时段且以特定方式编码的编码音频或视频数据的部分 (section)。

[0034] 特定时段的表示可指派给由MPD中的属性指示的群组,所述属性指示表示所属的适配集。相同适配集中的表示通常被视为彼此的替代,因为客户端设备可在这些表示之间动态且无缝地切换,例如以执行带宽适配。例如,特定时段内的视频数据的每一表示可被指派到相同适配集,使得表示中的任一者可被选择用于解码以呈现对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。在一些示例中,一个时段内的媒体内容可由来自群组0 (如果存在) 的一个表示或来自每一非零群组的至多一个表示的组合表示。时段的每一表示的定时数据可相对于时段的开始时间来表达。

[0035] 表示可包括一个或多个区段。每一表示可包括初始化区段,或表示的每一区段可为自初始化的。当存在时,初始化区段可以包含用于访问表示的初始化信息。一般来说,初始化区段不含有媒体数据。区段可以由标识符唯一地引用,诸如统一资源定位符 (URL)、统一资源名称 (URN) 或统一资源标识符 (URI)。MPD可以提供每个区段的标识符。在一些示例中,MPD还可提供呈范围 (range) 属性形式的字节范围,其可对应于可由URL、URN或URI访问的文件内的区段的数据。

[0036] 可选择不同表示以用于基本上同时检索不同类型的媒体数据。例如,客户端设备

可选择从中检索区段的音频表示、视频表示和计时文本表示。在一些示例中,客户端设备可选择用于执行带宽适配的特定适配集。即,客户端设备可选择包含视频表示的适配集、包含音频表示的适配集和/或包含计时文本的适配集。替代地,客户端设备可选择用于某些类型的媒体(例如,视频)的适配集,并且直接选择用于其它类型的媒体(例如,音频和/或计时文本)的表示。

[0037] 图1是示出实施用于通过网络流式传输媒体数据的技术的示例系统10的框图。在该示例中,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通信地耦合,网络74可以包括互联网。在一些情况下,经由网络74或其它连接到设备的连接可涵盖无线广域网(WWAN)、无线局域网(WLAN)、蓝牙或其它长或短程无线连接。在一些示例中,内容准备设备20和服务器设备60还可以通过网络74或另一网络耦合,或者可以直接通信地耦合。在一些示例中,内容准备设备20和服务器设备60可以包括相同的设备。

[0038] 在图1的示例中,内容准备设备20包括音频源22和视频源24。音频源22可包括(例如)麦克风,麦克风产生表示待由音频编码器26编码的捕获音频数据的电信号。替代地,音频源22可包括存储先前记录的音频数据的存储介质、例如计算机化合成器的音频数据生成器、或任何其它音频数据源。视频源24可包括产生待由视频编码器28编码的视频数据的摄像机、用先前记录的视频数据编码的存储介质、例如计算机图形源的视频数据生成单元、或任何其它视频数据源。在所有示例中,内容准备设备20不一定通信地耦合到服务器设备60,而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。

[0039] 原始音频和视频数据可以包括模拟或数字数据。模拟数据可在由音频编码器26和/或视频编码器28编码之前被数字化。音频源22可在讲话参与者正在讲话时从讲话参与者获得音频数据,并且视频源24可同时获得讲话参与者的视频数据。在其它示例中,音频源22可包括包含存储的音频数据的计算机可读存储介质,并且视频源24可包括包含存储的视频数据的计算机可读存储介质。以此方式,本公开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存档、预先记录的音频和视频数据。

[0040] 对应于视频帧的音频帧通常为含有音频数据的音频帧,该音频数据与视频帧内所含有的视频源24捕获(或生成)的视频数据同时由音频源22捕获(或生成)。例如,虽然讲话参与者通常通过讲话产生音频数据,但是音频源22捕获音频数据,并且视频源24同时(即,在音频源22捕获音频数据的同时)捕获讲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视频帧。因此,对应于视频帧的音频帧大体上对应于同时捕获音频数据和视频数据且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的情形。

[0041] 在一些示例中,音频编码器26可对每个编码音频帧中的表示记录编码音频帧的音频数据的的时间的时间戳进行编码,且类似地,视频编码器28可对每个编码视频帧中的表示记录编码视频帧的视频数据的的时间的时间戳进行编码。在此类示例中,对应于视频帧的音频帧可包括包含时间戳的音频帧以及包含相同时间戳的视频帧。内容准备设备20可包括内部时钟,音频编码器26和/或视频编码器28可从该内部时钟生成时间戳,或音频源22和视频源24可使用该内部时钟以分别使音频和视频数据与时间戳相关联。

[0042] 在一些示例中,音频源22可将对应于记录音频数据的的时间的数据发送到音频编码器26,并且视频源24可将对应于记录视频数据的的时间的数据发送到视频编码器28。在一些

示例中,音频编码器26可对编码音频数据中的序列识别符进行编码以指示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还可使用序列识别符来指示编码视频数据的相对时间排序。类似地,在一些示例中,序列标识符可以被映射或以其他方式与时间戳相关。

[0043] 音频编码器26通常产生编码音频数据流,而视频编码器28产生编码视频数据流。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。基本流为表示的单个数字译码(可能压缩)分量。例如,表示的译码视频或音频部分可为基本流。基本流可在封装于视频文件内之前转换成包化基本流(PES)。在相同表示内,流ID可用于区分属于一个基本流的PES分组与属于另一基本流的PES分组。基本流的数据的基本单元为包化基本流(PES)分组。因此,译码视频数据大体上对应于基本视频流。类似地,音频数据对应于一个或多个相应基本流。

[0044] 许多视频译码标准(例如,ITU-T H.264/AVC、高效率视频译码(HEVC)、多功能视频译码(VVC)、增强型视频译码(EVC)和AOMedia视频1(AV1))定义无错误比特流的语法、语义和解码过程,其中的任一者符合某一简档或级别。视频译码标准通常不指定编码器,但是编码器的任务是保证所生成的比特流对于解码器是符合标准的。在视频译码标准的上下文中,“简档”对应于算法、特征或工具和应用用于其的约束的子集。例如,如由H.264标准定义的,“简档”是由H.264标准指定的整个比特流语法的子集。“级别”对应于解码器资源消耗的限制,例如像解码器存储器和计算,其与图片的分辨率、比特率和块处理速率有关。简档可以用profile\_idc(简档指示符)值来信令通知,而级别可以用level\_idc(级别指示符)值来信令通知。

[0045] 例如,H.264标准认识到,在由给定简档的语法强加的界限内,取决于比特流中的语法元素所取的值(例如解码图片的指定大小),仍可能需要编码器和解码器的性能的大变化。H.264标准还认识到,在许多应用中,实现能够处理特定简档内的语法的所有假设使用的解码器既不实际也不经济。因此,H.264标准将“级别”定义为强加于比特流中的语法元素的值的指定约束集合。这些约束可以是对值的简单限制。替代地,这些约束可采取对值的算术组合(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标准还提供,单独实施方式可支持每个所支持的简档的不同级别。

[0046] 符合简档的解码器通常支持简档中定义的所有特征。例如,作为译码特征,在H.264/AVC的基线简档中不支持B图片译码,但是在H.264/AVC的其它简档中支持B图片译码。符合级别的解码器应该能够对不需要超出级别中定义的限制的资源的任何比特流解码。简档和级别的定义可能有助于可解释性。例如,在视频传输期间,可以针对整个传输会话协商和同意一对简档和级别定义。更具体地,在H.264/AVC中,级别可定义对需要处理的宏块的数目、解码图片缓冲器(DPB)大小、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以及B块是否具有小于 $8 \times 8$ 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比特流解码。

[0047] 在图1的示例中,内容准备设备20的封装单元30从视频编码器28接收包括译码视频数据的基本流,并且从音频编码器26接收包括译码音频数据的基本流。在一些示例中,视频编码器28和音频编码器26可各自包括用于从编码数据形成PES分组的包化器。在其它示例中,视频编码器28和音频编码器26可各自与用于从编码数据形成PES分组的相应包化器

接口。在其它示例中,封装单元30可包括用于从编码音频和视频数据形成PES分组的包化器。

[0048] 视频编码器28可以多种方式编码多媒体内容的视频数据,从而以各种比特速率且以各种特性产生多媒体内容的不同表示,所述特性例如像素分辨率、帧速率、对各种译码标准的符合性、对各种译码标准的各种简档和/或简档级别的符合性、具有一个或多个视图的表示(例如,用于二维或三维回放),或其它此类特性。如本公开中所使用的,表示可包括音频数据、视频数据、文本数据(例如,用于隐藏式(closed)字幕)或其它此类数据中的一者。表示可包括基本流,例如音频基本流或视频基本流。每个PES分组可包括识别PES分组所属的基本流的stream\_id。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)。

[0049] 封装单元30从音频编码器26和视频编码器28接收表示的基本流的PES分组,并且从PES分组形成对应网络抽象层(NAL)单元。译码视频区段可组织成NAL单元,NAL单元提供寻址例如视频电话、存储、广播或流式传输的应用的“网络友好”视频表示。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括块、宏块和/或切片级数据。其它NAL单元可为非VCL NAL单元。在一些示例中,一个时间实例中的译码图片(通常呈现为主要译码图片)可包含于接入单元中,该接入单元可包括一个或多个NAL单元。

[0050] 非VCL NAL单元可包括参数集NAL单元和SEI NAL单元等。参数集可含有序列级标头信息(在序列参数集(SPS)中)和不频繁改变的图片级标头信息(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不频繁改变的信息;因此,可以改进译码效率。此外,参数集的使用可以实现重要标头信息的带外传输,从而避免对用于错误复原的冗余传输的需要。在带外传输示例中,参数集NAL单元可在与其它NAL单元(例如SEI NAL单元)不同的信道上发送。

[0051] 补充增强信息(SEI)可含有对来自VCL NAL单元的译码图片样本进行解码并非必需的信息,但可辅助与解码、显示、错误复原和其它目的相关的过程。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部分,并且因此对于标准兼容解码器实施方式并不总是强制性的。SEI消息可以是序列级SEI消息或图片级SEI消息。一些序列级信息可包含于SEI消息中,例如SVC的示例中的可缩放性信息SEI消息和MVC中的视图可缩放性信息SEI消息。这些示例SEI消息可传达关于例如操作点的提取及操作点的特性的信息。另外,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可以根据可扩展标记语言(XML)来格式化MPD。

[0052] 封装单元30可将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)一起提供到输出接口32。输出接口32可包括网络接口或用于写入到存储介质的接口,例如通用串行总线(USB)接口、CD或DVD写入器或刻录机、到磁性或快闪存储介质的接口,或用于存储或发射媒体数据的其它接口。封装单元30可将多媒体内容的表示中的每一者的数据提供到输出接口32,输出接口32可经由网络传输或存储介质将数据发送到服务器设备60。在图1的示例中,服务器设备60包括存储各种多媒体内容64的存储介质62,每一多媒体内容64包括相应清单文件66和一个或多个表示68A到68N(表示68)。在一些示例中,输出接口32还可将数据直接发送到网络74。

[0053] 在一些示例中,表示68可分离成适配集。即,表示68的各种子集可包括相应共同特性集,例如编解码器、简档和级别、分辨率、视图数目、区段的文件格式、可识别待与待解码及呈现(例如由扬声器)的表示和/或音频数据一起显示的文本的语言或其它特性的文本类型信息、可描述适配集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、描述特定观众的内容适合性的分级信息或类似者。

[0054] 清单文件66可包括指示对应于特定适配集的表示68的子集以及适配集的共同特性的数据。清单文件66还可包括表示适配集的个别表示的个别特性(例如比特速率)的数据。以此方式,适配集可提供简化的网络带宽适配。可使用清单文件66的适配集元素的子元素来指示适配集中的表示。

[0055] 服务器设备60包括请求处理单元70和网络接口72。在一些示例中,服务器设备60可以包括多个网络接口。此外,服务器设备60的任何或所有特征可以在内容递送网络的其他设备上实现,诸如路由器、网桥、代理设备、交换机或其他设备。在一些示例中,内容递送网络的中间设备可高速缓存多媒体内容64的数据,并且包括基本上符合服务器设备60的组件的组件。通常,网络接口72被配置为经由网络74发送和接收数据。

[0056] 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)接收对存储介质62的数据的网络请求。例如,请求处理单元70可实施超文本传输协议(HTTP)版本1.1,如RFC 2616,R.Fielding等人的“Hypertext Transfer Protocol-HTTP/1.1”,Network Working Group,IETF,1999年6月中所描述。即,请求处理单元70可被配置为接收HTTP GET或部分GET请求,并且响应于请求而提供多媒体内容64的数据。请求可例如使用区段的URL来指定表示68中的一者的区段。在一些示例中,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可进一步被配置为服务HTTP HEAD请求以提供表示68中的一者的区段的标头数据。在任何情况下,请求处理单元70可被配置为处理请求以将请求的数据提供到请求设备,例如客户端设备40。

[0057] 另外或替代地,请求处理单元70可被配置为经由广播或多播协议(例如,eMBMS)递送媒体数据。内容准备设备20可以以与所描述的基本上相同的方式创建DASH区段和/或子区段,但是服务器设备60可以使用eMBMS或另一广播或多播网络传输协议来递送这些区段或子区段。例如,请求处理单元70可被配置为从客户端设备40接收多播群组加入请求。也就是说,服务器设备60可以向与特定媒体内容(例如,直播事件的广播)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又可以提交加入多播群组的请求。该请求可以在整个网络74(例如,构成网络74的路由器)中传播,使得路由器将去往与多播群组相关联的IP地址的业务引导到订阅客户端设备,例如客户端设备40。

[0058] 如图1的示例中所说明的,多媒体内容64包括清单文件66,其可对应于媒体呈现描述(MPD)。清单文件66可含有不同替代表示68(例如,具有不同质量的视频服务)的描述,并且描述可包括例如编解码器信息、简档值、级别值、比特速率和表示68的其它描述性特性。客户端设备40可检索媒体呈现的MPD以确定如何接入表示68的区段。

[0059] 具体地,检索单元52可检索客户端设备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由客户端设备40的用户选择的语言偏好、与由客户端设备40的用户设置的深度偏好相对应的一个或多个相机视角和/

或由客户端设备40的用户选择的评级偏好中的任何一个或全部。检索单元52可包括例如被配置为提交HTTP GET和部分GET请求的网络浏览器或媒体客户端。检索单元52可对应于由客户端设备40的一个或多个处理器或处理单元(未示出)执行的软件指令。在一些示例中,关于检索单元52所描述的功能性的全部或部分可实施于硬件,或硬件、软件和/或固件的组合中,其中可提供必要硬件以执行用于软件或固件的指令。

[0060] 检索单元52可将客户端设备40的解码和渲染能力与由清单文件66的信息指示的表示68的特性进行比较。检索单元52可最初检索清单文件66的至少一部分以确定表示68的特性。例如,检索单元52可请求描述一个或多个适配集的特性的清单文件66的一部分。检索单元52可选择具有可由客户端设备40的译码和渲染能力满足的特性的表示68的子集(例如,适配集)。检索单元52然后可确定适配集中的表示的比特速率,确定网络带宽的当前可用量,并且从具有可由网络带宽满足的比特速率的表示中的一者检索区段。

[0061] 一般来说,较高比特速率表示可产生较高质量视频回放,而较低比特速率表示可在可用网络带宽减小时提供足够质量视频回放。因此,当可用网络带宽相对高时,检索单元52可从相对高比特速率表示检索数据,而当可用网络带宽低时,检索单元52可从相对低比特速率表示检索数据。以这种方式,客户端设备40可以通过网络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。

[0062] 另外或替代地,检索单元52可被配置为根据广播或多播网络协议(例如,eMBMS或IP多播)接收数据。在此类示例中,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发出进一步请求。当不再需要多播群组的数据时,检索单元52可提交离开多播群组的请求,例如以停止回放或将信道改变到不同多播群组。

[0063] 网络接口54可接收选定表示的区段的数据且将该数据提供到检索单元52,检索单元52又可将区段提供到解封装单元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流解包化以检索编码数据,并且取决于编码数据是音频流还是视频流的部分(例如,如由流的PES分组标头所指示)而将编码数据发送到音频解码器46或视频解码器48。音频解码器46对编码音频数据进行解码并且将解码的音频数据发送到音频输出42,而视频解码器48对编码视频数据进行解码并且将解码的视频数据(其可包括流的多个视图)发送到视频输出44。

[0064] 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各自可实施为适用的多种合适处理电路中的任一者,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)、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中的每一者可包括在一个或多个编码器或解码器中,该编码器或解码器中的任一者可集成为组合式视频编码器/解码器(CODEC)的部分。同样地,音频编码器26和音频解码器46中的每一者可包括在一个或多个编码器或解码器中,该编码器或解码器中的任一者可集成为组合式CODEC的部分。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、检索单元52和/或解封装单元50的装置可包括集成电路、微处理器和/或无线通信设备,例如蜂窝式电话。

[0065] 客户端设备40、服务器设备60和/或内容准备设备20可被配置成根据本公开的技



术来操作。出于示例的目的,本公开关于客户端设备40和服务设备60描述了这些技术。然而,应理解,代替服务器设备60(或除了服务器设备60之外),内容准备设备20还可被配置为执行这些技术。

[0066] 封装单元30可形成NAL单元,该NAL单元包括识别NAL单元所属的程序的标头以及有效载荷,例如音频数据、视频数据或描述NAL单元所对应的传输或程序流的数据。例如,在H.264/AVC中,NAL单元包括1字节标头和不同大小的有效载荷。在其有效载荷中包括视频数据的NAL单元可包括各种粒度水平的视频数据。例如,NAL单元可包括视频数据的块、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可从视频编码器28接收呈基本流的PES分组的形式编码的视频数据。封装单元30可使每一基本流与对应节目相关联。

[0067] 封装单元30还可以组装来自多个NAL单元的接入单元。一般来说,接入单元可包括用于表示视频数据的帧以及对应于该帧的音频数据(当此类音频数据可得时)的一个或多个NAL单元。接入单元通常包括用于一个输出时间实例的所有NAL单元,例如,用于一个时间实例的所有音频和视频数据。例如,如果每一视图具有20帧每秒(fps)的帧速率,则每一时间实例可对应于0.05秒的时间间隔。在该时间间隔期间,可以同时渲染相同接入单元(相同时间实例)的所有视图的特定帧。在一个示例中,接入单元可包括一个时间实例中的译码图片,其可呈现为主要译码图片。

[0068] 因此,接入单元可以包括公共时间实例的所有音频和视频帧,例如,对应于时间X的所有视图。本公开还将特定视图的编码图片称作“视图分量”,即,视图分量可包括特定时间处的特定视图的编码图片(或帧)。因此,接入单元可被定义为包括共同时间实例的所有视图分量。接入单元的解码顺序不一定需要与输出或显示顺序相同。

[0069] 媒体呈现可包括媒体呈现描述(MPD),其可含有不同替代表示(例如,具有不同质量的视频服务)的描述,并且描述可包括例如编解码器信息、简档值和级别值。MPD为清单文件(例如,清单文件66)的一个示例。客户端设备40可检索媒体呈现的MPD以确定如何接入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。

[0070] 清单文件66(其可包括例如MPD)可通告表示68的区段的可用性。即,MPD可包括指示表示68中的一者的第一区段变得可用的挂钟时间(wall-clock time)的信息,以及指示表示68内的区段的持续时间的信息。以此方式,客户端设备40的检索单元52可基于特定区段之前的区段的起始时间以及持续时间来确定每一区段何时可用。

[0071] 在封装单元30已基于接收的数据将NAL单元和/或接入单元组装成视频文件之后,封装单元30将视频文件传递到输出接口32以供输出。在一些示例中,封装单元30可本地存储视频文件或经由输出接口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设备40。输出接口32可包括(例如)发送器、收发器、用于将数据写入到计算机可读介质的设备,例如像光学驱动器、磁性介质驱动器(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介质,例如像发射信号、磁性介质、光学介质、存储器、快闪驱动器或其它计算机可读介质。

[0072] 网络接口54可经由网络74接收NAL单元或接入单元,并且经由检索单元52将NAL单元或接入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流解包以检索编码数据,并取决于编码数据是音频流还是视频流的部分(例如,如由流的PES分组标头所指示)而将编码数据发送到音频解码器46或视频解码器48。音频解



码器46对编码音频数据进行解码并且将解码的音频数据发送到音频输出42,而视频解码器48对编码视频数据进行解码并且将解码的视频数据(其可包括流的多个视图)发送到视频输出44。

[0073] 在一些示例中,单个用户装备(UE)设备可以包括内容准备设备20、服务器设备60和客户端设备40的每个组件。UE设备可以参与与另一UE设备的媒体电信会话,所述另一UE设备也包括内容准备设备20、服务器设备60和客户端设备40的组件或类似组件。以这种方式,每个UE设备可以发送和接收媒体数据(例如,视频、音频和/或图像数据)作为媒体电信会话的一部分。本公开的技术一般地关于内容准备设备20、服务器设备60或客户端设备40中的一者来描述,但是这些技术也可由包括内容准备设备20、服务器设备60和客户端设备40的组件或类似组件的UE设备来执行。

[0074] 客户端设备40可以例如表示蜂窝电话、智能电话、平板计算机、膝上型计算机、头戴式显示器(HMD)等中的任何一种。服务器设备60可以表示网络服务器、第二UE设备,诸如第二蜂窝电话、智能电话、平板计算机、膝上型计算机等。在一些示例中,服务器设备60可以被配置为UE,并且客户端设备40可以被配置为服务器设备60的伴随设备,使得网络74可以表示相对短距离的通信耦合,诸如服务器设备60和客户端设备40之间的无线局域网(WLAN)、WiFi或蓝牙无线链路。

[0075] 在一些示例中,服务器设备60(例如,网络接口72)或内容准备设备20可被配置为检测网络74内的网络中断。例如,包括内容准备设备20和/或服务器设备60的组件并且正在向客户端设备40发送媒体数据的UE设备可能经历从一个基站到另一基站的网络切换。作为响应,视频编码器28形成要发送到客户端设备40的指示网络中断的数据(例如,被称为元数据或非VCL数据的非视频内容数据)。例如,视频编码器28可形成指示客户端设备40应进入自动驾驶模式的非VCL数据(例如,补充增强信息(SEI)消息),在所述自动驾驶模式期间客户端设备40将生成替换数据。视频编码器28或内容准备设备20和/或服务器设备60的另一组件可形成数据(例如,SEI消息)以指示进入(例如,转变到)自动驾驶模式的开始时间和结束(例如,转变出)自动驾驶模式的结束时间。在一些示例中,数据还可以包括使客户端设备40发出提示以请求用户同意进入自动驾驶模式的消息收发。在其它示例中,网络接口72可根据通信协议(例如实时传输协议(RTP))形成包括自动驾驶模式的开始时间和结束时间的消息,而非发送SEI消息。

[0076] 在一些示例中,客户端设备40的应用可以包括指示是否启用自动驾驶模式的配置数据。假设根据该配置数据启用自动驾驶模式,响应于接收到指示进入自动驾驶模式的数据,视频解码器48或客户端设备40的另一组件可以提示其用户同意使用自动驾驶模式。在一些情况下,启用自动驾驶模式的配置数据可以由用户提供,使得仅使用启用自动驾驶模式的预授权,而在其他示例中,当通过配置数据启用自动驾驶模式时,用户仍然可以接收提示以提供关于是否进入自动驾驶模式的同意。如果客户端设备40从用户接收到使用自动驾驶模式的肯定同意或者确定用户先前已经授权同意使用自动驾驶模式,则视频解码器48(或客户端设备40的另一单元)可以生成替换视频内容(例如,使用AI/ML),在指定的开始时间开始并在指定的结束时间停止,或者至少直到接收的媒体数据已经及时解码。在一些示例中,替换媒体数据可以在媒体电信会话期间由视频解码器48连续地生成,但是仅在自动驾驶模式期间呈现。

[0077] 如上所述,在一些示例中,除了视频数据之外或替代视频数据,媒体电信会话还可以包括XR/AR/VR/MR数据。除了视频数据之外或作为视频数据的替代,本公开的技术还可以用于生成替换XR/AR/VR/MR数据。XR内容(例如,AR、VR或MR内容)可以是例如要呈现给用户的计算机生成的2D或3D图形数据。

[0078] 在另一示例中,客户端设备40可以被配置为检测网络74内的网络中断。例如,包括用于经由网络74接收和呈现媒体数据的客户端设备40的组件的UE设备可以确定UE设备正在经历从一个基站到另一基站的网络切换或其他网络中断。作为响应,视频解码器48(或客户端设备40的另一元件)可以提示客户端设备40的用户同意进入自动驾驶模式(即,生成用于确定在网络中断期间将不会接收到的媒体数据的替换媒体数据)。如果客户端设备40从用户接收到指示肯定同意进入自动驾驶模式的输入,并且如果启用自动驾驶模式,则客户端设备40可以例如使用AI/ML过程生成用于确定在网络中断期间将不会接收到的媒体数据的替换媒体数据。

[0079] 在一些示例中,客户端设备40还可以向内容准备设备20和/或服务器设备60发送指示已经发生网络中断的数据。数据可以指示网络中断的时间段的开始时间和结束时间(例如,结束时间的估计)。作为响应,内容准备设备20和/或服务器设备60可以在指示的时间段期间停止发送媒体数据。在一些示例中,视频编码器28可响应于来自客户端设备40的数据而生成待发送到客户端设备40的下一视频帧作为帧内预测编码帧(I帧)。在一些示例中,视频编码器28可将后续图片的参考图片改变为不间断帧(例如,最后不间断单向帧间预测帧(P帧))。在一些示例中,视频编码器28可使用与视频解码器48相同的AI/ML模型来生成参考图片,然后使用这些生成的参考图片来执行后续帧的帧间预测,一旦自动驾驶模式时段结束,所述后续帧要发送到客户端设备40。

[0080] 在一些示例中,在网络中断已经结束之后,视频解码器48可以将接收的媒体数据与生成的替换媒体数据混合,使得从替换媒体数据到接收的媒体数据的转换对于客户端设备40的用户而言不太可感知。

[0081] 虽然关于DASH解释了图1,但是本公开的技术可结合多种不同媒体通信协议中的任一者使用。例如,本公开的技术还可以(附加地或替代地)结合实时传输协议(RTP)、实时流协议(RTSP)、网络实时通信(WebRTC)、安全可靠传输(SRT)等或其任何组合来使用。

[0082] 图2是更详细地示出图1的检索单元52的组件的示例集合的框图。在此示例中,检索单元52包括eMBMS中间件单元100、DASH客户端110和媒体应用112。

[0083] 在此示例中,eMBMS中间件单元100进一步包括eMBMS接收单元106、高速缓冲存储器104和代理服务器单元102。在此示例中,eMBMS接收单元106被配置为经由eMBMS接收数据,例如根据单向传输上的文件递送(FLUTE),在T.Paila等人的“FLUTE-File Delivery over Unidirectional Transport”(Network Working Group,RFC 6726,2012年11月,可在[tools.ietf.org/html/rfc6726](http://tools.ietf.org/html/rfc6726)处获得)中描述的。即,eMBMS接收单元106可经由广播从(例如)可充当广播/多播服务中心(BM-SC)的服务器设备60接收文件。

[0084] 当eMBMS中间件单元100接收文件的数据时,eMBMS中间件单元可将接收的数据存储于高速缓冲存储器104中。高速缓冲存储器104可包括计算机可读存储介质,例如快闪存储器、硬盘、RAM或任何其它合适的存储介质。

[0085] 代理服务器单元102可以充当DASH客户端110的服务器。例如,代理服务器单元102

可将MPD文件或其它清单文件提供到DASH客户端110。代理服务器单元102可通告MPD文件中的区段的可用性时间,以及可从中检索区段的超链接。这些超链接可以包括对应于客户端设备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可使用HTTP GET或部分GET请求从代理服务器单元102请求区段。例如,对于可从链路http://127.0.0.1/rep1/seg3获得的区段,DASH客户端110可以构建包括对http://127.0.0.1/rep1/seg3的请求的HTTP GET请求,并将该请求提交给代理服务器单元102。代理服务器单元102可从高速缓冲存储器104检索请求的数据并且响应于此类请求而将数据提供到DASH客户端110。

[0086] 图3是示出示例多媒体内容120的元素的概概念图。多媒体内容120可对应于多媒体内容64(图1),或存储于存储介质62中的另一多媒体内容。在图3的示例中,多媒体内容120包括媒体呈现描述(MPD)122和多个表示124A到124N(表示124)。表示124A包括可选标头数据126和区段128A到128N(区段128),而表示124N包括可选标头数据130和区段132A到132N(区段132)。为方便起见,字母N用以指定表示124中的每一者中的最后电影片段。在一些示例中,表示124之间可存在不同数目个电影片段。

[0087] MPD 122可包括与表示124分离的数据结构。MPD 122可对应于图1的清单文件66。同样地,表示124可对应于图1的表示68。一般来说,MPD 122可包括大体上描述表示124的特性的数据,所述特性例如译码及显现特性、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(trick)模式信息(例如,指示包括时间子序列的表示的信息)和/或用于检索远程时段(例如,用于在回放期间将定向广告插入到媒体内容中)的信息。

[0088] 标头数据126(当存在时)可描述区段128的特性,例如随机接入点(RAP,也被称作流接入点(SAP)的时间位置、区段128中的哪一者包括随机接入点、与区段128内的随机接入点的字节偏移、区段128的统一资源定位符(URL)、或区段128的其它方面。标头数据130(当存在时)可描述区段132的类似特性。另外或替代地,此类特性可完全包括在MPD 122内。

[0089] 区段128、132包括一个或多个译码视频样本,其中的每一者可包括视频数据的帧或切片。区段128的译码视频样本中的每一者可具有类似特性,例如高度、宽度及带宽要求。此类特性可由MPD 122的数据描述,但图3的示例中未示出此类数据。MPD 122可包括如3GPP规范所描述的特性,其中添加本公开中所描述的用信号表示的信息中的任一者或全部。

[0090] 区段128、132中的每一者可与唯一统一资源定位符(URL)相关联。因此,区段128、132中的每一者可使用流式传输网络协议(例如DASH)独立地检索。以此方式,目的地设备(例如客户端设备40)可使用HTTP GET请求来检索区段128或132。在一些示例中,客户端设备40可使用HTTP部分GET请求来检索区段128或132的特定字节范围。

[0091] 图4是示出示例视频文件150的元素的框图,示例视频文件150可对应于表示的区段,例如图3的区段128、132中的一者。区段128、132中的每一者可包括基本上符合图4的示例中所说明的数据布置的数据。视频文件150可以说是封装区段。如上所述,根据ISO基本媒体文件格式的视频文件及其扩展存储一序列对象中的数据,被称为“框”。在图4的示例中,视频文件150包括文件类型(FTYP)框152、电影(MOOV)框154、区段索引(sidx)框162、电影片段(MOOF)框164、以及电影片段随机接入(MFRA)框166。尽管图4表示视频文件的示例,但应理解,根据ISO基本媒体文件格式及其扩展,其它媒体文件可包括类似于视频文件150的数

据结构化的其它类型的媒体数据(例如,音频数据、计时文本数据等)。

[0092] 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可包括识别描述视频文件150的最佳用途的规范的数据。文件类型框152可替代地放置在MOOV框154、电影片段框164和/或MFRA框166之前。

[0093] 在一些示例中,例如视频文件150的区段可包括FTYP框152之前的MPD更新框(未展示)。MPD更新框可包括指示对应于包括视频文件150的表示的MPD将被更新的信息,连同用于更新MPD的信息。例如,MPD更新框可提供待用于更新MPD的资源的URI或URL。作为另一示例,MPD更新框可包括用于更新MPD的数据。在一些示例中,MPD更新框可紧跟在视频文件150的区段类型(STYP)框(未示出)之后,其中STYP框可定义视频文件150的区段类型。

[0094] 在图4的示例中,MOOV框154包括电影标头(MVHD)框156、轨道(TRAK)框158、以及一个或多个电影扩展(MVEX)框160。一般来说,MVHD框156可描述视频文件150的一般特性。例如,MVHD框156可包括描述视频文件150最初创建的时间、视频文件150最后修改的时间、视频文件150的时标、视频文件150的回放持续时间的数据,或大体上描述视频文件150的其它数据。

[0095] TRAK框158可包括用于视频文件150的轨道的数据。TRAK框158可包括描述对应于TRAK框158的轨道的特性的轨道标头(TKHD)框。在一些示例中,TRAK框158可包括译码视频图片,而在其它示例中,轨道的译码视频图片可包括在电影片段164中,电影片段164可由TRAK框158和/或sidx框162的数据参考。

[0096] 在一些示例中,视频文件150可包括多于一个轨道。因此,MOOV框154可包括等于视频文件150中的轨道的数目的多个TRAK框。TRAK框158可描述视频文件150的对应轨道的特性。例如,TRAK框158可描述对应轨道的时间和/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,类似于MOOV框154的TRAK框158的TRAK框可描述参数集轨道的特性。封装单元30可在描述参数集轨道的TRAK框内的参数集轨道中用信号表示序列级SEI消息的存在。

[0097] MVEX框160可描述对应电影片段164的特性,例如,以发信号通知除了包括在MOOV框154内的视频数据(如果存在的话)之外,视频文件150还包括电影片段164。在流式传输视频数据的上下文中,译码视频图片可包括在电影片段164中而非包括在MOOV框154中。因此,所有译码视频样本可包括在电影片段164中,而非包括在MOOV框154中。

[0098] MOOV框154可包括等于视频文件150中的电影片段164的数目的多个MVEX框160。MVEX框160中的每一者可描述电影片段164中的对应一者的特性。例如,每一MVEX框可包括电影扩展标头框(MEHD)框,其描述电影片段164中的对应一者的持续时间。

[0099] 如上文所提及,封装单元30可将序列数据集存储在不包括实际译码视频数据的视频样本中。视频样本可大体上对应于接入单元,其为特定时间实例处的译码图片的表示。在AVC的上下文中,译码图片包括一个或多个VCL NAL单元,其含有用以构造接入单元及其它相关联非VCL NAL单元的所有像素的信息,例如SEI消息。因此,封装单元30可在电影片段164中的一者中包括序列数据集,序列数据集可包括序列级SEI消息。封装单元30可进一步将序列数据集和/或序列级SEI消息的存在发信号通知为存在于对应于电影片段164中的一者的MVEX框160中的一者内的电影片段164中的一者中。

[0100] SIDX框162为视频文件150的可选元素。即,符合3GPP文件格式或其它此类文件格式

式的视频文件未必包括SIDX框162。根据3GPP文件格式的示例, SIDX框可用于识别区段(例如, 视频文件150内含有的区段)的子区段。3GPP文件格式将子区段定义为“具有对应媒体数据框的一个或多个连续电影片段框的自含式集合, 且含有由电影片段框参考的数据的媒体数据框必须在该电影片段框之后且在含有关于同一轨道的信息的下一电影片段框之前”。3GPP文件格式还指示SIDX框“含有对由框记录的(子)区段的子区段的参考的序列。所引用的子区段在呈现时间上是连续的。类似地, 由区段索引框引用的字节在区段内总是连续的。所引用的大小给出了所引用的材料中的字节数的计数。”

[0101] SIDX框162通常提供表示包括在视频文件150中的区段的一个或多个子区段的信息。例如, 此信息可包括子区段开始和/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)流接入点(SAP)、SAP的类型(例如, SAP是否为瞬时解码器刷新(IDR)图片、清洁随机接入(CRA)图片、断链接入(BLA)图片等)、SAP在子区段中的位置(就回放时间和/或字节偏移来说)等。

[0102] 电影片段164可包括一个或多个译码视频图片。在一些示例中, 电影片段164可包括一个或多个图片群组(GOP), 其中的每一者可包括数个译码视频图片, 例如帧或图片。另外, 如上文所描述, 在一些示例中, 电影片段164可包括序列数据集。电影片段164中的每一者可包括电影片段标头框(MFHD, 图4中未示出)。MFHD框可描述对应电影片段的特性, 例如电影片段的序列号。电影片段164可按序列号的次序包括在视频文件150中。

[0103] MFRA框166可描述视频文件150的电影片段164内的随机接入点。这可以辅助执行特技模式, 例如执行对由视频文件150封装的区段内的特定时间位置(即, 回放时间)的搜寻。在一些示例中, MFRA框166通常为可选的且无需包括在视频文件中。同样地, 客户端设备(例如客户端设备40)未必需要参考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可包括数个轨道区段随机接入(TFRA)框(未示出), 其等于视频文件150的轨道的数目, 或在一些示例中, 等于视频文件150的媒体轨道(例如, 非提示轨道)的数目。

[0104] 在一些示例中, 电影片段164可包括一个或多个流接入点(SAP), 例如IDR图片。同样, MFRA框166可提供SAP的视频文件150内的位置的指示。因此, 视频文件150的时间子序列可由视频文件150的SAP形成。时间子序列还可包括其它图片, 例如取决于SAP的P帧和/或B帧。时间子序列的帧和/或切片可以布置在区段内, 使得可以适当地解码取决于子序列的其它帧/切片的时间子序列的帧/切片。例如, 在数据的级别式布置中, 用于其它数据的预测的数据也可包括在时间子序列中。

[0105] 图5是示出根据本公开的技术生成替换帧的示例的概念图。在图5的示例中, 源设备190经由网络192发送媒体数据180。媒体数据180包括帧182A到182D(帧182)。响应于切换时段186的检测或指示, 替换数据生成器194生成替换帧184A-184C(替换帧184)。

[0106] 替换数据生成器194可包括用于生成替换帧184的一个或多个AI/ML模型。特别地, 替换数据生成器194可以接收帧182A、182B中的至少一个, 并将帧182A、182B中的至少一个应用于AI/ML模型, 以使用帧182A、182B中的至少一个生成替换帧184中的一个或多个。

[0107] 在一些示例中, 替换数据生成器194可以包括在客户端设备40(图1)内。因此, 如果客户端设备40经历网络中断, 则客户端设备40可以生成替换帧184。在一些示例中, 替换数据生成器194可以被包括在网络74(图1)或网络192内的中间网络设备(诸如网络设备(诸如路由器)的服务设备或服务卡)中。因此, 如果源设备190经历网络中断, 则中间网络设备可

以生成替换帧184并对帧进行编码,然后将编码的替换帧发送到客户端设备40。

[0108] 如上所述,在一些示例中,客户端设备40可以向源设备190发送指示客户端设备40已经经历网络中断的数据。作为响应,在一些示例中,源设备190可以生成帧182C作为强制I帧。也就是说,如果帧182C最初将被编码为P帧或双向预测帧(B帧),则源设备190可以替代地将帧182C编码为I帧。在一些示例中,响应于网络中断,源设备190可以仅使用帧182A、182B中的一个或两个作为参考帧来对帧182C编码。也就是说,如果帧182B和182C之间的一个或多个帧可以另外被用作参考帧,则源设备190可以避免使用这些帧作为参考帧,而是仅使用帧182A、182B中的一个或两个作为用于帧间预测帧182C的参考帧。在一些示例中,响应于网络中断,源设备190还可以包括其自己的替换数据生成器,其将相同的AI/ML模型应用于帧182A、182B,并且还可以生成替换帧184以用作参考帧,使得可以从替换帧184预测帧182C、182D。

[0109] 图6是示出根据本公开的技术的生成和呈现替换媒体数据的示例方法的流程图。在图6的示例中,最初,服务器设备(诸如图1的服务器设备60)检测例如网络74的网络中断(200)。例如,服务器设备60可以确定它正在经历切换(handover),例如无线电接入网络的基站之间的转换(switch)。作为响应,服务器设备60生成自动驾驶模式指示消息(202)。自动驾驶模式指示消息可以包括指示自动驾驶模式的开始时间和结束时间的数据。在一些示例中,自动驾驶模式指示消息还可以包括指示客户端设备(诸如客户端设备40(图1))提示用户同意进入自动驾驶模式的数据。如上所述,自动驾驶模式通常可以使得客户端设备40生成用于否则将在指示的开始时间和指示的结束时间之间接收的媒体数据的替换媒体数据。

[0110] 服务器设备60可以确定开始时间是例如在无线电接入网络的基站之间的网络切换的开始的时间。服务器设备60可以将结束时间估计为完成网络切换的典型时间,例如50ms。因此,开始时间可以是切换过程开始的时间T,并且结束时间可以是T+S,其中S是完成网络切换的典型时间。如果S是50ms,则开始时间可以是T,并且结束时间可以是T+50ms。

[0111] 然后,服务器设备60可以向客户端设备40发送自动驾驶模式指示消息(204)。客户端设备40可以从服务器设备60接收自动驾驶模式指示消息(206)。尽管图6中未示出,但是假设服务器设备60已经向客户端设备40发送媒体数据作为媒体流会话的一部分。在该示例中,响应于接收到自动驾驶模式指示消息,客户端设备40进入自动驾驶模式(208),例如,以使用在自动驾驶模式指示消息中指示的时间T开始的先前接收的媒体数据来生成替换媒体数据(例如,视频数据)。客户端设备40然后可以例如使用一个或多个AI/ML模型和先前接收的媒体数据来生成替换媒体数据(210)。客户端设备40可以呈现替换媒体数据(212)。客户端设备40可以继续生成和呈现替换媒体数据,直到达到自动驾驶模式指示消息中指示的结束时间。在自动驾驶模式结束之后,客户端设备40可恢复正常操作,即,从服务器设备60接收、解码和呈现比特流的媒体数据。

[0112] 以此方式,图6的方法表示呈现媒体数据的方法的示例,该方法包括:接收媒体比特流的第一媒体数据集合;响应于确定将不会接收到媒体比特流的在第一媒体数据集合之后的第二媒体数据集合,使用第一媒体数据集合预测性地生成用于第二媒体数据集合的替换媒体数据;以及呈现第一媒体数据集合和替换媒体数据。

[0113] 图7是示出根据本公开的技术的生成和呈现替换媒体数据的另一示例方法的流程

图。在该示例中,客户端设备40检测网络中断(220)。例如,客户端设备40可以确定它正在经历从一个基站到另一个基站的切换。作为响应,客户端设备40可以生成自动驾驶模式指示消息(222)并将自动驾驶模式指示消息发送到服务器设备60(224)。再次,自动驾驶模式指示消息可以包括指示自动驾驶模式的开始时间和结束时间的数据。因此,服务器设备60可以接收自动驾驶模式指示消息(226)并在自动驾驶时段期间停止发送媒体数据(228),其中自动驾驶时段对应于自动驾驶模式的开始时间和结束时间之间的时间。

[0114] 在此时间期间,客户端设备40可以进入自动驾驶模式(230)。当处于自动驾驶模式时,客户端设备40可以例如使用一个或多个AI/ML模型和先前从服务器设备60接收的媒体数据来生成替换媒体数据(232)。同样,客户端设备40可以呈现替换媒体数据(234)。在自动驾驶时段结束之后,服务器设备60可以恢复发送媒体数据(236)。客户端设备40可以接收所发送的媒体数据(238)并呈现所接收的媒体数据(238)。

[0115] 以此方式,图7的方法表示呈现媒体数据的方法的示例,该方法包括:接收媒体比特流的第一媒体数据集合;响应于确定将不会接收到媒体比特流的在第一媒体数据集合之后的第二媒体数据集合,使用第一媒体数据集合预测性地生成用于第二媒体数据集合的替换媒体数据;以及呈现第一媒体数据集合和替换媒体数据。

[0116] 图8是示出根据本公开的技术的生成和呈现媒体数据的示例方法的流程图。最初,客户端设备40可以从服务器设备60接收用于媒体会话的第一媒体数据集合(250)。客户端设备40可以经由单独的信道(诸如信令信道和数据通信信道)与服务器设备60通信。客户端设备40和服务器设备60可以经由信令信道交换低带宽信令消息,并且经由数据通信信道交换较高带宽媒体数据。客户端设备40可以解码并呈现第一媒体数据集合(252)。

[0117] 客户端设备40然后可以确定将不接收媒体会话的第二媒体数据集合(254)。例如,客户端设备40可以确定它正在经历从一个基站到另一个基站的切换。作为另一示例,客户端设备40可以从服务器设备60(例如,经由信令信道)接收指示服务器设备60正在经历从一个基站到另一个基站的切换的消息。在一些示例中,客户端设备40或服务器设备60可以确定信号丢失或信号强度降低可能是由于客户端设备40、服务器设备60或两者的地理移动的位置和/或方向。例如,某些地理区域历史上可能导致信号丢失,诸如如果客户端设备40或服务器设备60中的任一个在车辆中移动通过隧道、通过具有低无线覆盖的区域、在具有高信号干扰的区域中等。

[0118] 客户端设备40然后可以进入自动驾驶模式并且预测性地生成用于第二媒体数据集合的替换媒体数据(256)。例如,客户端设备40可以将一个或多个AI/ML模型应用于第一媒体数据集合的至少一部分以生成替换媒体数据。客户端设备40还可以呈现替换媒体数据(258)。

[0119] 随后,客户端设备40可以开始从服务器设备60接收用于媒体会话的第三媒体数据集合(260)。因此,客户端设备40可以停止生成替换媒体数据,而是呈现所接收的第三媒体数据集合(282)。

[0120] 以此方式,图8的方法表示呈现媒体数据的方法的示例,该方法包括:接收媒体比特流的第一媒体数据集合;响应于确定将不会接收到媒体比特流的在第一媒体数据集合之后的第二媒体数据集合,使用第一媒体数据集合预测性地生成用于第二媒体数据集合的替换媒体数据;以及呈现第一媒体数据集合和替换媒体数据。



[0121] 图9是示出包括可被配置为执行本公开的技术的设备的示例系统280的框图。具体地,系统280包括用户装备 (UE) 设备284、292,基站282A、282B、290A、290B (基站282、290),以及由中间网络设备288形成的网络。

[0122] 中间网络设备288通常可以是构成基于计算机的网络的各种不同网络设备中的任何一种。中间网络设备288可以例如包括路由器、交换机、集线器、网关、服务设备、服务器、网桥、控制器等。基站282、290表示用于无线电接入网络 (RAN) (例如,5G RAN) 的小区的接入点的示例。例如,基站282、290可以是相应的gNodeB。UE设备284、292表示相应的UE设备,诸如蜂窝电话、平板计算机、移动设备等。

[0123] UE设备284、292可以包括用于发送和接收媒体数据的组件,所述媒体数据包括用于媒体或AR电话会议的音频、视频和/或AR/XR/MR/VR数据。例如,UE设备284、292可以包括用于生成AR/XR/MR/VR数据的相机、麦克风和/或图形处理单元 (GPU)。

[0124] 根据本公开的技术,UE设备284和UE设备292可以参与媒体电话会议会话,诸如视频电话会议会话和/或AR/XR/MR/VR电话会议会话。因此,UE设备284、292可以发送和接收媒体数据。当UE设备284捕获或生成媒体数据时,UE设备284可以对捕获/生成的媒体数据进行编码并将其发送给UE设备292。同样地,当UE设备292捕获或生成媒体数据时,UE设备292可以对捕获/生成的媒体数据进行编码并将其发送给UE设备284。

[0125] 根据本公开的技术,UE设备284可以从包括基站282B的小区移动到包括基站282A的小区。因此,可以发生从基站282B到基站282A的切换。作为响应,UE设备284可以例如经由信令信道向UE设备292发送自动驾驶模式指示消息。自动驾驶模式指示消息可以表示对于UE设备292当前正在接收的媒体数据,UE设备292应当生成替换媒体数据,并且对于UE设备292当前正在发送的媒体数据,UE设备292应当停止发送媒体数据,直到自动驾驶模式到期。UE设备284可以构造自动驾驶模式指示消息以指示自动驾驶时段的开始时间和结束时间。

[0126] 在一些示例中,中间网络设备288中的一个或多个可以是配置为检测这种自动驾驶模式指示消息的服务设备。中间网络设备288的服务设备可以代表UE设备292生成替换媒体数据并对替换媒体数据进行编码,使得UE设备292不需要改变其解码和呈现过程。

[0127] 以这种方式,UE设备284、292表示用于呈现媒体数据的设备的示例,包括被配置为存储媒体数据的存储器;以及一个或多个处理器,其实施于电路中且被配置为:接收媒体比特流的第一媒体数据集合;响应于确定将不会接收到媒体比特流的在第一媒体数据集合之后的第二媒体数据集合,使用第一媒体数据集合预测性地生成用于第二媒体数据集合的替换媒体数据;以及呈现第一媒体数据集合和替换媒体数据。

[0128] 图10是示出包括用户装备 (UE) 设备300和伴随设备302的示例系统的框图。UE设备300和伴随设备302参与通信会话304。通信会话304可以通过诸如无线广域网 (WWAN)、无线局域网 (WLAN)、WiFi、蓝牙等的无线通信协议来进行。UE设备300可以表示任何用户装备,而伴随设备302可以表示由同一用户结合UE设备300使用的近接近度设备。例如,UE设备300可以是蜂窝电话,而伴随设备302是无线显示器,诸如监视器、电视或AR/XR/MR/VR头戴式耳机。作为另一示例,UE设备300可以是平板计算机或膝上型计算机,而伴随设备302可以是蜂窝电话,或者反之亦然。

[0129] 尽管主要相对于通过网络彼此通信的两个UE设备的场景进行描述,但是本公开的技术也可由UE设备300和伴随设备302执行。例如,如果UE设备300例如经由与另一UE设备



(未示出)的媒体通信会话接收媒体数据,则UE设备300可经由伴随设备302无线地呈现媒体通信会话的媒体数据。如果通信会话304被短暂地中断,例如,因为UE设备300和伴随设备302在物理上彼此相距太远,物理对象阻碍通信会话,或者由于其他这样的原因,伴随设备302可以被配置为检测中断并应用本公开的技术来生成替换媒体数据,直到通信会话304恢复。

[0130] 在一些示例中,UE设备300可以确定与另一UE设备(未示出)的通信会话已经被中断,如上面讨论的各种示例中那样,但是将生成替换媒体数据的过程转移(offload)到伴随设备302。例如,伴随设备302可以包括一个或多个AI/ML模型和AI/ML过程。UE设备300可以向伴随设备302发送接收到的媒体数据和对替换媒体数据的请求。作为响应,伴随设备302可以通过将AI/ML模型应用于接收到的媒体数据来生成替换媒体数据,并经由通信会话304将替换媒体数据发送到UE设备300。

[0131] 以下条款表示本公开的技术的某些示例:

[0132] 条款1:一种呈现媒体数据的方法,所述方法包括:接收媒体比特流的第一媒体数据集合;响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0133] 条款2:根据条款1所述的方法,其中接收所述第一媒体数据集合包括经由网络接收所述第一媒体数据集合,并且其中确定将不接收所述第二媒体数据集合包括确定网络中断将阻止接收所述第二媒体数据集合。

[0134] 条款3:根据条款2所述的方法,还包括确定网络中断的开始时间和网络中断的结束时间,其中预测性地生成替换数据包括预测性地生成在网络中断的开始时间和网络中断的结束时间之间的时间段内的替换媒体数据。

[0135] 条款4:根据条款1所述的方法,其中接收所述第一媒体数据集合包括经由网络从源设备接收所述第一媒体数据集合,并且其中确定将不接收所述第二媒体数据集合包括从所述源设备接收指示将不接收所述第二媒体数据集合的数据。

[0136] 条款5:根据条款4所述的方法,其中从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据包括从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者。

[0137] 条款6:根据条款4所述的方法,其中接收第一媒体数据集合包括经由网络的媒体数据信道从源设备接收第一媒体数据集合,并且其中从源设备接收指示将不接收第二媒体数据集合的数据包括经由网络的信令信道从源设备接收指示将不接收第二媒体数据集合的数据,所述信令信道与所述媒体数据信道分离。

[0138] 条款7:根据条款1所述的方法,还包括:在将不接收媒体数据的情况下,向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及从所述用户接收输入数据,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

[0139] 条款8:根据条款1所述的方法,其中所述媒体数据包括视频数据或扩展现实(XR)数据中的至少一个,所述XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一个。

[0140] 条款9:根据条款1所述的方法,还包括:接收所述媒体比特流的第三媒体数据集

合;将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

[0141] 条款10:根据条款1所述的方法,其中预测性地生成替换媒体数据包括将第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能(AI)过程或机器学习(ML)过程。

[0142] 条款11:一种用于呈现媒体数据的设备,所述设备包括:存储器,被配置为存储媒体数据;以及一个或多个处理器,实施于电路中且被配置为:接收媒体比特流的第一媒体数据集合;响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0143] 条款12:根据条款11所述的设备,其中,所述一个或多个处理器被配置为经由网络接收所述第一媒体数据集合,并且其中,为了确定将不接收所述第二媒体数据集合,所述一个或多个网络设备被配置为确定网络中断将阻止接收所述第二媒体数据集合。

[0144] 条款13:根据条款12所述的设备,其中一个或多个处理器还被配置为确定网络中断的开始时间和网络中断的结束时间,其中预测性地生成所述替换数据包括预测性地生成在网络中断的开始时间和网络中断的结束时间之间的时间段内的所述替换媒体数据。

[0145] 条款14:根据条款11所述的设备,其中,所述一个或多个处理器被配置为经由网络从源设备接收所述第一媒体数据集合,并且其中,为了确定将不接收所述第二媒体数据集合,所述一个或多个处理器被配置为从所述源设备接收指示将不接收所述第二媒体数据集合的数据。

[0146] 条款15:根据条款14所述的设备,其中为了从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据,所述一个或多个处理器被配置为从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者。

[0147] 条款16:根据条款14所述的设备,其中,所述一个或多个处理器被配置为经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合,并且其中,所述一个或多个处理器被配置为经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的数据,所述信令信道与所述媒体数据信道分离。

[0148] 条款17:根据条款11所述的设备,还包括显示器,其中,所述一个或多个处理器还被配置为:在将不接收媒体数据的情况下,经由所述显示器向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及从用户接收指示用户同意针对所述媒体比特流预测性地生成的替换媒体数据的输入数据。

[0149] 条款18:根据条款11所述的设备,其中,所述媒体数据包括视频数据或扩展现实(XR)数据中的至少一个,XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一个。

[0150] 条款19:根据条款11所述的设备,其中所述一个或多个处理器还被配置为:接收所述媒体比特流的第三媒体数据集合;将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

[0151] 条款20:根据条款11所述的设备,其中为了预测性地生成所述替换媒体数据,所述

一个或多个处理器被配置为将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程。

[0152] 条款21:一种用于呈现媒体数据的设备,所述设备包括:用于接收媒体比特流的第一媒体数据集合的装置;用于响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合针对所述第二媒体数据集合预测性地生成替换媒体数据的装置;以及用于呈现所述第一媒体数据集合和所述替换媒体数据的装置。

[0153] 条款22:根据条款21所述的设备,其中用于接收所述第一媒体数据集合的装置包括用于经由网络接收所述第一媒体数据集合的装置,还包括用于确定网络中断将阻止接收所述第二媒体数据集合的装置。

[0154] 条款23:根据条款22所述的设备,还包括用于确定网络中断的开始时间和网络中断的结束时间的装置,其中用于预测性地生成所述替换数据的装置包括用于预测性地生成在网络中断的开始时间和网络中断的结束时间之间的时间段内的所述替换媒体数据的装置。

[0155] 条款24:根据条款21所述的设备,其中用于接收所述第一媒体数据集合的装置包括用于经由网络从源设备接收所述第一媒体数据集合的装置,并且其中用于确定将不接收所述第二媒体数据集合的装置包括用于从所述源设备接收指示将不接收所述第二媒体数据集合的数据的装置。

[0156] 条款25:根据条款24所述的设备,其中用于从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的装置包括用于从所述源设备接收补充增强信息 (SEI) 消息或实时传输协议 (RTP) 消息中的至少一者的装置。

[0157] 条款26:根据条款24所述的设备,其中用于接收第一媒体数据集合的装置包括用于经由网络的媒体数据信道从源设备接收第一媒体数据集合的装置,并且其中用于从源设备接收指示将不接收第二媒体数据集合的数据的装置包括用于经由网络的信令信道从源设备接收指示将不接收第二媒体数据集合的数据的装置,所述信令信道与所述媒体数据信道分离。

[0158] 条款27:根据条款21所述的设备,还包括:用于在将不接收媒体数据的情况下向用户呈现请求用户同意预测性地生成替换媒体数据的提示的装置;以及用于从所述用户接收输入数据的装置,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

[0159] 条款28:根据条款21所述的设备,其中所述媒体数据包括视频数据或扩展现实 (XR) 数据中的至少一者,所述XR数据包括增强现实 (AR) 数据、虚拟现实 (VR) 数据或混合现实 (MR) 数据中的至少一者。

[0160] 条款29:根据条款21所述的设备,还包括:用于接收媒体比特流的第三媒体数据集合的装置;用于将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合的装置;以及用于在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分的装置。

[0161] 条款30:根据条款21所述的设备,其中用于预测性地生成替换媒体数据的装置包括用于将第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工

智能(AI)过程或机器学习(ML)过程的装置。

[0162] 条款31:一种其上存储有指令的计算机可读存储介质,所述指令在被执行时使得处理器:接收媒体比特流的第一媒体数据集合;响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0163] 条款32:根据条款31所述的计算机可读存储介质,其中,使得所述处理器接收所述第一媒体数据集合的指令包括使得所述处理器经由网络接收所述第一媒体数据集合的指令,还包括使得所述处理器确定网络中断将阻止接收所述第二媒体数据集合的指令。

[0164] 条款33:根据条款32所述的计算机可读存储介质,还包括使得所述处理器确定网络中断的开始时间和网络中断的结束时间的指令,其中使得所述处理器预测性地生成所述替换数据的指令包括使得所述处理器预测性地生成在网络中断的开始时间和网络中断的结束时间之间的时间段内的替换媒体数据的指令。

[0165] 条款34:根据条款31所述的计算机可读存储介质,其中使得所述处理器接收所述第一媒体数据集合的所述指令包括使得所述处理器经由网络从源设备接收所述第一媒体数据集合的指令,还包括使得所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的数据的指令。

[0166] 条款35:根据条款34所述的计算机可读存储介质,其中使得所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的所述指令包括使得所述处理器从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者的指令。

[0167] 条款36:根据条款34所述的计算机可读存储介质,其中,使得所述处理器接收所述第一媒体数据集合的指令包括使得所述处理器经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合的指令,以及其中使得所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的所述指令包括使得所述处理器经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的指令,所述信令信道与所述媒体数据信道分离。

[0168] 条款37:根据条款31所述的计算机可读存储介质,还包括使得所述处理器执行以下操作的指令:在将不接收媒体数据的情况下,向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及从用户接收输入数据,所述输入数据指示用户同意针对媒体比特流预测性地生成的替换媒体数据。

[0169] 条款38:根据条款31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或扩展现实(XR)数据中的至少一者,所述XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一者。

[0170] 条款39:根据条款31所述的计算机可读存储介质,还包括使得所述处理器进行以下操作的指令:接收所述媒体比特流的第三媒体数据集合;将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及在替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

[0171] 条款40:根据条款31所述的计算机可读存储介质,其中,使得所述处理器预测性地

生成替换媒体数据的指令包括使得所述处理器将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程的指令。

[0172] 条款41:一种呈现媒体数据的方法,所述方法包括:接收媒体比特流的第一媒体数据集合;响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0173] 条款42:根据条款41所述的方法,其中接收所述第一媒体数据集合包括经由网络接收所述第一媒体数据集合,并且其中确定将不接收所述第二媒体数据集合包括确定网络中断将阻止接收所述第二媒体数据集合。

[0174] 条款43:根据条款42所述的方法,还包括确定所述网络中断的开始时间和所述网络中断的结束时间,其中预测性地生成所述替换数据包括预测性地生成在所述网络中断的开始时间和所述网络中断的结束时间之间的时间段内的所述替换媒体数据。

[0175] 条款44:根据条款41所述的方法,其中接收所述第一媒体数据集合包括经由网络从源设备接收所述第一媒体数据集合,并且其中确定将不接收所述第二媒体数据集合包括从所述源设备接收指示将不接收所述第二媒体数据集合的数据。

[0176] 条款45:根据条款44所述的方法,其中从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据包括从所述源设备接收补充增强信息 (SEI) 消息或实时传输协议 (RTP) 消息中的至少一者。

[0177] 条款46:根据条款44和45中任一项所述的方法,其中,接收所述第一媒体数据集合包括经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合,并且其中,从所述源设备接收指示将不接收所述第二媒体数据集合的数据包括经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的数据,所述信令信道与所述媒体数据信道分离。

[0178] 条款47:根据条款41-46中任一项所述的方法,还包括:在不接收媒体数据的情况下,向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及从所述用户接收输入数据,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

[0179] 条款48:根据条款41-47中任一项所述的方法,其中所述媒体数据包括视频数据或扩展现实 (XR) 数据中的至少一个,所述XR数据包括增强现实 (AR) 数据、虚拟现实 (VR) 数据或混合现实 (MR) 数据中的至少一个。

[0180] 条款49:根据条款41-48中任一项所述的方法,还包括:接收所述媒体比特流的第三媒体数据集合;将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

[0181] 条款50:根据条款41-49中任一项所述的方法,其中预测性地生成替换媒体数据包括将第一媒体数据集合的至少一部分提供给被配置为生成替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程。

[0182] 条款51:一种用于呈现媒体数据的设备,所述设备包括:存储器,被配置为存储媒体数据;以及一个或多个处理器,实施于电路中且被配置为:接收媒体比特流的第一媒体数据集合;响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体

数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0183] 条款52:根据条款51所述的设备,其中所述一个或多个处理器被配置为经由网络接收所述第一媒体数据集合,并且其中为了确定将不接收所述第二媒体数据集合,所述一个或多个网络设备被配置为确定网络中断将阻止接收所述第二媒体数据集合。

[0184] 条款53:根据条款52所述的设备,其中,所述一个或多个处理器还被配置为确定所述网络中断的开始时间和所述网络中断的结束时间,其中,预测性地生成所述替换数据包括预测性地生成在所述网络中断的开始时间和所述网络中断的结束时间之间的时间段内的所述替换媒体数据。

[0185] 条款54:根据条款51所述的设备,其中所述一个或多个处理器被配置为经由网络从源设备接收所述第一媒体数据集合,并且其中为了确定将不接收所述第二媒体数据集合,所述一个或多个处理器被配置为从所述源设备接收指示将不接收所述第二媒体数据集合的数据。

[0186] 条款55:根据条款54所述的设备,其中为了从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据,所述一个或多个处理器被配置为从所述源设备接收补充增强信息(SEI)消息或实时传输协议(RTP)消息中的至少一者。

[0187] 条款56:根据条款54和55中任一项所述的设备,其中,所述一个或多个处理器被配置为经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合,并且其中,所述一个或多个处理器被配置为经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的数据,所述信令信道与所述媒体数据信道分离。

[0188] 条款57:根据条款51-56中任一项所述的设备,还包括显示器,其中所述一个或多个处理器还被配置为:在不接收媒体数据的情况下,经由显示器向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及从用户接收输入数据,所述输入数据指示用户同意针对媒体比特流预测性地生成的替换媒体数据。

[0189] 条款58:根据条款51-57中任一项所述的设备,其中所述媒体数据包括视频数据或扩展现实(XR)数据中的至少一个,所述XR数据包括增强现实(AR)数据、虚拟现实(VR)数据或混合现实(MR)数据中的至少一个。

[0190] 条款59:根据条款51-58中任一项所述的设备,其中所述一个或多个处理器还被配置为:接收所述媒体比特流的第三媒体数据集合;将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及在所述替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分。

[0191] 条款60:根据条款51-59中任一项所述的设备,其中为了预测性地生成所述替换媒体数据,所述一个或多个处理器被配置为将所述第一媒体数据集合的至少一部分提供给被配置为生成所述替换媒体数据的人工智能(AI)过程或机器学习(ML)过程。

[0192] 条款61:一种用于呈现媒体数据的设备,所述设备包括:用于接收媒体比特流的第一媒体数据集合的装置;用于响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合针对所述第二媒体数据集合预测性地生成替换媒体数据的装置;以及用于呈现所述第一媒体数据集合和所述替换媒体数据的装置。

[0193] 条款62:根据条款61所述的设备,其中用于接收第一媒体数据集合的装置包括用于经由网络接收第一媒体数据集合的装置,还包括用于确定网络中断将阻止接收第二媒体数据集合的装置。

[0194] 条款63:根据条款62所述的设备,还包括用于确定网络中断的开始时间和网络中断的结束时间的装置,其中用于预测性地生成替换数据的装置包括用于预测性地生成在网络中断的开始时间和网络中断的结束时间之间的时间段内的替换媒体数据的装置。

[0195] 条款64:根据条款61所述的设备,其中用于接收所述第一媒体数据集合的装置包括用于经由网络从源设备接收所述第一媒体数据集合的装置,以及其中用于确定将不接收所述第二媒体数据集合的装置包括用于从所述源设备接收指示将不接收所述第二媒体数据集合的数据的装置。

[0196] 条款65:根据条款64所述的设备,其中用于从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的装置包括用于从所述源设备接收补充增强信息 (SEI) 消息或实时传输协议 (RTP) 消息中的至少一者的装置。

[0197] 条款66:根据条款64和65中任一项的设备,其中用于接收第一媒体数据集合的装置包括用于经由网络的媒体数据信道从源设备接收第一媒体数据集合的装置,并且其中用于从源设备接收指示将不接收第二媒体数据集合的数据的装置包括用于经由网络的信令信道从源设备接收指示将不接收第二媒体数据集合的数据的装置,所述信令信道与所述媒体数据信道分离。

[0198] 条款67:根据条款61-66中任一项所述的设备,还包括:用于在不接收媒体数据的情况下向用户呈现请求用户同意预测性地生成替换媒体数据的提示的装置;以及用于从所述用户接收输入数据的装置,所述输入数据指示所述用户同意针对所述媒体比特流预测性地生成的替换媒体数据。

[0199] 条款68:根据条款61-67中任一项所述的设备,其中所述媒体数据包括视频数据或扩展现实 (XR) 数据中的至少一个,所述XR数据包括增强现实 (AR) 数据、虚拟现实 (VR) 数据或混合现实 (MR) 数据中的至少一个。

[0200] 条款69:根据条款61-68中任一项所述的设备,还包括:用于接收所述媒体比特流的第三媒体数据集合的装置;用于将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合的装置;以及用于在替换媒体数据之后呈现所述第三媒体数据集合的混合的至少一部分的装置。

[0201] 条款70:根据条款61-69中任一项所述的设备,其中用于预测性地生成替换媒体数据的装置包括用于将第一媒体数据集合的至少一部分提供给被配置为生成替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程的装置。

[0202] 条款71:一种其上存储有指令的计算机可读存储介质,所述指令在被执行时使处理器:接收媒体比特流的第一媒体数据集合;响应于确定将不接收所述媒体比特流的在所述第一媒体数据集合之后的第二媒体数据集合,使用所述第一媒体数据集合预测性地生成用于所述第二媒体数据集合的替换媒体数据;以及呈现所述第一媒体数据集合和所述替换媒体数据。

[0203] 条款72:根据条款71所述的计算机可读存储介质,其中,使处理器接收第一媒体数据集合的指令包括使得处理器经由网络接收第一媒体数据集合的指令,还包括使处理器确



定网络中断将阻止接收第二媒体数据集合的指令。

[0204] 条款73:根据条款72所述的计算机可读存储介质,还包括使处理器确定网络中断的开始时间和网络中断的结束时间的指令,其中使处理器预测性地生成替换数据的指令包括使处理器预测性地生成在网络中断的开始时间和网络中断的结束时间之间的时间段内的替换媒体数据的指令。

[0205] 条款74:根据条款71所述的计算机可读存储介质,其中,使所述处理器接收所述第一媒体数据集合的指令包括使所述处理器经由网络从源设备接收所述第一媒体数据集合的指令,还包括使所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的数据的指令。

[0206] 条款75:根据条款74所述的计算机可读存储介质,其中使所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的所述指令包括使所述处理器从所述源设备接收补充增强信息 (SEI) 消息或实时传输协议 (RTP) 消息中的至少一者的指令。

[0207] 条款76:根据条款74和75中任一项所述的计算机可读存储介质,其中使所述处理器接收所述第一媒体数据集合的指令包括使所述处理器经由所述网络的媒体数据信道从所述源设备接收所述第一媒体数据集合的指令,以及其中使所述处理器从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的所述指令包括使所述处理器经由所述网络的信令信道从所述源设备接收指示将不接收所述第二媒体数据集合的所述数据的指令,所述信令信道与所述媒体数据信道分离。

[0208] 条款77:根据条款71-76中任一项所述的计算机可读存储介质,还包括使所述处理器执行以下操作的指令:在将不接收媒体数据的情况下,向用户呈现请求用户同意预测性地生成替换媒体数据的提示;以及从用户接收输入数据,所述输入数据指示用户同意针对媒体比特流预测性地生成的替换媒体数据。

[0209] 条款78:根据条款71-77中任一项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或扩展现实 (XR) 数据中的至少一个,所述XR数据包括增强现实 (AR) 数据、虚拟现实 (VR) 数据或混合现实 (MR) 数据中的至少一个。

[0210] 条款79:根据条款71-78中任一项所述的计算机可读存储介质,还包括使所述处理器执行以下操作的指令:接收所述媒体比特流的第三媒体数据集合;将所述第三媒体数据集合的至少一部分与所述替换媒体数据混合;以及在所述替换媒体数据之后呈现所述第三媒体数据集合的是混合的至少一部分。

[0211] 条款80:根据条款71-79中任一项所述的计算机可读存储介质,其中使所述处理器预测性地生成替换媒体数据的指令包括使所述处理器将第一媒体数据集合的至少一部分提供给被配置为生成替换媒体数据的人工智能 (AI) 过程或机器学习 (ML) 过程的指令。

[0212] 在一个或多个示例中,所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件实施,则功能可作为一个或多个指令或代码存储于计算机可读介质上或经由计算机可读介质发送,且由基于硬件的处理单元执行。计算机可读介质可包括计算机可读存储介质,其对应于例如数据存储介质的有形介质,或包括促进例如根据通信协议将计算机程序从一处传送到另一处的任何介质的通信介质。以此方式,计算机可读介质通常可对应于 (1) 非暂时性的有形计算机可读存储介质,或 (2) 通信介质,例如信号或载波。数据存储介质可为可由一个或多个计算机或一个或多个处理器存取以检索用于本公开中所描述的技术



术中的实施方式的指令、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] 借助于示例而非限制,此类计算机可读存储介质可包括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、快闪存储器或可用以存储呈指令或数据结构形式的期望程序代码且可由计算机存取的任何其它介质。此外,任何连接被适当地称为计算机可读介质。例如,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或无线技术(例如红外线、无线电和微波)从网站、服务器或其它远程源发送指令,则同轴电缆、光纤电缆、双绞线、DSL或无线技术(例如红外线、无线电和微波)包括在介质的定义中。然而,应理解,计算机可读存储介质和数据存储介质并不包括连接、载波、信号或其它暂时性介质,而是实际上针对非暂时性有形存储介质。如本文中所使用,磁盘和光盘包括压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软盘和蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘用激光以光学方式再现数据。上述的组合也应当被包括在计算机可读介质的范围内。

[0214] 指令可由一个或多个处理器执行,所述一个或多个处理器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)或其它等效集成或离散逻辑电路。因此,如本文中所使用,术语“处理器”可指代前述结构中的任一者或适用于本文中所描述的技术中的实施方式的任何其它结构。另外,在一些方面中,本文中所描述的功能性可提供于被配置用于编码和解码的专用硬件和/或软件模块内,或并入于组合式编解码器中。此外,这些技术可以在一个或多个电路或逻辑元件中完全实现。

[0215] 本公开的技术可实施于广泛多种设备或装置中,所述设备或装置包括无线手持机、集成电路(IC)或IC组(例如,芯片组)。本公开中描述各种组件、模块或单元以强调配置为执行所公开的技术的设备的功能方面,但未必需要由不同硬件单元实现。而是,如上文所描述,各种单元可结合合适的软件和/或固件在编解码器硬件单元中组合,或通过互操作硬件单元(包含如上文所描述的一个或多个处理器)的集合来提供。

[0216] 已经描述了各种示例。这些和其他示例在所附权利要求的范围内。

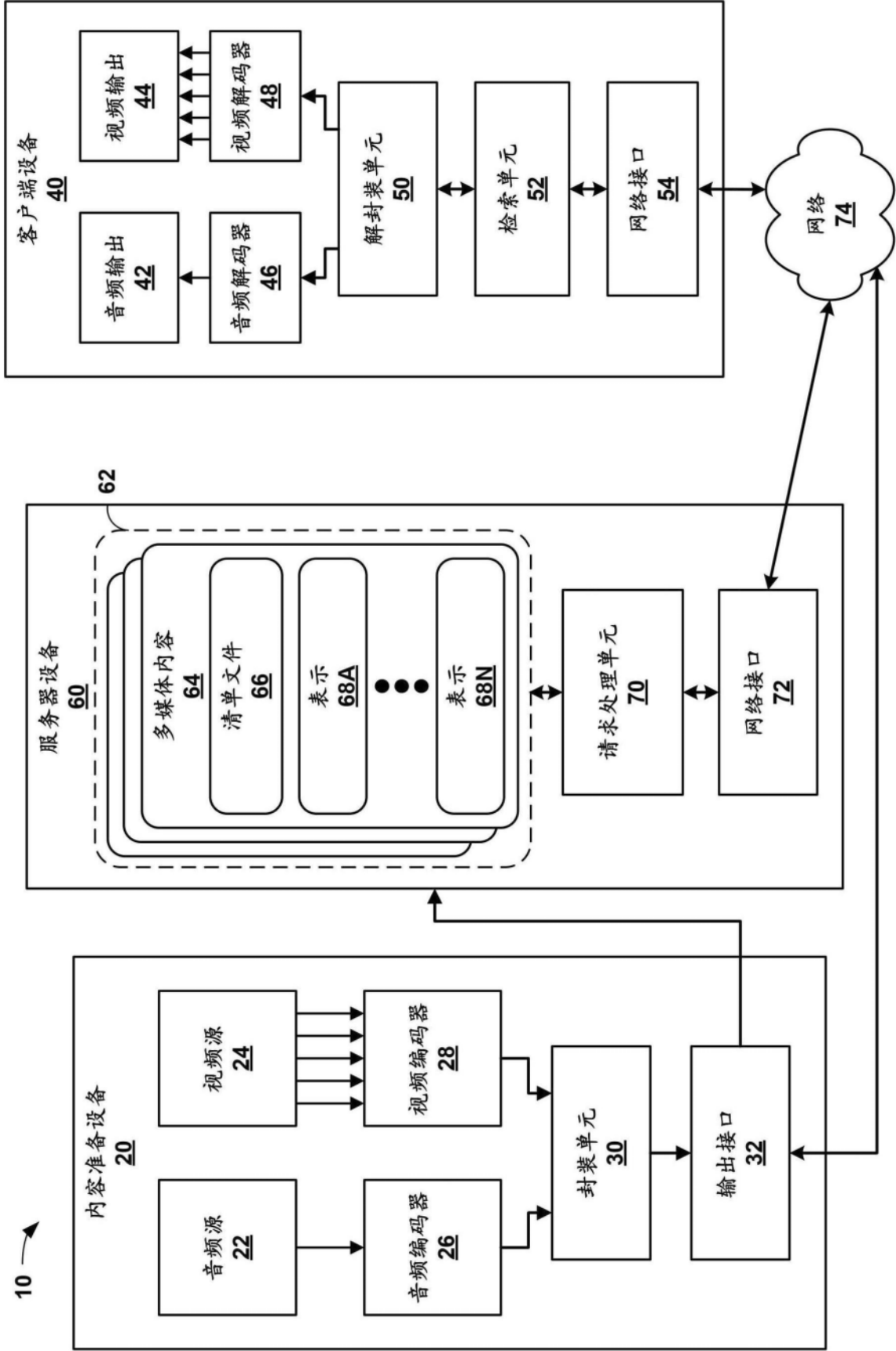


图1

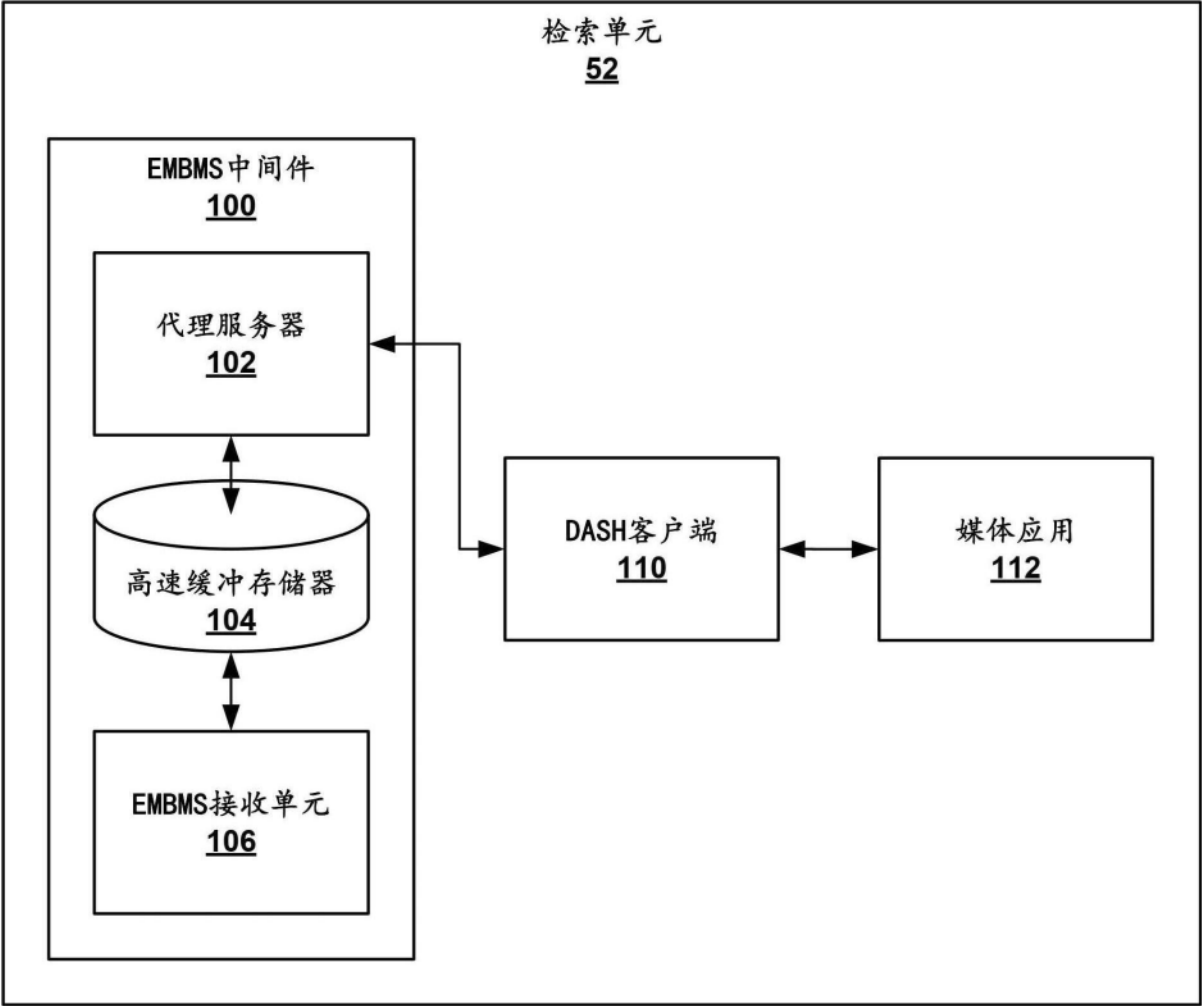


图2

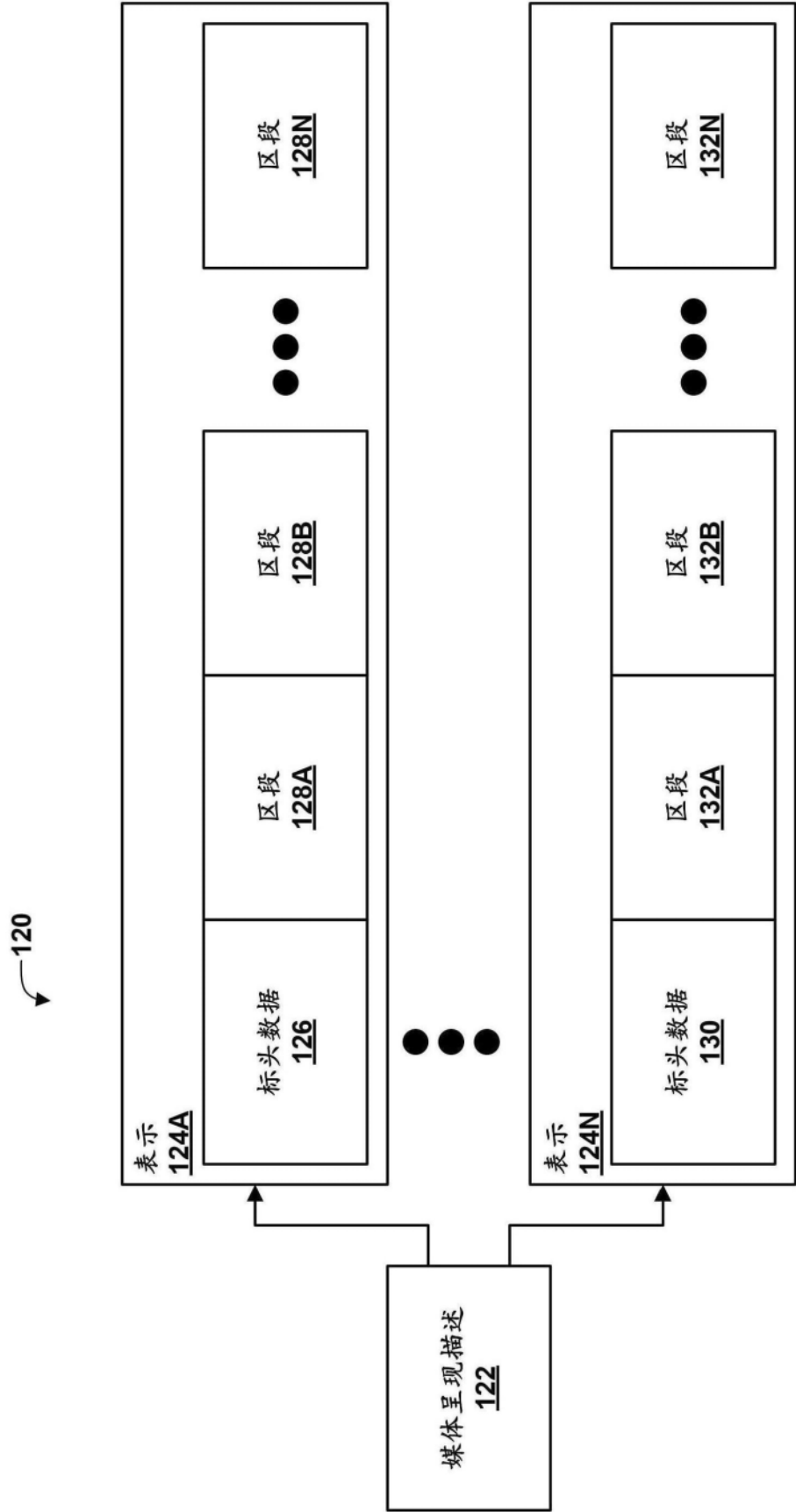


图3

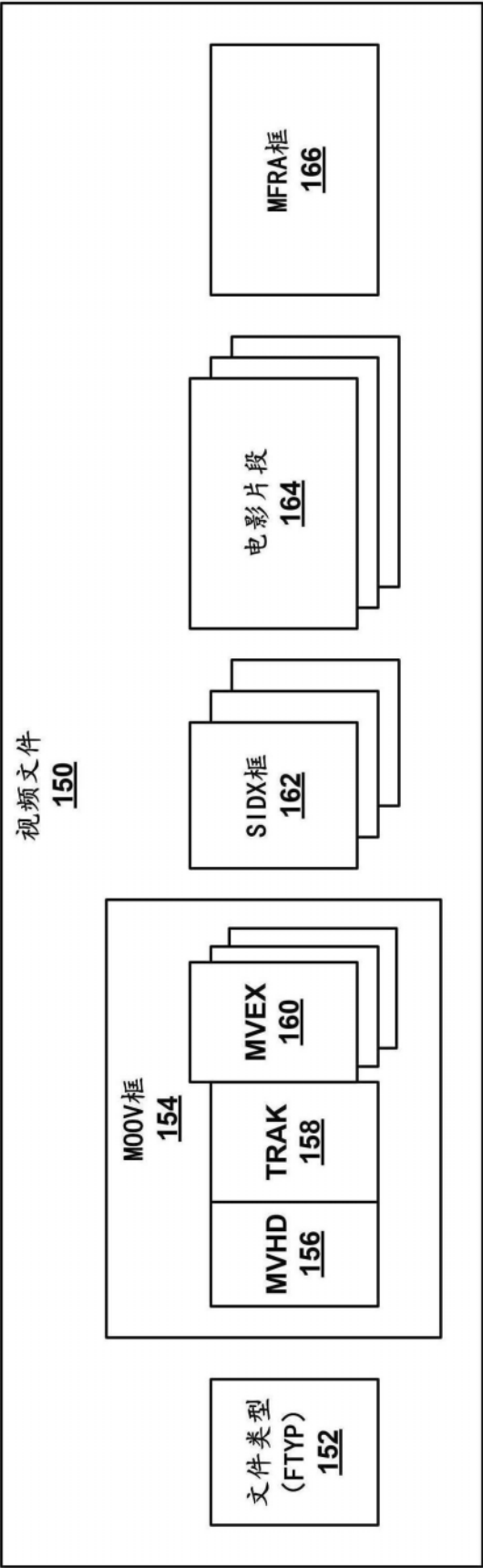


图4

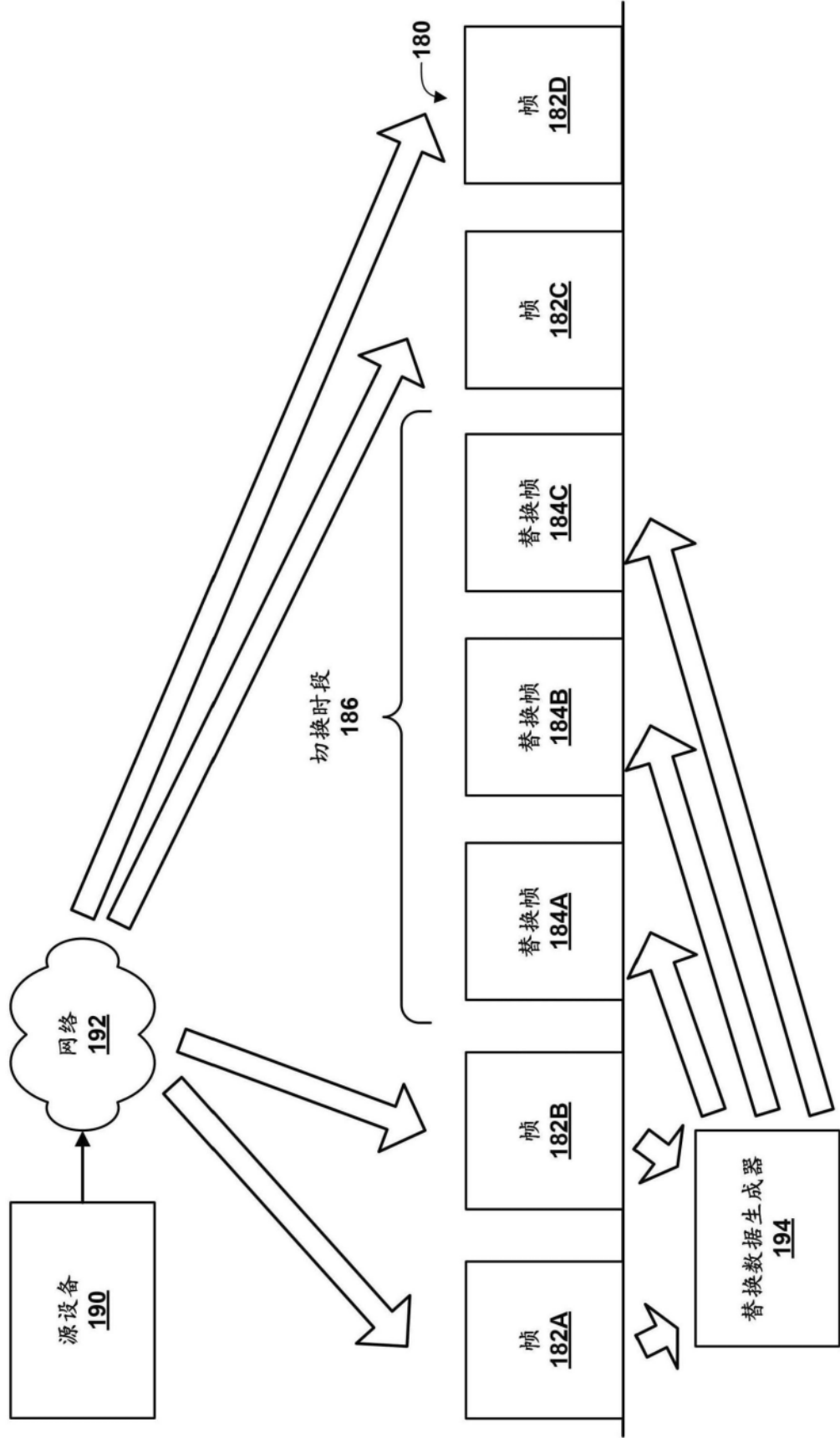


图5

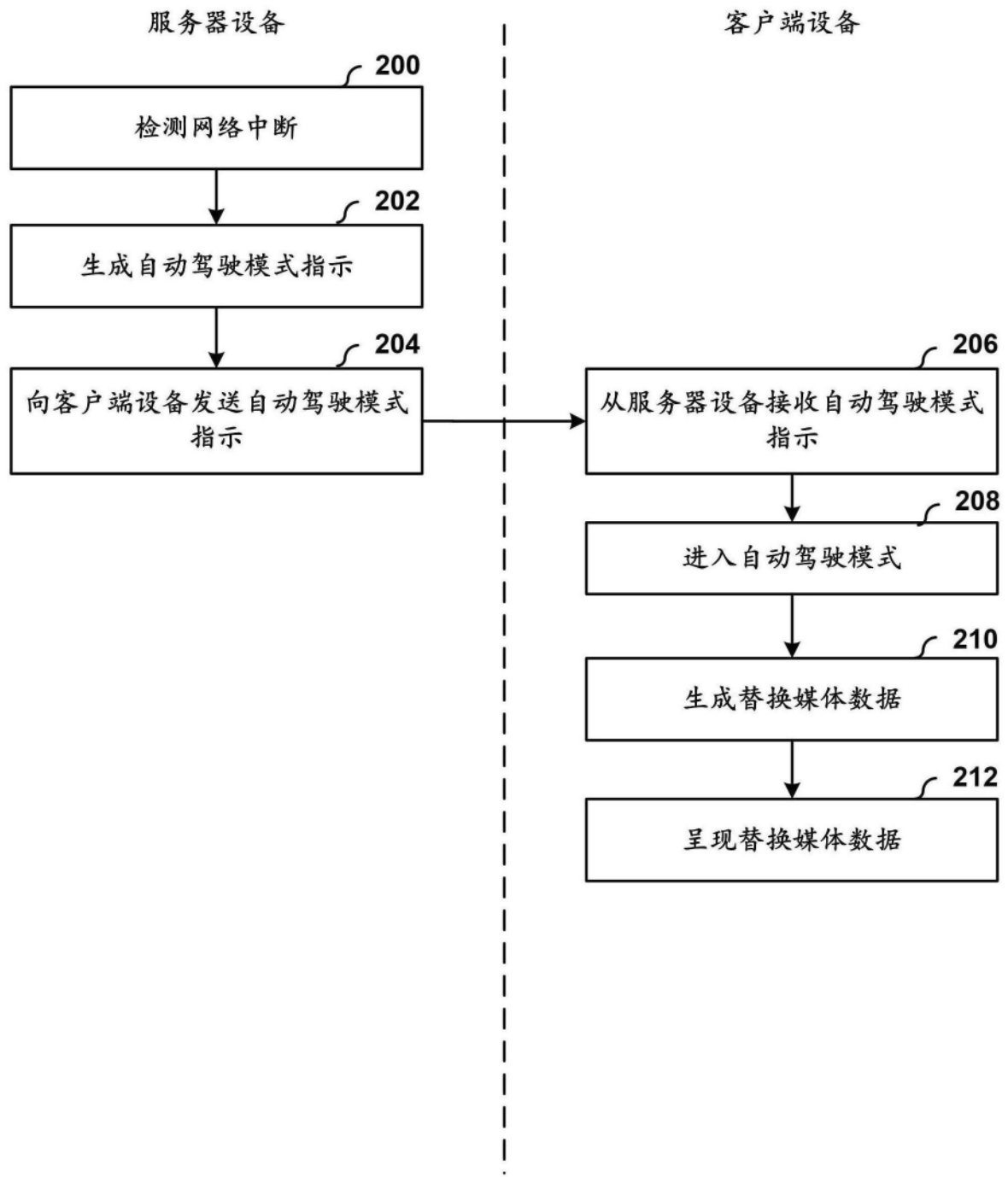


图6

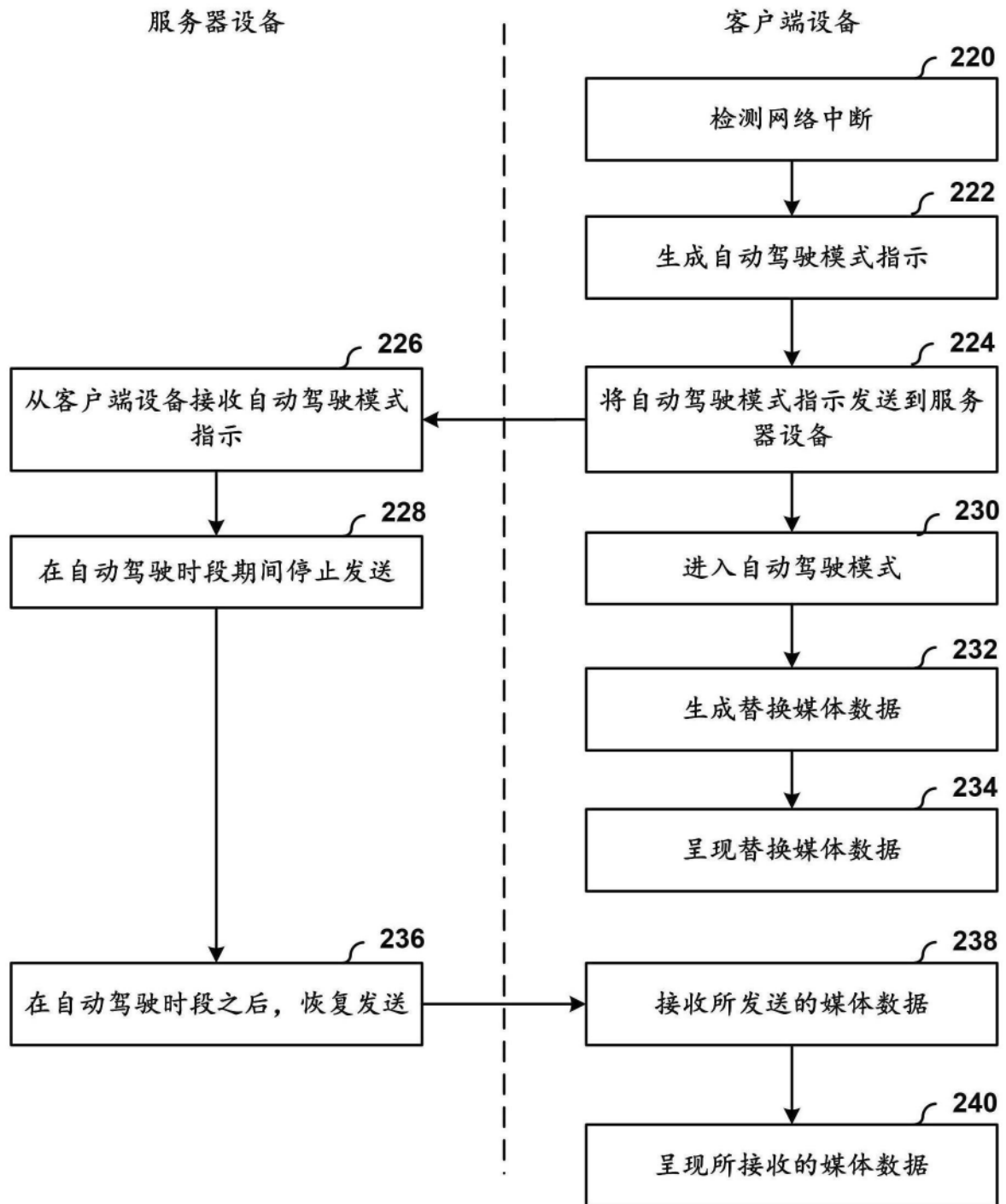


图7



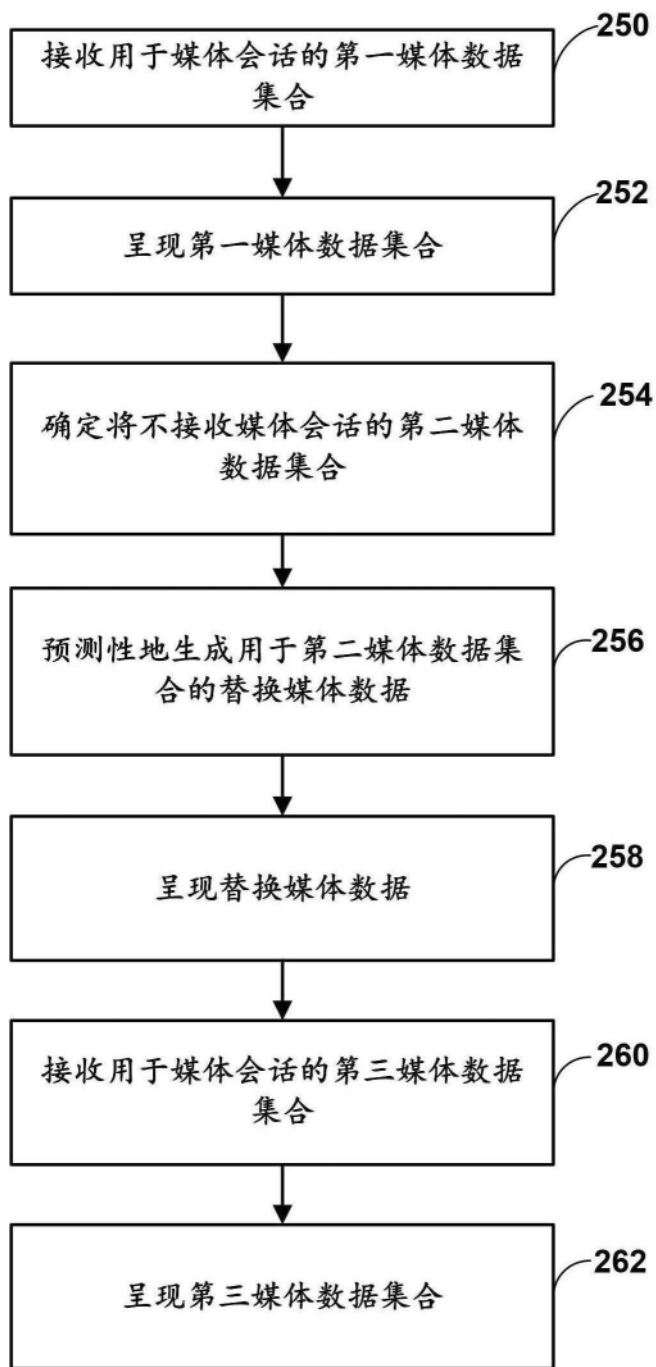


图8

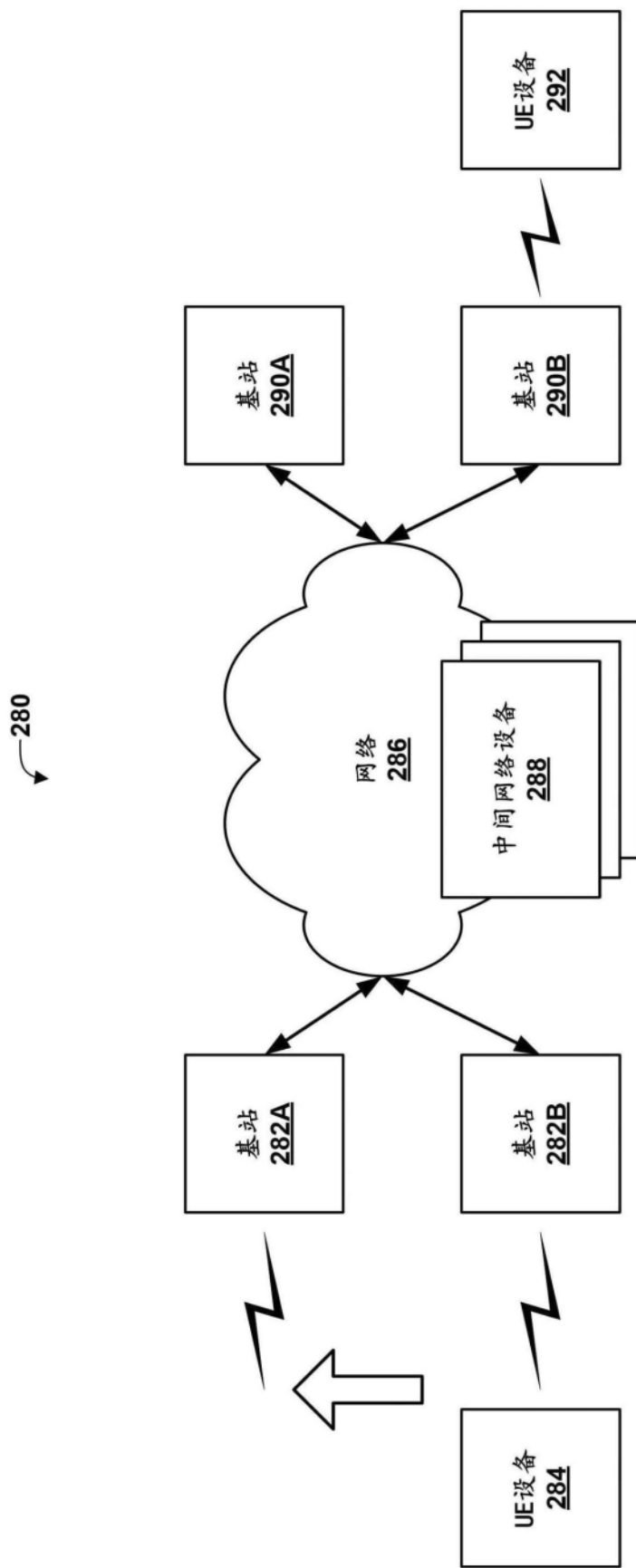


图9

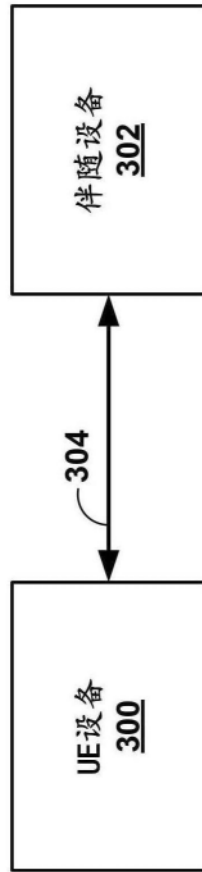


图10