



(12)发明专利

(10)授权公告号 CN 105744295 B

(45)授权公告日 2019.07.12

(21)申请号 201610086810.2

H04N 21/2381(2011.01)

(22)申请日 2011.07.20

H04N 21/435(2011.01)

(65)同一申请的已公布的文献号

H04N 21/436(2011.01)

申请公布号 CN 105744295 A

H04N 21/4363(2011.01)

(43)申请公布日 2016.07.06

H04N 21/438(2011.01)

(30)优先权数据

H04N 21/643(2011.01)

12/840,020 2010.07.20 US

H04N 21/84(2011.01)

(62)分案原申请数据

H04N 21/845(2011.01)

201180040329.5 2011.07.20

H04N 21/854(2011.01)

(73)专利权人 高通股份有限公司

(56)对比文件

地址 美国加利福尼亚州

CN 1666195 A, 2005.09.07,

(72)发明人 陈盈 马尔塔·卡切维奇

CN 1666526 A, 2005.09.07,

(74)专利代理机构 北京律盟知识产权代理有限公司
责任公司 11287

CN 101218828 A, 2008.07.09,

代理人 宋献涛

CN 105744295 A, 2016.07.06,

EP 1383333 A2, 2004.01.21,

US 2011228166 A1, 2011.09.22,

审查员 张露文

(51)Int.Cl.

H04N 21/235(2011.01)

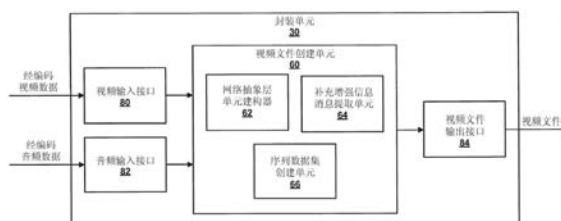
权利要求书4页 说明书23页 附图7页

(54)发明名称

提供用于串流传输视频数据的序列数据集

(57)摘要

本申请涉及提供用于串流传输视频数据的序列数据集。一种装置可封装视频数据以使得补充增强信息SEI消息与由所述SEI消息描述的经验码视频图片序列分离地存储。实例装置包含：控制单元，其经配置以产生与所述经验码视频图片分离的一个或一个以上SEI消息，其中所述SEI消息描述所述经验码视频图片序列中的相应者且包含所述经验码视频图片中的一者以上所共同的元素；以及输出接口，其经配置以与所述经验码视频图片序列分离地输出所述SEI消息。实例目的地装置可与所述经验码视频图片分离地接收所述SEI消息且使用所述SEI消息来再现所述经验码视频图片。



1. 一种接收经编码视频数据的方法,所述方法包括:

接收包括描述媒体表示的经译码视频图片序列的经封装补充增强信息SEI消息的初始化区段,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素,其中所述初始化区段包括所述媒体表示的多个文件中的第一文件,其中所述初始化区段与统一资源定位符URL相关联,且其中所述第一文件和所述初始化区段不包含任何媒体数据;

接收所述多个文件的与所述第一文件分离的一个或多个文件,其中所述一个或多个文件封装视频文件的所述经译码视频图片序列且与所述第一文件分离,且其中用于所述一个或多个文件的URL不同于和所述初始化区段相关联的所述URL;以及

在所述经译码视频图片序列已经解码之后基于所述SEI消息而再现所述经译码视频图片序列。

2. 根据权利要求1所述的方法,进一步包括:

提交对所述SEI消息的第一请求,其中接收所述SEI消息包括响应于所述第一请求而接收所述SEI消息;以及

仅在接收到所述SEI消息之后提交对所述经译码视频图片序列的第二请求,其中接收所述经译码视频图片序列包括响应于所述第二请求而接收所述经译码视频图片序列。

3. 根据权利要求2所述的方法,其中所述SEI消息包括第一组SEI消息,且其中所述经译码视频图片序列包括第一经译码视频图片序列,所述方法进一步包括:

在接收到所述第一经译码视频图片序列之后提交对第二组SEI消息的第三请求;

响应于所述第三请求而接收所述第二组SEI消息;

提交对由所述第二组SEI消息描述的第二经译码视频图片序列的第四请求;

在接收到所述第二组SEI消息之后接收所述第二经译码视频图片序列;以及

在所述第二经译码视频图片序列已经解码之后基于所述第二组SEI消息而再现所述第二经译码视频图片序列。

4. 根据权利要求1所述的方法,其进一步包括从参数集轨道、不包含经译码视频图片的轨道片段和不包含经译码视频数据的第三代合作伙伴计划3GPP格式文件中的至少一者提取所述SEI消息。

5. 根据权利要求1所述的方法,其进一步包括剖析和解码所述SEI消息,其中所述SEI消息是与包括所述经译码视频图片序列的存取单元分离地译码。

6. 一种用于接收经编码视频数据的设备,所述设备包括:

控制单元,其经配置以:

接收包括描述媒体表示的经译码视频图片序列的经封装补充增强信息SEI消息的初始化区段,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素,其中所述初始化区段包括所述媒体表示的多个文件中的第一文件,其中所述初始化区段与统一资源定位符URL相关联,且其中所述第一文件和所述初始化区段不包含任何媒体数据;以及

接收所述多个文件的与所述第一文件分离的一个或多个文件,其中所述一个或多个文件封装视频文件的所述经译码视频图片序列且与所述第一文件分离,且其中用于所述一个或多个文件的URL不同于所述初始化区段相关联的所述URL;以及

视频输出,其经配置以在所述经译码视频图片序列已经解码之后再现所述经译码视频图片序列。

7. 根据权利要求6所述的设备,其中所述控制单元经配置以提交对所述SEI消息的第一请求,响应于所述第一请求而接收所述SEI消息,仅在接收到所述SEI消息之后提交对所述经译码视频图片序列的第二请求,且响应于所述第二请求而接收所述经译码视频图片序列。

8. 根据权利要求7所述的设备,其中所述SEI消息包括第一组SEI消息,且其中所述经译码视频图片序列包括第一经译码视频图片序列,其中所述控制单元经配置在接收到所述第一经译码视频图片序列之后提交对第二组SEI消息的第三请求,响应于所述第三请求而接收所述第二组SEI消息,提交对由所述第二组SEI消息描述的第二经译码视频图片序列的第四请求,以及在接收到所述第二组SEI消息之后接收所述第二经译码视频图片序列,且

其中所述视频输出经配置以在所述第二经译码视频图片序列已经解码之后基于所述第二组SEI消息而再现所述第二经译码视频图片序列。

9. 根据权利要求6所述的设备,其中所述控制单元经配置以从参数集轨道、不包含经译码视频图片的轨道片段和不包含经译码视频数据的第三代合作伙伴计划3GPP文件中的至少一者提取所述SEI消息。

10. 根据权利要求6所述的设备,其中所述设备包括以下各项中的至少一者:

集成电路;

微处理器;以及

无线通信装置,其包含所述控制单元。

11. 一种用于接收经编码视频数据的设备,所述设备包括:

用于接收包括描述媒体表示的经译码视频图片序列的补充增强信息SEI消息的初始化区段的装置,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素,其中所述初始化区段包括所述媒体表示的多个文件中的第一文件,其中所述初始化区段与统一资源定位符URL相关联,且其中所述第一文件和所述初始化区段不包含任何媒体数据;

用于接收所述多个文件的与所述第一文件分离的一个或多个文件的装置,其中所述一个或多个文件封装视频文件的所述经译码视频图片序列且与所述第一文件分离,且其中用于所述一个或多个文件的URL不同于和所述初始化区段相关联的所述URL;以及

用于在所述经译码视频图片序列已经解码之后基于所述SEI消息而再现所述经译码视频图片序列的装置。

12. 根据权利要求11所述的设备,其进一步包括:

用于提交对所述SEI消息的第一请求的装置,其中所述用于接收所述SEI消息的装置包括用于响应于所述第一请求而接收所述SEI消息的装置;以及

用于仅在接收到所述SEI消息之后提交对所述经译码视频图片序列的第二请求的装置,其中所述用于接收所述经译码视频图片序列的装置包括用于响应于所述第二请求而接收所述经译码视频图片序列的装置。

13. 根据权利要求12所述的设备,其中所述SEI消息包括第一组SEI消息,且其中所述经译码视频图片序列包括第一经译码视频图片序列,所述设备进一步包括:

用于在接收到所述第一经译码视频图片序列之后提交对第二组SEI消息的第三请求的装置；

用于响应于所述第三请求而接收所述第二组SEI消息的装置；

用于提交对由所述第二组SEI消息描述的第二经译码视频图片序列的第四请求的装置；

用于在接收到所述第二组SEI消息之后接收所述第二经译码视频图片序列的装置；以及

用于在所述第二经译码视频图片序列已经解码之后基于所述第二组SEI消息而再现所述第二经译码视频图片序列的装置。

14. 根据权利要求13所述的设备，其进一步包括用于从参数集轨道、不包含经译码视频图片的轨道片段和不包含经译码视频数据的第三代合作伙伴计划3GPP文件中的至少一者提取所述SEI消息的装置。

15. 一种包括非暂时性计算机可读媒体的计算机程序产品，所述计算机可读媒体上存储有指令，所述指令在执行时致使用于接收经编码视频数据的设备的处理器进行以下操作：

接收包括描述媒体表示的经译码视频图片序列的补充增强信息SEI消息的初始化区段，其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素，其中所述初始化区段包括所述媒体表示的多个文件中的第一文件，其中所述初始化区段与统一资源定位符URL相关联，且其中所述第一文件和所述初始化区段不包含任何媒体数据；

接收所述多个文件的与所述第一文件分离的一个或多个文件，其中所述一个或多个文件封装视频文件的所述经译码视频图片序列且与所述第一文件分离，且其中用于所述一个或多个文件的URL不同于和所述初始化区段相关联的所述URL；以及

在所述经译码视频图片序列已经解码之后基于所述SEI消息而再现所述经译码视频图片序列。

16. 根据权利要求15所述的计算机程序产品，其进一步包括用以进行以下操作的指令：

提交对所述SEI消息的第一请求，其中所述用以接收所述SEI消息的指令包括用以响应于所述第一请求而接收所述SEI消息的指令；以及

仅在接收到所述SEI消息之后提交对所述经译码视频图片序列的第二请求，其中用以接收所述经译码视频图片序列的指令包括用以响应于所述第二请求而接收所述经译码视频图片序列的指令。

17. 根据权利要求16所述的计算机程序产品，其中所述SEI消息包括第一组SEI消息，且其中所述经译码视频图片序列包括第一经译码视频图片序列，所述计算机程序产品进一步包括用以进行以下操作的指令：

在接收到所述第一经译码视频图片序列之后提交对第二组SEI消息的第三请求；

响应于所述第三请求而接收所述第二组SEI消息；

提交对由所述第二组SEI消息描述的第二经译码视频图片序列的第四请求；

在接收到所述第二组SEI消息之后接收所述第二经译码视频图片序列；以及

在所述第二经译码视频图片序列已经解码之后基于所述第二组SEI消息而再现所述第

二经译码视频图片序列。

18. 根据权利要求15所述的计算机程序产品,其进一步包括用以从参数集轨道、不包含经译码视频图片的轨道片段和不包含经译码视频数据的第三代合作伙伴计划3GPP格式文件中的至少一者提取所述SEI消息的指令。

提供用于串流传输视频数据的序列数据集

[0001] 分案申请的相关信息

[0002] 本申请是国际申请号为PCT/US2011/044742,申请日为2011年7月20日,优先权日为2010年7月20日,发明名称为“提供用于串流传输视频数据的序列数据集”的PCT申请进入中国国家阶段后申请号为201180040329.5的中国发明专利申请的分案申请。

技术领域

[0003] 本发明涉及经编码的视频数据的输送。

背景技术

[0004] 可将数字视频能力并入到广泛范围的装置中,所述装置包括数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型计算机或桌上型计算机、摄像机、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、蜂窝式电话或卫星无线电电话、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,在由MPEG-2、MPEG-4、ITU-T H.263或ITU-T H.264/MPEG-4第10部分(高级视频译码(AVC))定义的标准及这些标准的扩展中所描述的视频压缩技术)以更有效率地发射及接收数字视频信息。

[0005] 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成宏块。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经帧内译码(I)的帧或切片中的宏块。在经帧间译码(P或B)的帧或切片中的宏块可使用相对于在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。

[0006] 在已编码视频数据之后,可将视频数据包化以用于传输或存储。可遵照各种标准中的任一者(例如,国际标准组织(ISO)基础媒体文件格式及其扩展(例如AVC))将视频数据组装成一视频文件。

[0007] 一直在努力开发基于H.264/AVC的新视频译码标准。一个此种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。AVC标准的一版本描述于JVT-AD007的“ITU-T推荐H.264|ISO/IEC 14496-10高级视频译码的编者草案修正——准备用于ITU-T SG 16 AAP批准(整合形式)(Editors'draft revision to ITU-T Rec.H.264|ISO/IEC 14496-10 Advanced Video Coding-in preparation for ITU-T SG 16 AAP Consent(in integrated form))”中(2009年2月瑞士日内瓦的第30届JVT会议),可在http://wftp3.itu.int/av-arch/jvt-site/2009_01_Geneva/JVT-AD007.zip处获得。此文献将SVC和MVC整合于AVC规范中。

发明内容

[0008] 大体上,本发明描述用于提供描述视频文件的经译码视频图片(例如)以支持根据HTTP串流传输的视频文件的串流传输的序列数据集的技术。序列数据集可包含例如序列参数集、图片参数集和一些补充增强信息(SEI)消息,例如序列层级SEI消息。这些技术可包含

将序列数据集与由序列数据集描述的经译码视频图片解耦。也就是说,并非与由序列层级SEI消息描述的经译码视频图片一起包含序列层级SEI消息,本发明的技术可包含将序列层级SEI消息与经译码视频图片解耦。因此,序列层级SEI消息可与经译码视频图片分离地从源装置传送到目的地装置。在一些实例中,包含序列层级SEI消息的序列数据集可在序列数据集所属的经译码视频图片的传送之前从源装置传送到目的地装置。

[0009] 在一个实例中,一种输出经封装视频数据的方法包含:产生与经译码视频图片分离的一个或一个以上补充增强信息(SEI)消息,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;以及与所述经译码视频图片序列分离地输出所述SEI消息。

[0010] 在另一实例中,一种用于输出经封装视频数据的设备包含:控制单元,其经配置以产生与经译码视频图片分离的一个或一个以上补充增强信息(SEI)消息,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;以及输出接口,其经配置以与所述经译码视频图片序列分离地输出所述SEI消息。

[0011] 在另一实例中,一种用于输出经封装视频数据的设备包含:用于产生与经译码视频图片分离的一个或一个以上补充增强信息(SEI)消息的装置,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;以及用于与所述经译码视频图片序列分离地输出所述SEI消息的装置。

[0012] 在另一实例中,一种包含计算机可读存储媒体的计算机程序产品包括指令,所述指令在执行时致使用于输出经封装视频数据的源装置的处理单元进行以下操作:产生与经译码视频图片分离的一个或一个以上补充增强信息(SEI)消息,其中所述SEI消息描述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;以及与所述经译码视频图片序列分离地输出所述SEI消息。

[0013] 在另一实例中,一种接收经编码视频数据的方法包含:接收描述视频文件的经译码视频图片序列的补充增强信息(SEI)消息,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;与所述SEI消息分离地接收所述视频文件的所述经译码视频图片序列;以及在所述经译码视频图片序列已经解码之后基于所述SEI消息而再现所述经译码视频图片序列。

[0014] 在另一实例中,一种用于接收经编码视频数据的设备包含:控制单元,其经配置以接收描述视频文件的经译码视频图片序列的补充增强信息(SEI)消息,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素,且与所述SEI消息分离地接收所述视频文件的所述经译码视频图片序列;以及视频输出,其经配置以在所述经译码视频图片序列已经解码之后再再现所述经译码视频图片序列。

[0015] 在另一实例中,一种用于接收经编码视频数据的设备包含:用于接收描述视频文件的经译码视频图片序列的补充增强信息(SEI)消息的装置,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;用于与所述SEI消息分离地接收所述视频文件的所述经译码视频图片序列的装置;以及用于在所述经译码视频图片序列已经解码之后基于所述SEI消息而再现所述经译码视频图片序

列的装置。

[0016] 在另一实例中,一种包含计算机可读存储媒体的计算机程序产品包括指令,所述指令在执行时致使用于接收经编码视频数据的目的地的装置的处理器的处理器进行以下操作:接收描述视频文件的经译码视频图片序列的补充增强信息(SEI)消息,其中所述SEI消息描述所述经译码视频图片序列中的相应者且包含所述经译码视频图片中的一者以上所共同的元素;与所述SEI消息分离地接收所述视频文件的所述经译码视频图片序列;以及在所述经译码视频图片序列已经解码之后基于所述SEI消息而再现所述经译码视频图片序列。

[0017] 一个或一个以上实例的细节陈述于随附图式及以下描述中。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目的地的装置的实例系统的框图。

[0019] 图2为说明实例封装单元的组件的框图。

[0020] 图3为说明可在各种位置中包含序列层级SEI消息的实例视频文件的元素的框图。

[0021] 图4是说明另一实例视频文件的元素的框图,所述另一实例视频文件包含含有序列层级SEI消息的参数集轨道。

[0022] 图5是说明包含第三代合作伙伴计划(3GPP)文件的简报的实例集合的框图,所述3GPP文件包含相应简报的相应序列数据集。

[0023] 图6是说明用于在视频文件内封装序列数据集且将视频文件从源装置提供到目的地的装置的实例方法的流程图。

[0024] 图7是说明包含多个序列数据集的视频文件的另一实例的框图。

具体实施方式

[0025] 大体上,本发明描述用于支持使用例如超文本传送协议(HTTP)串流传输等串流传输网络协议递送视频数据的技术。特定来说,本发明的技术是针对当例如以视频文件的形式发射视频数据时使某些补充增强信息(SEI)消息(例如序列层级SEI消息)与特定视频样本解耦。另外,所述技术可包含在目的地的装置处解译序列层级SEI消息以支持经编码视频图片的解码和再现。

[0026] 本发明还描述将SEI消息与编解码器层级中的视频样本解耦,其中SEI消息不一定需要与特定经译码图片相关联,例如当SEI消息与存取单元相关联时。在SEI消息与经译码图片去耦的情况下,指定SEI消息适用于哪一持续时间的时序信息可作为SEI消息的部分而用信号表示。

[0027] SEI消息可含有用以帮助解码器和/或再现器解码和/或显示某些经编码视频图片的信息。SEI消息还可用以提供错误恢复。一些SEI消息是序列层级(例如,与经译码图片群组(GOP)或经译码视频图片的其它序列相关),而其它SEI消息可与特定经译码图片相关。SEI消息通常与特定经译码图片一起发射。也就是说,为了检索SEI消息,解码器通常需要检索包含SEI消息的经译码图片。在解码器不需要经译码图片的数据的情况下,这大体上导致浪费带宽。而且,在一些情况下,用于不同经译码图片或两个GOP的两个SEI消息可能含有基

本上相同的信息,使得重复地发射SEI消息同样可造成带宽的浪费。

[0028] 本发明的技术包含单独于(例如,独立于)经译码视频数据来发射SEI消息,例如在SEI消息描述经译码视频图片序列的序列数据的情况下。序列数据可对应于经译码视频图片序列所共同的数据。在涉及H.264/AVC(高级视频译码)的一些实例中,当SEI消息包含与平移扫描矩形相关的信息和立体视频信息时,SEI消息可连同参数集基本串流一起发射。参数集基本串流可包含序列参数集(SPS)和/或图片参数集(PPS)消息。也就是说,根据这些技术,参数集基本串流除了一个或一个以上SEI消息外还可包含PPS和/或SPS消息。H.264/AVC中的此SEI消息的一个实例是帧包化布置SEI,其指定关于如何将一图像对的两个视图交错在一起成为一个帧的信息。作为另一实例,相对于可缩放视频译码(SVC),当SEI消息提供可缩放性信息时SEI消息可在参数集基本串流中发送。作为又一实例,在多视图视频译码(MVC)中,当SEI消息包含视图可缩放性信息、多视图获取信息和/或多视图场景信息时SEI消息可包含在参数集基本串流中。

[0029] 常规序列层级SEI消息已与某些存取单元相关联,例如包括瞬时解码器刷新(IDR)图片的存取单元。然而,本发明的技术是针对将序列层级SEI消息与存取单元分离。因此,本发明的技术可使得客户端装置能够与由序列层级SEI消息描述的经译码视频图片分离地检索序列层级SEI消息,而不是检索包含序列层级SEI消息的存取单元。因此,如果客户端装置经配置以解码和显示经译码视频图片序列的仅一部分,那么客户端装置可检索所述序列的仅所述部分,而无需检索其它图片,例如SEI消息原本将与之一起包含的IDR图片。根据本发明的技术,客户端装置可与经译码视频图片的所要部分分离地检索SEI消息。

[0030] 在一些实例中,并不是将序列层级SEI消息(和/或其它序列数据)包含在参数集轨道中,而是序列层级SEI消息可连续地包含在文件中,使得客户端可使用HTTP-Get操作来检索可用以在任一点解码和显示视频文件的信息。也就是说,用户可寻觅视频文件中的任一点,且解码器可检索SEI消息而无需检索额外的不必要的经译码视频数据,例如在视频文件内对应于寻觅请求的时间位置之前且将不会用来对在所述寻觅请求的时间位置处及之外的视频数据进行解码的经译码视频数据。

[0031] 视频或音频表示可包含一个或一个以上区段。每一表示可含有一初始化区段,或所述表示中的每一媒体区段可为自初始化的。初始化区段可含有用于存取对应表示(例如,由初始化区段描述的媒体区段)的数据的初始化信息。初始化区段大体上不含任何媒体数据,而是可仅包含初始化数据。每一表示可包含一个或一个以上媒体分量,其中每一媒体分量可为例如音频、视频或经计时文本等个别媒体类型的经编码版本。媒体分量在一个表示内的连续媒体区段的边界上可为时间连续的。

[0032] 作为又一实例,本发明的技术可包含发送序列数据(例如,PPS、SPS和序列层级SEI消息)以及(特定来说)不含有经译码视频数据(即,视频样本)的文件区段。源装置和目的地装置可以包含序列数据的视频样本的位置来配置,和/或以用于确定这些视频样本的位置的方法来配置。在一些实例中,序列数据可在经编码视频数据发送之前从服务器发送到客户端。也就是说,源装置可在发送由序列数据参考的经编码视频数据之前将例如PPS、SPS和序列层级SEI消息等序列数据提供到客户端装置。在基于文件区段的HTTP串流传输中,SEI消息(例如,序列层级SEI消息)连同参数集(例如序列参数集和图片参数集)可放置于初始化区段中,所述初始化区段不含有任何经译码音频或视频样本。

[0033] 本发明的技术可应用于遵照以下各项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,和/或多视图视频译码(MVC)文件格式。ISO基础媒体文件格式经设计以含有供以灵活、可扩展的格式呈现的定时媒体信息,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/IEC 14496-12:2004)在MPEG-4第12部分中指定,MPEG-4第12部分界定基于时间的媒体文件的一般结构。ISO基础媒体文件格式用作所述家族中的例如以下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/IEC 14496-15)、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定时期序列(例如,视听呈现)的时序、结构及媒体信息。所述文件结构是面向对象的。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隐含。

[0034] 遵照ISO基础媒体文件格式(及其扩展)的文件可形成一系列对象,称作“盒(box)”。呈ISO基础媒体文件格式的数据可含于盒中,且在盒外无其它数据需要含于文件内。这包含特定文件格式所需要的任何初始签名。“盒”可为通过唯一类型识别符及长度定义的面向对象的建置块。通常,一简报(presentation)含于一个文件中,且媒体简报为自含式的。电影容器(电影盒)含有媒体的元数据,且视频及音频帧含于媒体数据容器中且可在其它文件中。

[0035] 一简报(运动序列)可含于若干文件中。时序及成帧(位置及大小)信息通常在ISO基础媒体文件中,且辅助文件可基本上使用任何格式。此简报可在含有所述简报的系统的“本地”,或可经由网络或其它串流递送机制来提供。

[0036] 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(潜在地包含视频和音频数据两者)的逻辑结构,其又含有一组时间并行的轨道。文件的时间结构可为轨道含有在时间上的样本序列,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。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离。此结构信息可集中于一电影盒中,可能在时间上由电影片段盒扩展。所述电影盒可记载样本的逻辑及时序关系,且也可含有到其所处位置的指针。那些指针可指向同一文件中或另一文件中(例如,由URL参考)。

[0037] 每一媒体串流可含于专门用于所述媒体类型(音频、视频等)的轨道中,且可进一步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(解码串流所需要的解码器的类型)的“名称”及所述所需解码器的任何参数化。所述名称也可采取四字符码(例如,“moov”或“trak”)的形式。存在不仅用于MPEG-4媒体而且也用于使用此文件格式家族的其它组织所使用的媒体类型的经定义的样本条目格式。

[0038] 对元数据的支持通常采取两种形式。首先,可将定时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,可存在对附加到电影或个别轨道的非定时元数据的一般支持。结构支持为一般性的,且以类似于媒体数据(即,经译码视频图片)的存储的方式允许元数据资源存储于文件中的别处或存储于另一文件中。另外,这些资源可加以命名且可受保护。

[0039] 在ISO基础媒体文件格式中,样本分组为将轨道中的样本中的每一者指派为一个样本群组中的成员。样本群组中的样本不需要为连续的。举例来说,在呈现呈AVC文件格式

的H.264/AVC时,处于一个时间层级中的视频样本可被取样于一个样本群组中。可通过两个数据结构来表示样本群组:SampleToGroup盒(sbdp)及SampleGroupDescription盒。SampleToGroup盒表示将样本指派给样本群组。对于每一样本群组条目来说,可存在SampleGroupDescription盒的一个实例以描述对应群组的性质。

[0040] 任选的元数据轨道可用以用其所具有的“有趣特性”来标记每一轨道,对于所述轨道来说,其值可不同于群组的其它成员(例如,其位速率、屏幕大小或语言)。轨道内的一些样本可具有特殊特性或可个别地识别。特性的一个实例为同步点(常为视频I帧)。可在每一轨道中通过特殊表来识别这些点。更一般来说,也可使用元数据来记载轨道样本之间的相依性质。元数据可经结构化为文件格式样本的序列,正如视频轨道一样。此轨道可称作元数据轨道。每一元数据样本可经结构化为元数据语句。存在各种种类的语句,其对应于关于对应文件格式样本或其组成样本而可能被询问的各种问题。

[0041] 在经由串流传输协议递送媒体时,可能需要从其表示于文件中的方式变换媒体。此情形的一个实例发生于经由实时传输协议(RTP)传输媒体时。举例来说,在文件中,将视频的每一帧连续地存储为文件格式样本。在RTP中,必须遵守专门针对所使用的编解码器的包化规则以将这些帧置于RTP包中。串流传输服务器可经配置以在运行时间计算此包化。然而,存在对串流传输服务器的辅助的支持。可将被称为提示轨道的特殊轨道置于文件中。

[0042] 提示轨道含有用于串流传输服务器的关于如何针对特定协议从媒体轨道形成包串流的一般指令。因为这些指令的形式独立于媒体,所以在引入新编解码器时可能不需要修订服务器。另外,编码及编辑软件可不知晓串流传输服务器。一旦对文件完成编辑,则被称为提示程序(hinter)的一段软件可用以在将文件置于串流传输服务器上之前将提示轨道添加到文件。作为实例,在MPEG-4文件格式规范中存在针对RTP串流定义的提示轨道格式。

[0043] 术语“渐进下载”用以描述通常使用HTTP协议的数字媒体文件从服务器到客户端的传送。当从计算机起始时,所述计算机可在下载完成之前开始媒体的重放。串流传输媒体与渐进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。能够进行渐进下载重放的媒体播放器依赖于位于文件标头中的完整的元数据且在从网络服务器下载数字媒体文件时依赖于数字媒体文件的本地缓冲器。在指定量的经缓冲数据变得可用于本地重放装置时,所述装置可开始播放媒体。此指定量的经缓冲数据可通过编码器设置中的内容的产生者而嵌入于文件中,且可通过客户端计算机的媒体播放器所外加的额外缓冲器设置来加强。

[0044] 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/或第三代合作伙伴计划(3GPP)文件格式。所述技术可进一步应用于这些格式的这些及其它扩展,且可进一步应用以扩展其它文件格式从而与SEI消息描述的视频数据分离地提供SEI消息。因此,虽然为了实例目的而大体相对于AVC来描述,但应了解本发明的技术可大体上应用于任何其它文件格式。

[0045] 动画专家组(MPEG)已开发AVC文件格式作为ISO基本媒体文件格式的扩展。在AVC文件格式中,参数集(包含序列参数集和图片参数集)可与处于视频轨道中的视频基本串流相关联。还可能在称为参数集轨道的另一轨道中具有参数集,所述轨道包含含有作为SPS或

PPS网络抽象层(NAL)单元的样本的参数集基本串流。然而,SEI消息NAL单元的位置不限于在AVC文件格式中,只要任一SEI消息是在解码时间处于使用此SEI消息的任一图片之前的样本中呈现即可。

[0046] AVC文件格式具有SVC和MVC扩展以分别支持SVC和MVC的存储。在SVC和MVC中,多个层/视图可含于一个视频轨道或多个视频轨道中。在SVC和MVC规范中存在经定义以当在各种组织中呈现层/视图时构建操作点的工具。

[0047] SVC的一些功能性是从H.264/AVC继承。与先前的可缩放标准相比,SVC标准可提供各种优点,例如阶层式时间可缩放性、层间预测、单循环解码和灵活的输送接口。H.264/AVC提供灵活的阶层式B图片译码结构,其实现高级时间可缩放性。通过从H.264/AVC继承的此特征,SVC支持具有不同分辨率的层的时间可缩放性。在SVC中,图片群组(GOP)包含所谓的关键图片及在输出/显示次序上位于此关键图片与先前关键图片之间的所有图片。关键图片可以规则或不规则间隔来译码,且可使用先前关键图片作为运动补偿预测的参考而经帧内译码或帧间译码。可从具有较低时间层级的图片以阶层式方式预测非关键图片,且关键图片可具有最低时间层级。图片的时间层级可由NAL单元标头SVC扩展中的语法元素temporal_id指示。

[0048] SVC基于纹理、残余和运动而针对空间和信噪比(SNR)可缩放性引入层间预测。SVC中的空间可缩放性可经一般化到两个层之间的任一分辨率比率。SNR可缩放性可通过粗粒度可缩放性(CGS)、中等粒度可缩放性(MGS)或细粒度可缩放性(FGS)来实现。在SVC中,两个空间或CGS层可属于不同的相依性层(如NAL单元标头中的dependency_id可指示),而两个MGS层可在同一相依性层中。一个相依性层可包含具有从0到较高值的quality_id的质量层,所述值对应于质量增强层。SVC提供可用以减少层间冗余的层间预测方法。所述方法经概括为层间纹理预测、层间残余预测和层间运动预测。

[0049] SVC还提供灵活的系统 and 输送接口设计,其可实现编解码器到可缩放的多媒体应用系统的无缝集成。除了压缩和可缩放性提供之外,系统和输送接口集中于编解码器功能性,例如大体上对于视频编解码器来说是互操作性和一致性、可扩展性、随机存取、时序、缓冲管理以及错误恢复,且特定地对于可缩放译码来说是向后兼容性、可缩放性信息提供和可缩放性调适。

[0050] SVC进一步提供单循环解码方案。在单循环解码方案中,仅目标层需要运动补偿和完全解码。因此,与其中通常针对每个空间或SNR可缩放层执行运动补偿和完全解码的其它多循环解码方案相比,SVC的单循环解码方案可大大降低解码复杂性以及DPB大小。

[0051] SVC和MVC具有类似的高层级语法设计,例如在NAL单元标头、序列参数集和SEI消息方面。在SVC和MVC两者中,对于带宽或装置能力的调适,在某一程度上支持整个位流的提取。

[0052] 在一些实例中,本发明的技术可应用于H.264/AVC编解码器或基于AVC的编解码器,例如SVC、MVC或H.264/AVC的其它扩展。此类编解码器可经配置以当SEI消息在一个经译码图片(例如,存取单元)内相关联时辨识SEI消息,其中SEI消息可以ISO基础媒体文件格式或MPEG-2系统位流封装在所述存取单元内。所述技术还安排未来的译码标准(例如,H.265/HEVC)以允许与由SEI消息描述的任何存取单元分离地传递SEI消息。

[0053] 关于作为另一实例的3GPP,支持3GP文件的HTTP/TCP/IP输送以进行下载及渐进下

载。此外,使用HTTP进行视频串流传输可提供一些优点,且基于HTTP的视频串流传输服务正变得流行。HTTP串流传输可提供某些优点,包括可使用现有因特网组件及协议,使得不需要作新的努力来开发用于经由网络输送视频数据的新技术。其它传输协议(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。且,HTTP串流传输可为客户端驱动的,此可避免控制问题。

[0054] 举例来说,为了利用特征来获得最佳性能,服务器可追踪尚未得到确认的包的大小及内容。服务器也可分析文件结构并重建客户端缓冲器的状态以作出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服从所协商的简档。HTTP在已实施HTTP 1.1的网络服务器处不必需要新的硬件或软件实施。HTTP串流传输还提供TCP易用性及防火墙穿越。

[0055] 在HTTP串流传输中,频繁使用的操作包含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操作接收一字节范围作为输入参数且检索对应于所接收到的字节范围的文件的连续数目个字节。因此,可提供电影片段以进行HTTP串流传输,这是因为部分GET操作可取得一个或一个以上个别电影片段。请注意,在一电影片段中,可存在不同轨道的若干轨道片段。在HTTP串流传输中,媒体简报可为客户端可存取的数据的结构化集合。所述客户端可请求并下载媒体数据信息以向用户呈现串流传输服务。

[0056] 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。即,A/V源装置20及A/V目的地装置40可能够编码及解码(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可包括语音编码器(也称作声码器)。

[0057] 在图1的实例中,A/V源装置20包括音频源22及视频源24。举例来说,音频源22可包括麦克风,所述麦克风产生表示将由音频编码器26编码的所俘获音频数据的电信号。或者,音频源22可包括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可包括产生将由视频编码器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。

[0058] 原始音频及视频数据可包括模拟或数字数据。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。在讲话参与者正在讲话时,音频源22可获得来自讲话参与者的音频数据,且视频源24可同时获得所述讲话参与者的视频数据。在其它实例中,音频源22可包括包含所存储的音频数据的计算机可读存储媒体,且视频源24可包括包含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、串流传输的实时音频及视频数据或应用于经归档、预先记录的音频及视频数据。

[0059] 对应于视频帧的音频帧一般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由音频源22俘获的音频数据的音频帧。举例来说,当讲话参与者一般通过讲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)俘获讲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。因此,对应于视频帧的音频帧一般对应于其中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括音频数据及同时俘获的视频数据的情形。

[0060] 在一些实例中,音频编码器26可将表示记录每一经编码的音频帧的音频数据的时间的时戳编码于所述经编码的音频帧中,且类似地,视频编码器28可将表示记录每一经编码的视频帧的视频数据的时间的时戳编码于所述经编码的视频帧中。在这些实例中,对应于视频帧的音频帧可包括包含时戳的音频帧及包含同一时戳的视频帧。A/V源装置20可包含内部时钟,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一时戳相关联。

[0061] 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可将一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而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也可使用序列识别符来指示经编码视频数据的相对时间排序。类似地,在一些实例中,序列识别符可经映射或以其它方式与一时戳相关。

[0062] 本发明的技术大体上是针对经编码多媒体(例如,音频及视频)数据的输送,以及经输送的多媒体数据的接收及后续的解译及解码。如图1的实例中所展示,视频源24可将一场景的多个视图提供到视频编码器28。

[0063] A/V源装置20可将“服务”提供到A/V目的地装置40。服务大体上对应于MVC数据的可用视图的子集。举例来说,MVC数据可能可用于排序成零到七的八个视图。一个服务可对应于具有两个视图的立体视频,而另一服务可对应于四个视图,且又一服务可对应于所有八个视图。大体来说,服务对应于可用视图的任何组合(即,任何子集)。服务也可对应于可用视图的组合以及音频数据。操作点可对应于服务,使得A/V源装置20可对于由A/V源装置20提供的每一服务进一步提供操作点描述符。

[0064] 数据(无论是音频还是视频)的每一个别串流被称作基本串流。基本串流为一节目的单一、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串流。基本串流可在封装于视频文件内之前转换成包化基本串流(PES)。在同一节目内,使用串流ID来区分属于一个基本串流的PES包与属于另一基本串流的PES包。基本串流的数据的基本单元为包化基本串流(PES)包。因此,MVC视频数据的每一视图对应于相应基本串流。类似地,音频数据对应于一个或一个以上相应基本串流。根据本发明的技术,序列数据集(可包含例如序列参数集、图片参数集和序列层级SEI消息)可包含在参数集轨道或参数集基本串流中。本发明的技术可应用于符合MPEG-2系统的视频文件或视频串流,包含经包化基本串流(PES)、节目串流(PS)和输送串流(TS),及其AVC、SVC和MVC扩展。

[0065] 可将MVC经译码视频序列分离成若干子位流,所述子位流中的每一者为基本串流。可使用MVC view_id子集来识别每一子位流。基于每一MVC view_id子集的概念,定义MVC视频子位流。MVC视频子位流含有列于MVC view_id子集中的视图的NAL单元。节目串流通常仅含有来自基本串流的那些视图的NAL单元。还设计任何两个基本串流不可含有相同视图。

[0066] 在图1的实例中,封装单元30接收包括来自视频编码器28的视频数据的基本串流及包括来自音频编码器26的音频数据的基本串流。在一些实例中,视频编码器28及音频编

码器26可各自包含用于从经编码数据形成PES包的包化器。在其它实例中,视频编码器28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。在再其它实例中,封装单元30可包含用于从经编码音频及视频数据形成PES包的包化器。

[0067] 如本发明中所使用的“节目”可包括音频数据与视频数据(例如,由A/V源装置20的服务递送的音频基本串流及可用视图的子集)的组合。每一PES包包含识别PES包所属于的基本串流的stream_id。封装单元30负责将基本串流组装成视频文件。

[0068] 封装单元30从音频编码器26及视频编码器28接收一节目的基本串流的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,将经译码视频区段组织成NAL单元,此提供“网络友好”视频表示寻址应用,例如视频电话、存储、广播或串流传输。NAL单元可分类为视频译码层(VCL)NAL单元及非VCL NAL单元。VCL单元可含有核心压缩引擎,且可包含块、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(time instance)中的经译码图片(通常呈现为主要经译码图片)可含于一存取单元中,所述存取单元可包含一个或一个以上NAL单元。

[0069] 非VCL NAL单元可包含参数集NAL单元及SEI NAL单元,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)及很少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,无需对于每一序列或图片来重复很少改变的信息,因此可改善译码效率。此外,参数集的使用可实现重要标头信息的频带外传输,从而避免为了错误恢复而进行冗余传输的需要。在频带外传输实例中,参数集NAL单元可在与其它NAL单元(例如,SEI NAL单元)不同的信道上传输。

[0070] 补充增强信息(SEI)可含有对于从VCL NAL单元解码经译码图片样本来说并非必要但是可有助于与解码、显示、错误恢复和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消息是一些标准规范的标准化部分,且因此对于符合标准的解码器实施方案并非总是强制的。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实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。这些实例SEI消息可传递关于例如操作点的提取和操作点的特性的信息。

[0071] 根据本发明的技术,封装单元30可布置视频文件的数据以使得序列层级SEI消息与由序列层级SEI消息描述的经译码视频图片解耦。以此方式,封装单元30可将序列层级SEI消息与含有经译码视频数据的特定存取单元解耦。举例来说,封装单元30可在参数集基本串流中包含序列层级SEI消息,以及SPS和PPS数据,例如用于根据AVC或其扩展(例如MVC和SVC)编码的视频数据。另外,对于其中视频数据是根据MVC和/或SVC编码的实例,封装单元30可在参数集基本串流中包含子集SPS。

[0072] 在一些实例中,封装单元30可经配置以将仅某些SEI消息放置于参数集基本串流中。举例来说,对于H.264/AVC,封装单元30可在参数集基本串流中包含描述平移扫描矩形和立体视频信息的SEI消息。作为另一实例,对于SVC,封装单元30可在参数集基本串流中包含描述可缩放性信息的SEI消息。作为又一实例,对于MVC,封装单元30可包含描述视图可缩放性信息、多视图获取信息和多视图场景信息的SEI消息。

[0073] 大体上,SPS和PPS(和子集SPS,例如对于MVC和SVC)以及序列层级SEI消息可统称为“序列数据集”。如上文论述,序列数据集可包含在例如参数集基本串流等单独轨道中。在一些实例中,并非将序列数据集放置于单独参数轨道中,而是封装单元30可在文件中连续

地呈现一序列的序列数据集,使得目的地装置40可使用HTTP部分GET请求来检索整个序列数据集。举例来说,序列数据集可含于专门的视频样本中,所述视频样本不含经译码视频数据,例如经译码视频图片的部分或全体。

[0074] 通过将序列数据集与由序列数据集描述的视频数据分离,目的地装置40可与由序列数据集描述的视频数据分离地检索序列数据集。以此方式,源装置20可在将由序列数据集描述的视频数据发射到目的地装置40之前将序列数据集发射到目的地装置40。因此,源装置20可在将含有经译码视频图片的电影片段发送到目的地装置40之前将序列数据集发送到目的地装置40。

[0075] 在一些实例中,例如关于3GPP视频文件的使用,序列参数集可包含在单独的3GPP文件中,所述3GPP文件可在目的地装置40存取媒体简报描述符之后发送到目的地装置40。也就是说,目的地装置40可检索视频数据序列的媒体简报描述符,随后检索包含所述视频数据序列的序列数据集的3GPP文件,且随后检索所述视频数据序列自身。在一些实例中,序列数据集可形成媒体简报描述符自身的部分。在一些实例中,序列数据集可包含在不含经译码视频数据的电影盒的部分中。在一些实例中,序列数据集可形成放置于电影盒之后的电影片段。

[0076] 封装单元30可形成包括识别NAL所属于的节目的标头以及有效负载(例如,音频数据、视频数据或描述NAL单元所对应的输送串流或节目串流的数据)的NAL单元。举例来说,在H.264/AVC中,NAL单元包含1字节的标头及变化大小的有效负载。在一个实例中,NAL单元标头包括priority_id元素、temporal_id元素、anchor_pic_flag元素、view_id元素、non_idr_flag元素,及inter_view_flag元素。在常规MVC中,除了包含4字节的MVC NAL单元标头及NAL单元有效负载的前缀NAL单元及MVC经译码切片NAL单元外,保留H.264所定义的NAL单元。

[0077] NAL标头的priority_id元素可用于简单的单路径位流调适过程。temporal_id元素可用于指定对应NAL单元的时间层级,其中不同时间层级对应于不同帧速率。

[0078] anchor_pic_flag元素可指示图片为锚定图片还是非锚定图片。锚定图片及输出次序(即,显示次序)在锚定图片后的所有图片可正确地加以解码而无需对解码次序(即,位流次序)在前的图片进行解码,且因此,可用作随机存取点。锚定图片及非锚定图片可具有不同相依性,两者均在序列参数集中用信号表示。其它旗标将被论述并用于此章的以下部分中。此锚定图片也可称为开放GOP(图片群组)存取点,而在non_idr_flag元素等于零时也支持封闭GOP存取点。non_idr_flag元素指示图片为瞬时解码器刷新(IDR)还是视图IDR(V-IDR)图片。通常,可正确地解码IDR图片及输出次序或位流次序在其后的所有图片而无需对解码次序或显示次序在前的图片进行解码。

[0079] view_id元素可包括可用以识别视图的语法信息,其可用于MVC解码器内的数据互动功能(例如,用于视图间预测)且可用于解码器外的数据互动功能(例如,用于再现)。inter_view_flag元素可指定是否由其它视图使用对应NAL单元以用于视图间预测。为了传达符合AVC的用于基础视图的4字节NAL单元标头信息,在MVC中定义前缀NAL单元。在MVC的上下文中,基础视图存取单元包含视图的当前时间实例的VCL NAL单元以及其仅含有NAL单元标头的前缀NAL单元。H.264/AVC解码器可忽略前缀NAL单元。

[0080] 在有效负载中包含视频数据的NAL单元可包括各种粒度层级的视频数据。举例来

说,NAL单元可包括一视频数据块、一宏块、多个宏块、一视频数据切片,或视频数据的整个帧。封装单元30可从视频编码器28接收呈基本串流的PES包的形式经编码视频数据。封装单元30可使每一基本串流与对应节目相关联。

[0081] 封装单元30也可从多个NAL单元组装存取单元。通常,一存取单元可包括用于表示一视频数据帧以及对应于所述帧的音频数据(当此音频数据可用时)的一个或一个以上NAL单元。一存取单元通常包含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个帧/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,一存取单元可包括一个时间实例中的一经译码图片,所述经译码图片可呈现为主要经译码图片。因此,一存取单元可包括一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。本发明还将一特定视图的经编码图片称作“视图分量”。即,一视图分量可包括在特定时间的一特定视图的经编码图片(或帧)。因此,存取单元可定义为包括一共同时间实例的所有视图分量。存取单元的解码次序不必必须与输出或显示次序相同。

[0082] 如同大多数视频译码标准,H.264/AVC定义无错误位流的语法、语意及解码过程,所述各者中的任一者均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符合标准。在视频译码标准的上下文中,“简档”对应于算法、特征或工具及施加于其的约束的一子集。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一子集。“层级”对应于解码器资源消耗(例如,解码器存储器及计算)的限制,所述限制与图片的分辨率、位速率及宏块(MB)处理速率有关。可以profile_idc(简档指示符)值用信号表示一简档,而可以level_idc(层级指示符)值用信号表示一层级。

[0083] 举例来说,H.264标准认识到,在给定简档的语法外加的边界内,仍有可能需要编码器及解码器的性能的大变化,这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,实施能够应对一特定简档内的语法的所有假设使用的解码器既不实际也不经济。因此,H.264标准将“层级”定义为外加于位流中的语法元素的值的约束的一指定集合。这些约束可为对值的简单限制。或者,这些约束可采取对值的算术组合(例如,图片宽度乘以图片高度乘以每秒解码的图片的数目)的约束的形式。H.264标准进一步使得个别实施方案可对于每一受支持简档支持一不同层级。

[0084] 遵照一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,作为一译码特征,B图片译码在H.264/AVC的基线简档中不受支持,但在H.264/AVC的其它简档中受到支持。遵照一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,在视频传输期间,一对简档及层级定义可对于整个传输会话来加以协商并达成一致。更特定来说,在H.264/AVC中,层级可定义(例如)对以下各者的限制:需要处理的宏块的数目、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块是否可具有小于 8×8 像素的子宏块分割区。以此方式,一解码器可确定所述解码器是否能够适当地解码所述位流。

[0085] 参数集通常含有序列层标头信息(在序列参数集(SPS)中)及很少改变的图片层标

头信息(在图片参数集(PPS)中)。就参数集而论,无需对于每一序列或图片来重复此很少改变的信息,因此可改善译码效率。此外,参数集的使用可实现标头信息的频带外传输,从而避免为了错误恢复而进行冗余传输的需要。在频带外传输中,参数集NAL单元是在与其它NAL单元不同的信道上传输。

[0086] SVC和MVC的实例提供用于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和层级信息可包含于包含SPS扩展信息的NAL单元中。

[0087] 媒体简报可包含媒体简报描述符(MPD),其可含有不同的替代表示(例如,具有不同质量的视频服务)的描述且所述描述可包含例如编解码器信息、简档值和层级值。目的地装置40可检索媒体简报的MPD以确定如何存取各种简报的电影片段。电影片段可位于视频文件的电影片段盒(moof盒)中。

[0088] 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。编码器使用从一些先前经编码图片(本文中也被称作帧)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图片仅使用时间次序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。这些参考图片的时间次序可位于当前图片之前或之后。

[0089] 根据H.264译码标准,作为一实例,B图片使用先前经译码参考图片的两个列表,列表0及列表1。这两个列表可各自含有时间次序上为过去及/或未来的经译码图片。可以以下若干方式中的一者预测B图片中的块:根据列表0参考图片的运动补偿预测、根据列表1参考图片的运动补偿预测,或根据列表0参考图片与列表1参考图片两者的组合的运动补偿预测。为了得到列表0参考图片与列表1参考图片两者的组合,分别从列表0与列表1参考图片获得两个运动补偿参考区域。所述两个运动补偿参考区域的组合将用以预测当前块。

[0090] ITU-T H.264标准支持各种块大小的帧内预测(例如,对于明度分量的16乘16、8乘8或4乘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,对于明度分量的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大小)。在本发明中,“×”与“乘”可互换使用以指代按照垂直及水平尺寸的块的像素尺寸,例如,16×16像素或16乘16像素。一般来说,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, $N \times N$ 块一般在垂直方向上具有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。可将块中的像素排列成行及列。

[0091] 小于16乘16的块大小可称作16乘16宏块的分割区。视频块可包括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,以下变换应用,例如离散余弦变换(DCT)、整数变换、小波变换,或与残余视频块数据在概念上类似的变换,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。在一些状况下,视频块可包括变换域中的经量化变换系数的块。

[0092] 较小视频块可提供较好分辨率,且可用于包含高细节水平的视频帧的位置。一般来说,可将宏块及各种分割区(有时称作子块)视为视频块。另外,可将一切片视为多个视频块,例如宏块及/或子块。每一切片可为一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”可指代一视频帧的任何可独立解码单元,例如整个帧、帧的一切片、也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。

[0093] 术语宏块指用于根据包括 16×16 像素的二维像素阵列编码图片及/或视频数据的数据结构。每一像素包括一色度分量及一亮度分量。因此,宏块可界定四个亮度块(各自包括 8×8 像素的二维阵列)、两个色度块(各自包括 16×16 像素的二维阵列)及一标头,所述标头包括语法信息,例如经译码块模式(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。

[0094] 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及解封装单元38在适当时各自可实施为各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中的每一者可包含于一个或一个以上编码器或解码器中,其任一者可集成为组合式视频编码器/解码器(CODEC)的部分。同样,音频编码器26及音频解码器46中的每一者可包含于一个或一个以上编码器或解码器中,其任一者可集成为组合式CODEC的部分。包含视频编码器28、视频解码器48、音频编码器26、音频解码器46、封装单元30及/或解封装单元38的设备可包括集成电路、微处理器及/或无线通信装置,例如蜂窝式电话。

[0095] 在封装单元30已基于接收到的数据将NAL单元及/或存取单元组装成一视频文件之后,封装单元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,封装单元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可包括发射器、收发器、用于将数据写入到计算机可读媒体的装置,例如光驱、磁性媒体驱动器(例如,软驱动器)、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,例如传输信号、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。

[0096] 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可包括光驱、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封装单元38。解封装单元38可将一视频文件的元素解封装成组成的PES串流、将所述PES串流解包化以检索经编码数据,及将所述经编码数据发送到音频解码器46或视频解码器48,这取决于所述经编码数据是音频串流的部分还是视频串流的部分,例如由所述串流的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(其可包含一串流的多个视图)发送到视频输出44。

[0097] 另外,解封装单元38可将经解码序列数据集(可包含PPS数据、SPS数据和/或序列层级SEI消息)发送到视频输出44。视频输出44可包括使用一场景的单个视图或一场景的多个视图的显示,例如同时或快速交替连续地呈现一场景的每一视图的立体或自由立体显示。视频输出44可使用序列数据集来显示经解码视频图片。序列层级SEI消息也可用以辅助解码过程,但适当解码视频序列的样本并不需要SEI消息。因此,视频解码器48也可在解码

视频文件的视频样本时使用序列层级SEI消息。

[0098] 如上所述,SEI消息可用于多种目的。举例来说,序列层级SEI消息可描述平移扫描信息,所述信息用于调整视频数据以适应显示时的各种屏幕大小。因此,视频输出44(或其控制单元)可使用序列层级SEI消息来调整视频数据以由视频输出44显示。

[0099] 作为另一实例,序列层级SEI消息可描述立体视频信息。也就是说,序列层级SEI消息可描述哪些图片是用于三维视频应用的“左”或“右”图像。因此,视频输出44可使用序列层级SEI消息来确定如何呈现三维视频数据。举例来说,视频输出44可使用光的不同极化或以不同的滤色器(例如,红和绿)来显示左图像和右图像,或可在左图像与右图像之间快速地交替,使得与视频输出44相关联的作用中眼镜可在允许光通过的透镜之间交替。

[0100] 作为又一实例,序列层级SEI消息可提供可缩放性信息,例如对于其中根据SVC或MVC来构建视频文件的实例。因此,视频解码器48可使用序列层级SEI消息来确定增强层信息是否已被接收,且因此确定是否使用接收到的增强层数据或内插值以对基本层视频数据进行上取样。

[0101] 作为另一实例,序列层级SEI消息可提供多视图获取信息或多视图场景信息,例如对于其中根据MVC来构建视频文件的实例。大体上,多视图获取信息SEI消息指定获取环境的参数,例如固有和非固有相机参数、视图数目、焦距、主点、偏斜因数、旋转参数和/或平移参数。多视图场景信息SEI消息可在俘获视图时指定多个视图之间的最大相异性。此信息可在解码和/或再现三维视频数据时,例如当从现存图像构建额外图像时由视频解码器48和/或视频输出44使用。多视图场景信息还可描述在图像获取期间一个或一个以上相机或条件的特性,例如因(例如)高达1024个视图的相机设定引起的最大相异性是什么。多视图场景信息还可提供与相机校准和纠正相关的细节。

[0102] 图2为说明实例封装单元30的组件的框图。在图2的实例中,封装单元30包含视频输入接口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包含网络抽象层(NAL)单元建构器62、SEI消息提取单元64和序列数据集创建单元66。

[0103] 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。当数据经编码时,视频输入接口80及音频输入接口82可接收经编码视频及音频数据,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以组装成一视频文件。

[0104] 视频文件创建单元60可对应于一控制单元,所述控制单元包含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封装单元30所有的功能。对于其中视频文件创建单元60以软件和/或固件体现的实例,封装单元30可包含包括用于视频文件创建单元60的指令的计算机可读媒体,和用以执行所述指令的处理单元。视频文件创建单元60的子单元(在此实例中,为NAL单元建构器62、SEI消息提取单元64和序列数据集创建单元66)中的每一者可实施为个别硬件单元及/或软件模块,及/或可在功能上集成或进一步分离成额外子单元。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、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可进一步包含存储用

于NAL单元建构器62、SEI消息提取单元64和序列数据集创建单元66中的任一者或全部的指令的非暂时性计算机可读媒体,以及用于执行所述指令的处理器。

[0105] 通常,视频文件创建单元60可创建包含接收到的音频及视频数据的视频文件。NAL单元建构器62可形成包含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装包含特定时间实例的所有NAL单元的存取单元。此外,视频文件创建单元60可经配置以使序列层级SEI消息与由序列层级SEI消息描述的经编码视频图片解耦,且与由序列层级SEI消息描述的经编码视频图片分离地将序列层级SEI消息存储在创建的视频文件中。

[0106] 视频编码器28(图1)可与视频数据样本一起包含SEI消息。在一些情况下,针对若干个别样本从视频编码器28接收的SEI消息对于样本中的每一者来说可为相同的。因此,并非针对样本中的每一者发射相同SEI消息,在确定个别样本的SEI消息对于所述样本中的每一者来说相同(或非常类似)后,SEI消息提取单元64可将SEI消息与由SEI消息描述的样本分离。可与一个以上经编码视频图片相关的此类SEI消息可称为序列层级SEI消息。

[0107] 另外或替代地,视频文件创建单元60可经配置以产生经译码视频图片序列的序列层级SEI消息。以此方式,视频文件创建单元60可产生序列层级SEI消息或将现存SEI消息与视频样本分离以形成描述视频样本中的一者或一者以上的序列层级SEI消息。

[0108] 序列数据集创建单元66可创建由视频文件创建单元60产生的视频文件的序列数据集。序列数据集创建单元66可以序列参数集和/或图片参数集数据来分组序列层级SEI消息。举例来说,序列数据集创建单元66可在参数集基本串流中包含序列层级SEI消息,所述参数集基本串流包含SPS和PPS,且在一些实例(例如,关于SVC和MVC)中包含子集SPS。作为另一实例,序列数据集创建单元66可包含经创建视频文件的一连续部分中的经译码视频图片序列的序列层级SEI消息。也就是说,序列数据集创建单元66可确保特定经译码视频图片序列的序列层级SEI消息的数据在连续的一串一个或一个以上字节中分组在一起。举例来说,序列数据集创建单元66可将序列数据集的序列层级SEI消息和其它数据组装成特定视频样本,所述特定视频样本不含经译码视频图片的任一部分。

[0109] 作为另一实例,例如关于3GPP,序列数据集创建单元66可将经译码视频图片序列的序列数据集组装成相对小的3GPP文件,所述3GPP文件位于包含由序列数据集描述的经译码视频图片序列的一个或一个以上3GPP文件之前。替代地,序列数据集创建单元66可针对包含一个或一个以上3GPP文件的一个或一个以上相关简报在媒体简报描述(MPD)文件中包含序列数据集。

[0110] 作为又一实例,序列数据集创建单元66可在视频文件的电影盒(moov盒)不包含经译码视频图片时在所述电影盒内包含序列数据集。作为另一实例,序列数据集创建单元66可在发生于视频文件的电影盒(moov盒)之后但在包含经译码视频图片的任何电影片段盒之前的电影片段盒(moof盒)内包含序列数据集。

[0111] 通过将SPS、PPS、序列层级SEI消息和/或子集SPS分组为序列数据集,序列数据集创建单元66可简化将序列数据提供到客户端装置(例如目的地装置40)的任务。以此方式,序列层级SEI消息可在由序列层级SEI消息描述的经译码视频图片发射到目的地装置40之前发射到目的地装置40。

[0112] 大体上,序列数据集的SPS、PPS和序列层级SEI消息预期针对由序列数据集描述的

视频样本序列保持恒定。然而,在一些情况下,SPS、PPS和/或序列层级SEI消息的某些部分可针对个别经译码视频图片或针对经译码视频图片的集合或子序列稍微有所不同。举例来说,在封装有若干经译码视频图片的视频文件中,可存在不一定共享SPS、PPS或序列层级SEI消息的相同信息的一个或一个以上序列。

[0113] 当此情况发生时,序列数据集创建单元66可创建多个参数集轨道,其中每一轨道可包含一组SPS、PPS和/或SEI消息,以及指示轨道的序列数据集何时适用的时序信息的指示。举例来说,所述时序信息可指示视频序列的开始时间,其指示轨道开始应用于视频序列的时间和/或包含时序信息的轨道不再适用的视频序列结束时间。替代地,每一轨道的时序信息可仅指示结束时间,使得在到达特定参数集轨道的结束时间后,客户端装置可跳到下一参数集轨道。此时序信息可对应于协调世界时间(UTC)或参数集轨道在用于呈文件封装格式的整个视频的时间标度中适用于的最后一视频样本(对于结束时间来说)的持续时间。

[0114] 多个参数集轨道中的一者可适用于视频文件的所有样本。序列数据集创建单元66可通过将参数集轨道的时序信息设定为比任一电影样本或电影片段晚的UTC时间,来指示适用于视频文件的所有样本的参数集轨道。替代地,序列数据集创建单元66可针对适用于视频文件的所有样本的参数集轨道的时序信息指示零持续时间。

[0115] 对于其中序列数据集创建单元66创建多个参数集轨道的实例,序列数据集创建单元66可对参数集轨道中的每一者指派相应的唯一识别符。当例如目的地装置40等客户端装置例如通过本地重放或在HTTP串流传输期间请求寻觅特定时间位置时,目的地装置40可确定若干参数集轨道中的哪一者适用于所寻觅的时间位置。客户端装置可请求具有指示相应参数集轨道适用于比所寻觅时间位置晚的视频样本的时序信息的参数集轨道的识别符。服务器(例如,源装置20)可维持用以标记此类相关参数集轨道的数据结构以确保相关参数集轨道递送到客户端,例如在HTTP串流传输的上下文中。在本地重放期间,客户端可维持类似的数据结构以确保相关参数集轨道由客户端装置剖析。

[0116] 而且,并非包含若干参数集轨道,序列数据集创建单元66可在视频文件中包含若干序列数据集,其中每一者可连续地存储在文件内。也就是说,序列数据集中的每一者可连续地存储以使得所述序列数据集中的特定一者的所有数据是连续的,但所述序列数据集不需要彼此连续地存储。举例来说,序列数据集创建单元66可在一序列数据集相关的一系列片段之前的视频片段中存储所述序列数据集,且随后在所述系列片段之后存储一新序列数据集。所述新序列数据集可与在存储所述新序列数据集的片段之后的其它视频片段相关。当在文件中存储多个序列数据集时,序列数据集创建单元66可提供所述序列数据集中的每一者的时序信息,以及所述序列数据集中的每一者的唯一识别符。

[0117] 在视频文件创建单元60产生包含与由序列层级SEI消息描述的经译码视频图片分离的序列层级SEI消息的视频文件之后,视频文件创建单元60可将视频文件传递到视频文件输出接口84。视频文件输出接口84可将视频文件输出(例如)到源装置20的输出接口32。在一些实例中,视频文件输出接口84可将视频文件输出到源装置20的存储媒体(未图示)。视频文件可在本地存储于源装置20内、可存储到便携式存储媒体(例如,DVD、蓝光光盘、快闪驱动器、软磁盘或其它便携式存储媒体)、可经由网络输出(例如,根据例如HTTP串流传输等串流传输协议)或以其它方式输出,使得可由客户端装置(例如,目的地装置40)来接收视频文件。以此方式,视频数据(例如,序列层级SEI消息和由序列层级SEI消息描述的经译码

视频图片)的输出可包含视频数据的存储、传送、发射和广播。

[0118] 图3为说明可在各种位置中包含序列层级SEI消息的实例视频文件100的元素的框图。如上文所描述,根据ISO基础媒体文件格式及其扩展,视频文件将数据存储于被称为“盒”的一系列对象中。在图3的实例中,视频文件100包含文件类型(FTYP)盒102、电影(MOOV)盒104、电影片段(MOOF)盒112及电影片段随机存取(MFRA)盒114。

[0119] 文件类型盒102大体上描述视频文件100的文件类型。文件类型盒102可包含识别描述视频文件100的最佳用途的规范的数据。文件类型盒102可置于MOOV盒104、电影片段盒112及MFRA盒114之前。

[0120] 在图3的实例中,MOOV盒104包含电影标头(MVHD)盒106、轨道(TRAK)盒108及一个或一个以上电影扩展(MVEX)盒110。大体来说,MVHD盒106可描述视频文件100的一般特性。举例来说,MVHD盒106可包含描述在何时起初创建视频文件100、在何时最后修改视频文件100、视频文件100的时间标度、视频文件100的重放持续时间的数据,或大体上描述视频文件100的其它数据。

[0121] TRAK盒108可包含视频文件100的轨道的数据。TRAK盒108可包含描述对应于TRAK盒108的轨道的特性的轨道标头(TKHD)盒。在一些实例中,TRAK盒108可包含经译码视频图片,而在其它实例中,轨道的经译码视频图片可包含于可由TRAK盒108的数据参考的电影片段112中。

[0122] 在一些实例中,视频文件100可包含一个以上轨道。因此,MOOV盒104可包含等于视频文件100中的轨道的数目的数个TRAK盒。TRAK盒108可描述视频文件100的对应轨道的特性。举例来说,TRAK盒108可描述对应轨道的时间及/或空间信息。在封装单元30(图1)在视频文件(例如,视频文件100)中包含参数集轨道时,类似于MOOV盒104的TRAK盒108的TRAK盒可描述参数集轨道的特性。封装单元30可用信号表示描述参数集轨道的TRAK盒内的参数集轨道中的序列层级SEI消息的存在。

[0123] MVEX盒110可描述对应电影片段112的特性,例如用信号表示视频文件100除包含于MOOV盒104内的视频数据(如果存在)外还包含电影片段112。在串流传输视频数据的上下文中,经译码视频图片可包含于电影片段112中而非包含于MOOV盒104中。因此,所有经译码视频图片可包含于电影片段112中而非包括于MOOV盒104中。

[0124] MOOV盒104可包含等于视频文件100中的电影片段112的数目的数个MVEX盒110。MVEX盒110中的每一者可描述电影片段112中的对应一者的特性。举例来说,每一MVEX盒可包含电影扩展标头盒(MEHD)盒,其描述电影片段112中的对应一者的持续时间。

[0125] 如上所述,封装单元30可在不包含实际经译码视频数据的视频样本中存储序列数据集。视频样本可大体上对应于存取单元,所述存取单元为在特定时间实例处经译码图片的表示。在AVC的上下文中,经译码图片包含一个或一个以上VCL NAL单元,其含有用以构建存取单元和其它相关联非VCL NAL单元的所有像素的信息,例如SEI消息。因此,封装单元30可在电影片段112中的一者中包含序列数据集,其可包含序列层级SEI消息。封装单元30可进一步将序列数据集和/或序列层级SEI消息的存在用信号表示为存在于电影片段112中的一者中位于MVEX盒110中对应于电影片段112中的所述一者的一者内。

[0126] 电影片段112可包含一个或一个以上经译码视频图片。在一些实例中,电影片段112可包含一个或一个以上图片群组(GOP),所述图片群组中的每一者可包含若干经译码视

频图片,例如帧或图片。另外,如上文所描述,在一些实例中,电影片段112可包含序列数据集。电影片段112中的每一者可包含电影片段标头盒(MFHD)。MVHD盒可描述对应电影片段的特性,例如电影片段的序号。电影片段112可按序号次序包含于视频文件100中。

[0127] MFRA盒114可描述视频文件100的电影片段112内的随机存取点。这可辅助执行在视频文件100内寻觅特定时间位置。对于一些实例,MFRA盒114通常为任选的,且不需要包含于视频文件中。同样,例如目的地装置40的客户端装置不一定需要参考MFRA盒114来正确地解码并显示视频文件100的视频数据。MFRA盒114可包含等于视频文件100的轨道的数目或(在一些实例中)等于视频文件100的媒体轨道(例如,非提示轨道)的数目的数个轨道片段随机存取(TFRA)盒。

[0128] 图4是说明另一实例视频文件140的元素的框图,所述另一实例视频文件140包含含有序列层级SEI消息156的参数集轨道146。特定来说,视频文件140包含MOOV盒142和MDAT盒150。视频文件140可进一步包含其它盒,例如文件类型盒和/或MFRA盒等等。

[0129] 在图4的实例中,MOOV盒142包含视频数据轨道144和参数集轨道146。参数集轨道146对应于图片参数集(PPS) 152、序列参数集(SPS) 154和序列层级补充增强信息(SL SEI)消息156。视频数据轨道144对应于视频片段160A到160N(视频片段160)。大体上,视频片段160可各自包含一个或一个以上经译码视频图片。

[0130] 在图4的实例中,PPS 152、SPS 154、序列层级SEI消息156以及视频片段160的数据存储在MDAT盒150中。PPS 152、SPS 154和序列层级SEI消息156可统称为序列参数集。在一些实例中,参数集轨道146可包含一个以上图片参数集。PPS 152、SPS 154和序列层级SEI消息156大体上描述视频片段160的特性。这些特性可帮助例如目的地装置40等客户端装置选择、解码和/或再现视频片段160。

[0131] 作为一实例,当视频文件140对应于AVC文件格式时,序列层级SEI消息156可包含平移扫描矩形SEI消息、立体视频信息SEI消息和/或帧封装布置SEI消息中的一者或一者以上。作为另一实例,当视频文件140对应于SVC文件格式时,序列层级SEI消息156可包含可缩放性信息SEI消息。作为另一实例,当视频文件140对应于MVC文件格式时,序列层级SEI消息156可包含视图可缩放性信息SEI消息、多视图获取信息SEI消息和/或多视图场景信息SEI消息中的一者或一者以上。

[0132] 视频文件可包含类似于视频数据轨道144的若干媒体轨道。视频数据轨道中的每一者可具有不同特性,所述特性可对应于各种解码和/或再现要求。目的地装置40可尤其基于由序列数据集用信号表示的媒体轨道的视频片段的特性来选择媒体轨道中的一者。举例来说,目的地装置40可分别基于视频解码器48和视频输出44的解码和/或再现能力来选择轨道中的一者。目的地装置40可基于包含PPS 152、SPS 154和序列层级SEI消息156的参数集轨道146来确定视频片段160的特性。

[0133] 目的地装置40可经配置以检索序列参数集,例如MOOV盒104、电影片段112中包含序列参数集的一者、包含序列参数集但不包含经译码视频数据的视频样本,或在由序列参数集描述的任何经译码样本之前的视频文件100的其它元素。另外或替代地,源装置20可经配置以在将由序列参数集描述的任何经译码视频图片提供到目的地装置40之前将序列参数集提供到目的地装置40。而且,源装置20可经配置以在目的地装置40尚未首先请求电影片段112中被请求的一者的序列数据集的数据时拒绝来自目的地装置40的对电影片段112

中的所述一者的请求。

[0134] 图5是说明包含3GPP文件186、194的简报182、190的实例集合198的框图,所述3GPP文件186、194包含相应简报的相应序列数据集。也就是说,集合198可包含额外的简报,但图5为了实例目的而描绘两个简报182、190。

[0135] 媒体简报描述(MPD)180包含描述集合198的简报(例如,简报182、190)的数据。MPD 180可描述例如简报的统一资源定位符(URL)和/或统一资源名称(URN)。MPD 180可进一步描述集合198中的简报中的每一者的特性。举例来说,MPD 180可包含简报中的每一者的序列数据集。MPD 180还可描述存储在简报中的每一者中的若干3GPP视频文件。

[0136] 在图5的实例中,简报182包含标头数据184、具有序列数据集的3GPP文件186和3GPP文件188A到188N(3GPP文件188)。类似地,在此实例中,简报190包含标头数据192、具有序列数据集的3GPP文件194和3GPP文件196A到196N(3GPP文件196)。虽然使用字母“N”来指定3GPP文件188和3GPP文件196的基数,但N对于简报182和190来说不一定相同。也就是说,在简报182和简报190以及集合198的简报中的任一者中可存在不同数目的3GPP文件。

[0137] 标头数据184可描述简报182的某些特性,例如简报182中的3GPP文件的数目、对应于3GPP文件188中的每一者的开始的存储器地址、3GPP文件188的长度(例如,以字节为单位)、3GPP文件188的持续时间(例如,以秒、分钟等单位)、客户端装置适当解码和再现3GPP文件188的预期解码和/或再现能力,或其它数据。简报182可包含例如由ISO基本媒体文件格式标准和/或其扩展(例如3GPP标准)描述的额外数据。3GPP文件188可包含经译码视频图片。具有序列数据集的3GPP文件186包含描述经译码视频图片序列的经译码视频图片(例如,包含在3GPP文件188中的一者或一者以上中)的序列数据集。

[0138] 在图5的实例中,集合198的简报可对应于同一视频数据序列的各种简报。也就是说,简报中的每一者在重放期间(例如当解码和再现时)可为类似的。然而,视频数据在集合198的不同简报内可不同地译码或呈现。举例来说,简报190的视频样本可具有较高或较低的分辨率,以较高或较低的位速率编码,预期特定解码器缓冲器大小,具有特定帧速率,预期特定宏块解码速率,或基于其它特性而彼此不同。例如封装单元30(图1)的封装单元可将序列数据集放置于简报内的3GPP文件中,例如简报182的3GPP文件186和简报190的3GPP文件194。另外或替代地,封装单元30可将集合198的一个或一个以上简报的序列数据存储在MPD 180中。

[0139] 图6是说明用于在视频文件内封装序列数据集且将视频文件从源装置提供到目的地装置的实例方法的流程图。虽然为了实例和阐释目的关于源装置20和目的地装置40(图1)的组件来描述,但应了解,任何合适装置可实施图6的技术。

[0140] 最初,封装单元30可接收经编码视频图片序列(210)。例如视频编码器28等编码器可能已随经译码样本一起包含序列层级SEI消息。另外或替代地,封装单元30可与视频编码器28分离地创建序列层级SEI消息。在任一情况下,封装单元30可将序列层级SEI消息与由序列层级SEI消息描述的经译码视频图片解耦(212)。

[0141] 也就是说,封装单元30可创建包含序列层级SEI消息和由序列层级SEI消息描述的经译码视频图片的视频文件(214)。然而在如此做法时,封装单元30可与由序列层级SEI消息描述的经译码视频图片分离地存储序列层级SEI消息。以此方式,SEI消息可与经译码视频图片分离地发射和处理。举例来说,根据本发明的技术,封装单元30可将序列层级SEI消

息存储在经创建视频文件的参数集轨道中。作为另一实例,封装单元30可将序列层级SEI消息存储在经译码视频图片的轨道片段中、不包含经译码视频图片的单独3GPP文件中、类似于经译码视频图片但实际上不包含经译码视频数据的数据结构中、描述视频文件的媒体简报描述中,或以另外方式将序列层级SEI消息与由序列层级SEI消息描述的经译码视频图片分离。

[0142] 封装单元30可随后输出视频文件(216)。举例来说,封装单元30可致使源装置20将视频文件写入到存储媒体,例如光盘、软磁盘、快闪驱动器、硬驱动器、固态驱动器或其它存储媒体。此存储媒体可物理上输送到目的地装置40。替代地,源装置20可例如经由广播、网络发射或其它发射技术将视频文件发射到目的地装置40。在任一情况下,目的地装置40可最终接收视频文件(218)。

[0143] 在一些实例中,源装置20可例如响应于由目的地装置40发出到源装置20的一个或一个以上HTTP-Get或部分Get请求而将视频文件的相异部分提供到目的地装置40。目的地装置40可将第一HTTP-Get或部分Get请求发出到源装置20以检索序列数据集,例如包含序列层级SEI消息的参数集轨道的全部或一部分,且发出第二(或更多)HTTP-Get或部分Get请求以检索由序列数据集描述的经译码视频图片。

[0144] 在接收到视频文件之后,目的地装置40可基于序列层级SEI消息来解码视频文件(220)。也就是说,视频解码器48可使用序列层级SEI消息的数据来帮助解码过程。同样,视频显示器44可使用序列层级SEI消息来帮助再现经解码视频数据,以便显示经解码视频数据(222)。举例来说,视频解码器48和/或视频显示器44(或其再现单元或再现引擎)可使用序列层级SEI消息来确定如何使视频数据适于视频显示器44的原生分辨率,确定图像是左图像还是右图像(例如,对于立体视频数据来说),确定如何封装帧,确定视频数据的可缩放性信息,确定多视图获取信息,和/或确定多视图场景信息。

[0145] 图7是说明视频文件240的另一实例的框图。在图7的实例中,视频文件240包含电影(MOOV)盒242和电影数据(MDAT)盒250。MOOV盒242包含视频数据轨道和若干参数集轨道246A到246C。虽然在图7的实例中说明三个参数集轨道,但应了解,视频文件可大体上包含任一数目的参数集轨道。

[0146] 视频数据轨道244对应于若干各种视频片段254A到254C,其中每一者可存储在MDAT盒250中。在此实例中,视频片段254A、254B和254C的每一集合可具有不同的序列层级信息。举例来说,序列层级信息可在视频片段254A、254B和254C之间改变。因此,视频片段254A、254B和254C中的每一者可具有相应的序列数据集。在此实例中,序列数据集252A与视频片段254A相关,序列数据集252B与视频片段254B相关,且序列数据集252C与视频片段254C相关。

[0147] 视频片段的给定集合的序列数据集可应用于序列中的视频片段中的每一者。举例来说,序列数据集252A可应用于视频片段254A中的每一者。然而,视频片段254B的一个或一个以上特性可不同于视频片段254A。因此,序列数据集252B可描述视频片段254B的特性。

[0148] 每一序列数据集可对应于相应的参数集轨道。在图7的实例中,参数集轨道246A对应于序列数据集252A,参数集轨道246B对应于序列数据集252B,且参数集轨道246C对应于序列数据集252C。每一参数集轨道可包含描述一时间的数据,在所述时间期间,对应序列数据集应用于视频数据轨道244的视频片段。举例来说,参数集轨道246A可包含指示在对应于

视频片段254A中的最后一者的时间之后序列数据集252A不再适用的数据。此时序信息可为UTC时间或对应于重放持续时间。

[0149] 在一些实例中,视频文件可包含对应于所有视频片段254A到254C的额外参数集轨道。在此类实例中,目的地装置可利用来自适用于所有视频片段的参数集轨道的数据和来自特定适用于给定视频片段的参数集轨道的数据。

[0150] 虽然为了实例目的而说明为包含在M00V盒242中,但应了解,序列数据集252可包含在本发明中描述的其它位置中的任一者中。举例来说,序列数据集可包含在特定视频片段中、特定3GPP文件中、媒体简报描述内,或者视频文件内或与视频文件相关联的另外一或多个位置中。

[0151] 在一个或一个以上实例中,所描述的功能可以硬件、软件、固件或其任何组合来实施。如果以软件来实施,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,且可由基于硬件的处理单元执行。计算机可读媒体可包含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,或包含促进计算机程序(例如)根据一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。计算机程序产品可包含计算机可读媒体。

[0152] 作为实例而非限制,此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代码且可由计算机存取的任何其它媒体。且,任何连接均适当地称作计算机可读媒体。举例来说,如果使用同轴电缆、光缆、双绞线、数字订户线(DSL),或例如红外线、无线电及微波的无线技术而从网站、服务器或其它远程源传输指令,则同轴电缆、光缆、双绞线、DSL,或例如红外线、无线电及微波的无线技术包含于媒体的定义中。然而,应理解,计算机可读存储媒体及数据存储媒体不包含连接、载波、信号或其它瞬时媒体,而是针对非瞬时、有形的存储媒体。如本文中所使用,磁盘及光盘包含压缩光盘(CD)、激光光盘、光学光盘、数字多功能光盘(DVD)、软磁盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘通过激光以光学方式再现数据。以上各物的组合也应包含于计算机可读媒体的范围内。

[0153] 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)或其它等效集成电路或离散逻辑电路的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”可指代前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。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。

[0154] 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的集合(例如,芯片组)。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。实情为,如上文所描述,各种单元可组合于一编解码器硬件单元中或由互操作的硬件单元的集合(包含如上文

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。

[0155] 已描述各种实例。这些及其它实例属于所附权利要求书的范围内。

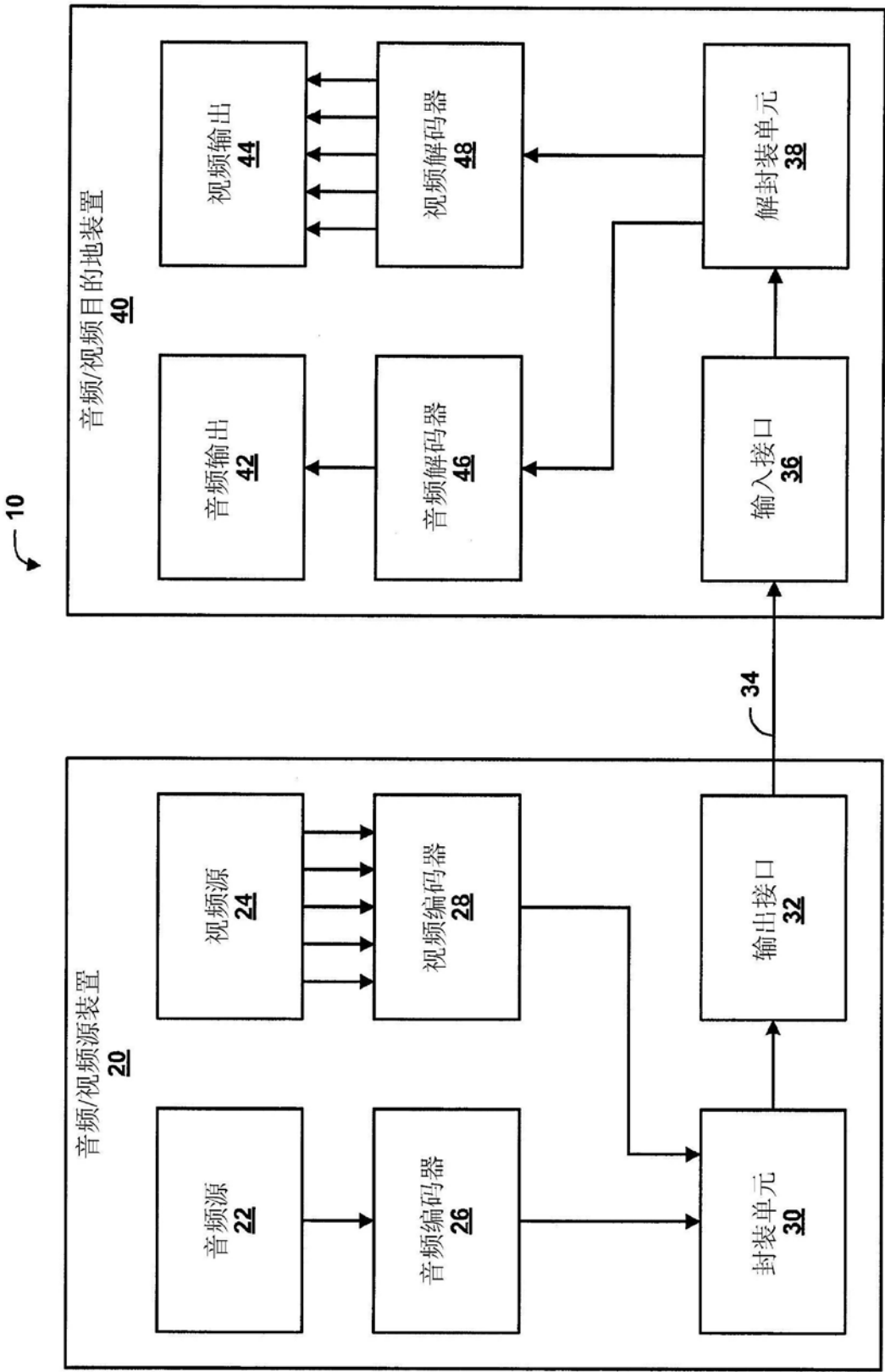


图1

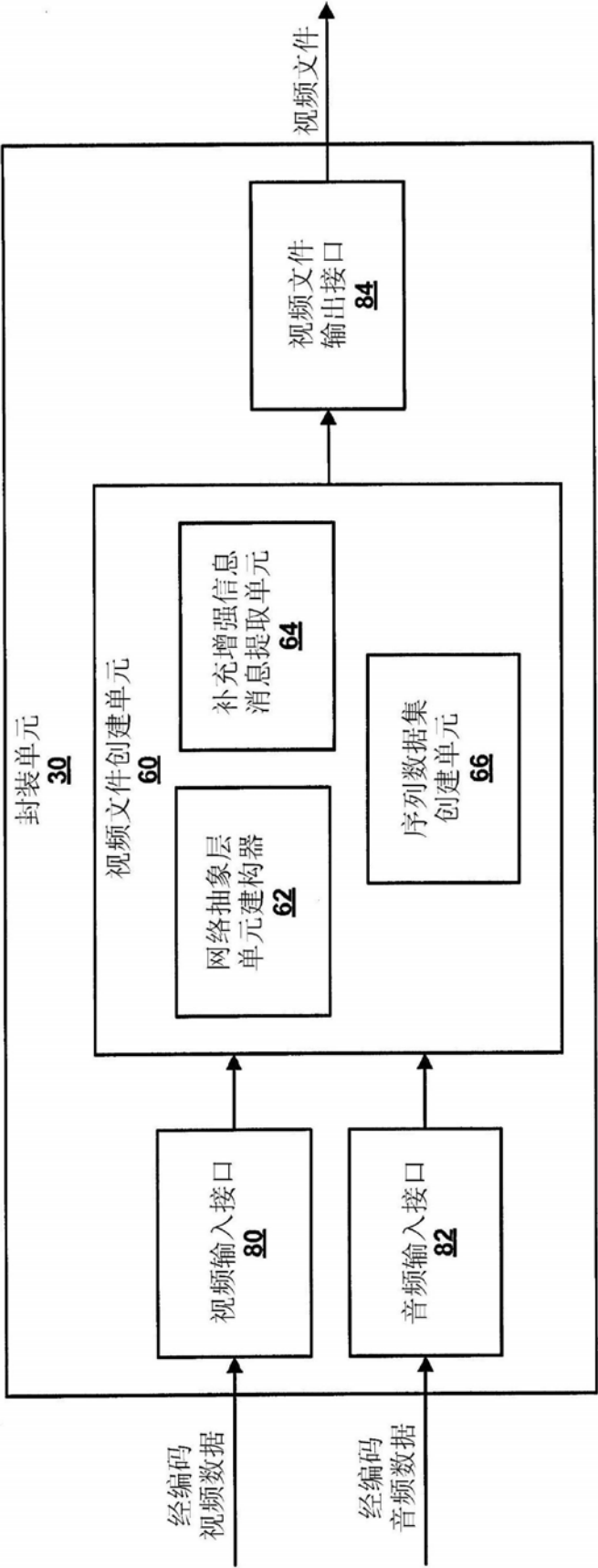


图2

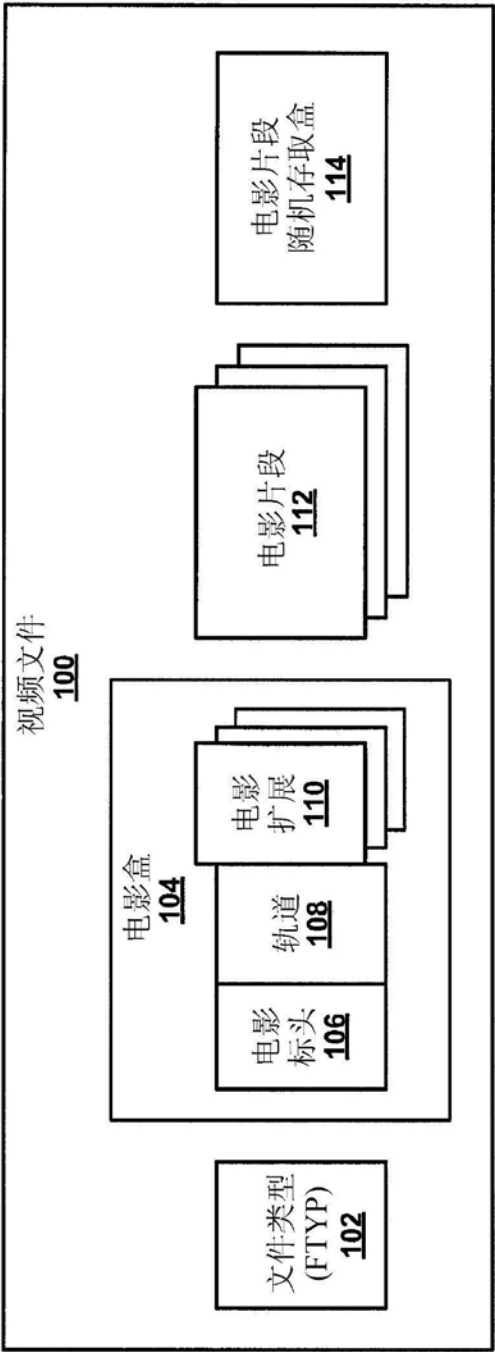


图3

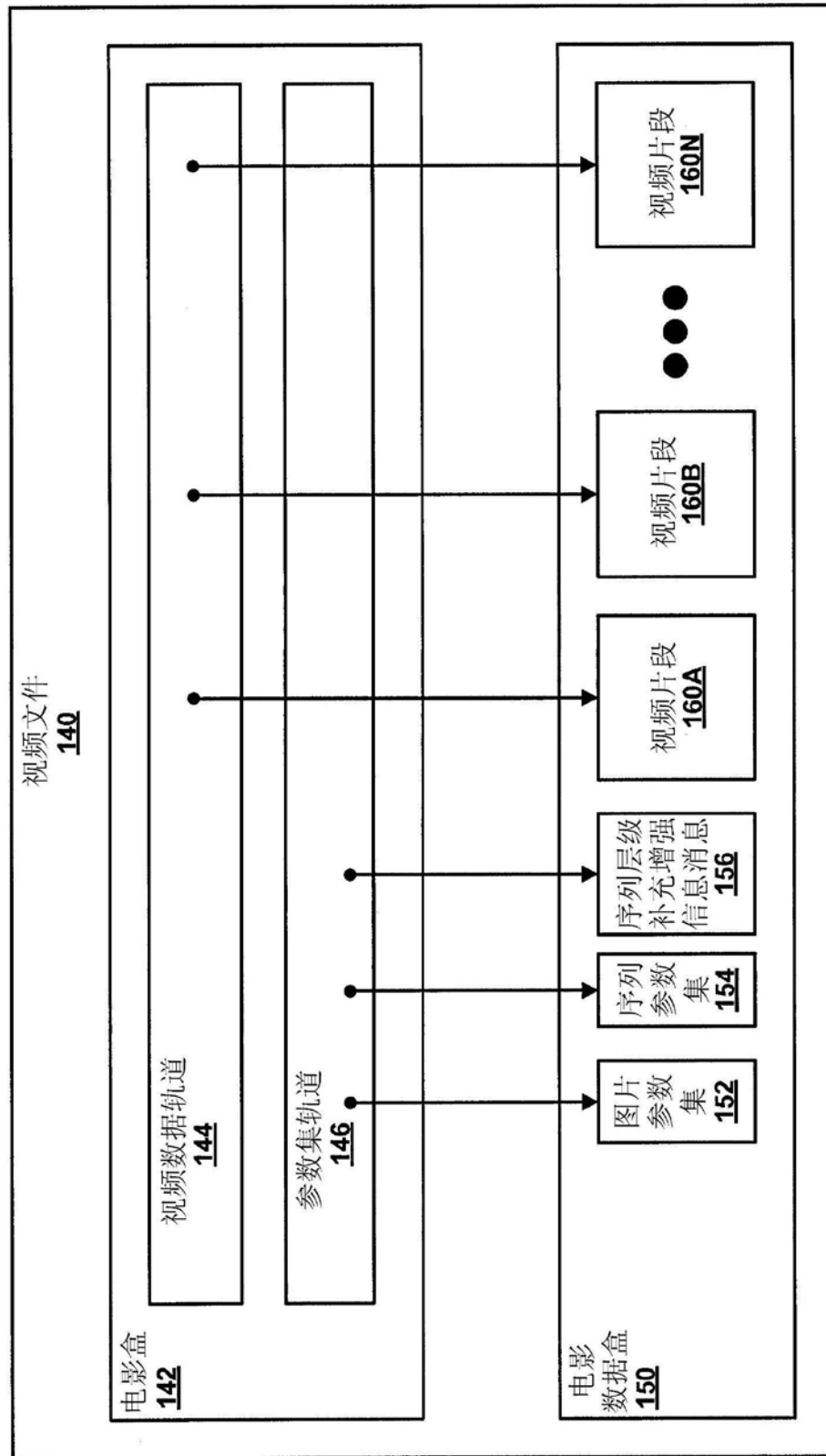


图4

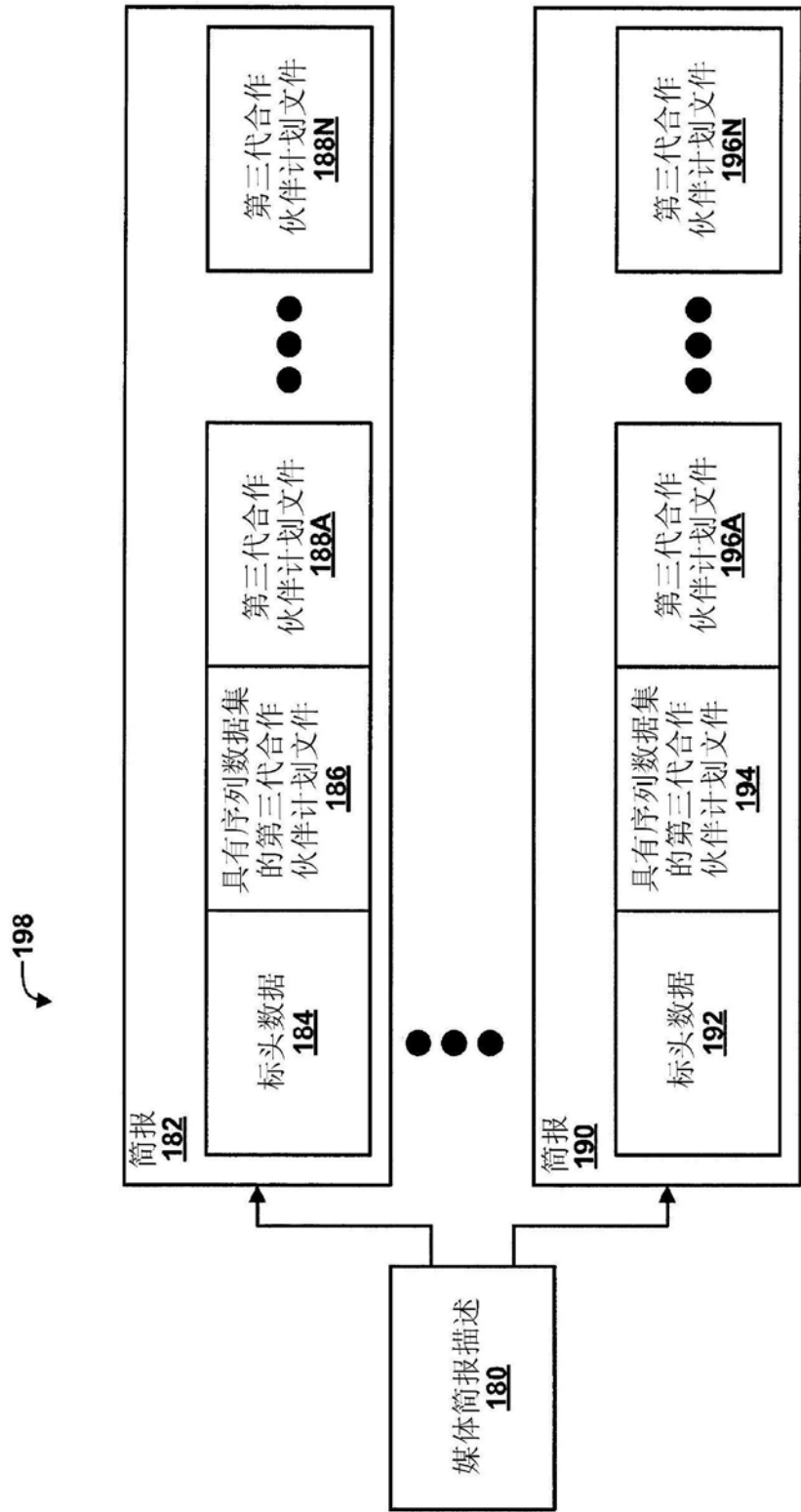


图5

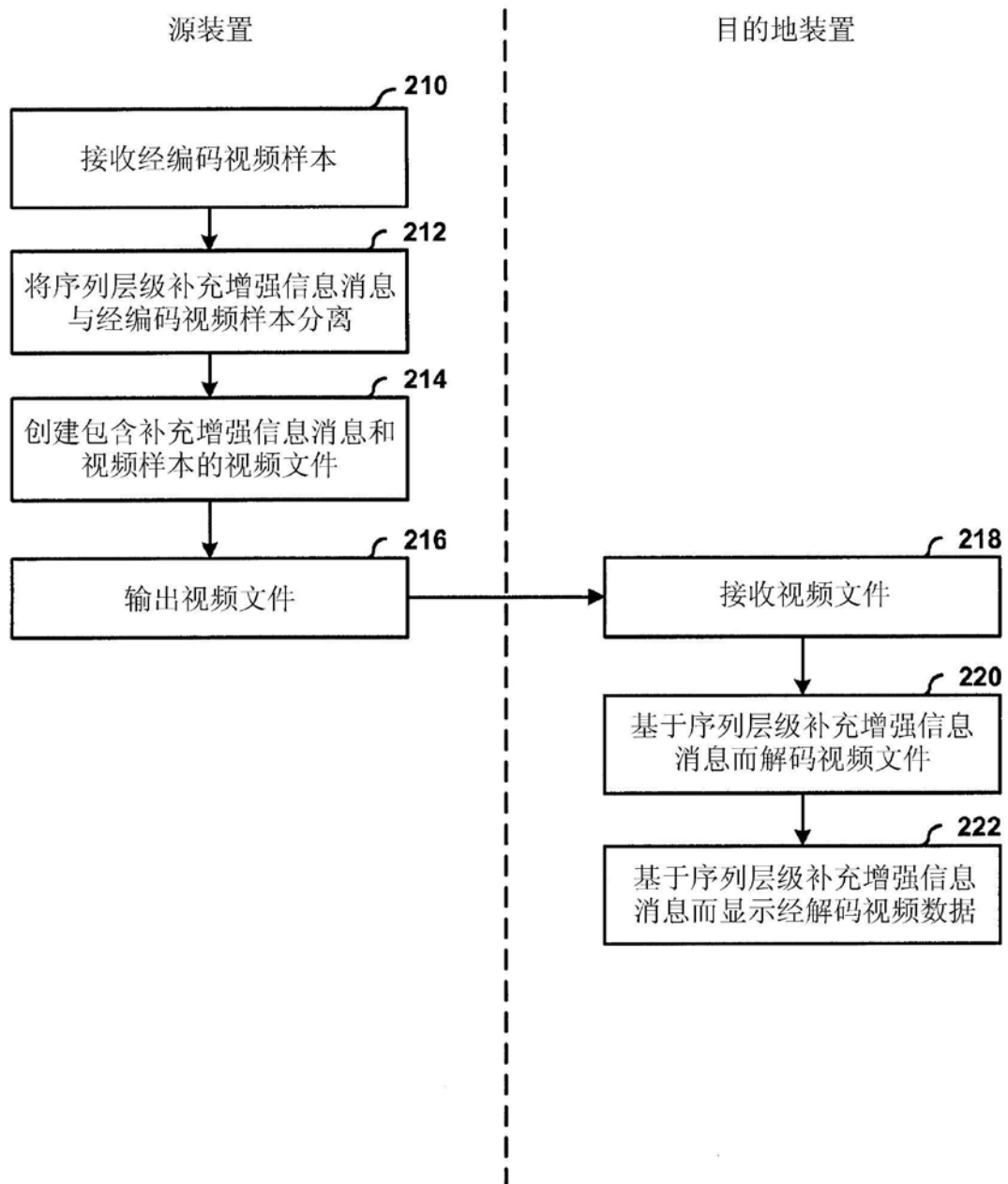


图6

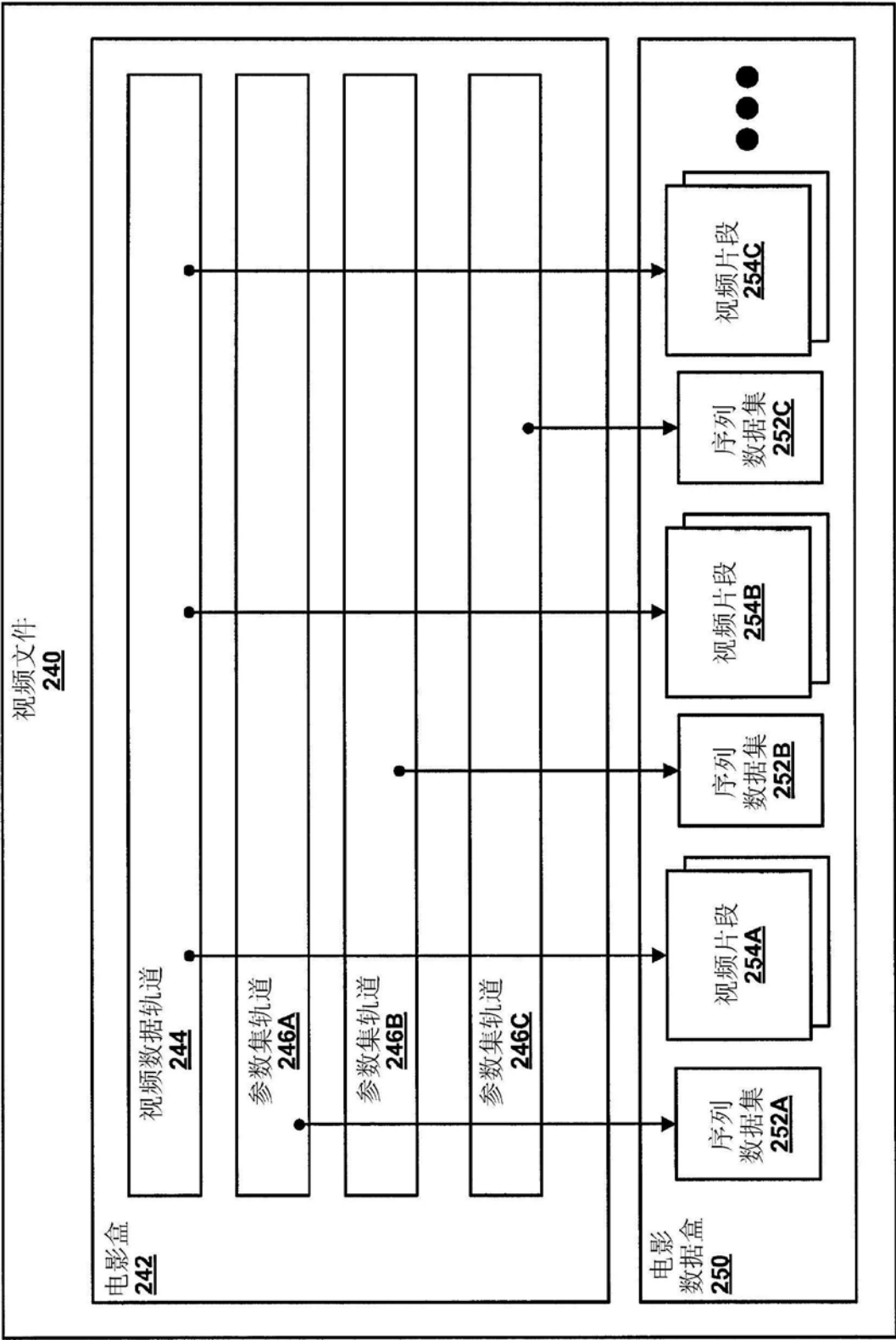


图7