



(12) 发明专利

(10) 授权公告号 CN 108702518 B

(45) 授权公告日 2020.12.18

(21) 申请号 201780011311.X

H04N 19/30 (2006.01)

(22) 申请日 2017.02.15

(56) 对比文件

(65) 同一申请的已公布的文献号  
申请公布号 CN 108702518 A

CN 105144719 A, 2015.12.09

CN 104604236 A, 2015.05.06

(43) 申请公布日 2018.10.23

US 2015110118 A1, 2015.04.23

US 2015373345 A1, 2015.12.24

(30) 优先权数据

MPEG MEETING.Text of ISO/IEC DIS

62/296,569 2016.02.17 US

14496-15 4th edition.《MOTION PICTURE

15/432,855 2017.02.14 US

EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11》

(85) PCT国际申请进入国家阶段日  
2018.08.15

.2015,

Miska M. Hannuksela, et al..MV-HEVC/  
SHVC HLS: On additional layer sets,  
rewriting of simulcast layers, and  
profile-tier-level indication for  
auxiliary picture layers.《Joint  
Collaborative Team on Video Coding (JCT-  
VC)》.2014,

(86) PCT国际申请的申请数据  
PCT/US2017/017979 2017.02.15

(87) PCT国际申请的公布数据  
W02017/142949 EN 2017.08.24

(73) 专利权人 高通股份有限公司  
地址 美国加利福尼亚州

Arturo A Rodriguez, et al..Prop SEI  
message to forewarn location of end\_of\_  
stream.《JOINT VIDEO TEAM OF ISO/IEC JTC1/  
SC29/WG11 AND ITU-T SG.16》.2008,

(72) 发明人 F·亨利 王业奎

审查员 张宁远

(74) 专利代理机构 北京律盟知识产权代理有限  
责任公司 11287

代理人 杨林勳

(51) Int.Cl.

H04N 19/70 (2006.01)

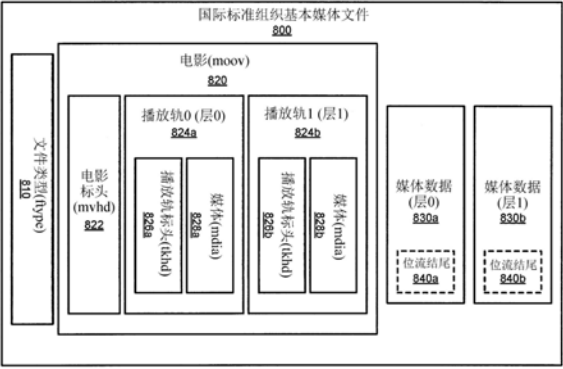
权利要求书5页 说明书30页 附图20页

(54) 发明名称

用于处理视频数据的方法、设备及计算机可读媒体

(57) 摘要

在各种实施方案中,本文中描述用以解决与ISO基本媒体文件中的EOB NAL单元的位置相关的问题的技术。在各种实施方案中,这些技术包含放宽文件包含仅仅一个EOB NAL单元的要求。这些技术进一步包含当重建位流时提供来自最高层的所述EOB NAL单元。这个EOB NAL单元可被添加到所述经重建位流中的最后存取单元。当来自所述最高层的所述EOB NAL单元不存在时,可改为提供来自较低层的EOB NAL单元。



CN 108702518 B

1. 一种用于处理经编码视频数据的方法,所述方法包括:

接收与经编码视频数据相关联的文件,其中所述经编码视频数据根据文件格式存储于所述文件中,其中所述经编码视频数据包含两个或多于两个层,且其中所述两个或多于两个层存储于所述文件中的两个或多于两个播放轨中;

从所述两个或多于两个层当中确定第一层,其中所述第一层包含第一位流结尾指示符,其中所述第一层经确定为所述两个或多于两个层当中的包含位流结尾指示符的最高层;

从所述两个或多于两个层当中识别一组层,其中所述一组层包含被选择以用于包含于经重建位流中的一或多个层;以及

产生所述经重建位流,其中产生所述经重建位流包含产生多个存取单元,其中所述多个存取单元包含来自所述一组层的一或多个样本,且其中所述经重建位流中的最后存取单元包含所述第一位流结尾指示符。

2. 根据权利要求1所述的方法,其进一步包括:

确定当产生所述多个存取单元时,所述经重建位流中的所述最后存取单元不包含位流结尾指示符;以及

添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

3. 根据权利要求1所述的方法,其进一步包括:

从所述一组层当中识别第二层,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层不包含所述第一层;

确定所述第一层为比所述第二层高的层;以及

基于所述第一层为比所述第二层高的层,添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

4. 根据权利要求1所述的方法,其进一步包括:

从所述一组层当中识别第二层,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层包含所述第一层;

确定所述第二层为比所述第一层高的层;以及

基于所述第二层为比所述第一层高的层,移动所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

5. 根据权利要求1所述的方法,其进一步包括:

确定所述经重建位流中的所述最后存取单元包含多于一个位流结尾指示符,其中所述多于一个位流结尾指示符包含所述第一位流结尾指示符;以及

从所述最后存取单元去除不同于所述第一位流结尾指示符的每一位流结尾指示符。

6. 根据权利要求1所述的方法,其中所述两个或多于两个层当中的至少两个层各自包含位流结尾指示符。

7. 根据权利要求1所述的方法,其进一步包括:

识别所述一组层是基于来自所述一组层中的经选择用于显示的层的样本而选择。

8. 根据权利要求1所述的方法,其中所述文件格式是从国际标准组织ISO基本媒体文件格式衍生。

9. 一种用于处理经编码视频的设备,其包括:

存储器,其经配置以存储文件,其中所述文件包含经编码视频数据,其中所述经编码视频数据根据文件格式存储于所述文件中,其中所述经编码视频数据包含两个或多于两个层,且其中所述两个或多于两个层存储于所述文件中的两个或多于两个播放轨中;以及处理器,其经配置以:

从所述两个或多于两个层当中确定第一层,其中所述第一层包含第一位流结尾指示符,其中所述第一层经确定为所述两个或多于两个层当中的包含位流结尾指示符的最高层;

从所述两个或多于两个层当中识别一组层,其中所述一组层包含被选择以用于包含于经重建位流中的一或多个层;以及

产生所述经重建位流,其中产生所述经重建位流包含产生多个存取单元,其中所述多个存取单元包含来自所述一组层的一或多个样本,且其中所述经重建位流中的最后存取单元包含所述第一位流结尾指示符。

10. 根据权利要求9所述的设备,其中所述处理器经进一步配置以:

确定当产生所述多个存取单元时,所述经重建位流中的所述最后存取单元不包含位流结尾指示符;以及

添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

11. 根据权利要求9所述的设备,其中所述处理器经进一步配置以:

从所述一组层当中识别第二层,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层不包含所述第一层;

确定所述第一层为比所述第二层高的层;以及

基于所述第一层为比所述第二层高的层,添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

12. 根据权利要求9所述的设备,其中所述处理器经进一步配置以:

从所述一组层当中识别第二层,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层包含所述第一层;

确定所述第二层为比所述第一层高的层;以及

基于所述第二层为比所述第一层高的层,移动所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

13. 根据权利要求9所述的设备,其中所述处理器经进一步配置以:

确定所述经重建位流中的所述最后存取单元包含多于一个位流结尾指示符,其中所述多于一个位流结尾指示符包含所述第一位流结尾指示符;以及

从所述最后存取单元去除不同于所述第一位流结尾指示符的每一位流结尾指示符。

14. 根据权利要求9所述的设备,其中所述两个或多于两个层当中的至少两个层各自包含位流结尾指示符。

15. 根据权利要求9的设备,其中所述一组层是基于来自所述一组层中的经选择用于显示的层的所述样本而选择。

16. 根据权利要求9所述的设备,其中所述文件格式是从ISO基本媒体文件格式衍生。

17. 一种上面存储有指令的非暂时性计算机可读媒体,所述指令在由一或多个处理器执行时使得所述一或多个处理器执行以下操作:

接收与经编码视频数据相关联的文件,其中所述经编码视频数据根据文件格式存储于所述文件中,其中所述经编码视频数据包含两个或多于两个层,且其中所述两个或多于两个层存储于所述文件中的两个或多于两个播放轨中;

从所述两个或多于两个层当中确定第一层,其中所述第一层包含第一位流结尾指示符,其中所述第一层经确定为所述两个或多于两个层当中的包含位流结尾指示符的最高层;

从所述两个或多于两个层当中识别一组层,其中所述一组层包含被选择以用于包含于经重建位流中的一或多个层;以及

产生所述经重建位流,其中产生所述经重建位流包含产生多个存取单元,其中所述多个存取单元包含来自所述一组层的一或多个样本,且其中所述经重建位流中的最后存取单元包含所述第一位流结尾指示符。

18. 根据权利要求17所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器执行以下操作:

确定当产生所述多个存取单元时,所述经重建位流中的所述最后存取单元不包含位流结尾指示符;以及

添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

19. 根据权利要求17所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器执行以下操作:

从所述一组层当中识别第二层,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层不包含所述第一层;

确定所述第一层为比所述第二层高的层;以及

基于所述第一层为比所述第二层高的层,添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

20. 根据权利要求17所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器执行以下操作:

从所述一组层当中识别第二层,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层包含所述第一层;

确定所述第二层为比所述第一层高的层;以及

基于所述第二层为比所述第一层高的层,移动所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元。

21. 根据权利要求17所述的非暂时性计算机可读媒体,其进一步包括指令,所述指令在由所述一或多个处理器执行时使得所述一或多个处理器执行以下操作:

确定所述经重建位流中的所述最后存取单元包含多于一个位流结尾指示符,其中所述多于一个位流结尾指示符包含所述第一位流结尾指示符;以及

从所述最后存取单元去除不同于所述第一位流结尾指示符的每一位流结尾指示符。

22. 根据权利要求17所述的非暂时性计算机可读媒体,其中所述两个或多于两个层当中的至少两层各自包含位流结尾指示符。

23. 根据权利要求17所述的非暂时性计算机可读媒体,其中所述一组层是基于来自所述一组层中的经选择用于显示的层的所述样本而选择。

24. 根据权利要求17所述的非暂时性计算机可读媒体,其中所述文件格式是从ISO基本媒体文件格式衍生。

25. 一种用于处理经编码视频数据的设备,其包括:

用于接收与经编码视频数据相关联的文件的装置,其中所述经编码视频数据根据文件格式存储于所述文件中,其中所述经编码视频数据包含两个或多于两个层,且其中所述两个或多于两个层存储于所述文件中的两个或多于两个播放轨中;

用于从所述两个或多于两个层当中确定第一层的装置,其中所述第一层包含第一位流结尾指示符,其中所述第一层经确定为所述两个或多于两个层当中的包含位流结尾指示符的最高层;

用于从所述两个或多于两个层当中识别一组层的装置,其中所述一组层包含被选择以用于包含于经重建位流中的一或多个层;以及

用于产生经重建位流的装置,其中产生所述经重建位流包含产生多个存取单元,其中所述多个存取单元包含来自所述一组层的一或多个样本,且其中所述经重建位流中的最后存取单元包含所述第一位流结尾指示符。

26. 根据权利要求25所述的设备,其进一步包括:

用于确定当产生所述多个存取单元时所述经重建位流中的所述最后存取单元不包含位流结尾指示符的装置;以及

用于添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元的装置。

27. 根据权利要求25所述的设备,其进一步包括:

用于从所述一组层当中识别第二层的装置,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层不包含所述第一层;

用于确定所述第一层为比所述第二层高的层的装置;以及

用于基于所述第一层为比所述第二层高的层添加所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元的装置。

28. 根据权利要求25所述的设备,其进一步包括:

用于从所述一组层当中识别第二层的装置,其中所述第二层经识别为所述一组层当中的最高层,其中所述第二层与所述第一层不同,且其中所述一组层包含所述第一层;

用于确定所述第二层为比所述第一层高的层的装置;以及

用于基于所述第二层为比所述第一层高的层移动所述第一位流结尾指示符到所述经重建位流中的所述最后存取单元的装置。

29. 根据权利要求25所述的设备,其进一步包括:

用于确定所述经重建位流中的所述最后存取单元包含多于一个位流结尾指示符的装置,其中所述多于一个位流结尾指示符包含所述第一位流结尾指示符;以及

用于从所述最后存取单元去除不同于所述第一位流结尾指示符的每一位流结尾指示符的装置。

30. 根据权利要求25所述的设备,其中所述两个或多于两个层当中的至少两个层各自包含位流结尾指示符。

31. 根据权利要求25所述的设备,其进一步包括:

用于识别所述一组层是基于来自所述一组层中的经选择用于显示的层的样本而选择

的装置。

32. 根据权利要求25所述的设备,其中所述文件格式是从国际标准组织ISO基本媒体文件格式衍生。

## 用于处理视频数据的方法、设备及计算机可读媒体

### 技术领域

[0001] 本申请案涉及在分层高效视频译码(L-HEVC, 也称作LHEVC) 位流存储于(例如, 写入到) 文件中时位流结尾(EOB) 网络抽象层(NAL) 单元的处理。另外, 本申请案也描述对HEVC和L-HEVC频块播放轨的改善, 其中包含用于避免HEVC和L-HEVC频块播放轨中的循环参考的方法和用于用信号传送频块播放轨与含有完整图片的视频的相应播放轨的相对位置的方法。

### 背景技术

[0002] 视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1Visual、ITU-T H.262或ISO/IEC MPEG-2Visual、ITU-T H.263、ISO/IEC MPEG-4Visual、ITU-T H.264或ISO/IEC MPEG-4 AVC, 包含其可调式视频译码(SVC) 和多视图视频译码(MVC) 扩展, 和高效视频译码(HEVC), 也被称作ITU-T H.265和ISO/IEC 23008-2, 包含其可调式译码扩展(即, 可调式高效视频译码, SHVC) 和多视图扩展(即, 多视图高效视频译码, MV-HEVC)。

### 发明内容

[0003] 在一些实施例中, 描述用于在分层高效视频译码(L-HEVC) 写入到文件时提供位流结尾(EOB) 网络存取层(NAL) 单元, 以及用于读取此文件的技术。

[0004] 根据至少一个实例, 提供一种用于处理经编码视频数据的方法, 所述方法包含接收与经编码视频数据相关联的文件。经编码视频数据可根据文件格式存储于文件中。经编码视频数据可包含两个或多于两个层。两个或多于两个层可存储于文件中的两个或多于两个播放轨中。方法进一步包含从两个或多于两个层当中确定第一层。第一层可经确定为两个或多于两个层当中的包含位流结尾指示符的最高层。方法进一步包含确定第一位流结尾指示符。第一位流结尾指示符可为来自第一层的位流结尾指示符。方法进一步包含从两个或多于两个层当中选择一组层。方法进一步包含产生一系列存取单元。所述系列存取单元可包含来自所述一组层的样本。方法进一步包含使用所述系列存取单元和第一位流结尾指示符产生经重建位流。经重建位流中的最后存取单元可包含第一位流结尾指示符。

[0005] 在另一实例中, 提供一种包含经配置以存储视频数据的存储器和处理器的设备。处理器经配置以且可接收与经编码视频数据相关联的文件。经编码视频数据可根据文件格式存储于文件中。经编码视频数据可包含两个或多于两个层。两个或多于两个层可存储于文件中的两个或多于两个播放轨中。处理器经配置以且可从两个或多于两个层当中确定第一层。第一层可经确定为两个或多于两个层当中的包含位流结尾指示符的最高层。处理器经配置以且可确定第一位流结尾指示符。第一位流结尾指示符可为来自第一层的位流结尾指示符。处理器经配置以且可从两个或多于两个层当中选择一组层。处理器经配置以且可产生一系列存取单元。所述系列存取单元可包含来自所述一组层的样本。处理器经配置以且可使用所述系列存取单元和第一位流结尾指示符产生经重建位流。经重建位流中的最后存取单元可包含第一位流结尾指示符。

[0006] 在另一实例中,提供一种其上存储有指令的计算机可读媒体,所述指令在由处理器执行时执行包含以下操作的方法:接收与经编码视频数据相关联的文件。经编码视频数据可根据文件格式存储于文件中。经编码视频数据可包含两个或多于两个层。两个或多于两个层可存储于文件中的两个或多于两个播放轨中。方法进一步包含从两个或多于两个层当中确定第一层。第一层可经确定为两个或多于两个层当中的包含位流结尾指示符的最高层。方法进一步包含确定第一位流结尾指示符。第一位流结尾指示符可为来自第一层的位流结尾指示符。方法进一步包含从两个或多于两个层当中选择一组层。方法进一步包含产生一系列存取单元。所述系列存取单元可包含来自所述一组层的样本。方法进一步包含使用所述系列存取单元和第一位流结尾指示符产生经重建位流。经重建位流中的最后存取单元可包含第一位流结尾指示符。

[0007] 在另一实例中,提供一种包含用于接收与经编码视频数据相关联的文件的装置的设备的装置。经编码视频数据可根据文件格式存储于文件中。经编码视频数据可包含两个或多于两个层。两个或多于两个层可存储于文件中的两个或多于两个播放轨中。设备进一步包括用于从两个或多于两个层当中确定第一层的装置。第一层可经确定为两个或多于两个层当中的包含位流结尾指示符的最高层。设备进一步包括用于确定第一位流结尾指示符的装置。第一位流结尾指示符可为来自第一层的位流结尾指示符。设备进一步包含用于从两个或多于两个层当中选择一组层的装置。设备进一步包含用于产生一系列存取单元的装置。所述系列存取单元可包含来自所述一组层的样本。设备进一步包含用于使用所述系列存取单元和第一位流结尾指示符产生经重建位流的装置。经重建位流中的最后存取单元可包含第一位流结尾指示符。

[0008] 在一些方面中,上文所描述的方法、设备和计算机可读媒体进一步包括确定来自所述系列存取单元的最后存取单元不包含位流结尾指示符。这些方面进一步包含添加第一位流结尾指示符到所述系列存取单元中的最后存取单元。

[0009] 在一些方面中,上文所描述的方法、设备和计算机可读媒体进一步包括识别所述一组层当中的第二层,其中第二层经识别为所述一组层当中的最高层。这些方面进一步包含确定第一层为比第二层高的层。这些方面进一步包含添加第一位流结尾指示符到所述系列存取单元中的最后存取单元。

[0010] 在一些方面中,上文所描述的方法、设备和计算机可读媒体进一步包括识别来自所述一组层当中的第二层,其中第二层经识别为所述一组层当中的最高层。这些方面进一步包含确定第二层为比第一层高的层。这些方面进一步包含将第一位流结尾指示符移动到所述系列存取单元中的最后存取单元。

[0011] 在一些方面中,上文所描述的方法、设备和计算机可读媒体进一步包括确定经重建位流中的最后存取单元包含多于一个位流结尾指示符。在这些方面中,多于一个位流结尾指示符包含第一位流结尾指示符。这些方面进一步包含从最后存取单元中去除除第一位流结尾指示符以外的所有多于一个位流结尾指示符。

[0012] 在一些方面中,两个或多于两个层当中的至少两个层包含位流结尾指示符。在一些方面中,所述一组层是基于来自所述一组层中的经选择用于显示的层的样本而选择。在一些方面中,文件格式是从ISO基本媒体文件格式衍生。

[0013] 此发明内容并不希望识别所主张的标的物的关键或基本特征,也不希望孤立使用



以确定所主张的标的物的范围。应参照此专利的整个说明书的合适部分、任何或所有图式和每一权利要求书来理解标的物。

[0014] 当参考以下说明书、权利要求书和随附图式时,前述内容连同其它特征和实施例将变得更显而易见。

## 附图说明

[0015] 下文参考如下图式详细描述本发明的说明性实施例:

[0016] 图1为说明包含编码装置和解码装置的系统的实例的框图。

[0017] 图2说明分层HEVC位流的简化实例。

[0018] 图3说明含有根据ISOBMFF格式化的用于视频呈现的数据和元数据的ISO基本媒体文件的实例。

[0019] 图4说明用于视频呈现的包含两层的ISO基本媒体文件的实例。

[0020] 图5A说明其中已存储多层位流的ISO基本媒体文件的一个实例。

[0021] 图5B和图5C说明在图5A中所说明的ISO基本媒体文件通过不同解码器装置读取时可产生的位流的实例。

[0022] 图6A说明其中已存储多层位流的ISO基本媒体文件的另一实例。

[0023] 图6B和图6C说明可从读取图6A中所说明的ISO基本媒体文件产生的位流的实例。

[0024] 图7说明其中已存储多层位流的ISO基本媒体文件的另一实例。

[0025] 图8A说明根据这些技术产生的ISO基本媒体文件的实例。

[0026] 图8B-8E说明可从图8A中所说明的文件构建的位流的实例。

[0027] 图9说明可包含于ISO基本媒体文件中的媒体框和媒体数据框的实例。

[0028] 图10说明用于处理经编码视频数据的过程的实例,其中包含经编码视频数据的文件可具有多于一个EOB NAL单元。

[0029] 图11为说明实例编码装置的框图。

[0030] 图12为说明实例解码装置的框图。

## 具体实施方式

[0031] 下文提供本发明的某些方面和实施例。这些方面和实施例中的一些可独立地应用且其中的一些可以将对所属领域的技术人员显而易见的组合来应用。在以下描述中,出于解释的目的,阐述众多特定细节以便提供对本发明的实施例的透彻理解。然而,各种实施例可在无这些特定细节的情况下得以实践将是显而易见的。图式和描述不希望为限定性的。

[0032] 随后描述仅提供示范性实施例且并不希望限制本发明的范围、可应用性或配置。相反地,示范性实施例的随后描述将为所属领域的技术人员提供用于实施示范性实施例的能够实现的描述。应理解,可在不脱离如所附权利要求书中所阐述的本发明的精神和范围的情况下对元件的功能和布置做出各种变化。

[0033] 在以下描述中给定特定细节从而提供实施例的透彻理解。然而,所属领域的一般技术人员应了解,实施例可在无需这些特定细节的情况下实践。举例来说,电路、系统、网络、过程和其它组件可在框图形式中展示为组件以便不以不必要细节混淆实施例。在其它情况下,已在无不必要的细节的情况下展示众所周知电路、过程、算法、结构和技术以便避

免混淆实施例。

[0034] 又,应注意,个别实施例可描述为经描绘为流程图、流程图式、数据流程图、结构图或框图的过程。尽管流程图可能将操作描述为顺序过程,但许多操作可并行地或同时来执行。另外,可重新排列操作的次序。当过程的操作完成时终止过程,但过程可具有不包含于图中的额外步骤。过程可对应于方法、函数、程序、次例程、子程序,等等。当过程对应于函数时,其终止可对应于函数返回到调用函数或主函数。

[0035] 术语“计算机可读媒体”包含(但不限于)便携式或非便携式存储装置、光学存储装置和能够存储、含有或携载指令和/或数据的各种其它媒体。计算机可读媒体可包含非暂时性媒体,其中可存储数据并且不包含载波和/或无线地或经由有线连接传播的暂时电子信号。非暂时性媒体的实例可包含(但不限于)磁盘或磁带、光学存储媒体,例如紧密光盘(CD)或数字化通用光盘(DVD)、闪存、存储器或存储器装置。计算机可读媒体可具有存储于其上的代码和/或机器可执行指令,代码和/或机器可执行指令可表示程序(procedure)、函数、子程序、程序(program)、例程、次例程、模块、软件套件、类别,或指令、数据结构或程序陈述式的任何组合。可通过传递和/或接收信息、数据、引数、参数或存储器内容将码段耦合到另一码段或硬件电路。信息、引数、参数、数据等可经由包含存储器共享、消息传递、符记传递、网络传输等的任何合适方式传递、转递或传输。

[0036] 此外,可由硬件、软件、固件、中间软件、微码、硬件描述语言或其任何组合实施方法的实施例。当以软件、固件、中间软件或微码实施时,用以执行必要任务的程序代码或码段(例如,计算机程序产品)可存储于计算机可读或机器可读媒体中。处理器可执行必要任务。

[0037] 文件格式标准可定义用于将视频(和也可能音频)数据打包和解包成一或多个文件的格式。文件格式标准包含国际标准化组织(ISO)基本媒体文件格式(ISOBMFF,定义于ISO/IEC 14496-12中)和从ISOBMFF衍生的其它文件格式,包含运动图片专家组(MPEG)MPEG-4文件格式(定义于ISO/IEC 14496-15中)、第三代合作伙伴计划(3GPP)文件格式(定义于3GPP TS 26.244中)和先进视频译码(AVC)文件格式(定义于ISO/IEC 14496-15中)。ISO/IEC 14496-12和14496-15的最新版本的草案本文可分别在[http://phenix.int-evry.fr/mpeg/doc\\_end\\_user/documents/111\\_Geneva/wg11/w15177-v6-w15177.zip](http://phenix.int-evry.fr/mpeg/doc_end_user/documents/111_Geneva/wg11/w15177-v6-w15177.zip)和[http://phenix.int-evry.fr/mpeg/doc\\_end\\_user/documents/112\\_Warsaw/wg11/w15479-v2-w15479.zip](http://phenix.int-evry.fr/mpeg/doc_end_user/documents/112_Warsaw/wg11/w15479-v2-w15479.zip)处得到。

[0038] ISOBMFF用作许多编码解码器封装格式(例如,AVC文件格式或任何其它合适的编码解码器封装格式)以及许多多媒体容器格式(例如,MPEG-4文件格式、3GPP文件格式(3GP)、DVB文件格式或任何其它合适的多媒体容器格式)的基础。ISOBMFF基本文件格式可用于连续媒体,其也称作流式传输媒体。

[0039] 除连续媒体(例如,音频和视频)以外,静态媒体(例如,图像)和元数据可存储于遵循ISOBMFF的文件中。根据ISOBMFF结构化的文件可用于许多目的,包含本地媒体文件播放、远程文件的渐进式下载、用作HTTP动态自适应流式传输(DASH)的片段、用作待流式传输的内容的容器(在此状况下容器包含封装指令)、用于记录接收的实时媒体流式传输,或其它用途。

[0040] 使用HEVC或HEVC的扩展编码的位流大体包含在位流结尾处的指示符,指示符可向

解码器指示已达到位流结尾。指示符可采取位流结尾 (EOB) 网络抽象层 (NAL) 单元的形式。

[0041] 当位流包含多层时,EOB NAL单元将仍放置于位流的结尾处,但当位流写入到文件时EOB NAL单元所放置的处可变化,且可引起解码器装置的问题。举例来说,当EOB NAL单元在最高层中且解码器装置仅仅读取基层时,EOB NAL单元可丢失。作为另一实例,当EOB NAL单元在基层中且解码器装置经配置以再现较高层时,EOB NAL单元可放置于经重建位流中的除经重建位流的结尾以外的位置中。作为另一实例,当基层是使用除HEVC以外的编码解码器(例如,AVC)编码且也包含EOB NAL单元时,EOB NAL单元将已经编码为非HEVC NAL单元。在此实例中,当解码器装置经配置以再现来自较高层的HEVC样本时,解码器装置可接收并非为HEVC NAL单元的EOB NAL单元。

[0042] 在各种实施方案中,本文中描述用以解决与ISO基本媒体文件中的EOB NAL单元的定位(例如,位置)相关的问题的技术。在各种实施方案中,这些技术包含放宽文件包含仅仅一个EOB NAL单元的要求。这些技术进一步包含当重建位流时提供来自最高层的EOB NAL单元。当来自最高层的EOB NAL单元不存在时,可改为提供来自较低层的EOB NAL单元。

[0043] 当ISO基本媒体文件包含频块播放轨时,各种问题可产生。举例来说,循环提取器参考、由于提取器的复制数据、不必要提取器、当仅仅需要频块播放轨的视频部分时缺少用于重建位流的规范、缺少用于频块播放轨的偏移信息和其它问题可发生。下文论述这些问题的各种解决方案。

[0044] 图1为说明包含编码装置104和解码装置112的系统100的实例的框图。编码装置104可为源装置的部分,且解码装置112可为接收装置的部分。源装置和/或接收装置可包含电子装置,例如,移动或静止电话手机(例如,智能电话、蜂窝式电话等)、台式计算机、膝上型计算机或笔记本电脑、平板电脑、机顶盒、电视、相机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置或任何其它合适的电子装置。在一些实例中,源装置和接收装置可包含用于无线通信的一或多个无线收发器。本文中所描述的译码技术适用于各种多媒体应用中的视频译码,包含流式传输视频传输(例如,经由因特网)、电视广播或传输、编码数字视频以供存储于数据存储媒体上、解码存储于数据存储媒体上的数字视频或其它应用。在一些实例中,系统100可支持单向或双向视频传输以支持例如视频会议、视频流式传输、视频播放、视频广播、游戏和/或视频电话的应用。

[0045] 通过使用视频译码标准或协议以产生经编码视频位流,编码装置104(或编码器)可用以编码视频数据,包含虚拟实境视频数据。视频译码标准包含ITU-T H.261、ISO/IEC MPEG-1Visual、ITU-T H.262或ISO/IEC MPEG-2Visual、ITU-T H.263、ISO/IEC MPEG-4 Visual和ITU-T H.264(也称为ISO/IEC MPEG-4AVC),包含其可调式视频译码和多视图视频译码扩展(分别称为SVC和MVC)。已由ITU-T视频译码专家组(VCEG)和ISO/IEC 动画专家组(MPEG)的视频译码联合协作小组(JCT-VC)完成较近期的视频译码标准(高效率视频译码(HEVC))。对HEVC的各种扩展处理多层视频译码,且也正由JCT-VC开发,包含对HEVC的多视图扩展(被称作MV-HEVC)和对HEVC的可调式扩展(被称作SHVC)或任何其它合适的译码协议。

[0046] 本文中所描述的实施方案使用HEVC标准或其扩展来描述实例。然而,本文中所描述的技术和系统也可适用于其它译码标准,例如,AVC、MPEG、其扩展或已可用或尚不可用或尚未开发的其它合适译码标准。因此,虽然可参考特定视频译码标准描述本文中所描述的

技术和系统,但一般所属领域的技术人员将了解,描述不应解译为仅适用于彼特定标准。

[0047] 视频源102可将视频数据提供到编码装置104。视频源102可为源装置的一部分,或可为除源装置以外的装置的一部分。视频源102可包含视频捕获装置(例如,摄像机、摄影机电话、视频电话等)、含有经存储视频的视频存档、提供视频数据的视频服务器或内容提供者、从视频服务器或内容提供者接收视频的视频馈入接口、用于产生计算机图形视频数据的计算机图形系统、这些源的组合或任何其它合适的视频源。视频源102的一个实例可包含因特网协议摄影机(IP摄影机)。IP摄影机为可用于监测、住家安全或其它合适应用的数字摄像机的类型。不同于模拟闭路电视(CCTV)摄影机,IP摄影机可经由计算机网络和因特网发送和接收数据。

[0048] 来自视频源102的视频数据可包含一或多个输入图片或帧。图片或帧为作为视频的部分的静态图像。编码装置104的编码器引擎106(或编码器)编码视频数据以产生经编码视频位流。在一些实例中,经编码视频位流(或“视频位流”或“位流”)为一系列的一或多个经译码视频序列。经译码视频序列(CVS)包含一系列存取单元(AU),其始于具有在基层中且具有某些性质的随机存取点图片的AU,直到且不包含具有在基层中且具有某些性质的随机存取点图片的下一AU。举例来说,开始CVS的随机存取点图片的某些性质可包含等于1的RASL旗标(例如,NoRaslOutputFlag)。否则,随机存取点图片(具有等于0的RASL旗标)并不开始CVS。存取单元(AU)包含一或多个经译码图片以及对应于共享相同输出时间的经译码图片的控制信息。图片的经译码图块在位流层级包封到称为网络抽象层(NAL)单元的数据单元中。举例来说,HEVC视频位流可包含一或多个 CVS,所述一或多个CVS包含NAL单元。两种类别的NAL单元存在于HEVC标准中,包含视频译码层(VCL)NAL单元和非VCL NAL单元。VCL NAL单元包含经译码图片数据的一个图块或图块片段(如下描述),且非VCL NAL单元包含关于一或多个经译码图片的控制信息。

[0049] NAL单元可含有形成视频数据的经译码表示(例如,视频中的图片的经译码表示)的位序列(例如,经编码视频位流、位流的CVS等)。编码器引擎106通过将每一图片分割成多个图块而产生图片的经译码表示。图块接着分割成亮度样本和色度样本的译码树型块(CTB)。亮度样本的CTB和色度样本的一或多个CTB连同样本的语法被称为译码树型单元(CTU)。CTU为用于HEVC编码的基本处理单元。CTU可经分裂成具有不同大小的多个译码单元(CU)。CU含有被称为译码块(CB)的亮度和色度样本阵列。

[0050] 亮度和色度CB可进一步分裂成预测块(PB)。PB为使用用于帧间预测的相同运动参数的亮度或色度分量的样本的块。亮度PB和一或多个色度PB以及相关语法形成预测单元(PU)。在位流中针对每个PU用信号传送运动参数集合,且所述运动参数集合用于亮度PB和一或多个色度PB的帧间预测。CB也可分割成一或多个变换块(TB)。TB表示色彩分量的样本的正方形块,对所述正方形块应用同一二维变换以用于译码预测残余信号。变换单元(TU)表示亮度和色度样本的TB以及相应语法元素。

[0051] CU的大小对应于译码节点的大小,且可为正方形形状。举例来说,CU的大小可为 $8 \times 8$ 样本、 $16 \times 16$ 样本、 $32 \times 32$ 样本、 $64 \times 64$ 样本或达到相应CTU的大小的任何其它适当大小。短语“ $N \times N$ ”在本文中用于指就垂直和水平尺寸来说视频块的像素尺寸(例如,8像素 $\times$ 8像素)。可按行和列来排列块中的像素。在一些实施例,块在水平方向上可不具有与在垂直方向上相同的数目个像素。与CU相关联的语法数据可描述(例如)将CU分割成一或多个

PU。分割模式可在CU经帧内预测模式编码抑或经帧间预测模式编码之间有所不同。PU可分割成非正方形形状。与CU相关联的语法数据也可描述(例如)根据CTU将CU分割成一或多个TU。TU可为正方形或非正方形形状。

[0052] 根据HEVC标准,可使用变换单元(TU)来执行变换。TU可针对不同CU而变化。可基于给定CU内的PU的大小而对TU设定大小。TU可与PU大小相同或小于PU。在一些实例中,可使用被称为“残余四分树”(RQT)的四分树结构将对应于CU的残余样本再分成较小单元。RQT的叶节点可对应于TU。可变换与TU相关联的像素差值以产生变换系数。变换系数可接着由编码器引擎106量化。

[0053] 一旦视频数据的图片被分割成CU,编码器引擎106即使用预测模式来预测每一PU。接着从原始视频数据中减去预测以得到残余(下文描述)。对于每一CU,可使用语法数据在位流内部用信号传送预测模式。预测模式可包含帧内预测(或图片内预测)或帧间预测(或图片间预测)。使用帧内预测时,使用(例如)DC预测以发现PU的平均值、使用平坦预测以使平坦表面拟合于PU、使用方向预测以从相邻数据外插或使用任何其它合适类型的预测,从同一图片中的相邻图像数据预测每一PU。使用帧间预测时,使用运动补偿预测从一或多个参考图片(按输出次序在当前图片之前或之后)中的图像数据预测每一PU。可(例如)在CU层级作出使用图片间预测抑或图片内预测来译码图片区域的决策。在一些实例中,图片的一或多个图块被指派有图块类型。图块类型包含I图块、P图块和B图块。I图块(帧内,可独立地解码)为仅仅通过帧内预测译码的图片的图块,且因此可独立地解码,这是由于I图块仅仅需要帧内的数据以预测图块的任一块。P图块(单向预测帧)为可通过帧内预测和单向帧间预测译码的图片的图块。P图块内的每一块是通过帧内预测或帧间预测而译码。当帧间预测应用时,块仅仅通过一个参考图片而预测,且因此参考样本仅仅来自一个帧的一个参考区。B图块(双向预测性帧)为可通过帧内预测和帧间预测译码的图片的图块。B图块的块可从两个参考图片双向预测,其中每一图片贡献一个参考区且两个参考区的样品集经加权(例如,使用相等权重)以产生双向预测块的预测信号。如上文所解释,一个图片的图块被独立译码。在一些状况下,图片可仅作为一个图块而被译码。

[0054] PU可包含与预测过程有关的数据。举例来说,当使用帧内预测来编码PU时,PU可包含描述用于PU的帧内预测模式的数据。作为另一实例,当使用帧间预测来编码PU时,PU可包含定义PU的运动向量的数据。界定PU的运动向量的数据可描述(例如)运动向量的水平分量、运动向量的垂直分量、运动向量的分辨率(例如,四分之一像素精度或八分之一像素精度)、运动向量所指向的参考图片,和/或运动向量的参考图片列表(例如,列表0、列表1或列表C)。

[0055] 编码装置104接着可执行变换和量化。举例来说,在预测之后,编码器引擎106可计算对应于PU的残余值。残余值可包括像素差值。在预测执行之后可能剩余的任何残余数据是使用块变换进行变换,所述块变换可基于离散余弦变换、离散正弦变换、整数变换、小波变换或其它合适的变换功能。在一些状况下,一或多个块变换(例如,大小为 $32 \times 32$ 、 $16 \times 16$ 、 $8 \times 8$ 、 $4 \times 4$ 等)可应用于每一CU中的残余数据。在一些实施例中,TU可用于通过编码器引擎106实施的变换和量化过程。给定的具有一或多个PU的CU也可包含一或多个TU。如下文进一步详细描述,可使用块变换将残余值变换成变换系数,且接着可使用TU来量化和扫描残余值以产生用于熵译码的序列化变换系数。

[0056] 在一些实施例中,在使用CU的PU进行帧内预测性或帧间预测性译码之后,编码器引擎106可计算CU的TU的残余数据。PU可包括空间域(或像素域)中的像素数据。在应用块变换之后,TU可包括变换域中的系数。如先前所提及,残余数据可对应于未经编码图片的像素与对应于PU的预测值之间的像素差值。编码器引擎106可形成包含 CU的残余数据的TU,且接着可变换TU以产生CU的变换系数。

[0057] 编码器引擎106可执行变换系数的量化。量化通过量化变换系数以减少用以表示系数的数据的量而提供进一步压缩。举例来说,量化可减少与系数中的一些或所有相关联的位深度。在一个实例中,具有n位值的系数可在量化期间降值舍位为m位值,其中n 大于m。

[0058] 在执行量化后,经译码视频位流包含经量化变换系数、预测信息(例如,预测模式、运动向量等)、分割信息和任何其它合适的的数据(例如其它语法数据)。接着可通过编码器引擎106对经译码视频位流的不同元素进行熵编码。在一些实例中,编码器引擎106可利用预定义扫描次序扫描经量化变换系数以产生可经熵编码的序列化向量。在一些实例中,编码器引擎106可执行自适应扫描。在扫描经量化变换系数以形成向量(例如,一维向量)之后,编码器引擎106可熵编码向量。举例来说,编码器引擎106可使用上下文自适应性可变长度译码、上下文自适应性二进制算术译码、基于语法的上下文自适应性二进制算术译码、概率区间分割熵译码或另一合适的熵编码技术。

[0059] 编码装置104的输出端110可通过通信链路120将组成经编码视频位流数据的NAL单元发送到接收装置的解码装置112。解码装置112的输入端114可接收NAL单元。通信链路120可包含由无线网络、有线网络或有线网络与无线网络的组合提供的信道。无线网络可包含任何无线接口或无线接口的组合,且可包含任何适合的无线网络(例如,因特网或其它广域网、基于封包的网路、WiFi™、射频(RF)、UWB、WiFi-Direct、蜂窝式、长期演进(LTE)、WiMax™等)。有线网络可包含任何有线接口(例如,光纤、以太网、电力线以太网、经由同轴电缆的以太网、数字信号线(DSL)等)。可使用各种装备来实施有线和/或无线网络,所述装备例如基站、路由器、存取点、桥接器、网关、交换器等。可根据通信标准(例如,无线通信协议)调制经编码视频位流数据,且将其传输到接收装置。

[0060] 在一些实例中,编码装置104可将经编码视频位流数据存储在存储装置108中。输出端110可从编码器引擎106或从存储装置108检索经编码视频位流数据。存储装置108 可包含多种分散式或本地存取的数据存储媒体中的任一个。举例来说,存储装置108可包含硬盘机、存储光盘、闪存、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适的数字存储媒体。

[0061] 解码装置112的输入端114接收经编码视频位流数据,且可将视频位流数据提供到解码器引擎116,或提供到存储装置118以供稍后由解码器引擎116使用。解码器引擎 116可通过熵解码(例如,使用熵解码器)和提取组成经编码视频数据的一或多个经译码视频序列的元素来解码经编码视频位流。解码器引擎116可接着重新按比例调整经编码视频位流数据且对经编码视频位流数据执行反变换。残余数据接着传递到解码器引擎 116的预测级。解码器引擎116接着预测像素的块(例如,PU)。在一些实例中,预测被添加到反变换的输出(残余数据)。

[0062] 解码装置112可将经解码视频输出到视频目的地装置122,视频目的地装置可包含用于将经解码视频数据显示给内容的消费者的显示器或其它输出装置。在一些方面中,视

频目的地装置122可为包含解码装置112的接收装置的部分。在一些方面中,视频目的地装置122可为不同于接收装置的单独装置的部分。

[0063] 补充增强信息(SEI)消息可包含于视频位流中。举例来说,SEI消息可用于携带并非为由解码装置112解码位流所必需的信息(例如,元数据)。此信息用于改善经解码输出的显示或处理(例如,此类信息可由解码器侧实体使用以改善内容的可视性)。

[0064] 在一些实施例中,视频编码装置104和/或视频解码装置112可分别与音频编码装置和音频解码装置集成。视频编码装置104和/或视频解码装置112也可包含实施上文所描述的译码技术所必需的其它硬件或软件,例如一或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、离散逻辑、软件、硬件、固件或其任何组合。视频编码装置104和视频解码装置112可集成为相应装置中的组合式编码器/解码器(编码解码器)的部分。

[0065] 对HEVC标准的扩展包含多视图视频译码扩展(被称作MV-HEVC)和可调式视频译码扩展(被称作SHVC)。MV-HEVC和SHVC扩展共享分层译码的概念,其中不同层包含于经编码视频位流中。经译码视频序列中的每一层是由唯一层识别符(ID)定址。层ID可存在于NAL单元的标头中以识别NAL单元所相关联的层。在MV-HEVC中,不同层可表示视频位流中的同一场景的不同视图。在SHVC中,提供以不同空间分辨率(或图片分辨率)或不同重建保真度表示视频位流的不同可调式层。可调式层可包含基层(层ID=0)和一或多个增强层(层ID=1、2、...n)。基层可符合HEVC的第一版本的简档,且表示位流中的最低可用层。与基层相比,增强层具有增加的空间分辨率、时间分辨率或帧速率和/或重建保真度(或质量)。增强层经阶层式组织,且可(或可不)取决于较低层。在一些实例中,可使用单一标准编码解码器来译码不同层(例如,使用HEVC、SHVC或其它译码标准编码所有层)。在一些实例中,可使用多标准编码解码器来译码不同层。举例来说,可使用AVC来译码基层,而可使用对HEVC标准的SHVC和/或MV-HEVC扩展来译码一或多个增强层。大体来说,层包含一组VCL NAL单元和对应的一组非VCL NAL单元。NAL单元被指派有特定层ID值。在层可取决于较低层的意义上,层可为阶层式的。

[0066] 大体来说,层包含一组VCL NAL单元和对应的一组非VCL NAL单元。NAL单元被指派有特定层ID值。在层可取决于较低层的意义上,层可为阶层式的。层集合指表示于位流内的独立的层的集合,意味着层集合内的层在解码过程中可取决于层集合中的其它层,但并不取决于任何其它层来进行解码。因此,层集合中的层可形成可表示视频内容的独立位流。可通过子位流提取过程的操作从另一位流获得层集合中的层的集合。层集合可对应于待在解码器希望根据某些参数操作时被解码的层集合。

[0067] 图2说明分层HEVC位流200的简化实例。位流200可为MV-HEVC位流和/或SHVC位流。在所说明的实例中,位流包含基层202和两个增强层(层1 212和层2 214)。基层202可根据HEVC的特定版本(例如,低于当前版本)编码。替代地,基层202可使用另一编码解码器(例如,AVC)来编码。在这些和其它实例中,使用不同于HEVC的最近版本的编码解码器编码基层202可向不支持HEVC的最近版本或不支持HEVC的解码装置(例如,仅仅支持AVC的装置)提供反向兼容性。通过添加增强层,现有位流无需经重新编码用于另一编码解码器。

[0068] 在图2的实例中,第一增强层(层1 212)提供与基层202中提供的相同但在更高分辨率下的视频。在一些状况下,层1 212中的图片214可使用来自基层202中的图片204的数

据而编码。在一些状况下,层1 212中的图片214可经独立地编码。支持更高分辨率的解码装置因此可选择以再现基层202或层1 212。在其它实例中,层1 212可提供与基层202相同的图片214,但使用不同编码解码器。

[0069] 层2 222在所说明实例中经提供为增强层的另一实例。在此实例中,层2 222提供与层1 212中提供的视频相同但在较高帧速率下的视频。层2 222因此包含比层1 212中所包含的更多的图片224。在一些状况下,层2 222中的图片224可使用层1 212中的图片214和/或基层202中的图片204编码。在一些状况下,层2 222中的图片224可经独立地编码。在一些状况下,层2 222中的一些图片224可使用来自其它层的数据编码而其它图片222经独立地编码。

[0070] 在不同于图2中所说明的实例的实例中,位流200中的层可提供(例如)一个视频的不同视图、可频块化成一个图片的不同视频、在不同位速率、不同帧速率、不同分辨率等下编码的图片。

[0071] 在包含多层的位流中,存取单元可包含针对特定时间实例的所有图片。举例来说,参看图2中所说明的实例,一个存取单元可包含来自基层202的图片和增强层212、214的相应图片,每一相应图片对应与来自基层202的图片相同的时刻。在一些状况下,存取单元可具有来自仅仅一些层的图片。举例来说,由于层2 222中的图片224是以较高帧速率编码,因此来自层2 222的一些图片224可驻存于其自身存取单元中。在一些状况下,基层202中的图片204可并不具有增强层212、214中的相应图片,且因此包含这些图片204的存取单元可不包含任何其它图片。

[0072] 如先前所描述,HEVC位流包含一组NAL单元,NAL单元包含VCL NAL单元和非VCL NAL单元。除了其它信息以外,非VCL NAL单元也可含有具有与经编码视频位流相关的高层级信息的参数集。举例来说,参数集可包含视频参数集(VPS)、序列参数集(SPS)和图片参数集(PPS)。参数集的目标的实例包含位速率效率、错误复原(error resiliency)和提供系统层接口。每一图块参考单个作用中PPS、SPS和VPS以存取解码装置112可用于解码图块的信息。可针对每一参数集译码识别符(ID),包含VPS ID、SPS ID和PPS ID。SPS包含SPS ID和VPS ID。PPS包含PPS ID和SPS ID。每一图块标头包含PPS ID。使用ID,可识别针对给定图块的作用中参数集。

[0073] VCL NAL单元包含形成经译码视频位流的经译码图片数据。各种类型的VCL NAL单元经定义于HEVC标准中。在单层位流中,如第一HEVC标准中所定义,含于AU中的VCL NAL单元具有相同NAL单元类型值,其中NAL单元类型值定义AU的类型和AU内的经译码图片的类型。举例来说,特定AU的VCL NAL单元可包含瞬时解码再新(IDR)NAL单元(值19),使得AU为IDR AU且AU的经译码图片为IDR图片。给定类型的VCL NAL单元是关于含于VCL NAL单元(例如,VCL NAL单元中的图片的图块或图块片段)中的图片,或其部分。三类图片经定义于HEVC标准中,包含前置图片、后置图片和帧内随机存取(IRAP)图片(也称作“随机存取图片”)。在多层位流中,AU内的图片的VCL NAL单元具有相同NAL单元类型值和相同类型经译码图片。举例来说,含有类型IDR的VCL NAL单元的图片称为AU中的IDR图片。在另一实例中,当AU含有为基层(层ID等于0)处的IRAP图片的图片时,AU为IRAP AU。

[0074] HEVC位流可进一步包含位流结尾(EOB)NAL单元。EOB NAL单元指示已达到位流结尾。编码器大体产生EOB NAL单元,且对于HEVC位流,EOB NAL单元为位流的最后AU中的最后



NAL单元。

[0075] 如上文所论述编码的视频位流可写入或打包到一或多个文件中以便将位流从编码装置104转移到解码装置112。举例来说,输出端110可包含文件写入引擎,其经配置以产生含有位流的一或多个文件。输出端110可通过通信链路120传输一或多个文件到解码装置112。替代地或另外,一或多个文件可存储于存储媒体(例如,磁带、磁盘或硬盘机,或一些其它媒体)上以供稍后传输到解码装置112。

[0076] 解码装置112可包含(例如)在输入端114中的文件剖析引擎。文件剖析引擎可读取通过通信链路120或从存储媒体接收的文件。文件剖析引擎可进一步从文件提取样本,并重建位流以用于由解码器引擎116解码。在一些状况下,经重建位流可与由编码器引擎106产生的位流相同。在一些状况下,编码器引擎106可能已产生具有用于解码位流的若干可能选项的位流,在所述状况下,经重建位流可包含所有可能选项中的仅一者或少于所有可能选项。

[0077] 如上文所论述编码的位流可使用ISOBMFF、从ISOBMFF衍生的文件格式、某一其它文件格式和/或包含ISOBMFF的文件格式的组合写入到文件。如上文所提及,当写入到文件时,位流可经存储和/或传输到解码装置,解码装置可再现并显示位流的内容。

[0078] 图3说明含有根据ISOBMFF格式化的用于视频呈现的数据和元数据的ISO基本媒体文件300的实例。ISOBMFF经设计而以灵活并可扩展格式含有计时媒体信息,所述计时媒体信息促进媒体的互换、管理、编辑和呈现。媒体的呈现可在含有呈现的系统“本地”,或呈现可经由网络或其它流式传输传递机构。

[0079] 如ISOBMFF规范中所定义的“呈现”为图片序列,其常常由于由视频捕获装置依序捕获而相关或由于某其它原因而相关。本文中,呈现也可被称作电影或视频呈现。呈现可包含音频。单一呈现可含于一或多个文件中,其中一个文件含有用于整个呈现的元数据。元数据包含信息,例如时序和成框数据、描述符、指标、参数,和描述呈现的其它信息。元数据自身并不包含视频和/或音频数据。除含有元数据的文件外的文件不需要根据ISOBMFF而格式化,且仅需要经格式化使得这些文件可由元数据参考。

[0080] ISO基本媒体文件的文件结构为物件定向式结构,且文件中的个别物件的结构可直接从物件的类型来推断。ISOBMFF规范将ISO基本媒体文件中的物件规范称作“框”。ISO基本媒体文件构建为可含有其它框的一系列框。框大体包含提供框的大小和类型的标头。大小描述框的整个大小,包含标头、字段和含有于帧内的所有框。播放器装置不能识别的类型的框通常被忽略并跳过。

[0081] 如由图3的实例所说明,在文件的顶部层级处,ISO基本媒体文件300可包含文件类型框310、电影框320,和一或多个电影片段330a、330n。在此层级可被包含但在此实例中并未表示的其它框包含自由空间框、元数据框和媒体数据框外加其它。

[0082] ISO基本媒体文件可包含通过框类型“ftyp”识别的文件类型框310。文件类型框310识别最适合于剖析文件的ISOBMFF规范。“最”在此情况下意味着,ISO基本媒体文件300可能已根据特定ISOBMFF规范格式化,但很可能与规范的其它反复兼容。此最合适规范被称作主流品牌。播放器装置可使用主流品牌来确定装置是否能够解码并显示文件的内容。文件类型框310也可包含版本号,所述版本号可用以指示ISOBMFF规范的版本。文件类型框310也可包含兼容品牌的列表,所述列表包含文件兼容的其它品牌的列表。ISO基本媒体文件可

与多于一个主流品牌兼容。

[0083] 当ISO基本媒体文件300包含文件类型框310时,存在仅一个文件类型框。ISO基本媒体文件300可省略文件类型框310以便与其它早期播放器装置兼容。当ISO基本媒体文件300不包含文件类型框310时,播放器装置可假定默认主流品牌(例如“mp41”)、次要版本(例如,“0”)和兼容的品牌(例如,“mp41”)。文件类型框310通常尽可能早地放置于ISO基本媒体文件300中。

[0084] ISO基本媒体文件可进一步包含电影框320,其可含有元数据以供呈现。电影框320通过框类型“moov”识别。ISO/IEC 14496-12规定,含于一个文件或多个文件中的呈现可包含仅一个电影框320。经常地,电影框320靠近ISO基本媒体文件的开始。电影框 320包含电影标头框322,且可包含一或多个播放轨框324以及其它框。

[0085] 通过框类型“mvhd”识别的电影标头框322可包含与媒体无关且与作为整体的呈现相关的信息。举例来说,电影标头框322可包含信息,例如产生时间、修改时间、时间标度和/或针对呈现的持续时间外加其它。电影标头框322也可包含识别呈现中之下一播放轨的识别符。举例来说,在所说明的实例中,识别符可指向由电影框320含有的播放轨框324。

[0086] 通过框类型“trak”识别的播放轨框324可含有用于呈现的播放轨的信息。呈现可包含一或多个播放轨,其中每一播放轨独立于呈现中的其它播放轨。每一播放轨可包含对于播放轨中的内容特定的时间和空间信息,且每一播放轨可与媒体框相关联。播放轨中的数据可为媒体数据,在此状况下,播放轨为媒体播放轨;或数据可为用于流式传输协议的打包化信息,在此状况下,播放轨为提示播放轨。媒体数据包含(例如)视频和音频数据。在所说明的实例中,实例播放轨框324包含播放轨标头框324a和媒体框324b。播放轨框可包含其它框,例如播放轨参考框、播放轨群组框、编辑框、用户数据框、元框和其它框。

[0087] 由框类型“tkhd”识别的播放轨标头框324a可指定含于播放轨框324中的播放轨的特性。举例来说,播放轨标头框324a可包含播放轨的产生时间、修改时间、持续时间、播放轨识别符、层识别符、群组识别符、音量、宽度和/或高度,外加其它。对于媒体播放轨,播放轨标头框324a可进一步识别播放轨是否经启用,播放轨是否应作为呈现的部分而播放,或播放轨是否可用以预览呈现,外加其它。播放轨的呈现通常假定为处于呈现的开始处。播放轨框324可包含此处未说明的编辑列表框,所述编辑列表框可包含显式时刻表图。时刻表图可指定播放轨的偏移时间外加其它,其中偏移指示在呈现开始之后播放轨的开始时间。

[0088] 在所说明的实例中,播放轨框324也包含由框类型“mdia”识别的媒体框324b。媒体框324b可含有物件和关于播放轨中的媒体数据的信息。举例来说,媒体框324b可含有处理例程参考框,其可识别播放轨的媒体类型以及藉以呈现播放轨中的媒体的过程。作为另一实例,媒体框324b可含有媒体信息框,其可指定播放轨中媒体的特性。媒体信息框可进一步包含样本表,其中每一样本描述包含(例如)样本的数据的位置的媒体数据(例如,视频或音频数据)的块。样本的数据存储于下文进一步论述的媒体数据框中。如同大部分其它框一般,媒体框324b也可包含媒体标头框。

[0089] 在所说明的实例中,实例ISO基本媒体文件300也包含呈现的多个片段330a、330b、330c、330n。片段330a、330b、330c、330n并非为ISOBMFF框,而是描述电影片段框 332和通过电影片段框332参考的媒体数据框338。电影片段框332和媒体数据框338 为顶部层级框,但此处分组在一起以指示电影片段框332与媒体数据框338之间的关系。

[0090] 由框类型“moof”识别的电影片段框332可通过包含将以其它方式存储于电影框320中的额外信息而扩展呈现。使用电影片段框332,呈现可经递增地建置。电影片段框332可包含电影片段标头框334和播放轨片段框336以及此处未说明的其它框。

[0091] 由框类型“mfhd”识别的电影片段标头框334可包含序号。播放器装置可使用序号来验证片段330a包含数据的下一段以供呈现。在一些状况下,文件的内容或用于呈现的文件可无序地被提供到播放器装置。举例来说,网络封包可频繁地以不同于最初传输封包的次序的次序到达。在这些状况下,序号可辅助播放器装置确定片段的正确次序。

[0092] 电影片段框332也可包含由框类型“traf”识别的一或多个播放轨片段框336。电影片段框332可包含一组播放轨片段(每播放轨零或多个)。播放轨片段可含有零或多个播放轨行程,其中的每一个描述播放轨的样本的连续行程。除添加样本到播放轨外,播放轨片段可用以将空的时间添加到播放轨。

[0093] 由框类型“mdat”识别的媒体数据框338含有媒体数据。在视频播放轨中,媒体数据框338将含有视频帧。媒体数据框可替代地或另外包含音频数据。呈现可包含含于一或多个个别文件中的零或多个媒体数据框。媒体数据由元数据描述。在所说明的实例中,媒体数据框338中的媒体数据可由包含于播放轨片段框336中的元数据来描述。在其它实例中,媒体数据框中的媒体数据可由电影框320中的元数据来描述。元数据可通过在文件300内的绝对偏移参考特定媒体数据,以使得媒体数据标头和/或媒体数据框338 内的自由空间可被跳过。

[0094] ISO基本媒体文件300中的其它片段330b、330c、330n可含有类似于针对第一片段330a说明的框的框,和/或可含有其它框。

[0095] 除支持媒体的本地播放外,ISOBMFF包含对通过网络流式传输媒体数据的支持。包含一个电影呈现的文件可包含称为提示播放轨的额外播放轨,其含有可辅助流式传输服务器形成并传输文件作为封包的指令。举例来说,这些指令可包含供服务器发送的数据(例如,标头信息)或对媒体数据的片段的参考。文件可包含用于不同流式传输协议的单独提示播放轨。提示播放轨也可在不需重新格式化文件的情况下添加到文件。

[0096] 用于对媒体数据进行流式传输的一种方法为经由超文字传送协议(HTTP)或DASH(定义于ISO/IEC 33009-1:2014)的动态自适应流式传输。被称为MPEG-DASH的DASH 为允许实现使用常规HTTP网页服务器的高质量媒体内容流式传输的自适应位速率流式传输技术。DASH通过将媒体内容打断成一系列小的基于HTTP的文件片段来操作,其中每一片段含有短时间间隔的内容。使用DASH,服务器可以不同位速率提供媒体内容。正播放媒体的客户端装置可在下载下一片段时在替代性位速率中做出选择,且因此适应于改变的网络条件。DASH使用因特网的HTTP网页服务器基础架构来经由全球信息网递送内容。DASH独立于用以编码并解码媒体内容的编码解码器,且因此与例如H.264 和HEVC的编码解码器外加其它一起操作。

[0097] ISOBMFF规范指定适用于DASH的六种类型流式传输存取点(SAP)。前两种SAP 类型(类型1和2)对应于H.264/AVC和HEVC中的瞬时解码再新(IDR)图片。举例来说, IDR图片为图片内(I图片)图片,其在解码器处完全再新或重新初始化解码程序且开始新的经译码视频序列。在一些实例中, IDR图片和按解码次序在IDR图片之后的任何图片不可取决于按解码次序在IDR图片之前出现的任何图片。

[0098] 第三SAP类型(类型3)对应于开放GOP(图片群组)随机存取点,因此对应于HEVC 中的断链存取(BLA)或清洁随机存取(CRA)图片。举例来说,CRA图片也为I图片。CRA 图片可能并不再新解码器,且可能不开始新CVS,从而允许CRA图片的前置图片取决于按解码次序在CRA图片之前出现的图片。随机存取可通过解码以下各者在CRA图片处进行:CRA图片、与并不取决于按解码次序在CRA图片之前出现的任何图片的CRA 图片相关联的前置图片,和按解码次序和输出次序两者在CRA之后的所有关联图片。在一些状况下,CRA图片可能并不具有关联前置图片。在一些实施例中,在多层状况下,属于具有大于0的层ID的层的IDR或CRA图片可为P图片或B图片,但这些图片可仅使用来自属于与IDR或CRA图片相同的存取单元且具有小于含有IDR或CRA图片的层的层ID的其它图片的层间预测。

[0099] 第四种SAP类型(类型4)对应于渐次性解码再新(GDR)随机存取点。

[0100] 图4说明用于视频呈现的包含两层的ISO基本媒体文件400的实例。如上文所论述,视频呈现可包含多层,包含基层和一或多个增强层。在图4的实例中,视频包含层0中的基层和层1中的增强层。在其它实例中,视频可包含多于一个增强层。

[0101] 在文件400的顶部层级处,ISO基本媒体文件400包含文件类型框410、电影框420和两个媒体数据框430a、430b,针对层0和层1的样本各有一个媒体数据框。文件类型框410可识别最适合于读取文件400的品牌(例如,ISOBMFF规范的反复或衍生)。文件类型框410也可识别兼容品牌(即,适合于读取文件400的其它品牌)的列表。

[0102] 电影框420可包含用于视频呈现的元数据。电影框420可包含电影标头框422,其可含有关于作为整体的呈现的信息。电影框420也可包含一或多个播放轨框。在所说明的实例中,电影框420包含两个播放轨框424a、424b,针对层0的元数据和针对层1的元数据各有一个播放轨框。在其它实例中,层0和层1两者的元数据可包含于一个播放轨框中。在其中呈现包含多于两个层的实例中,所有层可包含于一个播放轨中或一些播放轨可具有仅仅一个层而其它播放轨具有两个或多于两个层。

[0103] 播放轨框424a、424b中的每一个包含播放轨标头框426a、426b。播放轨标头框426a、426b可描述由相应播放轨框424a、424b参考的播放轨的特性。举例来说,播放轨标头框426a、426b可包含用于播放轨中的图片的宽度、高度和层ID外加其它。

[0104] 播放轨框424a、424b中的每一个可各自进一步包含媒体数据框428a、428b。媒体数据框428a、428b包含含于每一相应播放轨中的层中的样本的描述。对于样本的描述可包含(例如)样本在文件400内(或在另一文件内)的位置、在时间上共置样本的描述,和/或参考样本的描述外加其它。

[0105] 媒体数据框430a、430b含有两个层的样本。样本含有含于ISO基本媒体文件400 中的媒体的数据。在所说明的实例中,样本可含有视频数据和音频数据。在此实例中,层0的样本定位于一个媒体数据框430a中且层1的样本定位于第二媒体数据框430b中。在其它实例中,两个层的样本可包含于相同媒体数据框中。在其它实例中,一个层的样本可包含于多个媒体数据框中。在其它实例中,样本可含于其它文件中。

[0106] 如上文所论述,HEVC位流可包含作为位流的最后存取单元的最后NAL单元的位流结尾(EOB)NAL单元。EOB NAL单元可向解码器用信号传送位流结尾已达到。当位流为多层位流时,编码器将使EOB NAL单元正确地放置于位流中的最后存取单元中。然而,当位流随后写入到文件(例如,存储到文件或存储于文件内时),可产生与EOB NAL 单元的位置有关的

问题。

[0107] 图5A说明其中已存储多层位流的ISO基本媒体文件500的一个实例。在所说明的实例中,位流包含层0中的基层和层1中的增强层。实例文件500进一步包含文件类型框510,其可指定ISOBMFF的品牌或特定反复或ISOBMFF的文件500与的兼容的衍生。文件500也可包含电影框520,其可含有用于位流的元数据。实例文件500也包含媒体数据框530a、530b,其可含有位流的样本。

[0108] 电影框520在此实例中包含电影标头框522和播放轨0和播放轨1的两个播放轨框524a、524b。电影标头框522可包含描述作为整体的呈现的信息。在此实例中,层0在播放轨0中;因此第一播放轨框524a包含用于层0的元数据。另外,层1在播放轨1中;因此第二播放轨框524b包含用于层1的元数据。在其它实例中,用于两个层的元数据可在一个播放轨框中。在其它实例中,位流可包含多于两个层。在这些实例中,一些播放轨可包含一个层和/或一些播放轨可包含两个或多于两个层。

[0109] 在此实例中,每一播放轨框524a、524b包含播放轨标头框526a、526b和媒体框528a、528b。播放轨标头框526a、526b可描述播放轨,而媒体框528a、528b可包含播放轨中的样本的描述。

[0110] 如上文所论述,存储于文件500中的位流可具有EOB NAL单元540。在所说明的实例中,编码器将EOB NAL单元540放置在增强层(层1)中的数据的结尾处。在此实例中,因此可在播放轨1中发现EOB NAL单元540。

[0111] 在一些状况下,EOB NAL单元540在播放轨1中的放置对于解码装置的文件剖析引擎可能有问题。图5B和图5C说明当图5A中所说明的ISO基本媒体文件500由不同解码器装置读取时可产生的位流550、552的实例。在图5B中,解码装置仅仅读取播放轨0。从文件500提取的位流550因此包含各自包含来自基层的基本图片564a、564b、564c、564n的一系列存取单元562a、562b、562c、562n。然而,在图5A的实例中,EOB NAL单元540在播放轨1中;因此,EOB NAL单元540不包含于经重建位流550中。

[0112] 在图5B的实例中,解码装置(例如)可能已仅仅读取播放轨0,这是由于解码装置经配置以仅仅播放基层(例如,增强层未经选择用于显示)。替代地或另外,解码装置可仅已能够播放基层(例如,解码装置仅仅为与AVC兼容,且基层包含以AVC编码的位流)。解码装置可出于其它原因以其它方式仅仅读取播放轨0。

[0113] 在图5C的实例中,解码装置从文件500提取基层和增强层两者。所得位流552因此包含各自包含来自基层的基本图片564a、564b、564c、564n和来自增强层的增强型图片566a、566b、566c、566n的一系列存取单元562a、562b、562c、562n。在此实例中,由于EOB NAL单元540在增强层中,因此当从文件提取增强层时,也提取EOB NAL单元540,且被正确地放置在最后存取单元562n中。

[0114] 在图5C的实例中,解码装置可能已读取播放轨0和播放轨1两者,这是由于(例如)增强层经选择用于显示,且增强层需要来自基层的数据。解码装置可能出于其它原因以其它方式读取两个播放轨。

[0115] 图6A说明其中已存储多层位流的ISO基本媒体文件600的另一实例。在所说明的实例中,位流包含层0中的基层和层1中的增强层。实例文件600进一步包含文件类型框610,其可指定ISOBMFF的品牌或特定反复或ISOBMFF的文件600与的兼容的衍生。文件600也可包含

电影框620,其可含有用于位流的元数据。实例文件600也包含媒体数据框630a、630b,其可含有位流的样本。

[0116] 电影框620在此实例中包含电影标头框622和播放轨0和播放轨1的两个播放轨框624a、624b。电影标头框622可包含描述作为整体的呈现的信息。在此实例中,层0在播放轨0中,且因此第一播放轨框624a包含用于层0的元数据。另外,层1在播放轨1中,因此第二播放轨框624b包含用于层1的元数据。在其它实例中,用于两个层的元数据可在一个播放轨框中。在其它实例中,位流可包含多于两个层。在这些实例中,一些播放轨可包含一个层和/或一些播放轨可包含两个或多于两个层。

[0117] 在此实例中,每一播放轨框624a、624b包含播放轨标头框626a、626b和媒体框628a、628b。播放轨标头框626a、626b可描述播放轨,而媒体框628a、628b可包含播放轨中的样本的描述。

[0118] 如上文所论述,存储于文件600中的位流可具有EOB NAL单元640。在所说明的实例中,编码器装置已将EOB NAL单元640放置在基层中,且因此可在播放轨0中发现EOB NAL单元640。在各种实施方案中,可由于编码约束条件需要编码器装置将EOB NAL单元640放置在基层中。在一些状况下,编码器可出于另一原因而已将EOB NAL单元640放置在基层中。举例来说,位流中的最后存取单元可将仅仅一个图片包含于基层中。

[0119] 当EOB NAL单元640在基层中时,可产生针对解码器装置的问题。图6B和图6C说明可从读取图6A中所说明的ISO基本媒体文件600产生的位流650、652的实例。在图6B中,解码装置仅仅需要位流的基层,且因此指示文件剖析引擎仅仅读取播放轨0。所得位流650因此包含各自包含从播放轨0读取的基本图片664a、664b、664c、664n的一系列存取单元662a、662b、662c、662n。在此实例中,基层包含EOB NAL单元640;因此,图6B中所说明的位流650将EOB NAL单元640包含于最后存取单元662n中。

[0120] 在图6C的实例中,解码装置的文件剖析引擎从文件600读取基层和增强层两者。所得位流652因此包含各自包含从播放轨0读取的基本图片664a、664b、664c、664n和从播放轨1读取的增强型图片666a、666b、666c、666n的一系列存取单元662a、662b、662c、662n。在此实例中,每一存取单元首先以来自基层的NAL单元填充且接着以来自增强层的NAL单元填充。由于EOB NAL单元640在实例文件600中的基层中,因此在实例位流652中,EOB NAL单元640并非为最后存取单元中的最后NAL单元。位流652因此可不被正确地解码。

[0121] 图7说明其中已存储多层位流的ISO基本媒体文件700的另一实例。在所说明的实例中,位流包含层0中的AVC基层和层1中的包含AVC基层的HEVC版本的增强层。基层可已使用AVC编码,这是由于位流为未经重新编码用于HEVC的旧版位流或由于位流被制作成与具有早期解码器引擎的解码器装置兼容,或出于某一其它原因。可进一步提供增强层,以使得可通过支持HEVC的解码器装置检视相同位流。在其它实例中,可已使用不同于AVC的编码解码器编码基层,使得基层包含未使用HEVC编码解码器编码的样本。

[0122] 实例文件700包含文件类型框710,其可指定ISOBMFF的品牌或特定反复或ISOBMFF的文件700与的兼容的衍生。文件700也可包含电影框720,其可含有用于位流的元数据。实例文件700也包含媒体数据框730a、730b,其可含有位流的样本。

[0123] 电影框720在此实例中包含电影标头框722和播放轨0和播放轨1的两个播放轨框724a、724b。电影标头框722可包含描述作为整体的呈现的信息。在此实例中,层0在播放轨0

中,且因此第一播放轨框724a包含用于层0的元数据。另外,层1在播放轨1 中;因此第二播放轨框724b包含用于层1的元数据。在其它实例中,用于两个层的元数据可在一个播放轨框中。在其它实例中,位流可包含多于两个层。在这些实例中,一些播放轨可包含一个层和/或一些播放轨可包含两个或多于两个层。

[0124] 在此实例中,每一播放轨框724a、724b包含播放轨标头框726a、762b和媒体框728a、728b。播放轨标头框726a、762b可描述播放轨,而媒体框728a、728b可包含播放轨中的样本的描述。举例来说,播放轨0的框724a中的播放轨标头框726a和/或媒体框728a 可指示播放轨0的样本是使用AVC编码。类似地,播放轨1的框724b中的播放轨标头框726b和/或媒体框728b可指示播放轨1的样本是使用HEVC而编码。

[0125] 如上文所论述,存储于文件700中的位流可具有EOB NAL单元740。在所说明的实例中,编码器装置已将EOB NAL单元740放置在基层中,以使得可在播放轨0中发现EOB NAL单元740。在各种实施方案中,可由于编码约束条件需要编码器装置将EOB NAL单元740放置在基层中。在一些状况下,编码器可出于另一原因而已将EOB NAL 单元740放置在基层中。举例来说,位流中的最后存取单元可将仅仅一个图片包含于基层中。在这些和其它实例中,将根据AVC编码解码器产生EOB NAL单元740。

[0126] 在此实例中,可产生针对解码器装置的问题。举例来说,当解码器装置为与HEVC兼容时,解码器装置可指示文件剖析引擎从文件700读取基层和增强层两者。然而,EOB NAL单元740由于在基层中将为使用AVC产生的EOB NAL单元。因此,即使从文件 700读取并放置于所得位流结尾处,所得位流可并不完全与HEVC兼容,且可并不由 HEVC解码器装置正确地处理。

[0127] 在各种实施方案中,各种技术可用以克服上文关于EOB NAL单元所论述的问题。图8A说明根据本发明的技术产生的ISO基本媒体文件800的实例。在所说明的实例中,位流包含层0中的基层和层1中的增强层。实例文件800进一步包含文件类型框810,其可指定ISOBMFF的品牌或特定反复或ISOBMFF的文件800与的兼容的衍生。文件 800也可包含电影框820,其可含有用于位流的元数据。实例文件800也包含媒体数据框830a、830b,其可含有位流的样本。

[0128] 电影框820在此实例中包含电影标头框822和播放轨0和播放轨1的两个播放轨框824a、824b。电影标头框822可包含描述作为整体的呈现的信息。在此实例中,层0在播放轨0中,且因此第一播放轨框824a包含用于层0的元数据。另外,层1在播放轨1 中,因此第二播放轨框824b包含用于层1的元数据。在其它实例中,用于两个层的元数据可在一个播放轨框中。在其它实例中,位流可包含多于两个层。在这些实例中,一些播放轨可包含一个层和/或一些播放轨可包含两个或多于两个层。

[0129] 在此实例中,每一播放轨框824a、824b包含播放轨标头框826a、862b和媒体框828a、828b。播放轨标头框826a、862b可描述播放轨,而媒体框828a、828b可包含播放轨中的样本的描述。

[0130] 在本发明的各种实施方案中,用于克服与EOB NAL单元在文件800中的放置相关的问题的技术包含放宽EOB NAL单元在所述文件中的存在,以使得EOB NAL单元840a、840b可存在于一个或多个播放轨中。如图8中所说明,播放轨0因此可或可不包含 EOB NAL单元840a。类似地,播放轨1也可或可不包含EOB NAL单元840b。在一些状况下,EOB NAL单元840a 840b

可为类似的。在一些状况下,播放轨0中的EOB NAL 单元840a可不同于播放轨1中的EOB NAL 单元840b。举例来说,可使用不同于HEVC 的编码解码器编码播放轨0中的EOB NAL单元,而使用HEVC编码解码器编码播放轨 1中的EOB NAL单元。

[0131] 图8B和图8C说明可从图8A中所说明的文件800构建的位流850、852的实例。在图8B和图8C两者中,解码器装置仅仅需要基层,且因此指示文件剖析引擎仅仅读取播放轨0。每一实例位流850、852因此包含各自包含来自基层的基本图片864a、864b、864c、864n的一系列存取单元862a、862b、862c、862n。

[0132] 在图8B的实例中,播放轨0包含基层的EOB NAL单元840a,且播放轨1并不包含增强层的EOB NAL单元840b。在此实例中,解码器装置的文件剖析引擎可确定基层为包含EOB NAL单元的最高层,意味着无其它与比另一层(在此实例中,基层)更大(例如,“更高”)的层ID值相关联的“较高”层(例如,增强层)包含EOB NAL单元。文件剖析引擎因此可将来自基层的EOB NAL单元840a包含于重建位流850的最后存取单元 862n中。所得位流850因此可在没有任何与EOB NAL单元840a相关的错误的情况下被解码。

[0133] 图8B中所说明的技术可确保当仅仅从文件800读取播放轨0时,EOB NAL单元 840a包含于经重建位流650中。在此实例中,尽管考虑其它层,但文件剖析引擎可确定来自基层的EOB NAL单元840a适于在位流650中使用。

[0134] 在图8C的实例中,有可能播放轨0并不包含EOB NAL单元,且播放轨1中的增强层包含EOB NAL单元840b。在此实例中,当解码器装置的文件剖析引擎读取文件800 时,剖析引擎可经配置以检测到基层不包含EOB NAL单元。文件剖析引擎可经进一步配置以读取(例如,处理)最高层的播放轨(在此实例中,播放轨1)以确定播放轨中的层是否包含EOB NAL单元。在各种实施方案中,当最高层播放轨不包含EOB NAL单元时,文件剖析引擎可尝试用于下一最高层的播放轨,等等,直到EOB NAL单元经定位或文件剖析引擎确定文件800不包含EOB NAL单元为止。在所说明的实例中,发现播放轨 1中的EOB NAL单元时,文件剖析引擎可将此EOB NAL单元840b放置(例如,包含) 于重建位流852的最后存取单元862n中(例如,内)。所得位流852因此可在没有任何与 EOB NAL单元相关的错误的情况下被解码。

[0135] 替代地,在图8C的实例中,有可能播放轨0和播放轨1两者包含EOB NAL单元840a、840b。在此状况下,在一些实施方案中,文件剖析引擎可经配置以定位最高层(例如,在此实例中,增强层)中的EOB NAL单元840b,并将此EOB NAL单元840b放置于经重建位流852的最后存取单元862n中。可丢弃任何其它EOB NAL单元(包含来自基层的EOB NAL单元840a)。

[0136] 图8C中所说明的技术可避免可在解码器装置仅仅需要基层时产生的问题,基层不包含EOB NAL单元或包含与解码器装置的解码器引擎可能不兼容的EOB NAL单元。举例来说,当实例文件800不包含基层中的EOB NAL单元时,可实际上提供来自较高层的EOB NAL单元。作为另一实例,当使用AVC编码基层中的EOB NAL单元但解码器装置包含HEVC解码器引擎时,来自较高层的HEVC EOB NAL单元可取代基层中的 EOB NAL单元。

[0137] 图8D和图8E说明可从图8A中所说明的文件800构建的位流854、856的额外实例。在图8D和图8E两者中,解码器装置需要基层和增强层两者,且因此指示文件剖析引擎读取播放轨0和播放轨1两者。每一实例位流850、852因此包含各自包含来自基层的基本图片864a、864b、864c、864n和来自增强层的增强型图片868a、868b、868c、868n的一系列存取单元862a、862b、862c、862n。



[0138] 在图8D的实例中,基层和增强层两者包含EOB NAL单元840a、840b。在此实例中,在各种实施方案中,解码器装置的文件剖析引擎可经配置以包含仅仅来自最高层(在此实例中,增强层)的EOB NAL单元840b并丢弃任何其它EOB NAL单元(例如,来自基层的EOB NAL单元840a)。经重建文件854a因此仅仅包含来自增强层(例如,最高层)的EOB NAL单元840b。

[0139] 图8D中所说明的技术可避免产生包含多于一个EOB NAL单元或包含仅仅来自基层的EOB NAL单元的位流。举例来说,当文件将以其它方式仅仅具有基层中的EOB NAL单元时,允许增强层中的EOB NAL单元,以使得包含两个层的经重建位流不仅仅具有不恰当地定位于最后存取单元的中间的来自基层的EOB NAL单元。解码器装置的文件剖析引擎可进一步确保位流不具有多于一个EOB NAL单元。

[0140] 在图8E的实例中,基层包含EOB NAL单元840a,但增强层不包含EOB NAL单元。在此实例中,在各种实施方案中,解码器装置的文件剖析引擎识别在最高层(在此实例中,基层)中的EOB NAL单元,并重新排序从文件读取的NAL单元,以使得来自最高层的EOB NAL单元移动到最后存取单元862n的结尾。可丢弃任何其它EOB NAL单元。

[0141] 图8E中所说明的技术可避免产生具有不在最后存取单元的结尾处的EOB NAL单元的位流。举例来说,当基层包含EOB NAL单元但增强层不包含EOB NAL单元时,那么在不使用所说明技术的情况下,EOB NAL单元将放置于最后存取单元862n的中间某处,而非结尾处。

[0142] 上文所论述的各种技术因此可解决可在位流写入到(例如,存储)文件时关于EOB NAL单元的放置产生的问题。更一般来说,以上技术规定,当EOB NAL单元是图片的并非来自最高层的部分时,EOB NAL单元应在重建位流期间放置于最后存取单元的结尾处,除非在最高层中有作为图片的部分存在的EOB NAL单元。在后一状况中,来自最高层的EOB NAL单元应包含于最后存取单元中。替代地或另外,当多于一个EOB NAL单元存在于存取单元中时,应使用为具有最高层ID值的图片的部分的EOB NAL单元,且应丢弃任何其它EOB NAL单元。

[0143] 图9说明可包含于ISO基本媒体文件中的媒体框940和媒体数据框938a、938b的实例。如上文所论述,媒体框可包含于播放轨框中,且可含有描述播放轨中的媒体数据的物件和信息。在所说明的实例中,媒体框940包含媒体信息框942。媒体框940也可包含此处未说明的其它框。

[0144] 媒体信息框942可含有描述关于播放轨中的媒体的特性信息的物件。举例来说,媒体信息框942可包含描述播放轨中的媒体信息的位置的数据信息框。作为另一实例,当播放轨包含视频数据时,媒体信息框942可包含视频媒体标头。视频媒体标头可含有独立于视频媒体的译码的大体呈现信息。当播放轨包含音频数据时,媒体信息框942也可包含声音媒体标头。

[0145] 媒体信息框942也可包含样本表框944,如所说明实例中所提供。由框类型“stbl”识别的样本表框944可提供播放轨中的媒体样本的位置(例如,具有文件的位置),以及样本的时间信息。使用由样本表框944提供的信息,播放器装置可以正确时间次序定位样本,确定样本的类型,和/或确定大小、容器和容器内的样本的偏移外加其它。

[0146] 样本表框944可包含由框类型“stsd”识别的样本描述框946。样本描述框946可提供关于(例如)用于样本的译码类型的详细信息,和针对彼译码类型需要的任何初始化信息。存储于样本描述框中的信息对于包含样本的播放轨的类型可为特定的。举例来说,一个格式在播放轨为视频播放轨时可用于样本描述,且在播放轨为提示播放轨时可使用不同格

式。作为另一实例,样本描述的格式也可取决于提示播放轨的格式而发生变化。

[0147] 样本描述框946可包含一或多个样本项948a、948b、948c。样本项类型为摘要类别,且因此通常样本描述框946包含特定样本项,例如视频数据的可见样本项或音频样本的音频样本项外加其它实例。样本项框可存储针对特定样本的参数。举例来说,对于视频样本,样本项框可包含视频样本的宽度、高度、水平分辨率、垂直分辨率、帧计数和/或深度外加其它。作为另一实例,对于音频样本,样本项可包含信道计数、信道布局和/或采样速率外加其它。

[0148] 由样本项948a、948b、948c描述的样本970a、970b、970c存储于媒体数据框938a、938b中。媒体数据框938a、938b可包含于文件的顶部层级中。样本描述框946中参考的样本可包含于不同媒体数据框938a、938b中,或可包含于一个媒体数据框中。媒体数据框938a、938b可在与媒体框940相同的文件中和/或可在其它文件中。

[0149] 除上文所论述的问题以外,产生在将频块播放轨写入到根据ISOBMFF或从ISOBMFF衍生的格式而格式化的文件方面的各种问题。

[0150] HEVC标准的某些版本提供对于称为频块(tile)的矩形区的译码的支持。这些区可具有与来自先前经译码图片的其它频块的译码相依性或可经独立地解码。频块可经描述为对应于帧中的一或多个图块的单个矩形区。频块可成群成频块集合,其对应于在单个图块中译码的若干频块。频块集合可用以描述频块之间的相依性和/或描述跨越若干频块的关注区。

[0151] 包含频块的样本的ISO基本媒体文件中的播放轨通常被称作频块播放轨。频块播放轨可具有区分这些播放轨与其它播放轨的要求。来自ISO/IEC 14496-15的条款10.6.1的以下摘录提供这些要求的一些实例:

[0152] “HEVC(分别地LHEVC)频块播放轨为存在对携带频块所属的相关联HEVC层的NALU的HEVC(分别地LHEVC)播放轨的‘tbas’参考的视频播放轨。HEVC频块播放轨的样本描述类型应为‘hvt1’。LHEVC频块播放轨的样本描述类型应为‘lht1’。

[0153] “频块播放轨中的样本与样本描述框都不应含有VPS、SPS或PPS NAL单元,这些NAL单元应在样本中或在含有相关联层的播放轨的样本描述框中,如通过播放轨参考所识别。HEVC/LHEVC频块播放轨和含有相关联层的播放轨(如由‘tbas’播放轨参考所指示)两者可使用如附录B中所定义的提取器以指示原始位流如何被重建;在某些应用领域中,提取器在这些播放轨中的存在可受限制。”

[0154] 在上文中,包含频块播放轨所属的层的播放轨(其由频块播放轨的‘tbas’参考所参考)可被称作基本播放轨。

[0155] 如上述摘录中所提及,提取器可用以重建频块化的位流。提取器为ISO基本媒体文件中的使得能够从不同于含有提取器的播放轨的播放轨提取NAL单元的结构。明确地说,提取器含有指令以从另一播放轨提取数据,所述另一播放轨连结到提取器驻存于其中的播放轨。ISO/IEC 14496-15的条款10.6.4将提取器关于频块播放轨的使用描述如下:

[0156] “使用频块播放轨的HEVC或L-HEVC播放轨可使用提取器以参考来自其频块播放轨的数据,在此状况下播放轨应具有对频块播放轨的‘scal’播放轨参考。”

[0157] 在一些状况下,循环提取器参考可在ISO基本媒体文件包含HEVC和L-HEVC频块播放轨时发生可为可能的。举例来说,如上文所提及,参数集(例如,VPS、SPS或PPS NAL单元)

未存储于频块播放轨中,且实际上可在样本描述框或含有与频块播放轨(例如,基本播放轨)相关联的层的播放轨的样本中发现。频块播放轨中的样本因此使用提取器以参考参数集。然而,在一些状况下,基本播放轨自身可使用提取器以参考频块播放轨中的NAL单元,如由来自条款10.6.4的上述引用规定。因此可产生频块播放轨的样本与基本播放轨中的样本之间的循环参考。

[0158] 各种技术可用以克服频块播放轨中的循环提取器参考(在这些参考发生时)。举例来说,约束条件可经定义以使得当播放轨含有对频块播放轨的‘scal’播放轨参考时,参考频块播放轨不应含有提取器。在此实例中,参数集应仅仅存在于样本项中,且不在由‘tbas’播放轨参考所参考的基本播放轨的样本中。

[0159] 在一些状况下,一或多个频块化的存取单元从ISO基本媒体文件的重建也可遇到问题。频块化的存取单元的重建描述于IOS/IEC 14496-14的条款10.6.4中,且可概述如下:对于隐式重建,使用频块播放轨的HEVC或L-HEVC播放轨应使用对频块播放轨的‘sabt’播放轨参考来指示频块定序;HEVC或LHEVC播放轨的样本是通过按播放轨参考的次序将在此播放轨的‘sabt’播放轨参考中指示的所有播放轨中的具有相同解码时间(例如,仅仅使用时间到样本表而不考虑编辑列表)的样本中所含有的NAL单元附加到样本数据而重建。

[0160] 当HEVC或LHEVC频块播放轨使用提取器(例如,以参考含于HEVC或LHEVC播放轨样本中的参数集NAL单元)时,问题可随上述重建方法一起发生。在一些状况下,由提取器参考的数据将在最终经重建的频块化的存取单元中复制。

[0161] 各种技术可用以避免数据在频块化的存取单元中的复制。举例来说,可施加约束条件以使得HEVC或L-HEVC频块播放轨不应使用提取器以参考频块播放轨的‘tbas’播放轨。作为另一实例,在频块存取单元的重建中,当使用隐式重建时,应忽略对‘tbas’播放轨的提取器。作为另一实例,可施加约束条件以使得对于频块化的存取单元的隐式重建,应忽略频块播放轨中的提取器(如果存在)。

[0162] 如由来自上文引用的ISO/IEC 14496-15的条款10.6.1的摘录规定,HEVC和L-HEVC频块播放轨可使用提取器。举例来说,频块播放轨可从由频块播放轨的‘tbas’参考所参考的播放轨(例如,基本播放轨)提取样本。频块播放轨的解码和/或播放因此需要基本播放轨的存在。由于提取器的一个目的是提供播放轨的更紧凑形式,因此使频块播放轨使用提取器似乎与此目的相反。

[0163] 各种技术可用以避免在无利可图的情况下使用频块播放轨中的提取器。举例来说,在各种实施方案中,可禁止使用任一频块播放轨中的提取器。

[0164] 与频块播放轨相关的可能出现的另一问题发生在仅仅需要来自HEVC或L-HEVC频块播放轨的视频部分时。在此情况下,ISO/IEC 14496-15规范不指定错误剖析器应产生用于解码器引擎的数据。

[0165] 各种技术可用以克服在仅仅需要HEVC或L-HEVC频块播放轨的视频部分时 ISO/IEC 14496-15中关于流式传输或存取单元构建的遗漏规范。举例来说,可使用以下过程:首先,可输出含于样本项中的参数集和补充增强信息(SEI)NAL单元。接下来,可按以下次序输出其它NAL单元:必须在VCL NAL单元(例如,参数集NAL单元、前置SEI NAL单元等)之前的在基本播放轨(如通过‘tbas’播放轨参考所指示)中的第一样本(下文称为“firstSampleInBaseTrack”)中的NAL单元;频块播放轨中的 firstSampleInBaseTrack的

相应样本中的NAL单元;必须在VLC NAL单元(例如,EOS NAL单元、EOB NAL单元)之后的在firstSampleInBaseTrack中的NAL单元;必须在VLC NAL单元之前的在基本播放轨中的第二样本(下文称为“secondSampleInBaseTrack”)中的NAL单元;在频块播放轨中的secondSampleInBaseTrack的相应样本中的NAL单元;必须在VLC NAL单元之后的在secondSampleInBaseTrack中的NAL单元,等等。

[0166] 在上述过程中,基本播放轨和频块播放轨中的样本是通过解码时间而同步。换句话说,基本播放轨中的样本和其在频块播放轨中的相应样本具有相同解码时间。

[0167] 对于HEVC和L-HEVC频块播放轨的一个使用状况为解码图片内的关注区(ROI),而不是解码整个图片。在此状况下,图片的存储在频块播放轨中的部分可被解码。为如此执行,可需要知晓图片内的频块的相对位置。

[0168] 各种技术可用以用信号传送图片内频块的位置。举例来说,偏移信息可包含于HEVC 或L-HEVC频块播放轨的样本项中。偏移信息可包含水平和/或垂直偏移。可以像素单位、派卡(picas)、厘米、几分之一英吋、宏块或某一其它测量单位提供偏移。

[0169] 为提供频块播放轨的偏移信息,提议对ISO/IEC 14496-15的章节10.6.2.2和10.6.3.2 的修改。这些章节的本文提供于下,其中被添加到这些章节的文字使用带下划线的文字(添加文字的实例)指示。

[0170] 10.6.2.2语法

```
class HEVCTileConfigurationBox extends Box('hvtC') {
[0171]     unsigned int(16) horizontal_offset;
           unsigned int(16) vertical_offset;
[0172]     HEVCTileTierLevelConfigurationRecord() HEVCTileTierLevelConfig;
}
```

[0173] horizontal\_offset和vertical\_offset按亮度样本的单位给出由此播放轨表示的频块所表示的矩形区的左上亮度样本相对于由‘tbas’参考播放轨表示的图片的左上亮度样本的水平和垂直偏移。

[0174] 10.6.3.2语法

```
class LHEVCTileSampleEntry() extends VisualSampleEntry ('lht1'){
           unsigned int(16) min_horizontal_offset;
           unsigned int(16) min_vertical_offset;
           unsigned int(16) max_horizontal_offset;
[0175]    unsigned int(16) max_vertical_offset;
           MPEG4BitRateBox ();                      // 可选
           Box extra_boxes[];                          // 可选
}
```

[0176] min\_horizontal\_offset和min\_vertical\_offset按亮度样本的单位给出最低层的由此播放轨表示的频块所表示的矩形区的左上亮度样本相对于由‘tbas’参考播放轨表

示的最低层的图片的左上亮度样本的水平 and 垂直偏移。

[0177] max\_horizontal\_offset和max\_vertical\_offset按亮度样本的单位给出最高层的由此播放轨表示的频块所表示的矩形区的左上亮度样本相对于由‘tbas’参考播放轨表示的最低层的图片的左上亮度样本的水平 and 垂直偏移。

[0178] 图10说明用于处理经编码视频数据的过程1000的实例,其中包含经编码视频数据的文件可具有多于一个EOB NAL单元。在1002处,过程1000包含接收与经编码视频数据相关联的文件,其中所述经编码视频数据根据文件格式存储于文件中,其中经编码视频数据包含两个或多于两个层,且其中两个或多于两个层存储于文件中的两个或多于两个播放轨中。在一些状况下,文件中的至少两层可各自具有EOB NAL单元。在一些状况下,仅仅基层包含EOB NAL单元。在一些状况下,一或多个增强层可各自具有EOB NAL单元。

[0179] 在1004处,过程1000包含从两个或多于两个层当中确定第一层,其中第一层经确定为两个或多于两个层当中包含位流结尾指示符的最高层(例如,具有最高层ID值的层)。即,当文件包含具有EOB指示符的多于一个层时,在步骤1004处识别(例如,确定)具有EOB指示符的最高层。在一些状况下,可能在文件中存在甚至更高层(例如,与大于具有EOB指示符之前述最高层的层ID值相关联的层),但这些更高层不包含EOB。

[0180] 在1006处,过程1000包含确定第一位流结尾指示符,其中第一位流结尾指示符为来自第一层的位流结尾指示符。

[0181] 在1008处,过程1000包含从两个或多于两个层当中选择(例如,识别)一组层。在各种实施方案中,所述一组层可基于经选择用于显示的层中的样本而选择。在一些状况下,所述一组层可包含比文件中的层的数目少的层。在一些状况下,所述一组层可包含仅仅一个层。

[0182] 在1010处,过程1000包含产生一系列存取单元,其中所述系列存取单元包含来自所述一组层的样本。

[0183] 在1012处,过程1000包含使用所述系列存取单元和第一位流结尾指示符产生经重建位流,其中经重建位流中的最后存取单元包含第一位流结尾指示符。举例来说,在一些状况下,如上文所描述,一系列存取单元的最后存取单元可不包含位流结尾指示符。在这些状况下,第一位流结尾指示符可被添加到所述系列存取单元中的最后存取单元。在各种实施方案中,第一位流结尾指示符可在位流重建期间被添加到所述系列存取单元中的最后存取单元。所述系列存取单元中的最后存取单元接着可变成经重建位流中的最后存取单元。

[0184] 在一些状况下,重建位流所源于的所述一组层可不包含在步骤1004处确定的第一层。换句话说,与经重建位流相关联的所述一组层可不包含与最大层ID值相关联的也包含位流结尾指示符的层。在这些状况下,第一层可为与高于存在于步骤1008中所选择或识别的所述一组层中的最高层的层ID值相关联的层。当情况就是这样时,包含于第一层内的第一位流结尾指示符可经添加(例如,被包含)到所述系列存取单元中的最后存取单元,而不管所述一组层中的任何层是否将位流结尾指示符贡献给此最后存取单元。

[0185] 在一些状况下,在步骤1004确定的第一层可为比在步骤1008处确定的所述一组层中的最高层低的层。在这些状况下,来自第一层的第一位流结尾指示符可移动到所述系列存取单元中的最后存取单元。第一层可或可不在所述一组层中。

[0186] 在一些状况下,在产生经重建位流中的最后存取单元后,此最后存取单元可具有

多于一个位流结尾指示符。当情况就是这样时,可去除位流结尾指示符中的不同于来自第一层的第一位流结尾指示符的每一个。因此,经重建位流中的最后存取单元将仅仅包含第一位流结尾指示符。

[0187] 在一些实例中,过程1000可由计算装置或设备(例如,系统100)执行。举例来说,过程1000可由图1中展示的系统100和/或存储装置108或输出端110执行。在一些状况下,计算装置或设备可包含处理器、微处理器、微计算机或经配置以实施过程1000的步骤的装置的其它组件。在一些实例中,计算装置或设备可包含经配置以捕获包含视频帧的视频数据(例如,视频序列)的摄影机。举例来说,计算装置可包含摄影机装置(例如,IP摄影机或其它类型的摄影机装置),所述摄影机装置可包含视频编码解码器。在一些实例中,捕获视频数据的摄影机或其它捕获装置与计算装置分离,在此状况下,计算装置接收所捕获视频数据。计算装置可进一步包含经配置以传达视频数据的网络接口。网络接口可经配置以传达基于因特网协议(IP)的数据。

[0188] 过程1000经说明为逻辑流程图,其操作表示可在硬件、计算机指令或其组合中实施的操作的序列。在计算机指令的上下文中,操作表示存储于一或多个计算机可读存储媒体上的当由一或多个处理器执行时执行所叙述的操作的计算机可执行指令。通常,计算机可执行指令包含执行特定功能或实施特定数据类型的例程、程序、物件、组件、逻辑、数据结构和其类似者。描述操作的次序并不希望被理解为限制,且任何数目个经描述操作可按任何次序组合和/或与实施程序并行。

[0189] 另外,过程1000可在经配置有可执行指令的一或多个计算机系统的控制下执行,且可被实施为共同在一或多个处理器上执行的代码(例如,可执行指令、一或多个计算机程序或一或多个应用程序)、由硬件实施或其组合。如上文所提及,代码可存储于计算机可读或机器可读存储媒体上,例如,呈包括可由一或多个处理器执行的多个指令的计算机程序的形式。计算机可读或机器可读存储媒体可为非暂时性的。

[0190] 编码装置1004和解码装置1212的特定细节分别展示于图11和图12中。图11为说明可实施本发明中所描述的技术中的一或多个的实例编码装置1104的框图。编码装置1104可(例如)产生本文中所描述的语法结构(例如,VPS、SPS、PPS或其它语法元素的语法结构)。编码装置1104可执行视频图块内的视频块的帧内预测和帧间预测译码。如先前所描述,帧内译码至少部分地依赖于空间预测以减少或去除给定视频帧或图片内的空间冗余。帧间译码至少部分地依赖于时间预测以减少或去除视频序列的邻近或周围帧内的时间冗余。帧内模式(I模式)可指若干基于空间的压缩模式中的任一个。帧间模式(例如,单向预测(P模式)或双向预测(B模式))可指若干基于时间的压缩模式中的任一个。

[0191] 编码装置1104包含分割单元35、预测处理单元41、滤波器单元63、图片存储器64、求和器50、变换处理单元52、量化单元54和熵编码单元56。预测处理单元41包含运动估计单元42、运动补偿单元44和帧内预测处理单元46。对于视频块重建,编码装置1104也包含反量化单元58、反变换处理单元60和求和器62。滤波器单元63希望表示一或多个回路滤波器,例如解块滤波器、自适应回路滤波器(ALF)和样本自适应偏移(SAO)滤波器。尽管滤波器单元63在图11中展示为回路滤波器,但在其它配置中,滤波器单元63可实施为后回路滤波器。后处理装置57可对由编码装置1104产生的经编码视频数据执行额外处理。本发明的技术可在一些情况下由编码装置1104实施。然而,在其它情况下,本发明的技术中的一或多个可由后

处理装置57实施。

[0192] 如图11中所展示,编码装置1104接收视频数据,且分割单元35将数据分割成视频块。分割也可包含(例如)根据LCU和CU的四分树结构分割成图块、图块片段、频块或其它较大单元,以及视频块分割。编码装置1104通常说明编码待编码的视频图块内的视频块的组件。图块可划分成多个视频块(且可能划分成被称作频块的视频块集合)。预测处理单元41可基于错误结果(例如,译码速率和失真等级等)选择多个可能译码模式中的一个(例如,多个帧内预测译码模式中的一个或多个帧间预测译码模式中的一个)以用于当前视频块。预测处理单元41可将所得帧内或帧间译码块提供到求和器50以产生残余块数据并提供到求和器62以重建经编码块以用于用作参考图片。

[0193] 预测处理单元41内的帧内预测处理单元46可执行当前视频块相对于与待译码的当前块在相同的帧或图块中的一或多个相邻块的帧内预测译码,以提供空间压缩。预测处理单元41内的运动估计单元42和运动补偿单元44执行当前视频块相对于一或多个参考图片中的一或多个预测性块的帧间预测性译码,以提供时间压缩。

[0194] 运动估计单元42可经配置以根据视频序列的预定图案来确定用于视频图块的帧间预测模式。预定图案可将序列中的视频图块指定为P图块、B图块或GPB图块。运动估计单元42和运动补偿单元44可高度集成,但出于概念目的而单独说明。由运动估计单元42执行的运动估计为产生运动向量的过程,所述运动向量估计视频块的运动。运动向量(例如)可指示当前视频帧或图片内的视频块的预测单元(PU)相对于参考图片内的预测性块的位移。

[0195] 预测性块为就像素差来说被发现紧密地匹配待译码的视频块的PU的块,所述像素差可由绝对差和(SAD)、平方差和(SSD)或其它差度量确定。在一些实例中,编码装置1104可计算存储于图片存储器64中的参考图片的次整数像素位置的值。举例来说,编码装置1104可内插参考图片的四分之一像素位置、八分之一像素位置或其它分数像素位置的值。因此,运动估计单元42可执行关于全像素位置和分数像素位置的运动搜索且输出具有分数像素精确度的运动向量。

[0196] 运动估计单元42通过将PU的位置与参考图片的预测性块的位置比较而计算经帧间译码图块中的视频块的PU的运动向量。参考图片可选自第一参考图片列表(列表0)或第二参考图片列表(列表1),所述参考图片列表中的每一个识别存储在图片存储器64中的一或多个参考图片。运动估计单元42将所计算的运动向量发送到熵编码单元56和运动补偿单元44。

[0197] 由运动补偿单元44执行的运动补偿可涉及基于通过运动估计确定的运动向量提取或产生预测性块,可能执行内插达子像素精确度。在接收到当前视频块的PU的运动向量之后,运动补偿单元44可在参考图片列表中定位运动向量所指向的预测性块。编码装置1104通过从正经译码的当前视频块的像素值减去预测性块的像素值从而形成像素差值来形成残余视频块。像素差值形成用于块的残余数据,且可包含亮度和色度差分量两者。求和器50表示执行此减法运算的一或多个组件。运动补偿单元44也可产生与视频块和视频图块相关联的语法元素,以供解码装置1212用于解码视频图块的视频块。

[0198] 作为如上文所描述由运动估计单元42和运动补偿单元44执行的帧间预测的替代,帧内预测处理单元46可对当前块进行帧内预测。特定来说,帧内预测处理单元46可确定帧内预测模式以用以编码当前块。在一些实例中,帧内预测处理单元46可(例如)在单独的编

码遍次期间使用各种帧内预测模式来编码当前块,且帧内预测处理单元46(或在一些实例中,模式选择单元40)可从所测试的模式选择待使用的适当帧内预测模式。举例来说,帧内预测处理单元46可使用对各种所测试帧内预测模式的速率-失真分析来计算速率-失真值,且可在所测试模式间选择具有最佳速率-失真特性的帧内预测模式。速率-失真分析大体上确定经编码块与原始未经编码块(其经编码以产生经编码块)之间的失真(或误差)量,以及用以产生经编码块的位速率(即,位的数目)。帧内预测处理单元 46可根据各种经编码块的失真和速率来计算比率以确定哪一帧内预测模式展现所述块的最佳速率-失真值。

[0199] 在任何情况下,在选择用于块的帧内预测模式之后,帧内预测处理单元46可将指示用于块的所选择帧内预测模式的信息提供到熵编码单元56。熵编码单元56可编码指示所选定帧内预测模式的信息。编码装置1104可将各种块的编码上下文的定义以及待用于上下文中的每一个的最可能的帧内预测模式、帧内预测模式索引表和经修改帧内预测模式索引表的指示包含于经传输位流配置数据中。位流配置数据可包含多个帧内预测模式索引表和多个经修改帧内预测模式索引表(也被称作码字映射表)。

[0200] 在预测处理单元41经由帧间预测或帧内预测产生当前视频块的预测性块之后,编码装置1104通过从当前视频块减去预测性块而形成残余视频块。残余块中的残余视频数据可包含于一或多个TU中且被应用于变换处理单元52。变换处理单元52使用例如离散余弦变换(DCT)或概念上类似变换的变换将残余视频数据变换成残余变换系数。变换处理单元52可将残余视频数据从像素域转换到变换域(例如,频域)。

[0201] 变换处理单元52可将所得变换系数发送到量化单元54。量化单元54量化变换系数以进一步减小位速率。量化过程可减小与系数中的一些或所有相关联的位深度。可通过调整量化参数来修改量化程度。在一些实例中,量化单元54可接着执行对包含经量化变换系数的矩阵的扫描。替代性地,熵编码单元56可执行扫描。

[0202] 在量化之后,熵编码单元56对经量化变换系数进行熵编码。举例来说,熵编码单元56可执行上下文自适应性可变长度译码(CAVLC)、上下文自适应性二进制算术译码(CABAC)、基于语法的上下文自适应性二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码技术。在由熵编码单元56进行熵编码之后,可将经编码位流传输到解码装置1212,或加以存档以供稍后由解码装置1212传输或检索。熵编码单元56也可对正经译码的当前视频图块的运动向量和其它语法元素进行熵编码。

[0203] 反量化单元58和反变换处理单元60分别应用反量化和反变换以重建像素域中的残余块以供稍后用作参考图片的参考块。运动补偿单元44可通过将残余块添加到参考图片列表内的参考图片中的一个的预测性块来计算参考块。运动补偿单元44也可将一或多个内插滤波器应用到经重建残余块,以计算次整数像素值以用于运动估计中。求和器 62将经重建残余块添加到由运动补偿单元44所产生的运动补偿预测块以产生用于存储于图片存储器64中的参考块。参考块可由运动估计单元42和运动补偿单元44用作参考块以帧间预测后续视频帧或图片中的块。

[0204] 以此方式,图11的编码装置1104表示经配置以产生经编码视频位流的语法的视频编码器的实例。编码装置1104可(例如)产生如上文所描述的VPS、SPS和PPS参数集。编码装置1104可执行本文中所描述的技术中的任一个,包含上文关于图11和图12描述的过程。大体上已关于编码装置1104描述本发明的技术,但如上文所提及,本发明的技术中的一些也



可由后处理装置57实施。

[0205] 图12为说明实例解码装置1212的框图。解码装置1212包含熵解码单元80、预测处理单元81、反量化单元86、反变换处理单元88、求和器90、滤波器单元91和图片存储器92。预测处理单元81包含运动补偿单元82和帧内预测处理单元84。解码装置 1212可在一些实例中执行与关于来自图11的视频装置1104描述的编码遍次大体上互逆的解码遍次。

[0206] 在解码过程期间,解码装置1212接收表示由编码装置1104发送的经编码视频图块的视频块和相关联语法元素的经编码视频位流。在一些实施例中,解码装置1212可从编码装置1104接收经编码视频位流。在一些实施例中,解码装置1212可从网络实体79 (例如,服务器、媒体感知网络元件(MANE)、视频编辑器/编接器或经配置以实施上文所描述的技术中的一或多个的其它此类装置)接收经编码视频位流。网络实体79可或可不包含编码装置1104。在网络实体79将经编码视频位流传输到解码装置1212之前,本发明中所描述的技术中的一些可由网络实体79实施。在一些视频解码系统中,网络实体 79和解码装置1212可为独立装置的部分,而在其它情况下,关于网络实体79描述的功能性可由包括解码装置1212的同一装置执行。

[0207] 解码装置1212的熵解码单元80对位流进行熵解码以产生经量化系数、运动向量和其它语法元素。熵解码单元80将运动向量和其它语法元素转递到预测处理单元81。解码装置1212可接收视频图块层级和/或视频块层级的语法元素。熵解码单元80可处理并剖析例如VPS、SPS和PPS的一或多个参数集中的固定长度语法元素和可变长度语法元素两者。

[0208] 当将视频图块译码为经帧内译码(I)图块时,预测处理单元81的帧内预测处理单元 84可基于用信号传送的帧内预测模式和来自当前帧或图片的先前经解码块的数据而产生当前视频图块的视频块的预测数据。当视频帧被译码为经帧间译码(即,B、P或GPB)图块时,预测处理单元81的运动补偿单元82基于从熵解码单元80接收的运动向量和其它语法元素而产生当前视频图块的视频块的预测性块。所述预测性块可从参考图片列表内的参考图片中的一个产生。解码装置1212可基于存储于图片存储器92中的参考图片使用默认构建技术来构建参考帧列表(列表0和列表1)。

[0209] 运动补偿单元82通过剖析运动向量和其它语法元素来确定当前视频图块的视频块的预测信息,且使用预测信息产生正经解码的当前视频块的预测性块。举例来说,运动补偿单元82可使用参数集中的一或多个语法元素来确定用于译码视频图块的视频块的预测模式(例如,帧内或帧间预测)、帧间预测图块类型(例如,B图块、P图块或GPB图块)、图块的一或多个参考图片列表的构建信息、图块的每一经帧间编码视频块的运动向量、图块的每一经帧间译码视频块的帧间预测状态和用以解码当前视频图块中的视频块的其它信息。

[0210] 运动补偿单元82也可执行基于内插滤波器的内插。运动补偿单元82可使用如在编码视频块期间由编码装置1104使用的内插滤波器来计算参考块的次整数像素的内插值。在此状况下,运动补偿单元82可从所接收语法元素确定由编码装置1104使用的内插滤波器,且可使用内插滤波器来产生预测性块。

[0211] 反量化单元86反量化(或解量化)位流中所提供,并由熵解码单元80解码的经量化的变换系数。反量化过程可包含使用由编码装置1104针对视频图块中的每一视频块计算的量化参数以确定应应用的量化程度和(同样地)反量化程度。反变换处理单元88将反变换(例如,反DCT或其它合适的反变换)、反整数变换或概念上类似的反变换过程应用于变换系

数以便在像素域中产生残余块。

[0212] 在运动补偿单元82基于运动向量和其它语法元素而产生当前视频块的预测性块之后,解码装置1212通过将来自反转换处理单元88的残余块与由运动补偿单元82所产生的相应预测性块求和,而形成经解码视频块。求和器90表示执行此求和运算的组件或多个组件。如果需要,也可使用回路滤波器(在译码回路中或在译码回路后)以使像素转变平滑,或以其它方式改善视频质量。滤波器单元91希望表示一或多个回路滤波器(例如,解块滤波器、可调适回路滤波器(ALF)和样本可适性偏移(SAO)滤波器)。尽管滤波器单元91在图12中展示为回路滤波器,但在其它配置中,滤波器单元91可实施为后回路滤波器。给定帧或图片中的经解码视频块接着存储于图片存储器92中,图片存储器92存储用于后续运动补偿的参考图片。图片存储器92也存储经解码视频以供稍后呈现于显示装置(例如,图1中所展示的视频目的地装置122)上。

[0213] 在前述描述中,申请案的方面是参考特定实施例而描述,但所属领域的技术人员将认识到本发明不限于此。因此,尽管申请案的说明性实施例已经在本文中详细地描述,应理解,本发明概念可以其它方式不同地体现并采用,且希望解释所附权利要求书以包含除现有技术所限制外的这些变化。上文所描述的发明的各种特征和方面可单独地或联合地使用。另外,实施例可用于超出本文所描述的那些环境和应用的任何数目个环境和应用,而不脱离本说明书的更广精神和范围。因此,本说明书和图式被视为说明性而不是限定性。出于说明的目的,以特定次序描述方法。应了解,在替代实施例中,可以与所描述的次序不同的次序来执行所述方法。

[0214] 在组件被描述为“经配置以”执行某些操作的情况下,可(例如)通过设计电子电路或其它硬件以执行操作、通过编程可编程电子电路(例如,微处理器或其它适合的电子电路)以执行操作或其任何组合来实现此配置。

[0215] 结合本文中所揭示的实施例而描述的各种说明性逻辑块、模块、电路和算法步骤可被实施为电子硬件、计算机软件或两者的组合。为了清楚地说明硬件与软件的此可互换性,各种说明性组件、块、模块、电路和步骤已在上文大体按其功能性加以描述。将此功能性实施为硬件还是软件视特定应用和外加于整个系统上的设计约束而定。对于每一特定应用来说,所属领域的技术人员可以变化的方式实施所描述的功能性,但不应将此实施决策解释为导致脱离本发明的范围。

[0216] 本文中描述的技术可以硬件、软件、固件或其任何组合实施。这些技术可实施于多种装置中的任何者中,例如,通用计算机、无线通信装置手机或具有多种用途(包含在无线通信装置手机和其它装置中的应用)的集成电路装置。可将描述为模块或组件的任何特征一起实施于集成式逻辑装置中或分开来实施为离散但可互操作的逻辑装置。如果以软件实施,那么所述技术可至少部分由包括包含当经执行时执行以上描述方法中的一或多个的指令的程序代码的计算机可读数据存储媒体实现。计算机可读数据存储媒体可形成计算机程序产品的部分,计算机程序产品可包含包装材料。计算机可读媒体可包括存储器或数据存储媒体,例如,随机存取存储器(RAM),例如,同步动态随机存取存储器(SDRAM)、只读存储器(ROM)、非易失性随机存取存储器(NVRAM)、电可抹除可编程只读存储器(EEPROM)、闪存、磁性或光学数据存储媒体和类似者。所述技术另外或替代地可至少部分由计算机可读通信媒体实现,所述计算机可读通信媒体承载或传达呈指令或数据结构的形式且可由计算机存

取、读取和/或执行的程序代码,例如,传播的信号或波。

[0217] 程序代码可由可包含一或多个处理器的处理器执行,例如,一或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程逻辑阵列(FPGA)或其它等效集成式或离散逻辑电路。此处理器可经配置以执行本发明中所描述的技术中的任何者。通用处理器可为微处理器;但在替代例中,处理器可为任何常规处理器、控制器、微控制器或状态机。处理器也可实施为计算装置的组合,例如DSP与微处理器的组合、多个微处理器、结合DSP核心的一或多个微处理器,或任何其它此配置。因此,如本文中所使用的术语“处理器”可指前述结构、前述结构的任何组合或适合于实施本文中描述的技术的任何其它结构或设备中的任一个。此外,在一些方面中,本文中描述的功能性可提供于经配置用于编码和解码的专用软件模块或硬件模块,或并入于组合的视频编码器-解码器(编码解码器)中。

[0218] 本文中所论述的译码技术可体现于实例视频编码和解码系统中。系统包含由目的地装置提供待稍后经解码的经编码视频数据的源装置。明确地说,源装置经由计算机可读媒体将视频数据提供到目的地装置。源装置和目的地装置可包括广泛范围的装置中的任一个,包含台式计算机、笔记本(即,膝上型)计算机、平板计算机、机顶盒、电话手持机(例如所谓的“智能”电话)、所谓“智能”平板计算机、电视机、摄影机、显示装置、数字媒体播放器、视频游戏控制台、视频流式传输装置等。在一些状况下,源装置和目的地装置可经装备以用于无线通信。

[0219] 目的地装置可经由计算机可读媒体接收待解码的经编码视频数据。计算机可读媒体可包括能够将经编码视频数据从源装置移动到目的地装置的任何类型的媒体或装置。在一个实例中,计算机可读媒体可包括通信媒体以使得源装置能够实时地将经编码视频数据直接传输到目的地装置。可根据通信标准(例如,无线通信协议)调制经编码视频数据,且将经编码视频数据传输到目的地装置。通信媒体可包括任何无线或有线通信媒体,例如,射频(RF)频谱或一或多个物理传输线。通信媒体可形成基于封包的网路(例如局域网、广域网或全球网路,例如因特网)的一部分。通信媒体可包含路由器、交换器、基站或可用于促进从源装置到目的地装置的通信的任何其它装备。

[0220] 在一些实例中,经编码数据可从输出接口输出到存储装置。类似地,可由输入接口从存储装置存取经编码数据。存储装置可包含多种分布式或本地存取的数据存储媒体中的任一个,例如,硬盘机、蓝光光盘、DVD、CD-ROM、闪存、易失性或非易失性存储器或用于存储经编码视频数据的任何其它合适的数字存储媒体。在另一实例中,存储装置可对应于文件服务器或可存储由源装置产生的经编码视频的另一中间存储装置。目的地装置可经由流式传输或下载从存储装置存取所存储的视频数据。文件服务器可为能够存储经编码视频数据且将彼经编码视频数据传输到目的地装置的任何类型的服务器。实例文件服务器包含网页服务器(例如,用于网站)、FTP服务器、网络附加存储(NAS)装置或本地磁盘机。目的地装置可经由任何标准数据连接(包含因特网连接)而存取经编码的视频数据。此可包含无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、缆线调制解调器等)或适合于存取存储于文件服务器上的经编码视频数据的两者的组合。来自存储装置的经编码视频数据的传输可为流式传输、下载传输,或其组合。

[0221] 本发明的技术不必限于无线应用或设定。所述技术可应用于视频译码以支持多种多媒体应用中的任一个,例如,空中电视广播、有线电视传输、卫星电视传输、因特网流式传

输视频传输(例如,经由HTTP的动态自适应流式传输(DASH))、经编码到数据存储媒体上的数字视频、存储于数据存储媒体上的数字视频的解码或其它应用。在一些实例中,系统可经配置以支持单向或双向视频传输以支持例如视频流式传输、视频播放、视频广播和/或视频电话的应用。

[0222] 在一个实例中,源装置包含视频源、视频编码器和输出接口。目的地装置可包含输入接口、视频解码器和显示装置。源装置的视频编码器可经配置以应用本文中所揭示的技术。在其它实例中,源装置和目的地装置可包含其它组件或布置。举例来说,源装置可从外部视频源(例如,外部摄影机)接收视频数据。同样,目的地装置可与外部显示装置介接,而非包含集成式显示装置。

[0223] 以上实例系统仅为一个实例。用于并行地处理视频数据的技术可由任何数字视频编码和/或解码装置执行。尽管本发明的技术通常由视频编码装置执行,但所述技术也可由视频编码器/解码器(通常被称为“编码解码器(CODEC)”)执行。此外,本发明的技术也可由视频预处理器执行。源装置和目的地装置仅为源装置产生经译码视频数据以供传输到目的地装置的此类译码装置的实例。在一些实例中,源装置和目的地装置可以大体上对称的方式操作,使得所述装置中的每一个包含视频编码和解码组件。因此,实例系统可支持视频装置之间的单向或双向视频传输,例如用于视频流式传输、视频播放、视频广播或视频电话。

[0224] 视频源可包含视频捕获装置,例如摄像机、含有先前捕获的视频的视频存档和/或用以从视频内容提供者接收视频的视频馈入接口。作为另一替代方案,视频源可产生基于计算机图形的数据作为源视频,或实况视频、存档视频和计算机产生的视频的组合。在一些状况下,如果视频源为摄像机,那么源装置和目的地装置可形成所谓摄影机电话或视频电话。然而,如上文所提及,本发明所描述的技术一般可适用于视频译码,且可适用于无线和/或有线应用。在每一状况下,可由视频编码器编码所捕获、经预捕获或计算机产生的视频。经编码视频信息接着可由输出接口输出到计算机可读媒体上。

[0225] 如所提及,计算机可读媒体可包含暂态媒体,例如无线广播或有线网络传输;或存储媒体(即,非暂时性存储媒体),例如硬盘、闪存盘、紧密光盘、数字视频光盘、蓝光光盘或其它计算机可读媒体。在一些实例中,网络服务器(未图示)可从源装置接收经编码视频数据,且(例如)经由网络传输将经编码视频数据提供到目的地装置。类似地,媒体生产设施(例如光盘冲压设施)的计算装置可从源装置接收经编码视频数据且生产含有经编码视频数据的光盘。因此,在各种实例中,计算机可读媒体可理解为包含各种形式的一或多个计算机可读媒体。

[0226] 目的地装置的输入接口从计算机可读媒体接收信息。计算机可读媒体的信息可包含由视频编码器定义的语法信息(其也由视频解码器使用),所述语法信息包含描述块和其它经译码单元(例如,图片群组(GOP))的特性和/或处理的语法元素。显示装置将经解码视频数据显示给用户,且可包括多种显示装置中的任一个,例如阴极射线管(CRT)、液晶显示器(LCD)、等离子显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。已描述了本发明的各种实施例。

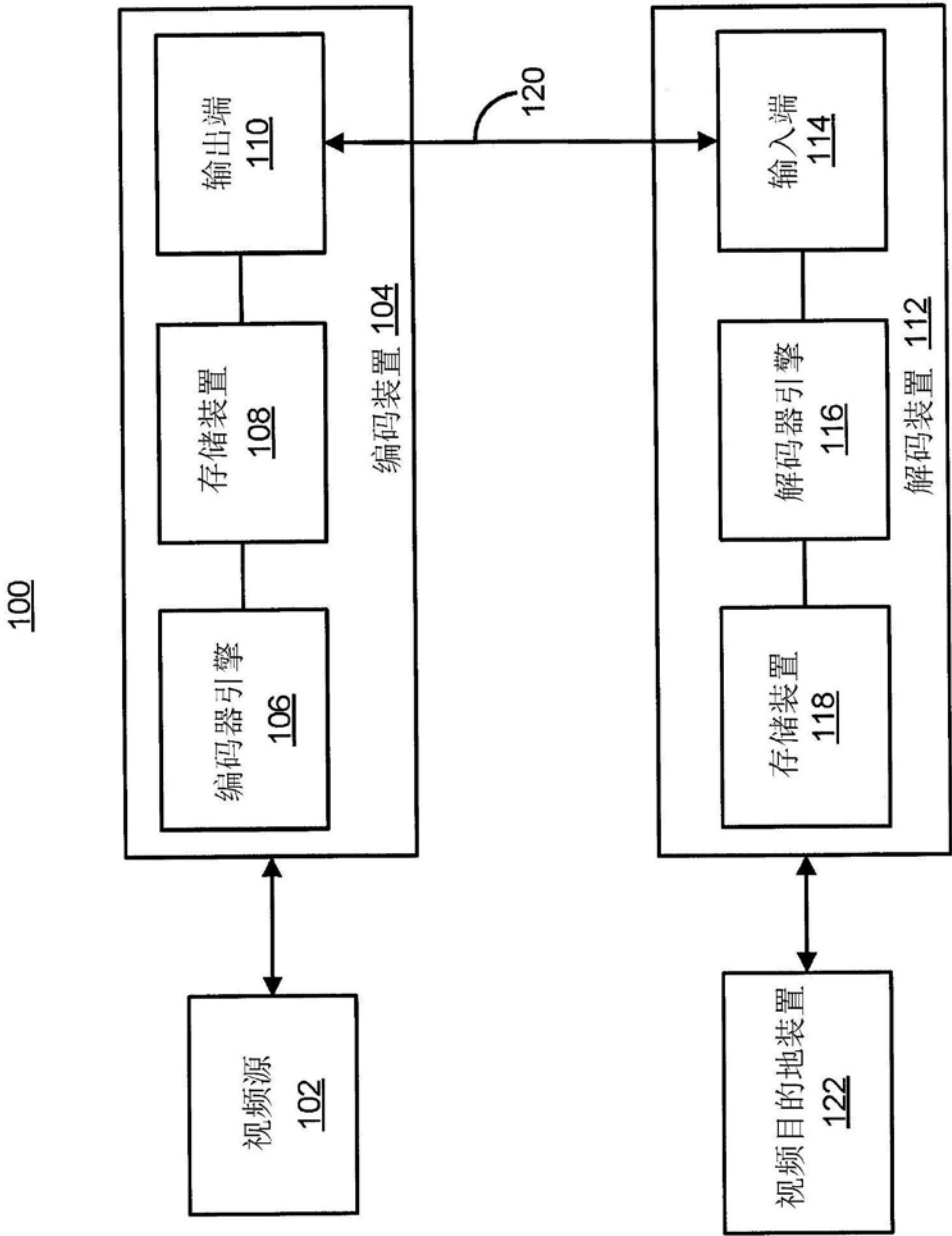


图1

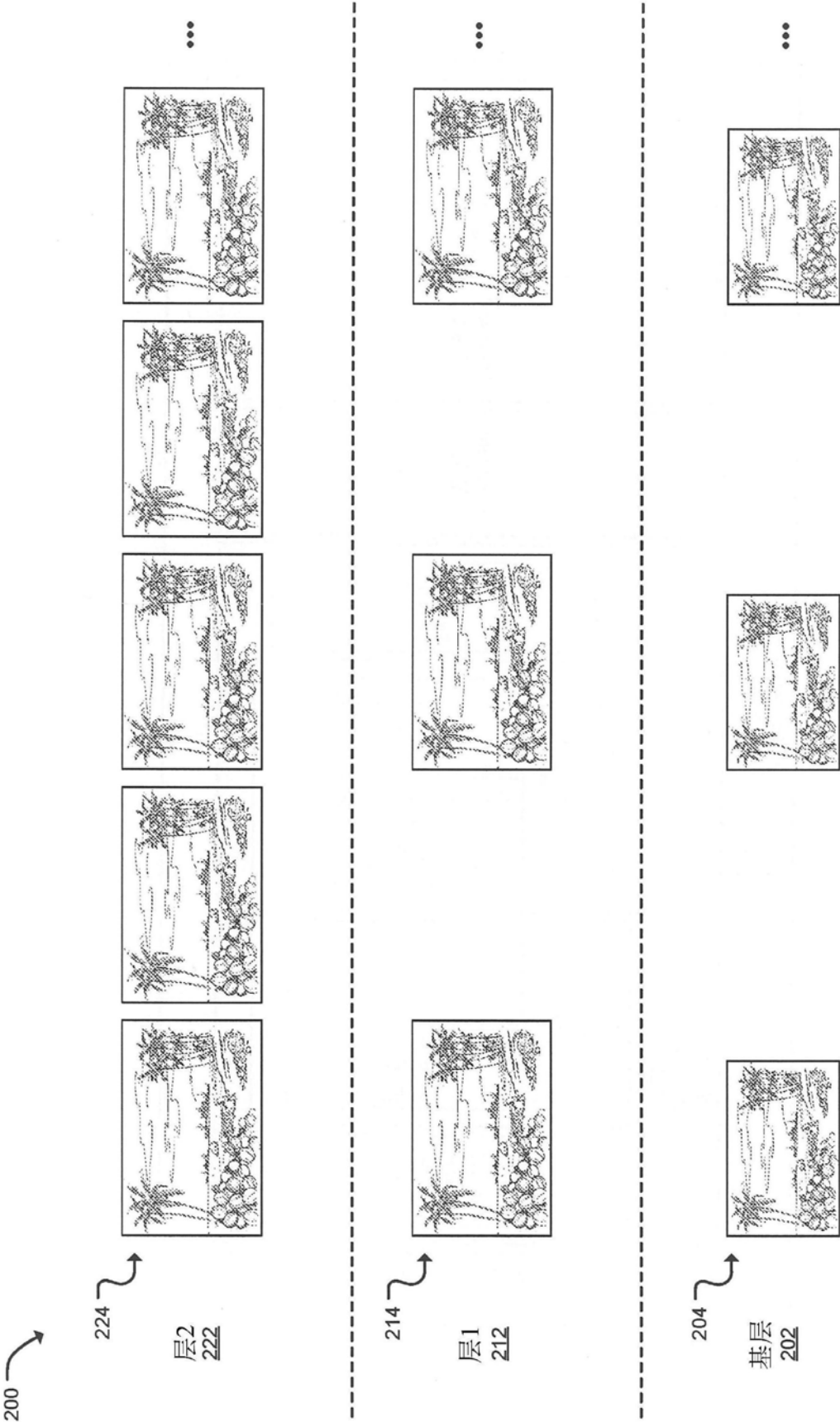


图2

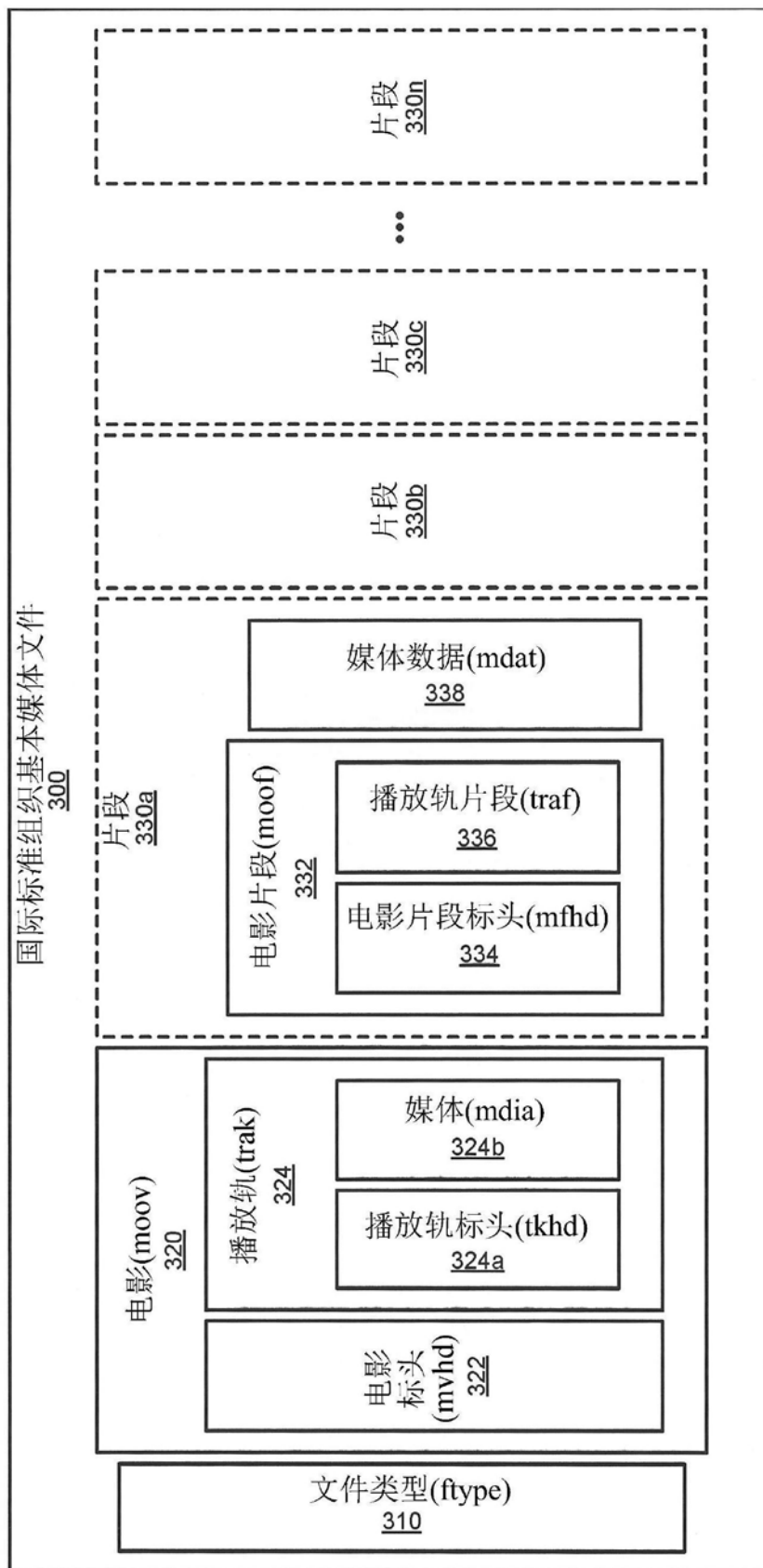


图3

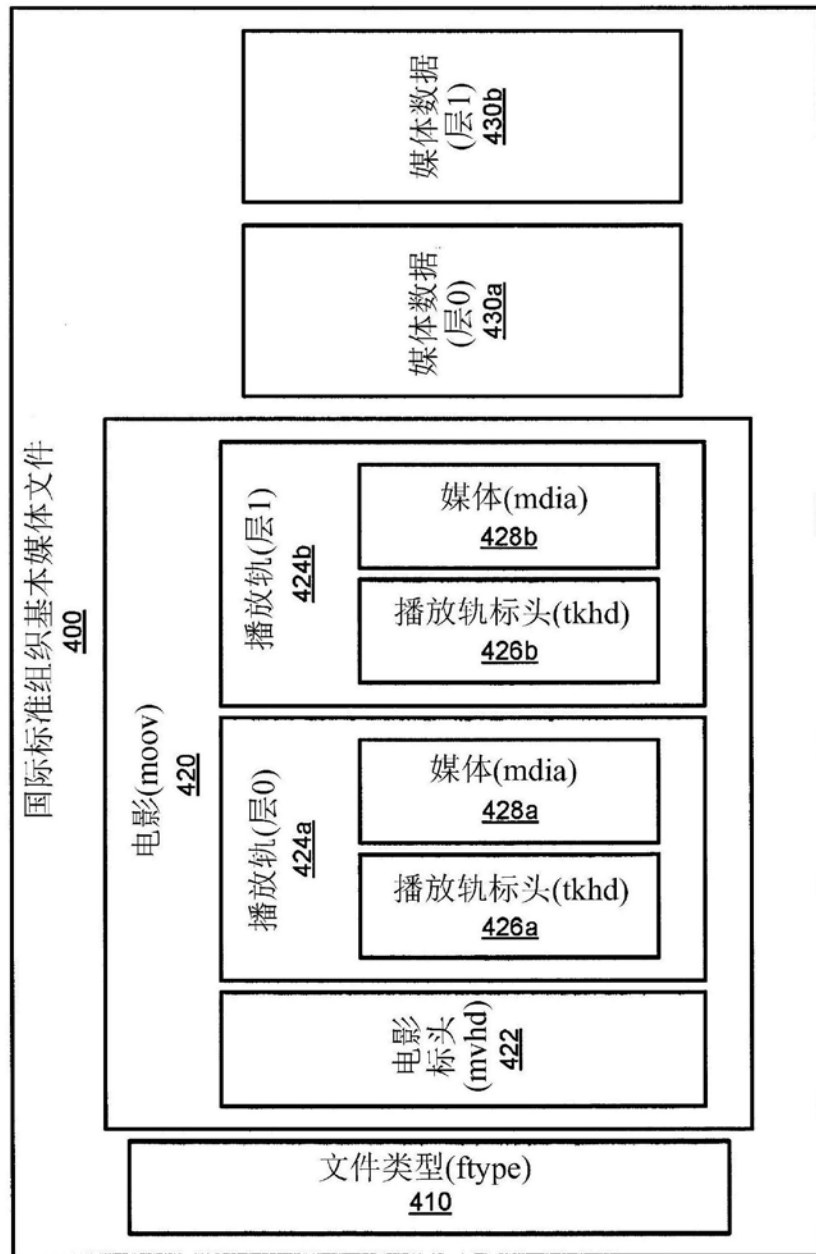


图4



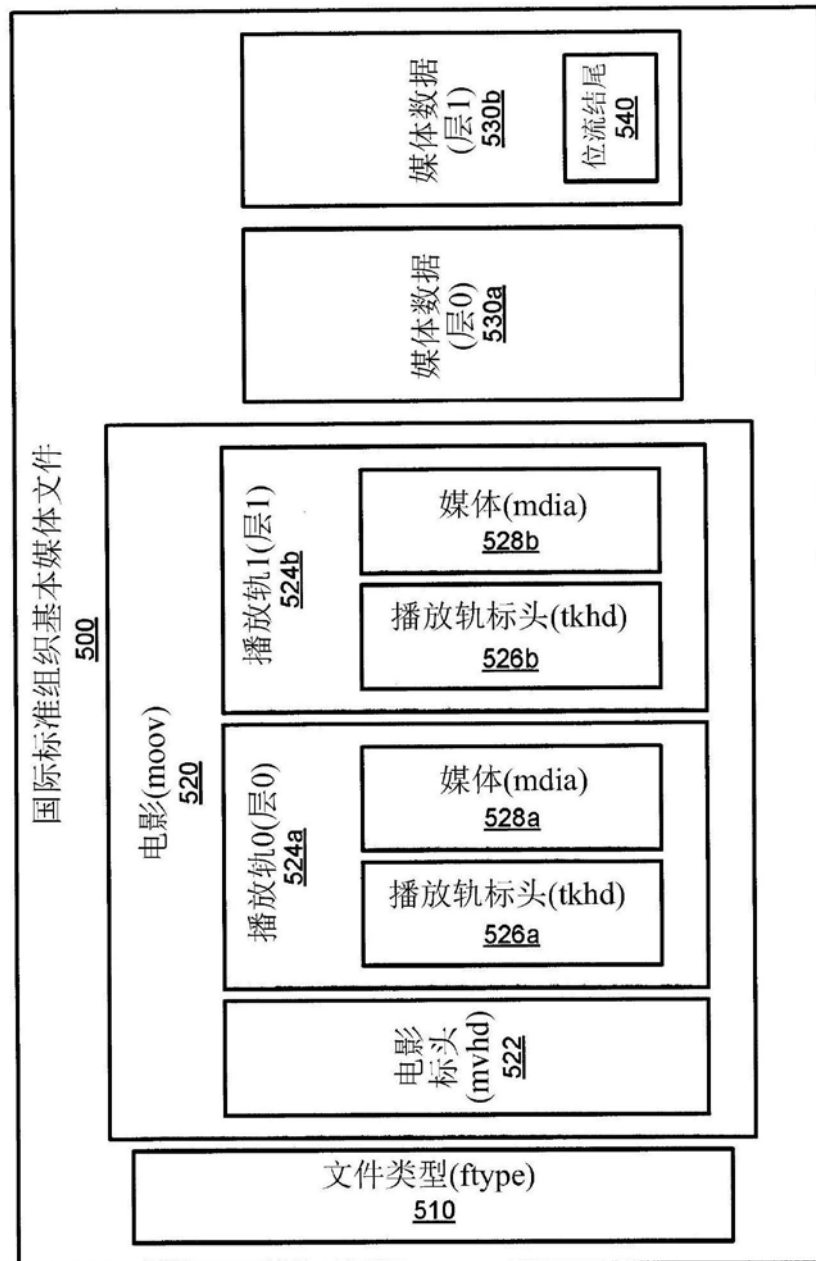


图5A

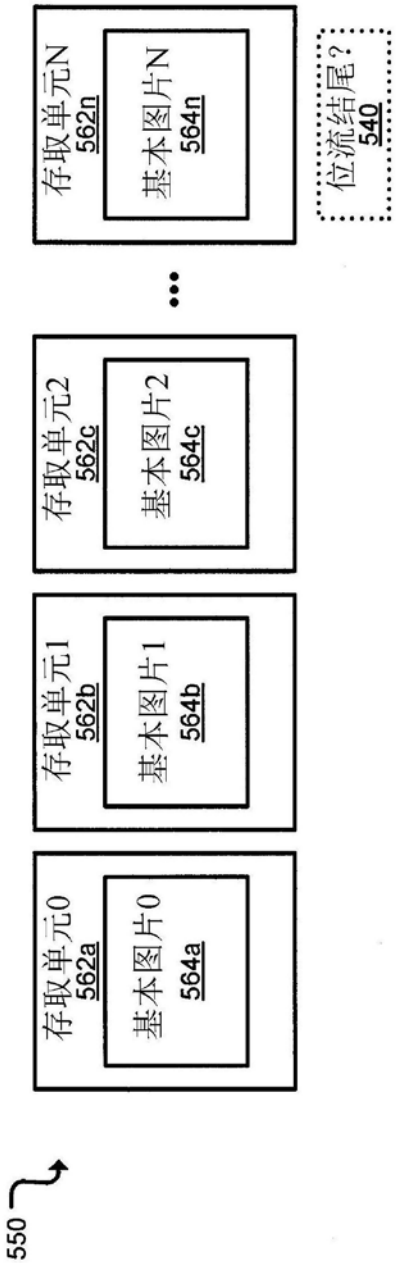


图5B

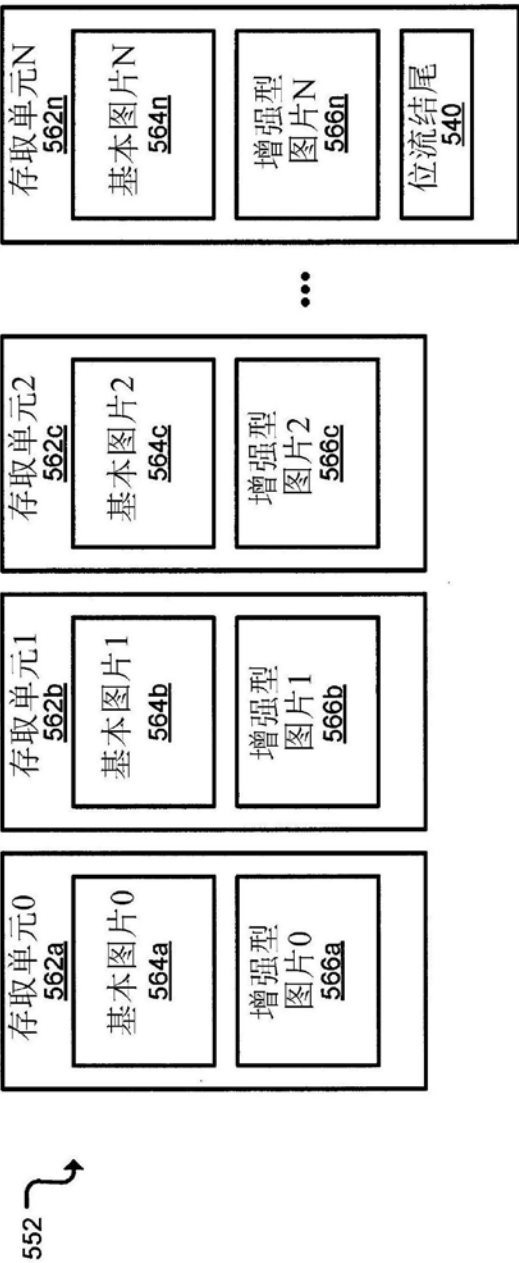


图5C

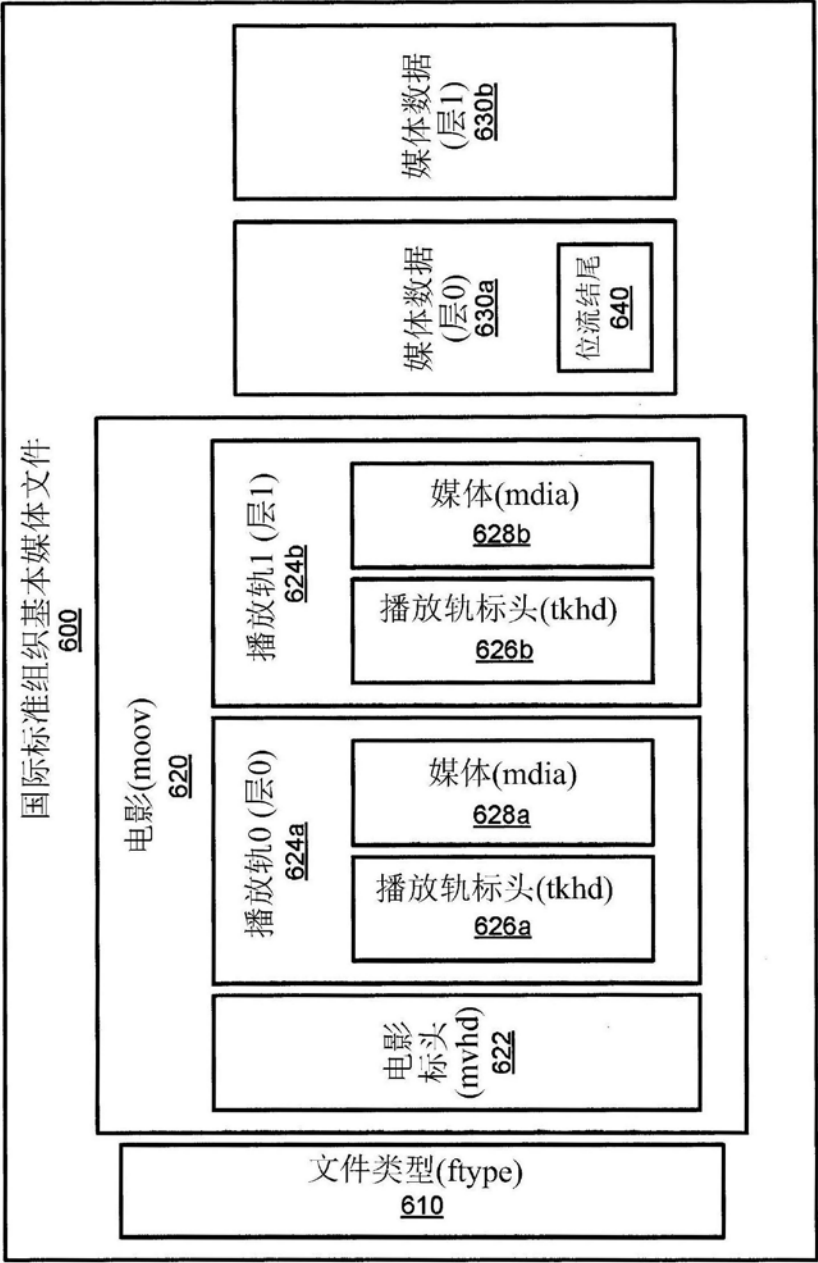


图6A

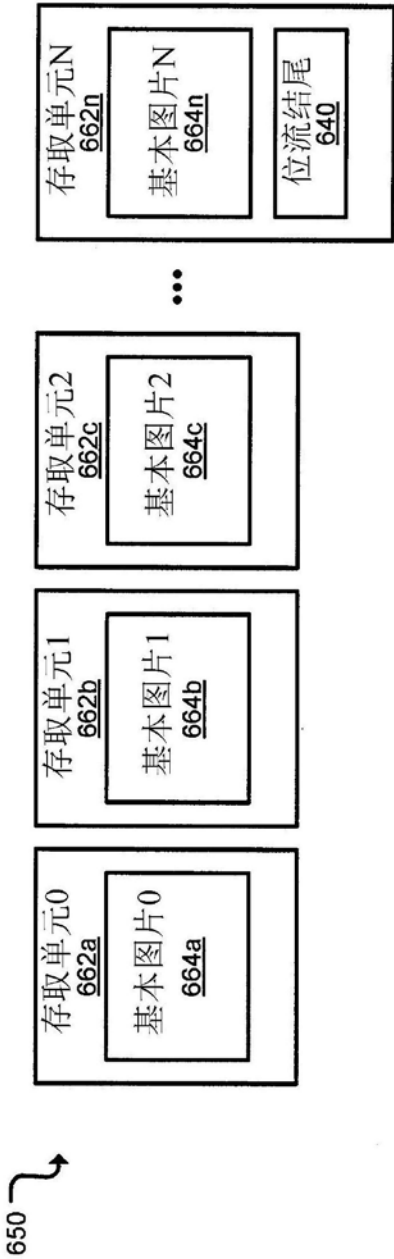


图6B

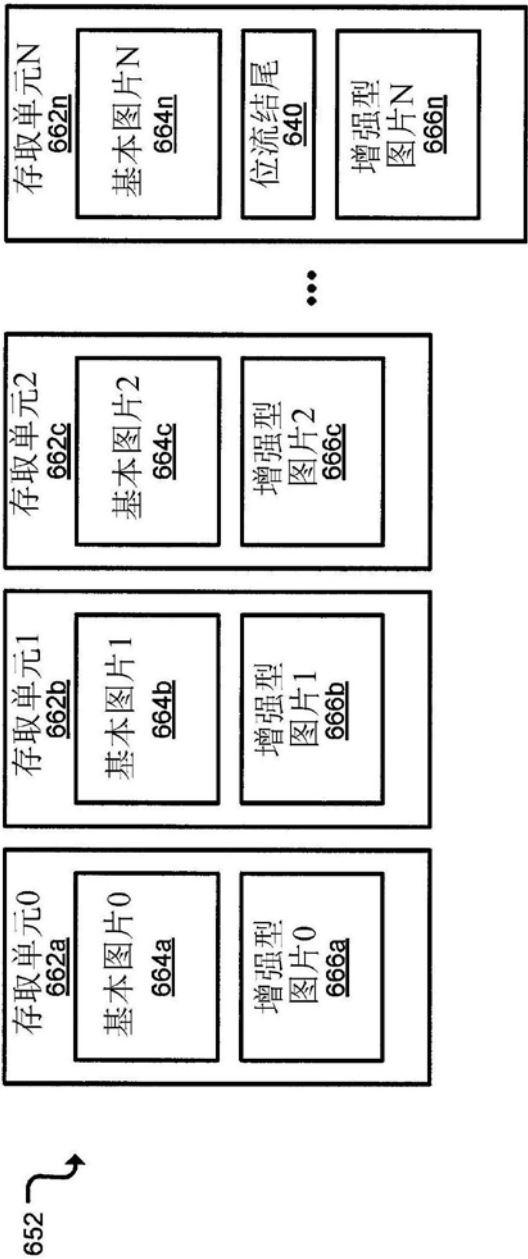


图6C

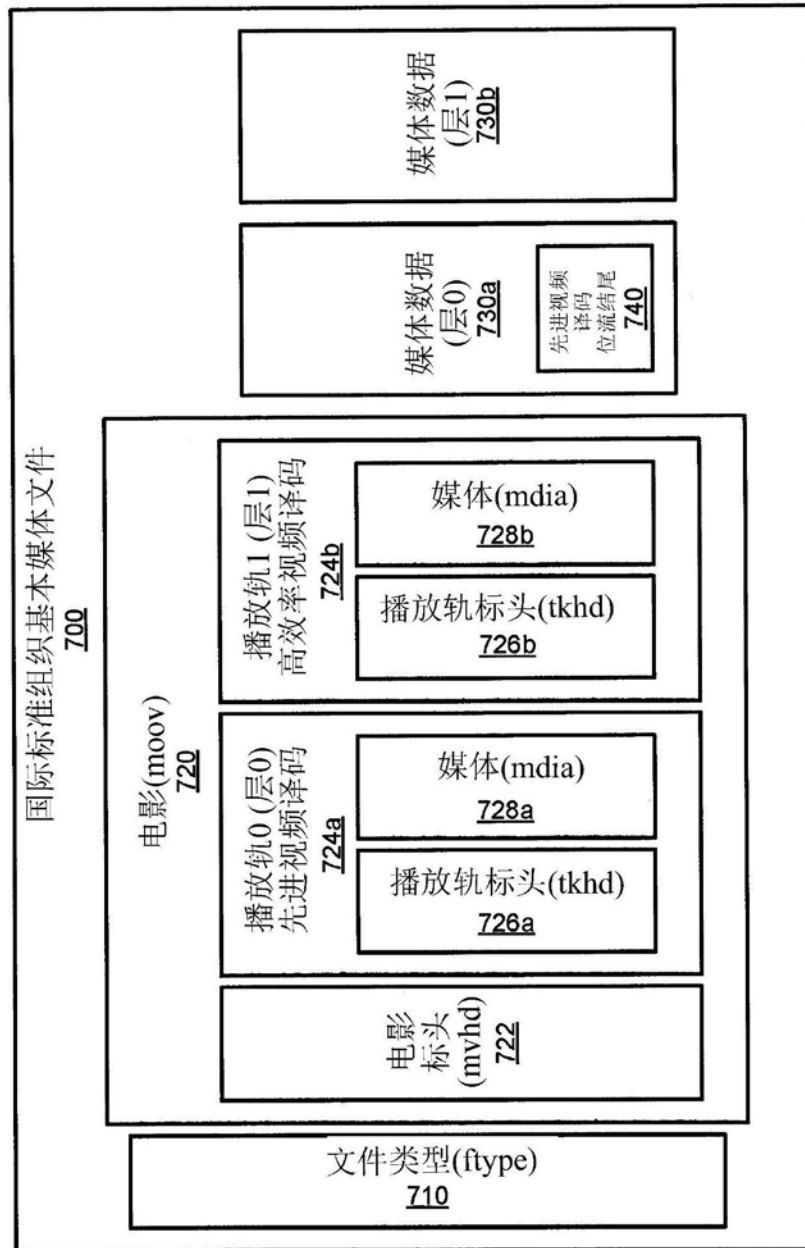


图7

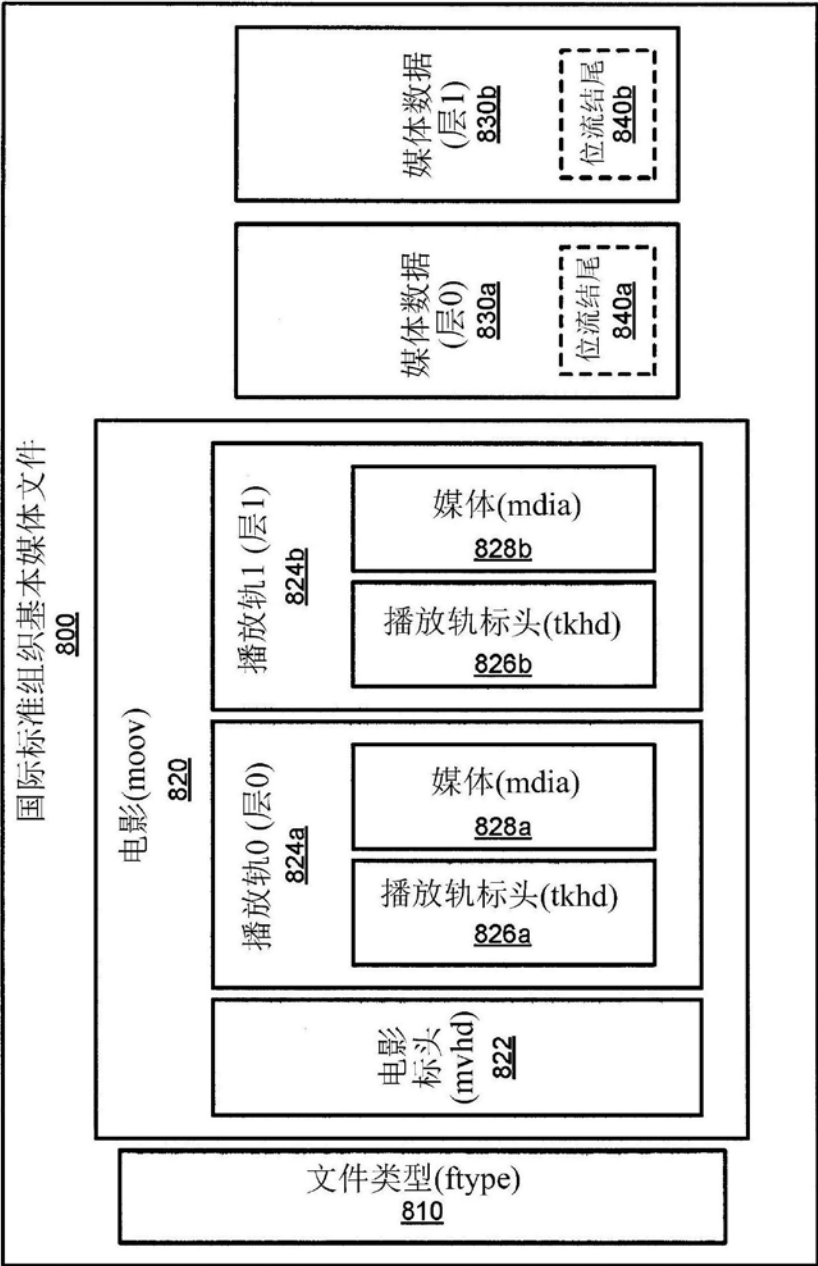


图8A



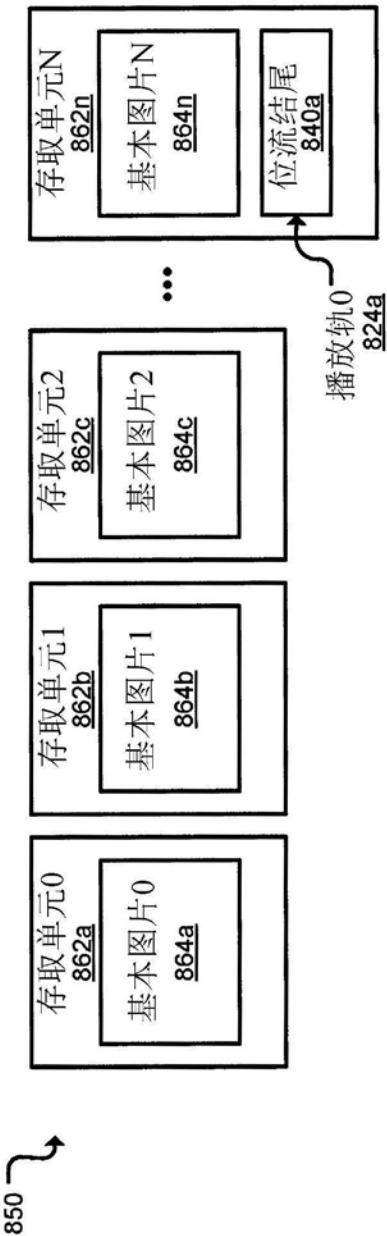


图8B

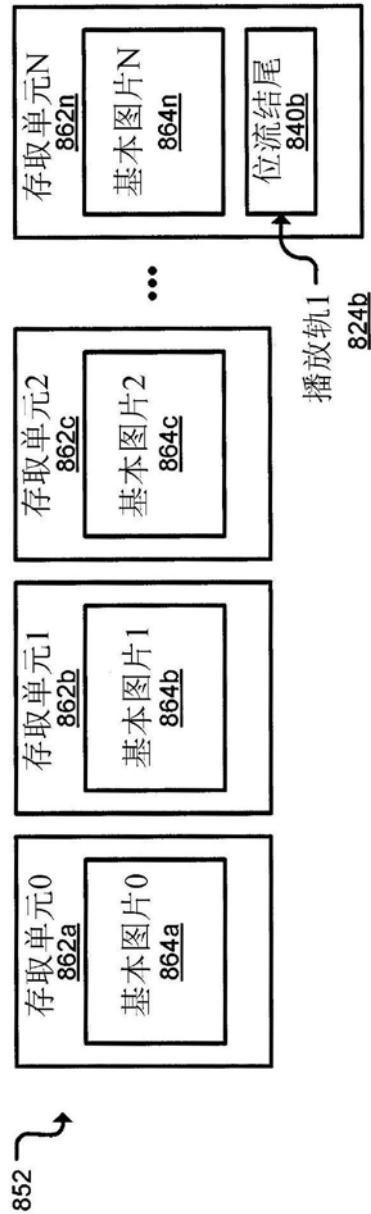


图8C

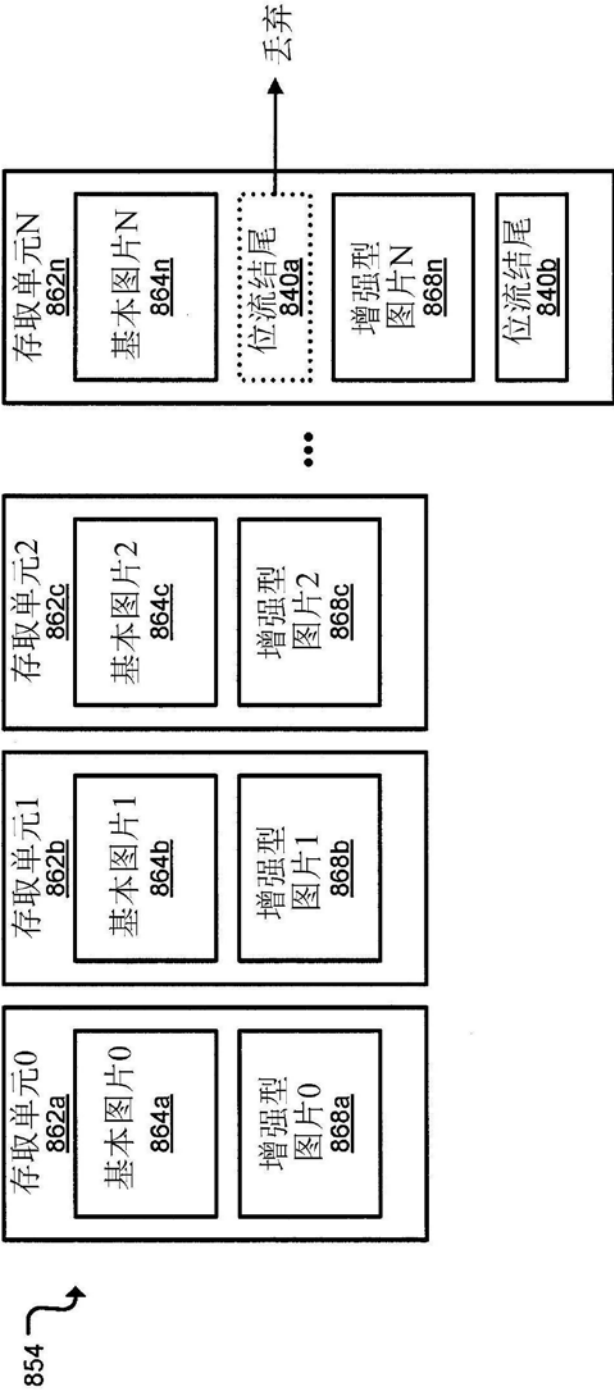


图8D

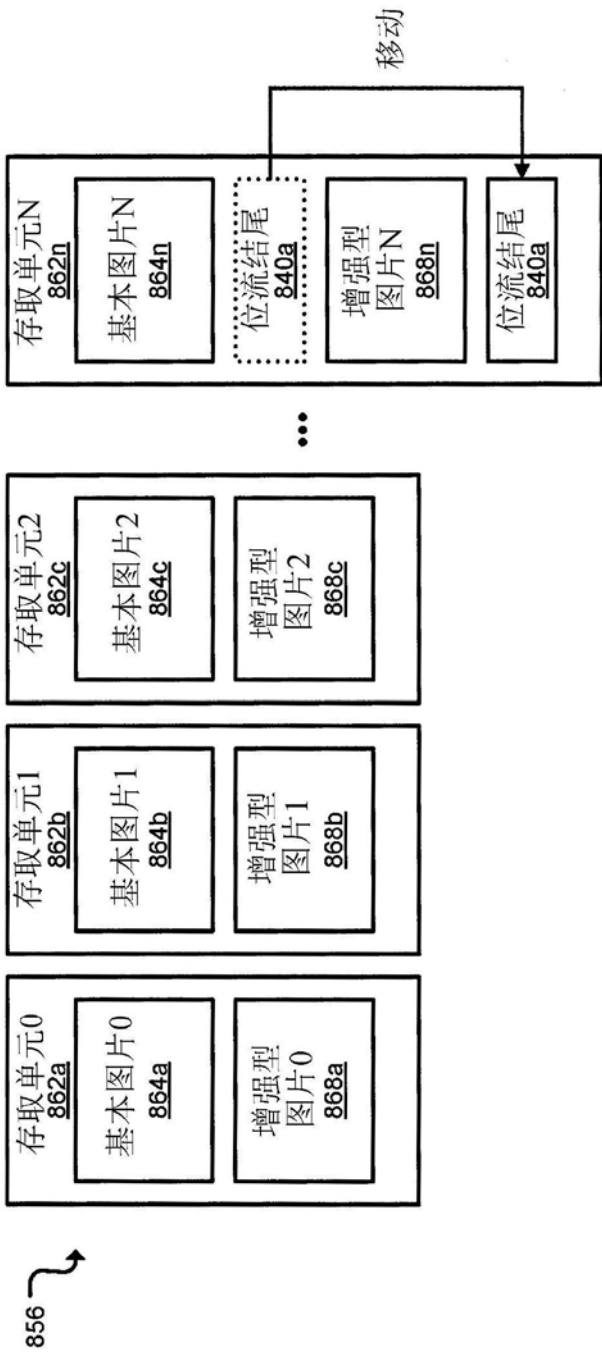


图8E

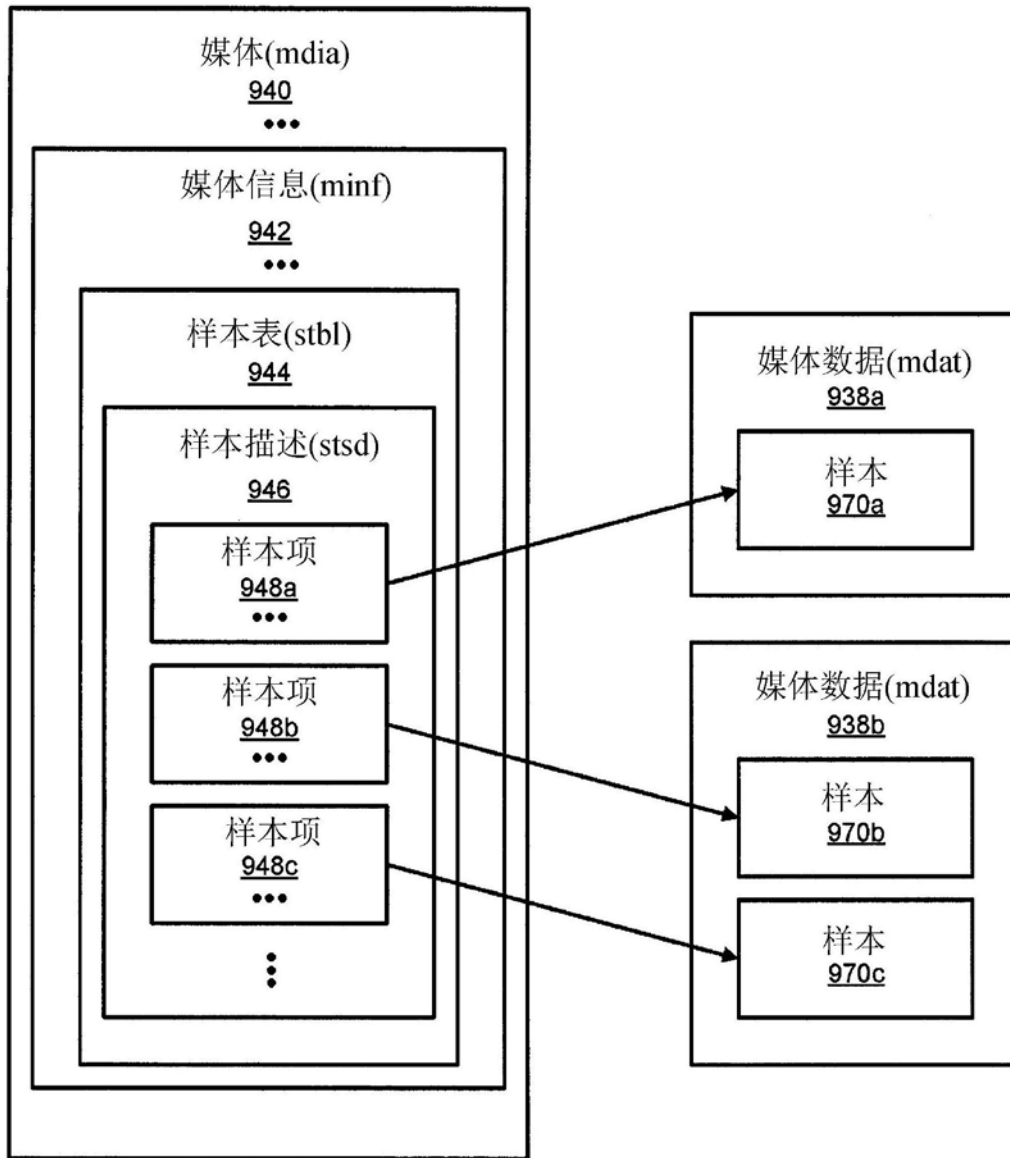


图9

1000

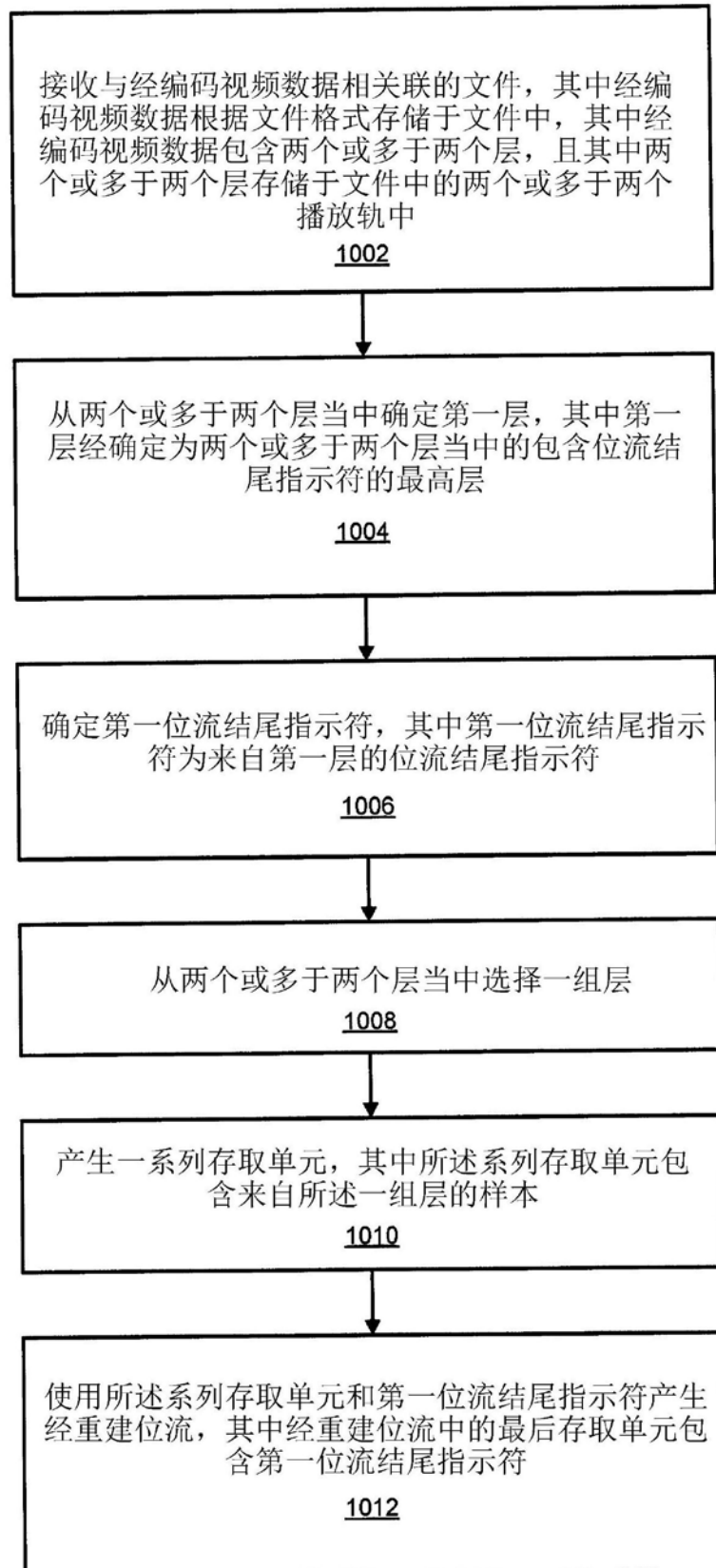


图10

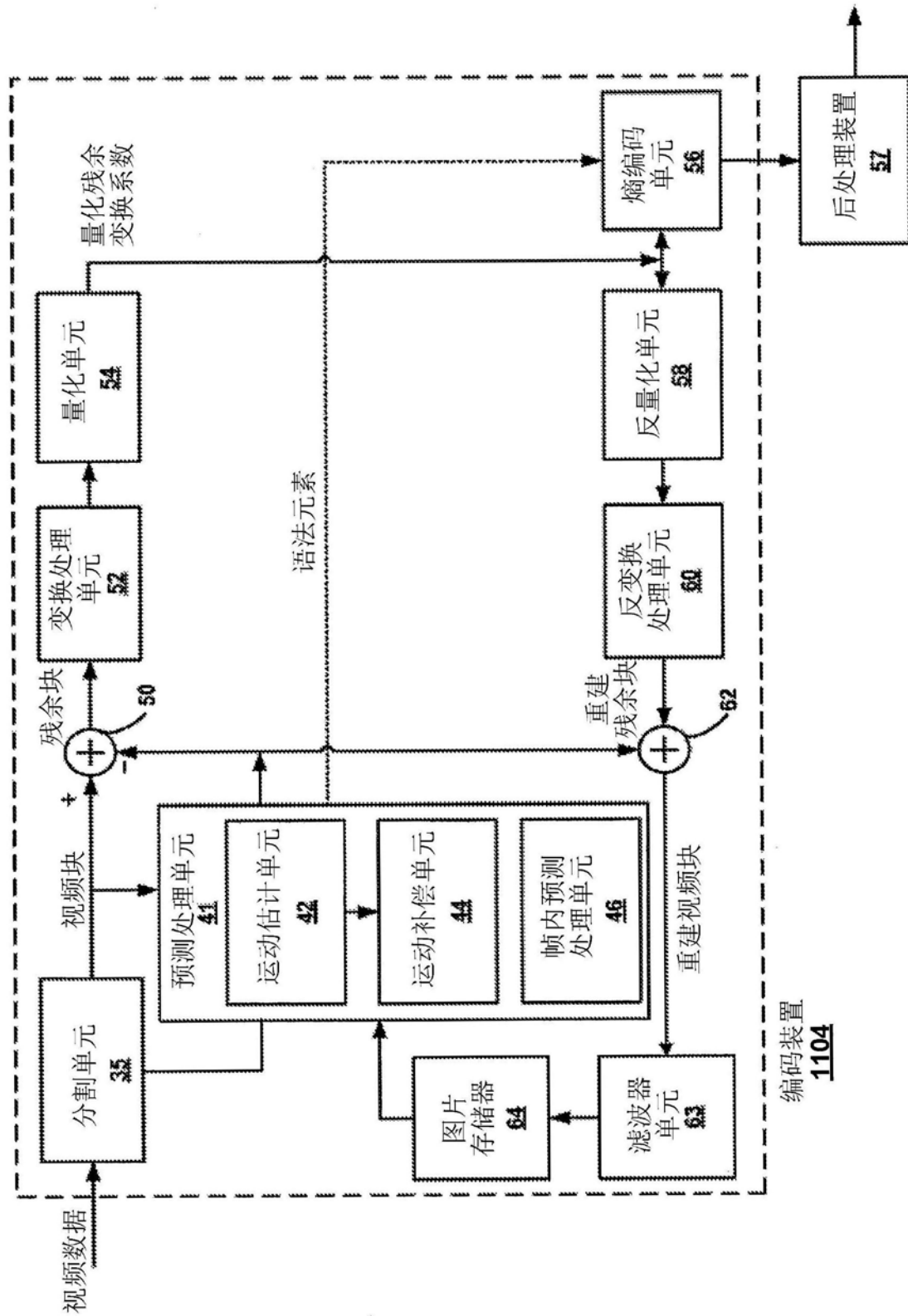


图11

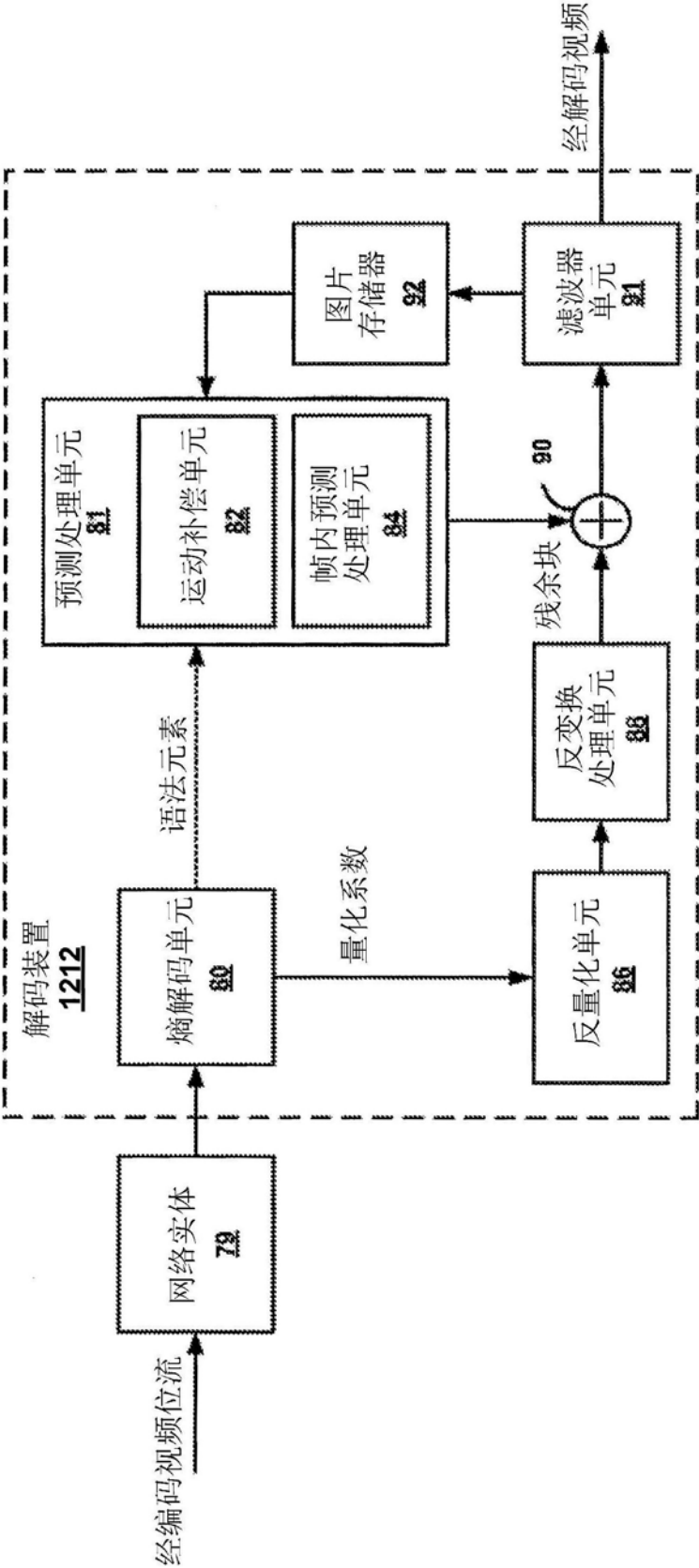


图12