



(45) 授权公告日 2016.06.08

代理人 宋献涛

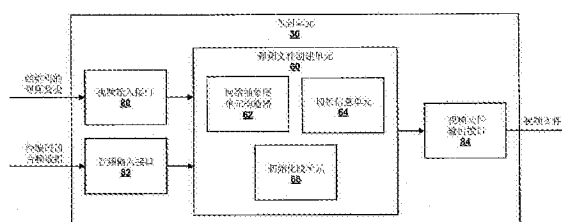
H04N 19/00(2014.01)

地址 美国加利福尼亚州

权利要求书5页 说明书26页 附图8页

用于串流视频数据的视频切换

在一个实例中,一种设备包括处理器,所述处理器经配置以进行以下操作:将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置;及响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求,将指示以下情形的信息提供到所述客户端装置:所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示,而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据。以此方式,所述客户端装置可使用所述信息来执行所述表示之间的无缝切换。



1. 一种发送视频数据的方法,所述方法包含:

将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置;

将指示以下情形的信息提供到所述客户端装置:所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示,而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据,其中将指示以下情形的信息提供到所述客户端装置包含将用于从所述第一表示切换到所述第二表示的延迟值发送到所述客户端装置;

响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求,将来自所述第一表示的所述所请求数据发送到所述客户端装置;及

在发送来自所述第一表示的所述所请求数据之后响应于来自所述客户端装置的检索来自所述第二表示的数据的第二请求,将来自所述第二表示的所述所请求数据发送到所述客户端装置。

2. 根据权利要求1所述的方法,其进一步包含:

确定所述第一表示及所述第二表示的经译码视频样本的最大解码延迟,其中经译码视频样本的解码延迟包含用于解码所述经译码视频样本的延迟;

确定所述第一表示及所述第二表示的所述经译码视频样本的最大图片重排延迟,其中经译码视频图片的图片重排延迟包含所述经译码视频图片的解码时间与所述经译码视频图片的呈现时间之间的差;及

计算包含所述最大解码延迟与所述最大图片重排延迟的总和的延迟值。

3. 根据权利要求2所述的方法,其进一步包含:

确定所述第一表示及所述第二表示的所述经译码视频样本是否在时间上对准;且

在所述第一表示及所述第二表示的所述经译码视频样本未在时间上对准时,其中计算所述延迟值进一步包含将时间对准延迟值相加到所述最大解码延迟与所述最大图片重排延迟的所述总和,其中所述时间对准延迟值包含所述第一表示的最后所请求片段的最后经译码视频样本与所述第二表示的第一片段的第一经译码视频样本之间的时间差。

4. 根据权利要求1所述的方法,其进一步包含:

确定所述表示的经译码视频样本的随机存取点图片的最大大小值;及

将所述最大大小值发送到所述客户端装置。

5. 根据权利要求1所述的方法,其进一步包含:

接收所述第一表示的初始化段;

确定所述第二表示可基于所述第一表示的所述初始化段来加以解码;

将群组识别符值指派给所述第一表示及所述第二表示;及

将所述群组识别符值发送到所述客户端装置,其中发送来自所述第一表示的所述所请求数据包含将所述初始化段发送到所述客户端装置;且

其中将所述第二表示的所述所请求数据发送到所述客户端装置包含将所述第二表示的所述所请求数据发送到所述客户端装置而无需将所述第二表示的初始化段发送到所述客户端装置。

6. 一种用于提供视频数据的设备,所述设备包含:

用于将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置的装置；

用于将指示以下情形的信息提供到所述客户端装置的装置：所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示，而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据；

用于响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求而将来自所述第一表示的所述所请求数据发送到所述客户端装置的装置；及

用于在发送来自所述第一表示的所述所请求数据之后响应于来自所述客户端装置的检索来自所述第二表示的数据的第二请求而将来自所述第二表示的所述所请求数据发送到所述客户端装置的装置；及

用于将用于从所述第一表示切换到所述第二表示的延迟值发送到所述客户端装置的装置。

7. 根据权利要求6所述的设备，其进一步包含：

用于确定所述第一表示及所述第二表示的经译码视频样本的最大解码延迟的装置，其中经译码视频样本的解码延迟包含用于解码所述经译码视频样本的延迟；

用于确定所述第一表示及所述第二表示的所述经译码视频样本的最大图片重排延迟的装置，其中经译码视频图片的图片重排延迟包含所述经译码视频图片的解码时间与所述经译码视频图片的呈现时间之间的差；及

用于计算包含所述最大解码延迟与所述最大图片重排延迟的总和的延迟值的装置。

8. 根据权利要求7所述的设备，其进一步包含：

用于确定所述第一表示及所述第二表示的所述经译码视频样本是否在时间上对准的装置；且

其中在所述第一表示及所述第二表示的所述经译码视频样本未在时间上对准时，所述用于计算所述延迟值的装置进一步包含用于将时间对准延迟值相加到所述最大解码延迟与所述最大图片重排延迟的所述总和的装置，其中所述时间对准延迟值包含所述第一表示的最后所请求片段的最后经译码视频样本与所述第二表示的第一片段的第一经译码视频样本之间的时间差。

9. 根据权利要求6所述的设备，其进一步包含：

用于确定所述表示的经译码视频样本的随机存取点图片的最大大小值的装置；及

用于将所述最大大小值发送到所述客户端装置的装置。

10. 根据权利要求6所述的设备，其进一步包含：

用于接收所述第一表示的初始化段的装置；

用于确定所述第二表示可基于所述第一表示的所述初始化段来加以解码的装置；

用于将群组识别符值指派给所述第一表示及所述第二表示的装置；及

用于将所述群组识别符值发送到所述客户端装置的装置，

其中所述用于发送来自所述第一表示的所述所请求数据的装置包含用于将所述初始化段发送到所述客户端装置的装置，且

其中所述用于发送来自所述第二表示的所述所请求数据的装置包含用于将所述第二

表示的数据发送到所述客户端装置而无需将所述第二表示的初始化段发送到所述客户端装置的装置。

11. 根据权利要求6所述的设备, 其中所述设备包含以下各者中的至少一者:

集成电路;

微处理器; 及

无线通信装置, 其包括所述处理器。

12. 一种接收视频数据的方法, 所述方法包含:

接收指示用于在视频序列的第一表示与所述视频序列的第二表示之间切换的延迟的延迟值;

根据串流网络协议请求来自视频序列的所述第一表示的数据;

响应于确定要从所述第一表示切换到所述第二表示, 基于所述延迟值来缓冲来自所述第一表示的适量的数据; 及

在缓冲来自所述第一表示的所述适量的数据之后, 根据所述串流网络协议请求来自所述视频序列的第二表示的数据。

13. 根据权利要求12所述的方法, 其进一步包含:

在请求所述第一表示的所述数据之前, 请求所述第一表示的初始化段,

其中请求来自所述第二表示的所述数据包含请求来自所述第二表示的所述数据而无需请求所述第二表示的初始化段。

14. 根据权利要求12所述的方法, 其进一步包含:

检测网络条件的改变; 及

基于所述网络条件的所述改变, 确定要从所述第一表示切换到所述第二表示。

15. 根据权利要求12所述的方法, 其进一步包含:

接收指示所述第一表示与所述第二表示是否在时间上对准的信息; 及

基于所述信息来选择所述第二表示的随机存取点。

16. 根据权利要求15所述的方法, 其中在所述第一表示与所述第二表示在时间上对准时, 选择所述随机存取点包含:

定位所述第一表示的具有跟在所述第一表示的所述经缓冲数据的样本的呈现时间值之后的最早呈现时间值的第一随机存取点; 及

选择所述第二表示的具有与所述第一随机存取点相同的呈现时间值的第二随机存取点,

其中请求来自所述第二表示的所述数据包含请求所述第二表示的包含所述第二随机存取点的片段。

17. 根据权利要求15所述的方法, 其中在所述第一表示与所述第二表示未在时间上对准时, 所述方法包含:

选择所述随机存取点包含选择所述第二表示的呈现时间值迟于所述第一表示的所述经缓冲数据的至少一个样本的呈现时间值的随机存取点,

其中缓冲所述适量的数据包含请求并存储所述第一表示的数据直到开始呈现时间值早于所述选定随机存取点的所述呈现时间值且结束呈现时间值迟于所述选定随机存取点的所述呈现时间值的第一片段, 且

其中请求来自所述第二表示的数据包含以所述第二表示的包含所述选定随机存取点的第二片段开始来请求所述第二表示的数据。

18. 根据权利要求17所述的方法, 其进一步包含:

解码所述第一表示的所述数据直到到达所述第一片段的图片群组为止, 所述图片群组包含开始呈现时间值早于所述选定随机存取点的呈现时间值且结束呈现时间值迟于所述选定随机存取点的所述呈现时间值的第一图片, 及

在到达所述第一片段的所述图片群组之后, 以所述选定随机存取点开始来解码所述第二表示的所述数据。

19. 一种用于接收视频数据的设备, 所述设备包含:

用于接收指示用于在视频序列的第一表示与所述视频序列的第二表示之间切换的延迟的延迟值的装置;

用于根据串流网络协议请求来自视频序列的所述第一表示的数据的装置;

用于响应于确定要从所述第一表示切换到所述第二表示而基于所述延迟值来缓冲来自所述第一表示的适量的数据的装置; 及

用于在缓冲来自所述第一表示的所述适量的数据之后根据所述串流网络协议请求来自所述视频序列的第二表示的数据的装置。

20. 根据权利要求19所述的设备, 其进一步包含:

用于在请求所述第一表示的所述数据之前请求所述第一表示的初始化段的装置,

其中所述用于请求来自所述第二表示的所述数据的装置包含用于请求来自所述第二表示的所述数据而无需请求所述第二表示的初始化段的装置。

21. 根据权利要求19所述的设备, 其进一步包含:

用于检测网络条件的改变的装置; 及

用于基于所述网络条件的所述改变而确定要从所述第一表示切换到所述第二表示的装置。

22. 根据权利要求19所述的设备, 其进一步包含:

用于接收指示所述第一表示与所述第二表示是否在时间上对准的信息的装置; 及

用于基于所述信息来选择所述第二表示的随机存取点的装置。

23. 根据权利要求22所述的设备, 其中在所述第一表示与所述第二表示在时间上对准时, 所述用于选择所述随机存取点的装置包含:

用于定位所述第一表示的具有跟在所述第一表示的所述经缓冲数据的样本的呈现时间值之后的最早呈现时间值的第一随机存取点的装置; 及

用于选择所述第二表示的具有与所述第一随机存取点相同的呈现时间值的第二随机存取点的装置,

其中所述用于请求来自所述第二表示的所述数据的装置包含用于请求所述第二表示的包含所述第二随机存取点的片段的装置。

24. 根据权利要求22所述的设备, 其中在所述第一表示与所述第二表示未在时间上对准时:

所述用于选择所述随机存取点的装置包含用于选择所述第二表示的呈现时间值迟于所述第一表示的所述经缓冲数据的至少一个样本的呈现时间值的随机存取点的装置,

所述用于缓冲所述适量的数据的装置包含用于请求并存储所述第一表示的数据直到开始呈现时间值早于所述选定随机存取点的所述呈现时间值且结束呈现时间值迟于所述选定随机存取点的所述呈现时间值的第一片段的装置,且

所述用于请求来自所述第二表示的数据的装置包含用于以所述第二表示的包含所述选定随机存取点的第二片段开始来请求所述第二表示的数据的装置。

25.根据权利要求24所述的设备,其进一步包含:

用于解码所述第一表示的所述数据直到到达所述第一片段的图片群组为止的装置,所述图片群组包含开始呈现时间值早于所述选定随机存取点的呈现时间值且结束呈现时间值迟于所述选定随机存取点的所述呈现时间值的第一图片,及

用于在到达所述第一片段的所述图片群组之后以所述选定随机存取点开始来解码所述第二表示的所述数据的装置。

26.根据权利要求19所述的设备,其中所述设备包含以下各者中的至少一者:

集成电路;

微处理器;及

无线通信装置。

用于串流视频数据的视频切换

[0001] 本申请案主张以下申请案的优先权：2010年7月13日申请的第61/363,884号美国临时申请案、2010年7月21日申请的第61/366,436号美国临时申请案、2010年8月16日申请的第61/374,222号美国临时申请案及2011年1月14日申请的第61/433,110号美国临时申请案，所述申请案中的每一者的全文特此以引用的方式并入本文中。

技术领域

[0002] 本发明涉及经编码视频数据的存储及输送。

背景技术

[0003] 数字视频能力可并入到广泛范围的装置中，包括数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型或桌上型计算机、数码相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电电话、视频电话会议装置及其类似者。数字视频装置实施视频压缩技术(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(AVC)所定义的标准及这些标准的扩展中所描述的视频压缩技术)，以较有效地发射及接收数字视频信息。

[0004] 视频压缩技术执行空间预测及/或时间预测，以减少或移除视频序列中所固有的冗余。对于基于块的视频译码来说，可将视频帧或切片分割为多个宏块。可进一步分割每一宏块。使用相对于相邻宏块的空间预测来编码帧内译码(I)帧或切片中的宏块。经帧间译码(P或B)帧或切片中的宏块可使用相对于同一帧或切片中的相邻宏块的空间预测或相对于其它参考帧的时间预测。

[0005] 在视频数据已被编码之后，可将视频数据包化以供发射或存储。可遵照多种标准中的任一者(例如，国际标准化组织(ISO)基础媒体文件格式及其扩展(例如，AVC))将视频数据组合成视频文件。

[0006] 一直在努力开发出基于H.264/AVC的新视频译码标准。一个此类标准为可缩放视频译码(SVC)标准，其为H.264/AVC的可缩放扩展。另一标准为多视图视频译码(MVC)，其已变成H.264/AVC的多视图扩展。MVC的联合草案描述于JVT-AB204的“关于多视图视频译码的联合草案8.0(Joint Draft 8.0 on Multiview Video Coding)”(2008年7月在德国汉诺威举办的第28届JVT会议)中，此文献可从http://wftp3.itu.int/av-arch/jvt-site/2008_07_Hannover/JVT-AB204.zip获得。AVC标准的版本描述于JVT-AD007的“对ITU-T Rec.H.264|ISO/IEC14496-10高级视频译码的编者修正草案-为ITU-T SG16AAP同意做准备(以综合形式)(Editors'draft revision to ITU-T Rec.H.264|ISO/IEC14496-10 Advanced Video Coding-in preparation for ITU-T SG 16AAP Consent(in integrated form))”(2009年2月在瑞士日内瓦举办的第30届JVT会议)中，此文献可从http://wftp3.itu.int/av-arch/jvt-site/2009_01_Geneva/JVT-AD007.zip获得。本文献提供SVC及MVC标准与AVC规范的综合。

发明内容

[0007] 大体来说,本发明描述用于在使用网络串流协议来串流多媒体内容的表示时在所述表示之间切换的技术。同一多媒体内容可以多种方式来编码,从而产生所述多媒体内容的各种表示。本发明的技术可通过改进客户端装置在不中断所述客户端装置处的视频重放的情况下在表示之间切换的能力来改进多媒体内容的数据的串流(例如,根据网络串流协议,例如HTTP版本1.1,还称作HTTP串流)。

[0008] 根据本发明的技术,服务器装置可用信号发出(即,响应于请求而提供)表示解码延迟及图片重排延迟的延迟值。解码延迟可表示解码表示的一个或一个以上经编码视频样本所需的时间量。图片重排延迟可表示为实现显示目的而重排图片所需的时间量,因为图片未必按与其解码相同的次序来显示。客户端装置可使用所述延迟值来确定客户端装置可从一个表示切换到另一表示的点。也就是说,客户端装置可确定客户端装置停止请求第一表示的数据且改为开始请求第二不同表示的数据的点。考虑所述用信号发出的延迟值,客户端可确定此点以使得重放(在解码及重排之后)不受表示之间的切换所中断。

[0009] 在一个实例中,一种发送视频数据的方法包括:将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置;将指示以下情形的信息提供到所述客户端装置:所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示,而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据;响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求,将来自所述第一表示的所述所请求数据发送到所述客户端装置;及在发送来自所述第一表示的所述所请求数据之后响应于来自所述客户端装置的检索来自所述第二表示的数据的第二请求,将来自所述第二表示的所述所请求数据发送到所述客户端装置。

[0010] 在另一实例中,一种用于发送视频数据的设备包括处理器,所述处理器经配置以进行以下操作:将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置;将指示以下情形的信息提供到所述客户端装置:所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示,而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据;响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求,将来自所述第一表示的所述所请求数据发送到所述客户端装置;及在发送来自所述第一表示的所述所请求数据之后响应于来自所述客户端装置的检索来自所述第二表示的数据的第二请求,将来自所述第二表示的所述所请求数据发送到所述客户端装置。

[0011] 在另一实例中,一种用于发送视频数据的设备包括:用于将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置的装置;用于将指示以下情形的信息提供到所述客户端装置的装置:所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示,而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据;用于响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求而将来自所述第一表示的所述所请求数据发送到所述客户端装置的装置;及用于在发送来自所述第

一表示的所述所请求数据之后响应于来自所述客户端装置的检索来自所述第二表示的数据的第二请求而将来自所述第二表示的所述所请求数据发送到所述客户端装置的装置。

[0012] 在另一实例中,一种计算机程序产品包括计算机可读存储媒体,所述计算机可读存储媒体包含在执行时致使用于发送视频数据的源装置的处理器的指令:将用于存取视频序列的第一表示及所述视频序列的第二表示的数据的信息提供到客户端装置;将指示以下情形的信息提供到所述客户端装置:所述客户端装置可在所述第一表示的所述数据的发射期间切换到所述第二表示,而不会在显示所述视频序列时经历呈现间断且无需同时执行一个以上解码器以在所述切换期间解码所述第一表示及所述第二表示的所述数据;响应于来自所述客户端装置的检索来自所述第一表示的数据的第一请求,将来自所述第一表示的所述所请求数据发送到所述客户端装置;及在发送来自所述第一表示的所述所请求数据之后响应于来自所述客户端装置的检索来自所述第二表示的数据的第二请求,将来自所述第二表示的所述所请求数据发送到所述客户端装置。

[0013] 在另一实例中,一种接收视频数据的方法包括:接收指示用于在视频序列的第一表示与所述视频序列的第二表示之间切换的延迟的延迟值;根据串流网络协议请求来自视频序列的所述第一表示的数据;响应于确定待从所述第一表示切换到所述第二表示,基于所述延迟值来缓冲来自所述第一表示的适量的数据;及在缓冲来自所述第一表示的所述适量的数据之后,根据所述串流网络协议请求来自所述视频序列的第二表示的数据。

[0014] 在另一实例中,一种用于接收视频数据的设备包括处理器,所述处理器经配置以进行以下操作:接收指示用于在视频序列的第一表示与所述视频序列的第二表示之间切换的延迟的延迟值;根据串流网络协议请求来自视频序列的所述第一表示的数据;响应于确定待从所述第一表示切换到所述第二表示,基于所述延迟值来缓冲来自所述第一表示的适量的数据;及在缓冲来自所述第一表示的所述适量的数据之后,根据所述串流网络协议请求来自所述视频序列的第二表示的数据。

[0015] 在另一实例中,一种用于接收视频数据的设备包括:用于接收指示用于在视频序列的第一表示与所述视频序列的第二表示之间切换的延迟的延迟值的装置;用于根据串流网络协议请求来自视频序列的所述第一表示的数据的装置;用于响应于确定待从所述第一表示切换到所述第二表示而基于所述延迟值来缓冲来自所述第一表示的适量的数据的装置;及用于在缓冲来自所述第一表示的所述适量的数据之后根据所述串流网络协议请求来自所述视频序列的第二表示的数据的装置。

[0016] 在另一实例中,一种计算机程序产品包括计算机可读存储媒体,所述计算机可读存储媒体包含致使用于接收视频数据的目的地装置的处理器的指令:接收指示用于在视频序列的第一表示与所述视频序列的第二表示之间切换的延迟的延迟值;根据串流网络协议请求来自视频序列的所述第一表示的数据;响应于确定待从所述第一表示切换到所述第二表示,基于所述延迟值来缓冲来自所述第一表示的适量的数据;及在缓冲来自所述第一表示的所述适量的数据之后,根据所述串流网络协议请求来自所述视频序列的第二表示的数据。

[0017] 一个或一个以上实例的细节陈述于附图及以下描述中。其它特征、目标及优势将从所述描述及所述图式以及从权利要求书显而易见。

附图说明

[0018] 图1为说明其中音频/视频(A/V)源装置将音频及视频数据传送到A/V目的地装置的实例系统的框图。

[0019] 图2为说明实例包封单元的组件的框图。

[0020] 图3为说明经译码视频帧序列的概念图。

[0021] 图4为说明可切换表示的概念图。

[0022] 图5A为说明在时间上对准的两个表示的实例的概念图。

[0023] 图5B为说明未在时间上对准的两个表示的实例的概念图。

[0024] 图6为说明用于执行两个可切换表示之间的无缝切换的实例方法的流程图。

[0025] 图7为说明服务器装置将描述可切换表示的信息提供到客户端装置的实例方法的流程图。

[0026] 图8为说明客户端装置自从服务器装置检索第一表示的数据切换到从所述服务器装置检索第二表示的数据的实例方法的流程图。

具体实施方式

[0027] 大体来说,本发明描述用于在网络串流上下文中支持多媒体内容的表示之间的切换的技术。当在表示之间进行切换时,可能存在由多种因素引起的自然延迟。如下文更详细地解释,本发明的技术包括用信号发出(例如,服务器装置响应于来自客户端装置请求而提供数据)指示延迟的信息以供在从多媒体内容的一个表示切换到所述多媒体内容的另一表示时考虑。以此方式,客户端装置可缓冲来自所述第一表示的足够量的数据,以允许在接收到来自所述第二表示的数据且开始解码来自所述第二表示的数据之前解码来自所述第一表示的数据。因此,客户端装置可能能够在多个表示之间切换而不会中断多媒体内容的重放。以此方式,本发明的技术可改进用户当在客户端装置上观看内容时的体验。

[0028] 用信号发出的延迟可表示解码延迟及图片重排延迟两者。明确地说,用信号发出的延迟值可包含用于从一个表示切换到另一表示的最大解码延迟与最大重排延迟的组合。可(例如)基于每一帧的编码模式、位速率、分辨率或其它因素而存在与图片序列相关联的不同解码延迟。用信号发出的延迟值可包括任何表示的任何存取单元的最大解码延迟,或切换所涉及的两个表示的任何存取单元的最大解码延迟。

[0029] 图片重排延迟可对应于与从解码到显示次序重排图片相关联的延迟。在H.264/AVC的实例中,位流可包括各种类型的经译码图片,例如,经帧内译码图片(例如,I图片)及经帧间译码图片(例如,P图片及B图片)。I图片的数据大体上仅相对于所述I图片的其它数据来译码,且因此经图片内译码。P图片及B图片的数据可相对于一个或一个以上其它先前经译码图片来译码。举例来说,P图片可相对于在显示次序上在所述P图片之前的一个或一个以上图片来译码。作为另一实例,B图片可相对于在显示次序上在所述B图片之前的一个或一个以上图片及在显示次序上在所述B图片之后的一个或一个以上图片来译码。用于参考的图片可经编码且早于所述图片而放置于所述位流中以供那些图片参考。

[0030] 作为一个实例,B图片可相对于I图片及P图片两者来编码,且P图片可相对于I图片来编码。在所述三个图片中,所述P图片可具有最早的显示时间,所述I图片可具有最迟的显

示时间,且所述B图片可具有在所述P图片的显示时间与所述I图片的显示时间之间的显示时间。归因于这些图片的预测结构,视频编码器可将所述I图片作为所述三个图片中的最早者而放置于所述位流中,将所述B图片作为所述三个图片中的最迟者而放置于所述位流中,且将所述P图片放置于所述位流中在所述I图片与所述B图片之间。因此,解码器可按次序来解码所述位流,即,先解码所述I图片,接着解码所述P图片,且再接着解码所述B图片。然而,所述解码器可将所述图片重排成显示次序,使得所述P图片在所述B图片之前显示,且所述B图片在所述I图片之前显示。

[0031] 如上文所论述,用信号发出的延迟值可包括解码延迟与图片重排延迟的组合。所述图片重排延迟可对应于任何表示中的任何图片的呈现时间与解码时间之间的最大差。在一些实例中,图片重排延迟可对应于切换中所涉及的两个表示中的任何图片的呈现时间与解码时间之间的最大差。呈现时间对应于显示图片的相对时间,而解码时间对应于解码图片的相对时间。如上文所提,呈现时间与解码时间未必相等,且因此,在解码之后,可将经解码的图片重排成呈现时间排序。

[0032] 服务器装置可向客户端装置用信号发出用于无缝切换的延迟值。所述延迟值可包含任何表示中的任何图片的最大解码延迟与同一多媒体内容的任何表示中的任何图片的呈现时间与解码时间之间的最大差的总和。客户端装置可使用此用信号发出的延迟值来执行两个表示之间的无缝切换。举例来说,客户端装置可确定在切换到不同表示之前当前表示中待缓冲的数据的量,使得可实现无缝切换。也就是说,假定可用带宽足够用于新表示,那么客户端装置可具有流畅重放及无缝解码。

[0033] 为实现无缝切换,客户端装置可提供在表示之间切换期间的无缝解码以及经解码图片的无缝显示两者。无缝解码可确保不需要第二解码器来解码来自新表示的图片(但客户端装置仍可包括一个以上解码器)。无缝显示可对应于使图片准备好在表示之间切换期间连续重放。

[0034] 本发明的技术还包括从一组表示形成切换群组,且用信号发出一组表示属于同一切换群组。表示可包括初始化段,所述初始化段包括用于初始化目的地装置的解码器及/或其它元件的数据。举例来说,所述初始化段可描述ISO基础媒体文件的电影盒、H.264/AVC位流的含有序列参数集及图片参数集的样本,以及可能描述序列层级补充增强信息(SEI)消息。在一些情况下,同一初始化段可用以针对一个以上表示初始化所述目的地装置。切换群组可包括可各自使用同一初始化段来初始化的一组表示。以此方式,客户端装置可检索切换群组中的第一表示的初始化段,接着切换到所述切换群组中的第二表示而无需检索所述第二表示特有的初始化段。也就是说,客户端装置可使用所述第一表示的初始化段来解码所述第二表示。切换群组可包括具有在时间上对准的段的表示,因此从一个表示切换到另一表示,此时所有段均以随机存取点图片开始。在此情况下,如果此群组之间发生切换,则需要请求不同表示的时间上重叠的段。

[0035] 大体来说,I图片可称作随机存取点(RAP)。也就是说,解码器可以I图片开始来开始解码视频数据且能够解码位流中在所述I图片之后的一些图片(如果不是所有的话)。然而,可能存在不可正确解码的一些图片,例如显示次序早于I图片但解码次序迟于I图片的B图片。这些B图片可取决于位流中较早的I图片。在位流中其所有后续图片可在不检索位流中的较早图片的情况下加以适当地解码的I图片可称作瞬时解码刷新(IDR)随机存取点。在

位流中其至少一个后续图片无法在不检索位流中的较早图片的情况下加以适当地解码的I图片可称作开放解码刷新(ODR)随机存取点。

[0036] 换句话说, IDR图片可仅含有经帧内译码切片且可致使除了IDR图片之外的所有参考图片被标记为“未用于参考”。经译码视频序列可被定义为按解码次序从一个IDR存取单元(包括在内)到下一IDR存取单元(排除在外)或到位流的末尾(取决于哪一个出现较早)的连续存取单元的序列。可适当地解码包括IDR图片的图片群组(GOP)中的每一图片,而不管是否解码了任何先前图片。此GOP可称作封闭GOP。也就是说,封闭GOP为其中所有图片可适当地加以解码的GOP,假定所述GOP中的所有图片可用于解码。在H.264/AVC的实例中,封闭GOP可从IDR存取单元(即,包括IDR图片的存取单元)开始。此外,可适当地解码经译码视频序列中的每一图片,假定所述经译码视频序列中的每一图片可用于解码(包括初始IDR图片)。

[0037] 开放GOP可被定义为其中在显示次序上在所述GOP的I图片之前的图片不可正确地解码(甚至在所述开放GOP中的所有图片均可用于解码时)的图片群组。编码器可将指示GOP为开放GOP的信息(例如恢复点补充增强信息(SEI)消息)包括在位流中。开放GOP的I图片可称作ODR图片。虽然在解码次序上在所述开放GOP之后的图片可为可正确地解码的,但所述开放GOP中的某些图片可参考在解码次序上先前的I图片,且因此不可适当地解码,除非所述先前I图片可用于解码。

[0038] 大体上,解码器可从IDR或ODR图片开始来开始解码位流,且因此将IDR图片及ODR图片两者视为所述位流内的随机存取点。然而,因为当在ODR图片处开始存取位流时存在不可适当地解码的在所述位流中在ODR图片之后的一些图片,所以解码器在尝试解码这些图片时可能会白费功夫。根据本发明的技术,服务器可用信号发出随机存取点图片是IDR图片还是ODR图片。服务器还可利用信号发出所述随机存取点的字节范围,使得客户端装置可提交仅存取所述IDR及/或ODR图片的一个或一个以上HTTP部分GET请求。以此方式,客户端可检索所述IDR图片及/或所述ODR图片以实现多媒体内容的特技模式显示,例如快进或倒退。

[0039] 本发明的技术可应用于遵照以下格式中的任一者的视频文件:ISO基础媒体文件格式、可缩放视频译码(SVC)文件格式、高级视频译码(AVC)文件格式、第三代合作伙伴计划(3GPP)文件格式,及/或多视图视频译码(MVC)文件格式,或其它类似视频文件格式。

[0040] 所述ISO基础媒体文件格式经设计以含有呈灵活、可扩展的格式的供呈现的时控媒体信息,所述灵活、可扩展的格式促进媒体的互换、管理、编辑及呈现。ISO基础媒体文件格式(ISO/IEC14496-12:2004)在MPEG-4第12部分中指定,MPEG-4第12部分定义基于时间的媒体文件的一般结构。ISO基本媒体文件格式用作所述系列中的例如以下各项的其它文件格式的基础:经定义以支持H.264/MPEG-4AVC视频压缩的AVC文件格式(ISO/IEC14496-15)、3GPP文件格式、SVC文件格式,及MVC文件格式。3GPP文件格式及MVC文件格式为AVC文件格式的扩展。ISO基础媒体文件格式含有用于媒体数据的时控序列(例如,音频视觉呈现)的时序、结构及媒体信息。所述文件结构可为面向对象的。可极其简单地将文件分解成基本对象,且对象的结构由其类型暗示。

[0041] 遵照ISO基础媒体文件格式(及其扩展)的文件可形成为一连串对象,称作“盒”。呈ISO基础媒体文件格式的数据可含于盒中,使得文件内不需要含有其它数据,且文件内不需要有在盒之外的数据。此包括特定文件格式所需的任何初始签名。“盒”可为通过唯一类型

识别符及长度定义的面向对象的构建块。通常,呈现含于一个文件中,且媒体呈现为自含式的。电影容器(电影盒)可含有媒体的元数据,且视频及音频帧可含于媒体数据容器中且可在其它文件中。

[0042] 表示(运动序列)可含于若干文件中。时序及成帧(位置及大小)信息大体上在ISO基础媒体文件中,且辅助文件可基本上使用任何格式。此表示可在含有所述表示的系统的“本地”,或可经由网络或其它流递送机制来提供。

[0043] 所述文件可具有逻辑结构、时间结构及物理结构,且这些结构不需要耦合。文件的逻辑结构可为电影或视频剪辑(可能包括视频数据及音频数据两者),其又含有一组时间并行的轨道。文件的时间结构可使得轨道含有按时间的样本的序列,且那些序列通过任选的编辑列表而映射到整个电影的时间线中。文件的物理结构可将逻辑、时间及结构分解所需的数据与媒体数据样本自身分开。此结构信息可集中于电影盒中,可能在时间上由电影片段盒扩展。所述电影盒可评注样本的逻辑及时序关系,且还可含有到其所处位置的指针。那些指针可指向同一文件或另一文件(例如,通过URL参考)。

[0044] 每一媒体流可含于专门用于那个媒体类型(音频、视频等)的轨道中,且可通过样本条目来参数化。所述样本条目可含有确切媒体类型(解码所述流所需的解码器的类型)的“名称”及那个所需解码器的任何参数化。所述名称还可采取四字符码(例如,“moov”或“trak”)的形式。存在不仅用于MPEG-4媒体而且用于使用此文件格式系列的其它组织所使用的媒体类型的经定义的样本条目格式。

[0045] 对元数据的支持大体上采取两种形式。首先,可将时控元数据存储于适当轨道中,且在需要时与其所描述的媒体数据同步。其次,可存在对附接到电影或个别轨道的非时控元数据的一般支持。对元数据的结构支持为一般的,且允许元数据资源以与媒体数据(即,经译码视频图片)的存储类似的方式存储于所述文件中的别处或存储于另一文件中。另外,可为这些资源指定名称且可保护这些资源。

[0046] 术语“渐进式下载”用以描述通常使用HTTP协议(例如,HTTP的版本1.1)的数字媒体文件从服务器到客户端的传送。渐进式下载的一个实例为HTTP串流。当由计算机起始时,所述计算机可在下载完成之前开始媒体重放。串流媒体与渐进式下载之间的一个差别在于,正存取数字媒体的最终用户装置如何接收并存储数字媒体数据。能够进行渐进式下载重放的媒体播放器依赖于位于文件的标头中的元数据为完整的以及在从网络服务器下载数字媒体文件时对其进行本地缓冲。在指定量的经缓冲数据变得可为本地重放装置所用时,所述装置可开始播放所述媒体。此指定量的经缓冲数据可由编码器设定中的内容制作者嵌入到文件中,且可通过客户端计算机的媒体播放器所强加的额外缓冲器设定来加强。

[0047] 在HTTP串流中,并非为表示提供电影(MOOV)盒中的所有数据(包括经编码音频及视频样本),而是支持电影片段(MOOF盒),所述电影片段可与电影盒分开地含有样本。每一电影片段可含有在某一时间段内的样本。大体来说,客户端可使用所述电影片段来快速地搜寻所要时间。电影片段可含有文件的连续字节,使得在HTTP串流(举例来说)中,客户端可发出检索电影片段的单一部分GET请求。部分GET请求指定待检索的文件的识别符(例如,统一资源定位符(URL)或统一资源名称(URN))以及所述文件内的字节范围。在一些实例中,经编码音频及视频样本可完全含于电影片段中,使得电影盒自身不含有任何经编码音频或视

频样本。

[0048] AVC及3GPP为ISO基础媒体文件格式的扩展,而SVC及MVC为AVC文件格式的扩展。因此,本发明的技术可应用于遵照以下各者的视频文件:ISO基础媒体文件格式、AVC文件格式及其扩展(例如,SVC及MVC),及/或3GPP文件格式。所述技术可进一步应用于这些格式的这此及其它扩展,且可进一步应用以扩展其它文件格式以提供描述可切换表示的信息及用以允许两个或两个以上表示之间的无缝切换的延迟值。

[0049] 运动图片专家组(MPEG)已开发出AVC文件格式,作为ISO基础媒体文件格式的扩展。在AVC文件格式中,参数集(包括序列参数集及图片参数集)可与视频基本流相关联,所述视频基本流位于视频轨道中。还有可能使参数集位于称作参数集轨道的另一轨道中,所述参数集轨道包括含有作为SPS或PPS网络抽象层(NAL)单元的样本的参数集基本流。然而,SEI消息NAL单元的位置在AVC文件格式中并不受限制,只要任何SEI消息存在于解码时间在使用此SEI消息的任何图片之前的样本中便可。

[0050] AVC文件格式通过SVC及MVC扩展来扩展以分别支持SVC及MVC的存储。在SVC及MVC中,多个层/视图可含于一个视频轨道或多个视频轨道中。存在SVC及MVC规范中所定义的使用在所述层/视图以各种组织来呈现时构造操作点的工具。

[0051] 关于作为另一实例的3GPP,支持3GPP文件的HTTP/TCP/IP输送以进行下载及渐进式下载。此外,使用HTTP进行视频串流可提供一些优点,且基于HTTP的视频串流服务正变得流行。HTTP串流可提供某些优点,包括可使用现有因特网组件及协议,使得不需要做出新的努力来开发用于经由网络输送视频数据的新技术。其它输送协议(例如,RTP有效负载格式)需要中间网络装置(例如,中间盒)知晓媒体格式及信令上下文。而且,HTTP串流可为客户端驱动的,此可避免控制问题。

[0052] 举例来说,为了利用特征来获得最佳性能,服务器可持续跟踪尚未得到应答的包的大小及内容。服务器还可分析文件结构并重构客户端缓冲器的状态以做出RD最佳切换/精简决策。另外,可满足对位流变化的约束以便保持遵从所协商的简档。HTTP在实施了HTTP1.1的网络服务器处未必需要新的硬件或软件实施方案。HTTP串流还提供TCP友好性及防火墙穿越。

[0053] 本发明的技术可适用于网络串流协议,例如,(例如)根据HTTP动态自适应性串流(DASH)的HTTP串流。在HTTP串流中,频繁使用的操作包括GET及部分GET。GET操作检索与给定统一资源定位符(URL)或统一资源名称(URN)相关联的整个文件。部分GET操作接收字节范围作为输入参数且检索文件的对应于所接收到的字节范围的多个连续字节。因此,可提供电影片段以进行HTTP串流,因为部分GET操作可取得一个或一个以上个别电影片段。请注意,在电影片段中,可存在不同轨道的若干轨道片段。在HTTP串流中,媒体表示可为客户端可存取的数据的结构化集合。客户端可请求并下载媒体数据信息以向用户呈现串流服务。

[0054] 在使用HTTP串流来串流3GPP数据的实例中,可存在多媒体内容的视频及/或音频数据的多个表示。这些表示的清单可定义于媒体呈现描述(MPD)数据结构中。媒体呈现可对应于HTTP串流客户端装置可存取的数据的结构化集合。所述HTTP串流客户端装置可请求并下载媒体数据信息以向所述客户端装置的用户呈现串流服务。媒体呈现可描述于MPD数据结构中,所述MPD数据结构可包括MPD的更新。

[0055] 媒体呈现可含有一个或一个以上周期的序列。周期可由MPD中的Period(周期)元

素来定义。每一周期可在MPD中具有属性start(开始)。所述MPD可包括针对每一周期的start属性及availableStartTime(可用开始时间)属性。对于实况服务,周期的start属性与MPD属性availableStartTime的总和可以UTC格式指定周期的可用性时间,尤其是对应周期中每一表示的第一媒体段。对于随选服务,第一周期的start属性可为0。对于任何其它周期,所述start属性可指定对应周期的开始时间相对于第一周期的开始时间之间的时间偏移。每一周期可延长直到下一周期的开始为止,或直到媒体呈现的末尾为止(在最后一个周期的情况下)。周期开始时间可为精确的。周期开始时间可反映由播放所有先前周期的媒体产生的实际时序。

[0056] 每一周期可含有同一媒体内容的一个或一个以上表示。表示可为音频或视频数据的许多个替代性经编码版本中的一者。所述表示可通过编码类型来区分,例如,通过针对视频数据的位速率、分辨率及/或编解码器以及针对音频数据的位速率、语言及/或编解码器。术语“表示”可用以指代对应于多媒体内容的特定周期且以特定方式来编码的经编码音频或视频数据的一部分。

[0057] 特定周期的表示可指派给通过MPD中的group(群组)属性指示的群组。同一群组中的表示大体上被视为彼此的替代物。举例来说,特定周期的视频数据的每一表示可指派给同一群组,使得可选择所述表示中的任一者以进行解码从而显示对应周期的多媒体内容的视频数据。在一些实例中,一个周期内的媒体内容可通过来自群组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来表示。一个周期的每一表示的时序数据可相对于所述周期的开始时间来表达。

[0058] 表示可包括一个或一个以上段。每一表示可包括初始化段,或表示的每一段可为自我初始化的。在存在时,所述初始化段可含有用于存取所述表示的初始化信息。大体来说,所述初始化段不含有媒体数据。段可唯一地由识别符来参考,例如统一资源定位符(URL)、统一资源名称(URN)或统一资源识别符(URI)。所述MPD可提供每一段的识别符。在一些实例中,所述MPD还可提供呈range(范围)属性的形式的字节范围,所述字节范围可对应于文件内的可通过URL、URN或URI来存取的段的数据。

[0059] 每一表示还可包括一个或一个以上媒体分量,其中每一媒体分量可对应于一个个别媒体类型(例如,音频、视频或时控文本(例如,对于隐藏字幕))的经编码版本。媒体分量可跨越一个表示内的连续媒体段的边界为时间连续的。

[0060] MPD还可(例如)通过提供group属性来支持部分表示。部分表示可对应于多媒体文件的视频或音频轨道。MPD可指示:在两个部分表示可彼此切换时,所述两个部分表示具有相同group属性。

[0061] 服务器装置还可利用信号发出两个表示是否可有效地彼此切换。如果两个表示中的任一者的初始化段可在解码另一表示时使用,则可认为所述两个表示可有效地切换。也就是说,客户端装置可接收当前表示的初始化段,接着切换到新表示而无需请求新表示的初始化段。

[0062] 服务器装置可在MPD数据结构中设定旗标,所述旗标指示表示是否为部分表示。举例来说,所述MPD数据结构可包括每一表示的切换识别符。服务器装置可将两个或两个以上可有效地切换的表示的切换识别符设定为相同值。以此方式,客户端装置可基于具有相同切换识别符的表示中的每一者而确定彼此可有效地切换的一组表示。在MPD中,服务器可进

一步指示所述表示的段是否在时间上对准。

[0063] MPD还可含有minBuf ferrTime(最小缓冲时间)属性,所述属性指示确保流畅播放所需的最初缓冲的媒体的最小量,限制条件为每一表示是以MPD的bandwidth(带宽)属性的值或高于所述值来递送的。以此方式,客户端装置可确定在开始重放之前待缓冲多少媒体,假定某一量的可用带宽。客户端可进一步基于可用带宽的量来选择表示,例如,最初或在网络条件改变时基于bandwidth属性。这些属性可在用以在各种网络条件下实现内容的连续解码及呈现的解决方案中提供简单缓冲器模型。

[0064] 在每一周期中,表示可为可彼此切换的。MPD可提供bitstreamSwitchingFlag(位流切换旗标)属性。当此旗标为真时,其指示来自含有相同媒体类型的任何两个不同表示的在一个周期内的任何两个时间相继的媒体段的位流层级的拼接结果遵从媒体段格式。当此旗标为真时,所有位流可假定为在时间上对准的且具有相同的简档及层级指示符,且许多序列参数为相同的(在H.264/AVC译码的实例中)。

[0065] 在串流及重放期间,客户端装置可在表示之间进行切换。举例来说,客户端装置可确定网络条件已改变而使得可用带宽已增加或已减小。当带宽已增加时,客户端装置可切换到较高质量的表示,假定客户端装置能够解码并再现较高质量表示的数据。另一方面,当带宽已减小时,客户端装置可切换到消耗较少带宽的较低质量的表示。大体来说,基于在正在进行的媒体呈现期间经更新的信息,客户端装置可切换表示。切换到新表示可包括与从前一个表示停止的时间点(例如,时间位置)开始检索或搜寻表示内的特定位置类似的程序。

[0066] 在确定到新表示的切换为适当之后,客户端装置可以随机存取点开始来开始串流所述新表示。也就是说,客户端装置可在新表示中在迟于但接近于当前呈现时间的所要呈现时间处搜寻随机存取点。通过呈现来自前一个表示的数据直到所述随机存取点,接着在所述随机存取点处以及在所述随机存取点后呈现来自所述新表示的数据,客户端装置可实现无缝切换。根据本发明的技术,客户端装置可使用用信号发出的延迟时间来确定新表示的随机存取点,从所述随机存取点开始串流所述新表示的数据。服务器装置可针对媒体内容在MPD数据结构中用信号发出延迟时间。在一些实例中,源装置可跨越不同表示来对准随机存取点,此可提供在表示中定位随机存取点方面的优点。

[0067] 在一些情况下,两个可切换表示可能并不为时间对准的。举例来说,所述两个表示中的随机存取点可出现在不同时间位置处。在此情况下,服务器可设定旗标或提供其它信息来指示所述两个表示并不对准。在此情况下,从一个表示切换到另一表示可为困难的。举例来说,当前表示中的片段的末尾与所切换到的表示的新片段中的在所述随机存取点处开始的帧可具有显示时间重叠。本发明还提供用于在两个非对准表示之间进行无缝切换的技术。明确地说,客户端装置及/或服务器装置可引入额外延迟,以致使客户端装置在于随机存取点处切换到新表示之前继续检索来自当前表示的数据。此延迟可取决于当前表示的GOP结构以及图片在GOP中的位置。

[0068] 根据本发明的技术,服务器装置可用信号发出重叠延迟值,所述值表示非随机存取点图片在一个表示中的时间位置与随机存取点图片在另一表示中的时间位置之间的差。服务器装置可针对所述表示在MPD中用信号发出重叠延迟值。在一些实例中,服务器装置可用信号发出整个周期或甚至整个服务内每一表示的最大重叠延迟值。在一些实例中,服务

器装置可用信号发出作为重叠延迟与上文针对对准的片段所论述的延迟值的总和的最大延迟值。因此,服务器装置可用信号发出单一延迟值,所述单一延迟值可包括两个或两个以上可切换表示的最大解码延迟、最大图片重排延迟及最大重叠延迟的总和。

[0069] 服务器装置可进一步用信号发出表示的最小延迟值。假定所述表示是以其带宽属性的值或高于所述值来递送的,如果缓冲了足够的数据以覆盖所述最小延迟,则可确保客户端装置流畅地重放所述表示。服务器装置可经配置以修改最小缓冲时间值的计算以覆盖无缝切换的情况。所述最小缓冲时间值(minBufferTime)可提供确保流畅播放及无缝切换所需的最初缓冲的媒体的最小量,限制条件为每一表示是以其带宽属性的值或高于所述值来递送的。以此方式,所述最小缓冲时间值可在用以在各种网络条件下实现内容的连续解码及呈现的解决方案中对应于简单缓冲器模型。

[0070] 大体来说,客户端装置的用户可在重放期间暂停或停止媒体呈现。当用户暂停或停止呈现时,客户端装置可停止请求媒体段或其部分,直到所述用户重新继续重放为止。为了重新继续重放,客户端装置可以紧跟在上次所请求片段后的片段开始来请求媒体段。任何特定表示的MPD可提供所述表示的TrickMode(特技模式)元素,在此情况下,可假定所述表示被明确地启用以用于特技模式。特技模式大体上包括快进及倒退模式,或相对于“正常”重放来更改重放的其它此类模式。MPD可提供alternatePlayoutRate(交替播放速率)属性,在此情况下,客户端装置可以对应于alternatePlayoutRate的速度来播放表示,其中解码器简档及层级与正常播放速率相同。客户端装置还可使用多个表示来支持特技模式行为。

[0071] 图1为说明其中音频/视频(A/V)源装置20将音频及视频数据输送到A/V目的地装置40的实例系统10的框图。图1的系统10可对应于视频会议系统、服务器/客户端系统、广播台/接收器系统,或其中将视频数据从源装置(例如,A/V源装置20)发送到目的地装置(例如,A/V目的地装置40)的任何其它系统。在一些实例中,A/V源装置20及A/V目的地装置40可执行双向信息交换。也就是说,A/V源装置20及A/V目的地装置40可能进行以下两者:编码及解码(以及发射及接收)音频及视频数据。在一些实例中,音频编码器26可包含语音编码器,其还称作声码器。

[0072] 在图1的实例中,A/V源装置20包含音频源22及视频源24。举例来说,音频源22可包含麦克风,所述麦克风产生表示待由音频编码器26编码的所俘获音频数据的电信号。或者,音频源22可包含存储先前记录的音频数据的存储媒体、例如计算机化合成器的音频数据产生器,或任何其它音频数据源。视频源24可包含产生待由视频编码器28编码的视频数据的摄像机、编码有先前记录的视频数据的存储媒体、视频数据产生单元,或任何其它视频数据源。

[0073] 大体来说,源装置20可充当经编码视频内容的服务器。视频编码器28可以不同方式编码同一视频内容,例如使用不同编码标准及/或以不同位速率、帧速率或版本(例如,二维或三维)来进行编码。因此,同一视频内容可以各种不同表示而可用。源装置20可广告视频内容的不同表示的可用性。目的地装置40可能(例如)基于所述表示的特性及目的地装置40的解码及再现能力以及当前网络条件(例如,可用带宽)而选择所述表示中的一者。

[0074] 源装置20可大体上充当服务器,例如,根据HTTP版本1.1。因此,源装置20可根据串流网络协议及基于网络的串流标准(例如,DASH标准)将表示的数据提供到目的地装置40。

在一些情况下,如下文更详细地论述,表示可为可切换的。也就是说,目的地装置40可开始检索一个表示的数据,接着切换到不同的表示中间流。根据本发明的技术,目的地装置40可在两个表示之间进行无缝切换,而无需并行地解码两个表示且不会在重放过程中遇到间断。

[0075] 音频源22及视频源24可俘获原始音频及视频数据。原始音频及视频数据可包含模拟或数字数据。在通过音频编码器26及/或视频编码器28编码之前,模拟数据可经数字化。在说话参与者正在说话时,音频源22可获得来自说话参与者的音频数据,且视频源24可同时获得所述说话参与者的视频数据。在其它实例中,音频源22可包含包括所存储的音频数据的计算机可读存储媒体,且视频源24可包含包括所存储的视频数据的计算机可读存储媒体。以此方式,本发明中所描述的技术可应用于实况、串流、实时音频及视频数据或应用于经存档、预先记录的音频及视频数据。

[0076] 对应于视频帧的音频帧大体上为含有与在视频帧内所含有的由视频源24所俘获的视频数据同时发生的由音频源22所俘获的音频数据的音频帧。举例来说,当说话参与者大体上通过说话而产生音频数据时,音频源22俘获音频数据,且视频源24同时(即,在音频源22正俘获音频数据的同时)俘获说话参与者的视频数据。因此,音频帧可在时间上对应于一个或一个以上特定视频帧。因而,音频帧对应于视频帧大体上对应于同时俘获到音频数据及视频数据且音频帧及视频帧分别包含同时俘获到的音频数据及视频数据的情形。

[0077] 在一些实例中,音频编码器26可将表示记录每一经编码音频帧的音频数据的时间的时戳编码于所述经编码音频帧中,且类似地,视频编码器28可将表示记录每一经编码视频帧的视频数据的时间的时戳编码于所述经编码视频帧中。在这些实例中,音频帧对应于视频帧可包含音频帧包含时戳且视频帧包含相同时戳。A/V源装置20可包括内部时钟,音频编码器26及/或视频编码器28可从所述内部时钟产生时戳,或音频源22及视频源24可使用所述内部时钟来使音频及视频数据分别与时间戳相关联。

[0078] 在一些实例中,音频源22可将对应于记录音频数据的时间的数据发送到音频编码器26,且视频源24可将对应于记录视频数据的时间的数据发送到视频编码器28。在一些实例中,音频编码器26可将序列识别符编码于经编码音频数据中以指示经编码音频数据的相对时间排序但不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经编码视频数据的相对时间排序。类似地,在一些实例中,序列识别符可经映射或以其它方式与时间戳相关。

[0079] 本发明的技术大体上是针对于经编码多媒体(例如,音频及视频)数据的存储及输送,以及经输送多媒体数据的接收以及后续解译及解码。如图1的实例中所展示,视频源24可将场景的多个视图提供到视频编码器28。

[0080] 数据(无论是音频还是视频)的每一个别流称作基本流。基本流为节目的单一、经数字译码(可能经压缩)的分量。举例来说,所述节目的经译码视频或音频部分可为基本流。基本流可在包封于视频文件内之前转换成包化基本流(PES)。在同一节目内,使用流ID来区分属于一个基本流的PES包与属于另一基本流的PES包。基本流的数据的基本单元为包化基本流(PES)包。因此,MVC视频数据的每一视图对应于相应基本流。类似地,音频数据对应于一个或一个以上相应基本流。

[0081] 在图1的实例中,包封单元30接收来自视频编码器28的包含视频数据的基本流及

来自音频编码器26的包含音频数据的基本流。在一些实例中,视频编码器28及音频编码器26可各自包括用于从经编码数据形成PES包的包化器。在其它实例中,视频编码器28及音频编码器26可各自与用于从经编码数据形成PES包的相应包化器介接。在另外其它实例中,包封单元30可包括用于从经编码音频及视频数据形成PES包的包化器。

[0082] 表示(在所述术语在本发明中使用时)可包括音频数据与视频数据(例如,由A/V源装置20递送的音频基本流与视频基本流)的组合。多媒体内容可具有一个或一个以上表示,例如在视频数据以不同帧速率、不同位速率或不同分辨率来编码时。类似地,多媒体内容的音频数据可对应于各种语言,例如英语、法语及中文。表示可对应于音频基本流及一个或一个以上视频基本流(例如,以支持三维视频数据)的选择。

[0083] 包封单元30从音频编码器26及视频编码器28接收节目的基本流的PES包且从所述PES包形成对应网络抽象层(NAL)单元。在H.264/AVC(高级视频译码)的实例中,经译码视频段组织成NAL单元,其提供“网络友好的”视频表示寻址应用,例如视频电话、存储、广播或串流。NAL单元可分类为视频译码层(VCL)NAL单元及非VCL NAL单元。VCL单元可对应于由核心压缩引擎产生的数据,且可包括块、宏块及/或切片层级数据。其它NAL单元可为非VCL NAL单元。在一些实例中,一个时间实例中的经译码图片(通常呈现为主要经译码图片)可含于存取单元中,所述存取单元可包括一个或一个以上NAL单元。

[0084] 非VCL NAL单元可包括参数集NAL单元及SEI NAL单元,以及其它单元。参数集可含有序列层级标头信息(在序列参数集(SPS)中)及偶尔改变的图片层级标头信息(在图片参数集(PPS)中)。就参数集(例如,PPS及SPS)而论,无需针对每一序列或图片来重复偶尔改变的信息,因此可改进译码效率。此外,参数集的使用可实现重要标头信息的带外发射,从而避免为了错误恢复而进行冗余发射的需要。在带外发射实例中,参数集NAL单元可在与其它NAL单元(例如,SEI NAL单元)不同的信道上发射。

[0085] 补充增强信息(SEI)可含有并非为解码来自VCL NAL单元的经译码图片样本所必需但可辅助与解码、显示、错误恢复及其它目的有关的过程的信息。SEI消息可含于非VCL NAL单元中。SEI消息为一些标准规范的定标准部分,且因此对于符合标准的解码器实施方案来说并非总是必选的。SEI消息可为序列层级SEI消息或图片层级SEI消息。一些序列层级信息可含于SEI消息中,例如在SVC的实例中为可缩放性信息SEI消息及在MVC中为视图可缩放性信息SEI消息。这些实例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。

[0086] 根据本发明的技术,包封单元30可用信号发出两个或两个以上表示可有效地彼此切换。举例来说,包封单元30可确定所述两个或两个以上表示的初始化段是否可互换地使用以解码及/或再现所述表示的经编码样本。包封单元30可产生用以指示表示是否可有效地彼此切换的数据。因此,目的地装置40可请求所述数据(例如)以便在串流期间从一个表示切换到另一表示。

[0087] 包封单元30还可计算延迟值,所述延迟值表示当在两个可有效地切换的表示之间切换时的最大延迟。也就是说,所述延迟值可大体上表示当在两个表示之间切换时可预期的延迟的量(假定恒定的带宽可用性)。因此,当所述表示在时间上对准时,包封单元30可将所述延迟值计算为最大解码延迟与最大图片重排延迟的总和。当所述表示并不在时间上对准时(例如,当随机存取点出现在所述表示内的不同时间位置处时),包封单元30可进一步

计算重叠延迟且将所述重叠延迟包括于所述计算出的延迟值中。源装置20可(例如)响应于来自目的地装置40的对所述延迟值的请求而将所述延迟值提供到目的地装置40。

[0088] 目的地装置40可经配置以基于所述延迟值来检索并缓冲表示的足量数据以便在两个表示之间执行无缝切换。因此,目的地装置40可缓冲足量的数据,使得目的地装置40能够开始请求不同表示的数据且开始解码新表示的数据而不遭受显示中断。也就是说,在检索来自所述新表示的数据时,目的地装置40可解码并再现来自前一个表示的数据。此外,目的地装置40可经配置以使得两个表示的同时解码为不必要的。也就是说,在存取新表示的随机存取点之后,目的地装置40无需解码对应于视频重放中的同一时间位置的来自两个或两个以上表示的数据。此可改进电池性能,因为视频解码器48无需解码同一时间位置的多个样本。

[0089] 在自身有效负载中包括视频数据的NAL单元可包含各种粒度等级的视频数据。举例来说,NAL单元可包含视频数据的块、宏块、多个宏块、视频数据的切片,或视频数据的整个帧。包封单元30可从视频编码器28接收呈基本流的PES包的形式的经编码视频数据。包封单元30可使每一基本流与对应节目相关联。

[0090] 包封单元30还可组合来自多个NAL单元的存取单元。大体来说,存取单元可包含用于表示视频数据的帧以及对应于所述帧的音频数据(当此音频数据为可用时)的一个或一个以上NAL单元。存取单元大体上包括一个输出时间实例的所有NAL单元,例如,一个时间实例的所有音频及视频数据。举例来说,如果每一视图具有20帧/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。在此时间间隔期间,可同时再现同一存取单元(同一时间实例)的所有视图的特定帧。在对应于H.264/AVC的实例中,存取单元可包含一个时间实例中的经译码图片,所述经译码图片可呈现为主要经译码图片。因此,存取单元可包含共同时间实例的所有音频及视频帧,例如对应于时间X的所有视图。本发明还将特定视图的经编码图片称作“视图分量”。也就是说,视图分量可包含在特定时间的特定视图的经编码图片(或帧)。因此,存取单元可被定义为包含共同时间实例的所有视图分量。存取单元的解码次序无需必定与输出或显示次序相同。

[0091] 如同多数视频译码标准一样,H.264/AVC定义无错误位流的语法、语义及解码过程,其中的任一者遵照某一简档或层级。H.264/AVC不指定编码器,但编码器的任务是保证所产生的位流对于解码器来说为符合标准的。在视频译码标准的上下文中,“简档”对应于算法、特征或工具及施加到算法、特征或工具的约束的子集。如由H.264标准所定义,例如,“简档”为由H.264标准指定的整个位流语法的子集。“层级”对应于解码器资源消耗的限制,例如与图片的分辨率、位速率及宏块(MB)处理速率有关的解码器存储器及计算。可以profile_idc(简档指示符)值用信号发出简档,而可以level_idc(层级指示符)值用信号发出层级。

[0092] 举例来说,H.264标准认识到,在由给定简档的语法强加的边界内,仍有可能需要编码器及解码器的性能的大变化,此取决于位流中的语法元素所采取的值(例如,经解码图片的指定大小)。H.264标准进一步认识到,在许多应用中,实施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用也不经济。因此,H.264标准将“层级”定义为强加于位流中的语法元素的值的约束的指定集合。这些约束可为对各值的简单限制。或者,这些约束可采取对各值的算术组合(例如,图片宽度乘以图片高度乘以每秒经解码的图片的数目)的约

束的形式。H.264标准进一步提供,个别实施方案可针对每一所支持简档支持不同层级。

[0093] 遵照简档的解码器通常支持所述简档中定义的所有特征。举例来说,作为译码特征,B图片译码在H.264/AVC的基线简档中未受到支持,但在H.264/AVC的其它简档中受到支持。遵照层级的解码器应能够解码不需要超出所述层级中定义的限制的资源位的任何位流。简档及层级的定义可有助于可解译性。举例来说,在视频发射期间,可针对整个发射会话来协商一对简档及层级定义并实现一致。更具体来说,在H.264/AVC中,层级可定义(例如)对以下各者的限制:需要加以处理的宏块的数目、经解码图片缓冲器(DPB)大小、经译码图片缓冲器(CPB)大小、垂直运动向量范围、每两个连续MB的运动向量的最大数目,及B块是否具有少于 8×8 个像素的子宏块分割区。以此方式,解码器可确定所述解码器是否能够适当地解码所述位流。

[0094] 参数集大体上含有在序列参数集(SPS)中的序列层标头信息及在图片参数集(PPS)中的偶尔改变的图片层标头信息。就参数集而论,无需针对每一序列或图片来重复此偶尔改变的信息;因此,可改进译码效率。此外,参数集的使用可实现标头信息的带外发射,从而避免为实现错误恢复而进行冗余发射的需要。在带外发射中,参数集NAL单元在与其它NAL单元不同的信道上发射。

[0095] SVC及MVC的实例提供可缩放视频或多视图视频的序列参数集(SPS)扩展。SPS扩展信息可含于子SPS NAL单元中,所述子SPS NAL单元可具有与其它NAL单元不同的NAL单元类型。受支持操作点的简档及层级信息可包括于包括SPS扩展信息的NAL单元中。

[0096] 媒体表示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,且所述描述可包括(例如)编解码器信息、简档值及层级值。目的地装置40可检索媒体表示的MPD以确定如何存取各种呈现的电影片段。电影片段可位于视频文件的电影片段盒(moof盒)中。

[0097] 视频压缩标准(例如,ITU-T H.261、H.262、H.263、MPEG-1、MPEG-2及H.264/MPEG-4第10部分)利用运动补偿时间预测来减少时间冗余。编码器使用来自一些先前经编码图片(本文中还可称作帧)的运动补偿预测以根据运动向量来预测当前经译码图片。在典型视频译码中,存在三种主要图片类型。所述三种主要图片类型为经帧内译码图片(“I图片”或“I帧”)、预测图片(“P图片”或“P帧”)及双向预测图片(“B图片”或“B帧”)。P图片仅使用在时间次序上在当前图片之前的参考图片。在B图片中,可从一个或两个参考图片来预测所述B图片的每一块。这些参考图片在时间次序上可位于当前图片之前或之后。

[0098] 根据H.264译码标准,作为实例,B图片使用先前经译码参考图片的两个列表,即列表0及列表1。这两个列表可各自含有在时间次序上的过去及/或未来的经译码图片。可按以下若干方式中的一者来预测B图片中的块:从列表0参考图片的运动补偿预测、从列表1参考图片的运动补偿预测,或从列表0参考图片及列表1参考图片两者的组合的运动补偿预测。为了得到列表0参考图片及列表1参考图片两者的组合,分别从列表0参考图片及列表1参考图片获得两个运动补偿参考区域。可使用其组合来预测当前块。

[0099] ITU-T H.264标准支持各种块大小的帧内预测(例如,对于亮度分量的 16×16 、 8×8 或 4×4 ,以及对于色度分量的 8×8)以及各种块大小的帧间预测(例如,对于亮度分量的 16×16 、 16×8 、 8×16 、 8×8 、 8×4 、 4×8 及 4×4 ,以及对于色度分量的对应经缩放的大小)。在本发明中,“ $N \times N$ ”与“ N 乘 N ”可互换使用以指代就垂直及水平维度来说的块的像素尺寸,例

如, 16×16 个像素或 16 乘 16 个像素。大体来说, 16×16 块将在垂直方向上具有 16 个像素 ($y=16$) 且在水平方向上具有 16 个像素 ($x=16$)。同样, $N \times N$ 块大体上在垂直方向上具有 N 个像素且在水平方向上具有 N 个像素, 其中 N 表示非负整数值。可将块中的像素布置成列及行。块可在水平及垂直维度上具有不同数目个像素。也就是说, 块可包括 $N \times M$ 个像素, 其中 N 未必等于 M 。

[0100] 小于 16 乘 16 的块大小可称作 16 乘 16 宏块的分割区。视频块可包含像素域中的像素数据的块, 或变换域中的变换系数的块 (例如, 在应用例如以下各者的变换之后: 离散余弦变换 (DCT)、整数变换、小波变换, 或针对残余视频块数据的概念上类似的变换, 所述残余视频块数据表示经译码视频块与预测性视频块之间的像素差)。在一些情况下, 视频块可包含变换域中的经量化变换系数的块。

[0101] 较小视频块可提供较好分辨率, 且可用于视频帧的包括高细节等级的位置。大体来说, 可将宏块及各种分割区 (有时称作子块) 视为视频块。另外, 可将切片视为多个视频块, 例如宏块及/或子块。每一切片可为视频帧的可独立解码单元。或者, 帧自身可为可解码单元, 或帧的其它部分可被定义为可解码单元。术语“经译码单元”或“译码单元”可指代视频帧的任何可独立解码单元, 例如整个帧、帧的切片、还称作序列的图片群组 (GOP), 或根据适用的译码技术定义的另一可独立解码单元。

[0102] 术语“宏块”指代用于根据包含 16×16 个像素的二维像素阵列编码图片及/或视频数据的数据结构。每一像素包含色度分量及亮度分量。因此, 宏块可界定四个亮度块 (各自包含 8×8 个像素的二维阵列)、两个色度块 (各自包含 16×16 个像素的二维阵列) 及标头, 所述标头包含语法信息, 例如经译码块型样 (CBP)、编码模式 (例如, 帧内 (I) 或帧间 (P 或 B) 编码模式)、经帧内编码块的分割区的分割区大小 (例如, 16×16 、 16×8 、 8×16 、 8×8 、 8×4 、 4×8 或 4×4)、或经帧间编码宏块的一个或一个以上运动向量。

[0103] 视频编码器 28、视频解码器 48、音频编码器 26、音频解码器 46、包封单元 30 及解包封单元 38 在适当时各自可实施为多种合适处理电路中的任一者, 例如一个或一个以上微处理器、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 及视频解码器 48 中的每一者可包括于一个或一个以上编码器或解码器中, 其中的任一者可集成为组合式视频编码器/解码器 (编解码器) 的部分。同样, 音频编码器 26 及音频解码器 46 中的每一者可包括于一个或一个以上编码器或解码器中, 其中的任一者可集成为组合式编解码器的部分。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、包封单元 30 及/或解包封单元 38 的设备可包含集成电路、微处理器及/或无线通信装置 (例如, 蜂窝式电话)。

[0104] 在包封单元 30 已基于接收到的数据将 NAL 单元及/或存取单元组合为视频文件之后, 包封单元 30 将所述视频文件传递到输出接口 32 以供输出。在一些实例中, 包封单元 30 可在本地存储所述视频文件或经由输出接口 32 将所述视频文件发送到远程服务器, 而非直接将所述视频文件发送到目的地装置 40。举例来说, 输出接口 32 可包含发射器、收发器、用于将数据写入到计算机可读媒体的装置 (例如光盘驱动器、磁性媒体驱动器 (例如, 软盘驱动器))、通用串行总线 (USB) 端口、网络接口或其它输出接口。输出接口 32 将所述视频文件输出到计算机可读媒体 34, 例如发射信号、磁性媒体、光学媒体、存储器、快闪驱动器或其它计算机可读媒体。

[0105] 最终,输入接口36从计算机可读媒体34检索数据。举例来说,输入接口36可包含光盘驱动器、磁性媒体驱动器、USB端口、接收器、收发器或其它计算机可读媒体接口。输入接口36可将NAL单元或存取单元提供到解包单元38。解包单元38可将视频文件的元素解包封装成PES流,将所述PES流解包以检索经编码数据,且将所述经编码数据发送到音频解码器46或视频解码器48,此取决于所述经编码数据是音频流的部分还是视频流的部分,例如由所述流的PES包标头所指示。音频解码器46解码经编码音频数据且将所述经解码音频数据发送到音频输出42,而视频解码器48解码经编码视频数据且将所述经解码视频数据(其可包括流的多个视图)发送到视频输出44。

[0106] 图2为说明实例包单元30的组件的框图。在图2的实例中,包单元30包括视频输入接口80、音频输入接口82、视频文件创建单元60及视频文件输出接口84。在此实例中,视频文件创建单元60包括NAL单元构造器62、切换信息单元64及初始化段单元66。

[0107] 视频输入接口80及音频输入接口82分别接收经编码视频及音频数据。视频输入接口80及音频输入接口82可在视频及音频数据经编码时接收所述数据,或可从计算机可读媒体检索经编码视频及音频数据。在接收到经编码视频及音频数据后,视频输入接口80及音频输入接口82即刻将经编码视频及音频数据传递到视频文件创建单元60以组合成视频文件。

[0108] 视频文件创建单元60可对应于控制单元,所述控制单元包括经配置以执行归其所有的功能及程序的硬件、软件及/或固件。所述控制单元可进一步执行大体上归包单元30所有的功能。对于其中视频文件创建单元60以软件及/或固件来体现的实例来说,包单元30可包括:计算机可读媒体,其包含用于视频文件创建单元60的指令;及用以执行所述指令的基于硬件的处理单元。视频文件创建单元60的子单元(在此实例中为NAL单元构造器62、切换信息单元64及初始化段单元66)中的每一者可实施为个别硬件单元及/或软件模块,且可在功能上集成或进一步分离成额外子单元。

[0109] 视频文件创建单元60可对应于任何合适处理单元或处理电路,例如一个或一个以上微处理器(其可对应于单一集成电路的多个核心)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、数字信号处理器(DSP)或其任何组合。视频文件创建单元60可进一步包括:非暂时性计算机可读媒体,其存储用于NAL单元构造器62、切换信息单元64及初始化段单元66中的任一者或全部的指令;以及用于执行所述指令的处理器。

[0110] 大体来说,视频文件创建单元60可创建包括接收到的音频及视频数据的视频文件。NAL单元构造器62可形成包括经编码视频及音频样本的NAL单元。视频文件创建单元60可进一步经配置以组合包括经编码视频及音频样本的表示。在一些实例中,视频文件创建单元60可创建多个表示,所述多个表示可在一些实例中可彼此切换。

[0111] 切换信息单元64可经配置以产生用于在所述表示中的两者或两者以上之间进行切换的信息。举例来说,切换信息单元64可经配置以基于最大解码延迟、最大图片重排延迟及(在所述表示并不在时间上对准时)最大重叠延迟的总和来计算延迟值。客户端装置可使用所述延迟值来确定待缓冲的任何表示的数据的量以便从一个表示无缝切换到另一表示。因此,客户端装置可缓冲足量的数据以考虑与以下各者相关联的延迟:解码所缓冲的图片、重排所述图片,且确保呈现时间早于新表示的随机存取点的图片可在客户端装置检索、解码并再现新表示的数据的时间期间加以解码及再现。

[0112] 初始化段单元66可产生用于存取特定表示的初始化信息。大体来说,初始化段中的初始化信息不含有媒体数据。初始化段单元66可产生用于所述表示中的每一者的初始化段。然而,在一些情况下,可使用相同的初始化段来初始化多个表示。当此情形成立时,初始化段单元66可产生指示表示可有效地切换的信息。也就是说,如果第一表示与第二表示可有效地切换,则客户端装置可检索所述第一表示的初始化段,检索并解码所述第一表示的数据,接着开始基于所述第一表示的所述初始化段来检索并解码所述第二表示的数据,而无需检索所述第二表示的初始化段。

[0113] 在一些实例中,视频文件创建单元60可产生描述媒体呈现的媒体呈现描述(MPD)数据结构。根据本发明的技术,视频文件创建单元60可产生指示两个或两个以上表示是否可有效地彼此切换的经修改MPD数据结构及用于在两个表示之间无缝切换的延迟值。以下表1提供根据本发明的技术的实例MPD数据结构。表1为其中视频文件创建单元60以周期层级来用信号发出信息的实例。在其它实例中,可以其它层级来用信号发出信息。

[0114] “元素或属性名称”栏描述语法对象的名称。“类型”栏描述所述语法对象是元素还是属性。“基数”栏描述语法对象的基数,即,在对应于表1的数据结构的实例中所述语法对象的实例的数目。“任选性”栏描述语法对象是否为任选的,在此实例中,其中“M”指示必选,“O”指示任选,“OD”指示具有默认值的任选,且“CM”指示有条件的必选。“描述”栏描述对应语法对象的语义。

[0115] 在此实例中,视频文件创建单元60可设定单一旗标(seamlessSwitchingFlag(无缝切换旗标))以指示表示是否为可切换的。如果所述旗标设定为真,则客户端装置可假定所述表示可无缝切换,假设初始延迟等于或大于在MPD中用信号发出的minBufferTime。当此实例的seamlessSwitchingFlag设定为真时,客户端装置可进一步假定执行无缝切换不需要播放间断及表示的并行解码。客户端可遵照本发明的技术(例如,如下文关于图8所描述)来执行无缝切换。

[0116] 表1

[0117]

元素或属性名称	类型	基数	任选性	描述
MPD	E	1	M	携载媒体呈现的媒体呈现描述的根元素。
type	A		OD 默认: 随选	“随选”或“实况”。指示媒体呈现的类型。目前,定义了随选类型及实况类型。如果不存在,则应将呈现的类型推断为随选。
...				
minBufferTime	A		M	提供确保流畅播放所需的最初缓冲的媒体的最小量,限制条件为每一表示是以其 <i>bandwidth</i> 属性的值或高于所述值来递送的。
minBufferSwitching	A		M	提供确保表示之间的无缝切换所需的最

[0118]

元素或属性名称	类型	基数	任选性	描述
				初缓冲的媒体的最小量, 限制条件为每一表示是以其 <i>bandwidth</i> 属性的值或高于所述值来递送的。
...	A			
Period	E	1..N	M	提供周期的信息
...	A			...
Representation	E	1..N	M	可用以提供关于媒体呈现的原始来源(例如, 内容提供者)的信息。
...	A			
efficientSwitchGroup	A		O 默认: 假	当存在时, 指示所述表示为多视图视频加深度表示。
reUseInitSegmentFlag	A	0, 1	O 默认: 假	当为真时, 指示所述表示利用具有相同 repID 且具有最接近的呈现时间的前一个呈现的初始化段。

[0119] 图3为说明经译码视频帧100到132的序列的概念图。所述帧以不同方式来划阴影以指示在分级预测结构内的位置。举例来说, 帧100、116及132用黑色来划阴影以表示帧110、116、132处于所述分级预测结构的顶部处。帧100、116、132可包含(例如)在单一方向上从其它帧(例如, P帧)预测的经帧内译码帧或经帧间译码帧。当经帧内译码时, 帧100、116、132仅从同一帧内的数据来预测。当经帧间译码时, 例如, 帧116可相对于帧100的数据来译码, 如由从帧116到帧100的虚线箭头所指示。帧116、132分别形成图片群组(GOP)134、136的关键帧。

[0120] 帧108、124用暗黑色来划阴影以指示其为在编码分级结构中紧跟在帧100、116及132后面的帧。帧108、124可包含双向、帧间模式预测经编码帧。举例来说, 帧108可从帧100及116的数据来预测, 而帧124可从帧116及132来预测。帧104、112、120及128用较淡色来划阴影以指示其为在编码分级结构中紧跟在帧108及124后面的帧。帧104、112、120及128还可包含双向、帧间模式预测经编码帧。举例来说, 帧104可从帧100及108来预测, 帧112可从帧108及116来预测, 帧120可从帧116及124来预测, 且帧128可从帧124及132来预测。

[0121] 最后, 帧102、106、110、114、118、122、126及130用白色来划阴影以指示这些帧在编码分级结构中为最后的。帧102、106、110、114、118、122、126及130可为双向、帧间模式预测经编码帧。帧102可从帧100及104来预测, 帧106可从帧104及108来预测, 帧110可从帧108及112来预测, 帧114可从帧112及116来预测, 帧118可从帧116及120来预测, 帧122可从帧120及124来预测, 帧126可从帧124及128来预测, 且帧130可从帧128及132来预测。

[0122] 帧100到132是以显示次序来说明的。也就是说, 在解码之后, 帧100在帧102之前显示, 帧102在帧104之前显示, 等等。然而, 归因于编码分级结构, 帧100到132可以不同次序来加以解码。此外, 在经编码之后, 帧100到132可按解码次序来布置于位流中, 所述位流包括帧100到132的经编码数据。举例来说, 在GOP134的帧当中, 帧116可最后显示。然而, 归因于编码分级结构, 在GOP134中, 可首先解码帧116。也就是说, 例如, 为了适当地解码帧108, 可需要首先解码帧116, 以便使之充当帧108的参考帧。同样, 帧108可充当帧104、106、110及112的参考帧, 且因此可需要在帧104、106、110及112之前加以解码。

[0123] 显示帧的时间可称作呈现时间, 而解码帧的时间可称作解码时间。解码时间及呈现时间大体上提供相对于同一序列中的其它帧的时间排序的指示。帧的解码时间与帧的呈

现时间之间的差可称作帧的图片重排延迟。如上文所论述,源装置(例如,源装置20)可计算一组可切换表示的最大图片重排延迟以计算所述表示的将用信号发出的延迟值。

[0124] 以此方式,目的地装置(例如,目的地装置40)可使用所述用信号发出的延迟值来缓冲足量的数据以提供两个可切换表示之间的无缝切换。举例来说,假设帧100到132为第一表示的帧。目的地装置40可确定网络条件已改变且第二表示更适合于新的网络条件。目的地装置40可进一步确定在所述第二表示中存在具有与帧132相同的呈现时间的随机存取点。目的地装置40可接着检索来自所述第一表示的帧100到116中的每一者,接着检索来自所述第二表示的呈现时间等于帧118到132的呈现时间的帧。目的地装置40还可检索所述第二表示的具有与帧116相同的呈现时间的帧(如果所述随机存取点为开放解码刷新随机存取点),使得可适当地解码所述第二表示的对应于帧118到130的帧。

[0125] 图4为说明可切换表示154、160的概念图。图4说明原始视频序列150分别与表示154、160的片段158、164之间的对应性。在此实例中,表示154包括初始化段156及片段158A到158N,且表示160包括初始化段162及片段164A到164N。

[0126] 初始化段156、162可包括类似或相同的数据。因此,在一些实例中,客户端装置可检索初始化段156且基于表示154的初始化段156来解码表示160的片段164中的一者或一者以上的数据,而无需检索表示160的初始化段162。目的地装置可因此在表示154的片段158与表示160的片段164之间切换,而无需检索初始化段156、162两者。

[0127] 作为实例,在3GPP HTTP串流中,初始化段含有用于存取对应表示的初始化信息,但初始化段不含有媒体数据。当串流以H.264/AVC译码的视频时,可在视频的第一IDR图片之前存在若干个参数,例如序列参数集、图片参数集及/或一些序列层级补充增强信息(SEI)消息。那些参数可适用于整个表示,因此根据本发明的技术,在初始化段正在使用中时,所述参数可包括于初始化段中,而非将其包括在含有第一IDR的媒体段中。因此,初始化段156、162可分别包括表示154、160的序列参数集、图片参数集及/或序列层级SEI消息。

[0128] 图4的实例将片段158及片段164说明为具有不同大小。此将用以说明片段158所占用的存储器的量未必与片段164相同。也就是说,例如,片段158A在存储器消耗方面(及同样在经由网络来输送时在带宽使用方面)并不具有与片段164A相同的大小。然而,虚线说明片段158A及片段164A对应于同一视频段152A。虽然片段158A与164A为不同大小,但其对应于视频序列150的同一时间区。换句话说,片段158A的样本的开始呈现时间及结束呈现时间对应于片段164的样本的开始呈现时间及结束呈现时间。

[0129] 以此方式,片段158及164可被说成在时间上对准,假定随机存取点在表示154及160中出现在相同点中。也就是说,当随机存取点在片段158中的一者中时,片段164中的对应一者也具有随机存取点,且所述随机存取点共享解码次序值。表示的在时间上对准可简化在两个表示之间切换的过程。

[0130] 图5A为说明在时间上对准的两个表示180、190的实例的概念图。表示180包括随机存取点(RAP)182A、182B及样本184A到184D。随机存取点为具有以下额外性质的样本:解码器可解码所述随机存取点而无需接收到所述表示中领先于所述随机存取点的任何数据。因此,随机存取点可大体来说为经帧内模式编码的图片。

[0131] 表示180的随机存取点182及样本184以解码时间次序来加以说明,例如,在样本将布置于从服务器装置发送到客户端装置的位流中时。应理解,随机存取点182及样本184的

呈现时间未必匹配解码时间。

[0132] 图5A的表示190包括随机存取点192A、192B及样本194A到194D。再次,在此实例中,随机存取点192及样本194以解码时间次序来加以说明。如可看出,随机存取点182具有对应的随机存取点192。也就是说,随机存取点182A具有与随机存取点192A相同的解码时间,且随机存取点182B具有与随机存取点192B相同的解码时间。

[0133] 为从表示180切换到表示190,客户端装置可基于从服务器装置接收到的延迟值而缓冲表示180的某一数目个样本(其可包括随机存取点)。客户端装置可接着(例如)通过分析从服务器装置用信号发出的数据来确定表示180的下一随机存取点何时出现,所述数据指示表示180中的随机存取点的位置、解码时间及/或呈现时间。举例来说,假设表示180的下一随机存取点为随机存取点182B。客户端装置可接着以表示190的具有与表示180的所确定随机存取点相同的解码时间的随机存取点开始来请求表示190的数据。在此实例中,客户端装置将以随机存取点192B开始来开始请求表示190的数据,随机存取点192B具有与随机存取点182B相同的解码时间。

[0134] 以此方式,客户端装置可继续解码表示180的数据且开始检索表示190的数据,同时解码并重排表示180的经缓冲数据。在解码完表示180的所有经缓冲数据之后,客户端装置可立即开始解码并重排表示190的所检索数据。以此方式,客户端装置可执行两个表示之间的无缝切换。也就是说,客户端装置一次可仅解码一个表示(因此不需要两个或两个以上解码器来实现所述切换),且还可避免重放间断,所述间断原本可能因未缓冲足以执行所述切换的数据而引起。

[0135] 在上述实例中,表示180与190在时间上对准,从而使客户端装置能够简单地确定表示180的在经缓冲数据之后的随机存取点,接着在具有与表示180的所确定随机存取点相同的解码时间的随机存取点处请求表示190的数据。在其它实例中,两个表示并不在时间上对准,但仍可使用本发明的技术来在所述表示之间进行切换。

[0136] 图5B为说明并不在时间上对准的两个表示200、210的实例的概念图。表示200包括随机存取点202A、202B及样本204A到204D,而表示210包括随机存取点212A、212B及样本214A到214D。再次,图5B的实例中的随机存取点及样本以解码时间次序来加以说明。

[0137] 在图5B的实例中,表示200与210并不在时间上对准。具体来说,在此实例中,随机存取点202B在表示210中不具有对应的随机存取点(即,具有相同解码时间的随机存取点)。同样,随机存取点212B在表示200中不具有对应的随机存取点。大体来说,两个表示无需具有相同数目个随机存取点,且所述随机存取点无需在时间上对准。然而,本发明的技术仍可应用于执行并不在时间上对准的两个表示之间的切换。

[0138] 举例来说,为在表示200与表示210之间进行切换,客户端装置可基于从服务器装置接收到的延迟值而缓冲表示200的某一数目个样本(其可包括随机存取点)。所述服务器装置可将所述延迟值计算为最大解码延迟、最大图片重排延迟及(因为表示200与210并不在时间上对准)重叠延迟的总和。因此,通过基于所述延迟值来缓冲数据,客户端装置可缓冲足量的数据以确保表示200与210之间的无缝切换,此可考虑解码延迟、图片重排延迟及与表示并不在时间上对准相关联的延迟。

[0139] 表示200的经缓冲数据可包括与表示210的随机存取点212B重叠的片段。举例来说,所述片段可包括随机存取点202B以及样本204C及204D(且可能包括图5B中未说明的表

示200的后续样本)。所述片段在以下意义上可与随机存取点212B“重叠”:所述片段中的样本的最早呈现时间可早于随机存取点212B的呈现时间,但所述片段中的样本的最迟呈现时间可迟于随机存取点212B的呈现时间。此片段被说成与所述随机存取点重叠。本发明还可将此片段称作“横跨”所述随机存取点。所述片段可对应于表示200的最后缓冲的片段。也就是说,客户端装置可请求来自表示200的数据直到且包括与表示210的随机存取点重叠的片段。

[0140] 客户端装置还可定位表示210的包括随机存取点212B的片段。客户端装置可以包括随机存取点212B的片段开始来请求来自表示210的数据。客户端装置可解码从表示200检索的数据直到与随机存取点212B重叠的片段中的图片群组。在到达所述图片群组后,客户端装置即刻可开始解码表示210的数据。客户端装置可确定表示200中是否存在呈现时间早于随机存取点212B的呈现时间的任何图片,且在解码随机存取点212B之前解码任何这些图片。以此方式,客户端装置可在表示200与210之间进行无缝切换。也就是说,客户端装置可避免视频重放间断,同时还避免并行解码,即,一次操作两个或两个以上解码器来解码所述表示。

[0141] 虽然在此实例中可能会检索最终未解码的某一量的数据,但此对少量额外数据的检索被无需执行表示的并行解码的益处胜过且还允许无缝重放。因此,检索额外数据的成本可被视为可忽略的。

[0142] 图6为说明用于执行两个可切换表示之间的无缝切换的实例方法的流程图。虽然关于图1的源装置20及目的地装置40来描述,但应理解,图6的技术可通过任何服务器装置及客户端装置来执行。此外,所述方法中的步骤无需必定以图6中所说明的精确次序来发生,且可添加额外步骤或可移除一些步骤,而不偏离本发明的技术。

[0143] 最初,源装置20可将描述两个或两个以上可切换表示的数据提供到目的地装置40(250)。举例来说,此数据可描述流畅重放的最小网络带宽要求、用以解码及/或再现所述表示的经编码视频数据的解码器装置的简档及/或层级要求、随机存取点的位置(例如,片段及/或字节范围)、随机存取点的解码时间、随机存取点的呈现时间、所述表示的片段的解码及/或呈现时间、所述片段的字节范围、包括所述随机存取点的统一资源定位符,或用于解码及/或再现经编码视频数据的其它此类数据。

[0144] 所述数据可进一步指示所述两个或两个以上表示(或其某子集)是否形成有效切换群组,且所述两个或两个以上表示中的任一者是否在时间上对准。此外,所述数据可包括用于在可切换表示之间进行切换的延迟值。所述延迟值可对应于所述可切换表示的最大解码延迟与最大图片重排延迟的总和。另外,对于并不在时间上对准的表示,所述延迟值可考虑重叠延迟。

[0145] 目的地装置40可接收描述所述两个或两个以上表示的数据(252)。目的地装置40可接着确定当前网络条件(254)(例如,可用带宽)以最初选择所述表示中的一者。另外,目的地装置40可确定其自己的解码及/或再现能力以在所述表示之间进行选择。也就是说,在目的地装置40能够解码及再现的表示中,目的地装置40可基于当前网络条件来选择所述表示中的一者。举例来说,目的地装置40可选择具有最接近于但不超过当前可用网络带宽的带宽要求的第一表示。

[0146] 目的地装置40可接着请求来自第一表示的数据(256)。举例来说,目的地装置40可

产生指定所述第一表示的片段且可能指定所述片段的字节范围的HTTP GET或部分GET请求。响应于所述请求,源装置20可将由目的地装置40请求的第一表示的经编码样本发送到目的地装置40(258)。目的地装置40可解码并显示接收到的样本(260)。假定网络条件不改变,则目的地装置40可继续请求所述第一表示的样本,接着在所述样本被接收到之后解码并显示所述样本。

[0147] 在图6的实例中,目的地装置40可确定网络条件的改变(262)。举例来说,目的地装置40可确定可用网络带宽已增加或减小。响应于确定带宽增加,目的地装置40可选择具有较高带宽要求及对应的较高质量的第二表示。另一方面,响应于确定带宽减小,目的地装置40可选择具有较低带宽要求但具有较低质量的第二表示。

[0148] 在确定网络条件已改变之后,目的地装置40可基于从源装置20接收到的延迟值来缓冲所述第一表示的某量的数据(264)。举例来说,目的地装置40可基于所述延迟值来确定待从所述第一表示检索的样本的数目或片段的数目以允许从所述第一表示到所述第二表示的无缝切换。缓冲数据可包括产生对来自所述第一表示的数据的一个或一个以上请求及将所述请求发送到源装置20。在缓冲来自所述第一表示的此数据之后,目的地装置40可开始请求来自选定第二表示的数据(266)。

[0149] 响应于所述请求,源装置20可将第二表示的经编码样本发送到目的地装置40(268)。在此时间期间,目的地装置40可继续解码第一表示的经缓冲数据。在接收到来自所述第二表示的所请求数据之后且在解码所述第一表示的具有在所述第二表示的随机存取点之前的呈现时间的样本之后,目的地装置40可开始解码并显示所述第二表示的接收到的样本(270)。在一些实例中,目的地装置40可请求来自所述第二表示的数据而不请求所述第二表示的初始化段。因此,目的地装置40可使用从所述第一表示检索的初始化段来解码并显示所述第二表示的样本。

[0150] 图7为说明服务器装置(例如,源装置20)将描述可切换表示的信息提供到客户端装置(例如,目的地装置40)的实例方法的流程图。虽然关于图1的源装置20及目的地装置40来描述,但应理解,图7的技术可通过任何服务器装置及客户端装置来执行。此外,所述方法中的步骤无需必定以图7中所说明的精确次序来发生,且可添加额外步骤或可移除一些步骤,而不偏离本发明的技术。

[0151] 最初,源装置20可接收两个或两个以上可切换表示的经编码视频样本(280)。在一些实例中,源装置20可接收原始视频数据并编码两个或两个以上可切换表示。在其它实例中,源装置20可从包括一个编码器或多个编码器的中间装置接收经编码表示,其产生所述可切换表示。所述表示在一些实例中可在时间上对准,或在其它实例中可并不在时间上对准。

[0152] 源装置20可确定所述可切换表示的接收到的样本的最大解码延迟(282)。举例来说,源装置20可分析每一表示的样本以确定每一表示的最大解码延迟。接着,基于这些所确定的最大值,源装置20可计算所有所述表示的最大解码延迟。举例来说,假设源装置20接收三个表示的样本。源装置20可确定第一表示的第一最大解码延迟、第二表示的第二最大解码延迟及第三表示的第三最大解码延迟。源装置20可接着将所述第一最大解码延迟、所述第二最大解码延迟及所述第三最大解码延迟中的最大值计算为可切换表示的最大解码延迟。所述样本可包含表示的存取单元,使得所述最大解码延迟对应于与解码所述可切换表

示中的任一者的任何存取单元相关联的最大延迟。

[0153] 源装置20还可确定所述样本的最大重排延迟(284)。再次,源装置20可分析所述可切换表示的样本以计算每一表示的最大重排延迟,接着计算这些最大重排延迟中的最大值以确定与所述可切换表示相关联的最大重排延迟。大体来说,图片重排延迟为表示的帧的呈现时间与解码时间之间的差。因此,所述最大图片重排延迟可对应于所述可切换表示中的任一者的任一帧的呈现时间与解码时间之间的最大差。

[0154] 源装置20可接着将与所述可切换表示之间的无缝切换相关联的延迟值计算(例如)为最大解码延迟与最大图片重排延迟的总和(286)。源装置20可进一步确定所述表示是否在时间上对准(288)。举例来说,源装置20可仅在以下时间时才确定所述表示在时间上对准:对于所述表示中的一者中的任何随机存取点来说,其它表示中的每一者具有解码时间与所述表示中的所述一者的随机存取点相同的随机存取点。源装置20可从解码装置接收关于所述表示是否在时间上对准的指示,或当源装置20编码所述表示时,可持续跟踪是否产生将在时间上对准的表示。

[0155] 如果所述表示在时间上对准(288的“是”分支),则源装置20可简单地存储计算出的延迟值及经编码视频样本(294)。另一方面,如果所述表示并不在时间上对准(288的“否”分支),则源装置20可将在286处计算出的延迟值视为中间延迟值。源装置20可接着计算所述可切换表示的最大重叠延迟(290)。图片的重叠延迟可基于当前表示的GOP结构以及所述图片在所述GOP中的位置。源装置20可计算所述可切换表示的最大重叠延迟,接着将所述最大重叠延迟加到在步骤286处计算出的延迟值(292)。源装置20可接着存储此计算出的延迟值连同所述可切换表示的视频样本(294)。举例来说,源装置20可产生信令信息,源装置20可响应于此信令信息的请求而将所述信令信息提供到目的地装置40。

[0156] 图8为说明客户端装置(例如,目的地装置40)自从服务器装置(例如,源装置20)检索第一表示的数据切换到从所述服务器装置检索第二表示的数据的实例方法的流程图。虽然关于图1的源装置20及目的地装置40来描述,但应理解,图8的技术可通过任何服务器装置及客户端装置来执行。此外,所述方法中的步骤无需必定以图8中所说明的精确次序来发生,且可添加额外步骤或可移除一些步骤,而不偏离本发明的技术。此外,图8的实例假设所述两个表示并不在时间上对准。

[0157] 最初,目的地装置40可(例如)从源装置20接收两个或两个以上可切换表示的延迟值(300)。目的地装置40可接收描述所述可切换表示的额外信息,例如随机存取点的位置、所述表示是否在时间上对准、描述客户端装置成功地解码/再现表示的解码及/或再现要求的简档及/或层级指示符、所述表示的带宽限制,或其它此类信息。

[0158] 目的地装置40可接收、解码并显示第一表示的数据(302)。举例来说,目的地装置40可基于当前网络条件(例如,可用带宽及目的地装置40的解码及再现能力)来选择第一表示。目的地装置40可接着提交对来自所述第一表示的数据的一个或一个以上请求,且在基于接收到的延迟值缓冲了足量的数据之后开始解码并显示所述第一表示的数据。

[0159] 在某点处,目的地装置40可确定网络条件已存在改变(304),例如,可用带宽已增加或减小。因此,基于新网络条件,目的地装置40可选择可切换到所述第一表示的第二表示(306)。如上文所提,假定在此实例中所述第一表示与所述第二表示并不在时间上对准。

[0160] 目的地装置40可确定包括所述第二表示中的下一随机存取点(RAP)的片段(308),

在此实例中被标记为“片段M”。也就是说,目的地装置40可确定所述第二表示的呈现时间迟于所述第一表示的经缓冲数据的呈现时间的随机存取点。目的地装置40可接着确定片段M,其包括所述随机存取点。

[0161] 目的地装置40可接着在所述第一表示中定位横跨所述第二表示的随机存取点的片段(310),在此实例中被标记为“片段N”。也就是说,目的地装置40可识别具有以下两个呈现时间的片段N:早于所述第二表示的随机存取点的呈现时间的最早呈现时间;及迟于所述第二表示的随机存取点的呈现时间的最迟呈现时间。

[0162] 目的地装置40可请求来自所述第一表示的数据直到且包括片段N(312)。在请求片段N之后,目的地装置40可以片段M开始来请求来自所述第二表示的数据(314)。目的地装置40可解码第一表示的所有样本直到横跨片段M的随机存取点的呈现时间的图片群组(316)。

[0163] 目的地装置40可确保呈现时间早于片段M中的随机存取点的呈现时间的所有图片已经解码并显示,所述图片可包括上文所提及的图片群组内的图片。此后,目的地装置40可以所述随机存取点开始来开始解码所述第二表示的图片(318)。目的地装置40可进一步显示所述第二表示的经解码图片,此可能需要某量的图片重排。

[0164] 另外,目的地装置40可请求并解码所述第二表示的片段而无需请求所述第二表示特有的初始化段。举例来说,当接收到的信息指示所述第一表示及所述第二表示可有效地切换时,目的地装置40可基于针对所述第一表示所接收到的初始化段来解码并再现所述第二表示的样本。因此,图8的方法可进一步包括在接收、解码并显示所述第一表示的数据之前(例如,在步骤302之前)请求并接收所述第一表示的初始化段。

[0165] 在一个或一个以上实例中,可以硬件、软件、固件或其任何组合来实施所描述的功能。如果以软件来实施,则功能可作为一个或一个以上指令或代码而存储于计算机可读媒体上或经由计算机可读媒体来传输且由基于硬件的处理单元来执行。计算机可读媒体可包括对应于例如数据存储媒体的有形媒体的计算机可读存储媒体,或包括促进计算机程序(例如)根据通信协议从一处传送到另一处的任何媒体的通信媒体。以此方式,计算机可读媒体大体上可对应于(1)非暂时性的有形计算机可读存储媒体或(2)例如信号或载波的通信媒体。数据存储媒体可为可由一个或一个以上计算机或一个或一个以上处理器存取以检索用于实施本发明中所描述的技术的指令、代码及/或数据结构的任何可用媒体。短语“一个或一个以上处理器”可包括多核心处理器的一个或一个以上核心,其中的每一者可包括于共同芯片(例如,单一集成电路)内。计算机程序产品可包括计算机可读媒体。

[0166] 借助于实例而非限制,所述计算机可读存储媒体可包含RAM、ROM、EEPROM、CD-ROM或其它光盘存储装置、磁盘存储装置或其它磁性存储装置、快闪存储器,或可用于存储呈指令或数据结构的形式的所要程序代码且可由计算机存取的任何其它媒体。而且,任何连接被适当地称作计算机可读媒体。举例来说,如果使用同轴电缆、光纤电缆、双绞线、数字订户线(DSL)或例如红外线、无线电及微波等无线技术而从网站、服务器或其它远程源发射指令,则同轴电缆、光纤电缆、双绞线、DSL或例如红外线、无线电及微波等无线技术包括于媒体的定义中。然而,应理解,计算机可读存储媒体及数据存储媒体不包括连接、载波、信号或其它瞬时媒体,而是改为针对非瞬时、有形的存储媒体。如本文中所使用,磁盘及光盘包括压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软性磁盘及蓝光光盘,其中磁盘通常以磁性方式再生数据,而光盘通过激光以光学方式再生数据。以上各者的组合也应

包括于计算机可读媒体的范围内。

[0167] 可由例如一个或一个以上数字信号处理器(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等一个或一个以上处理器来执行指令。因此,如本文中所使用,术语“处理器”可指代上述结构或适于实施本文中所描述的技术的任何其它结构中的任一者。另外,在一些方面中,本文所描述的功能性可提供于经配置以用于编码及解码的专用硬件及/或软件模块内,或并入于组合式编解码器中。而且,所述技术可完全实施于一个或一个以上电路或逻辑元件中。

[0168] 本发明的技术可实施于广泛多种装置或设备中,包括无线手持机、集成电路(IC)或IC集合(例如,芯片组)。在本发明中描述各种组件、模块或单元以强调经配置以执行所揭示的技术的装置的功能方面,但未必需要通过不同硬件单元来实现。而是,如上文所描述,各种单元可组合于编解码器硬件单元中或由交互操作的硬件单元的集合(包括如上文所描述的一个或一个以上处理器)结合合适的软件及/或固件来提供。

[0169] 已描述了各种实例。这些及其它实例属于所附权利要求书的范围内。

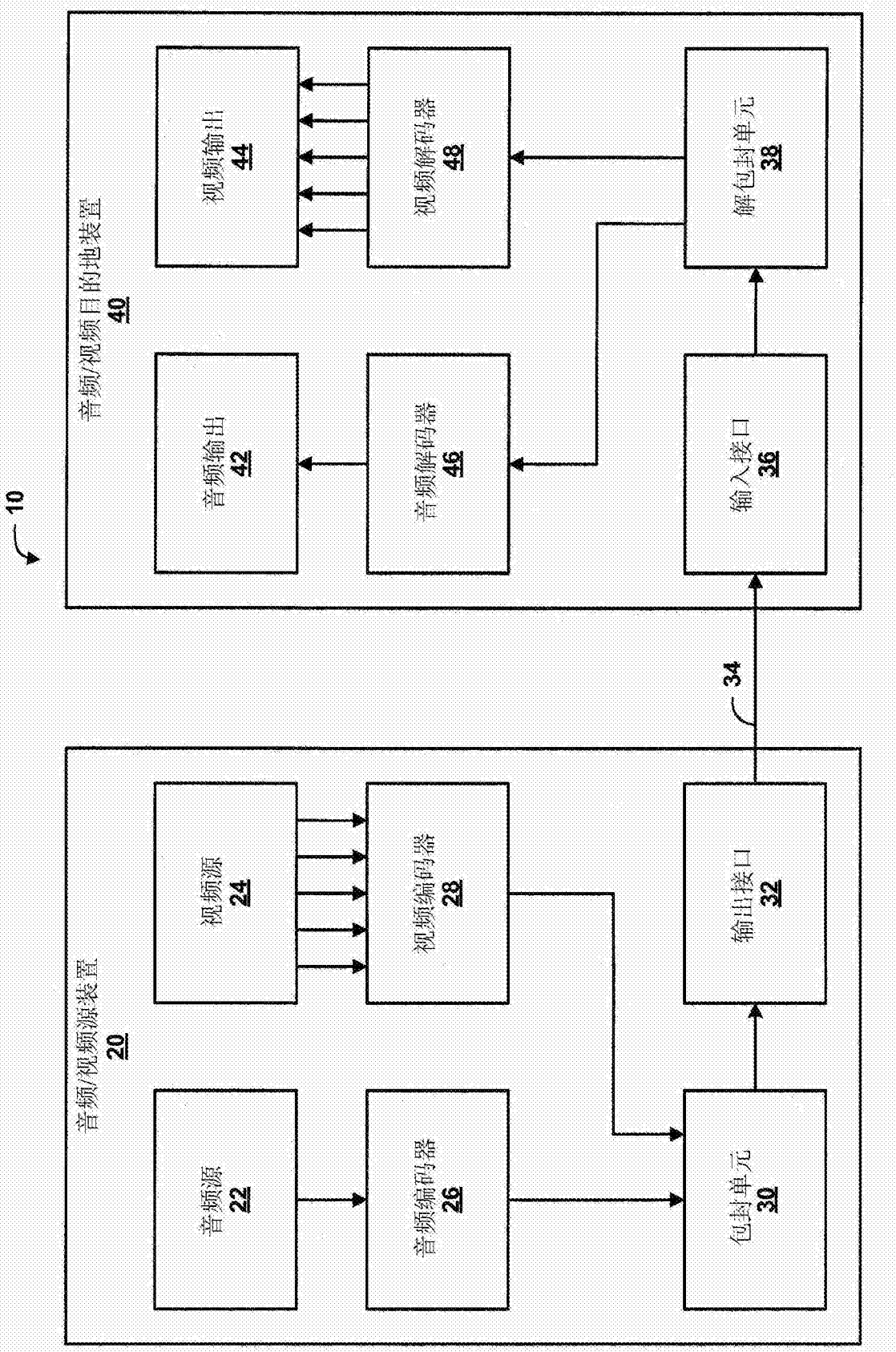


图1

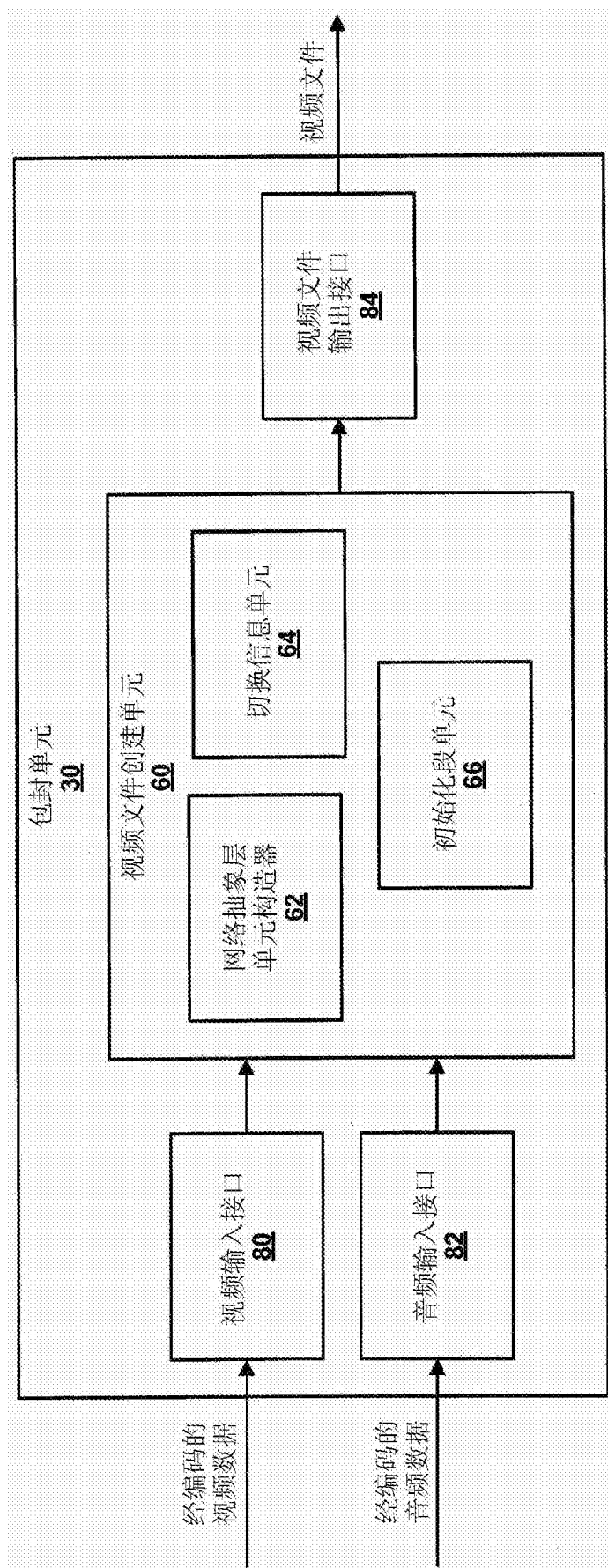


图2

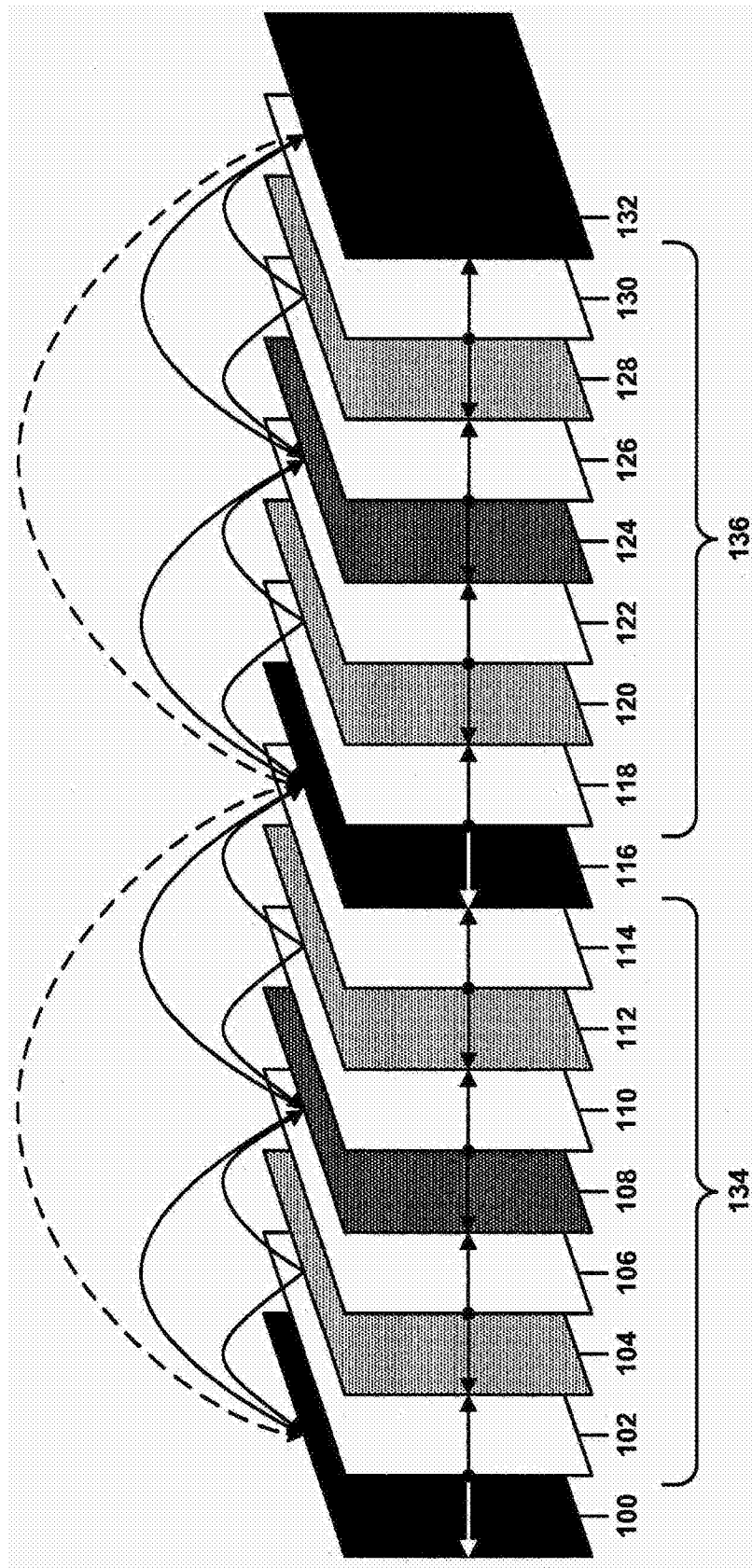


图3

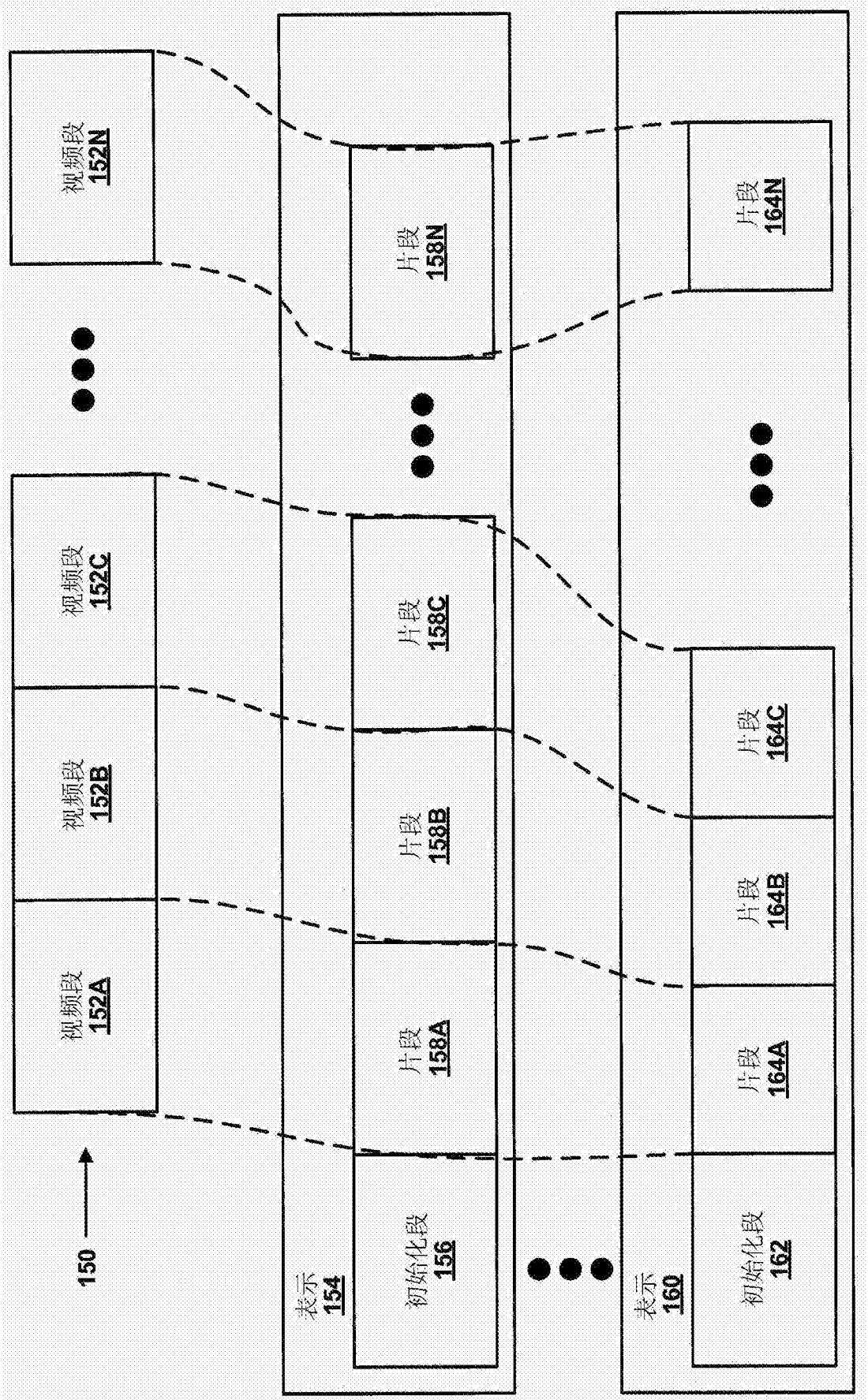


图4

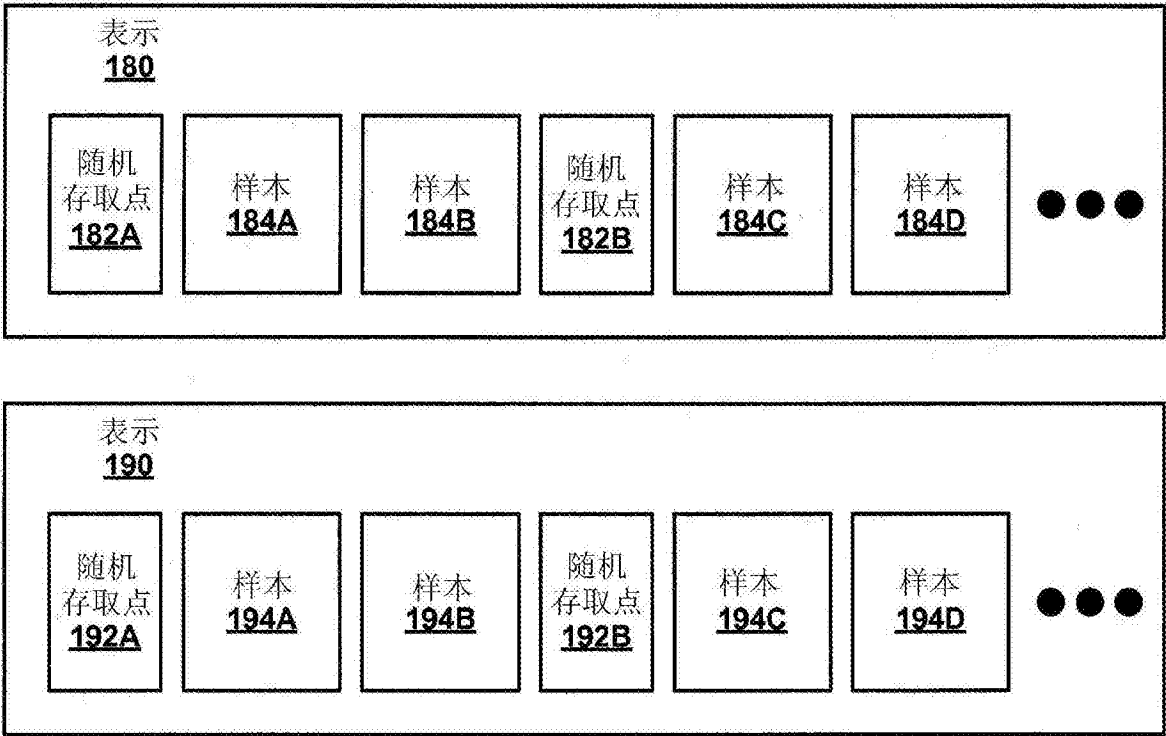


图5A

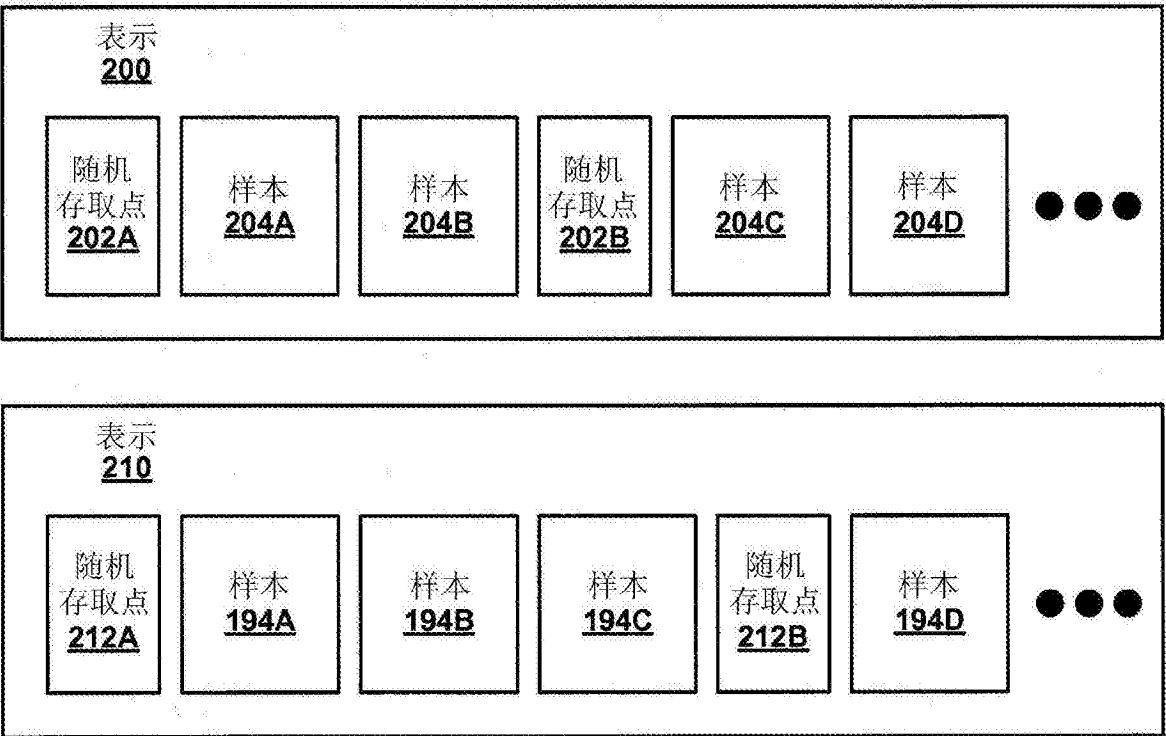


图5B

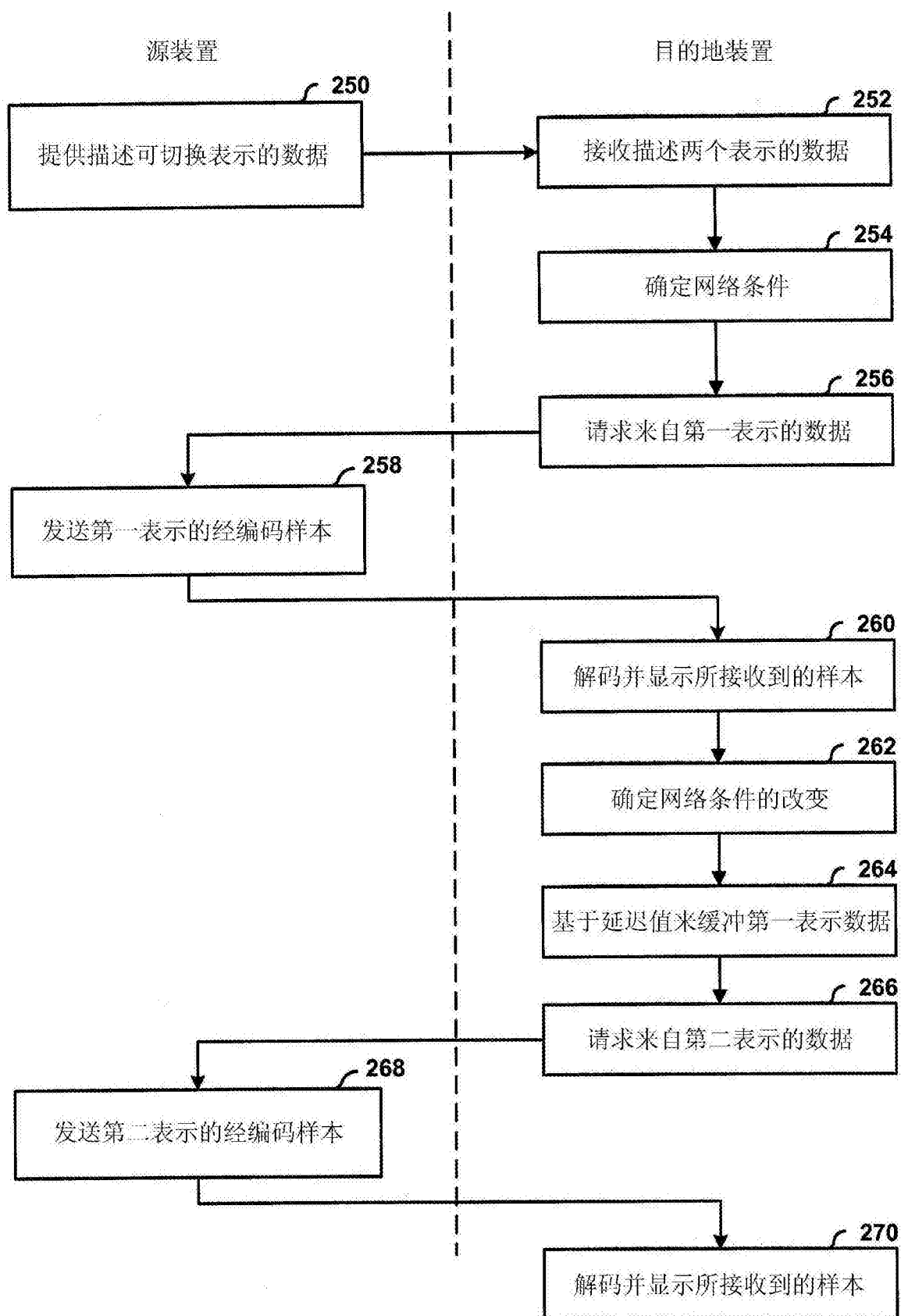


图6

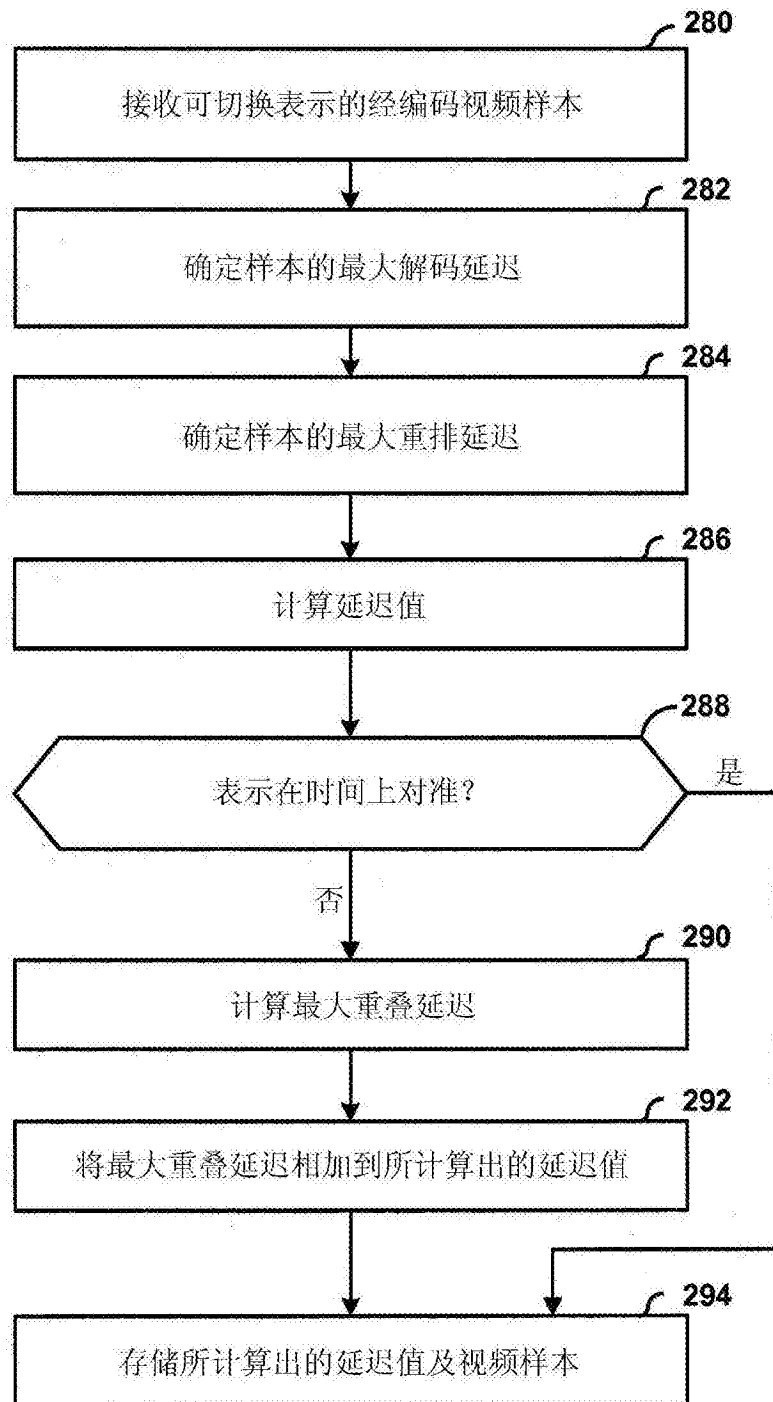


图7

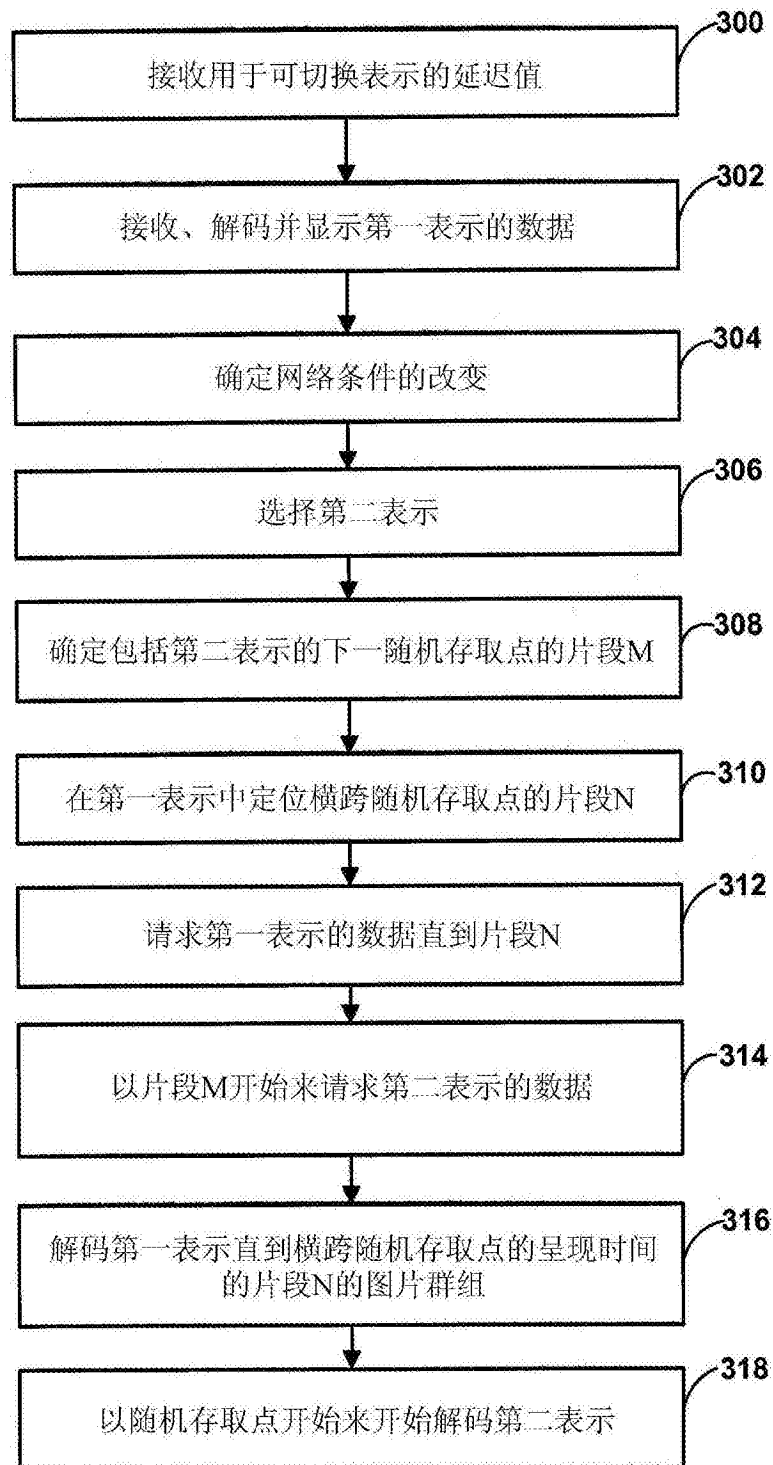


图8