(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 114430911 B
(45)授权公告日 2025.06.24

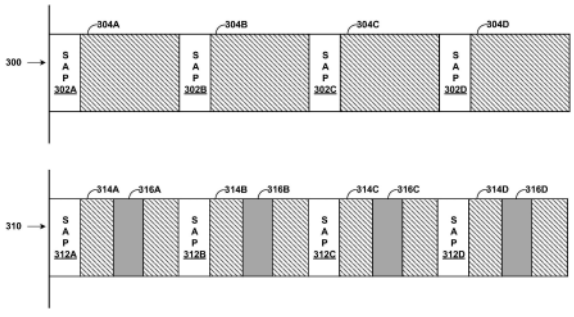
(21)申请号 202080066431.1
(22)申请日 2020.10.02
(65)同一申请的已公布的文献号
申请公布号 CN 114430911 A
(43)申请公布日 2022.05.03
(30)优先权数据
62/909,642 2019.10.02 US
17/061,152 2020.10.01 US
(85)PCT国际申请进入国家阶段日
2022.03.22
(86)PCT国际申请的申请数据
PCT/US2020/054026 2020.10.02
(87)PCT国际申请的公布数据
WO2021/067768 EN 2021.04.08
(73)专利权人 高通股份有限公司
地址 美国加利福尼亚
(72)发明人 T·施托克哈默 I·布阿齐齐 W·齐亚
(74)专利代理机构 永新专利商标代理有限公司 72002
专利代理师 赵腾飞
(51)Int.Cl.
H04N 21/262(2011.01)
H04N 21/845(2011.01)
H04L 65/60(2022.01)
H04L 65/611(2022.01)
H04L 65/80(2022.01)
(56)对比文件
3GPP.Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats,AMENDMENT 4: Segment Independent SAP Signalling (SISSI), MPD chaining, MPD reset and other extensions.《ISO/IEC 23009-1:2014/PDAM 4 and ISO/IEC JTC 1/SC 29/WG 11》.2015,第1-17页.
审查员 黄碧云
权利要求书4页 说明书28页 附图9页

(54)发明名称

在DASH分段的再同步点处的随机存取

(57)摘要

一种用于检索媒体的设备包括:存储器,其被配置为存储媒体呈现的媒体数据;以及一个或多个处理器,其在电路中实现并且被配置为:检索用于媒体呈现的清单文件,该清单文件指示能够在媒体呈现的表示的分段的再同步点处开始对比特流的媒体数据的容器解析,再同步点位于除了分段的开始以外的位置并且表示能够开始对比特流的媒体数据的容器解析的点;使用清单文件来形成用于在再同步点处开始检索表示的媒体数据的请求;发送该请求,以在再同步点处开始发起对媒体呈现的媒体数据的检索;以及呈现所检索到的媒体数据。



CN 114430911 B

1.一种检索媒体数据的方法,所述方法包括:

检索用于媒体呈现的清单文件,所述清单文件指示在所述媒体呈现的表示的分段中的再同步点的位置,所述清单文件还指示能够在所述再同步点处开始对比特流的媒体数据的容器解析,所述再同步点位于除了所述分段的开始以外的位置并且表示能够开始对所述比特流的所述媒体数据的所述容器解析的点并且表示随机存取点(RAP)跟随在所述再同步点之后;

使用所述清单文件,来形成用于在所述再同步点处开始检索所述表示的所述媒体数据的请求;

发送所述请求,以在所述再同步点处开始发起对所述媒体呈现的所述媒体数据的检索;

解析在所述再同步点处开始的所检索到的媒体数据并且确定所述媒体数据的所述RAP跟随在所述再同步点之后;以及

呈现包括所述媒体数据的所述RAP并且跟随在所述媒体数据的所述RAP之后的所检索到的媒体数据并且不呈现在从所述再同步点到所述媒体数据的所述RAP的所检索到的媒体数据。

2.根据权利要求1所述的方法,其中,呈现所检索到的媒体数据包括:解析在所述再同步点处所检索到的媒体数据的文件级别媒体数据容器。

3.根据权利要求2所述的方法,其中,解析包括:

解析所述文件级别媒体数据容器,直到检测到所述媒体呈现的所述RAP为止;以及

向媒体解码器发送所述RAP。

4.根据权利要求1所述的方法,其中,所述再同步点包括组块边界的开始。

5.根据权利要求4所述的方法,其中,所述组块边界包括组块的开始,所述组块包括零或一分段类型值、零或一制作者参考时间值、零或更多事件消息、至少一个电影片段盒和至少一个媒体数据容器盒。

6.根据权利要求1所述的方法,其中,所述清单文件指示在所述表示的所述分段中的所述再同步点的可用性。

7.根据权利要求1所述的方法,其中,检索所述清单文件包括在所述分段完全形成之前检索所述清单文件,其中,发送用于发起对所述媒体数据的检索的所述请求包括:将发起对所述媒体数据的检索的所述请求作为实时流发送,其中在检索到所述分段中的至少一些分段之前,所述媒体呈现的分段未完全形成。

8.根据权利要求6所述的方法,其中,所述清单文件指示能够在所述再同步点处执行的随机存取的类型。

9.根据权利要求6所述的方法,其中,所述清单文件指示所述再同步点的位置和定时以及所述位置和定时信息是准确的还是估计的。

10.根据权利要求1所述的方法,其中,所述清单文件包括媒体呈现描述(MPD)。

11.一种用于检索媒体数据的设备,所述设备包括:

存储器,其被配置为存储媒体呈现的媒体数据;以及

一个或多个处理器,其在电路中实现并且被配置为:

检索用于媒体呈现的清单文件,所述清单文件指示在所述媒体呈现的表示的分段中的

再同步点的位置,所述清单文件还指示能够在所述再同步点处开始对比特流的媒体数据的容器解析,所述再同步点位于除了所述分段的开始以外的位置并且表示能够开始对所述比特流的所述媒体数据的所述容器解析的点并且表示随机存取点(RAP)跟随在所述再同步点之后;

使用所述清单文件来形成用于在所述再同步点处开始检索所述表示的所述媒体数据的请求;

发送所述请求,以在所述再同步点处开始发起对所述媒体呈现的所述媒体数据的检索;

解析在所述再同步点处开始的所检索到的媒体数据并且确定所述媒体数据的所述RAP跟随在所述再同步点之后;以及

呈现包括所述媒体数据的所述RAP并且跟随在所述媒体数据的所述RAP之后的所检索到的媒体数据并且不呈现在从所述再同步点到所述媒体数据的所述RAP的所检索到的媒体数据。

12.根据权利要求11所述的设备,其中,为了呈现所检索到的媒体数据,所述一个或多个处理器被配置为:解析在所述再同步点处的所检索到的媒体数据的文件级别媒体数据容器。

13.根据权利要求12所述的设备,其中,为了解析所述文件级媒体数据容器,所述一个或多个处理器被配置为:

解析所述文件级别媒体数据容器,直到检测到所述媒体呈现的所述RAP为止;以及

向媒体解码器发送所述RAP。

14.根据权利要求11所述的设备,其中,所述再同步点包括组块边界的开始。

15.根据权利要求14所述的设备,其中,所述组块边界包括组块的开始,所述组块包括零或一分段类型值、零或一制作者参考时间值、零或更多事件消息、至少一个电影片段盒和至少一个媒体数据容器盒。

16.根据权利要求11所述的设备,其中,所述清单文件指示在所述表示的所述分段中的所述再同步点的可用性。

17.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在所述分段完全形成之前检索所述清单文件,并且其中,所述一个或多个处理器被配置为将发起对所述媒体数据的检索的所述请求作为实时流发送,其中在检索到所述分段中的至少一些分段之前,所述媒体呈现的分段未完全形成。

18.根据权利要求16所述的设备,其中,所述清单文件指示能够在所述再同步点处执行的随机存取的类型。

19.根据权利要求16所述的设备,其中,所述清单文件指示所述再同步点的位置和定时以及所述位置和定时信息是准确的还是估计的。

20.根据权利要求11所述的设备,其中,所述清单文件包括媒体呈现描述(MPD)。

21.一种在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进行以下操作:

检索用于媒体呈现的清单文件,所述清单文件指示在所述媒体呈现的表示的分段中的再同步点的位置,所述清单文件还指示能够在所述再同步点处开始对比特流的媒体数据的

容器解析,所述再同步点位于除了所述分段的开始以外的位置并且表示能够开始对所述比特流的所述媒体数据的所述容器解析的点并且表示随机存取点(RAP)跟随在所述再同步点之后;

使用所述清单文件来形成用于在所述再同步点处开始检索所述表示的所述媒体数据的请求;

发送所述请求,以在所述再同步点处开始发起对所述媒体呈现的所述媒体数据的检索;

解析在所述再同步点处开始的所检索到的媒体数据并且确定所述媒体数据的所述RAP跟随在所述再同步点之后;以及

呈现包括所述媒体数据的所述RAP并且跟随在所述媒体数据的所述RAP之后的所检索到的媒体数据并且不呈现在从所述再同步点到所述媒体数据的所述RAP的所检索到的媒体数据。

22.根据权利要求21所述的计算机可读存储介质,其中,所述再同步点包括组块边界的开始。

23.根据权利要求22所述的计算机可读存储介质,其中,所述组块边界包括组块的开始,所述组块包括零或一分段类型值、零或一制作者参考时间值、零或更多事件消息、至少一个电影片段盒和至少一个媒体数据容器盒。

24.根据权利要求21所述的计算机可读存储介质,其中,所述清单文件指示在所述表示的所述分段中的所述再同步点的可用性。

25.根据权利要求21所述的计算机可读存储介质,其中,使得所述处理器检索所述清单文件的所述指令包括使得所述处理器在所述分段完全形成之前检索所述清单文件的指令,其中,使得所述处理器发送用于发起对所述媒体数据的检索的所述请求的所述指令包括:使得所述处理器将发起对所述媒体数据的检索的所述请求作为实时流发送的指令,其中在检索到所述分段中的至少一些分段之前,所述媒体呈现的分段未完全形成。

26.根据权利要求24所述的计算机可读存储介质,其中,所述清单文件指示能够在所述再同步点处执行的随机存取的类型。

27.根据权利要求24所述的计算机可读存储介质,其中,所述清单文件指示所述再同步点的位置和定时以及所述位置和定时信息是准确的还是估计的。

28.根据权利要求21所述的计算机可读存储介质,其中,所述清单文件包括媒体呈现描述(MPD)。

29.一种用于检索媒体数据的设备,所述设备包括:

用于检索用于媒体呈现的清单文件的单元,所述清单文件指示在所述媒体呈现的表示的分段中的再同步点的位置,所述清单文件还指示能够在所述再同步点处开始对比特流的媒体数据的容器解析,所述再同步点位于除了所述分段的开始以外的位置并且表示能够开始对所述比特流的所述媒体数据的所述容器解析的点并且表示随机存取点(RAP)跟随在所述再同步点之后;

用于使用所述清单文件来形成用于在所述再同步点处开始检索所述表示的所述媒体数据的请求的单元;

用于发送所述请求以在所述再同步点处开始发起对所述媒体呈现的所述媒体数据的

检索的单元;

用于解析在所述再同步点处开始的所检索到的媒体数据并且确定所述媒体数据的所述RAP跟随在所述再同步点之后的单元;以及

用于呈现包括所述媒体数据的所述RAP并且跟随在所述媒体数据的所述RAP之后的所检索到的媒体数据并且不呈现在从所述再同步点到所述媒体数据的所述RAP的所检索到的媒体数据的单元。

30.根据权利要求29所述的设备,其中,所述再同步点包括组块边界的开始。

31.根据权利要求30所述的设备,其中,所述组块边界包括组块的开始,所述组块包括零或一分段类型值、零或一制作者参考时间值、零或更多事件消息、至少一个电影片段盒和至少一个媒体数据容器盒。

32.根据权利要求29所述的设备,其中,所述清单文件指示在所述表示的所述分段中的所述再同步点的可用性。

33.根据权利要求29所述的设备,其中,用于检索所述清单文件的单元包括用于在所述分段完全形成之前检索所述清单文件的单元,其中,用于发送用于发起对所述媒体数据的检索的所述请求的单元包括:用于将发起对所述媒体数据的检索的所述请求作为实时流发送的单元,其中在检索到所述分段中的至少一些分段之前,所述媒体呈现的分段未完全形成。

34.根据权利要求32所述的设备,其中,所述清单文件指示能够在所述再同步点处执行的随机存取的类型。

35.根据权利要求32所述的设备,其中,所述清单文件指示所述再同步点的位置和定时以及所述位置和定时信息是准确的还是估计的。

36.根据权利要求29所述的设备,其中,所述清单文件包括媒体呈现描述(MPD)。

在DASH分段的再同步点处的随机存取

[0001] 本申请要求享受于2020年10月1日递交的、编号为17/061,152的美国申请的权益,于2019年10月2日递交的、编号为62/909,642的美国临时申请的权益,据此将上述申请的全部内容通过引用的方式并入。

技术领域

[0002] 本公开内容涉及经编码的视频数据的存储和传输。

背景技术

[0003] 数字视频能力可以被合并到各种各样的设备中,各种各样的设备包括数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型计算机或台式计算机、数字相机、数字记录设备、数字媒体播放器、视频游戏设备、视频游戏机、蜂窝或卫星无线电电话、视频电话会议设备等。数字视频设备实现视频压缩技术(诸如在通过MPEG-2、MPEG-4、ITU-T H.263或ITU-T H.264/MPEG-4(第10部分,高级视频译码(AVC))、ITU-T H.265(还被称为高效率视频译码(HEVC))定义的标准以及此类标准的扩展中描述的技术),以更加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,以减少或去除在视频序列中固有的冗余。对于基于块的视频译码,可以将视频帧或切片划分为宏块。每个宏块可以被进一步分割。在经帧内译码的(I)帧或切片中的宏块是使用相对于相邻宏块的空间预测来编码的。在经帧间译码的(P或B)帧或切片中的宏块可以使用相对于在相同帧或切片中的相邻宏块的空间预测,或者相对于其它参考帧的时间预测。

[0005] 在视频数据已经被编码之后,视频数据可以被分组化以用于传输或存储。视频数据可以被组装成符合各种标准(诸如国际标准化组织(ISO)的基媒体文件格式以及其扩展(诸如AVC))中的任何一种标准的视频文件。

发明内容

[0006] 概括而言,本公开内容描述了用于存取在基于HTTP的动态自适应流式传输(DASH)和/或通用媒体存取格式(CMAF)中的分段(不仅在分段的开始处,而且在分段内的其它地方)的数据(例如,用于随机存取)的技术。本公开内容还描述了涉及用信号通知在分段内执行随机存取的能力的技术。本公开内容描述了与这些技术相关的各种用例。例如,本公开内容定义了在DASH和ISO基媒体文件格式(BMFF)中的再同步(resync)点以及再同步点的信令。

[0007] 在一个示例中,一种检索媒体数据的方法包括:检索用于媒体呈现的清单文件,该清单文件指示能够在媒体呈现的表示的分段的再同步点处开始对比特流的媒体数据的容器解析,该再同步点位于除了分段的开始以外的位置并且表示能够开始对比特流的媒体数据的容器解析的点;使用清单文件,来形成用于在再同步点处开始检索表示的媒体数据的请求;发送该请求,以在再同步点处开始发起对媒体呈现的媒体数据的检索;以及呈现所检

索到的媒体数据。

[0008] 在另一示例中,一种用于检索媒体数据的设备包括:存储器,其被配置为存储媒体呈现的媒体数据;以及一个或多个处理器,其在电路中实现并且被配置为:检索用于媒体呈现的清单文件,该清单文件指示能够在媒体呈现的表示的分段的再同步点处开始对比特流的媒体数据的容器解析,该再同步点位于除了分段的开始以外的位置并且表示能够开始对比特流的媒体数据的容器解析的点;使用清单文件来形成用于在再同步点处开始检索表示的媒体数据的请求;发送该请求,以在再同步点处开始发起对媒体呈现的媒体数据的检索;以及呈现所检索到的媒体数据。

[0009] 在另一示例中,一种计算机可读存储介质,其上存储有指令,该指令在被执行时使得处理器进行以下操作:检索用于媒体呈现的清单文件,该清单文件指示能够在媒体呈现的表示的分段的再同步点处开始对比特流的媒体数据的容器解析,该再同步点位于除了分段的开始以外的位置并且表示能够开始对比特流的媒体数据的容器解析的点;使用清单文件来形成用于在再同步点处开始检索表示的媒体数据的请求;发送该请求,以在再同步点处开始发起对媒体呈现的媒体数据的检索;以及呈现所检索到的媒体数据。

[0010] 在另一示例中,一种用于检索媒体数据的设备包括:用于检索用于媒体呈现的清单文件的单元,该清单文件指示能够在媒体呈现的表示的分段的再同步点处开始对比特流的媒体数据的容器解析,该再同步点位于除了分段的开始以外的位置并且表示能够开始对比特流的媒体数据的容器解析的点;用于使用清单文件来形成用于在再同步点处开始检索表示的媒体数据的请求的单元;用于发送该请求以在再同步点处开始发起对媒体呈现的媒体数据的检索的单元;以及用于呈现所检索到的媒体数据的单元。

[0011] 在附图和下文描述中阐述了一个或多个示例的细节。根据说明书和附图以及根据权利要求书,其它特征、对象和优点将是显而易见的。

附图说明

[0012] 图1是示出实现用于在网络上对媒体数据进行流式传输的技术的示例系统的框图。

[0013] 图2是检索单元的示例组件集合的框图。

[0014] 图3是示出示例多媒体内容的元素的概念图。

[0015] 图4是示出可以对应于表示的分段的示例视频文件的元素的框图。

[0016] 图5是示出根据本公开内容的可以在第一用例中使用的示例低时延架构的概念图。

[0017] 图6是进一步详细地示出关于图5描述的用例的示例的概念图。

[0018] 图7是示出在广播协议的上下文中使用DASH和CMAF随机存取的示例第二用例的概念图。

[0019] 图8是示出清单文件中的流存取点(SAP)的示例信令的概念图。

[0020] 图9是示出根据本公开内容的技术的检索媒体数据的示例方法的流程图。

具体实施方式

[0021] 本公开内容的技术可以应用于符合根据以下各项中的任何一项封装的视频数据

的视频文件:ISO基媒体文件格式、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、或其它类似的视频文件格式。

[0022] 在HTTP流式传输中,频繁使用的操作包括HEAD、GET和部分GET。HEAD操作检索与给定的统一资源定位符(URL)或统一资源名称(URN)相关联的文件的报头,而不检索与URL或URN相关联的有效载荷。GET操作检索与给定的URL或URN相关联的整个文件。部分GET操作接收作为输入参数的字节范围,并且检索文件的连续数量的字节,其中字节数量对应于所接收的字节范围。因此,可以提供电影片段(fragment)以用于HTTP流式传输,这是因为部分GET操作可以获得一个或多个个体电影片段。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存取的结构化数据集合。客户端可以请求并且下载媒体数据信息以向用户呈现流式服务。

[0023] 在使用HTTP流式传输来对3GPP数据进行流式传输的示例中,针对多媒体内容的视频和/或音频数据可以存在多个表示。如下文所解释的,不同的表示可以对应于不同的译码特性(例如,视频译码标准的不同的简档或级别)、不同的译码标准或译码标准的扩展(诸如多视图和/或可缩放扩展)、或不同的比特率。这样的表示的清单可以是在媒体呈现描述(MPD)数据结构中定义的。媒体呈现可以对应于HTTP流式传输客户端设备可存取的结构化数据集合。HTTP流式传输客户端设备可以请求并且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。媒体呈现可以是在MPD数据结构中描述的,MPD数据结构可以包括MPD的更新。

[0024] 媒体呈现可以包含一个或多个时段的序列。每个时段可以延长直到下一时段的开始为止,或者直到媒体呈现的结束为止(在最后一个时段的情况下)。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、视频、定时文本或其它此类数据的数个替代的经编码的版本中的一个版本。表示可以在编码类型(例如,对于视频数据而言,比特率、分辨率和/或编解码器,以及对于音频数据而言,比特率、语言和/或编解码器)方面不同。术语表示可以用于指代经编码的音频或视频数据中的与多媒体内容的特定时段相对应并且以特定方式编码的部分。

[0025] 特定时段的表示可以被指派给通过MPD中的指示该表示所属的适配集的属性指示的组。同一适配集中的表示通常被认为是彼此的替代,这是因为客户端设备可以在这些表示之间动态地并且无缝地切换,例如以执行带宽适配。例如,针对特定时段的视频数据的每个表示可以被指派给相同的适配集,使得可以选择这些表示中的任何表示进行解码以呈现针对对应的时段的多媒体内容的媒体数据,诸如视频数据或音频数据。在一些示例中,在一个时段内的媒体内容可以通过来自组0的一个表示(如果存在的话)或者来自每个非零组的至多一个表示的组合来表示。针对时段的每个表示的定时数据可以是相对于该时段的开始时间来表达的。

[0026] 表示可以包括一个或多个分段。每个表示可以包括初始化分段,或者表示的每个分段可以是自初始化的。当存在时,初始化分段可以包含用于存取表示的初始化信息。通常,初始化分段不包含媒体数据。分段可以由标识符(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))唯一地引用。MPD可以针对每个分段提供标识符。在一些示例中,MPD还可以以range属性的形式提供字节范围,字节范围可以对应于用于在文件

内通过URL、URN或URI可存取的分段的数据。

[0027] 对于不同类型的媒体数据,可以选择不同的表示以用于基本上同时地检索。例如,客户端设备可以选择要从其检索分段的音频表示、视频表示和定时文本表示。在一些示例中,客户端设备可以选择特定的适配集以用于执行带宽适配。也就是说,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,客户端设备可以针对某些类型的媒体(例如,视频)选择适配集,并且针对其它类型的媒体(例如,音频和/或定时文本)直接地选择表示。

[0028] 基于HTTP的低时延动态自适应流式传输(LL-DASH)是用于DASH的简档,其尝试以低时延向DASH客户端提供媒体数据。下文简要地概述了用于LL-DASH的某些技术:

[0029] • 编码是基于片段化的ISO BMFF文件的,并且通常假设CMAF片段和CMAF组块(chunk)的。

[0030] • 每个组块是由DASH打包器(packager)单独地可存取的,并且被映射到被上传到源服务器的HTTP组块。这种1对1映射是针对低时延操作的建议,但是不是要求。无论如何,客户端都不应当假设这种1对1映射被保留给客户端。

[0031] • 使用了部分可用分段的低延迟协议(例如,HTTP组块传输编码),使得客户端可以在这些分段完成之前存取它们。针对能够利用这种特征的客户端来调整可用性开始时间。

[0032] • 允许两种操作模式:

[0033] o通过应用@duration信令和基于$Number$的模板化来使用简单的实时产品

[0034] o通过在DASH第四版中提议的更新来支持具有作为$Number$或者$Time$的SegmentTimeline的主要实时产品。

[0035] • 可以使用MPD有效性到期事件,但是客户端不必理解这些事件。

[0036] • 通常,带内事件消息可以存在,但是仅预期客户端恢复在分段的开始处(不是在任意组块处)的消息。DASH打包器可以在组块边界处或者使用定时元数据轨道完全异步地接收来自编码器的通知。

[0037] • 允许在单个媒体呈现中以及在媒体呈现的一个时段内具有使用组块低时延模式的适配集以及针对不同媒体类型使用短分段的适配集。

[0038] • DASH客户端在媒体管道上的特定量的播放控制可以是可用的,并且应当用于DASH客户端的稳健性。例如,回放可以在某个时间段内被加速或减速,或者DASH客户端可以执行对分段的搜索。

[0039] • 该系统被设计为与标准HTTP/1.1一起可工作,但是还应当适用于HTTP扩展和其它协议,以用于改进的低时延操作。

[0040] • MPD包括关于服务配置以及服务特性(例如,包括服务的目标时延)的显式信令。

[0041] • MPD以及可能还有分段包括锚定时间,其允许DASH客户端测量与实时相比的当前时延,并且进行调整以满足服务期望。

[0042] • 例如,在编码器故障的情况下,解决了操作稳健性。

[0043] • 现有的DRM和加密模式是与提议的低时延操作兼容的。

[0044] 基于以上高层面概述,定义了如下内容:分段可以用于不仅在分段边界处、而且在分段内随机地存取表示。如果提供了这样的随机存取,则应当在MPD中对此用信号通知。

[0045] 图1是示出实现用于在网络上对媒体数据进行流式传输的技术的示例系统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通过可以包括互联网的网络74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。

[0046] 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可以包括例如麦克风,其产生表示被捕获的要由音频编码器26编码的音频数据的电信号。替代地,音频源22可以包括存储先前记录的音频数据的存储介质、音频数据生成器(诸如计算机化的合成器)、或任何其它音频数据源。视频源24可以包括产生要由视频编码器28编码的视频数据的摄像机、利用先前记录的视频数据来编码的存储介质、视频数据生成单元(诸如计算机图形源)、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,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。

[0047] 原始音频和视频数据可以包括模拟数据或数字数据。模拟数据可以在被音频编码器26和/或视频编码器28编码之前被数字化。音频源22可以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可以同时获得讲话参与者的视频数据。在其它示例中,音频源22可以包括计算机可读存储介质,计算机可读存储介质包括存储的音频数据,以及视频源24可以包括计算机可读存储介质,计算机可读存储介质包括存储的视频数据。以这种方式,在本公开内容中描述的技术可以被应用于实况的(live)、流式传输的、实时的(real-time)音频和视频数据或者被应用于被存档的、预先记录的音频和视频数据。

[0048] 与视频帧相对应的音频帧通常是包含音频数据的音频帧,音频数据是与由视频源24捕获的(或生成的)被包含在视频帧内的视频数据同时地、由音频源22捕获的(或生成的)。例如,当讲话参与者通常通过讲话产生音频数据时,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。相应地,与视频帧相对应的音频帧通常对应于音频数据和视频数据同时被捕获的情况,并且对于该情况,音频帧和视频帧分别包括同时被捕获的音频数据和视频数据。

[0049] 在一些示例中,音频编码器26可以将表示用于每个经编码的音频帧的音频数据被记录的时间的时间戳编码在该经编码的音频帧中,并且类似地,视频编码器28可以将表示用于每个经编码的视频帧的视频数据被记录的时间的时间戳编码在该经编码的视频帧中。在这样的示例中,与视频帧相对应的音频帧可以包括含有时间戳的音频帧和含有相同时间戳的视频帧。内容准备设备20可以包括内部时钟,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将音频数据和视频数据分别与时间戳进行关联。

[0050] 在一些示例中,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,以及视频源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。在一些示例中,音频编码器26可以将序列标识符编码在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,但是不一定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,在一些示例中,序列标识符可以被映射或以其它方式与时间戳相关。

[0051] 音频编码器26通常产生经编码的音频数据的流,而视频编码器28产生经编码的视频数据的流。每个个别的数据流(无论是音频还是视频)都可以被称为基本流。基本流是表示的单个的、经数字译码的(可能被压缩的)分量。例如,表示的经译码的视频或音频部分可以是基本流。基本流在被封装在视频文件内之前,可以被转换为分组化基本流(PES)。在同一表示内,流ID可以用于将属于一个基本流的PES分组与属于另一基本流的PES分组区分开来。基本流的基本数据单元是分组化基本流(PES)分组。因此,经译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相应的基本流。

[0052] 许多视频译码标准(诸如ITU-T H.264/AVC、以及即将产生的高效率视频译码(HEVC)标准)定义了用于无错误比特流的语法、语义和解码过程,其中的任何一者符合某个简档或级别。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标准兼容的。在视频译码标准的上下文中,"简档"对应于应用于它们的算法、特征、或工具和约束的子集。如由H.264标准所定义的,例如,"简档"是由H.264标准指定的整个比特流语法的子集。"级别"对应于与图片的分辨率、比特率和块处理率有关的解码器资源消耗的限制,诸如例如,解码器存储器和计算。可以利用profile_idc(简档指示符)值来用信号通知简档,而可以利用level_idc(级别指示符)值来用信号通知级别。

[0053] 例如,H.264标准认识到,在由给定简档的语法施加的界限内,仍然可能需要编码器和解码器的性能的大变化,这取决于比特流中的语法元素所取的值,诸如经解码的图片的指定大小。H.264标准进一步认识到,在许多应用中,实现能够处理在特定简档内的语法的所有假设使用的解码器是既不实际也不经济的。因此,H.264标准将"级别"定义为对在比特流中的语法元素的值施加的指定的约束集合。这些约束可以是对值的简单限制。替代地,这些约束可以采取对值的算术组合的约束的形式(例如,图片宽度乘以图片高度乘以每秒经解码的图片数量)。H.264标准还规定,各个实现针对每个支持的简档可以支持不同级别。

[0054] 符合简档的解码器通常支持在简档中定义的所有特征。例如,作为译码特征,B图片译码在H.264/AVC的基准简档中是不支持的,但是在H.264/AVC的其它简档中是支持的。符合级别的解码器应当能够对不需要超出在该级别中定义的限制的资源的任何比特流进行解码。简档和级别的定义可以有助于可解释性。例如,在视频传输期间,可以针对整个传输会话协商并且约定一对简档和级别定义。更具体地,在H.264/AVC中,级别可以定义对以下各项的限制:需要被处理的宏块的数量、经解码图片缓冲器(DPB)大小、经译码图片缓冲器(CPB)大小、垂直运动矢量范围、每两个连续MB的运动矢量的最大数量、以及B块是否可以具有小于8x8个像素的子宏块分割。以这种方式,解码器可以确定该解码器是否能够对比特流进行正确地解码。

[0055] 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接收包括经译码的视频数据的基本流,并且从音频编码器26接收包括经译码的音频数据的基本流。在一些示例中,视频编码器28和音频编码器26可以各自包括用于从经编码的数据形成PES分组的分组化器。在其它示例中,视频编码器28和音频编码器26可以各自与用于从经编码的数据形成PES分组的相应的分组化器进行对接。在其它示例中,封装单元30可以包括用于从经编码的音频数据和视频数据形成PES分组的分组化器。

[0056] 视频编码器28可以以各种方式对多媒体内容的视频数据进行编码,以产生多媒体内容的处于各种比特率并且具有各种特性(诸如像素分辨率、帧速率、符合各种译码标准、

符合用于各种译码标准的各个简档和/或简档的级别、具有一个或多个视图的表示(例如,用于二维或三维回放)或其它这样的特性)的不同表示。如在本公开内容中使用的表示可以包括音频数据、视频数据、文本数据(例如,用于隐藏式字幕)或其它这样的数据中的一者。表示可以包括基本流,诸如音频基本流或视频基本流。每个PES分组可以包括标识该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个表示的视频文件(例如,分段)。

[0057] 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,并且从PES分组形成对应的网络抽象层(NAL)单元。可以将经译码的视频分段组织为NAL单元,这些NAL单元提供了寻址到诸如视频电话、存储、广播或流式传输的应用的“网络友好”视频表示。NAL单元可以被分类为视频译码层(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/或切片级别数据。其它NAL单元可以是非VCL NAL单元。在一些示例中,在一个时间实例中经译码的图片(通常被呈现为基本经译码的图片)可以被包含在存取单元中,存取单元可以包括一个或多个NAL单元。

[0058] 除此之外,非VCL NAL单元可以包括参数集NAL单元和SEI NAL单元。参数集可以包含序列级别报头信息(在序列参数集(SPS)中)和不频繁变化的图片级别报头信息(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不频繁变化的信息;因此可以提高译码效率。此外,使用参数集可以实现对重要报头信息的带外传输,从而避免对用于错误恢复的冗余传输的需求。在带外传输示例中,可以在与其它NAL单元(诸如SEI NAL单元)不同的信道上发送参数集NAL单元。

[0059] 补充增强信息(SEI)可以包含对于从VCL NAL单元对经译码的图片样本进行解码不必要的信息,但是可能有助于与解码、显示、错误恢复和其它目的有关的过程。SEI消息可以被包含在非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并非总是强制的。SEI消息可以是序列级别SEI消息或图片级别SEI消息。一些序列级别信息可以被包含在SEI消息(诸如在SVC的示例中的可缩放性信息SEI消息和在MVC中的视图可缩放性信息SEI消息)中。这些示例SEI消息可以传递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,诸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)来将MPD格式化。

[0060] 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以及清单文件(例如,MPD)提供给输出接口32。输出接口32可以包括网络接口、或用于写入存储介质的接口(诸如通用串行总线(USB)接口、CD或DVD刻录机或烧录机、去往磁存储介质或闪速存储介质的接口、或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,输出接口32可以经由网络传输或存储介质将数据发送给服务器设备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,每种多媒体内容64包括相应的清单文件66和一个或多个表示68A-68N(表示68)。在一些示例中,输出接口32还可以直接向网络74发送数据。

[0061] 在一些示例中,表示68可以被分成适配集。也就是说,表示68的各个子集可以包括相应的公共特性集,诸如编解码器、简档和级别、分辨率、视图数量、用于分段的文件格式、可以标识将与表示和/或要被解码和例如通过扬声器呈现的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、可以描述针对在适配集中的表示的场景的相机角度或现实

世界视角的相机角度信息、描述对于特定观众的内容适合性的评级信息等。

[0062] 清单文件66可以包括指示与特定的适配集相对应的表示68的子集以及用于适配集的公共特性的数据。清单文件66还可以包括表示用于适配集的个别表示的个别特性(诸如比特率)的数据。以这种方式,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的适配集的子元素来指示在适配集中的表示。

[0063] 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它设备(诸如路由器、桥接器、代理设备、交换机或其它设备)上实现。在一些示例中,内容递送网络的中间设备可以对多媒体内容64的数据进行高速缓存,并且包括基本上与服务器设备60的组件一致的组件。通常,网络接口72被配置为经由网络74发送和接收数据。

[0064] 请求处理单元70被配置为从诸如客户端设备40的客户端设备接收对存储介质62的数据的网络请求。例如,请求处理单元70可以实现如在RFC 2616(1999年6月,IETF,网络工作组,R.Fielding等人的“Hypertext Transfer Protocol-HTTP/1.1”)中描述的超文本传输协议(HTTP)版本1.1。也就是说,请求处理单元70可以被配置为接收HTTP GET或部分GET请求,并且响应于该请求来提供多媒体内容64的数据。请求可以例如使用表示68中的一个表示68的分段的URL来指定该分段。在一些示例中,请求还可以指定分段的一个或多个字节范围,由此包括部分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中的一个表示的分段的报头数据。在任何情况下,请求处理单元70可以被配置为处理请求以将所请求的数据提供给进行请求的设备,诸如客户端设备40。

[0065] 另外或替代地,请求处理单元70可以被配置为经由诸如eMBMS的广播或多播协议来递送媒体数据。内容准备设备20可以以与所描述的基本相同的方式来创建DASH分段和/或子分段,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分段或子分段。例如,请求处理单元70可以被配置为从客户端设备40接收多播组加入请求。也就是说,服务器设备60可以向客户端设备(包括客户端设备40)通告与多播组相关联的互联网协议(IP)地址,该多播组与特定的媒体内容(例如,实况事件的广播)相关联。客户端设备40进而可以提交用于加入多播组的请求。该请求可以在整个网络74(例如,组成网络74的路由器)中传播,使得致使路由器将去往与多播组相关联的IP地址的业务引导到订制客户端设备(诸如客户端设备40)。

[0066] 如在图1的示例中所示,多媒体内容64包括清单文件66,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表示68的编解码器信息、简档值、级别值、比特率和其它描述性特性。客户端设备40可以检索媒体呈现的MPD以确定如何存取表示68的分段。

[0067] 具体地,检索单元52可以检索客户端设备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以下各项中的任何一项或全部: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。检索单元52可以包括例如被配置为提交HTTP GET和部分GET请求的网页浏览器或媒体客户端。检索单元52可以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描述的功能中的全部或部分功能可以用硬件、或者用硬

件、软件和/或固件的组合来实现，其中可以提供必需的硬件以执行针对软件或固件的指令。

[0068] 检索单元52可以将客户端设备40的解码和渲染能力与通过清单文件66的信息指示的表示68的特性进行比较。检索单元52可以初始地检索清单文件66的至少一部分以确定表示68的特性。例如，检索单元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一部分。检索单元52可以选择表示68的具有能够通过客户端设备40的译码和渲染能力满足的特性的子集(例如，适配集)。检索单元52然后可以确定用于在适配集中的表示的比特率，确定当前可用的网络带宽量，并且从表示中的具有能够通过网络带宽满足的比特率的一个表示中检索分段。

[0069] 通常，较高比特率的表示可以产生较高质量的视频回放，而较低比特率的表示可以在可用网络带宽减小时提供足够质量的视频回放。相应地，当可用网络带宽是相对高的时，检索单元52可以从相对高比特率的表示中检索数据，而当可用网络带宽是低的时，检索单元52可以从相对低比特率的表示中检索数据。以这种方式，客户端设备40可以在网络74上对多媒体数据进行流式传输，同时还适应于网络74的变化的网络带宽可用性。

[0070] 另外或替代地，检索单元52可以被配置为根据诸如eMBMS或IP多播的广播或多播网络协议来接收数据。在这样的示例中，检索单元52可以提交用于加入与特定的媒体内容相关联的多播网络组的请求。在加入多播组之后，检索单元52可以接收该多播组的数据，而无需向服务器设备60或内容准备设备20发出另外的请求。当不再需要多播组的数据(例如，停止回放或者将信道改变到不同的多播组)时，检索单元52可以提交用于离开该多播组的请求。

[0071] 网络接口54可以接收所选择的表示的分段的数据并且将其提供给检索单元52，检索单元52进而可以将分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的PES流，对PES流进行解分组化以检索经编码的数据，并且向音频解码器46或者视频解码器48(取决于经编码的数据是音频流还是视频流的一部分，例如如通过流的PES分组报头指示的)发送经编码的数据。音频解码器46对经编码的音频数据进行解码并且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且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。

[0072] 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均可以在适用的情况下被实现为各种适当的处理电路中的任何一者，诸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、固件或其任何组合。视频编码器28和视频解码器48中的每一者可以被包括在一个或多个编码器或解码器中，其中的任一者可以被集成为组合的视频编码器/解码器(CODEC)的一部分。同样，音频编码器26和音频解码器46中的每一者可以被包括在一个或多个编码器或解码器中，其中的任一者可以被集成为组合的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、检索单元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(诸如蜂窝电话)。

[0073] 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进行操作。出于示例的目的，本公开内容关于客户端设备40和服务器设备60描述

了这些技术。然而,应当理解的是,内容准备设备20可以被配置为代替服务器设备60(或者除了服务器设备60之外)执行这些技术。

[0074] 封装单元30可以形成NAL单元,NAL单元包括标识该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字节的报头和可变大小的有效载荷。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级别的视频数据。例如,NAL单元可以包括视频数据的块、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可以以基本流的PES分组的形式从视频编码器28接收经编码的视频数据。封装单元30可以将每个基本流与对应的节目进行关联。

[0075] 封装单元30还可以从多个NAL单元组装存取单元。通常,存取单元可以包括一个或多个NAL单元,其用于表示视频数据的帧、以及与该帧相对应的音频数据(当这样的音频数据是可用的时)。存取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,如果每个视图具有20帧每秒(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,可以同时渲染用于同一存取单元(同一时间实例)的所有视图的特定帧。在一个示例中,存取单元可以包括在一个时间实例中的经译码的图片,其可以被呈现为基本经译码图片。

[0076] 相应地,存取单元可以包括公共时间实例的所有音频帧和视频帧,例如,与时间X相对应的所有视图。本公开内容还将特定视图的经编码的图片称为“视图分量”。也就是说,视图分量可以包括针对在特定时间处的特定视图的经编码的图片(或帧)。相应地,存取单元可以被定义为包括公共时间实例的所有视图分量。存取单元的解码顺序不一定需要与输出顺序或显示顺序相同。

[0077] 媒体呈现可以包括媒体呈现描述(MPD),其可以包含不同替代表示(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、简档值和级别值。MPD是清单文件(诸如清单文件66)的一个示例。客户端设备40可以检索媒体呈现的MPD,以确定如何存取各个呈现的电影片段。电影片段可以位于视频文件的电影片段盒(box)(moof盒)中。

[0078] 清单文件66(其可以包括例如MPD)可以通告表示68的分段的可用性。也就是说,MPD可以包括指示表示68中的一个表示的第一分段变得可用的挂钟时间的信息、以及指示在表示68内的分段的持续时间的信息。以这种方式,客户端设备40的检索单元52可以基于在特定分段之前的分段的开始时间以及持续时间来确定每个分段何时可用。

[0079] 在封装单元30已经基于所接收的数据将NAL单元和/或存取单元组装为视频文件之后,封装单元30将视频文件传递到输出接口32以进行输出。在一些示例中,封装单元30可以将视频文件进行本地存储或者经由输出接口32将视频文件发送给远程服务器,而不是将视频文件直接发送给客户端设备40。输出接口32可以包括例如发射机、收发机、用于将数据写入计算机可读介质的设备(诸如例如,光学驱动器、磁介质驱动器(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,诸如例如,传输信号、磁介质、光学介质、存储器、闪存驱动器或其它计算机可读介质。

[0080] 网络接口54可以经由网络74接收NAL单元或存取单元,并且经由检索单元52将NAL

单元或存取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,将PES流进行解分组化以检索经编码的数据,并且向音频解码器46或者视频解码器48(取决于经编码的数据是音频流还是视频流的一部分,例如如通过流的PES分组报头指示的)发送经编码的数据。音频解码器46对经编码的音频数据进行解码并且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且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。

[0081] 根据本公开内容的技术,内容准备设备20和/或服务器设备60可以在DASH/CMAF分段中添加额外的随机存取点。随机存取包括干净的随机存取以及开放的或逐步的解码器刷新,一直到仅提供对文件格式解析的再同步。这可以通过提供以下各项来解决:提供能够在该点处开始再同步和解密的信息的组块边界;以及关于后续随机存取点的类型的信令。tfdt的可用性连同moof报头信息以及潜在地对初始化分段的使用允许在呈现时间级别上的时间再同步。本公开内容将这个新点称为“再同步点”。也就是说,再同步点表示可以正确地解析文件级别容器(例如,在ISO BMFF中的盒)并且跟随其之后将出现在媒体数据(例如,I帧)中的随机存取点的点。因此,客户端设备40可以例如在这些随机存取点中的一个随机存取点处随机地存取多媒体内容64。

[0082] 内容准备设备20和/或服务器设备60还可以在清单文件66(例如,MPD)中添加适当的信令,该信令指示在每个DASH分段中的随机存取点和再同步的可用性以及提供关于随机存取点的位置、类型和定时的信息。内容准备设备20和/或服务器设备60可以在清单文件66(MPD)中提供指示分段中的额外再同步点可用的信令,可能添加关于随机存取的位置、定时和类型的特性、以及信息是准确的还是估计的。因此,客户端设备40可以使用该用信号通知的数据来确定这样的随机存取点是否是可用的,并且相应地对检索和回放进行再同步。

[0083] 客户端设备40可以被配置具有如下的能力:在任意起点的情况下,通过找到再同步点来再同步以解封装、解密和解码。内容准备设备20和/或服务器设备60可以提供满足如在CMAF TUC中提出的以上要求的适当的组块。稍后可以定义不同的类型。

[0084] 客户端设备40可以被配置为在如例如在基于HTML-5/MSE的回放中可用的受限的接收器环境中开始处理。这个问题可以通过接收器实现来解决。然而,向具有获得数据结构中的再同步点的映射、定时和类型的能力的接收器管道提供再同步触发和信息将是合适的,并且向所述接收器管道提供再同步触发和信息允许解码管道的用户在随机存取点处对回放进行初始化。

[0085] 此外,可以提供在清单文件66中的后向兼容的信令。未被配置为具有解析信令的能力的客户端设备可以忽略该信令并且执行如上所讨论的方法。此外,清单文件66可以包括将位置与@bandwidth的值进行关联的信令,以便在适配集级别上启用信令。

[0086] 图2是更详细地示出图1的检索单元52的示例组件集合的框图。在该示例中,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。

[0087] 在该示例中,eMBMS中间件单元100还包括eMBMS接收单元106、高速缓存104和代理服务器单元102。在该示例中,eMBMS接收单元106被配置为例如根据基于单向传输的文件递送(FLUTE)经由eMBMS来接收数据,FLUTE是在2012年11月、RFC 6726、网络工作组、T.Paila等人的“FLUTE—File Delivery over Unidirectional Transport”中描述的,其在tools.ietf.org/html/rfc6726处可获得。也就是说,eMBMS接收单元106可以经由广播从例

如服务器设备60接收文件,服务器设备60可以充当广播/多播服务中心(BM-SC)。

[0088] 随着eMBMS中间件单元100接收用于文件的数据,eMBMS中间件单元可以将所接收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,诸如闪存、硬盘、RAM或任何其它适当的存储介质。

[0089] 代理服务器单元102可以充当用于DASH客户端110的服务器。例如,代理服务器单元102可以向DASH客户端110提供MPD文件或其它清单文件。代理服务器单元102可以在MPD文件中通告针对分段的可用性时间以及分段可以从其中被检索的超链接。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,对于IPv4,127.0.0.1)。以这种方式,DASH客户端110可以使用HTTP GET或部分GET请求从代理服务器单元102请求分段。例如,对于从链接http://127.0.0.1/rep1/seg3可得到的分段,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,并且向代理服务器单元102提交该请求。代理服务器单元102可以响应于这样的请求来从高速缓存104检索所请求的数据,并且将该数据提供给DASH客户端110。

[0090] 图3是示出示例多媒体内容120的元素的概念图。多媒体内容120可以对应于多媒体内容64(图1)或在存储介质62中存储的另一多媒体内容。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-124N(表示124)。表示124A包括可选的报头数据126和分段128A-128N(分段128),而表示124N包括可选的报头数据130和分段132A-132N(分段132)。为了方便起见,字母N用于指定在表示124中的每个表示124中的最后一个电影片段。在一些示例中,在表示124之间可以存在不同数量的电影片段。

[0091] MPD 122可以包括与表示124分开的数据结构。MPD 122可以对应于图1的清单文件66。同样,表示124可以对应于图1的表示68。通常,MPD 122可以包括通常描述表示124的特性(诸如译码和渲染特性、适配集、MPD 122所对应的简档、文本类型信息、相机角度信息、评级信息、特技模式信息(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,用于在回放期间将目标广告插入到媒体内容中)的信息)的数据。

[0092] 报头数据126(当存在时)可以描述分段128的特性,例如,随机存取点(RAP,还被称为流存取点(SAP))的时间位置、分段128中的哪个分段128包括随机存取点、在分段128内的随机存取点的字节偏移、分段128的统一资源定位符(URL)、或分段128的其它方面。报头数据130(当存在时)可以描述分段132的类似特性。另外或替代地,这样的特性可以被完全包括在MPD 122内。

[0093] 分段128、分段132包括一个或多个经译码的视频样本,其中的每个视频样本可以包括视频数据的帧或切片。分段128的经译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、宽度和带宽要求。这样的特性可以通过MPD 122的数据来描述,尽管在图3的示例中未示出这样的数据。MPD 122可以包括如通过3GPP规范描述的特性,其中添加了在本公开内容中描述的用信号通知的信息中的任何或全部信息。

[0094] 分段128、分段132中的每个分段可以与唯一的统一资源定位符(URL)相关联。因此,分段128、分段132中的每个分段可以是使用诸如DASH的流式传输网络协议来独立地可检索的。以这种方式,诸如客户端设备40的目的地设备可以使用HTTP GET请求来检索分段128或分段132。在一些示例中,客户端设备40可以使用HTTP部分GET请求来检索分段128或分段132的特定字节范围。

[0095] 根据本公开内容的技术,MPD 122(其再次可以对应于图1的清单文件66)可以包括用于解决上文讨论的问题的信令。例如,对于表示124中的每个表示124(并且可能在适配集级别上默认的),MPD 122可以包括一个或多个再同步元素(其可以实现后向兼容性)。每个再同步元素可以指示:对于在表示124中的对应表示124中的分段128、分段132中的每个分段,以下各项成立:

[0096] • 流存取点(SAP)类型@type或更小(但是大于0)的再同步点存在于每个分段中,其中,最大deltaT是通过@dT用信号通知的,并且最大字节偏移差是通过@dImax用信号通知的,并且最小字节偏移差是通过@dImin用信号通知的,其中,这些值中的两个值需要乘以被指派给该表示的@bandwidth属性的值以获得该值。“deltaT”指代跟随在再同步点之后的任何数据的最早的呈现时间的差,其以表示的@timescale为单位。如果@type被设置为0,则仅保证在容器和解密级别上的再同步。

[0097] • 可以使用如通过正在使用的分段格式定义的再同步模式来设置再同步标记标志@marker,以指示在每个再同步点处包括再同步点。

[0098] • 对于不同的SAP类型,可以存在多个再同步元素。

[0099] • 再同步点要求对媒体的处理能够发生在ISO BMFF和结合CMAF报头/初始化分段的解密信息上。

[0100] 下文关于图8描述了再同步点的示例使用。

[0101] 对于分段128、分段132中的任何分段(针对其,在MPD 122中提供再同步元素并且该分段是基于ISO BMFF或CMAF的),以下各项可以成立:

[0102] • 分段可以是如下定义的一个或多个组块的序列。

[0103] • 此外,对于在元素中指定的@type的分段中的任何两个连续再同步点,以下各项可以成立:

[0104] o两者的最早呈现时间的差至多可以是@dT的值。

[0105] o从开始的字节偏移的差至多可以是利用@bandwidth值进行归一化的@dImax。

[0106] o从开始的字节偏移的差至少可以是利用@bandwidth值进行归一化的@dImin。

[0107] • 如果设置了再同步标记标志,则每个再同步点可以包括再同步盒/styp。

[0108] 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可以对应于表示的分段,诸如图3的分段128、分段132中的一者。分段128、分段132中的每一者可以包括基本上符合在图4的示例中示出的数据的布置的数据。视频文件150可以被认为是封装分段。如上所述,根据ISO基媒体文件格式以及其扩展的视频文件将数据存储在被称为“盒”的一系列对象中。在图4的示例中,视频文件150包括文件类型(FTYP)盒152、电影(MOOV)盒154、分段索引(sidx)盒162、电影片段(MOOF)盒164、以及电影片段随机存取(MFRA)盒166。尽管图4表示视频文件的示例,但是应当理解的是,其它媒体文件可以包括根据ISO基媒体文件格式以及其扩展与视频文件150的数据类似地构造的其它类型的媒体数据(例如,音频数据、定时文本数据等)。

[0109] 文件类型(FTYP)盒152通常描述用于视频文件150的文件类型。文件类型盒152可以包括标识描述用于视频文件150的最佳使用的规范的数据。文件类型盒152可以替代地被放置在MOOV盒154、电影片段盒164和/或MFRA盒166之前。

[0110] 在一些示例中,诸如视频文件150的分段可以包括在FTYP盒152之前的MPD更新盒

(未示出)。MPD更新盒可以包括指示与包括视频文件150的表示相对应的MPD要被更新的信息以及用于更新MPD的信息。例如,MPD更新盒可以提供用于要用于更新MPD的资源的URI或URL。作为另一示例,MPD更新盒可以包括用于更新MPD的数据。在一些示例中,MPD更新盒可以紧跟在视频文件150的分段类型(STYP)盒(未示出)之后,其中STYP盒可以定义用于视频文件150的分段类型。

[0111] 在图4的示例中,MOOV盒154包括电影报头(MVHD)盒156、轨道(TRAK)盒158和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可以包括描述视频文件150最初何时被创建、视频文件150最近何时被修改、用于视频文件150的时间标度、用于视频文件150的回放的持续时间的数据、或者通常描述视频文件150的其它数据。

[0112] TRAK盒158可以包括用于视频文件150的轨道的数据。TRAK盒158可以包括描述与TRAK盒158相对应的轨道的特性的轨道报头(TKHD)盒。在一些示例中,TRAK盒158可以包括经译码的视频图片,而在其它示例中,轨道的经译码的视频图片可以被包括在电影片段164中,电影片段164可以通过TRAK盒158和/或sidx盒162的数据来引用。

[0113] 在一些示例中,视频文件150可以包括多于一个轨道。因此,MOOV盒154可以包括一数量的TRAK盒,TRAK盒的数量等于视频文件150中的轨道的数量。TRAK盒158可以描述视频文件150的对应轨道的特性。例如,TRAK盒158可以描述用于对应轨道的时间信息和/或空间信息。当封装单元30(图3)在诸如视频文件150的视频文件中包括参数集轨道时,类似于MOOV盒154的TRAK盒158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盒内用信号通知序列级别SEI消息在参数集轨道中的存在。

[0114] MVEX盒160可以描述对应的电影片段164的特性,例如,以用信号通知除了被包括在MOOV盒154内的视频数据(如果有的话)之外,视频文件150包括电影片段164。在流式传输视频数据的上下文中,经译码的视频图片可以被包括在电影片段164中,而不是在MOOV盒154中。因此,所有经译码的视频样本可以被包括在电影片段164中,而不是在MOOV盒154中。

[0115] MOOV盒154可以包括一数量的MVEX盒160,MVEX盒160的数量等于视频文件150中的电影片段164的数量。MVEX盒160中的每一者可以描述电影片段164中的对应的电影片段的特性。例如,每个MVEX盒可以包括电影扩展报头盒(MEHD)盒,其描述用于电影片段164中的对应的电影片段的时间上的持续时间。

[0116] 如上所述,封装单元30可以将序列数据集存储在不包括实际经译码的视频数据的视频样本中。视频样本通常可以对应于存取单元,存取单元是在特定时间实例处的经译码的图片的表示。在AVC的上下文中,经译码的图片包括一个或多个VCL NAL单元和其它相关联的非VCL NAL单元(诸如SEI消息),该一个或多个VCL NAL单元包括构造存取单元的所有像素的信息。相应地,封装单元30可以在电影片段164中的一个电影片段中包括序列数据集,序列数据集可以包括序列级别SEI消息。封装单元30还可以在MVEX盒160中的与电影片段164中的一个电影片段164相对应的MVEX盒160内将序列数据集和/或序列级别SEI消息的存在用信号通知为存在于电影片段164中的该电影片段164中。

[0117] SIDX盒162是视频文件150的可选元素。也就是说,符合3GPP文件格式或其它这样的文件格式的视频文件不一定包括SIDX盒162。根据3GPP文件格式的示例,SIDX盒可以用于标识分段(例如,被包含在视频文件150内的分段)的子分段。3GPP文件格式将子分段定义为

“具有对应媒体数据盒的一个或多个连续电影片段盒的自包含集,并且包含由电影片段盒引用的数据的媒体数据盒必须跟随在该电影片段盒之后并且在包含关于相同轨道的信息的下一电影片段盒之前。”3GPP文件格式还指示SIDX盒“包含对由该盒记载的(子)分段的子分段的引用的序列。所引用的子分段在呈现时间上是连续的。类似地,由分段索引盒引用的字节在分段内始终是连续的。所引用的大小给出了在所引用的材料中的字节的数量的计数。”

[0118] SIDX盒162通常提供表示在视频文件150中包括的分段的一个或多个子分段的信息。例如,此类信息可以包括子分段开始和/或结束的回放时间、针对子分段的字节偏移、子分段是否包括流存取点(SAP)(例如,从SAP开始)、用于SAP的类型(例如,SAP是即时解码器刷新(IDR)图片、干净随机存取(CRA)图片、断链存取(BLA)图片等)、SAP在子分段中的位置(依据回放时间和/或字节偏移)等。

[0119] 电影片段164可以包括一个或多个经译码的视频图片。在一些示例中,电影片段164可以包括一个或多个图片组(GOP),其中的每个图片组可以包括数个经译码的视频图片,例如,帧或图片。另外,如上所述,在一些示例中,电影片段164可以包括序列数据集。电影片段164中的每个电影片段可以包括电影片段报头盒(MFHD,在图4中未示出)。MFHD盒可以描述对应电影片段的特性,诸如用于电影片段的序列号。电影片段164可以按照序列号的顺序被包括在视频文件150中。

[0120] MFRA盒166可以描述在视频文件150的电影片段164内的随机存取点。这可以辅助执行特技模式,诸如对在通过视频文件150封装的分段内的特定时间位置(即,回放时间)执行搜索。MFRA盒166通常是可选的,并且在一些示例中不需要被包括在视频文件中。同样,客户端设备(诸如客户端设备40)不一定需要引用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可以包括一数量的轨道片段随机存取(TFRA)盒(未示出),TFRA盒的数量等于视频文件150的轨道的数量,或者在一些示例中,等于视频文件150的媒体轨道(例如,非提示轨道)的数量。

[0121] 在一些示例中,电影片段164可以包括一个或多个流存取点(SAP),诸如IDR图片。同样,MFRA盒166可以提供对SAP在视频文件150内的位置的指示。相应地,视频文件150的时间子序列可以从视频文件150的SAP形成。时间子序列还可以包括其它图片,诸如依赖于SAP的P帧和/或B帧。可以将时间子序列的帧和/或切片布置在分段内,使得可以正确地对时间子序列的、依赖于该子序列的其它帧/切片的帧/切片进行解码。例如,在数据的分层布置中,用于针对其它数据的预测的数据还可以被包括在时间子序列中。

[0122] 在一个示例中,本公开内容如下在下面的表中定义了“组块”。该表提供了组块的基数和序数两者的示例定义:

[0123]

NL 0	基数	规范	约束	描述
styp	0/1	ISO/IEC 14496-12	DASH/CMAF 约束	分段类型 与 CMAF 组块的信令兼容性
prft	0/1	ISO/IEC 14496-12	DASH/CMAF 约束	制作者参考时间
emsg	*	ISO/IEC 23009-1	DASH/CMAF 约束	事件消息
moof	1	ISO/IEC 14496-12	DASH/CMAF 约束	电影片段盒和其包含的盒
mdat	1	ISO/IEC 14496-12	DASH/CMAF 约束	用于媒体样本的媒体数据容器

[0124] 在一个示例中,本公开内容将再同步点定义为组块的开始。此外,再同步点可以被指派以下特性:

[0125] • 其具有从分段的开始的字节偏移,该字节偏移指向组块的第一字节。

[0126] • 其具有最早呈现时间,该最早呈现时间是从moof中的信息以及可能指派给其的电影报头推导的。

[0127] • 其具有指派给其的SAP类型,如在ISO/IEC 14496-12中定义的。

[0128] • 存在对组块是否包含再同步盒的指示(上文中的styp)。

[0129] • 从再同步点开始连同在电影报头中的信息,可以进行文件格式解析和解密。

[0130] 在一些示例中,可以定义再同步标记盒,该再同步标记盒通过针对标记扫描字节流来实现与分段(例如,视频文件150)的开始的同步。该再同步盒可以具有以下特性:

[0131] • 其定义了用于再同步的具有非常高的可能性的独特模式。

[0132] • 其定义了SAP类型。

[0133] 再同步标记盒可以是新盒,或者其可以重用现有盒,诸如styp盒。在本公开内容中,假设具有特定限制的styp可以用作再同步标记盒。目前正在对这种方法的稳健性进行研究。

[0134] 下面的图5-图7用于描述若干用例,对于这些用例,在除了分段的开始以外的点处对DASH分段的随机存取可以是有用的。

[0135] 图5是示出在根据本公开内容的可以第一用例中使用的示例低时延架构200的概念图。也就是说,图5示出了用于根据DASH-IF IOP来操作低时延DASH服务的基本信息流。低时延架构200包括DASH打包器202、编码器216、内容递送网络(CDN)220、常规DASH客户端230和低时延DASH客户端232。编码器216通常可以对应于图1的音频编码器26和视频编码器28中的任一者或两者,而DASH打包器202可以对应于图1的封装单元30。

[0136] 在该示例中,编码器216对所接收的媒体数据进行编码以形成CMAF报头(CH)(诸如CH 208)、CMAF初始组块206A、206B(CIC 206)和CMAF非初始组块204A-204D(CNC 204)。编码器216向DASH打包器202提供CH 208、CIC 206和CNC 204。DASH打包器202还接收服务描述,其包括关于服务的一般描述和编码器216的编码器配置的信息。

[0137] DASH打包器202使用服务描述、CH 208、CIC 206和CNC 204来形成媒体呈现描述

(MPD)210和初始化分段212。DASH打包器202还产生将CH 208、CIC 206和CNC 204映射到分段214A、214B(分段214)中,并且以增量方式将分段214提供给CDN 220。随着分段214被生成,DASH打包器202可以以组块的形式递送分段214。CDN 220包括用于存储MPD 210、IS 212和分段214的分段存储222。例如,响应于来自常规DASH客户端230和低时延DASH客户端232的HTTP Get或部分Get请求,CDN 220向常规DASH客户端230递送完整分段,但是向低时延DASH客户端232递送个体组块(例如,CH 208、CIC 206和CNC 204)。

[0138] 图6是进一步详细地示出关于图5描述的用例的示例的概念图。图6的示例示出了分段250A-250E(分段250),其包括组块252A-252E(组块252)的相应集合。客户端设备(诸如图1的客户端设备40)可以检索完整分段250或者个体组块252。例如,如图5所示,常规DASH客户端230可以检索分段250,而低时延DASH客户端232可以检索个体组块252(至少最初)。

[0139] 图6进一步描绘了可以如何通过检索个体组块252而不是完整分段250来减少时延。例如,在当前时间处检索完整分段可能导致较高的时延。简单地检索最新的完全可用的分段可以减少时延,但是仍然产生相对高的时延。

[0140] 替代地通过检索组块,可以大大地减少这些时延。例如,在如图6中的“现在”所指示的当前时间处,未完全形成分段250E。然而,即使在完全形成分段250E之前,客户端设备也可以检索所形成的组块,诸如分段250E的组块252E-1和252E-2,假设组块252E-1和252E-2被形成并且可供检索。

[0141] 当加入实时流时,通常应当实现低时延和快速启动两者。然而,这并非小事,并且下文基于图6讨论了一些策略:

[0142] • 在第一种情况下,在实时边缘处,将作为在时间历史中落后3个分段的分段(即,分段250B、250C和250D)加载到缓冲器中。一旦一个分段可用,回放就开始。这导致显著的时延,但是回放可以相对快地开始,这是因为加载了在分段的开始处的随机存取。

[0143] • 在第二种情况下,选择最近的可用分段(分段250D),而不是三分段旧分段。在这种情况下,回放时延至少是分段持续时间,但是可以更高。启动可能与以上情况类似。

[0144] • 在其它三种情况下,包括多个组块的分段(例如,分段250E)在仍然被产生的同时被回放。这减少了时延,但是存在回放的开始可能受到影响的问题,尤其是如果最近发布的分段的分段可用性开始时间与挂钟时间的差大于目标时延。在这种情况下,客户端设备可能必须进行等待,直到接下来的第二分段发布为止。在6秒分段的情况下,这可能导致4-5秒的启动时延。

[0145] • 存在其它技术和用例。例如,客户端可以在开始处存取旧分段,下载所有内容,并且加速回放并且进行快进解码。然而,这样的方法具有如下的缺点:在可以发生加速解码之前,需要下载大量数据。此外,它在解码器接口中没有得到广泛支持。

[0146] 合适的解决方案可以是:

[0147] • 适配集的至少一个表示在分段/片段中包含更频繁的随机存取点和非初始组块。

[0148] • DASH客户端可以使用来自MPD的信息来确定存在这样的随机存取方法,但是可能不准确地用信号通知随机存取点的位置/字节偏移。

[0149] • DASH客户端可以在启动时存取这种表示,但是只能从最近可用的非初始组块的字节范围开始下载,或至少接近该范围。

[0150] • 一旦下载,DASH客户端就可以确定随机存取点,并且开始处理数据以及相同表示的同样下载的初始化分段/CMAF报头。下文讨论随机存取点的定位。

[0151] 然而,后一种方法可能遇到各种问题,如下文概述的。

[0152] 因此,如在图6的示例中所示以及如下文所讨论的,使用如在本公开内容中讨论的组块可以大幅减少时延。提前用信号通知组块的开头可以允许提前产生清单文件,该清单文件不需要频繁更新,但是仍然能够指示在组块内的流存取点(SAP)的大致位置。以这种方式,客户端设备能够使用清单文件来确定组块边界的位置,而不需要连续的清单文件更新,并且同时仍然允许客户端设备在组块边界的开头(例如,再同步点)处发起媒体流式传输。也就是说,客户端设备可以甚至在已经完全形成分段之前,从清单文件确定包括再同步点的分段的字节范围,这是因为清单文件可以用信号通知字节范围或表示在该分段中的再同步点的大致位置的其它数据。

[0153] 图7是示出在广播协议的上下文中使用DASH和CMAF随机存取的示例第二用例的概念图。图7示出了包括媒体编码器280、CMAF/文件格式(FF)打包器282、DASH打包器284、ROUTE发送器286、CDN源服务器288、ROUTE接收器290、DASH客户端292、CMAF/FF解析器294和媒体解码器296的示例。媒体编码器280对媒体数据(诸如音频数据或视频数据)进行编码。媒体编码器280可以对应于图1的音频编码器26或视频编码器28、或图5的编码器216。媒体编码器280向CMAF/FF打包器282提供经编码的媒体数据,该CMAF/FF打包器282根据CMAF和特定文件格式(诸如ISO BMFF或其扩展)将经编码的媒体数据格式化为文件。

[0154] CMAF/FF打包器282将这些文件(例如,组块)提供给DASH打包器284,DASH打包器284将这些文件/组块聚合为DASH分段。DASH打包器284还可以形成清单文件(诸如MPD),其包括描述文件/组块/分段的数据。此外,根据本公开内容的技术,DASH打包器284可以确定未来的流存取点(SAP)或随机存取点(RAP)的近似位置,并且在MPD中用信号通知该近似位置。CMAF/FF打包器282和DASH打包器284可以对应于图1的封装单元30或图5的DASH打包器202。

[0155] DASH打包器284向ROUTE发送器286和CDN源服务器288提供分段以及MPD。ROUTE发送器286和CDN源服务器288可以对应于图1的服务器设备60或图5的CDN 220。通常,在该示例中,ROUTE发送器286可以根据ROUTE来向ROUTE接收器290发送媒体数据。在其它示例中,其它基于文件的递送协议可以用于广播或多播,诸如FLUTE。另外或替代地,CDN源服务器288可以例如根据HTTP来向ROUTE接收器290发送媒体数据和/或直接地向DASH客户端292发送媒体数据。

[0156] ROUTE接收器290可以在中间件(诸如图2的eMBMS中间件单元100)中实现。ROUTE接收器290可以例如在如图2所示的高速缓存104中缓冲所接收的媒体数据。DASH客户端292(其可以对应于图2的DASH客户端110)可以使用HTTP从ROUTE接收器290检索经高速缓存的媒体数据。替代地,DASH客户端292可以根据HTTP直接从CDN源服务器288检索媒体数据,如上文所讨论的。

[0157] 此外,根据本公开内容的技术,DASH客户端292可以使用清单文件(诸如MPD)来确定例如跟随在清单文件中用信号通知的再同步点之后的SAP或RAP的位置。DASH客户端292可以从下一个最快的再同步点开始发起对媒体呈现的检索。再同步点通常可以指示比特流的能够正确地解析文件容器级别数据的位置。因此,DASH客户端292可以在再同步点处开始

发起流式传输,并且将从再同步点开始接收的媒体数据递送到CMAF/FF解析器294。

[0158] CMAF/FF解析器294可以开始对从再同步点开始的媒体数据的解析。CMAF/FF解析器294可以对应于图1的解封装单元50。此外,CMAF/FF解析器294可以从经解析的数据中提取可解码的媒体数据,并且将可解码的媒体数据递送到媒体解码器296,媒体解码器296可以对应于图1的音频解码器46或视频解码器48。媒体解码器296可以对媒体数据进行解码,并且将经解码的媒体数据递送到对应的输出设备,诸如图1的音频输出42或视频输出44。

[0159] 在广播的情况下,在图7中示出了用于DASH/CMAF和ROUTE的组合的示例。在低时延DASH模式与ROUTE的组合中(如例如针对ABR多播上的DVB TM-IPI任务组以及针对ATSC简档所考虑的),可能出现以下问题。如果ROUTE接收器290在DASH/CMAF低时延分段的中间加入,则其不能开始处理数据,这是因为没有同步是可用的,也不存在用于其它目的的任何随机存取。因此,即使将在分段的中间提供更频繁的随机存取,启动也会被延迟。

[0160] 合适的解决方案可以是:

[0161] • 广播/多播表示在分段/片段中包含更频繁的随机存取点和非初始组块。

[0162] • DASH客户端292使用MPD的信息和/或可能通过来自ROUTE接收器290的信息来确定存在这样的随机存取方法。DASH客户端292可以使用这样的信息来精确地定位随机存取点,但也可能不是这样。

[0163] • DASH客户端292可以在启动时存取该表示,但是可能无法从开始存取所有信息。

[0164] • 一旦开始存取分段的所接收的部分,DASH客户端292可以找到随机存取点,并且开始处理数据以及相同表示的同样下载的初始化非分段/CMAF报头。下文讨论了对随机存取点的定位。

[0165] 然而,后一种方法可能遇到如下文概述的各种问题。

[0166] 在如上文在第二用例中讨论的类似情况下,分组丢失也可能是有问题的,不仅仅是在随机存取再同步时。在该示例第三用例中,如上文讨论的类似过程可以适用。除此之外,甚至可以是如下情况:不仅尝试干净随机存取,而且在充分盒解析是可能的之后,可以尝试非随机存取组块(例如,没有IDR帧)、解码和呈现的事件。因此,不仅再同步到干净随机存取很重要,而且随机存取以进行文件格式解析也很重要。

[0167] 如果实时媒体内容通常以低时延分发,但是然后在时间平移中使用相同的媒体内容进行延迟回放,则可能发生另一第四用例。客户端可能希望在特定时间处存取媒体呈现,但是该时间可能不(并且通常将不会)与分段/CMAF片段开始重合。

[0168] 合适的解决方案可以是:

[0169] • 适配集的至少一个表示可以在分段/片段中包含更频繁的随机存取点和非初始组块。

[0170] • DASH客户端292可以使用来自MPD的信息来确定存在这样的随机存取方法,但是随机存取点的位置/字节偏移可能不是精确地已知的。

[0171] • DASH客户端292可以在寻求时存取该表示,但是可以仅允许从最近可用的非初始组块的字节范围或至少接近该字节范围开始下载。

[0172] • 一旦下载,DASH客户端292就可以找到随机存取点,并且开始处理数据以及相同表示的同样下载的初始化分段/CMAF报头。下文讨论对随机存取点的定位。

[0173] 然而,后一种方法可能遇到如下文概述的各种问题。

[0174] 在ISO BMFF/DASH/CMAF分段的情况下的再同步通常涉及如下文概述的多个过程：

[0175] 1)找到盒结构。

[0176] 2)找到具有所有相关信息的CMAF组块/片段。

[0177] 3)通过mdat和tfdt找到定时。

[0178] 4)获得所有解密相关信息(如果适用的话)。

[0179] 5)可能处理事件消息。

[0180] 6)在基本流级别上开始解码。

[0181] 下文概述了在特定时间处在盒结构中找到再同步点的示例方式：

[0182] • 如果存在分段索引(SIDX盒)，则提供这样的再同步点作为呈现时间和字节偏移。然而，由于没有提前完全形成分段，因此分段索引通常不可用于低时延实时。

[0183] • 如果分段的开始是可用的，则客户端可以下载字节范围的最小集合，使得能够处理盒结构。

[0184] • 再同步由底层协议提供，该底层协议例如规定用信号通知组块的边界并且客户端能够开始解析。

[0185] • 如果分段的开始不能通过用信号通知的数据容易地确定，则客户端可以找到允许客户端随机地存取数据的同步模式。然后，客户端可以开始解析并且找到允许处理的适当的盒结构，诸如例如emsg、prft、mdat、moof和/或mdat。

[0186] 本公开内容描述了可以适用于以上第四示例的技术。前三者表示在对应信息可用的情况下的示例简化。

[0187] 基于以上讨论，本公开内容认识到以下问题，并且这些问题需要解决方案：

[0188] 1)在DASH/CMAF分段中添加额外的随机存取点。随机存取可以包括干净随机存取以及开放的或逐步的解码器刷新，一直到仅提供关于文件格式解析的再同步。

[0189] 2)在MPD(或其它清单文件)中添加适当的信令，其指示在每个DASH分段中的随机存取点和再同步的可用性以及提供关于随机存取点的位置、类型和定时的信息。信息可以是精确的，或者可能在一范围内。

[0190] 3)在任意起点的情况下，通过找到再同步点来再同步以解封装、解密和解码的能力。

[0191] 4)在如例如在基于HTML-5/MSE的回放中可用的受限的接收器环境中开始处理的能力。

[0192] 图8是示出清单文件中的流存取点(SAP)的示例信令的概念图。具体而言，图8示出了包括SAP 302A-302D(SAP 302)和分段304A-304D(分段304)的比特流300、以及包括SAP 312A-312D(SAP 312)、SAP 316A-316D(SAP 316)和分段314A-314D(分段314)的比特流310。也就是说，在该示例中，比特流310的分段314包括与比特流300的分段304相比更频繁的SAP 312、SAP 316。SAP 302、SAP 312中的每一者可以对应于分段304、分段314中的对应分段的开始以及这些分段的第一组块。SAP 316可以对应于在对应分段316内(但不是在对应分段316的开头处)的组块的开始。

[0193] 为了提供用于实现恒定比特率表示的简单技术(具有等距为1000个样本的组块(并且@timescale＝1000，以样本持续时间为单位)和SAP类型1(这可以例如是音频表示))，可以添加再同步元素，其中：

[0194] • @type＝1

[0195] • @dT＝1000

[0196] • @dImin＝100

[0197] • @dImax＝100

[0198] 接收这样的信息的客户端(例如,图1的客户端设备40)可能不能针对@duration＝10000的分段识别能够在精确的字节范围内每秒存取随机存取点。如果比特率是可变的,则接收器(例如,客户端设备40)可以使用@dIMin和@dIMax来识别要在其中找到随机存取点的范围。作为用信号通知最大的@dT的替代,它还可以用信号通知标称组块持续时间。

[0199] 清单文件的再同步元素还可以包括URL@index,其使用与常规分段相同的模板函数指向每个分段中的再同步点的二进制再同步索引。如果存在,该再同步可以以与分段索引类似的方式提供分段中的所有再同步点的准确位置。如果该索引存在,则再同步索引可以可用于在清单文件/MPD的发布时间处可用的时段的所有分段。

[0200] 在一种方法中,再同步索引可以与分段索引相同,但是可以改变。

[0201] 图1的客户端设备40可以使用ISO BMFF 4字符盒类型作为再同步到媒体文件(例如,图4的视频文件150,其可以是分段)中的基础。在一个示例中,所选择的盒类型是“styp”盒,但是它还可以是“moof”盒本身。盒字符串类型的随机仿真非常罕见。下文描述了styp仿真的测试报告。然后通过针对已知的预期盒类型进行检查来避免这种仿真。客户端设备40可以执行如下概括的再同步机制:

[0202] 1)找到分段中的“styp”字节字符串例如在字节偏移B1处的出现。

[0203] 2)如下针对随机仿真进行验证:将下一个盒类型与预期盒类型列表进行比较:“styp”、“sidx”、“ssix”、“prft”、“moof”、“mdat”、“free”、“mfra”、“skip”、“meta”、“meco”。

[0204] a.如果找到已知盒类型中的一种盒类型,则字节偏移B1-4字节是再同步点的字节偏移。

[0205] b.如果这不是之前提到的已知盒类型中的一种盒类型,则styp盒的出现被视为无效同步点并且被忽略。从上面的步骤1重新开始。

[0206] 对来自被扫描的DASH-IF测试资产的30282个分段测试了本公开内容的技术。该扫描揭示了“styp”字符串在文件中的28408次出现,并且这28408次出现中的仅10次出现(2840次出现中大约1次)是仿真,一旦确定以下盒不是预期盒类型中的一种盒类型,该仿真就被丢弃:“styp”、“sidx”、“ssix”、“prft”、“moof”、“mdat”、“free”、“mfra”、“skip”、“meta”、“meco”。

[0207] 基于这些结果,认为使用styp再同步点检测以及组块结构是足够的。仅限于可以跟随在styp之后的盒子集(诸如prft、emsg、free、skip和moof)将是合适的。

[0208] 剩下的问题是确定SAP类型和最早呈现时间。后者通过使用tfdt和电影片段报头中的其它信息而容易地实现。记录该算法将是合适的。

[0209] 存在如下文描述的用于确定SAP类型的若干选项:

[0210] • 基于moof中的信息进行检测。可以记录并且执行简单技术。

[0211] • 使用用于SAP类型的兼容性品牌。已经通过使用CMAF,可以得出以下各项:

[0212] ocmff:指示SAP是1或2

[0213] ocmfl:指示SAP是0(对于加密而言是正确的?)

[0214] ocmfr::指示SAP是1、2或3

[0215] • 如果持续地使用,该信令可以是足够的。可以定义用于其它SAP类型的兼容性品牌。

[0216] • 其它技术可以用于指示SAP类型。

[0217] 现有选项可以用于确定SAP类型。

[0218] 以这种方式,本公开内容的技术可以概述如下,并且可以由诸如图1的内容准备设备20、服务器设备60和/或客户端设备40的设备来执行,如上文所讨论的:

[0219] 在DASH上下文中,在某些情况下,分段被视为用于下载、存取媒体呈现的单个单元以及还通过被寻址的URL。然而,分段可以被构造成在容器级别上实现再同步以及甚至在分段内对相应的表示的随机存取。再同步机制是通过再同步元素来支持的并且用信号通知的。

[0220] 再同步元素用信号通知分段中的再同步点。再同步点是组块的开始(以字节位置为单位),其中,组块被定义为在分段内的包含特定呈现持续时间的媒体数据的结构化连续字节范围,并且可以是在容器格式上独立地存取(包括潜在的解密)的。分段中的再同步点可以定义如下:

[0221] • 再同步点是组块的开始。

[0222] • 此外,再同步点已经指派以下特性:

[0223] o它具有指向组块的第一字节的从分段的开始的字节偏移或索引值。

[0224] o它在表示中具有指派的最早呈现时间。

[0225] o它具有例如通过ISO/IEC 14496-12中的SAP类型定义的指派的SAP类型。

[0226] o它已经指派了标记特性,其指示是否能够在通过特定标记解析分段时检测到再同步点,或者是否需要通过外部单元用信号通知再同步点。

[0227] • 从再同步点开始处理以及初始化分段中的信息(如果存在的话)允许容器解析和解密。是否以及如何存取所包含的基本视频流的能力是通过SAP类型定义的。

[0228] 由于因果原因,用信号通知MPD中的每个再同步点可能是困难的,这是因为再同步点可以是由分段打包器独立于MPD更新而添加的。例如,再同步点可以是由编码器和打包器独立于MPD而生成的。此外,在低时延情况下,MPD信令可能对于DASH客户端(例如,图2的DASH客户端110或图7的DASH客户端292)是不可用的。因此,存在两种方式来用信号通知在MPD中的分段中提供的再同步点:

[0229] • 通过在用于每个分段的再同步索引分段中提供用于再同步点的二进制映射。这最容易用于在网络上完全可用的分段。

[0230] • 通过用信号通知再同步点在分段中的存在以及还有一些额外信息,这些信息允许依据字节位置和呈现时间容易地定位再同步点。

[0231] 为了用信号通知以上特性,再同步元素具有不同的属性,这些属性在DASH规范的第5.3.12.2条中进行了更详细的解释。

[0232] 随机存取指代通过利用初始化分段(如果存在的话)对表示进行初始化并且从用信号通知的分段向前解码和呈现表示,来从时间t处的随机存取点开始向前处理、解码和呈现表示。随机存取点可以是利用RandomAccess元素用信号通知的,如在下面的表10中所定义的。

[0233] 表10—随机存取信令

元素或属性名称	使用	描述
ResyncPoint		再同步点信令
@type	OD 默认: "0"	如果大于 0，则指定在分段中存在一个或多个再同步点，其中 SAP 类型等于或小于在该值中指示的 SAP 类型。 如果为 0，则指定在具有任何 SAP 类型的分段中存在再同步点。 对于关于 SAP 类型的细节，参照 DASH 的第 4.5.2 条。
@dT	O	指定任何两个连续再同步点的时间值的最大差，这两个再同步点被包括在用于该表示的任何媒体分段中的上述@type 定义中。 该值以对应表示的@timescale 的尺度来表达。 如果不存在，则该值是未知的。
@dImax	O	指定任何两个连续再同步点的索引值的最大差（通过表示的@bandwidth 值进行归一化的），这两个再同步点被包括在用于该表示的任何媒体分段中的上述@type 定义中。 为了获得最大差（以字节为单位），@dImax 的值乘以@bandwidth 值。 如果不存在，则该值是未知的。

[0234]

元素或属性名称	使用	描述
@dImin	OD 默认: 0	指定任何两个连续再同步点的索引值的最小差（通过表示的@bandwidth 值进行归一化的），这两个再同步点被包括在用于该表示的任何媒体分段中的上述@type 定义中。 为了获得最小差（以字节为单位），@dImax 的值乘以@bandwidth 值。 如果不存在，则假设该值为 0。
@marker	OD 默认: FALSE	如果被设置为 TRUE，则它指定每个再同步点包括用于再同步点检测的特定标记。 如果不存在或为 FALSE，则无法预期用于每个用信号通知的再同步点的标记的存在。
@index	OD	指定用于为所定义的@type 创建再同步点分段列表的模板。 如果存在，则再同步点索引分段应该可用于每个对应的可用媒体分段。 对于更多细节，参照 DASH 的第 5.3.9.4.4 款。

[0235]

[0236] 表11提供了不同的随机存取点类型。

[0237] 表11

[0238]

```
<!-- Resync Point  -->
<xs:complexType name="ResyncType">
    <xs:attribute name="type" type="SAPType" default="0"/>
    <xs:attribute name="dT" type="xs:unsignedInt"/>
    <xs:attribute name="dImax" type="xs:float"/>
    <xs:attribute name="dImin" type="xs:float" default="0"/>
    <xs:attribute name="marker" type="xs:boolean" default="false"/>
    <xs:attribute name="index" type="xs:anyURI"/>
    <xs:anyAttribute namespace="##other" processContents="lax"/>
</xs:complexType>
```

[0239] 再同步索引分段包含与媒体分段相关的信息。再同步索引分段以与分段索引类似的方式提供分段中的所有再同步点的准确位置。在DASH规范的第5.3.12.1条中定义了再同步点。

[0240] 用于ISO BMFF的再同步点可以被定义为ISO BMFF分段的开始,在基数和序数两者方面具有以下限制:

[0241]

NL 0	基数	规范	约束	描述

[0242]

NL 0	基数	规范	约束	描述
styp	0/1	ISO/IEC 14496-12	DASH/CMAF 约束	分段类型 对CMAF组块的信令兼容性
prft	0/1	ISO/IEC 14496-12	DASH/CMAF约束	制作者参考时间
emsg	*	ISO/IEC 23009-1	DASH/CMAF约束	事件消息
moof	1	ISO/IEC 14496-12	DASH/CMAF约束	电影片段盒以及它包含的盒
mdat	1	ISO/IEC 14496-12	DASH/CMAF约束	用于媒体样本的媒体数据容器

[0243] 对于基于ISO BMFF的再同步点,特性可以定义如下:

[0244] • 索引Index被定义为上述受约束的ISO BMFF分段的第一字节的偏移。

[0245] • 最早呈现时间Time被定义为组块中的任何样本的解码时间、合成偏移和编辑列表的组合的最小时间。

[0246] • SAP类型是根据DASH规范的第4.5.2条来定义的。

[0247] • 如果styp以“cmfl”作为主要兼容性品牌来呈现,则存在标记。

[0248] 再同步索引分段可以对一个表示的一个媒体分段进行索引,并且可以定义如下:

[0249] • 每个表示索引分段应当以“styp”盒开始,并且品牌“risg”应当存在于“styp”盒

中。品牌“risg”的合规性要求是通过该子条款定义的。

[0250] • 每个媒体分段是通过一个或多个分段索引盒进行索引的;用于给定媒体分段的盒是连续的。

[0251] 图9是示出根据本公开内容的技术的检索媒体数据的示例方法的流程图。相对于图1的客户端设备40解释了图9的方法。然而,图5的低时延DASH客户端232或包括媒体解码器296、CMAF/FF解析器294、DASH客户端292和图7的ROUTE接收器290的客户端设备还可以被配置为执行该方法或类似方法。

[0252] 最初,客户端设备40可以检索媒体呈现的清单文件(350),诸如MPD。客户端设备40可以从例如服务器设备60检索清单文件。清单文件可以包括指示媒体呈现包括在媒体呈现的表示的分段内的组块边界处的再同步点的数据。因此,客户端设备40可以确定媒体呈现的再同步点(352),例如,最近可用的再同步点。通常,再同步点可以指示组块边界的开始,组块边界是表示的能够正确地解析文件级别容器(例如,如上所讨论的数据结构,诸如盒)的随机可存取的点。

[0253] 具体地,清单文件可以指示再同步点的位置,诸如从分段的开头的字节偏移。该信息可能不会准确地标识再同步点在分段中的位置,但是可以保证再同步点将在从字节偏移的字节范围内是可用的。因此,客户端设备40可以形成指定所指示的字节偏移以在再同步点处开始检索的请求(诸如HTTP部分Get请求)(354)。然后,客户端设备40可以向服务器设备60发送该请求(356)。

[0254] 响应于该请求,客户端设备40可以接收所请求的媒体数据(358),其包括再同步点。如上所述,字节偏移可能不会精确地标识再同步点的位置,并且因此,客户端设备40可以解析数据,直到检测到再同步点的实际位置为止。客户端设备40可以从再同步点开始,解析文件级别数据结构(诸如文件格式盒),以确定所检索到的媒体数据的对应组块的媒体数据的位置。具体地,客户端设备40可以通过检测例如分段类型值、制作者参考时间值、事件消息、电影片段和媒体数据容器盒来将再同步点识别为组块的开始。电影片段可以包括经编码的媒体数据。

[0255] 解封装单元50可以例如从电影片段中提取对应组块的经编码的媒体数据(360),并且将经编码的媒体数据提供给例如视频解码器48。组块可以从随机存取点(RAP)(诸如视频数据的经帧内预测的帧(I帧))开始。清单文件还可以指示RAP是闭合图片组(GOP)的开始还是开放GOP的开始,并且由此指示能够从RAP开始执行的随机存取的类型(例如,I帧的前导图片是可解码的还是不可解码的)。视频解码器48可以进而对经编码的媒体数据进行解码(362),并且将经解码的媒体数据发送给例如视频输出44,以呈现媒体数据(364)。

[0256] 以这种方式,图9的方法表示检索媒体数据的方法的示例,该方法包括:检索用于媒体呈现的清单文件,该清单文件指示能够在媒体呈现的表示的分段的再同步点处开始对比特流的媒体数据的容器解析,再同步点位于除了分段的开始以外的位置并且表示能够开始对比特流的媒体数据的容器解析的点;使用清单文件,来形成用于在再同步点处开始检索表示的媒体数据的请求;发送该请求以在再同步点处开始发起对媒体呈现的媒体数据的检索;以及呈现所检索到的媒体数据。

[0257] 在以下示例中概述了本公开内容的某些技术:

[0258] 示例1:一种检索媒体数据的方法,所述方法包括:检索用于媒体呈现的清单文件,

所述清单文件指示能够在所述媒体呈现的表示的再同步点处开始再同步和解密;在所述再同步点处开始检索所述表示的媒体数据;以及呈现所检索到的媒体数据。

[0259] 示例2:根据示例1所述的方法,其中,所述再同步点包括组块边界的开始。

[0260] 示例3:根据示例2所述的方法,其中,所述组块边界包括组块的开始,所述组块包括零或一分段类型值、零或一制作者参考时间值、零或更多事件消息、至少一个电影片段盒和至少一个媒体数据容器盒。

[0261] 示例4:根据示例1-3中任一项所述的方法,其中,所述清单文件指示在所述表示的分段中的所述再同步点的可用性。

[0262] 示例5:根据示例4所述的方法,其中,所述再同步点位于除了所述分段的开头以外的位置。

[0263] 示例6:根据示例4和示例5中任一项所述的方法,其中,所述清单文件指示能够在所述再同步点处执行的随机存取的类型。

[0264] 示例7:根据示例4-6中任一项所述的方法,其中,所述清单文件指示所述再同步点的位置和定时以及所述位置和定时信息是准确的还是估计的。

[0265] 示例8:根据示例1-7中任一项所述的方法,其中,所述清单文件包括媒体呈现描述(MPD)。

[0266] 示例9:一种用于检索媒体数据的设备,所述设备包括用于执行根据示例1-8中任一项所述的方法的一个或多个单元。

[0267] 示例10:根据示例9所述的设备,其中,所述一个或多个单元包括在电路中实现的一个或多个处理器和被配置为存储媒体数据的存储器。

[0268] 示例11:根据示例9所述的设备,其中,所述设备包括以下各项中的至少一项:集成电路;微处理器;或者无线通信设备。

[0269] 示例12:一种在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执行根据示例1-8中任一项所述的方法。

[0270] 示例13:一种用于检索媒体数据的设备,所述设备包括:用于检索用于媒体呈现的清单文件的单元,所述清单文件指示能够在所述媒体呈现的表示的再同步点处开始再同步和解密;用于在所述再同步点处开始检索所述表示的媒体数据的单元;以及用于呈现所检索到的媒体数据的单元。

[0271] 示例14:一种发送媒体数据的方法,所述方法包括:向客户端设备发送用于媒体呈现的清单文件,所述清单文件指示能够在所述媒体呈现的表示的再同步点处开始再同步和解密;从所述客户端设备接收针对在所述再同步点处开始的媒体数据的请求;以及响应于所述请求,来向所述客户端设备发送所述表示的在所述再同步点处开始请求的媒体数据。

[0272] 示例15:根据示例14所述的方法,还包括:生成所述清单文件。

[0273] 示例16:根据示例14和示例15中任一项所述的方法,其中,所述再同步点包括组块边界的开始。

[0274] 示例17:根据示例16所述的方法,其中,所述组块边界包括组块的开始,所述组块包括零或一分段类型值、零或一制作者参考时间值、零或更多事件消息、至少一个电影片段盒和至少一个媒体数据容器盒。

[0275] 示例18:根据示例14-17中任一项所述的方法,其中,所述清单文件指示在所述表

示的分段中的所述再同步点的可用性。

[0276] 示例19:根据示例18所述的方法,其中,所述再同步点位于除了所述分段的开头以外的位置。

[0277] 示例20:根据示例18和示例19中任一项所述的方法,其中,所述清单文件指示能够在所述再同步点处执行的随机存取的类型。

[0278] 示例21:根据示例18-20中任一项所述的方法,其中,所述清单文件指示所述再同步点的位置和定时以及所述位置和定时信息是准确的还是估计的。

[0279] 示例22:根据示例14-21中任一项所述的方法,其中,所述清单文件包括媒体呈现描述(MPD)。

[0280] 示例23:一种用于发送媒体数据的设备,所述设备包括用于执行根据示例14-22中任一项所述的方法的一个或多个单元。

[0281] 示例24:根据示例23所述的设备,其中,所述一个或多个单元包括在电路中实现的一个或多个处理器和被配置为存储媒体数据的存储器。

[0282] 示例25:根据示例23所述的设备,其中,所述设备包括以下各项中的至少一项:集成电路;微处理器;或者无线通信设备。

[0283] 示例26:一种在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执行根据示例1-8中任一项所述的方法。

[0284] 示例27:一种用于发送媒体数据的设备,所述设备包括:用于向客户端设备发送用于媒体呈现的清单文件的单元,所述清单文件指示能够在所述媒体呈现的表示的再同步点处开始再同步和解密;用于从所述客户端设备接收针对在所述再同步点处开始的媒体数据的请求的单元;以及用于响应于所述请求,来向所述客户端设备发送所述表示的在所述再同步点处开始请求的媒体数据的单元。

[0285] 在一个或多个示例中,所描述的功能可以用硬件、软件、固件或其任何组合来实现。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被存储在计算机可读介质上或者通过计算机可读介质进行发送并且由基于硬件的处理单元执行。计算机可读介质可以包括计算机可读存储介质,其对应于诸如数据存储介质的有形介质,或者包括例如根据通信协议来促进计算机程序从一个地方传送到另一个地方的任何介质的通信介质。以这种方式,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)诸如信号或载波的通信介质。数据存储介质可以是能够由一个或多个计算机或者一个或多个处理器存取以检索用于实现在本公开内容中描述的技术的指令、代码和/或数据结构的任何可用的介质。计算机程序产品可以包括计算机可读介质。

[0286] 通过举例而非限制性的方式,这样的计算机可读存储介质可以包括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、或者能够用于以指令或数据结构形式存储期望的程序代码以及能够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、光纤光缆、双绞线、数字用户线(DSL)或者无线技术(诸如红外线、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、光纤光缆、双绞线、DSL或者无线技术(诸如红外线、无线电和微波)被包括在介质的定义中。然而,应当理解的是,计算机可读存储介质和数据存储介质不包括连接、载波、信号或其它暂时性介质,而是替代地针对非暂时性的有形存储介质。如本文所使用

的,磁盘和光盘包括压缩光盘(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。上述各项的组合还应当被包括在计算机可读介质的范围之内。

[0287] 指令可以由一个或多个处理器(诸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)、或其它等效的集成或分立逻辑电路)来执行。因此,如本文所使用的术语“处理器”可以指代前述结构中的任何一者或者适于实现本文描述的技术的任何其它结构。另外,在一些方面中,本文描述的功能可以是在被配置用于编码和解码的专用硬件和/或软件模块内提供的,或者被并入经组合的编解码器中。此外,所述技术可以在一个或多个电路或逻辑元件中完全地实现。

[0288] 本公开内容的技术可以在多种多样的设备或装置(包括无线手机、集成电路(IC)或一组IC(例如,芯片组))中实现。在本公开内容中描述了各种组件、模块或单元以强调被配置为执行所公开的技术的设备的功能性方面,但是不一定需要通过不同的硬件单元来实现。确切而言,如上所述,各种单元可以被组合在编解码器硬件单元中,或者由互操作的硬件单元的集合(包括如上所述的一个或多个处理器)结合适当的软件和/或固件来提供。

[0289] 已经描述了各个示例。这些示例和其它示例在所附的权利要求的范围内。

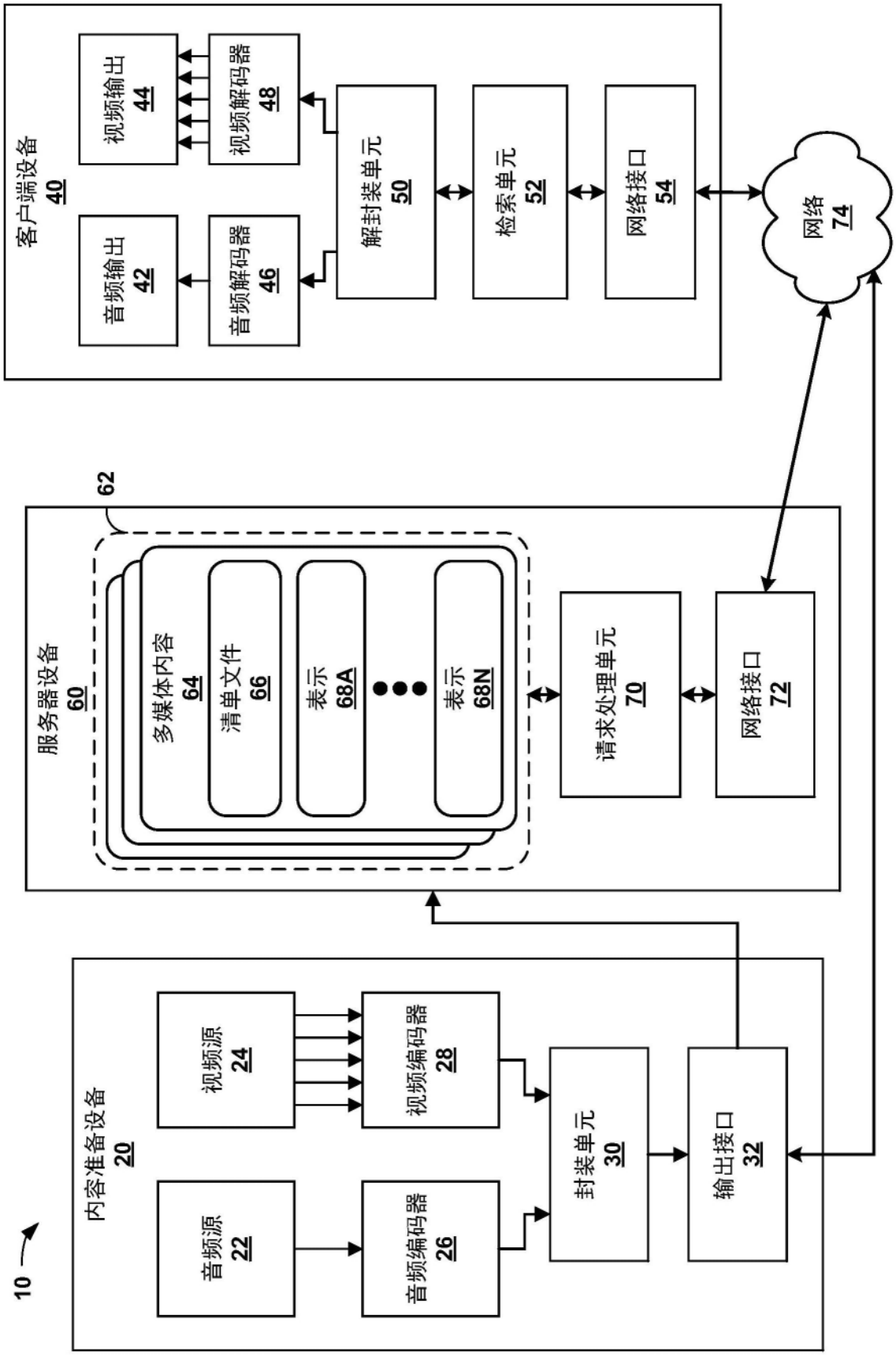

10
内容准备设备
20
音频源
22
视频源
24
音频编码器
26
视频编码器
28
封装单元
30
输出接口
32
服务器设备
60
62
多媒体内容
64
清单文件
66
表示
68A
表示
68N
请求处理单元
70
网络接口
72
网络
74
客户端设备
40
音频输出
42
视频输出
44
音频解码器
46
视频解码器
48
解封装单元
50
检索单元
52
网络接口
54

图1

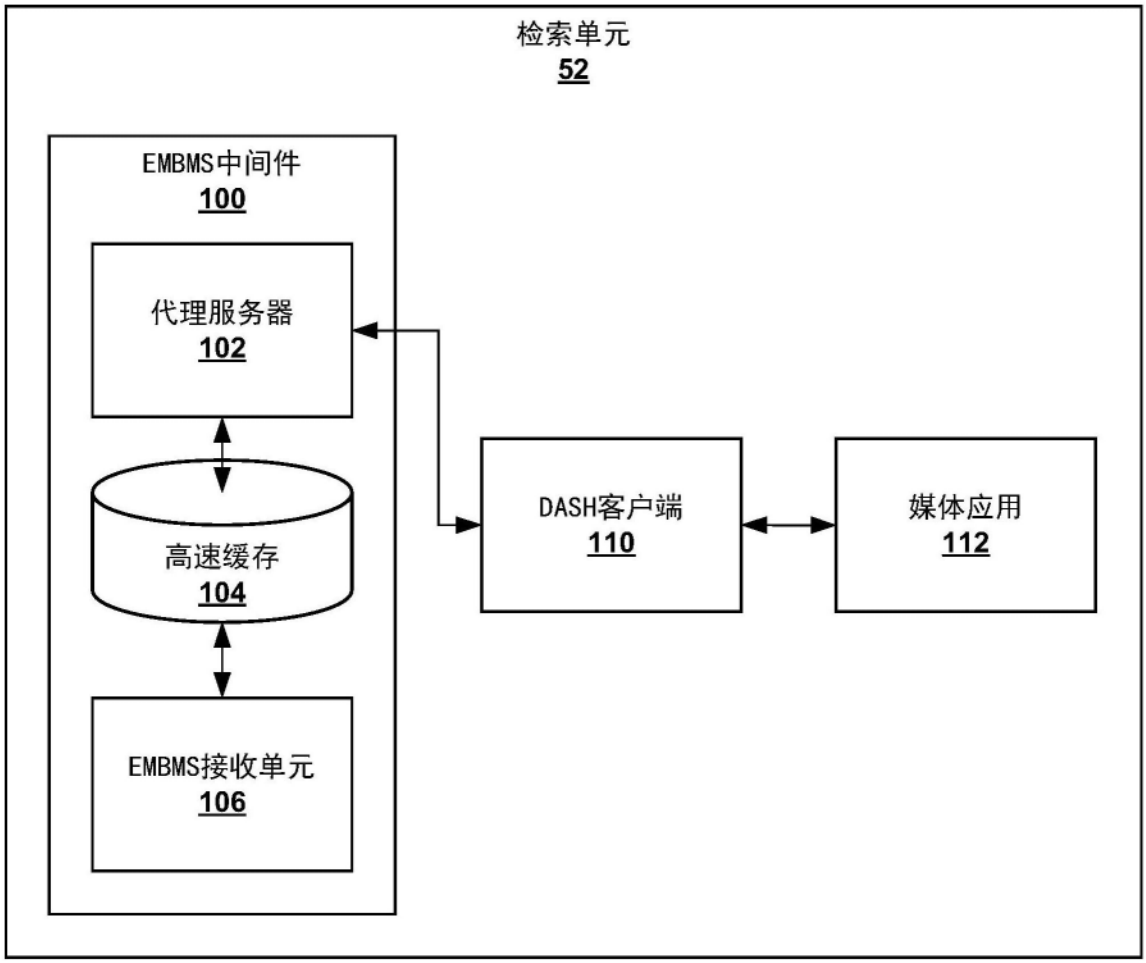

检索单元
52
EMBMS中间件
100
代理服务器
102
高速缓存
104
EMBMS接收单元
106
DASH客户端
110
媒体应用
112

图2

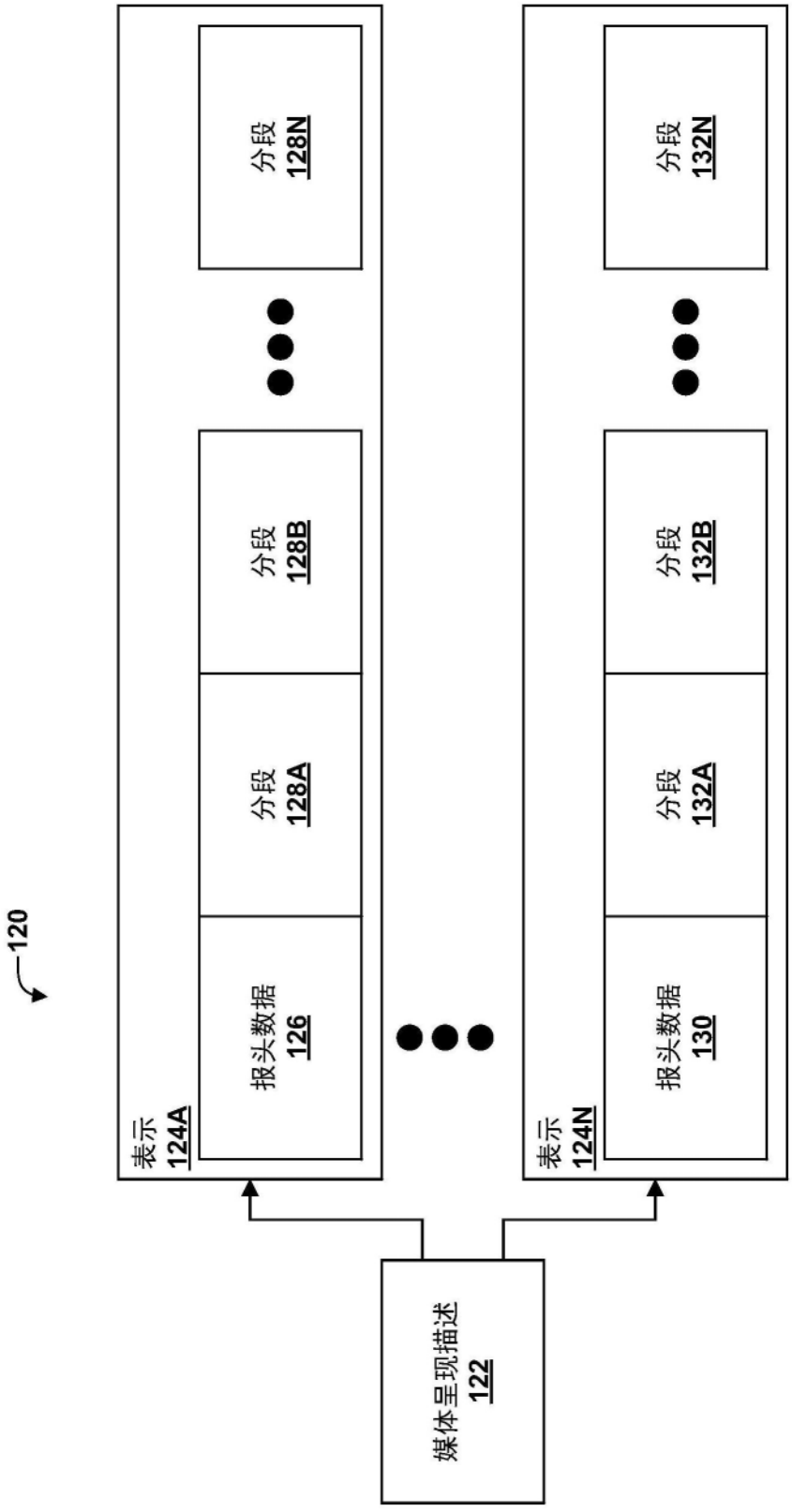

120
表示
124A
报头数据
126
分段
128A
分段
128B
分段
128N
表示
124N
报头数据
130
分段
132A
分段
132B
分段
132N
媒体呈现描述
122

图3

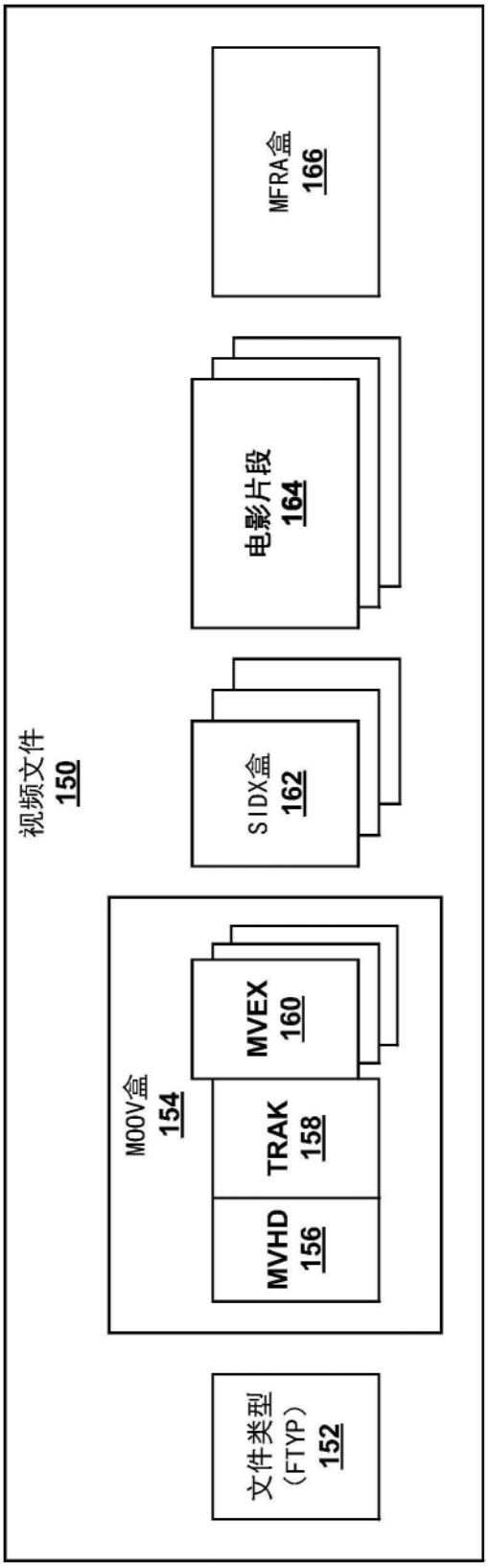

视频文件
150
文件类型
(FTYP)
152
MOOV盒
154
MVHD
156
TRAK
158
MVEX
160
SIDX盒
162
电影片段
164
MFRA盒
166

图4

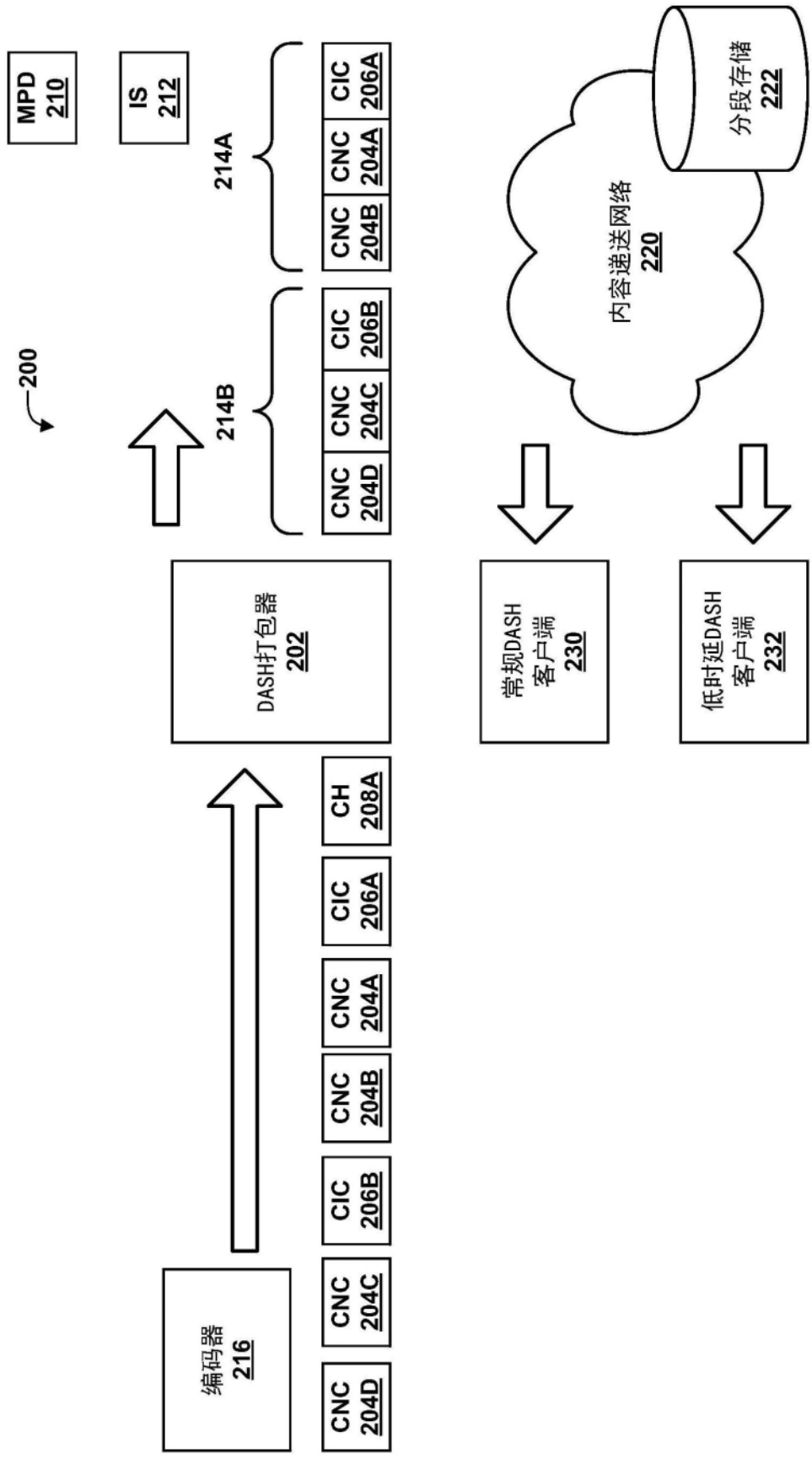

MPD 210
IS 212
200
214A
214B
CNC 204D
CNC 204C
CIC 206B
CNC 204B
CNC 204A
CIC 206A
CH 208A
编码器 216
DASH打包器 202
CNC 204D
CNC 204C
CIC 206B
CNC 204B
CNC 204A
CIC 206A
CH 208A
内容递送网络 220
分段存储 222
常规DASH 客户端 230
低时延DASH 客户端 232

图5

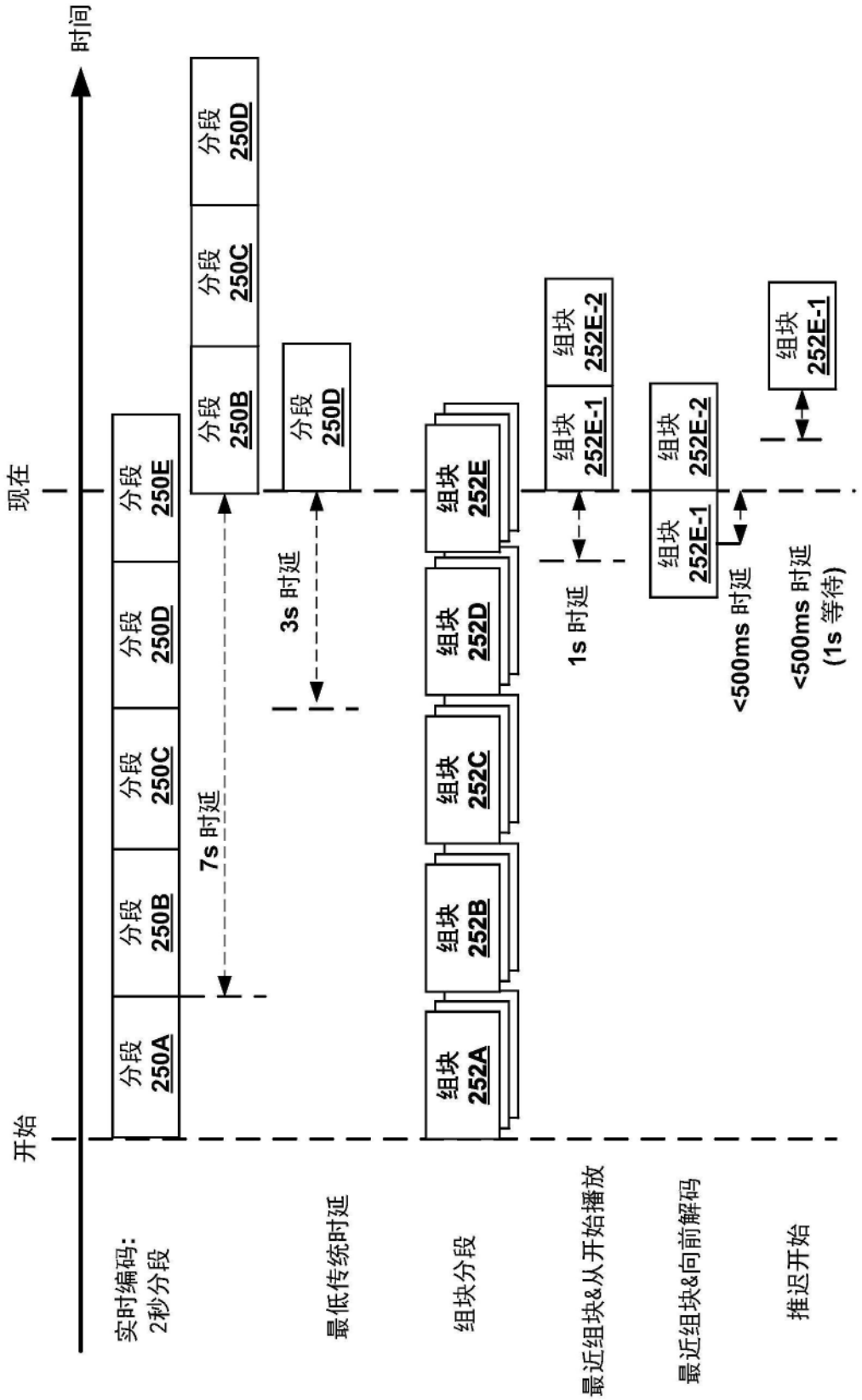

时间
现在
开始
分段 250A
分段 250B
分段 250C
分段 250D
分段 250E
7s 时延
分段 250B
分段 250C
分段 250D
3s 时延
分段 250D
组块 252A
组块 252B
组块 252C
组块 252D
组块 252E
1s 时延
组块 252E-1
组块 252E-2
组块 252E-1
组块 252E-2
<500ms 时延
组块 252E-1
<500ms 时延
(1s 等待)
实时编码:
2秒分段
最低传统时延
组块分段
最近组块&从开始播放
最近组块&向前解码
推迟开始

图6

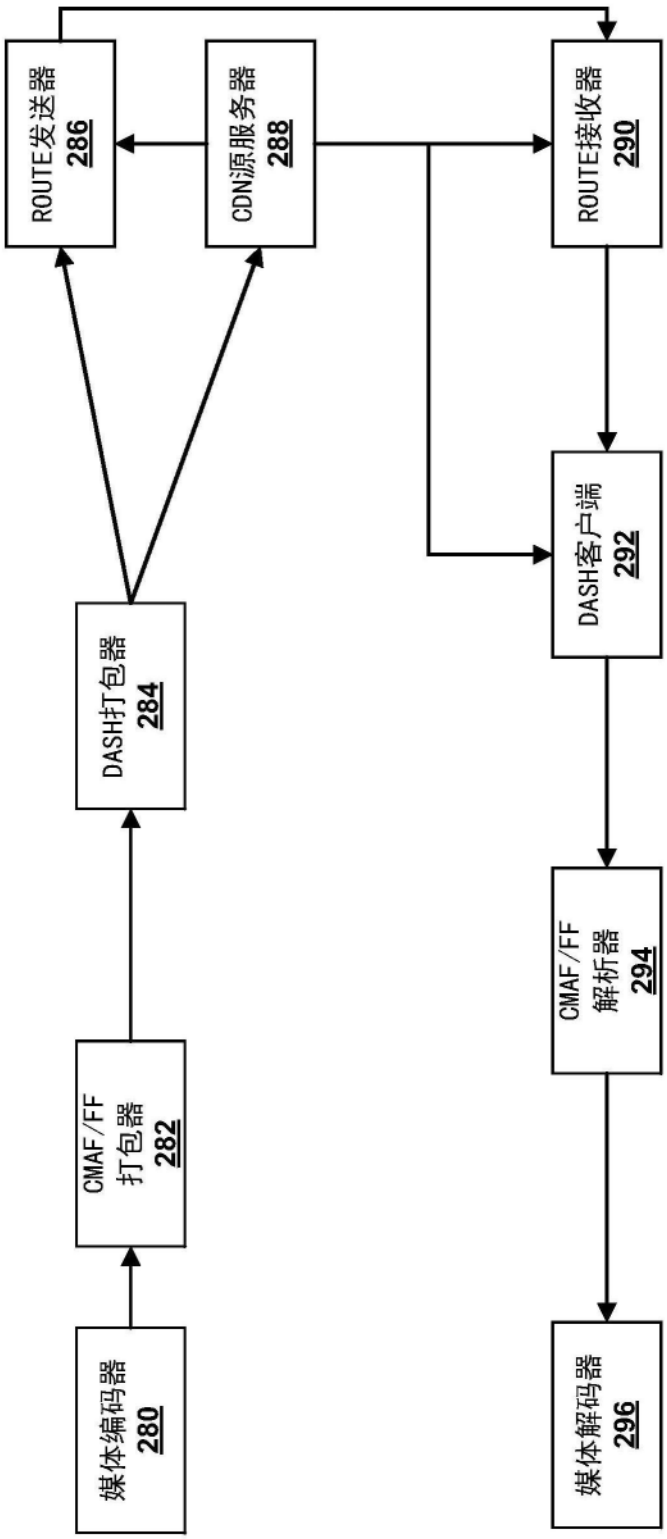

媒体编码器
280
CMAF/FF
打包器
282
DASH打包器
284
ROUTE发送器
286
CDN源服务器
288
ROUTE接收器
290
DASH客户端
292
CMAF/FF
解析器
294
媒体解码器
296

图7

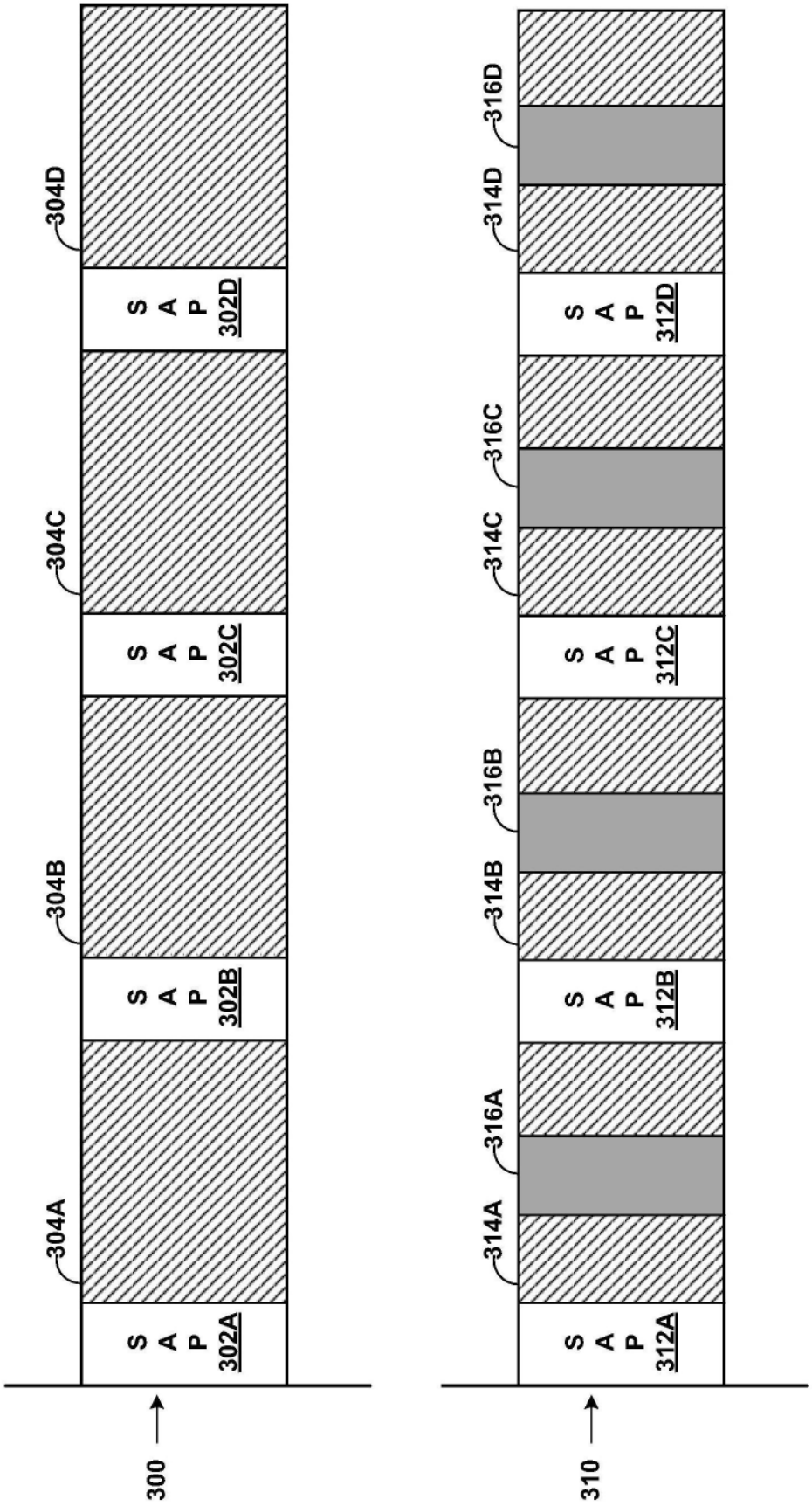

304D
SAP 302D
304C
SAP 302C
304B
SAP 302B
304A
SAP 302A
300
316D
314D
SAP 312D
316C
314C
SAP 312C
316B
314B
SAP 312B
316A
314A
SAP 312A
310

图8

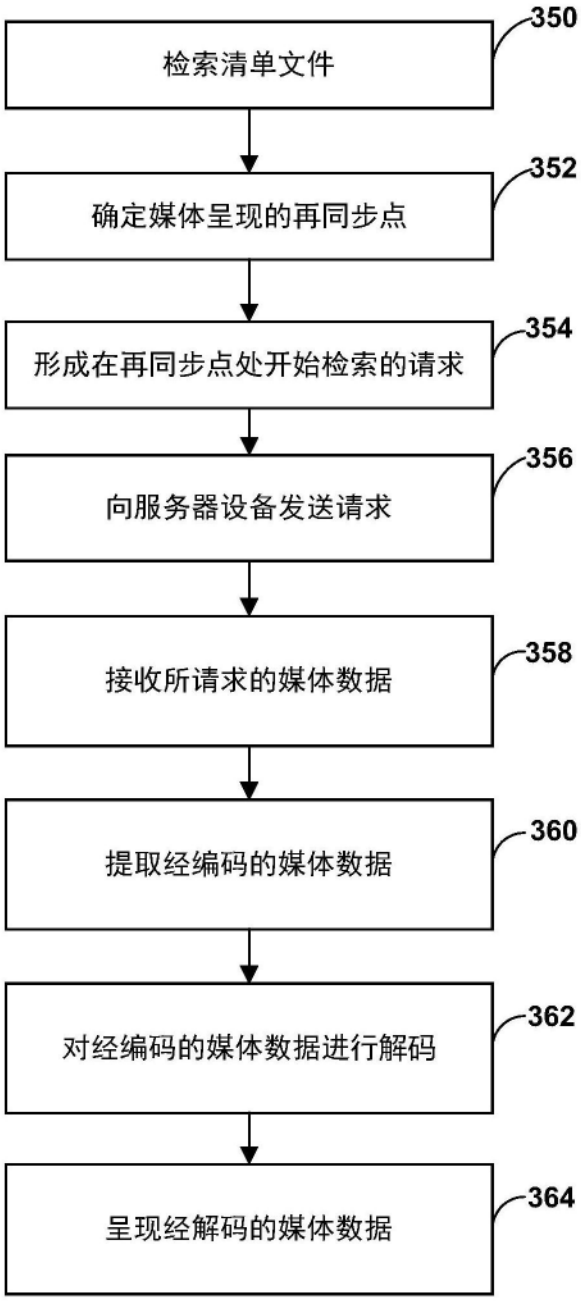

检索清单文件
350
确定媒体呈现的再同步点
352
形成在再同步点处开始检索的请求
354
向服务器设备发送请求
356
接收所请求的媒体数据
358
提取经编码的媒体数据
360
对经编码的媒体数据进行解码
362
呈现经解码的媒体数据
364

图9