



(12)发明专利

(10)授权公告号 CN 105637885 B

(45)授权公告日 2019.12.20

(21)申请号 201480056622.4

尼安

(22)申请日 2014.10.23

伏努·亨利

(65)同一申请的已公布的文献号

(74)专利代理机构 北京律盟知识产权代理有限公司 11287

申请公布号 CN 105637885 A

代理人 宋献涛

(43)申请公布日 2016.06.01

(51)Int.Cl.

(30)优先权数据

H04N 19/70(2006.01)

61/894,886 2013.10.23 US

H04N 19/30(2006.01)

14/521,099 2014.10.22 US

(85)PCT国际申请进入国家阶段日

(56)对比文件

2016.04.14

CN 103141069 A, 2013.06.05,

(86)PCT国际申请的申请数据

ISO/IEC.Text of ISO/IEC FDIS 14496-

PCT/US2014/061955 2014.10.23

15:2010 3rd edition Carriage of NAL unit
structured video in the ISO BMFF.《MOTION
PICTURE EXPERT GROUP OR ISO/IEC JTC1/
SC29/WG11》.2013,

(87)PCT国际申请的公布数据

M. M. Hannuksela.MV-HEVC/SHVC HLS:

W02015/061561 EN 2015.04.30

Layer-wise startup of the decoding
process.《13.JCT-VC MEETING》.2013,

(73)专利权人 高通股份有限公司

审查员 张真玮

地址 美国加利福尼亚州

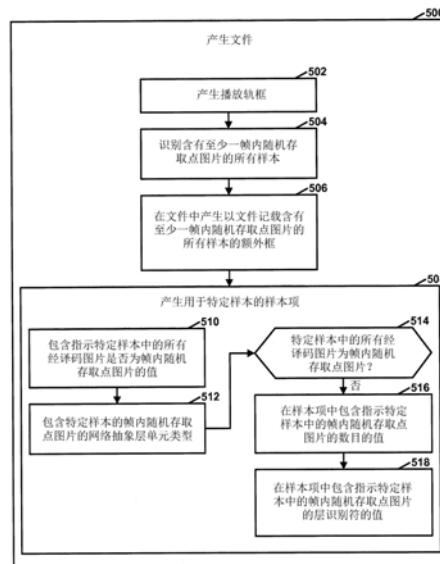
权利要求书6页 说明书54页 附图12页

(54)发明名称

一种处理多层视频数据的方法、视频装置和
计算机可读数据存储媒体

(57)摘要

计算装置产生文件，所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框。用于所述播放轨的媒体数据包括一连串样本。所述样本中的每一者为多层视频数据的视频存取单元。作为产生所述文件的部分，所述计算装置在所述文件中产生以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框。



1. 一种处理多层视频数据的方法,所述方法包括:

产生文件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:

用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的多个视频存取单元中的视频存取单元,且

产生所述文件包括:

在所述文件中定义以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片。

2. 根据权利要求1所述的方法,其中产生所述文件包括:

在所述样本项中包含指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值。

3. 根据权利要求1所述的方法,其中:

产生所述文件包括在所述文件中产生包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框,

所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

4. 根据权利要求1所述的方法,其进一步包括编码所述多层视频数据。

5. 一种处理多层视频数据的方法,所述方法包括:

从文件获得含有用于所述文件中的播放轨的元数据的播放轨框,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的多个视频存取单元中的视频存取单元;

从所述文件获得以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框的定义,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片;及

基于所述额外框确定含有至少一IRAP图片的所述样本中的一或多者。

6.根据权利要求5所述的方法,其包括:

从所述样本项获得指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值。

7.根据权利要求5所述的方法,其进一步包括:

从所述文件获得包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框,所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

8.根据权利要求5所述的方法,其进一步包括在含有至少一IRAP图片的所述一或多个样本中的一者处开始对所述多层视频数据的解码。

9.一种视频装置,其包括:

数据存储媒体,其经配置以存储多层视频数据;及

一或多个处理器,其经配置以:

产生文件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:

用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的多个视频存取单元中的视频存取单元,且

产生所述文件包括在所述文件中定义以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片。

10.根据权利要求9所述的视频装置,其中所述一或多个处理器经配置以:

在所述样本项中包含指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值。

11.根据权利要求9所述的视频装置,其中所述一或多个处理器经配置以在所述文件中产生包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框,

所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片，那么多视图视频译码样本为同步样本。

12. 根据权利要求9所述的视频装置，其中所述一或多个处理器经配置以编码所述多层视频数据。

13. 根据权利要求9所述的视频装置，其中所述装置包含集成电路，微处理器，或无线手持机。

14. 根据权利要求9所述的视频装置，其进一步包括经配置以俘获所述多层视频数据的一或多个相机。

15. 一种视频装置，其包括：

数据存储媒体，其经配置以存储多层视频数据；及

一或多个处理器，其经配置以：

从文件获得含有用于所述文件中的播放轨的元数据的播放轨框，其中用于所述播放轨的媒体数据包括一连串样本，所述样本中的每一者为所述多层视频数据的多个视频存取单元中的视频存取单元；

从所述文件获得以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框的定义，其中：

定义所述额外框以包含样本项，所述样本项包含：

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值，

其中，如果所述特定样本中并非所有经译码图片皆为IRAP图片，所述样本项进一步包含：

指示所述特定样本中的IRAP图片的数目的值，及

指示所述特定样本中的所述IRAP图片的层识别符的值，且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片，且还含有其它层中的一或多个非IRAP图片；及

基于所述额外框中的信息确定含有至少一IRAP图片的所述样本中的一或多者。

16. 根据权利要求15所述的视频装置，其中所述一或多个处理器经配置以：

从所述样本项获得指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值。

17. 根据权利要求15所述的视频装置，其中所述一或多个处理器经配置以：

从所述文件获得包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框，所述播放轨的每一同步样本为所述播放轨的随机存取样本，

如果存取单元中的每一经译码图片为IRAP图片，那么可缩放视频译码样本为同步样本，且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片，那么多视图视频译码样本为同步样本。

18. 根据权利要求15所述的视频装置，其中所述一或多个处理器经配置以在含有至少一IRAP图片的所述一或多个样本中的一者处开始对所述多层视频数据的解码。

19. 根据权利要求15所述的视频装置，其中所述装置包含集成电路，微处理器，或无线手持机。

20. 根据权利要求15所述的视频装置,其进一步包括经配置以显示经解码视频数据的显示器。

21. 一种视频装置,其包括:

用于存储经编码多视图视频数据的装置;及

用于产生文件的装置,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:

用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为多层视频数据的多个视频存取单元中的视频存取单元,及

产生所述文件包括:在所述文件中定义以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片。

22. 根据权利要求21所述的视频装置,其包括:

用于在所述样本项中包含指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值的装置。

23. 根据权利要求21所述的视频装置,其中:

产生所述文件包括在所述文件中产生包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框,

所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

24. 根据权利要求21所述的视频装置,其进一步包括用于编码所述多层视频数据的装置。

25. 一种视频装置,其包括:

用于从文件获得含有用于所述文件中的播放轨的元数据的播放轨框的装置,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为多层视频数据的多个视频存取单元中的视频存取单元;

用于从所述文件获得以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框的定义的装置,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片;及

用于基于所述额外框确定含有至少一IRAP图片的所述样本中的一或者者的装置。

26.根据权利要求25所述的视频装置,其包括:

用于从所述样本项获得指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值的装置。

27.根据权利要求25所述的视频装置,其进一步包括:

用于从所述文件获得包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框的装置,

所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

28.根据权利要求25所述的视频装置,其进一步包括用于在含有至少一IRAP图片的所述一或多个样本中的一者处开始对所述多层视频数据的解码的装置。

29.一种计算机可读数据存储媒体,其具有存储于其上的指令,所述指令当经执行时使一或多个处理器:

产生文件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:

用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为多层视频数据的多个视频存取单元中的视频存取单元,

为了产生所述文件,所述一或多个处理器在所述文件中定义以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片。

30.根据权利要求29所述的计算机可读数据存储媒体,其中所述指令使所述一或多个处理器:

在所述样本项中包含指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值。

31. 根据权利要求29所述的计算机可读数据存储媒体,其中所述指令进一步使所述一或多个处理器:

在所述文件中产生包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框,

所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

32. 一种计算机可读数据存储媒体,其具有存储于其上的指令,所述指令当经执行时使一或多个处理器:

从文件获得含有用于所述文件中的播放轨的元数据的播放轨框,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为多层视频数据的多个视频存取单元中的视频存取单元;

从所述文件获得以文件记载含有至少一帧内随机存取点IRAP图片的所有所述样本的额外框的定义,其中:

定义所述额外框以包含样本项,所述样本项包含:

指示所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值,

其中,如果所述特定样本中并非所有经译码图片皆为IRAP图片,所述样本项进一步包含:

指示所述特定样本中的IRAP图片的数目的值,及

指示所述特定样本中的所述IRAP图片的层识别符的值,且

其中所述所产生的文件中的所述视频存取单元中的至少一者含有基础层中的IRAP图片,且还含有其它层中的一或多个非IRAP图片;及

基于所述额外框确定含有至少一IRAP图片的所述样本中的一或多者。

33. 根据权利要求32所述的计算机可读数据存储媒体,其中所述指令使所述一或多个处理器:

从所述样本项获得指示在所述特定样本的所述IRAP图片中的视频译码层VCL网络抽象层NAL单元的NAL单元类型的值。

34. 根据权利要求32所述的计算机可读数据存储媒体,其中所述指令使所述一或多个处理器:

从所述文件获得包含以文件记载所述播放轨的同步样本的同步样本表的同步样本框,

所述播放轨的每一同步样本为所述播放轨的随机存取样本,

如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且

如果所述存取单元中的每一经译码图片为无随机存取跳过前置RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

一种处理多层视频数据的方法、视频装置和计算机可读数据存储媒体

[0001] 本申请案主张2013年10月23日申请的第61/894,886号美国临时专利申请案的权利,所述申请案的全部内容被以引用的方式并入本文中。

技术领域

[0002] 本发明涉及视频译码。

背景技术

[0003] 数字视频能力可并入到广泛范围的装置中,包含数字电视、数字直播系统、无线广播系统、个人数字助理(PDA)、膝上型或桌上型计算机、平板计算机、电子书阅读器、数字相机、数字记录装置、数字媒体播放器、视频游戏装置、视频游戏控制台、蜂窝式或卫星无线电话(所谓的“智能电话”)、视频电话会议装置、视频流式传输装置及其类似者。数字视频装置实施视频压缩技术,例如由MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4第10部分高级视频译码(AVC)所定义的标准、目前正在开发的高效率视频译码(HEVC)标准及此些标准的扩展中所描述的那些视频压缩技术。视频装置通过实施此些视频压缩技术可更有效地发射、接收、编码、解码及/或存储数字视频信息。

[0004] 视频压缩技术执行空间(图片内)预测及/或时间(图片间)预测来减少或移除视频序列中固有的冗余。对于基于块的视频译码,可将视频切片(即,视频帧或视频帧的一部分)分割成视频块(其也可被称作树块)、译码单元(CU)及/或译码节点。可使用相对于同一图片中的相邻块中的参考样本的空间预测来编码图片的经帧内译码(I)的切片中的视频块。图片的经帧间译码(P或B)切片中的视频块可使用相对于同一图片中的相邻块中的参考样本的空间预测或相对于其它参考图片中的参考样本的时间预测。图片可被称作帧,且参考图片可被称作参考帧。

[0005] 空间或时间预测导致用于待编码的块的预测性块。残余数据表示经译码的原始块与预测性块之间的像素差。根据指向形成预测性块的参考样本的块的运动向量及指示经译码块与预测性块之间的差异的残余数据来编码经帧间译码块。经帧内译码块是根据帧内译码模式及残余数据来编码。为进行进一步压缩,可将残余数据从像素域变换到变换域,从而导致可接着进行量化的残余变换系数。可扫描最初排列成二维阵列的经量化变换系数以便产生变换系数的一维向量,且可应用熵译码以达成甚至较多压缩。

发明内容

[0006] 一般来说,本发明涉及基于国际标准组织(ISO)基本媒体文件格式(ISOBMFF)将视频内容存储于文件中。本发明的一些实例涉及用于存储含有多个经译码层的视频流的方法,其中每一层可为可缩放层、纹理视图、深度视图等,且所述方法可适用于存储多视图高效率视频译码(MV-HEVC)、可缩放HEVC(SHVC)、三维HEVC(3D-HEVC)及其它类型的视频数据。

[0007] 在一个方面中,本发明描述一种处理多层视频数据的方法,所述方法包括:产生文

件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元,且产生所述文件包括:在所述文件中产生以文件记载含有至少一帧内随机存取点(IRAP)图片的所有所述样本的额外框。

[0008] 在另一方面中,本发明描述一种处理多层视频数据的方法,所述方法包括:从文件获得含有用于所述文件中的播放轨的元数据的播放轨框,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元;从所述文件获得以文件记载含有至少一IRAP图片的所有所述样本的额外框;及基于所述额外框中的信息确定含有至少一IRAP图片的所述样本中的一或者者。

[0009] 在另一方面中,本发明描述一种视频装置,所述视频装置包括:数据存储媒体,其经配置以存储多层视频数据;及一或多个处理器,其经配置以:产生文件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元,且产生所述文件包括在所述文件中产生以文件记载含有至少一IRAP图片的所有所述样本的额外框。

[0010] 在另一方面中,本发明描述一种视频装置,所述视频装置包括:数据存储媒体,其经配置以存储多层视频数据;及一或多个处理器,其经配置以从文件获得含有用于所述文件中的播放轨的元数据的播放轨框,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元;从所述文件获得以文件记载含有至少一IRAP图片的所有所述样本的额外框;及基于所述额外框中的信息确定含有至少一IRAP图片的所述样本中的一或者者。

[0011] 在另一方面中,本发明描述一种视频装置,其包括:用于产生文件的装置,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元,且产生所述文件包括:在所述文件中产生以文件记载含有至少一IRAP图片的所有所述样本的额外框。

[0012] 在另一方面中,本发明描述一种视频装置,其包括:用于从文件获得含有用于所述文件中的播放轨的元数据的播放轨框的装置,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元;用于从所述文件获得以文件记载含有至少一IRAP图片的所有所述样本的额外框的装置;及用于基于所述额外框中的信息确定含有至少一IRAP图片的所述样本中的一或者者的装置。

[0013] 在另一方面中,本发明描述一种计算机可读数据存储媒体,其具有存储于其上的指令,所述指令当经执行时使一或多个处理器:产生文件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框,其中:用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元,为了产生所述文件,所述一或多个处理器在所述文件中产生以文件记载含有至少一IRAP图片的所有所述样本的额外框。

[0014] 在另一方面中,本发明描述一种计算机可读数据存储媒体,其具有存储于其上的指令,所述指令当经执行时使一或多个处理器:从文件获得含有用于所述文件中的播放轨的元数据的播放轨框,其中用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为所述多层视频数据的视频存取单元;从所述文件获得以文件记载含有至少一IRAP

图片的所有所述样本的额外框；及基于所述额外框中的信息确定含有至少一IRAP 图片的所述样本中的一或多者。

[0015] 在附图及以下描述中阐明本发明的一或多个实例的细节。其它特征、目标及优势将从描述、图式及权利要求书显而易见。

附图说明

- [0016] 图1为说明可使用本发明中描述的技术的实例视频编码及解码系统的框图。
- [0017] 图2为说明可实施本发明中所描述的技术的实例视频编码器的框图。
- [0018] 图3为说明可实施本发明中所描述的技术的实例视频解码器的框图。
- [0019] 图4为说明形成网络的部分的一组实例装置的框图。
- [0020] 图5为说明根据本发明的一或多个技术的文件的实例结构的概念图。
- [0021] 图6为说明根据本发明的一或多个技术的文件的实例结构的概念图。
- [0022] 图7为说明根据本发明的一或多个技术的文件产生装置的实例操作的流程图。
- [0023] 图8为说明根据本发明的一或多个技术的计算装置执行随机存取及/或等级切换的实例操作的流程图。
- [0024] 图9为说明根据本发明的一或多个技术的文件产生装置的实例操作的流程图。
- [0025] 图10为说明根据本发明的一或多个技术的计算装置的实例操作的流程图。
- [0026] 图11为说明根据本发明的一或多个技术的文件产生装置的实例操作的流程图。
- [0027] 图12为说明根据本发明的一或多个技术的目的地装置的实例操作的流程图。

具体实施方式

[0028] ISO基本媒体文件格式 (ISOBMFF) 为用于存储媒体数据的文件格式。ISOBMFF可扩展以支持符合特定视频译码标准的视频数据的存储。举例来说，ISOBMFF先前已经扩展以支持符合H.264/AVC及高效率视频译码 (HEVC) 视频译码标准的视频数据的存储。此外，ISOBMFF先前已经扩展以支持符合H.264/AVC的多视图译码 (MVC) 及可缩放视频译码 (SVC) 扩展的视频数据的存储。MV-HEVC、3D-HEVC及SHVC为HEVC视频译码标准的支持多层视频数据的扩展。添加到ISOBMFF用于符合H.264/AVC的MVC及 SVC扩展的视频数据的存储的特征不足以用于符合MV-HEVC、3D-HEVC及SHVC的视频数据的有效存储。换句话说，如果想要试图将用于符合H.264/AVC的MVC及SVC 扩展的视频数据的存储的ISOBMFF的扩展用于符合MV-HEVC、3D-HEVC及SHVC的视频数据的有效存储，那么可能出现各种问题。

[0029] 举例来说，不同于符合H.264/AVC的MVC或SVC扩展的位流，符合MV-HEVC、3D-HEVC或SHVC的位流可包含含有帧内随机存取点 (IRAP) 图片及非IRAP图片的存取单元。含有IRAP图片及非IRAP图片的存取单元可用于MV-HEVC、3D-HEVC及SHVC 中的随机存取。然而，ISOBMFF及其现有扩展不提供识别此些存取单元的方式。此可妨碍计算装置执行随机存取及层切换的能力。

[0030] 因此，根据本发明的一个实例，计算装置可产生文件，所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框。用于播放轨的媒体数据包括一连串样本。样本中的每一者可为多层视频数据 (例如，MV-HEVC、3D-HEVC或SHVC视频数据) 的视频存取单元。作为产生文件的部分，计算装置可在文件中产生以文件记载含有至少一IRAP 图片的所有所述

样本的额外框。能够基于额外框中的信息确定含有IRAP图片的样本可使接收文件的计算装置能够在不剖析及解译NAL单元的情况下执行随机存取及层切换。此可减小复杂度且减少处理时间。

[0031] 此外,例如MV-HEVC、3D-HEVC及SHVC视频数据的多层视频数据可包含用于每一存取单元的多个经译码图片。然而,当在存取单元中存在多个经编码图片时,ISOBMFF及其现有扩展不提供关于存取单元内的个别经译码图片的信息。因此,在计算装置(例如,流式传输服务器)正确定是否转递文件中的NAL单元的实例中,计算装置可能需要剖析且解译存储于NAL单元中的信息以便确定是否转递NAL单元。剖析且解译存储于NAL单元中的信息可增大计算装置的复杂度且可增加流式传输延迟。

[0032] 因此,根据本发明的一个实例,计算装置可产生文件,所述文件包括含有用于所述文件中的播放轨的元数据的播放轨框。用于播放轨的媒体数据包括一连串样本。样本中的每一者为多层视频数据的视频存取单元。作为产生文件的部分,计算装置在文件中产生子样本信息框,所述子样本信息框含有指定在所述子样本信息框中给出的子样本信息的类型的旗标。当旗标具有特定值时,对应于子样本信息框的子样本含有正好一个经译码图片及与所述经译码图片相关联的零或多个非视频译码层(VCL)NAL单元。以此方式,接收文件的计算装置可能能够使用在子样本信息框中给出的子样本信息进行关于文件的样本内的个别经译码图片的确定。与经译码图片相关联的非VCL NAL单元可包含用于适用于经译码图片的参数集(例如,PPS、SPS、VPS)及SEI的NAL单元。

[0033] 在多层视频数据中,存取单元可包含标记为用于输出的经译码图片及标记为不用于输出的经译码图片。视频解码器可使用标记为不用于输出的经译码图片作为用于解码标记为用于输出的经译码图片的参考图片。用于图片的切片的NAL单元的NAL单元标头可包含图片输出旗标(例如,HEVC中的pic_output_flag),其指示是否将所述图片标记为用于输出。在ISOBMFF文件中,需要每一样本与指示样本将输出的时间的输出时间(例如,组成时间)相关联。然而,标记为不用于输出的图片不具有输出时间。因此标记为不用于输出的图片的存在可能违反ISOBMFF的此要求,或可能需要非标准暂时解决方案技术。

[0034] 因此,根据本发明的一或多个技术,计算装置可产生文件,所述文件包括围封媒体内容的媒体数据框。媒体内容包括一连串样本。样本中的每一者包括多层视频数据的存取单元。作为产生文件的部分,计算装置可响应于多层视频数据的位流的至少一存取单元包含具有等于第一值(例如,1)的图片输出旗标的经译码图片及具有等于第二值(例如,0)的图片输出旗标的经译码图片,使用至少两个播放轨将位流存储于文件中。对于来自至少两个播放轨的每一相应播放轨,相应播放轨的每一样本中的所有经译码图片具有相同的图片输出旗标值。允许输出具有等于第一值(例如,1)的图片输出旗标的图片,且允许将具有等于第二值(例如,0)的图片输出旗标的图片用作参考图片,但不允许将其输出。至少两个播放轨的使用可解决以上描述的问题,这是因为可对每一播放轨中的每一样本指派恰当输出时间,且视频解码器可不输出不允许输出的含有所述样本的播放轨中的图片。

[0035] 虽然本发明的所述技术的描述中的许多者描述MV-HEVC、3D-HEVC及SHVC,但读者应了解,本发明的所述技术可适用于其它视频译码标准及/或其扩展。

[0036] 图1为说明可使用本发明中描述的技术的实例视频编码及解码系统10的框图。如图1中所展示,系统10包含源装置12,源装置12产生稍后待由目的地装置14解码的经编码视

频数据。源装置12及目的地装置14可包括广泛范围的装置中的任一者,包含桌上型计算机、笔记型(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)、所谓的“智能”板、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类似者。在一些情况下,源装置12和目的地装置14可经装备以用于无线通信。源装置12及目的地装置14可被考虑为视频装置。

[0037] 在图1的实例中,源装置12包含视频源18、视频编码器20及输出接口22。在一些情况下,输出接口22可包含调制器/解调器(调制解调器)及/或发射器。在源装置12中,视频源18可包含例如视频俘获装置(例如,视频相机)、含有先前所俘获的视频的视频存档、从视频内容提供者接收视频的视频馈入接口及/或用于将计算机图形数据产生为源视频的计算机图形系统的源,或此些源的组合。然而,本发明中所描述的技术可大体上适用于视频译码,且可应用于无线及/或有线应用。

[0038] 视频编码器20可编码经俘获、经预俘获或计算机产生的视频。源装置12可经由源装置12的输出接口22将经编码视频数据直接发射到目的地装置14。经编码视频数据也可(或替代地)存储到存储装置33上,用于稍后由目的地装置14或其它装置存取,以用于解码及/或播放。

[0039] 目的地装置14包含输入接口28、视频解码器30及显示装置32。在一些情况下,输入接口28可包含接收器及/或调制解调器。目的地装置14的输入接口28经由链路16接收经编码视频数据。经由链路16传达或在存储装置33上所提供的经编码视频数据可包含由视频编码器20所产生的多种语法元素,其供例如视频解码器30等视频解码器在解码所述视频数据时使用。此些语法元素可与在通信媒体上发射、存储于存储媒体上或存储于文件服务器上的经编码视频数据包含在一起。

[0040] 显示装置32可与目的地装置14集成在一起或在目的地装置14的外部。在一些实例中,目的地装置14可包含集成式显示装置且也可经配置以与外部显示装置介接。在其它实例中,目的地装置14可为显示装置。一般来说,显示装置32向用户显示经解码视频数据,且可包括多种显示装置中的任一者,例如,液晶显示器(LCD)、等离子体显示器、有机发光二极管(OLED)显示器或另一类型的显示装置。

[0041] 视频编码器20及视频解码器30各自可实施为多种合适的编码器电路中的任一者,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。当技术部分以软件实施时,装置可将用于软件的指令存储于合适的非暂时性计算机可读媒体中,且使用一或多个处理器执行硬件中的所述指令,从而执行本发明的技术。视频编码器20及视频解码器30中的每一者可包含在一或多个编码器或解码器中,编码器或解码器中的任一者可集成为相应装置中的组合式编码器/解码器(编解码器)的部分。

[0042] 目的地装置14可经由链路16接收待解码的经编码视频数据。链路16可包括能够将经编码视频数据从源装置12移动到目的地装置14的任何类型的媒体或装置。在一个实例中,链路16可包括使源装置12能够实时将经编码视频数据直接发射到目的地装置14的通信媒体。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发射到目的地装置14。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包的网络(例如,局域网、广域网或例如因特网等

全球网络)的部分。通信媒体可包含路由器、交换器、基站或任何其它可以用于促进从源装置12到目的地装置14的通信的设备。

[0043] 替代地,输出接口22可将经编码数据输出到存储装置33。类似地,输入接口28可存取经编码数据存储装置33。存储装置33可包含多种分布式或局部存取的数据存储媒体中的任何者,例如,硬盘驱动器、蓝光(Blu-ray)光盘、DVD、CD-ROM、快闪存储器、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数字存储媒体。在再一实例中,存储装置33可对应于文件服务器或可固持由源装置12产生的经编码视频的另一中间存储装置。目的地装置14可经由流式传输或下载存取来自存储装置33的经存储的视频数据。文件服务器可为能够存储经编码视频数据并将所述经编码视频数据发射到目的地装置14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、网络附接存储(NAS)装置及本地磁盘驱动器。目的地装置14可经由任何标准数据连接(包含因特网连接)而存取经编码视频数据。此可包含适合于存取存储于文件服务器上的经编码视频数据的无线信道(例如,Wi-Fi连接)、有线连接(例如,DSL、缆线调制解调器等)或两者的结合。经编码视频数据从存储装置33的传输可为流式传输、下载传输或两者的组合。

[0044] 本发明的技术不必限于无线应用或设定。所述技术可适用于支持多种多媒体应用(例如,(例如)经由因特网的空中电视广播、有线电视传输、卫星电视传输、流式视频传输)中的任一者的视频译码、供存储于数据存储媒体上的数字视频的编码、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10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、视频广播及/或视频电话等应用。

[0045] 此外,在图1的实例中,视频译码系统10包含文件产生装置34。文件产生装置34可接收由源装置12产生的经编码视频数据。文件产生装置34可产生包含经编码视频数据的文件。目的地装置14可接收由文件产生装置34产生的文件。在各种实例中,文件产生装置34可包含各种类型的计算装置。举例来说,文件产生装置34可包括媒体感知网络元件(MANE)、服务器计算装置、个人计算装置、专用计算装置、商用计算装置或另一类型的计算装置。在一些实例中,文件产生装置34为内容传递网络的部分。文件产生装置34可经由例如链路16等信道从源装置12接收经编码视频数据。此外,目的地装置14可经由例如链路16等信道从文件产生装置34接收文件。文件产生装置34可被考虑为视频装置。

[0046] 在其它实例中,源装置12或另一计算装置可产生包含经编码视频数据的文件。然而,为了易于解释,本发明将文件产生装置34描述为产生文件。然而,应理解,一般来说,这些描述适用于计算装置。

[0047] 视频编码器20及视频解码器30可根据例如高效率视频译码(HEVC)标准或其扩展等视频压缩标准操作。HEVC标准也可被称作ISO/IEC 23008-2。最近,已由ITU-T视频译码专家组(VCEG)及ISO/IEC动画专家组(MPEG)的视频译码联合协作小组(JCT-VC)完成HEVC的设计。最近的HEVC草案规格且下文被称作HEVC WD可从http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1003-v1.zip中获得。对HEVC的多视图扩展(即,MV-HEVC)也正由JCT-3V开发。题为“MV-HEVC 草案文本5(MV-HEVC Draft Text 5)”且下文被称作MV-HEVC WD5的MV-HEVC的最近工作草案(WD)可从http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1004-v6.zip中获得。对HEVC的可缩放扩展(即,SHVC)也正由JCT-VC 开发。题为“高效率视频译码(HEVC)可缩

放扩展草案3(High efficiency video coding (HEVC) scalable extension draft 3)”且下文被称作SHVC WD3的SHVC的最近工作草案 (WD) 可从http://phenix.it-sudparis.eu/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1008-v3.zip中获得。HEVC的范围扩展的最近工作草案 (WD) 可从http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1005-v3.zip中获得。题为“3D-HEVC草案文本1(3D-HEVC Draft Text 1)”的HEVC的3D扩展的最近工作草案 (WD) (即, 3D-HEVC) 可从http://phenix.int-evry.fr/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1001-v3.zip中获得。视频编码器20及视频解码器30 可根据此些标准中的一或者操作。

[0048] 替代地, 视频编码器20及视频解码器30可根据其它专属或行业标准(例如, ITU-T H.264标准, 替代地被称作MPEG-4, 第10部分, 高级视频译码 (AVC))或此些标准的扩展而操作。然而, 本发明的技术不限于任何特定译码标准。视频压缩标准的其它实例包含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及ITU-T H.264(也称为ISO/IEC MPEG-4AVC), 包含其可缩放视频译码 (SVC) 及多视图视频译码 (MVC) 扩展。

[0049] 尽管图1中未展示, 但在一些方面中, 视频编码器20及视频解码器30可各自与音频编码器及解码器集成, 且可包含适当MUX-DEMUX单元或其它硬件及软件, 以处置共同数据流或单独数据流中的音频及视频两者的编码。如果适用, 那么在一些实例中, MUX-DEMUX单元可符合ITU H.223多路复用器协议或其它协议(例如, 用户数据报协议 (UDP))。

[0050] JCT-VC正致力于HEVC标准的开发。HEVC标准化努力是基于视频译码装置的演进型模型(被称作HEVC测试模型 (HM))。HM根据(例如)ITU-T H.264/AVC假定视频译码装置相对于现有装置的若干额外能力。举例来说, 尽管H.264/AVC提供九个帧内预测编码模式, 但HM可提供多达三十三个帧内预测编码模式。

[0051] 一般来说, HM的工作模型描述视频帧或图片可划分成包含明度样本及色度样本两者的树块或最大译码单元 (LCU) 的序列。树块也可被称作译码树单元 (CTU)。树块具有与H.264/AVC标准的宏块类似的用途。切片包含按译码次序的许多连续树块。视频帧或图片可分割成一或多个切片。每一树块可根据四分树而分裂成若干译码单元 (CU)。举例来说, 作为四分树的根节点的树块可分裂成四个子节点, 且每一子节点又可为母节点并分裂成另四个子节点。作为四分树的叶节点的最后未分裂的子节点包括译码节点(即, 经译码视频块)。与经译码位流相关联的语法数据可定义树块可分裂的最大次数, 且也可定义译码节点的最小大小。

[0052] CU包含译码节点以及与所述译码节点相关联的预测单元 (PU) 及变换单元 (TU)。CU的大小对应于译码节点的大小且形状必须为正方形。CU的大小可在从8x8像素高达具有最大64x64像素或大于64x64像素的树块的大小的范围内。每一CU可含有一或多个 PU及一或多个TU。与CU相关联的语法数据可描述(例如) CU到一或多个PU的分割。分割模式可在CU经跳过或直接模式编码、帧内预测模式编码或是帧间预测模式编码之间不同。PU可经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) CU根据四分树到一或多个TU的分割。TU的形状可为正方形或非正方形。

[0053] HEVC标准允许根据TU进行变换, 所述变换对于不同CU可不同。通常基于针对经分割LCU所定义的给定CU内的PU的大小来对TU设定大小, 但情况可能并非始终如此。TU的大小

通常与PU相同或比PU小。在一些实例中,可使用被称为“残余四分树”(RQT)的四分树结构而将对应于CU的残余样本再分为较小单元。RQT的叶节点可被称作TU。与TU相关联的像素差值可经变换以产生可加以量化的变换系数。

[0054] 一般来说,PU包含与预测过程有关的数据。举例来说,当PU经帧内模式编码时,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PU经帧间模式编码时,PU可包含定义PU的运动向量的数据。定义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及/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] 一般来说,TU用于变换及量化过程。具有一或多个PU的给定CU也可包含一或多个变换单元(TU)。在预测之后,视频编码器20可计算对应于PU的残余值。残余值包含像素差值,所述像素差值可转换成变换系数、经量化且使用TU进行扫描以产生串行化变换系数用于熵译码。本发明通常使用术语“视频块”来指CU的译码节点(即,译码块)。在一些特定情况下,本发明也可使用术语“视频块”指树块(即,LCU)或CU,其包含译码节点及PU与TU。

[0056] 视频序列通常包含一系列视频帧或图片。图片群组(GOP)大体上包括一系列视频图片中的一或多者。GOP可包含GOP的标头、图片中的一或多者的标头或别处的语法数据,所述语法数据描述包含于GOP中的图片的数目。图片的每一切片可包含描述所述相应切片的编码模式的切片语法数据。视频编码器20通常对个别视频切片内的视频块进行操作,以便编码视频数据。视频块可对应于CU内的译码节点。视频块可具有固定或变化的大小,且可根据指定译码标准而大小不同。

[0057] 作为一实例,HM支持以各种PU大小的预测。假定特定CU的大小为 $2Nx2N$,那么HM支持以 $2Nx2N$ 或 NxN 的PU大小的帧内预测,及以 $2Nx2N$ 、 $2NxN$ 、 $Nx2N$ 或 NxN 的对称PU大小的帧间预测。HM也支持以 $2NxN$ 、 $2NxN$ 、 $Nx2N$ 及 NxN 的PU大小的帧间预测的不对称分割。在不对称分割中,CU的一个方向未分割,而另一方向分割成25%及75%。CU的对应于25%分割的部分由“n”其后接着“上(Up)”、“下(Down)”、“左(Left)”或“右(Right)”的指示来指示。因此,举例来说,“ $2NxN$ ”指水平上以顶部的 $2Nx0.5N$ PU及底部的 $2Nx1.5N$ PU分割的 $2Nx2N$ CU。

[0058] 在本发明中,“ NxN ”与“N乘N”可互换地使用以指视频块在垂直维度与水平维度方面的像素尺寸,例如,16x16像素或16乘16像素。一般来说,16x16块在垂直方向上具有16个像素($y=16$)且在水平方向上将具有16个像素($x=16$)。同样地, NxN 块通常在垂直方向上具有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。可按行及列来排列块中的像素。此外,块未必需要在水平方向上与垂直方向上具有相同数目个像素。举例来说,块可包括 NxM 像素,其中M未必等于N。

[0059] 在使用CU的PU的帧内预测性或帧间预测性译码之后,视频编码器20可计算CU的TU的残余数据。PU可包括空间域中的像素数据(也被称作像素域),且TU可包括在将变换(例如,离散余弦变换(DCT)、整数变换、小波变换或概念上类似的变换)应用于残余视频数据之后的变换域中的系数。所述残余数据可对应于未经编码的图片的像素与对应于PU的预测值之间的像素差。视频编码器20可形成包含CU的残余数据的TU,且接着变换所述TU以产生CU的变换系数。

[0060] 在进行用以产生变换系数的任何变换之后,视频编码器20可对变换系数执行量

化。量化通常指将变换系数量化以可能地减少用以表示所述系数的数据的量,从而提供进一步压缩的过程。量化过程可减少与系数中的一些或所有相关联的位深度。举例来说,可在量化期间将n位值降值舍位到m位值,其中n大于m。

[0061] 在一些实例中,视频编码器20可使用预定义扫描次序来扫描经量化变换系数以产生可经熵编码的经串行化向量。在其它实例中,视频编码器20可执行自适应扫描。在扫描经量化变换系数以形成一维向量之后,视频编码器20可(例如)根据上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法来熵编码一维向量。视频编码器20也可熵编码与经编码的视频数据相关联的供由视频解码器30在解码视频数据过程中使用的语法元素。

[0062] 为了执行CABAC,视频编码器20可将上下文模型内的上下文指派到待发射的符号。所述上下文可与(例如)符号的邻近值是否为非零有关。为了执行CAVLC,视频编码器20可选择用于待发射的符号的可变长度码。可变长度译码(VLC)中的码字可经建构使得相对较短代码对应于更可能的符号,而较长代码对应于较不可能的符号。以此方式,相对于(例如)针对待发射的每一符号使用相等长度码字,使用VLC可达成位节省。概率确定可基于指派给符号的上下文而进行。

[0063] 视频编码器20可输出包含形成经译码图片及相关联数据的表示的位序列的位流。术语“位流”可为用以指网络抽象层(NAL)单元流(例如,一连串NAL单元)或字节流(例如,含有开始码首码的NAL单元流及如由HEVC标准的附录B指定的NAL单元的封装)的集合性术语。NAL单元为含有NAL单元中的数据的类型的指示及含有所述数据的呈按需要穿插有仿真阻止位的原始字节序列有效负载(RBSP)的形式的字节的语法结构。NAL单元中的每一者可包含NAL单元标头且可封装RBSP。NAL单元标头可包含指示NAL单元类型码的语法元素。由NAL单元的NAL单元标头指定的NAL单元类型码指示NAL单元的类型。RBSP可为含有封装在NAL单元内的整数数目个字节的语法结构。在一些情况下,RBSP包含零个位。

[0064] 不同类型的NAL单元可封装不同类型的RBSP。举例来说,第一类型的NAL单元可封装图片参数集(PPS)的RBSP,第二类型的NAL单元可封装切片片段的RBSP,第三类型的NAL单元可封装补充增强信息(SEI)的RBSP,等等。封装视频译码数据的RBSP(如与参数集及SEI消息的RBSP相对)的NAL单元可被称作视频译码层(VCL)NAL单元。含有参数集(例如,视频参数集(VPS)、序列参数集(SPS)、PPS等)的NAL单元可被称作参数集NAL单元。

[0065] 本发明可将封装片段切片的RBSP的NAL单元称作经译码切片NAL单元。如在HEVC WD中所定义,切片片段为在平铺块扫描中连续定序且在单一NAL单元中含有的整数数目个CTU。相比之下,在HEVC WD中,切片可为一独立切片片段及(如果有)在同一存取单元内下一个独立切片片段(如果有)前的所有随后附属切片片段中含有的整数数目个CTU。独立切片片段为切片片段标头的语法元素的值并非从先前切片片段的值推断的切片片段。附属切片片段为切片片段标头的一些语法元素的值从按解码次序先前独立切片片段的值推断的切片片段。经译码切片NAL单元的RBSP可包含切片片段标头及切片数据。切片片段标头为含有涉及在切片片段中表示的第一或所有CTU的数据元素的经译码切片片段的一部分。切片标头为独立切片片段的切片片段标头,所述独立切片区是当前切片片段或按解码次序在当前附属切片片段前最近的独立切片片段。

[0066] VPS为含有适用于零或多个全部经译码视频序列 (CVS) 的语法元素的语法结构。SPS 为含有适用于零或多个全部CVS的语法元素的语法结构。SPS可包含识别在SPS在作用中时在作用中的VPS的语法元素。因此, VPS的语法元素可比SPS的语法元素更一般化地可适用。

[0067] 参数集(例如,VPS、SPS、PPS等)可含有直接或间接从切片的切片标头参考的识别。参考过程被称为“启动”。因此,当视频解码器30正解码特定切片时,由所述特定切片的切片标头中的语法元素直接或间接参考的参数集被称为“经启动”。取决于参数集类型,启动可基于每一图片或基于每一序列发生。举例来说,切片的切片标头可包含识别 PPS的语法元素。因此,当视频译码器译码切片时,可启动PPS。此外,PPS可包含识别SPS的语法元素。因此,当识别SPS的PPS经启动时,可启动SPS。SPS可包含识别 VPS的语法元素。因此,当识别VPS的SPS经启动时,启动VPS。

[0068] 视频解码器30可接收由视频编码器20产生的位流。此外,视频解码器30可剖析所述位流以从所述位流获得语法元素。视频解码器30可至少部分基于从位流获得的语法元素重建构视频数据的图片。重建构视频数据的过程可与由视频编码器20执行的过程大体上互逆。举例来说,视频解码器30可使用PU的运动向量确定当前CU的PU的预测性块。此外,视频解码器30可反量化当前CU的TU的系数块。视频解码器30可对系数块执行反变换,以重建构当前CU的TU的变换块。通过将当前CU的PU的预测性块的样本添加到当前CU的TU的变换块的对应样本,视频解码器30可重建构当前 CU的译码块。通过重建构图片的每一CU的译码块,视频解码器30可重建构图片。

[0069] 在HEVC WD中,CVS可开始于瞬时解码刷新 (IDR) 图片,或断链存取 (BLA) 图片,或为位流中的第一图片的清洁随机存取 (CRA) 图片,包含并非IDR或BLA图片的所有随后图片。IDR图片仅含有I切片(即,仅使用帧内预测的切片)。IDR图片可为按解码次序在位流中的第一图片,或可稍后出现在位流中。每一IDR图片为按解码次序CVS 的第一图片。在HEVC WD中, IDR图片可为每一VCL NAL单元具有等于IDR_W_RADL 或IDR_N_LP的nal_unit_type的帧内随机存取点 (IRAP) 图片。

[0070] IDR图片可用于随机存取。然而,按解码次序在IDR图片后的图片不能使用在IDR图片前解码的图片作为参考。因此,依赖于供随机存取的IDR图片的位流可具有比使用额外类型的随机存取图片的位流显著低的译码效率。在至少一些实例中, IDR存取单元为含有IDR图片的存取单元。

[0071] 在HEVC中引入CRA图片的概念以允许按解码次序在CRA图片之后但按输出次序在CRA图片之前的图片将在所述CRA图片前解码的图片用于参考。按解码次序在CRA 图片之后但按输出次序在CRA图片之前的图片被称作与CRA图片相关联的前置图片 (或CRA图片的前置图片)。即,为了改进译码效率,在HEVC中引入CRA图片的概念,以允许按解码次序在CRA图片之后但按输出次序在CRA图片之前的图片将在CRA图片前解码的图片用于参考。CRA存取单元为经译码图片为CRA图片的存取单元。在 HEVC WD中,CRA图片为每一VCL NAL单元具有等于CRA_NUT的nal_unit_type的帧内随机存取图片。

[0072] CRA图片的前置图片在解码从IDR图片或按解码次序在所述CRA图片前出现的 CRA 图片开始的情况下可正确地解码。然而,在发生从CRA图片的随机存取时,CRA 图片的前置图片可为不可解码的。因此,视频解码器在随机存取解码期间通常解码CRA 图片的前置图

片。为防止从取决于解码开始处而可能不可用的参考图片的误差传播,按解码次序及输出次序两者在CRA图片后的图片无一者可将按解码次序或输出次序在CRA图片前的任何图片(其包含前置图片)用于参考。

[0073] BLA图片的概念是在引入了CRA图片后在HEVC中引入的,且是基于CRA图片的概念。BLA图片通常源自于在CRA图片的位置处拼接的位流,且在所述拼接的位流中,将所述拼接点CRA图片改变到BLA图片。因此,BLA图片可在原始位流处的CRA图片,且CRA图片由位流拼接器改变为在所述CRA图片的位置处的位流拼接后的BLA图片。在一些情况下,含有RAP图片的存取单元可在本文中被称作RAP存取单元。BLA存取单元为含有BLA图片的存取单元。在HEVC WD中,BLA图片可为帧内随机存取图片,对于所述帧内随机存取图片,每一VCL NAL单元具有等于BLA_W_LP、BLA_W_RADL或BLA_N_LP的nal_unit_type。

[0074] 一般来说,IRAP图片含有仅I切片,且可为BLA图片、CRA图片或IDR图片。举例来说,HEVC WD指示IRAP图片可为每一VCL NAL单元具有在BLA_W_LP到RSV_IRAP_VCL23的范围内(包含性)的nal_unit_type的经译码图片。此外,HEVC WD指示按解码次序在位流的第一图片必须为IRAP图片。HEVC WD的表7-1展示NAL单元类型码及NAL单元类型类别。以下再现HEVC WD的表7-1。

[0075] 表7-1-NAL单元类型码及NAL单元类型类别

[0076]

nal_unit_type	nal_unit_type 的名称	NAL 单元的内容及 RBSP 语法结构	NAL 单元类型类别
0 1	TRAIL_N TRAIL_R	非 TSA、非 STSA 后置图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
2 3	TSA_N TSA_R	TSA 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
4 5	STSA_N STSA_R	STSA 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
6 7	RADL_N RADL_R	RADL 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
8 9	RASL_N RASL_R	RASL 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
10 12	RSV_VCL_N10 RSV_VCL_N12	保留的非IRAP子层非参考VCL NAL单元类型	VCL

[0077]

14	RSV_VCL_N14		
11	RSV_VCL_R11		
13	RSV_VCL_R13	保留的非 IRAP 子层参考 VCL NAL 单元类型	VCL
15	RSV_VCL_R15		
16	BLA_W_LP		
17	BLA_W_RADL	BLA 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
18	BLA_N_LP		
19	IDR_W_RADL	IDR 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
20	IDR_N_LP		
21	CRA_NUT	CRA 图片的经译码切片片段 slice_segment_layer_rbsp()	VCL
22	RSV_IRAP_VCL22	保留的 IRAP VCL NAL 单元类型	VCL
23	RSV_IRAP_VCL23		
24..31	RSV_VCL24.. RSV_VCL31	保留的非 IRAP VCL NAL 单元类型	VCL
32	VPS_NUT	视频参数集 video_parameter_set_rbsp()	非 VCL
33	SPS_NUT	序列参数集 seq_parameter_set_rbsp()	非 VCL
34	PPS_NUT	图片参数集 pic_parameter_set_rbsp()	非 VCL
35	AUD_NUT	存取单元定界符 access_unit_delimiter_rbsp()	非 VCL
36	EOS_NUT	序列的末尾 end_of_seq_rbsp()	非 VCL
37	EOB_NUT	位流的末尾 end_of_bitstream_rbsp()	非 VCL
38	FD_NUT	填充数据 filler_data_rbsp()	非 VCL
39	PREFIX_SEI_NUT	补充增强信息 sei_rbsp()	非 VCL
40	SUFFIX_SEI_NUT		
41..47	RSV_NVCL41.. RSV_NVCL47	保留	非 VCL
48..63	UNSPEC48.. UNSPEC63	未指定	非 VCL

[0078] BLA图片与CRA图片之间的一差异如下。对于CRA图片,如果解码开始于按解码次序在CRA图片前的RAP图片,那么相关联的前置图片可正确地解码。然而,当发生从CRA图片的随机存取时(即,当解码从所述CRA图片开始时,或换句话说,当所述CRA图片为位流中的第一图片时),与所述CRA图片相关联的前置图片不可正确地解码。相比之下,可能不存在与BLA图片相关联的前置图片可解码的情形,甚至当解码从按解码次序在BLA图片前的RAP图片开始时。

[0079] 与特定CRA图片或特定BLA图片相关联的前置图片中的一些可正确地解码,甚至当所述特定CRA图片或所述特定BLA图片为位流中的第一图片时。这些前置图片可被称作可解码前置图片(DLP)或随机存取可解码前置(RADL)图片。在HEVC WD中,RADL图片可为每一VCL NAL单元具有等于RADL_R或RADL_N的nal_unit_type的经译码图片。此外,HEVC WD指示

所有RADL图片为前置图片且不将RADL图片用作用于同一相关联的IRAP图片的后置图片的解码过程的参考图片。当存在时,所有RADL图片按解码次序先于同一相关联的IRAP图片的所有后置图片。HEVC WD指示RADL存取单元可为经译码图片为RADL图片的存取单元。后置图片可为按输出次序在相关联的IRAP图片后的图片(即,按解码次序的先前IRAP图片)。

[0080] 其它前置图片可被称作不可解码前置图片(NLP)或随机存取跳过前置(RASL)图片。在HEVC WD中,RASL图片可为每一VCL NAL单元具有等于RASL_R或RASL_N的nal_unit_type的经译码图片。所有RASL图片皆为相关联的BLA图片或CRA图片的前置图片。

[0081] 假设必要参数集在其需要启动时可用,那么IRAP图片及按解码次序所有随后非RASL图片可正确地解码,而不执行按解码次序在IRAP图片前的任何图片的解码过程。在位流中可存在仅含有并非IRAP图片的I切片的图片。

[0082] 在多视图译码中,可存在来自不同视点的同一场景的多个视图。术语“存取单元”可以用以指对应于相同时间实例的图片集。因此,视频数据可经概念化为一系列随时间过去而出现的存取单元。“视图分量”可为单一存取单元中的视图的经译码表示。在本发明中,“视图”可指与相同视图识别符相关联的一连串或一组视图分量。视图分量可含有纹理视图分量及深度视图分量。在本发明中,“视图”可指与相同视图识别符相关联的一组或一连串一或多个视图分量。

[0083] 纹理视图分量(即,纹理图片)可为单一存取单元中的视图的纹理的经译码表示。纹理视图可为与视图次序索引的相同值相关联的一连串纹理视图分量。视图的视图次序索引可指示所述视图相对于其它视图的相机位置。深度视图分量(即,深度图片)可为单一存取单元中的视图的深度的经译码表示。深度视图可为与视图次序索引的相同值相关联的一组或一连串一或多个深度视图分量。

[0084] 在MV-HEVC、3D-HEVC及SHVC中,视频编码器可产生包括一系列NAL单元的位流。位流的不同NAL单元可与位流的不同层相关联。可将层定义为具有相同层识别符的VCL NAL单元及相关联的非VCL NAL单元的集合。层可等效于多视图视频译码中的视图。在多视图视频译码中,层可含有具有不同时间实例的相同层的所有视图分量。每一视图分量可为属于特定时间实例的特定视图的视频场景的经译码图片。在3D视频译码的一些实例中,层可含有特定视图的所有经译码深层图片或特定视图的经译码纹理图片。在3D视频译码的其它实例中,层可含有特定视图的纹理视图分量及深度视图分量两者。类似地,在可缩放视频译码的上下文中,层通常对应于具有不同于其它层中的经译码图片的视频特性的经译码图片。这些视频特性通常包含空间分辨率及质量等级(例如,信噪比)。在HEVC及其扩展中,通过将具有特定时间等级的图片的群组定义为子层,可在一层内获得时间可缩放性。

[0085] 对于位流的每一相应层,可在不参考任何较高层中的数据的情况下解码较低层中的数据。在可缩放视频译码中,例如,可在不参考增强层中的数据的情况下解码基础层中的数据。一般来说,NAL单元可仅封装单一层的数据。因此,可从位流移除封装位流的最高剩余层的数据的NAL单元,而不影响位流的剩余层中的数据的可解码性。在多视图译码及3D-HEVC中,较高层可包含额外视图分量。在SHVC中,较高层可包含信噪比(SNR)增强数据、空间增强数据及/或时间增强数据。在MV-HEVC、3D-HEVC及SHVC中,如果视频解码器可在不参考任何其它层的数据的情况下解码层中的图片,那么所述层可被称作“基础层”。基础层可符合HEVC基础规格(例如,HEVC WD)。

[0086] 在SVC中,不同于基础层的层可被称作“增强层”,且可提供增强从位流解码的视频数据的视觉质量的信息。SVC可增强空间分辨率、信噪比(即,质量)或时间速率。在可缩放视频译码(例如,SHVC)中,“层表示”可为单一存取单元中的空间层的经译码表示。为了易于解释,本发明可将视图分量及/或层表示称作“视图分量/层表示”或简单地称作“图片”。

[0087] 为了实施所述层,NAL单元的标头可包含nuh_reserved_zero_6bits语法元素。在HEVC WD中,保留nuh_reserved_zero_6bits语法元素。然而,在MV-HEVC、3D-HEVC 及SVC中,nuh_reserved_zero_6bits语法元素被称作nuh_layer_id语法元素。nuh_layer_id 语法元素指定层的识别符。位流的具有指定不同值的nuh_layer_id语法元素的NAL单元属于位流的不同层。

[0088] 在一些实例中,如果NAL单元涉及多视图译码(例如,MV-HEVC)、3DV译码(例如,3D-HEVC)或可缩放视频译码(例如,SHVC)中的基础层,那么所述NAL单元的 nuh_layer_id 语法元素等于0。可在不参考位流的任何其它层中的数据的情况下解码位流的基础层中的数据。如果NAL单元与多视图译码、3DV或可缩放视频译码中的基础层不相关,那么所述NAL单元的nuh_layer_id语法元素可具有非零值。

[0089] 此外,在层内的一些视图分量/层表示可在不参考同一层内的其它视图分量/层表示的情况下加以解码。因此,封装层的某些视图分量/层表示的数据的NAL单元可从位流移除,而不影响所述层中的其它视图分量/层表示的可解码性。移除封装此些视图分量/ 层表示的数据的NAL单元可减小位流的帧速率。可在不参考在层内的其它视图分量/层表示的情况下解码的在所述层内的视图分量/层表示的子集可在本文中被称作“子层”或“时间子层”。

[0090] NAL单元可包含指定NAL单元的时间识别符(即,Temporal IDs)的temporal_id语法元素。NAL单元的时间识别符识别NAL单元属于的子层。因此,位流的每一子层可具有不同时间识别符。一般来说,如果层的第一NAL单元的时间识别符小于同一层的第二NAL单元的时间识别符,那么可在不参考由第二NAL单元封装的数据的情况下解码由第一NAL单元封装的数据。

[0091] 位流可与多个操作点相关联。位流的每一操作点与一组层识别符(例如,一组nuh_layer_id值)及时间识别符相关联。可将所述组层识别符表示为0pLayerIdSet,且可将时间识别符表示为TemporalID。如果NAL单元的层识别符在操作点的层识别符集合中且NAL单元的时间识别符小于或等于操作点的时间识别符,那么所述NAL单元与所述操作点相关联。因此,操作点可对应于所述位流中的NAL单元的子集。

[0092] 如上所介绍,本发明涉及基于ISO基本媒体文件格式(ISOBMFF)将视频内容存储于文件中。详言地说,本发明描述用于含有多个经译码层的视频流的存储的各种技术,其中每一层可为可缩放层、纹理视图、深度视图或其它类型的层或视图。本发明的所述技术可应用于(例如)MV-HEVC视频数据、SHVC视频数据、3D-HEVC视频数据及/或其它类型的视频数据的存储。

[0093] 现将简要地论述文件格式及文件格式标准。文件格式标准包含ISO基本媒体文件格式(ISOBMFF、ISO/IEC 14496-12,下文“ISO/IEC 14996-12”)及从ISOBMFF导出的其它文件格式标准,包含MPEG-4文件格式(ISO/IEC 14496-14)、3GPP文件格式(3GPP TS 26.244)及AVC文件格式(ISO/IEC 14496-15,下文“ISO/IEC 14996-15”)。因此,ISO/IEC 14496-12

指定ISO基本媒体文件格式。其它文件针对特定应用扩展ISO基本媒体文件格式。举例来说，ISO/IEC 14496-15描述呈ISO基本媒体文件格式的NAL单元结构化视频的载运。H.264/AVC及HEVC以及其扩展为NAL单元结构化视频的实例。ISO/IEC 14496-15包含描述H.264/AVC NAL单元的载运的章节。另外，ISO/IEC 14496-15的第8节描述HEVC NAL单元的载运。

[0094] 将ISOBMFF用作用于许多编解码器封装格式(例如，AVC文件格式)以及用于许多多媒体容器格式(例如，MPEG-4文件格式、3GPP文件格式(3GP)及DVB文件格式)的基础。除例如音频及视频等连续媒体之外，例如图像等静态媒体以及元数据可存储于符合ISOBMFF的文件中。根据ISOBMFF结构化的文件可用于许多用途，包含局部媒体文件播放、远程文件的逐渐下载、用于经由HTTP的动态自适应流式传输(DASH)的片段、用于待流式传输的内容及其包化指令的容器及接收的实时媒体流的记录。因此，虽然原先针对存储而设计，但ISOBMFF已证明用于流式传输(例如，用于逐渐下载或DASH)的价值。为了流式传输目的，可使用在ISOBMFF中定义的电影片段。

[0095] 符合HEVC文件格式的文件可包括一系列叫作框的对象。框可为由唯一类型识别符及长度定义的面向对象建构块。举例来说，框可为ISOBMFF中的基本语法结构，包含四字符译码框类型、框的字节计数及有效负载。换句话说，框可为包括经译码框类型、框的字节计数及有效负载的语法结构。在一些情况下，在符合HEVC文件格式的文件中的所有数据可包括于框内，且在并不在框中的文件中可能不存在数据。因此，ISOBMFF文件可由一连串框组成，且框可含有其它框。举例来说，框的有效负载可包含一或多个额外框。在本发明中其它处详细描述的图5及图6展示根据本发明的一或多个技术的在文件内的实例框。

[0096] 符合ISOBMFF的文件可包含各种类型的框。举例来说，符合ISOBMFF的文件可包含文件类型框、媒体数据框、电影框、电影片段框等等。在此实例中，文件类型框包含文件类型及兼容性信息。媒体数据框可含有样本(例如，经译码图片)。电影框(“moov”)含有用于存在于文件中的连续媒体流的元数据。可将连续媒体流中的每一者在文件中表示为播放轨。举例来说，电影框可含有关于电影的元数据(例如，样本之间的逻辑及时序关系，还有指向样本的位置的指标)。电影框可包含若干类型的子框。电影框中的子框可包含一或多个播放轨框。播放轨框可包含关于电影的个别播放轨的信息。播放轨框可包含指定单一播放轨的总信息的播放轨头框。此外，播放轨框可包含含有媒体信息框的媒体框。媒体信息框可包含含有媒体样本在播放轨中的数据索引的样本表框。样本表框中的信息可用以按时间(且对于播放轨的样本中的每一者，按类型、大小、容器及到样本的所述容器的偏移)定位样本。因此，将用于播放轨的元数据围封于播放轨框(“trak”)中，而将播放轨的媒体内容围封于媒体数据框(“mdat”)中或直接围封于单独文件中。用于播放轨的媒体内容包括一连串样本(例如，由一连串样本组成)，例如，音频或视频存取单元。

[0097] ISOBMFF指定以下类型的播放轨：媒体播放轨，其含有基本媒体流；提示播放轨，其包含媒体发射指令或表示接收的包流；及计时元数据播放轨，其包括时间同步的元数据。用于每一播放轨的元数据包含样本描述项的列表，每一项提供在播放轨中使用的译码或封装格式及对于处理所述格式需要的初始化数据。每一样本与播放轨的样本描述项中的一者相关联。

[0098] ISOBMFF实现通过各种机构指定样本特定元数据。样本表框(“stbl”)内的特定框已经标准化以响应普通需求。举例来说，同步样本框(“stss”)为样本表框内的框。同步样本

框用以列出播放轨的随机存取样本。本发明可将由同步样本框列出的样本称作同步样本。在另一实例中,样本分组机构实现根据四字符分组类型将样本映射成共享指定为文件中的样本群组描述项的同一性质的样本的群组。已在ISOBMFF中指定若干分组类型。

[0099] 样本表框可包含一或多个SampleToGroup框及一或多个样本群组描述框(即, SampleGroupDescription框)。SampleToGroup框可用以确定样本属于的样本群组,连同所述样本群组的相关联描述。换句话说,SampleToGroup框可指示样本属于的群组。SampleToGroup框可具有“sbgp”的框类型。SampleToGroup框可包含分组类型元件(例如, grouping_type)。分组类型元件可为识别样本分组的类型(即,用以形成样本群组的规则)的整数。此外,SampleToGroup框可包含一或多个项。SampleToGroup框中的每一项可与播放轨中的一系列不同的非重叠连续样本相关联。每一项可指示样本计数元素(例如, sample_count)及群组描述索引元素(例如, group_description_index)。项目的样本计数元素可指示与所述项目相关联的样本的数目。换句话说,项目的样本计数元素可为给出具有相同样本群组描述符的连续样本的数目的整数。群组描述索引元素可识别含有与所述项目相关联的样本的描述的SampleGroupDescription框。多个项目的群组描述索引元素可识别相同SampleGroupDescription框。

[0100] 当前文件格式设计可具有一或多个问题。为了基于ISOBMFF存储特定视频编解码器的视频内容,可能需要所述视频编解码器的文件格式规格。为了含有例如MV-HEVC及SHVC等多个层的视频流的存储,可重新使用来自SVC及MVC文件格式的概念中的一些。然而,许多部分不能直接用于SHVC及MV-HEVC视频流。HEVC文件格式的直接应用具有至少下列缺点:SHVC及MV-HEVC可开始于含有基础层中的IRAP图片但也可含有其它层中的其它非IRAP图片的存取单元,或反之亦然。同步样本当前不允许用于随机存取的此点的指示。

[0101] 本发明描述对以上问题的潜在解决方案,以及提供其它潜在改进,以实现含有多个层的视频流的高效且灵活存储。本发明中描述的技术潜在地适用于用于存储由任何视频编解码器译码的此视频内容的任何文件格式,但所述描述是针对基于HEVC文件格式存储SHVC及MV-HEVC视频流,其在ISO/IEC 14496-15的第8条中指定。

[0102] 以下将详细论述本发明的所述技术的详细实施方案。可在以下实例中概述本发明的技术。可分开来使用以下实例。替代地,可将以下实例的各种组合一起使用。

[0103] 在第一实例中,Compressorname为在VisualSampleEntry框中指定的值。如在ISO/IEC 14496-12的第8.5.2.1节中所描述,VisualSampleEntry框为存储关于使用的译码类型的详细信息及对于所述译码所需要的任何初始化信息的用于视频播放轨的样本表框的类型。Compressorname指示用以产生媒体数据的压缩器的名称。视频解码器可使用Compressorname的值确定如何及/或是否解码文件中的视频数据。如在ISO/IEC 14496-12的第8.5.3节中所定义,Compressorname在固定32字节字段中格式化,其中第一字节经设定到待显示的字节的数目,接着为可显示的数据的字节的所述数目,且接着填补完整的一共32个字节(包含大小字节)。

[0104] 第一实例允许Compressorname的两个新值。Compressorname的第一新值为针对含有SHVC视频流的文件的“\013SHVC译码”。Compressorname的第二新值为用于含有 MV-HEVC视频流的文件的“\016MV-HEVC译码”。可如在以下第9.5.3.1.3节及第10.5.3.2 节中所展示实施此第一实例。

[0105] 如上简要地描述,文件可包含含有用于文件的播放轨的元数据的电影框。电影框可包含用于所述文件的每一播放轨的播放轨框。此外,播放轨框可包含含有声明播放轨的媒体的特性信息的所有对象的媒体信息框。媒体信息框可包含样本表框。样本表框可指定样本特定元数据。举例来说,样本表框可包含多个样本描述框。样本描述框中的每一者可为样本项的实例。在ISO/IEC 14496-12中,VisualSampleEntry类别的实例可用作样本项。针对特定视频译码标准的样本项的类别可扩展VisualSampleEntry类别。举例来说,针对HEVC的样本项的类别可扩展VisualSampleEntry类别。因此,本发明可将扩展VisualSampleEntry类别的不同类别称作不同样本项类型。

[0106] 在第二实例中,针对HEVC播放轨定义两个新样本项(即,“样本”)类型——“hev2”及“hvc2”。两个新样本项类型允许使用聚集器及提取器。一般来说,聚集器聚集呈单一聚集的数据单元的形式的多个NAL单元。举例来说,聚集器可含有多个NAL单元及/或可实际上串接多个NAL单元。一般来说,提取器指示从其它播放轨获得的数据的类型。举例来说,经由多个播放轨存储媒体数据(例如,HEVC数据)可导致紧凑的文件,这是因为可通过使用叫作提取器(其作为NAL单元嵌入于媒体数据中)的相对较小数据单元跨媒体播放轨参考数据来避免数据的复制。可如在以下第9.5.3.1.1节、第9.5.3.1.2节、第9.5.4节、第9.5.6节、第10.4.5节、第10.5.3.1.1节及第10.5.3.2节中所展示来实施此第二实例。

[0107] 在第三实例中,与关于用于多层次流的参数集的存储的特定要求相关联的样本项的定义经修改,以便实现对特定层或特定操作点的方便的随机存取。举例来说,当SHVC、MV-HEVC或3D-HEVC播放轨具有样本项且当样本含有至少一IRAP图片时,对于解码所述样本所需要的所有参数应包含于所述样本项或所述样本自身中。在此实例中,当样本不含有任何IRAP图片时,对于解码所述样本所需要的所有参数集(例如,VPS、SPS、PPS)应包含于所述样本项中或自从含有至少一IRAP图片的先前样本到所述样本自身(包含性)的任何样本中。可如以下在第9.5.3.1.1节中所展示实施此第三实例。

[0108] 在第三实例的一个替代版本中,当SHVC、MV-HEVC或3D-HEVC播放轨具有样本项且当在样本中的图片为IRAP图片时,对于解码所述图片所需要的所有参数集应包含于所述样本项中或所述样本自身中。此外,在此替代例中,当样本不含有任何IRAP图片时,对于解码所述图片所需要的所有参数集应包含于样本项中或同一层中在含有至少一IRAP图片的先前样本后的样本到所述样本自身中的任何者中(包含性)。

[0109] 在第四实例中,定义用于现有样本项类型的以下情况。在此实例中,属于样本项类型“hev1”及“hvc1”的样本含有用于具有HEVC VCL NAL单元的SHVC及MV-HEVC播放轨的HEVC、SHVC及MV-HEVC配置。此外,含有SHVC及MV-HEVC配置的样本项类型“hev1”及“hvc1”是针对无HEVC NAL单元但具有nuh_layer_id大于0的VCL NAL单元的SHVC及MV-HEVC播放轨定义,其中不允许提取器。可如以下在第9.5.3.1.1节中所展示实施此第四实例。

[0110] 在第五实例中,将SHVC、MV-HEVC或3D-HEVC播放轨中的同步样本定义为含有皆为IRAP图片的图片的样本。可如以下在第9.5.5节及第10.4.3节中所展示实施此第五实例。如以下在第9.5.5节所指定,如果存取单元中的每一经译码图片为IRAP图片,那么将SHVC样本考虑为同步样本,如在HEVC WD中所定义。此外,如以下在第10.4.3节中所指定,如果存取单元中的每一经译码图片为无RASL图片的IRAP图片,那么将MV-HEVC样本考虑为同步样本,如在HEVC WD中所定义。

[0111] 因此,在第五实例中,作为产生文件的部分,文件产生装置34可产生同步样本框,所述同步样本框包含以文件记载多层视频数据的播放轨的同步样本的同步样本表。播放轨的每一同步样本为播放轨的随机存取样本。如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本。如果存取单元中的每一经译码图片为无RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

[0112] 在第五实例的替代版本中,将SHVC、MV-HEVC或3D-HEVC播放轨中的同步样本定义为含有皆为无RASL图片的IRAP图片的图片的样本。同步样本表以文件记载同步样本。视情况,同步样本样本群组以文件记载同步样本。换句话说,同步样本样本群组包含识别同步样本的信息。

[0113] 在第六实例中,将“rap”样本群组定义为含有含皆为IRAP图片(具有或无RASL图片)的图片的那些样本。可如以下在第9.5.5节中所示实施此第六实例。替代地,在第六实例中,将“rap”样本群组定义为含有含皆为IRAP图片的图片的那些样本,但不包含指示为同步样本的那些样本。

[0114] 在第七实例中,定义以文件记载含有以下各者的所有样本的新样本群组或新框:至少一IRAP图片、所述样本中的IRAP图片中的VCL NAL单元的NAL单元类型、样本中的所有经译码图片是否皆为IRAP图片及(如果不)样本中的IRAP图片的数目,及样本中的此些IRAP图片的层ID值。

[0115] 因此,在此第七实例中,文件产生装置34可产生包括含有用于文件中的播放轨的元数据的播放轨框的文件。用于播放轨的媒体数据包括一连串样本。样本中的每一者可为多层视频数据的存取单元。作为产生文件的部分,文件产生装置34在文件中产生以文件记载含有至少一IRAP图片的所有样本的额外框。

[0116] 部分如以下在第9.5.5.1节中所示实施此第七实例。如以下在第9.5.5.1节中所示,随机可存取样本项类别扩展VisualSampleGroupEntry类别。随机可存取样本项类别的实例(即,随机可存取样本项框)对应于含有至少一IRAP图片的样本。此外,随机可存取样本项框包含指定对应的样本中的所有经译码图片是否为IRAP图片的all_pics_are_IRAP值。

[0117] 因此,在第七实例中,文件产生装置34可产生包含值(例如,all_pics_are_IRAP)的样本项。值等于1指定样本中的每一经译码图片为IRAP图片。值等于0指定样本中并非所有经译码图片皆为IRAP图片。

[0118] 此外,根据第七实例,当样本的并非所有经译码图片皆为IRAP图片时,文件产生装置34可在对应于样本的样本项中包含指示样本群组的每一样本中的IRAP图片的数目的值。另外,当样本中并非所有经译码图片皆为IRAP图片时,文件产生装置34可在对应于样本的样本项中包含指示样本中的IRAP图片的层识别符的值。

[0119] 替代地,在第七实例中,新样本群组或新框以文件记载此些样本,但不包含指示为同步样本或“rap”样本群组的成员的样本。

[0120] 此第七实例可解决当使用ISOBMFF或其现有扩展存储多层视频数据时可引起的一或多个问题。举例来说,在单层视频译码中,通常每存取单元仅存在单一经译码图片。然而,在多层视频译码中,通常每存取单元存在一个以上经译码图片。ISOBMFF及其现有扩展不提供指示哪些样本包含一或多个IRAP图片的方式。此可妨碍计算装置定位文件中的随机存取点或执行层切换的能力。举例来说,在无指示样本中的哪些者含有一或多个IRAP图片的信

息的情况下,计算装置可能需要剖析及解译NAL单元以便确定存取单元是否可用作随机存取点及/或用于层切换。剖析及解释NAL单元可对计算装置增加复杂度,且可消耗时间及处理资源。此外,执行随机存取及/或层切换的一些计算装置(例如,流式传输服务器)并未经配置以剖析或解译NAL单元。

[0121] 在第八实例中,包含新类型的子样本的介绍,其中每一子样本含有一个经译码图片及其相关联的非VCL NAL单元。可如以下在第9.5.8节中所示实施此第八实例。因此,在此第八实例中,文件产生装置34可产生包括含有用于文件中的播放轨的元数据的播放轨框的文件。用于播放轨的媒体数据包括一连串样本。样本中的每一者为多层视频数据的存取单元。作为产生文件的部分,文件产生装置34在文件中产生子样本信息框,所述子样本信息框含有指定在所述子样本信息框中给定的子样本信息的类型的旗标。当旗标具有特定值时,对应于所述子样本信息框的子样本含有正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL单元。

[0122] 此第八实例可解决当使用ISOBMFF或其现有扩展存储多层视频数据时可引起的一或多个问题。举例来说,在多层视频译码中,每样本可存在多个经译码图片。举例来说,对于每一层,在样本中可存在一或多个图片。然而,在对于H.264/AVC及HEVC的ISOBMFF的扩展中,当样本包含多个图片时,子样本信息框不提供关于所述样本内的个别图片的信息。此第八实例的技术可通过提供新类型的子样本信息框来解决此问题,所述新类型的子样本信息框提供关于含有仅一个经译码图片及与所述经译码图片相关联的非VCL NAL单元的子样本的信息。如与仅提供在与经译码图片相关联的NAL单元内的此信息相反,提供关于文件结构中的个别经译码图片的信息可使计算装置能够确定关于经译码图片的信息,而不必解译所述NAL单元。在一些情况下,为了减小计算装置的复杂度及/或增大计算装置的处理量,计算装置未经配置以解译NAL单元。在计算装置正流式传输存储于文件中的NAL单元的一些实例中,计算装置可使用子样本信息框中的信息确定是否将子样本的NAL单元转递到客户端装置。

[0123] 第九实例涉及在多层上下文中的非输出样本的处置。特定来说,在第九实例中,当存取单元含有具有等于1的pic_output_flag的一些经译码图片及具有等于0的pic_output_flag的一些其它经译码图片时,必须使用至少两个播放轨存储流,使得在每一播放轨内,每一样本中的所有经译码图片具有相同的pic_output_flag值。可如以下在第9.5.9节中所示实施此第九实例。

[0124] 因此,在此第九实例中,文件产生装置34可产生包括围封媒体内容的媒体数据框的文件。媒体内容包括一连串样本。样本中的每一者为多层视频数据的存取单元。响应于多层视频数据的位流的至少一存取单元包含具有等于1的图片输出旗标(例如, pic_output_flag)的经译码图片及具有等于0的图片输出旗标的经译码图片,文件产生装置34可使用至少两个播放轨将位流存储于文件中。对于来自至少两个播放轨的每一相应播放轨,相应播放轨的每一样本中的所有经译码图片具有相同的图片输出旗标值。

[0125] 此第九实例可解决当使用ISOBMFF或其现有扩展存储多层视频数据时可引起的一或多个问题。举例来说,如果使用单一播放轨存储具有等于0的图片输出旗标及等于1的图片输出旗标的经译码图片,那么将违反各种文件格式化规则。举例来说,文件格式化规则通常需要每时刻在播放轨中仅存在一个样本。如果单一播放轨存储具有等于0的图片输出旗

标及等于1的图片输出旗标的经译码图片,那么每时刻在播放轨中将存在多个样本。迫使具有不同图片输出旗标值的经译码图片在文件的不同播放轨中可解决此问题。

[0126] 以下描述本发明的一些技术的实例实施方案。以下描述的实例实施方案是基于在MPEG输出文件W13478中的14496-15的最新近集成规格。以下包含对附录A的改变(通过下划线展示)及添加的章节(第9节针对SHVC,且第10节针对MV-HEVC)。换句话说,本发明的特定实例可修改ISO/IEC 14496-15的附录A,且可将第9节及/或第10节添加到ISO/IEC 14496-15。通过下划线及双下划线展示的文字可具有与本发明的实例的特定相关性。虽然在本文中描述的实例中各处使用术语SHVC,但本发明的设计实际上不仅将仅支持SHVC编解码器,而取而代之,可支持包含MV-HEVC、3D-HEVC的所有多层编解码器,除非另有明确地提到。

[0127] 9 SHVC基本流及样本定义

[0128] 9.1 介绍

[0129] 此条款指定SHVC数据的存储格式。其扩展在第8条中的HEVC的存储格式的定义。

[0130] 如在此条款及附录A到D中定义的用于SHVC内容的存储的文件格式使用ISO基本媒体文件格式的现有能力及普通HEVC文件格式(即,在第8条中指定的文件格式)。此外,使用尤其以下结构或扩展来支持SHVC特定特征。

[0131] 聚集器:

[0132] 通过将NAL单元的不规则模式改变成聚集的数据单元的规则模式来实现NAL 单元的高效可缩放分组的结构。

[0133] 提取器:

[0134] 实现从不同于含有媒体数据的播放轨的播放轨高效提取NAL单元的结构。

[0135] 时间元数据语句:

[0136] 用于存储媒体样本的经时间对准信息的结构。

[0137] HEVC兼容性:

[0138] 提供用于以HEVC兼容方式存储SHVC位流,使得HEVC兼容基础层可由任一遵照普通HEVC文件格式的读取器使用。

[0139] 9.2 基本流结构

[0140] 根据8.2存储SHVC流,具有SHVC视频基本流的以下定义:

[0141] ●SHVC视频基本流应含有所有视频译码有关的NAL单元(即,含有视频数据或传信视频结构的那些NAL单元)且可含有例如SEI消息及存取单元定界符NAL 单元的非视频译码有关的NAL单元。也可存在聚集器(见A.2)或提取器(见A.3)。聚集器及提取器应如在此国际标准中所定义来处理(例如,在存取文件时,不应直接置放于输出缓冲器中)。可存在未明确受到抑制的其它NAL单元,且如果其未经辨识,那么应将其忽略(例如,在存取文件时,不置放于输出缓冲器中)。不应使用相关联的参数集流来存储SHVC流。

[0142] 可存在等于0的nuh_layer_id的VCL NAL单元、大于0的nuh_layer_id的VCL NAL单元及SHVC视频基本流中存在的非VCL NAL单元。另外,可存在SHVC视频基本流中存在的聚集器NAL单元及提取器NAL单元。

[0143] 9.3 普通HEVC文件格式的使用

[0144] SHVC文件格式为在第8条中定义的普通HEVC文件格式的扩展。

[0145] 9.4 样本及配置定义

[0146] 9.4.1 介绍

[0147] SHVC样本:SHVC样本也为如在ISO/IEC 23008-2的附录H中定义的存取单元。

[0148] 9.4.2 标准次序及限制

[0149] 9.4.2.1 限制

[0150] 除8.3.2中的要求外,以下限制也应用于SHVC数据。

[0151] ●VCL NAL单元:在其组成时间为由一个存取单元表示的图片的组成时间的实例中应含有在所述存取单元中的所有VCL NAL单元。SHVC样本应含有至少一 VCL NAL单元。

[0152] ●聚集器/提取器:在聚集器中包含或由提取器参考的所有NAL单元的次序恰好为解码次序,如同此些NAL单元存在于不含有聚集器/提取器的样本中。在处理了聚集器或提取器后,所有NAL单元必须呈有效的解码次序,如在ISO/IEC 23008-2中所指定。

[0153] 9.4.2.2 解码器配置记录

[0154] 当将在8.3.3.1中定义的解码器配置记录用于可解译为SHVC或HEVC流的流时,HEVC解码器配置记录应反映HEVC兼容基础层的性质,例如,其应仅含有对于解码 HEVC基础层所需要的参数集。

[0155] SHVCDecoderConfigurationRecord在结构上与HEVCDecoderConfigurationRecord相同。语法如下:

[0156] aligned (8) class SHVCDecoderConfigurationRecord {

[0157] //与HEVCDecoderConfigurationRecord语法结构中的字段相同的字段

[0158] }

[0159] SHVCDecoderConfigurationRecord 中的字段的语义与针对HEVCDecoderConfigurationRecord所定义相同。

[0160] 9.5 从ISO基本媒体文件格式的导出

[0161] 9.5.1 SHVC播放轨结构

[0162] 可缩放视频流由文件中的一或多个视频播放轨表示。每一播放轨表示可缩放流的一或多个操作点。当然,如果需要,可缩放流可进一步变细。

[0163] 让最低操作点为含有具有仅等于0的nuh_layer_id及仅等于0的TemporalId的NAL单元的所有操作点中的一者。应将含有最低操作点的播放轨提名为“可缩放基础播放轨”。为相同可缩放编码信息的部分的所有其它播放轨应通过类型 ‘sbas’ (可缩放基础) 的播放轨参考而与此基础播放轨有联系。

[0164] 共享同一可缩放基础播放轨的所有播放轨必须共享与可缩放基础播放轨相同的时间标度。

[0165] 9.5.2 数据共享及提取

[0166] 不同的播放轨可逻辑上共享数据。此共享可呈以下两个形式中的一者:

[0167] a) 将样本数据从一个播放轨复制到另一播放轨(且可压紧或与例如音频等其它数据重新交错)。此建立较大总体文件,但为了易于提取,低位速率数据可经压紧及/或与其它材料交错。

[0168] b) 可存在关于如何在读取文件时执行此复制的指令。对于第二情况,使用提取器(在A.3中所定义)。

- [0169] 9.5.3 SHVC视频流定义
- [0170] 9.5.3.1 样本项名称及格式
- [0171] 9.5.3.1.1 定义
- [0172] 类型: 'hvc2'、'hev2'、'shc1'、'shv1'、'shcC'
- [0173] 容器: 样本描述框 ('stsd')
- [0174] 必选: 'hvc1'、'hev1'、'hvc2'、'hev2'、'shc1' 或 'shv1' 样本项为必选的
- [0175] 数量: 可存在一或多个样本项
- [0176] 当样本项名称为 'shc1' 时, array_completeness 的默认及必选值对于所有类型的参数集的阵列为 1, 且对于所有其它阵列为 0。当样本项名称为 'shv1' 时, array_completeness 的默认值对于所有阵列皆为 0。
- [0177] 当样本项名称为 'shv1' 时, 以下适用:
- [0178] ● 如果样本含有至少一IRAP图片(如在ISO/IEC 23008-2中所定义), 那么对于解码所述样本所需要的所有参数集应包含于样本项中或所述样本自身中。
- [0179] ● 否则(样本不含有IRAP图片), 对于解码所述样本所需要的所有参数集应包含于样本项中或自从含有至少一IRAP图片的先前样本到所述样本自身(包含性)的任何样本中。
- [0180] 替代地, 当样本项名称为 'shv1' 时, 以下适用:
- [0181] ● 如果样本中的经译码图片为IRAP图片(如在ISO/IEC 23008-2中所定义), 那么对于解码所述经译码图片所需要的所有参数集应包含于样本项中或所述样本自身中。
- [0182] ● 否则(样本中的经译码图片并非IRAP图片), 对于解码所述经译码图片所需要的所有参数集应包含于样本项中或自从与所述经译码图片相同的层中的含有IRAP图片的先前样本到所述样本自身(包含性)的任何样本中。
- [0183] 如果SHVC基本流含有可使用的HEVC兼容基础层, 那么应使用HEVC视觉样本项 ('hvc1' 或 'hev1')。此处, 所述项目一开始应含有HEVC配置框, 可能接着为如下所定义的SHVC配置框。HEVC配置框以文件记载涉及HEVC兼容基础层的简档、层、层级及可能也有参数集, 如由HEVCDriverConfigurationRecord所定义。SHVC配置框以文件记载涉及存储于SHVCConfigurationBox中的含有SHVC兼容增强层的整个流的简档、层、层级及可能也有参数集(如由HEVCDriverConfigurationRecord所定义)。
- [0184] 如果SHVC基本流不含有可使用的HEVC基础层, 那么应使用SHVC视觉样本项 ('shc1' 或 'shv1')。SHVC视觉样本项应含有如下所定义的SHVC配置框。此包含SHVCDriverConfigurationRecord, 如在此国际标准中所定义。
- [0185] 在任一给定样本项中的SHVC及HEVC配置中的lengthSizeMinusOne字段应具有相同值。
- [0186] 提取器或聚集器可用于在 'hvc1'、'hev1'、'hvc2'、'hev2'、'shc1' 或 'shv1' 播放轨中具有大于 0 的nuh_layer_id 的NAL单元。'hvc2' 或 'hev2' 样本项中的 'extra_boxes' 可为SHVCConfigurationBox 或其它扩展框。
- [0187] 注意 当指示HEVC兼容性时, 可能有必要指示用于HEVC基础层的不现实层级以适

应整个流的位速率,这是因为将所有NAL单元考虑为包含于HEVC基础层中,且因此,可将其馈入到解码器,期望解码器舍弃其未辨识的那些NAL单元。此情况在使用‘hvc1’或‘hev1’样本项且HEVC及SHVC两个配置皆存在时发生。

[0188] SHVCConfigurationBox可存在于‘hvc1’或‘hev1’样本项中。在此情况下,以下的HEVCSHVCSSampleEntry定义适用。

[0189] 下表展示对于视频播放轨的样本项、配置及SHVC工具(不包含计时元数据,其始终用于另一播放轨中)的所有可能使用:

[0190] 表10—用于HEVC及SHVC播放轨的样本项的使用

[0191]

样本项名称	具有配置记录	含义
‘hvc1’或 ‘hev1’	仅 HEVC 配置	无具有大于 0 的 nuh_layer_id 的 NAL 单元的普通 HEVC 播放轨; 提取器及聚集器不应存在。
‘hvc1’或 ‘hev1’	HEVC 及 SHVC 配置	具有 nuh_layer_id 等于 0 的 NAL 单元及 nuh_layer_id 大于 0 的 NAL 单元的 SHVC 播放轨; 提取器及聚集器可存在; 提取器不应参考具有等于 0 的 nuh_layer_id 的 NAL 单元; 聚集器不应含有但可参考具有等于 0 的 nuh_layer_id 的 NAL 单元。
‘hvc2’或 ‘hev2’	仅 HEVC 配置	无具有大于 0 的 nuh_layer_id 的 NAL 单元的普通 HEVC 播放轨; 提取器可存在且用以参考 NAL 单元; 聚集器可存在以含有且参考 NAL 单元。
‘hvc2’或 ‘hev2’	HEVC 及 SHVC 配置	具有 nuh_layer_id 等于 0 的 NAL 单元及 nuh_layer_id 大于 0 的 NAL 单元的 SHVC 播放轨; 提取器及聚集器可存在; 提取器可参考任何 NAL 单元; 聚集器可皆含有且参考任何 NAL 单元。
‘shc1’或 ‘shv1’	SHVC 配置	无具有等于 0 的 nuh_layer_id 的 NAL 单元的 SHVC 播放轨; 提取器可存在且用以参考 NAL 单元; 聚集器可存在以含有且参考 NAL 单元。

[0192] 9.5.3.1.2 语法

```
class SHVCConfigurationBox extends Box('shcC') {
    SHVCDecoderConfigurationRecord() SHVCCConfig;
}
class HEVC2SampleEntry() extends VisualSampleEntry ('hev2'or 'hvc2') {
```

[0193]

```

        HEVCCConfigurationBox hevconfig;
        MPEG4BitRateBox ();           //任选
        MPEG4ExtensionDescriptorsBox (); //任选
        extra_boxes boxes;           //任选
    }

    class HEVCSHVCSSampleEntry() extends HEVCSampleEntry {
        SHVCCConfigurationBox shvcconfig;
    }

    class HEVC2SHVCSSampleEntry() extends HEVC2SampleEntry() {
        SHVCCConfigurationBox shvcconfig; //任选
    }

    //在播放轨并不 HEVC 兼容的情况下使用这个
    class SHVCSampleEntry() extends VisualSampleEntry ('shv1' or 'shc1') {
        SHVCCConfigurationBox shvcconfig;
        MPEG4BitRateBox ();           //任选
        MPEG4ExtensionDescriptorsBox (); //任选
        extra_boxes boxes;           //任选
    }

```

[0194] [0195] 9.5.3.1.3 语义

[0196] 当样本项应用到的流含有具有大于0的nuh_layer_id的NAL单元时,基础类别VisualSampleEntry中的Compressorname指示与正推荐的值“\013SHVC Coding” (\013为11,以字节为单元的字符串“SHVC Coding”的长度)一起使用的压缩器的名称。

[0197] 9.5.4 SHVC视觉宽度及高度

[0198] 如果含有具有大于0的nuh_layer_id的NAL单元的流由类型‘hvc1’、‘hev1’、‘hvc2’、‘hev2’的样本项描述,那么在所述流的VisualSampleEntry中以文件记载的视觉宽度及高度为HEVC基础层的视觉宽度及高度;否则,其为通过解码整个流的最高层的经解码图片的视觉宽度及高度。

[0199] 9.5.5 同步样本

[0200] 如果存取单元中的每一经译码图片为IRAP图片,那么将SHVC样本考虑为同步样本,如在ISO/IEC 23008-2中所定义。同步样本由同步样本表以文件记载,且可额外由同步样本样本群组及‘rap’样本群组以文件记载。

[0201] 9.5.5.1 随机存取样本样本群组

[0202] 9.5.5.1.1 定义

[0203] 群组类型:‘ras’

[0204] 容器:样本群组描述框(‘ras’)

[0205] 必选:否

[0206] 数量:零或多个

[0207] 随机存取样本样本群组识别含有至少一IRAP图片的样本。

[0208] 9.5.5.1.2 语法

[0209]

```
class RandomAccessibleSampleEntry() extends VisualSampleGroupEntry ('ras ') {
    unsigned int(1) reserved = 0;
    unsigned int(1) all_pics_are_IRAP
    unsigned int(6) IRAP_nal_unit_type
    if( !all_pics_are_IRAP ) {
        unsigned int(2) reserved = 0;
        unsigned int(6) num_IRAP_pics;
        for(i=0; i< num_IRAP_pics; i++) {
            unsigned int(2) reserved = 0;
            unsigned int(6) IRAP_pic_layer_id
        }
    }
}
```

[0210] 9.5.5.1.3 语义

[0211] all_pics_are_IRAP等于1指定群组的每一样本中的所有经译码图片为IRAP图片。当值等于0时,以上约束可或可不适用。

[0212] IRAP_nal_unit_type指定群组的每一样本中的IRAP图片的NAL单元类型。IRAP_nal_unit_type的值应在16到23的范围内(包含性)。

[0213] num_IRAP_pics指定群组的每一样本中的IRAP图片的数目。

[0214] IRAP_pic_layer_id指定群组的每一样本中的第i个IRAP图片的nuh_layer_id 的值。

[0215] 9.5.6 关于随机存取恢复点及随机存取点的样本群组

[0216] 对于由类型 'hvc1'、'hev1'、'hvc2' 或 'hev2' 的样本项描述的视频数据,随机存取恢复样本群组及随机存取点样本群组分别识别用于对整个位流操作的HEVC解码及SHVC解码器(如果有)的随机存取恢复点及随机存取点。

[0217] 对于由类型 'shc1' 或 'shv1' 描述的视频数据,随机存取恢复样本群组识别整个SHVC位流中的随机存取恢复,且随机存取点样本群组识别整个SHVC位流中的随机存取点。

[0218] 如果存取单元中的每一经译码图片为IRAP图片(具有或无RASL图片),那么将SHVC样本考虑为随机存取点,如在ISO/IEC 23008-2中所定义,且ISO/IEC 14496-2中的前置样本为所有图片皆为RASL图片的样本,如在ISO/IEC 23008-2中所定义。

[0219] 9.5.7 独立抛弃式样本框

[0220] 如果其用于HEVC及SHVC两者皆兼容的播放轨中,那么应注意语句为真,不管使用

SHVC数据(可能仅HEVC数据)的何有效子集。如果信息变化,那么可能需要“未知”值(字段sample-depends-on、sample-is-depended-on及sample-has-redundancy的值0)。

[0221] 9.5.8 用于SHVC的子样本的定义

[0222] 此子条款扩展8.4.8中的HEVC的子样本的定义。

[0223] 对于在SHVC流中的子样本信息框(ISO/IEC 14496-12的8.7.7)的使用,基于如下指定的子样本信息框的旗标的值定义子样本。此框的存在是可选的;然而,如果存在于含有SHVC数据的播放轨中,那么其应具有此处定义的语义。

[0224] 旗标指定在此框中给出的子样本信息的类型,如下:

[0225] 0: 基于NAL单元的子样本。子样本含有一或多个相邻NAL单元。

[0226] 1: 基于解码单元的子样本。子样本含有正好一个解码单元。

[0227] 2: 基于平铺块的子样本。子样本含有一个平铺块及含有所述平铺块的VCL NAL单元的相关联的非VCL NAL单元(如果有),或含有一或多个非VCL NAL单元。

[0228] 3: 基于CTU行的子样本。子样本含有切片内的一个CTU行及含有所述CTU 行的VCL NAL单元的相关联的非VCL NAL单元(如果有),或含有一或多个非VCL NAL单元。当entropy_coding_sync_enabled_flag等于0时,不应使用此类型的子样本信息。

[0229] 4: 基于切片的子样本。子样本含有一个切片(其中每一切片可含有一或多个切片片段,其中的每一者为NAL单元)及相关联的非VCL NAL单元(如果有),或含有一或多个非VCL NAL单元。

[0230]

5: 基于图片的子样本。子样本含有一个经译码图片及相关联的非 VCL NAL 单元。

[0231] 其它旗标值是保留的。

[0232] subsample_priority字段应设定到根据在ISO/IEC 14496-12中的此字段的规格的值。

[0233] 仅如果在此子样本经舍弃的情况下仍可解码此样本时,应将此可舍弃字段设定到1(例如,子样本由SEI NAL单元组成)。

[0234] 当NAL单元的第一字节包含于子样本中时,先前长度字段必须也包含于同一子样本中。

```

if(flags == 0) {
    unsigned int(1) SubLayerRefNalUnitFlag;
    unsigned int(1) RapNalUnitFlag;
    unsigned int(1) VclNalUnitFlag;
    unsigned int(1) DiscardableFlag;
    unsigned int(1) NoInterLayerPredFlag;
    unsigned int(6) LayerId;
    unsigned int(3) TempId;
    unsigned int(18) reserved = 0;
} else if(flags == 1)
    unsigned int(32) reserved = 0;
else if(flags == 2) {
    unsigned int(2) vcl_idc;
    unsigned int(2) reserved = 0;
[0235]    unsigned int(4) log2_min_luma_ctb;
    unsigned int(12) ctb_x;
    unsigned int(12) ctb_y;
} else if(flags == 3 || flags == 4) {
    unsigned int(2) vcl_idc;
    unsigned int(30) reserved = 0;
} esle if(flags == 5) {
    unsigned int(1) DiscardableFlag;
    unsigned int(6) VclNalUnitType;
    unsigned int(6) LayerId;
    unsigned int(3) TempId;
    unsigned int(1) NoInterLayerPredFlag;
    unsigned int(1) SubLayerRefNalUnitFlag;
    unsigned int(14) reserved = 0;
}

```

[0236] }

[0237] SubLayerRefNalUnitFlag等于0指示子样本中的所有NAL单元为如在ISO/IEC 23008-2中指定的子层非参考图片的VCL NAL单元。值1指示子样本中的所有NAL 单元为如在ISO/IEC 23008-2中指定的子层参考图片的VCL NAL单元。

[0238] RapNalUnitFlag等于0指示子样本中的NAL单元中无一者具有等于如在ISO/IEC 23008-2中指定的IDR_W_RADL、IDR_N_LP、CRA_NUT、BLA_W_LP、BLA_W_RADL、BLA_N_LP、RSV_

IRAP_VCL22或RSV_IRAP_VCL23的 nal_unit_type。值1指示子样本中的所有NAL单元具有如在ISO/IEC 23008-2中指定的IDR_W_RADL、IDR_N_LP、CRA_NUT、BLA_W_LP、BLA_W_RADL、BLA_N_LP、RSV_IRAP_VCL22或RSV_IRAP_VCL23的nal_unit_type。

[0239] VclNalUnitFlag等于0指示子样本中的所有NAL单元为非VCL NAL单元。值 1指示子样本中的所有NAL单元为VCL NAL单元。

[0240]

DiscardableFlag 指示子样本中的 VCL NAL 单元的 discardable_flag 值。子样本中的所有 VCL NAL 单元具有相同的 discardable_flag 值。

注意此并非与子样本信息框中的可舍弃字段相同的定义。

NoInterLayerPredFlag 指示子样本中的 VCL NAL 单元的 inter_layer_pred_enabled_flag 的值。子样本中的所有 VCL NAL 单元具有相同的 inter_layer_pred_enabled_flag 值。

LayerId 指示子样本中的 NAL 单元的 nuh_layer_id 值。子样本中的所有 NAL 单元应具有相同的 nuh_layer_id 值。

TempId 指示子样本中的 NAL 单元的 TemporalId 值。子样本中的所有 NAL 单元应具有相同的 TemporalId 值。

[0241] vcl_idc指示子样本含有视频译码层(VCL)数据、非VCL数据或是两者,如下:

[0242] 0:子样本含有VCL数据且不含有非VCL数据

[0243] 1:子样本不含有VCL数据且含有非VCL数据

[0244] 2:子样本可含有应彼此相关联的VCL及非VCL数据两者。举例来说,子样本可含有解码单元信息SEI消息,接着为与SEI消息相关联的NAL单元的集合。

[0245] 3:保留

[0246] log2_min_luma_ctb指示ctb_x及ctb_y的单位,如下所指定:

[0247] 0:8个明度样本

[0248] 1:16个明度样本

[0249] 2:32个明度样本

[0250] 3:64个明度样本

[0251] ctb_x指定当旗标等于2且vcl_idc等于1或2时与子样本相关联的平铺块的最右边明度样本的0基坐标,按从如以上指定的log2_min_luma_ctb导出的单位。

[0252] ctb_y指定当旗标等于2且vcl_idc等于1或2时与子样本相关联的平铺块的最底部明度样本的0基坐标,按从如以上指定的log2_min_luma_ctb导出的单位。

[0253]

VclNalUnitType 指示子样本中的 VCL NAL 单元的 nal_unit_type 值。子样本中的所有 VCL NAL 单元应具有相同的 nal_unit_type 值。

[0254] 9.5.9 处置非输出样本

[0255] 8.4.9 中的规格适用,其中“HEVC”由“SHVC”替换,且将非输出样本定义为目标输出层的图片具有等于0的pic_output_flag的样本。当存取单元含有具有等于1的 pic_output_flag的一些经译码图片及具有等于0的pic_output_flag的一些其它经译码图片时,必须使用至少两个播放轨来存储流,使得在每一播放轨内,每一样本中的所有经译码图片具有相同的pic_output_flag值。

[0256] 10 10 MV-HEVC基本流及样本定义

[0257] 10.1 介绍

[0258] 此条款指定MV-HEVC数据的存储格式。其扩展在第8条中的HEVC的存储格式的定义。

[0259] 如在此条款及附录A到D中定义的用于MV-HEVC内容的存储的文件格式使用ISO 基本媒体文件格式的现有能力及普通HEVC文件格式(即,在第8条中指定的文件格式)。此外,使用尤其以下结构或扩展来支持MV-HEVC特定特征。

[0260] 聚集器:

[0261] 通过将NAL单元的不规则模式改变成聚集的数据单元的规则模式来实现NAL 单元的高效可缩放分组的结构。

[0262] 提取器:

[0263] 实现从不同于含有媒体数据的播放轨的播放轨高效提取NAL单元的结构。

[0264] HEVC兼容性:

[0265] 提供用于以HEVC兼容方式存储MV-HEVC位流,使得HEVC兼容基础层可由任一遵照普通HEVC文件格式的读取器使用。

[0266] 对MV-HEVC的支持包含许多工具,且存在可使用其的方式的各种“模型”。详言地说,可以许多方式将MV-HEVC流置放于播放轨中,其中有以下方式:

[0267] 1.一个播放轨中的所有视图,标有样本群组;

[0268] 2.其自己的播放轨中的每一视图,标于样本项中;

[0269] 3.混合,含有所有视图的一个播放轨,及各含有可独立地译码的视图的一或多个单一视图播放轨;

[0270] 4.各在播放轨中的预期的操作点(例如,HEVC基础、立体声对、多视图场景)。

[0271] MV-HEVC文件格式允许将一或多个视图存储到播放轨内,类似于第9条中对SHVC的支持。每播放轨可使用多个视图的存储,例如,当内容提供者想要提供并不意欲用于子集建构的多视图位流时,或当已针对少数预定义的输出视图的集合(例如,1个、2个、5个或9个视图)建立位流时(其中可相应地建立播放轨)。如果将一个以上视图存储于播放轨中且存在表示MV-HEVC位流的若干播放轨(一个以上),那么推荐样本分组机制的使用。

[0272] 当MV-HEVC位流由多个播放轨表示且播放器使用含有多个播放轨中的数据的操作点时,所述播放器必须重建MV-HEVC存取单元,之后将其递送到MV-HEVC解码器。MV-HEVC操作点可明确地由播放轨表示,即,仅通过解析样本的所有提取器及聚集器 NAL单元来重建存取单元。如果操作点的数目大,那么针对每一操作点建立播放轨可能消耗空间且不切实际。在此情况下,MV-HEVC存取单元如在10.5.2中所指定来重建。MV-HEVC解码器配置记录含有指示相关联的样本使用明确或是隐含存取单元重建的字段(见explicit_au_track字段)。

[0273] 10.2 MV-HEVC播放轨结构

[0274] 根据8.2存储MV-HEVC流,具有MV-HEVC视频基本流的以下定义:

[0275] ●MV-HEVC视频基本流应含有所有视频译码有关的NAL单元(即,含有视频数据或传信视频结构的那些NAL单元)且可含有例如SEI消息及存取单元定界符 NAL单元的非视频译码有关的NAL单元。也可存在聚集器(见A.2)或提取器(见 A.3)。聚集器及提取器应如在此国际标准中所定义来处理(例如,在存取文件时,不应直接置放于输出缓冲器中)。可存在未明确受到抑制的其它NAL单元,且如果其未经辨识,那么应将其忽略(例如,在存取文件时,不置放于输出缓冲器中)。当需要时,不应使用相关联的参数集流来存储MV-HEVC流。

[0276] 可存在等于0的nuh_layer_id的VCL NAL单元、大于0的nuh_layer_id的VCL NAL单元及MV-HEVC视频基本流中存在的非VCL NAL单元。另外,可存在MV-HEVC视频基本流中存在的聚集器及提取器NAL单元。

[0277] 10.3 普通HEVC文件格式的使用

[0278] MV-HEVC文件格式为在第8条中定义的普通HEVC文件格式的扩展。

[0279] 10.4 样本及配置定义

[0280] 10.4.1 介绍

[0281] MV-HEVC样本: MV-HEVC样本也为如在ISO/IEC 23008-2的附录F中定义的存取单元。

[0282] 10.4.2 标准次序及限制

[0283] 10.4.2.1 限制

[0284] 除8.3.2中的要求外,以下限制也应用于MV-HEVC数据。

[0285] ●VCL NAL单元:在其组成时间为由一个存取单元表示的图片的组成时间的实例中应含有在所述存取单元中的所有VCL NAL单元。MV-HEVC样本应含有至少一VCL NAL单元。

[0286] ●聚集器/提取器:在聚集器中包含或由提取器参考的所有NAL单元的次序恰好为解码次序,如同这些NAL单元存在于不含有聚集器/提取器的样本中。在处理了聚集器或提取器后,所有NAL单元必须呈有效的解码次序,如在ISO/IEC 23008-2中所指定。

[0287] 10.4.2.2 解码器配置记录

[0288] 当将在8.3.3.1中定义的解码器配置记录用于可解译为MV-HEVC或HEVC流的流时,HEVC解码器配置记录应反映HEVC兼容基础视图的性质,例如,其应仅含有对于解码HEVC基础视图所需要的参数集。

[0289] MVHEVCDecoderConfigurationRecord 在结构上与HEVCDecoderConfigurationRecord 相同。语法如下:

[0290] aligned (8) class MVHEVCDecoderConfigurationRecord {

[0291] //与HEVCDecoderConfigurationRecord语法结构中的字段相同的字段

[0292] }

[0293] MVHEVCDecoderConfigurationRecord 中的字段的语义与针对HEVCDecoderConfigurationRecord所定义相同。

[0294] 10.4.3 同步样本

[0295] 如果存取单元中的每一经译码图片为无RASL图片的IRAP图片,那么将MV-HEVC 样

本考虑为同步样本,如在ISO/IEC 23008-2中所定义。同步样本由同步样本表以文件记载,且可额外地由与在SHVC中类似地定义的同步样本样本群组及‘rap’样本群组以文件记载。

[0296] 10.4.4 独立且抛弃式样本框

[0297] 如果其用于HEVC及MV-HEVC两者皆兼容的播放轨中,那么应注意语句为真,不管使用MV-HEVC数据(可能仅HEVC数据)的何有效子集。如果信息变化,那么可能需要“未知”值(字段sample-depends-on、sample-is-depended-on及sample-has-redundancy 的值0)。

[0298] 10.4.5 关于随机存取恢复点及随机存取点的样本群组

[0299] 对于由类型‘hvc1’、‘hev1’、‘hvc2’或‘hev2’的样本项描述的视频数据,随机存取恢复样本群组及随机存取点样本群组分别识别用于对整个位流操作的HEVC解码及MV-HEVC解码器(如果有)的随机存取恢复点及随机存取点。

[0300] 对于由MV-HEVC样本项类型描述的视频数据,随机存取恢复样本群组识别整个MV-HEVC位流中的随机存取恢复,且随机存取点样本群组识别整个MV-HEVC位流中的随机存取点。

[0301] 10.5 从ISO基本媒体文件格式的导出

[0302] 10.5.1 MV-HEVC播放轨结构

[0303] 多视图视频流由文件中的一或多个视频播放轨表示。每一播放轨表示流的一或多个视图。

[0304] 存在一或多个播放轨的最小集合,当将所述一或多个播放轨放在一起时,其含有经编码信息的完整集合。所有这些播放轨应具有在所有其样本项中的旗标“complete_representation”集合。形成完整经编码信息的此群播放轨叫作“完整子集”。

[0305] 让最低操作点为含有具有仅等于0的nuh_layer_id及仅等于0的TemporalId的NAL单元的所有操作点中的一者。应将含有最低操作点的播放轨提名为“基础视图播放轨”。为相同流的部分的所有其它播放轨应通过类型‘sbas’(视图基础)的播放轨参考而与此基础播放轨有联系。

[0306] 共享同一基础视图播放轨的所有播放轨必须共享与基础视图播放轨相同的时间标度。

[0307] 如果由播放轨表示的视图将由另一播放轨表示的另一视图用作视图间预测参考,那么类型‘scal’的播放轨参考应包含于参考源播放轨的播放轨中,用于视图间预测。

[0308] 如果将编辑应用到含有MV-HEVC位流的视图分量,那么编辑列表应在受到编辑影响的所有播放轨上一致。

[0309] 10.5.2 存取单元的重建构

[0310] 为了从一或多个MV-HEVC播放轨的样本重建构存取单元,可能需要首先确定目标输出视图。

[0311] 对于解码确定的目标输出视图需要的视图可从在视图识别符框或‘scal’播放轨参考中包含的参考视图识别符得出。

[0312] 如果若干播放轨含有用于存取单元的数据,那么在解码时间执行播放轨中的相应样本的对准,即,仅使用时间到样本表,而不考虑编辑列表。

[0313] 通过按符合ISO/IEC 23008-02的次序排列其NAL单元从所需的播放轨中的相应样本重建构存取单元。以下次序提供形成符合的存取单元的程序的概括:

- [0314] ●所有参数集NAL单元(从相关联的参数集播放轨且从相关联的基本流播放轨)。
- [0315] ●所有SEI NAL单元(从相关联的参数集播放轨且从相关联的基本流播放轨)。
- [0316] ●按视图次序索引值的降序的视图分量。在视图分量内的NAL单元在所述样本内按其出现的次序。

[0317] 10.5.3 样本项

[0318] 10.5.3.1 用于样本项的框

[0319] 10.5.3.1.1 视图识别符框

[0320] 10.5.3.1.1.1 定义

[0321] 框类型: 'vwid'

[0322] 容器: 样本项('hev1'、'hvc1'、'hev2'、'hvc2'、'mhc1'、'mhv1')或 MultiviewGroupEntry

[0323] 必选: 是(对于样本项)

[0324] 数量: 正好一个(对于样本项)

[0325] 当包含于样本项中时,此框指示包含于播放轨中的视图。此框也指示针对每一列出的视图的视图次序索引。另外,当视图识别符框包含于样本项中时,所述框包含在播放轨中包含的temporal_id的最小值及最大值。此外,所述框指示对于解码播放轨中包含的视图所需要的参考的视图。

[0326] 10.5.3.1.1.2 语法

```
class ViewIdentifierBox extends FullBox ('vwid', version=0, flags)
```

{

[0327] unsigned int(2) reserved6 = 0;
 unsigned int(3) min_temporal_id;
 unsigned int(3) max_temporal_id;
 unsigned int(16) num_views;
 for (i=0; i<num_views; i++) {

 unsigned int(6) reserved1 = 0;

 unsigned int(6) layer_id[i];

 unsigned int(10) view_id[i];

 unsigned int(2) base_view_type;

 for (j = 0; j < layer_id[i]; j++) {

 unsigned int(1) depdnet_layer[i][j];

 }

}

}

[0329] 10.5.3.1.1.3 语义

[0330] 当视图识别符框包含于样本项中时,min_temporal_id,max_temporal_id分别选

取在映射到播放轨或层的NAL单元的NAL单元标头扩展中存在的temporal_id 语法元素的相应最小值及最大值。对于AVC流,此选取在或将在首码NAL单元中的值。

[0331] num_views指示当视图识别符框存在于样本项中时在播放轨中包含的视图的数目。

[0332] layer_id[i]指示当视图识别符框包含于样本项中时在包含于播放轨中的层的NAL单元标头中的nuh_layer_id语法元素的值。

[0333] view_id指示具有等于layer_id[i]的nuh_layer_id的第i层的视图识别符,如在ISO/IEC 23008-2的附录F中所指定。

[0334] base_view_type指示视图是否为基础视图(是否虚拟)。其选取以下值:

[0335] 0指示视图既非基础视图,也非虚拟基础视图。

[0336] 1应用以标注MV-HEVC位流的非虚拟基础视图。

[0337] 2为保留值且不应使用。

[0338] 3指示具有view_id[i]的视图为虚拟基础视图。具有view_id[i]的相应独立译码的非基础视图驻留于另一播放轨中,当base_view_type等于3时,随后 num_ref_views应等于0。

[0339] dependent_layer[i][j]指示具有等于j的nuh_layer_id的第j层可为具有等于layer_id[i]的nuh_layer_id的层的直接或是间接参考的层。当视图识别符框包含于样本项中时,推荐其指示同一样本项中的参考的视图。

[0340] 10.5.3.2 样本项定义

[0341] 样本项类型: 'hvc2'、'hev2'、'mhc1'、'mhv1'、'mhcC'

[0342] 容器:样本描述框('stsd')

[0343] 必选: 'hvc1'、'hev1'、'hvc2'、'hev2'、'mhc1' 或 'mhv1' 框中的一者为必选的

[0344] 数量:可存在一或多个样本项

[0345] 如果MV-HEVC基本流含有可使用的HEVC兼容基础层,那么应使用HEVC视觉样本项('hvc1'、'hev1'、'hvc2'、'hev2')。此处,所述项目一开始应含有HEVC配置框,可能接着为如下所定义的MV-HEVC配置框。HEVC配置框以文件记载涉及HEVC 兼容基础层的简档、层级及可能也有参数集,如由HEVCDecoderConfigurationRecord所定义。MV-HEVC配置框以文件记载涉及存储于MVHEVCConfigurationBox中的含有非基础视图的整个流的简档、层级及可能也有参数集信息(如由 MVHEVCDecoderConfigurationRecord所定义)。

[0346] 对于所有样本项 'hvc1'、'hev1'、'hvc2'、'hev2' ,样本项中的宽度及高度字段以文件记载HEVC基础层。对于MV-HEVC样本项 ('mhc1'、'mhv1') ,宽度及高度以文件记载通过解码整个流的任一单一视图所达成的分辨率。

[0347] 如果MV-HEVC基本流不含有可使用的HEVC基础层,那么应使用MV-HEVC视觉样本项 ('mhc1'、'mhv1') 。MV-HEVC视觉样本项应含有如下所定义的MV-HEVC配置框。此包含MVHEVCDecoderConfigurationRecord,如在此国际标准中所定义。

[0348] 在任一给定样本项中的MV-HEVC及HEVC配置中的lengthSizeMinusOne字段应具有相同值。

[0349] 此处也适用如在6.5.3.1.1中以文件记载的对于相同项类型 'hvc1' 及 'hev1' 的要求。

[0350] MVHEVCConfigurationBox 可存在于 ‘hvc1’、‘hev1’、‘hvc2’、‘hev2’ 样本项中。在此些情况下,以下HEVCMVHEVCSampleEntry或HEVC2MVHEVCSampleEntry定义分别适用。

[0351] 基础类别VisualSampleEntry中的Compressorname指示使用的压缩器的名称,其中推荐值“\014MV-HEVC Coding”(\016为14,以字节为单元的字符串“MV-HEVC coding”的长度)。

[0352] 直接或通过来自提取器的参考解码存在于视频流的样本数据中的NAL单元所需要的参数集应存在于所述视频流的解码器配置中或相关的参数集流(如果使用)中。

[0353] 下表展示对于视频播放轨的样本项(当MV-HEVC基本流存储于一或多个播放轨中时)、配置及MV-HEVC工具的所有可能使用。

[0354] 表14 用于HEVC及MV-HEVC播放轨的样本项的使用

[0355]

样本项名称	具有配置记录	含义
‘hvc1’ 或 ‘hev1’	仅 HEVC 配置	无具有大于 0 的 nuh_layer_id 的 NAL 单元的普通 HEVC 播放轨; 提取器及聚集器不应存在。
‘hvc1’ 或 ‘hev1’	HEVC 及 MV-HEVC 配置	具有 nuh_layer_id 等于 0 的 NAL 单元及 nuh_layer_id 大于 0 的 NAL 单元的 MV-HEVC 播放轨; 提取器及聚集器可存在; 提取器不应参考具有等于 0 的 nuh_layer_id 的 NAL 单元; 聚集器不应含有但可参考具有等于 0 的 nuh_layer_id 的 NAL 单元。
‘hvc2’ 或 ‘hev2’	仅 HEVC 配置	无具有大于 0 的 nuh_layer_id 的 NAL 单元的普通 HEVC 播放轨; 提取器可存在且用以参考 NAL 单元; 聚集器可存在以含有且参考 NAL 单元。
‘hvc2’ 或 ‘hev2’	HEVC 及 MV-HEVC 配置	具有 nuh_layer_id 等于 0 的 NAL 单元及 nuh_layer_id 大于 0 的 NAL 单元的 MV-HEVC 播放轨; 提取器及聚集器可存在; 提取器可参考任何 NAL 单元; 聚集器可皆含有且参考任何 NAL 单元。
‘mhv1’ 或 ‘mhc1’	MV-HEVC 配置	无具有等于 0 的 nuh_layer_id 的 NAL 单元的 MV-HEVC 播放轨; 提取器可存在且用以参考 NAL 单元; 聚集器可存在以含有且参考 NAL 单元。

[0356] 以下中的样本项mvhevc-type为{mhv1,mhc1}中的一者。

[0357] 10.5.3.3 语法

```

class MVHEVCCConfigurationBox extends Box('mhcC') {
    MVHEVCDecoderConfigurationRecord() MVHEVCCConfig;
}

class HEVCMVHEVCSampleEntry() extends HEVCSampleEntry() {
    ViewIdentifierBox view_identifiers; //任选
    MVHEVCCConfigurationBox mvhevcconfig; //任选
[0358]
}

class HEVC2MVHEVCSampleEntry() extends HEVC2SampleEntry() {
    ViewIdentifierBox view_identifiers; //任选
    MVHEVCCConfigurationBox mvhevcconfig; //任选
}

//在播放轨并不 HEVC 兼容的情况下使用这个
class MVHEVCSampleEntry() extends VisualSampleEntry(mvhevc-type) {
    MVHEVCCConfigurationBox mvhevcconfig; //必选
    ViewIdentifierBox view_identifiers; //必选
[0359]
    MPEG4BitRateBox bitrate; //任选
    MPEG4ExtensionDescriptorsBox descr; //任选
}

```

[0360] 10.5.4 用于MV-HEVC的子样本的定义

[0361] 类似于针对SHVC定义的定义来定义用于MV-HEVC的子样本的定义。

[0362] 10.5.5 处置非输出样本

[0363] 类似于针对SHVC定义的处置来处置用于MV-HEVC的非输出样本。

[0364] 以下展示到附录A的改变。

[0365] 附录A(标准)

[0366] 流中结构

[0367] A.1 介绍

[0368] 聚集器及提取器为实现NAL单元的高效分组或从其它播放轨提取NAL单元的文件格式内部结构。

[0369] 聚集器及提取器使用NAL单元语法。这些结构被看作样本结构的上下文中的NAL单元。在存取样本时,必须移除聚集器(留下其含有或参考的NAL单元)且提取器必须由其参考的数据替换。聚集器及提取器必须不存在于文件格式外的流中。

[0370] 这些结构使用由ISO/IEC 14496-10或ISO/IEC 23008-2针对应用/输送层保留的NAL 单元类型。

[0371] 注意以下来自ISO/IEC 14496-10:

[0372] “注意——可使用NAL单元类型0及24..31,如由应用程序所确定。在此推荐国际标

准中未指定针对这些nal_unit_type值的解码过程。”

[0373] 注意以下来自ISO/IEC 23008-2:

[0374] “注意1——可使用在UNSPEC48..UNSPEC63的范围中的NAL单元类型,如由应用程序所确定。在此规格中未指定针对这些nal_unit_type值的解码过程。由于不同应用程序可将这些NAL单元类型用于不同目的,因此必须在产生具有这些nal_unit_type值的NAL单元的编码器的设计中及在解译具有这些nal_unit_type值的NAL单元的内容的解码器的设计中格外小心。”

[0375] A.2 聚集器

[0376] A.2.1 定义

[0377] 此子条款描述使NALU映射群组项能够一致且重复的聚集器。(见附录B)。

[0378] 聚集器用以对同一样本的NAL单元分组。

[0379] 为了ISO/IEC 14496-10视频的存储,以下规则适用:

[0380] -聚集器使用与SVC VCL NAL单元或MVC VCL NAL单元相同但具有不同 NAL单元类型值的NAL单元标头。

[0381] -当聚集器的NAL单元语法(在ISO/IEC 14496-10的7.3.1中所指定)的 svc_extension_flag等于1时,SVC VCL NAL单元的NAL单元标头用于聚集器。否则,将MVC VCL NAL单元的NAL单元标头用于聚集器。

[0382] 为了ISO/IEC 23008-2视频的存储,聚集器使用如在ISO/IEC 23008-2中所定义的 NAL单元标头,其对于普通HEVC、SHVC及MV-HEVC具有相同语法。

[0383] 聚集器可通过包含来将NAL单元聚集于其内(在由其长度指示的大小内),且也通过参考聚集其后的NAL单元(在由其内的additional_bytes字段指示的区域内)。当流由AVC或HEVC文件读取器扫描时,仅将包含的NAL单元看作“在聚集器内”。此准许AVC或HEVC文件读取器跳过整个一组不需要的NAL单元(当其通过包含而经聚集时)。此也准许AVC或HEVC读取器不跳过需要的NAL单元,而让其保持在流中(当其通过参考而经聚集时)。

[0384] 聚集器可以用以将基础层或基础视图NAL单元分组。如果将这些聚集器用于‘avc1’、‘hvc1’或‘hev1’播放轨中,那么聚集器不应使用基础层或基础视图NAL单元的包含,而使用基础层或基础视图NAL单元的参考(聚集器的长度仅包含其标头,且由聚集器参考的NAL单元由additional_bytes指定)。

[0385] 当聚集器由具有等于零的data_length的提取器或由映射样本群组参考时,将聚集器作为聚集包含及参考的字节对待。

[0386] 聚集器可包含或参考提取器。提取器可从聚集器提取。聚集器必须不直接包含或参考另一聚集器;然而,聚集器可包含或参考参考聚集器的提取器。

[0387] 当扫描流时:

[0388] a) 如果聚集器未经辨识(例如,由AVC或HEVC读取器或解码器),那么易于与其包含的内容一起被舍弃;

[0389] b) 如果不需要聚集器(即,其属于不当层),那么其及其通过包含及参考两者的内容易于被舍弃(使用其长度及additional_bytes字段);

[0390] c) 如果需要聚集器,那么易于舍弃其标头且保留其内容。

[0391] 将聚集器存储于如任一其它NAL单元的样本内。

[0392] 所有NAL单元按解码次序保持处于聚集器内。

[0393] A.2.2 语法

```
class aligned(8) Aggregator (AggregatorSize) {
    NALUnitHeader();
    unsigned int i = sizeof(NALUnitHeader());
    unsigned int((lengthSizeMinusOne+1)*8)
    additional_bytes;
    i += lengthSizeMinusOne+1;
    while (i<AggregatorSize) {
        unsigned int((lengthSizeMinusOne+1)*8)
        NALUnitLength;
        unsigned int(NALUnitLength*8) NALUnit;
        i += NALUnitLength+lengthSizeMinusOne+1;
    };
}
```

[0396] A.2.3 语义

[0397] 变量AggregatorSize的值等于聚集器NAL单元的大小,且函数sizeof (X) 按字节返回字段X的大小。

[0398] NALUnitHeader () :SVC及MVC VCL NAL单元的首先四个字节,或ISO/IEC23008-2NAL单元的首先两个字节。

[0399] nal_unit_type应设定到聚集器NAL单元类型(对于ISO/IEC 14496-10视频为类型30且对于ISO/IEC 23008-2视频为类型48)。

[0400] 对于包含或参考SVC NAL单元的聚集器,以下应适用。

[0401] 应如在ISO/IEC 14496-10中所指定设定forbidden_zero_bit及 reserved_three_2bits。

[0402] 应如在A.4中所指定设定其它字段(nal_ref_idc、idr_flag、priority_id、no_inter_layer_pred_flag、dependency_id、quality_id、temporal_id、use_ref_base_pic_flag、discardable_flag及output_flag)。

[0403] 对于包含或参考MVC NAL单元的聚集器,以下应适用。

[0404] 应如在ISO/IEC 14496-10中所指定设定forbidden_zero_bit及reserved_one_bit。

[0405] 应如在A.5中所指定设定其它字段(nal_ref_idc、non_idr_flag、priority_id、view_id、temporal_id、anchor_pic_flag及inter_view_flag)。

[0406] 对于包含或参考ISO/IEC 23008-2NAL单元的聚集器,以下应适用。

[0407] 应如在ISO/IEC 23008-2中所指定设定forbidden_zero_bit。

[0408] 应如在A.6中所指定设定其它字段(nuh_layer_id及 nuh_temporal_id_plus1)。

[0409] additional_bytes:当此聚集器由具有等于零的data_length或映射样本群组的

提取器参考时应被考虑为聚集的在此聚集器NAL单元后的字节的数目。

[0410] NALUnitLength: 指定NAL单元遵循的大小(以字节计)。此字段的大小通过lengthSizeMinusOne字段指定。

[0411] NALUnit: 如在ISO/IEC 14496-10或ISO/IEC 23008-2中指定的NAL单元,包含NAL单元标头。NAL单元的大小由NALUnitLength指定。

[0412] A.3 提取器

[0413] A.3.1 定义

[0414] 此子条款描述实现通过参考从其它播放轨提取NAL单元数据的播放轨的紧密形成的提取器。

[0415] 聚集器可包含或参考提取器。提取器可参考聚集器。当提取器由需要其的文件读取器处理时,提取器由其参考的字节逻辑替换。那些字节不必含有提取器;提取器不必直接或间接参考另一提取器。

[0416] 注意参考的播放轨可含有提取器,即使由提取器参考的数据不必。

[0417] 提取器含有通过类型‘scal’的播放轨参考从另一播放轨提取数据的指令,所述另一播放轨与提取器驻留于其中的播放轨有联系。

[0418] 复制的字节应为以下中的一者:

[0419] a) 一整个NAL单元;注意,当参考聚集器时,复制包含的及参考的字节

[0420] b) 一个以上整个NAL单元

[0421] 在两个情况下,提取的字节开始于有效长度字段及NAL单元标头。

[0422] 仅从经由指示的‘scal’播放轨参考参考的播放轨中的单一识别的样本复制字节。对准是在解码时间,即,仅使用时间到样本表,接着为样本数目的计数的偏移。提取器为媒体级概念且因此在考虑任一编辑列表前适用于目的地播放轨。(然而,通常将预期两个播放轨中的编辑列表将相同)。

[0423] A.3.2 语法

```
class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
    data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
    data_length;
}
```

[0425] A.3.3 语义

[0426] NALUnitHeader():SVC及MVC VCL NAL单元的首先四个字节,或ISO/IEC23008-2NAL单元的首先两个字节。

[0427] 应将nal_unit_type设定到提取器NAL单元类型(对于ISO/IEC 14496-10视频为类

型31且对于ISO/IEC 23008-2视频为类型49)。

[0428] 对于参考SVC NAL单元的提取器,以下应适用。

[0429] 应如在ISO/IEC 14496-10中所指定设定forbidden_zero_bit及 reserved_three_2bits。

[0430] 应如在A.4中所指定设定其它字段(nal_ref_idc、idr_flag、priority_id、no_inter_layer_pred_flag、dependency_id、quality_id、temporal_id、use_ref_base_pic_flag、discardable_flag及output_flag)。

[0431] 对于参考MVC NAL单元的提取器,以下应适用。

[0432] 应如在ISO/IEC 14496-10中所指定设定forbidden_zero_bit及 reserved_one_bit。

[0433] 应如在A.5中所指定设定其它字段(nal_ref_idc、non_idr_flag、priority_id、view_id、temporal_id、anchor_pic_flag及inter_view_flag)。

[0434] 对于参考ISO/IEC 23008-2NAL单元的提取器,以下应适用。

[0435] 应如在ISO/IEC 23008-2中所指定设定forbidden_zero_bit。

[0436] 应如在A.6中所指定设定其它字段(nuh_layer_id及nuh_temporal_id_plus1)。

[0437] track_ref_index指定类型‘scal’的播放轨参考的索引以用以找到提取数据所来自的播放轨。数据提取自的所述播放轨中的样本在时间上对准或在媒体解码时间线中最紧接于前(即,仅使用时间到样本表),通过具有含有提取器的样本的sample_offset所指定的偏移来调整。第一播放轨参考具有索引值1;值0为保留的。

[0438] sample_offset给出应用作信息的源的有联系的播放轨中的样本的相对索引。样本0(零)为具有与含有提取器的样本的解码时间相比相同或最紧靠于前的解码时间的样本;样本1(一)为下一个样本,样本-1(负1)为先前样本,等等。

[0439] data_offset:在参考样本内的复制的第一字节的偏移。如果提取开始于所述样本中的数据的第一字节,那么偏移选取值0。偏移应参考NAL单元长度字段的开头。

[0440] data_length:复制的字节的数目。如果此字段选取值0,那么复制整个单一参考的NAL单元(即,复制的长度从由数据偏移参考的长度字段取得,在聚集器的情况下,由additional_bytes字段扩增)。

[0441] 注意如果两个播放轨使用不同lengthSizeMinusOne值,那么提取的数据将需要重新格式化以符合目的地播放轨的长度字段大小。

[0442] A.4 SVC的NAL单元标头值

[0443] 提取器及聚集器皆使用NAL单元标头SVC扩展。由提取器提取或由聚集器聚集的NAL单元为通过递归式检验聚集器或提取器NAL单元的内容而参考或包含的所有那些 NAL单元。

[0444] 字段nal_ref_idc、idr_flag、priority_id、temporal_id、dependency_id、quality_id、discardable_flag、output_flag、use_ref_base_pic_flag及no_inter_layer_pred_flag应选取以下值:

[0445] nal_ref_idc应设定到在所有提取的或聚集的NAL单元中的字段的最高值。

[0446] idr_flag应设定到在所有提取的或聚集的NAL单元中的字段的最高值。

[0447] priority_id、temporal_id、dependency_id及quality_id应分别设定到在所有提

取的或聚集的NAL单元中的字段的最低值。

[0448] 如果且仅如果所有提取的或聚集的NAL单元具有设定到1的discardable_flag,那么应将discardable_flag设定到1,且否则,将其设定到0。

[0449] 如果聚集的或提取的NAL单元中的至少一者具有设定到1的output_flag,将应将此旗标设定到1,且否则,将其设定到0。

[0450] 如果且仅如果提取的或聚集的VCL NAL单元中的至少一者具有设定到1的 use_ref_base_pic_flag,那么应将use_ref_base_pic_flag设定到1,且否则,将其设定到0。

[0451] 如果且仅如果所有提取的或聚集的VCL NAL单元具有设定到1的 no_inter_layer_pred_flag,那么应将no_inter_layer_pred_flag设定到1,且否则,将其设定到0。

[0452] 如果提取的或聚集的NAL的组合为空,那么此些字段中的每一者选取与映射的层描述一致的值。

[0453] 注意聚集器可将具有不同可缩放性信息的NAL单元分组。

[0454] 注意聚集器可用以将可不由NAL单元标头传信的属于可缩放性等级的 NAL单元(例如,属于相关区的NAL单元)分组。此些聚集器的描述可通过层描述及NAL单元映射群组进行。在此情况下,在一个实例中可出现具有相同可缩放性信息的一个以上聚集器。

[0455] 注意如果多个可缩放播放轨参考相同媒体数据,那么聚集器应仅将具有相同可缩放性信息的NAL单元分组。此确保所得模式可由播放轨中的每一者存取。

[0456] 注意如果特定层中无NAL单元于存取单元中,那么可存在空聚集器(其中聚集器的长度仅包含标头,且additional_bytes为零)。

[0457] A.5 MVC的NAL单元标头值

[0458] 聚集器及提取器皆使用NAL单元标头MVC扩展。由提取器提取或由聚集器聚集的NAL单元为通过递归式检验聚集器或提取器NAL单元的内容而参考或包含的所有那些 NAL单元。

[0459] 字段nal_ref_idc、non_idr_flag、priority_id、view_id、temporal_id、anchor_pic_flag 及inter_view_flag应选取以下值:

[0460] nal_ref_idc应设定到在所有聚集的或提取的NAL单元中的字段的最高值。

[0461] non_idr_flag应设定到在所有聚集的或提取的NAL单元中的字段的最低值。

[0462] priority_id及temporal_id应分别设定到在所有聚集的或提取的NAL单元中的字段的最低值。

[0463] view_id应设定到所有聚集的或提取的VCL NAL单元当中的具有最低视图次序索引的VCL NAL单元的view_id值。

[0464] anchor_pic_flag及inter_view_flag应分别设定到在所有聚集的或提取的VCL NAL单元中的字段的最高值。

[0465] 如果提取的或聚集的NAL的组合为空,那么此些字段中的每一者选取与映射的层描述一致的值。

[0466] A.6 用于ISO/IEC 23008-2的NAL单元标头值

[0467] 聚集器及提取器皆使用如在ISO/IEC 23008-2中指定的NAL单元标头。由提取器提取或由聚集器聚集的NAL单元为通过递归式检验聚集器或提取器NAL单元的内容而参考或包含的所有那些NAL单元。

- [0468] 字段nuh_layer_id及nuh_temporal_id_plus1应如下设定：
- [0469] nuh_layer_id应设定到所有聚集的或提取的NAL单元中的字段的最低值。
- [0470] nuh_temporal_id_plus1应设定到所有聚集的或提取的NAL单元中的字段的最低值。

[0471] 在一个替代性实例中,定义新结构、表或样本群组以文件记载所有IRAP存取单元,如在MV-HEVC WD5或SHVC WD3的附录A中所定义。替代地,定义所述新结构、表或样本群组以文件记载所有IRAP存取单元(如在MV-HEVC WD5或SHVC WD3的附录 F中所定义),但不包含所有经译码图片为IRAP图片的那些存取单元。在另一替代性实例中,重新定义同步样本样本群组项SyncSampleEntry包含在保留的位中的指定属于此群组的样本中的所有图片为IDR图片、CRA图片或BLA图片的一者中的 aligned_sync_flag。在另一替代性实例中,定义用于SHVC及MV-HEVC的共同文件格式包含来自SHVC及MV-HEVC文件格式的所有共同方面,且仅将SHVC及MV-HEVC 文件格式重新定义为仅包含仅与所述扩展有关的方面。在另一替代性实例中,定义 SHVC 元数据样本项 SHVC Metadata Sample Entry 及 SHVC Metadata Sample Config Box,且也定义元数据样本语句类型scalabilityInfoSHVCStatement。

[0472] 图2为说明可实施本发明中所描述的技术的实例视频编码器20的框图。视频编码器20可经配置以输出单一视图、多视图、可缩放、3D及其它类型的视频数据。视频编码器20可经配置以将视频输出到后处理实体27。后处理实体27意欲表示可处理来自视频编码器20的经编码视频数据的视频实体(例如,MANE或拼接/编辑装置)的实例。在一些情况下,后处理实体可为网络实体的实例。在一些视频编码系统,后处理实体27 及视频编码器20可为分开的装置的部分,而在其它情况下,关于后处理实体27描述的功能性可由包括视频编码器20的同一装置执行。后处理实体27可为视频装置。在一些实例中,后处理实体27可与图1的文件产生装置34相同。

[0473] 视频编码器20可执行视频切片内的视频块的帧内译码及帧间译码。帧内译码依赖于空间预测以减小或移除给定视频帧或图片内的视频中的空间冗余。帧间译码依赖于时间预测以减小或移除视频序列的邻近帧或图片内的视频的时间冗余。帧内模式(I模式) 可指若干基于空间的压缩模式中的任一者。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者。

[0474] 在图2的实例中,视频编码器20包含分割单元35、预测处理单元41、滤波器单元63、参考图片存储器64、求和器50、变换处理单元52、量化单元54及熵编码单元56。预测处理单元41包含运动估计单元42、运动补偿单元44及帧内预测处理单元46。为了视频块重建,视频编码器20也包含反量化单元58、反变换处理单元60及求和器 62。滤波器单元63意欲表示一或多个环路滤波器,例如,解块滤波器、自适应环路滤波器(ALF) 及样本自适应偏移(SAO) 滤波器。尽管滤波器单元63在图2中展示为环路滤波器,但在其它配置中,滤波器单元63可实施为后环路滤波器。

[0475] 视频编码器20的视频数据存储器可存储待由视频编码器20的组件编码的视频数据。存储于视频数据存储器中的视频数据可(例如)从视频源18获得。参考图片存储器 64可为存储参考视频数据用于由视频编码器20在编码视频数据过程中使用(例如,在帧内或帧间译码模式中的)的参考图片存储器。视频数据存储器及参考图片存储器64可由多种存储器

装置中的任一者形成,例如,动态随机存取存储器(DRAM)(包含同步 DRAM(SDRAM))、磁阻式 RAM(MRAM)、电阻式RAM(RRAM)或其它类型的存储器装置。视频数据存储器及参考图片存储器64可由相同的存储器装置或单独存储器装置来提供。在各种实例中,视频数据存储器可与视频编码器20的其它组件一起在芯片上,或相对于那些组件在芯片外。

[0476] 如图2中所展示,视频编码器20接收视频数据,且分割单元35将数据分割成视频块。此分割也可包含分割成切片、平铺块或其它较大单元以及视频块分割,例如,根据 LCU 及CU的四分树结构。视频编码器20大体上说明编码待编码视频切片内的视频块的组件。可将切片划分成多个视频块(且可能划分成被称作平铺块的视频块集合)。预测处理单元41可基于误差产生(例如,译码速率及失真的等级)选择用于当前视频块的多个可能译码模式中的一者,例如,多个帧内译码模式中的一者或多个帧间译码模式中的一者。预测处理单元41可将所得经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以重构经编码块以用于用作参考图片。

[0477] 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相同的帧或切片中的一或多个相邻块的帧内预测性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及运动补偿单元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。

[0478] 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切片的帧间预测模式。预定图案可将序列中的视频切片指明为P切片、B切片或GPB切片。运动估计单元42及运动补偿单元44可高度集成,但为概念目的而分开来说明。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,运动向量估计视频块的运动。举例来说,运动向量可指示当前视频帧或图片内的视频块的PU相对于参考图片内的预测性块的位移。

[0479] 预测性块为就像素差而言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,视频编码器 20 可计算存储于参考图片存储器64中的参考图片的次整数像素位置的值。举例来说,视频编码器20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可执行相对于全像素位置及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。

[0480] 运动估计单元42通过比较PU的位置与参考图片的预测性块的位置而计算经帧间译码切片中的视频块的PU的运动向量。参考图片可从第一参考图片列表(列表0)或第二参考图片列表(列表1)选择,所述列表中的每一者识别存储于参考图片存储器64中的一或多个参考图片。运动估计单元42将经计算运动向量发送到熵编码单元56及运动补偿单元44。

[0481] 由运动补偿单元44所执行的运动补偿可涉及基于由运动估计所确定的运动向量而提取或产生预测性块,可能执行子像素精确度的内插。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的一者中定位运动向量所指向的预测性块。视频编码器20可通过从正被译码的当前视频块的像素值减去预测性块的像素值来形成残余视频块,从而形成像素差值。像素差值形成用于块的残余数据,且可包含明度及色度差分量两者。求和器50表示执行此减法运算的所述组件。运动补偿单元44 也可产生与视频块及视频切片相关联的语法元素以供视频解码器30在解码视频切片的视频块时使用。

[0482] 如上文所描述,作为由运动估计单元42及运动补偿单元44所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。详言地说,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) 在分开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可使用用于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,并在所测试模式当中选择具有最佳速率-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可从各种经编码块的失真及速率计算比率以确定哪种帧内预测模式展现所述块的最佳速率-失真值。

[0483] 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可将指示用于块的选定帧内预测模式的信息提供到熵编码单元56。熵编码单元56可根据本发明的技术编码指示选定帧内预测模式的信息。视频编码器20可在所发射的位流中包含配置数据,其可包含以下各者:多个帧内预测模式索引表及多个经修改的帧内预测模式索引表(也称作码字映射表);各种块的编码上下文的定义;及待用于所述上下文中的每一者的最有可能的帧内预测模式、帧内预测模式索引表及经修改的帧内预测模式索引表的指示。

[0484] 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,视频编码器20可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)。

[0485] 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可接着执行包含经量化变换系数的矩阵的扫描。替代地,熵编码单元56可执行扫描。

[0486] 在量化后,熵编码单元56可熵编码表示经量化变换系数的语法元素。举例来说,熵编码单元56可执行上下文自适应可变长度译码(CAVLC)、上下文自适应二进制算术译码(CABAC)、基于语法的上下文自适应二进制算术译码(SBAC)、概率区间分割熵(PIPE)译码或另一熵编码方法或技术。在由熵编码单元56熵编码之后,经编码位流可被发射到视频解码器30,或经存档以供视频解码器30稍后发射或检索。熵编码单元56也可熵编码当前正译码的视频切片的运动向量及其它语法元素。

[0487] 反量化单元58及反变换处理单元60分别应用反量化及反变换以重建构像素域中的残余块,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中的一者内的参考图片中的一者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构残余块,以计算子整数像素值以用于运动估计中。求和器62将经重建构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参考图片存储器64中的参考块。参考块可由运动估计单元42及运动补偿单元44用作参考块以帧间预测随后视频帧或图片中的块。

[0488] 视频编码器20表示经配置以产生可使用本发明中所描述的文件格式技术存储的视频数据的视频译码器的实例。

[0489] 图3为说明可实施本发明中所描述的技术的实例视频解码器30的框图。视频解码器30可经配置以解码单一视图、多视图、可缩放、3D及其它类型的视频数据。在图3 的实例中,视频解码器30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及参考图片存储器92。预测处理单元81包含运动补偿单元82及帧内预测处理单元84。视频解码器30可在一些实例中执行与关于来自图2的视频编码器20描述的编码遍次大体上互逆的解码遍次。

[0490] 经译码图片缓冲器(CPB)79可接收且存储位流的经编码视频数据(例如,NAL单元)。存储于CPB 79中的视频数据可(例如)经由视频数据的有线或无线网络通信或通过存取物理数据存储媒体从链路16获得,例如,从例如相机等局部视频源。CPB 79可形成存储来自经编码视频位流的经编码视频数据的视频数据存储器。CPB 79可为存储参考视频数据用于由视频解码器30在解码视频数据过程中使用(例如,在帧内或帧间译码模式中)的参考图片存储器。CPB 79及参考图片存储器92可由多种存储器装置中的任一者形成,例如,动态随机存取存储器(DRAM)(包含同步DRAM(SDRAM))、磁阻式RAM(MRAM)、电阻式RAM(RRAM)或其它类型的存储器装置。CPB 79及参考图片存储器92可由同一存储器装置或分开的存储器装置提供。在各种实例中,CPB 79可与视频解码器30的其它组件一起在芯片上,或相对于那些组件在芯片外。

[0491] 在解码过程期间,视频解码器30从视频编码器20接收表示经编码视频切片的视频块的经编码视频位流及相关联的语法元素。视频解码器30可从网络实体29接收经编码视频位流。网络实体29可(例如)为服务器、MANE、视频编辑器/拼接器或经配置以实施以上描述的技术中的一或更多的其它此类装置。网络实体29可或可不包含视频编码器,例如,视频编码器20。本发明中所描述的技术中的一些可由网络实体29在网络实体29 将经编码视频位流发射到视频解码器30前实施。在一些视频解码系统中,网络实体29 及视频解码器30可为分开的装置的部分,而在其它情况下,关于网络实体29描述的功能性可由包括视频解码器30的同一装置执行。可将网络实体29考虑为视频装置。此外,在一些实例中,网络实体29为图1的文件产生装置34。

[0492] 视频解码器30的熵解码单元80熵解码位流的特定语法元素以产生经量化系数、运动向量及其它语法元素。熵解码单元80将运动向量及其它语法元素转递到预测处理单元81。视频解码器30可在视频切片层级及/或视频块层级接收语法元素。

[0493] 当视频切片经译码为经帧内译码(I)切片时,预测处理单元81的帧内预测处理单元 84可基于来自当前帧或图片的先前经解码块的经传信帧内预测模式及数据来产生用于当前视频切片的视频块的预测数据。当将视频帧经译码为经帧间译码(即,B、P或GPB)切片时,预测处理单元81的运动补偿单元82基于从熵解码单元80接收的运动向量及其它语法元素来产生当前视频切片的视频块的预测性块。可根据在参考图片列表中的一者内的参考图片中的一者产生所述预测性块。视频解码器30可基于存储于参考图片存储器92中的参考图片使用默认建构技术来建构参考帧列表:列表0及列表1。

[0494] 运动补偿单元82通过剖析运动向量及其它语法元素来确定用于当前视频切片的视频块的预测信息,并使用所述预测信息以产生正解码的当前视频块的预测性块。举例来说,运动补偿单元82使用接收的语法元素中的一些确定用以译码视频切片的视频块的预测模式(例如,帧内预测或帧间预测)、帧间预测切片类型(例如,B切片、P切片或GPB 切片)、切

片的参考图片列表中的一或更多的建构信息、切片的每一经帧间编码视频块的运动向量、切片的每一经帧间译码视频块的帧间预测状态及解码当前视频切片中的视频块的其它信息。

[0495] 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可使用如由视频编码器20在视频块的编码期间所使用的内插滤波器,以计算参考块的次整数像素的内插值。在此情况下,运动补偿单元82可从所接收语法元素确定由视频编码器20所使用的内插滤波器并可使用所述内插滤波器产生预测性块。

[0496] 反量化单元86将位流中所提供且由熵解码单元80解码的经量化变换系数反量化(即,解量化)。反量化过程可包含使用由视频编码器20计算的用于视频切片中的每一视频块的量化参数,以确定量化程度及(同样)应应用的反量化程度。反变换处理单元88将反变换(例如,反DCT、反整数变换或概念上类似的反变换过程)应用于变换系数,以便产生像素域中的残余块。

[0497] 在运动补偿单元82基于运动向量及其它语法元素产生当前视频块的预测性块后,视频解码器30通过将来自反变换处理单元88的残余块与由运动补偿单元82所产生的对应预测性块求和而形成经解码视频块。求和器90表示执行此求和操作的所述组件。如果需要,也可使用其它环路滤波器(在译码环路中或在译码环路后)使像素转变平滑,或以其它方式改进视频质量。滤波器单元91意欲表示一或多个环路滤波器,例如,解块滤波器、自适应环路滤波器(ALF)及样本自适应偏移(SAO)滤波器。尽管滤波器单元91在图3中展示为环路滤波器,但在其它配置中,滤波器单元91可实施为后环路滤波器。接着,将给定帧或图片中的经解码视频块存储于存储用于随后运动补偿的参考图片的参考图片存储器92中。参考图片存储器92也存储用于稍后在显示装置(例如,图1的显示装置32)上呈现的经解码视频。

[0498] 图3的视频解码器30表示经配置以解码可使用本发明中所描述的文件格式技术存储的视频数据的视频解码器的实例。

[0499] 图4为说明形成网络100的部分的装置的实例集合的框图。在此实例中,网络100包含路由装置104A、104B(路由装置104)及转码装置106。路由装置104及转码装置106意欲表示可形成网络100的部分的少量装置。例如交换器、集线器、网关、防火墙、桥接器及其它这些装置的其它网络装置也可包含于网络100内。此外,可沿着服务器装置102与客户端装置108之间的网络路径提供额外网络装置。在一些实例中,服务器装置102可对应于源装置12(图1),而客户端装置108可对应于目的地装置14(图1)。

[0500] 一般来说,路由装置104实施一或多个路由协议以经由网络100交换网络数据。在一些实例中,路由装置104可经配置以执行代理服务器或高速缓冲存储器操作。因此,在一些实例中,路由装置104可被称作代理装置。一般来说,路由装置104执行路由协议以发现经由网络100的路线。通过执行这些路由协议,路由装置104B可发现从自身经由路由装置104A到服务器装置102的网络路线。

[0501] 本发明的技术可由例如路由装置104及转码装置106等网络装置实施,但也可由客户端装置108实施。以此方式,路由装置104、转码装置106及客户端装置108表示经配置以执行本发明的技术的装置的实例。此外,图1的装置及图2中所说明的编码器20及图3中所说明的解码器30也为可经配置以执行本发明的技术中的一或更多的装置的实例。

[0502] 图5为说明根据本发明的一或多个技术的文件300的实例结构的概念图。在图5的

实例中,文件300包含电影框302及多个媒体数据框304。虽然在图5的实例中说明为在同一文件中,但在其它实例中,电影框302及媒体数据框304可在分开的文件中。如上所指示,框可为由唯一类型识别符及长度定义的面向对象建构块。举例来说,框可为 ISOBMFF中的基本语法结构,包含四字符译码框类型、框的字节计数及有效负载。

[0503] 电影框302可含有用于文件300的播放轨的元数据。文件300的每一播放轨可包括媒体数据的连续流。媒体数据框304中的每一者可包含一或多个样本305。样本305中的每一者可包括音频或视频存取单元。如在本发明中其它处所描述,在多视图译码(例如, MV-HEVC及3D-HEVC)及可缩放视频译码(例如,SHVC)中,每一存取单元可包括多个经译码图片。举例来说,存取单元可包含用于每一层的一或多个经译码图片。

[0504] 此外,在图5的实例中,电影框302包含播放轨框306。播放轨框306可围封用于文件300的播放轨的元数据。在其它实例中,电影框302可包含用于文件300的不同播放轨的多个播放轨框。播放轨框306包含媒体框307。媒体框307可含有声明关于播放轨内的媒体数据的信息的所有对象。媒体框307包含媒体信息框308。媒体信息框308 可含有声明播放轨的媒体的特性信息的所有对象。媒体信息框308包含样本表框309。样本表框309可指定样本特定元数据。

[0505] 在图5的实例中,样本表框309包含 SampleToGroup 框310 及 SampleGroupDescription 框312。在其它实例中,除 SampleToGroup 框310 及 SampleGroupDescription 框312之外,样本表框309也可包含其它框,及/或可包含多个 SampleToGroup 框及 SampleGroupDescription 框。SampleToGroup 框310 可将样本(例如,样本305中的特定者)映射到一群样本。SampleGroupDescription 框312 可指定由所述群样本(即,样本群组)中的样本共享的性质。此外,样本表框309可包含多个样本项框311。样本项框311中的每一者可对应于所述群样本中的样本。在一些实例中,样本项框311 为扩展基础样本群组描述类别(如在以上第9.5.5.1.2节中所定义)的随机可存取样本项类别的实例。

[0506] 根据本发明的一或多个技术,SampleGroupDescription 框312 可指定样本群组中的每一样本含有至少一IRAP图片。以此方式,文件产生装置34可产生包括含有用于文件300中的播放轨的元数据的播放轨框306的文件。用于播放轨的媒体数据包括一连串样本305。样本中的每一者可为多层视频数据(例如,SHVC、MV-HEVC或3D-HEVC视频数据)的视频存取单元。此外,作为产生文件300的部分,文件产生装置34可在文件300 中产生以文件记载含有至少一IRAP图片的所有样本305的额外框(即,样本表框309)。换句话说,额外框识别含有至少一IRAP图片的所有样本305。在图5的实例中,额外框定义以文件记载(例如,识别)含有至少一IRAP图片的所有样本305的样本群组。换句话说,所述额外框指定含有至少一IRAP图片的样本305属于样本群组。

[0507] 此外,根据本发明的一或多个技术,样本项框311中的每一者可包含指示对应的样本中的所有经译码图片是否为IRAP图片的值(例如,all_pics_are_IRAP)。在一些实例中,所述值等于1指定所述样本中并非所有经译码图片皆为IRAP图片。所述值等于0指定不需要样本群组中的每一样本中的每一经译码图片为IRAP图片。

[0508] 在一些实例中,当特定样本中并非所有经译码图片皆为IRAP图片时,文件产生装置34在用于所述特定样本的样本项框311中的一者中可包含指示所述特定样本中的 IRAP 图片的数目的值(例如, num_IRAP_pics)。另外,文件产生装置34在用于所述特定样本的样

本项中可包含指示所述特定样本中的IRAP图片的层识别符的值。文件产生装置34在用于所述特定样本的样本项中也可包含指示所述特定样本的IRAP图片中的VCL NAL单元的NAL单元类型的值。

[0509] 此外,在图5的实例中,样本表框309包含子样本信息框314。虽然图5的实例仅展示一个子样本信息框,但样本表框309可包含多个子样本信息框。一般来说,子样本信息框经设计以含有子样本信息。子样本为样本的一系列相邻字节。ISO/IEC 14496-12 指示应针对给定译码系统(例如,H.264/AVC或HEVC)供应子样本的特定定义。

[0510] ISO/IEC 14496-15的第8.4.8节指定用于HEVC的子样本的定义。特定来说,ISO/IEC 14496-15的第8.4.8节指定对于在HEVC流中的子样本信息框(ISO/IEC 14496-12的8.7.7) 的使用,基于子样本信息框的旗标字段的值定义子样本。根据本发明的一或多个技术,如果子样本信息框314中的旗标字段等于5,那么对应于子样本信息框314的子样本含有一个经译码图片及相关联的非VCL NAL单元。相关联的非VCL NAL单元可包含含有可适用于经译码图片的SEI消息的NAL单元及含有可适用于经译码图片的参数集(例如,VPS、SPS、PPS等)的NAL单元。

[0511] 因此,在一个实例中,文件产生装置34可产生文件(例如,文件300),所述文件包括含有用于文件中的播放轨的元数据的播放轨框(例如,播放轨框306)。在此实例中,用于播放轨的媒体数据包括一连串样本,所述样本中的每一者为多层次视频数据(例如,SHVC、MV-HEVC或3D-HEVC视频数据)的视频存取单元。此外,在此实例中,作为文件产生装置34产生文件的部分,文件产生装置34可在文件中产生子样本信息框(例如,子样本信息框314),所述子样本信息框含有指定在所述子样本信息框中给出的子样本信息的类型的旗标。当所述旗标具有特定值时,对应于子样本信息框的子样本含有正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL单元。

[0512] 此外,根据本发明的一或多个技术,如果子样本信息框314的旗标字段等于0,那么子样本信息框314进一步包含DiscardableFlag值,NoInterLayerPredFlag值、LayerId值及TempId值。如果子样本信息框314的旗标字段等于5,那么子样本信息框314可包含DiscardableFlag值、VclNalUnitType值、LayerId值、TempId值、NoInterLayerPredFlag值、SubLayerRefNalUnitFlag值及保留值。

[0513] SubLayerRefNalUnitFlag等于0指示子样本中的所有NAL单元为子层非参考图片的VCL NAL单元,如在ISO/IEC 23008-2(即,HEVC)中所指定。SubLayerRefNalUnitFlag等于1指示子样本中的所有NAL单元为子层参考图片的VCL NAL单元,如在ISO/IEC 23008-2(即,HEVC)中所指定。因此,当文件产生装置34产生子样本信息框314且旗标具有特定值(例如,5)时,文件产生装置34在子样本信息框314中包含指示子样本中的所有NAL单元是否为子层非参考图片的VCL NAL单元的额外旗标。

[0514] DiscardableFlag值指示子样本中的VCL NAL单元的discardable_flag值的值。如在 ISO/IEC 14496-15的第A.4节中所指定,如果且仅如果所有所述提取的或聚集的NAL单元具有设定到1的discardable_flag,那么应将discardable_flag值设定到1,且否则,将其设定到0。如果含有NAL单元的位流可在无NAL单元的情况下正确地解码,那么 NAL单元可具有设定到1的discardable_flag。因此,如果含有NAL单元的位流可在无 NAL单元的情况下正确地解码,那么NAL单元可为“可舍弃的”。子样本中的所有VCL NAL单元应具有相同

discardable_flag值。因此,当文件产生装置34产生子样本信息框314且旗标具有特定值(例如,5)时,文件产生装置34在子样本信息框314中包含指示子样本的所有VCL NAL单元是否可舍弃的额外旗标(例如,discardable_flag)。

[0515] NoInterLayerPredFlag值指示子样本中的VCL NAL单元的inter_layer_pred_enabled_flag的值。如果且仅如果所有提取的或聚集的VCL NAL单元具有设定到1的inter_layer_pred_enabled_flag,那么应将inter_layer_pred_enabled_flag设定到1,且否则,将其设定到0。子样本中的所有VCL NAL单元应具有相同inter_layer_pred_enabled_flag值。因此,当文件产生装置34产生子样本信息框314且旗标具有特定值(例如,5)时,文件产生装置34在子样本信息框314中包含指示是否针对子样本的所有VCL NAL单元启用层间预测的额外值(例如,inter_layer_pred_enabled_flag)。

[0516] LayerId指示子样本中的NAL单元的nuh_layer_id值。子样本中的所有NAL单元应具有相同nuh_layer_id值。因此,当文件产生装置34产生子样本信息框314且旗标具有特定值(例如,5)时,文件产生装置34在子样本信息框314中包含指示子样本的每一NAL单元的层识别符的额外值(例如,LayerId)。

[0517] TempId指示子样本中的NAL单元的TemporalId值。子样本中的所有NAL单元应具有相同TemporalId值。因此,当文件产生装置34产生子样本信息框314且旗标具有特定值(例如,5)时,文件产生装置34在子样本信息框314中包含指示子样本的每一NAL单元的时间识别符的额外值(例如,TempId)。

[0518] VclNalUnitType指示子样本中的VCL NAL单元的nal_unit_type语法元素。nal_unit_type语法元素为NAL单元的NAL单元标头中的语法元素。nal_unit_type语法元素指定NAL单元中含有的RBSP的类型。子样本中的所有nal_unit_type VCL NAL单元应具有相同nal_unit_type值。因此,当文件产生装置34产生子样本信息框314且旗标具有特定值(例如,5)时,文件产生装置34在子样本信息框314中包含指示子样本的VCL NAL单元的NAL单元类型的额外值(例如,VclNalUnitType)。子样本的所有VCL NAL单元具有相同NAL单元类型。

[0519] 图6为说明根据本发明的一或多个技术的文件300的实例结构的概念图。如在ISO/IEC 14496-15的第8.4.9节中所指定,HEVC允许仅用于参考且不用于输出的文件格式样本。举例来说,HEVC允许视频中的未显示的参考图片。

[0520] 此外,ISO/IEC 14496-15的第8.4.9节指定当任一此非输出样本存在于播放轨中时,应如下约束文件。

[0521] 1. 非输出样本应被给予在输出的样本的时间范围外的组成时间。

[0522] 2. 应使用不包含非输出样本的组成时间的编辑列表。

[0523] 3. 当播放轨包含CompositionOffsetBox('ctts')时,

[0524] a. 应使用CompositionOffsetBox的版本1,

[0525] b. 对于每一非输出样本,应将sample_offset的值设定为等于 -2^{31} ,

[0526] c. 播放轨的SampleTableBox('stbl')中应含有CompositionToDecodeBox('cslg'),且

[0527] d. 当对于所述播放轨存在CompositionToDecodeBox时,框中的leastDecodeToDisplayDelta字段的值应等于不包含用于非输出样本的sample_offset值

的CompositionOffsetBox中的最小组成偏移。

[0528] 注意:因此,leastDecodeToDisplayDelta大于-2³¹。

[0529] 如在ISO/IEC 14496-12中所指定,CompositionOffsetBox提供解码时间与组成时间之间的偏移。CompositionOffsetBox包含一组sample_offset值。sample_offset值中的每一者为给出组成时间与解码时间之间的偏移的非负整数。组成时间指将输出样本的时间。解码时间指将解码样本的时间。

[0530] 如上所指示,经译码切片NAL单元可包含切片片段标头。切片片段标头可为经译码切片片段的部分,且可含有涉及切片片段中的第一或所有CTU的数据元素。在HEVC中,切片片段标头包含pic_output_flag语法元素。一般来说,pic_output_flag语法元素包含于图片的切片的第一切片片段标头中。因此,本发明可将图片的切片的第一切片片段标头的pic_output_flag称作图片的pic_output_flag。

[0531] 如在HEVC WD的第7.4.7.1中所指定,pic_output_flag语法元素影响经解码图片输出及移除过程,如在HEVC WD的附录C中所指定。一般来说,如果用于切片片段的切片片段标头的pic_output_flag语法元素为1,那么输出包含对应于所述切片片段标头的切片的图片。另外,如果用于切片片段的切片片段标头的pic_output_flag语法元素为0,那么可解码包含对应于所述切片片段标头的切片的图片,用于用作参考图片,但不输出所述图片。

[0532] 根据本发明的一或多个技术,在ISO/IEC 14496-15的第8.4.9节中对HEVC的参考可由对应的对SHVC、MV-HEVC或3D-HEVC的参考替换。此外,根据一或多个技术的本发明,当存取单元含有具有等于1的pic_output_flag的一些经译码图片及具有等于0的pic_output_flag的一些其它经译码图片时,必须使用至少两个播放轨来存储流。用于所述播放轨中的每一相应者,相应播放轨的每一样本中的所有经译码图片具有相同pic_output_flag值。因此,播放轨的第一者中的所有经译码图片具有等于0的pic_output_flag,且播放轨的第二者中的所有经译码图片具有等于1的pic_output_flag。

[0533] 因此,在图6的实例中,文件产生装置34可产生文件400。类似于在图5的实例中的文件300,文件400包含电影框402及一或多个媒体数据框404。媒体数据框404中的每一者可对应于文件400的不同播放轨。电影框402可含有用于文件400的播放轨的元数据。文件400的每一播放轨可包括媒体数据的连续流。媒体数据框404中的每一者可包含一或多个样本405。样本405中的每一者可包括音频或视频存取单元。

[0534] 如上所指示,在一些实例中,当存取单元含有具有等于1的pic_output_flag的一些经译码图片及具有等于0的pic_output_flag的一些其它经译码图片时,必须使用至少两个播放轨来存储流。因此,在图6的实例中,电影框402包含播放轨框406及播放轨框408。播放轨框406及408中的每一者围封用于文件400的不同播放轨的元数据。举例来说,播放轨框406可围封用于具有pic_output_flag等于0的经译码图片且不具有pic_output_flag等于1的图片的播放轨的元数据。播放轨框408可围封用于具有pic_output_flag等于1的经译码图片且不具有pic_output_flag等于0的图片的播放轨的元数据。

[0535] 因此,在一个实例中,文件产生装置34可产生包括围封(例如,包括)媒体内容的媒体数据框(例如,媒体数据框404)的文件(例如,文件400)。媒体内容包括一连串样本(例如,样本405)。样本中的每一者可为多层视频数据的存取单元。在此实例中,当文件产生装置34响应于位流的至少一存取单元包含具有等于1的图片输出旗标的经译码图片及具有等于0

的图片输出旗标的经译码图片的确定产生文件时,文件产生装置34可使用至少两个播放轨将位流存储于文件中。对于来自至少两个播放轨的每一相应播放轨,在相应播放轨的每一样本中的所有经译码图片具有相同图片输出旗标值。允许输出具有等于1的图片输出旗标的图片,且允许将具有等于0的图片输出旗标的图片用作参考图片,但不允许将其输出。

[0536] 图7为说明根据本发明的一或多个技术的文件产生装置34的实例操作的流程图。图7的操作与本发明的其它流程图中所说明的操作一起为实例。根据本发明的技术的其它实例操作可包含更多、更少或不同动作。

[0537] 在图7的实例中,文件产生装置34产生文件(500)。作为产生文件的部分,文件产生装置34产生含有用于文件中的播放轨的元数据的播放轨框(502)。以此方式,文件产生装置34产生文件,所述文件包括含有用于在文件中的播放轨的元数据的播放轨框。用于所述播放轨的媒体数据包括一连串样本。所述样本中的每一者为多层视频数据的视频存取单元。在一些实例中,文件产生装置34编码所述多层视频数据。

[0538] 此外,作为产生文件的部分,文件产生装置34识别含有至少一IRAP图片的所有样本(504)。此外,文件产生装置34可在文件中产生以文件记载含有至少一IRAP图片的所有样本的额外框(506)。在一些实例中,所述额外框为未在ISOBMFF或其现有扩展中定义的新框。在一些实例中,所述额外框定义以文件记载含有至少一IRAP图片的所有样本的样本群组。举例来说,所述额外框可为或包括包含SampleToGroup框及 SampleGroupDescription框的样本表框。SampleToGroup框识别含有至少一IRAP图片的样本。SampleGroupDescription框指示所述样本群组为含有至少一IRAP图片的一群样本。

[0539] 此外,在图7的实例中,文件产生装置34可产生用于包含至少一IRAP图片的样本中的特定者的样本项(508)。在一些实例中,文件产生装置34可产生用于包含至少一IRAP图片的样本中的每一者相应者的样本项。所述样本项可为 RandomAccessibleSampleEntry,如在以上第9.5.5.1.2节中所定义。

[0540] 如在图7的实例中所说明,作为产生用于特定样本的样本项的部分,文件产生装置34在用于所述特定样本的样本项中可包含指示所述特定样本中的所有经译码图片是否为IRAP图片的值(510)。以此方式,文件产生装置34可在文件中产生包含指示在所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值的样本项。此外,文件产生装置34在用于所述特定样本的样本项中可包含指示所述特定样本的IRAP图片中的VCL NAL单元的NAL单元类型的值(512)。

[0541] 此外,文件产生装置34可确定特定样本中的所有经译码图片是否为IRAP图片(514)。当特定样本中并非所有经译码图片皆为IRAP图片(514的“否”)时,文件产生装置34在用于所述特定样本的样本项中可包含指示所述特定样本中的IRAP图片的数目的值(516)。另外,文件产生装置34在用于所述特定样本的样本项中可包含指示所述特定样本中的IRAP图片的层识别符(例如,nuh_layer_ids)的值。

[0542] 如上所指示,图7仅提供为实例。其它实例不包含图7的每一动作。举例来说,一些实例不包含步骤502、504及508。此外,一些实例不包含步骤510到518中的各者。此外,一些实例包含一或多个额外动作。举例来说,一些实例包含作为产生所述文件的部分而产生同步样本框的额外动作,所述同步样本框包含以文件记载多层视频数据的播放轨的同步样本的同步样本表。播放轨的每一同步样本为播放轨的随机存取样本。在此实例中,如果存取单

元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本。此外,在此实例中,如果存取单元中的每一经译码图片为无RASL图片的IRAP 图片,那么多视图视频译码样本为同步样本。

[0543] 图8为说明根据本发明的一或多个技术的计算装置执行随机存取及/或等级切换的实例操作的流程图。在图8的实例中,计算装置接收文件(550)。在图8的实例中,计算装置可为中间网络装置(例如,MANE、流式传输服务器)、解码装置(例如,目的地装置 14)或另一类型的视频装置。在一些实例中,计算装置可为内容传递网络的部分。

[0544] 在图8的实例中,计算装置可从文件获得含有用于文件中的播放轨的元数据的播放轨框(552)。用于所述播放轨的媒体数据包括一连串样本。在图8的实例中,样本中的每一者为多层视频数据的视频存取单元。

[0545] 此外,在图8的实例中,计算装置可从文件获得额外框(554)。额外框以文件记载含有至少一IRAP图片的所有样本。因此,计算装置可基于额外框中的信息确定含有至少一IRAP图片的所有样本(556)。

[0546] 此外,在一些实例中,计算装置可从文件获得包含指示在所述一连串样本中的特定样本中的所有经译码图片是否为IRAP图片的值的样本项。当特定样本中并非所有经译码图片皆为IRAP图片时,计算装置可从样本项获得指示特定样本中的IRAP图片的数目的值。另外,所述计算装置可从样本项获得指示特定样本中的IRAP图片的层识别符的值。此外,在一些实例中,计算装置可从样本项获得指示所述特定样本的IRAP图片中的VCL NAL单元的NAL单元类型的值。另外,在一些实例中,计算装置可从文件获得同步样本框,所述同步样本框包含以文件记载视频数据的播放轨的同步样本的同步样本表。在此些实例中,播放轨的每一同步样本为播放轨的随机存取样本,如果存取单元中的每一经译码图片为IRAP图片,那么可缩放视频译码样本为同步样本,且如果存取单元中的每一经译码图片为无RASL图片的IRAP图片,那么多视图视频译码样本为同步样本。

[0547] 另外,在图8的实例中,计算装置可开始传递或解码含有至少一IRAP图片的样本的NAL单元,而不传递或解码按解码次序在所述样本之前的文件的NAL单元(558)。以此方式,计算装置可执行随机存取或层切换。举例来说,计算装置可开始在含有至少一 IRAP图片的一或多个样本中的一者处的多层视频数据的解码。

[0548] 图9为说明根据本发明的一或多个技术的文件产生装置34的实例操作的流程图。在图9的实例中,文件产生装置34可产生包括含有用于文件中的播放轨的元数据的播放轨框的文件(600)。用于所述播放轨的媒体数据包括一连串样本。在图9的实例中,样本中的每一者为多层视频数据的视频存取单元。在一些实例中,文件产生装置34编码所述多层视频数据。

[0549] 作为产生文件的部分,文件产生装置34可确定子样本是否含有正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL单元(602)。响应于确定子样本含有正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL单元(602的“是”),文件产生装置34可在文件中产生子样本信息框,所述子样本信息框含有具有指示子样本含有正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL 单元的值(例如,5)的旗标(604)。否则(602的“否”),文件产生装置34可在文件中产生含有具有另一值(例如,0、1、2、3、4)的旗标的子样本信息框(606)。

[0550] 以此方式,文件产生装置34可产生文件,所述文件包括含有用于在文件中的播放轨的元数据的播放轨框。用于所述播放轨的媒体数据包括一连串样本,所述样本中的每一者为多层视频数据的视频存取单元。作为产生文件的部分,文件产生装置34在文件中产生子样本信息框,所述子样本信息框含有指定在所述子样本信息框中给出的子样本信息的类型的旗标。当所述旗标具有特定值时,对应于所述子样本信息框的子样本含有正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL单元。

[0551] 图10为说明根据本发明的一或多个技术的计算装置的实例操作的流程图。在图10的实例中,计算装置接收文件(650)。在图10的实例中,计算装置可为中间网络装置,例如,MANE或流式传输服务器。在一些实例中,计算装置可为内容传递网络的部分。此外,在图10的实例中,计算装置可从文件获得播放轨框(651)。播放轨框含有用于文件中的播放轨的元数据。用于所述播放轨的媒体数据包括一连串样本。在图10的实例中,样本中的每一者为多层视频数据的视频存取单元。

[0552] 此外,在图10的实例中,计算装置可从文件获得子样本信息框(652)。计算装置使用子样本信息框中的信息提取子位流(654)。子位流可包括存储于文件中的位流的操作点的每一NAL单元。换句话说,子位流的NAL单元可为存储于文件中的NAL单元的子集。计算装置可从文件获得子样本信息框,且可在不剖析或解译样本的序列中包含的 NAL单元的情况下提取子位流。当提取子位流时不剖析或解译NAL单元可减小计算装置的复杂度及/或可加快提取子位流的过程。

[0553] 此外,在一些实例中,当旗标具有特定值时,计算装置可从子样本信息框获得以下中的一或者者:

[0554] ●指示子样本的所有VCL NAL单元是否可舍弃的额外旗标,

[0555] ●指示子样本的VCL NAL单元的NAL单元类型的额外值,其中所述子样本的所有VCL NAL单元具有相同NAL单元类型,

[0556] ●指示子样本的每一NAL单元的层识别符的额外值,

[0557] ●指示子样本的每一NAL单元的时间识别符的额外值,

[0558] ●指示是否针对子样本的所有VCL NAL单元启用层间预测的额外旗标,或

[0559] ●指示子样本中的所有NAL单元是否为子层非参考图片的VCL NAL单元的额外旗标。

[0560] 在图10的实例中,作为提取子位流的部分,计算装置可确定子样本信息框的“旗标”值是否具有指示子样本信息框对应于正好一个经译码图片及与所述经译码图片相关联的零或多个非VCL NAL单元的特定值(例如,5)(656)。当子样本信息框的“旗标”值具有特定值(656的“是”)时,计算装置可基于在子样本信息框中指定的信息确定是否需要经译码图片以便解码操作点(658)。举例来说,计算装置可基于可舍弃旗标确定VCL NAL单元类型指示符、层识别符、时间识别符、无层间预测旗标及/或子层参考NAL单元旗标,不管是否需要经译码图片以便解码操作点。当需要经译码图片以解码操作点(658的“是”)时,计算装置可在子位流中包含子样本的NAL单元(660)。否则,在图10的实例中,当不需要经译码图片以解码操作点(658的“否”)时,计算装置不在子位流中包含子样本的NAL单元(662)。

[0561] 此外,在图10的实例中,计算装置可输出子位流(664)。举例来说,计算装置可将子位流存储到计算机可读存储媒体或将子位流发射到另一计算装置。

[0562] 如上所指示,图10为实例。其它实例可包含或省略图10的特定动作。举例来说,一些实例省略动作650、651、654及/或664。此外,一些实例省略动作656到662中的一或多者。

[0563] 图11为说明根据本发明的一或多个技术的文件产生装置34的实例操作的流程图。在图11的实例中,文件产生装置34可产生包括围封媒体内容的媒体数据框的文件(700)。所述媒体内容可包括一连串样本,所述样本中的每一者为多层视频数据的存取单元。在各种实例中,多层视频数据可为SHVC数据、MV-HEVC数据或3D-HEVC数据。在一些实例中,文件产生装置34编码所述多层视频数据。

[0564] 在图11的实例中,作为产生文件的部分,文件产生装置34可确定多层视频数据的位流的至少一存取单元是否包含具有等于第一值(例如,1)的图片输出旗标的经译码图片及具有等于第二值(例如,0)的图片输出旗标的经译码图片(702)。允许输出具有等于第一值(例如,1)的图片输出旗标的图片,且允许将具有等于第二值(例如,0)的图片输出旗标的图片用作参考图片,但不允许将其输出。在其它实例中,其它装置可进行确定多层视频数据的位流的至少一存取单元是否包含具有等于第一值的图片输出旗标的经译码图片及具有等于第二值的图片输出旗标的经译码图片。

[0565] 响应于多层视频数据的位流的至少一存取单元包含具有等于第一值的图片输出旗标的经译码图片及具有等于第二值的图片输出旗标的经译码图片(702的“是”),文件产生装置34使用至少一第一播放轨及第二播放轨将位流存储于文件中(704)。对于来自第一播放轨及第二播放轨的每一相应播放轨,在相应播放轨的每一样本中的所有经译码图片具有相同图片输出旗标值。

[0566] 此外,在图11的实例中,响应于确定位流中无存取单元包含具有等于第一值(例如,1)的图片输出旗标的经译码图片及具有等于第二值(例如,0)的图片输出旗标的经译码图片(702的“否”),文件产生装置34可使用单一播放轨将位流存储于文件中(706)。在其它实例中,文件产生装置34可产生具有多个播放轨的文件,甚至当位流中无存取单元包含具有等于第一值(例如,1)的图片输出旗标的经译码图片及具有等于第二值(例如,0)的图片输出旗标的经译码图片时。

[0567] 如上所指示,图11为实例。其它实例可包含较少动作。举例来说,一些实例省略动作702及706。

[0568] 图12为说明根据本发明的一或多个技术的目的地装置14的实例操作的流程图。在图12的实例中,目的地装置14接收文件(750)。所述文件可包括围封媒体内容的媒体数据框,所述媒体内容包括一连串样本。所述样本中的每一者可为多层视频数据的存取单元。在各种实例中,多层视频数据可为SHVC数据、MV-HEVC数据或3D-HEVC数据。此外,在图12的实例中,目的地装置14可从文件获得第一播放轨框及第二播放轨框(751)。第一播放轨框含有用于文件中的第一播放轨的元数据。第二播放轨框含有用于文件中的第二播放轨的元数据。对于来自第一播放轨及第二播放轨的每一相应播放轨,在相应播放轨的每一样本中的所有经译码图片具有相同图片输出旗标值。允许输出具有等于第一值(例如,1)的图片输出旗标的图片,且允许将具有等于第二值(例如,0)的图片输出旗标的图片用作参考图片,但不允许将其输出。

[0569] 目的地装置14的视频解码器30可针对具有等于第一值(例如,1)的图片输出旗标的图片解码播放轨中的图片,且可针对具有等于第二值(例如,0)的图片输出旗标的图片解

码播放轨中的图片(752)。在一些情况下,视频解码器30可使用具有等于1的图片输出旗标的图片解码具有等于0的图片输出旗标的图片,且反之亦然。目的地装置14可输出具有等于第一值的图片输出旗标的图片(754)。目的地装置14不输出具有等于第二值的图片输出旗标的图片(756)。以此方式,对于来自第一播放轨及第二播放轨的每一相应播放轨,目的地装置14可解码在相应播放轨的每一样本中的经译码图片,且输出具有等于第一值的图片输出旗标的经解码图片。

[0570] 如上所指示,图12仅提供为实例。其它实例可省略图12的特定动作,例如,动作752到756。

[0571] 在一或多个实例中,所描述功能可以硬件、软件、固件或其任何组合来实施。如果以软件实施,那么所述功能可作为一或多个指令或代码而在计算机可读媒体上存储或传输,且由基于硬件的处理单元执行。计算机可读媒体可包含计算机可读存储媒体(其对应于例如数据存储媒体等有形媒体),或包含促进将计算机程序从一处传送到另一处(例如,根据通信协议)的任何媒体的通信媒体。以此方式,计算机可读媒体大体上可对应于(1) 非暂时性的有形计算机可读存储媒体,或(2) 例如信号或载波的通信媒体。数据存储媒体可为可通过一或多个计算机或一或多个处理器存取以检索指令、代码及/或数据结构以用于实施本发明所描述的技术的任何可用媒体。计算机程序产品可包含计算机可读媒体。

[0572] 借助于实例而非限制,此些计算机可读存储媒体可包括RAM、ROM、EEPROM、CD-ROM或其它光盘存储器、磁盘存储器或其它磁性存储器、快闪存储器或可用以存储呈指令或数据结构形式的所要的程序代码且可由计算机存取的任何其它媒体。而且,任何连接被恰当地称为计算机可读媒体。举例来说,如果使用同轴缆线、光纤缆线、双绞线、数字订户线(DSL)或例如红外线、无线电及微波等无线技术从网站、服务器或其它远程源传输指令,那么同轴缆线、光纤缆线、双绞线、DSL或例如红外线、无线电及微波等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、信号或其它暂时性媒体,而取而代之,是有关非暂时性的有形存储媒体。如本文所使用,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)、软性磁盘及蓝光光盘,其中磁盘通常以磁性方式再现数据,而光盘用激光以光学方式再现数据。以上各者的组合也应包含于计算机可读媒体的范围内。

[0573] 可通过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)或其它等效集成或离散逻辑电路的一或多个处理器来执行指令。因此,如本文中所使用的术语“处理器”可指上述结构或适合于实施本文中所描述的技术的任何其它结构中的任一者。此外,在一些方面中,本文中所描述的功能性可在经配置用于编码及解码的专用硬件及/或软件模块内提供,或并入于组合式编解码器中。而且,所述技术可完全实施于一或多个电路或逻辑元件中。

[0574] 本发明的技术可以多种装置或设备实施,所述装置或设备包含无线手持机、集成电路(IC)或IC集合(例如,芯片集)。在本发明中描述各种组件、模块或单元以强调经配置以执行所揭示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。相反地,如上所述,各种单元可与合适的软件及/或固件一起组合在编解码器硬件单元中或由互操作硬件单元的集合提供,硬件单元包含如上文所描述的一或多个处理器。

[0575] 已描述各种实例。此些及其它实例处于以下权利要求书的范围内。

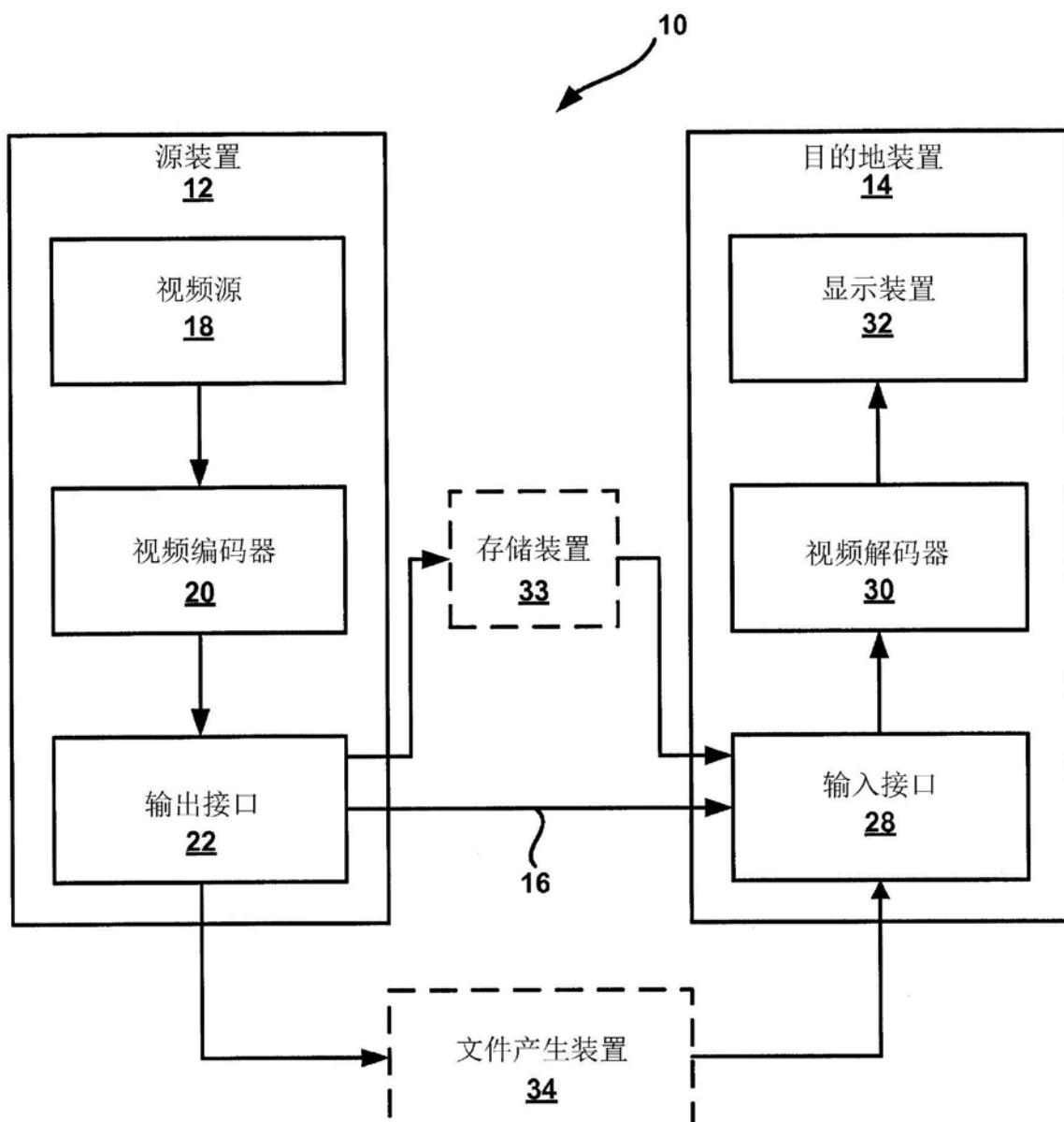


图1

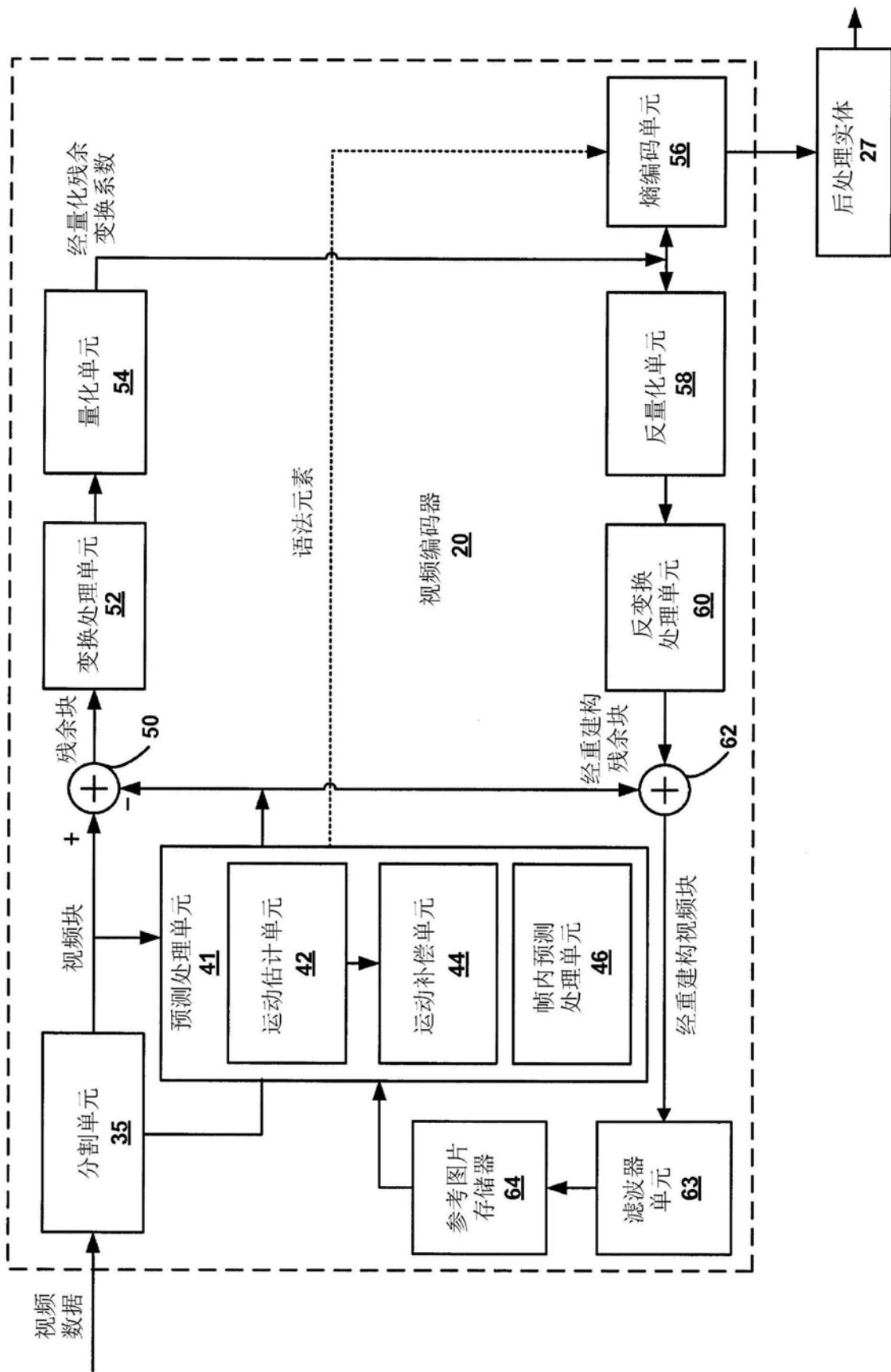


图2

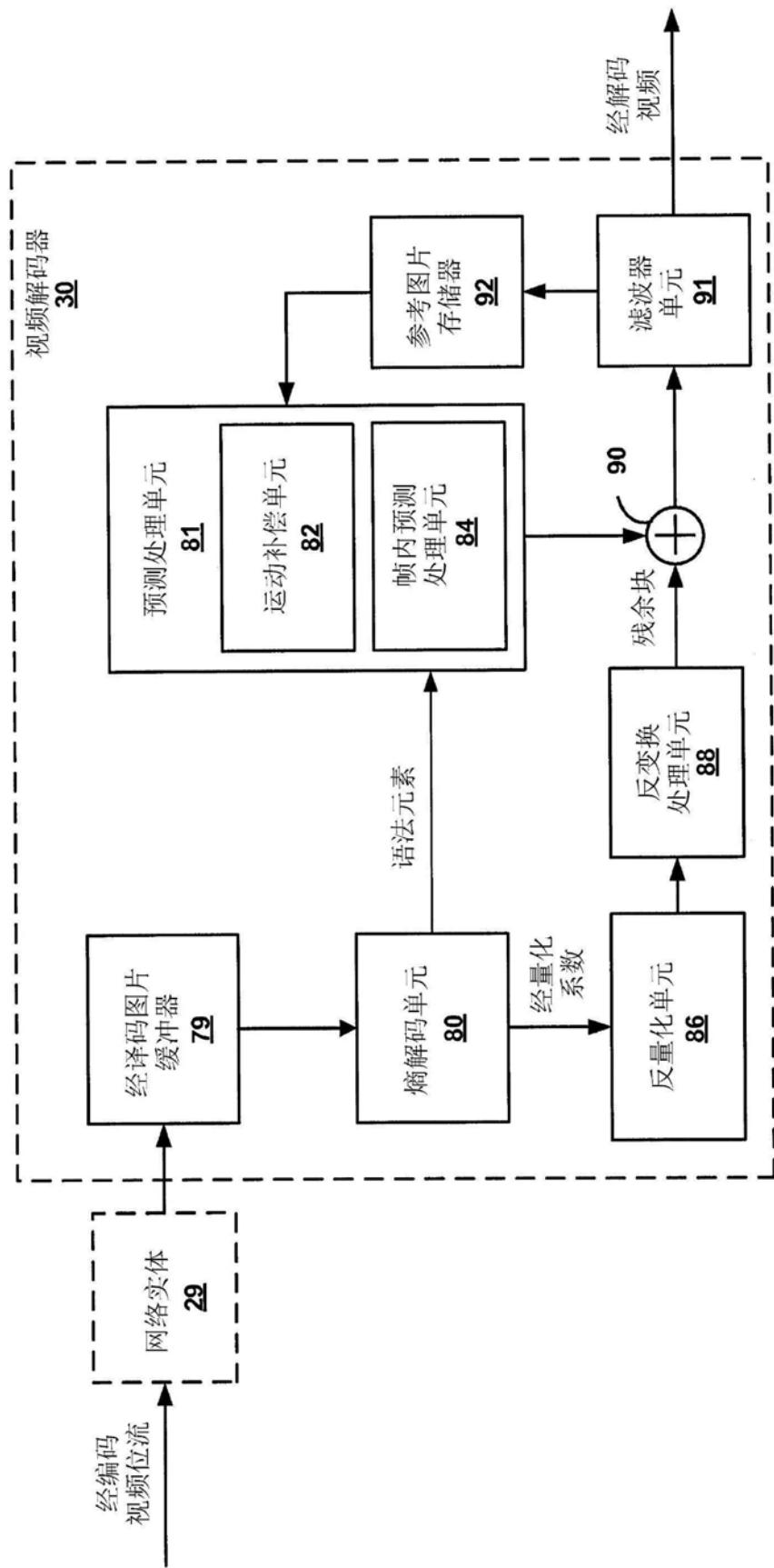


图3

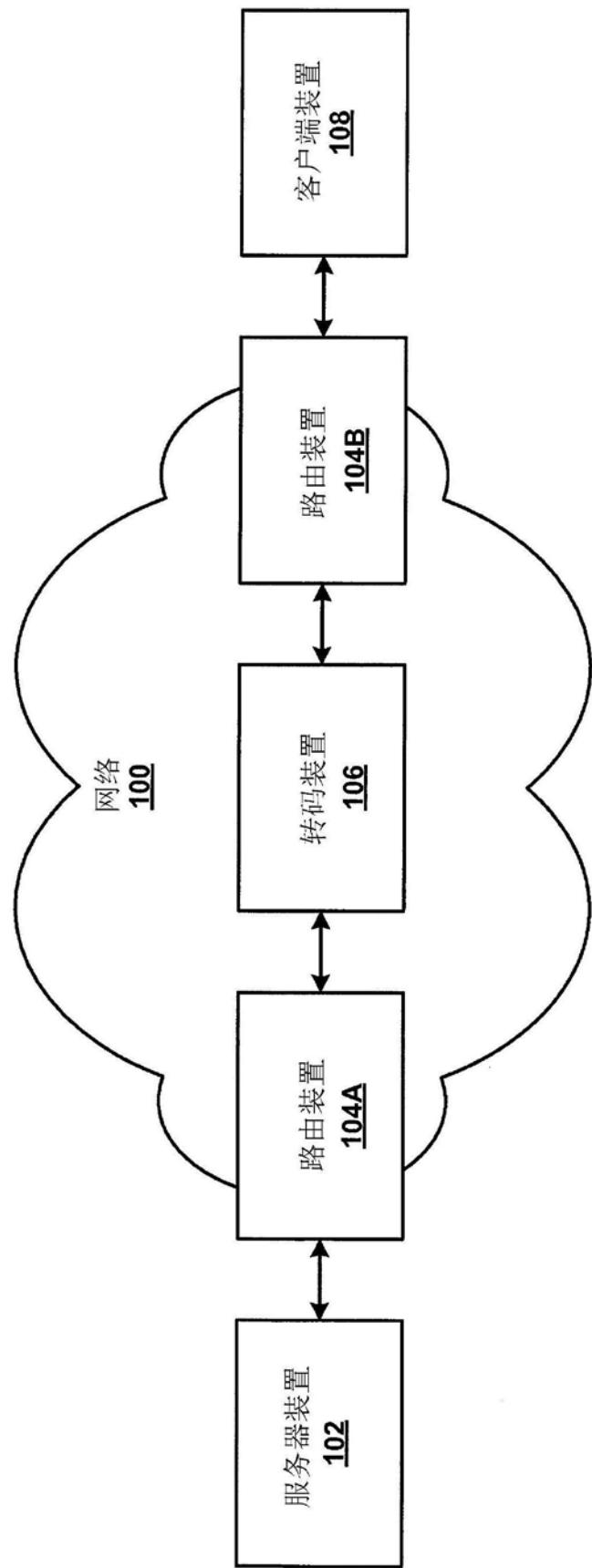


图4

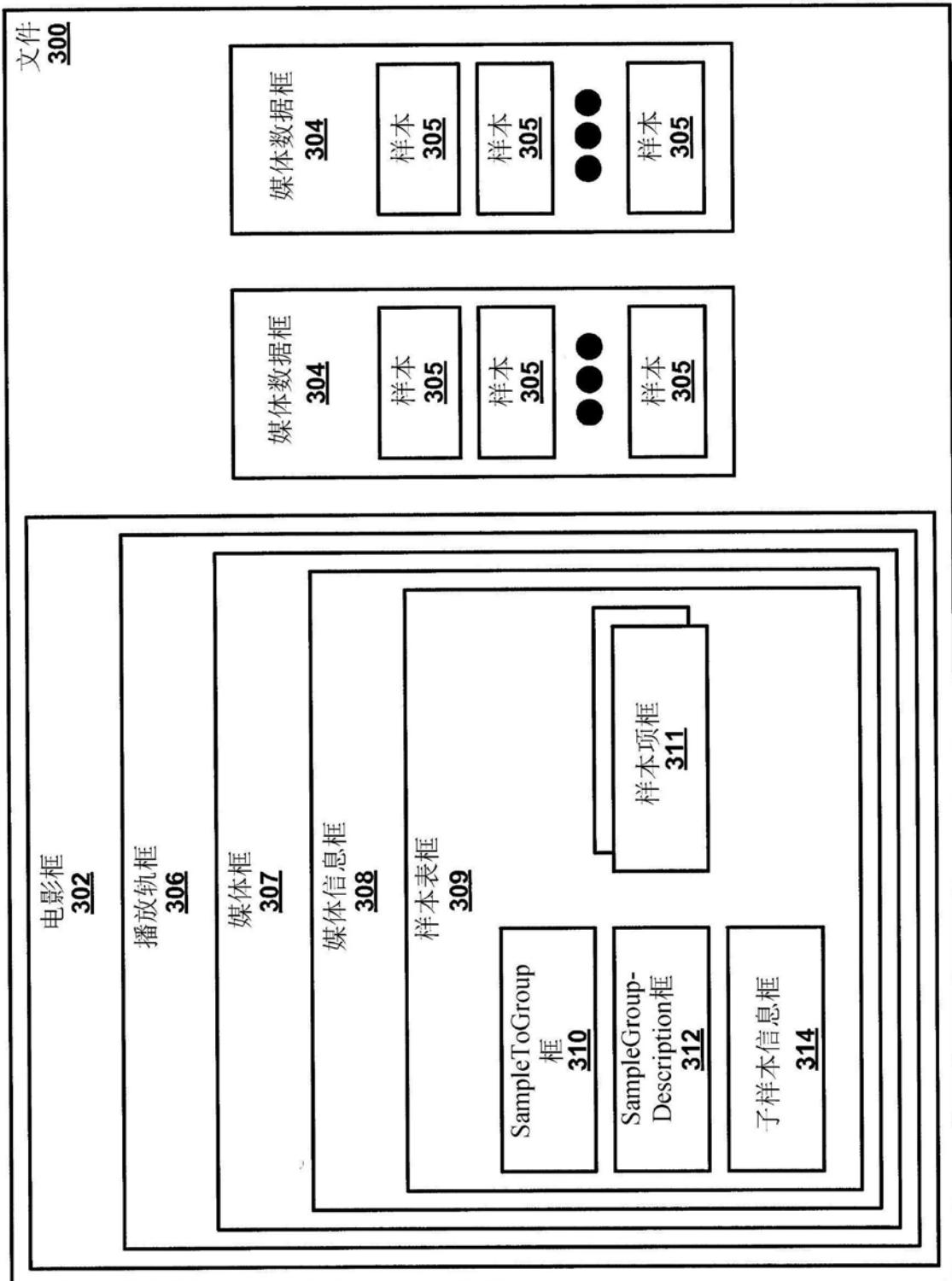


图5

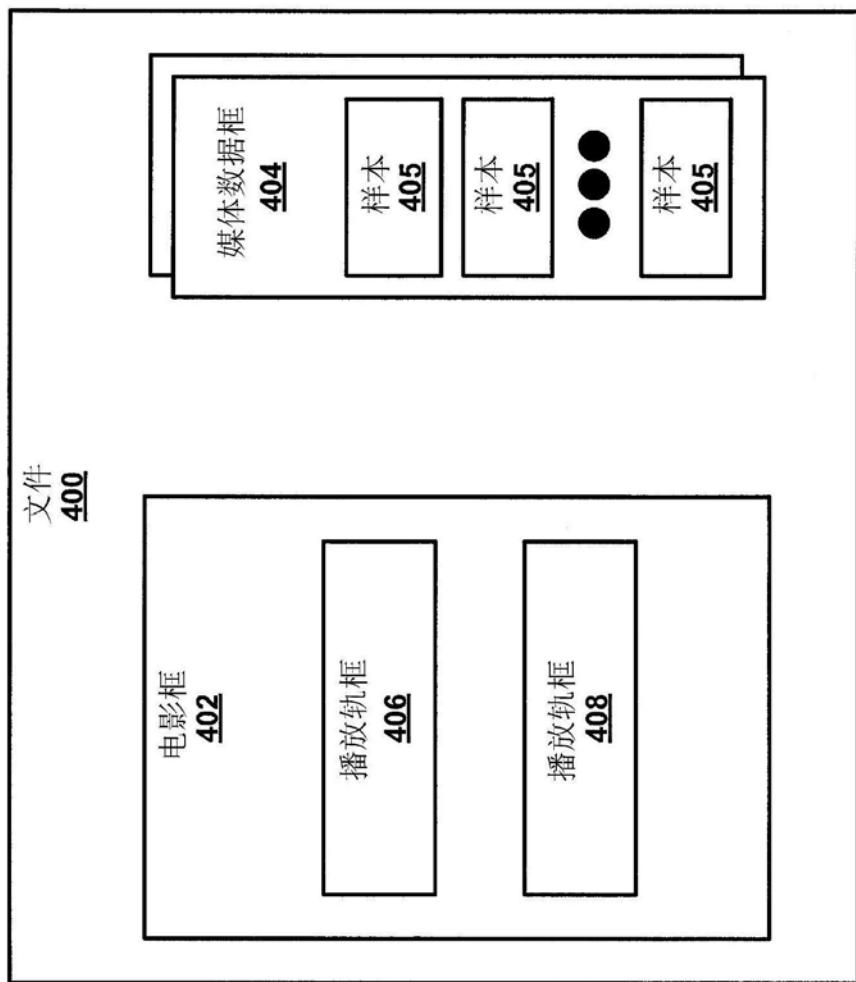


图6

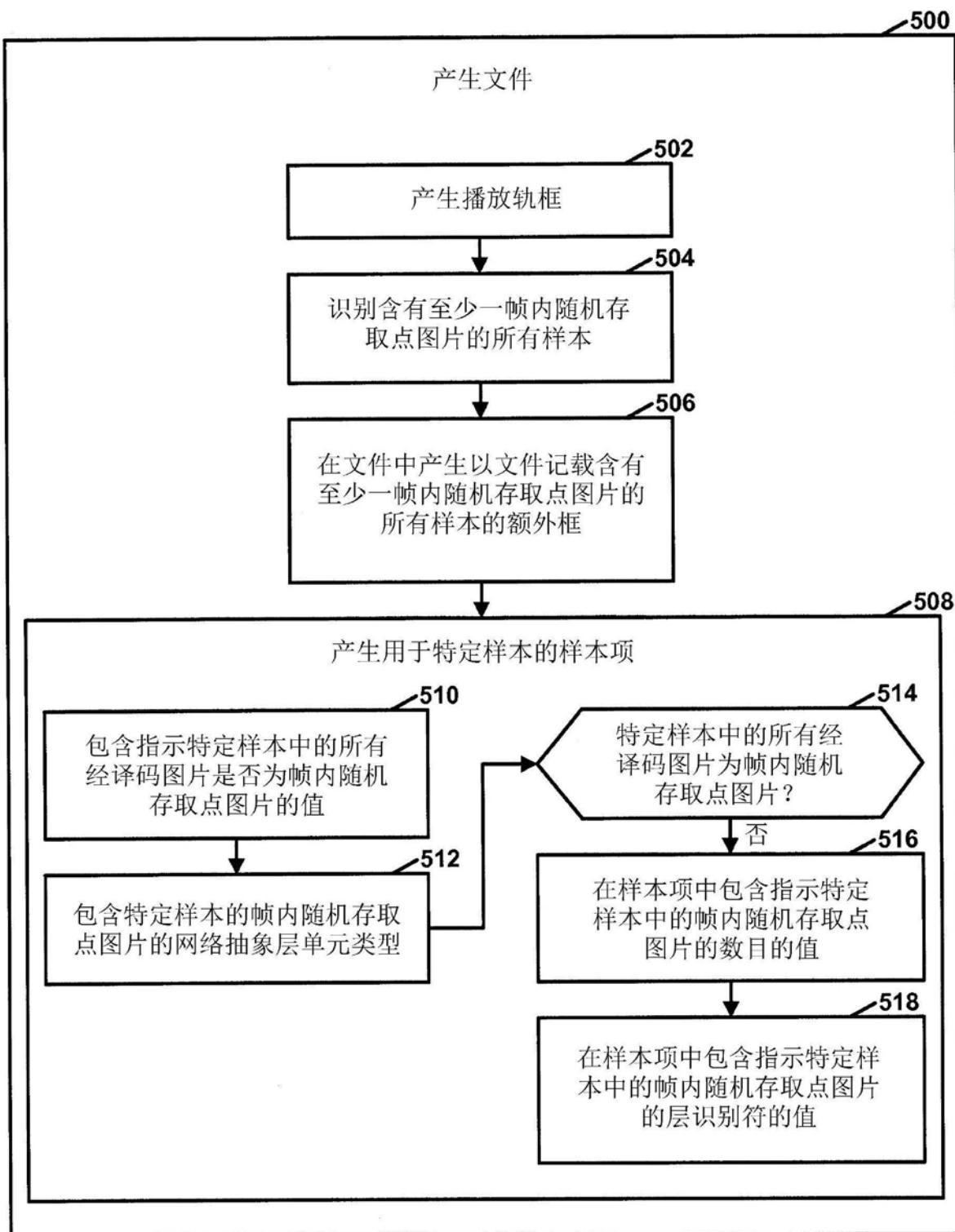


图7

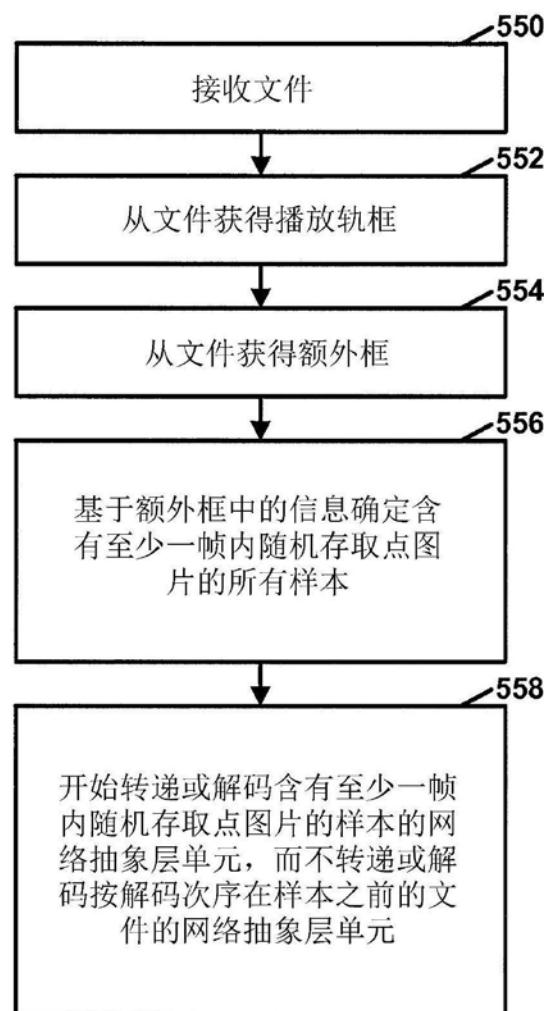


图8

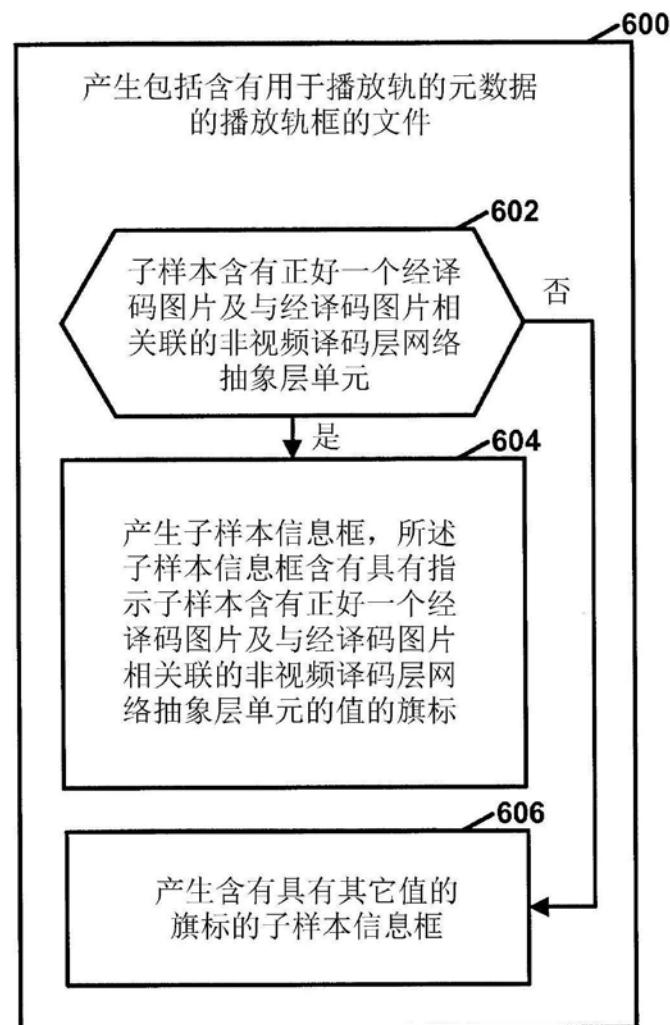


图9

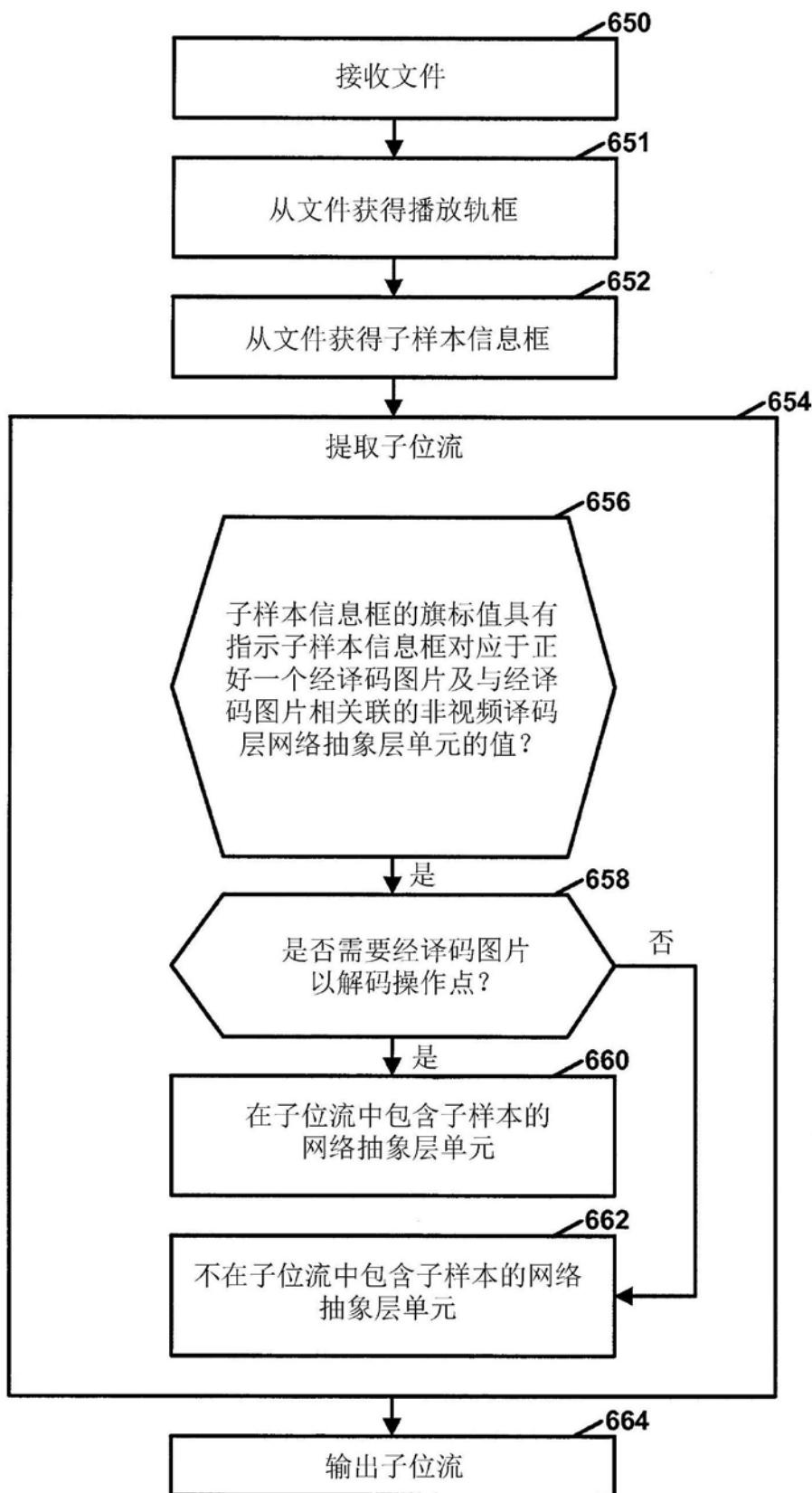


图10

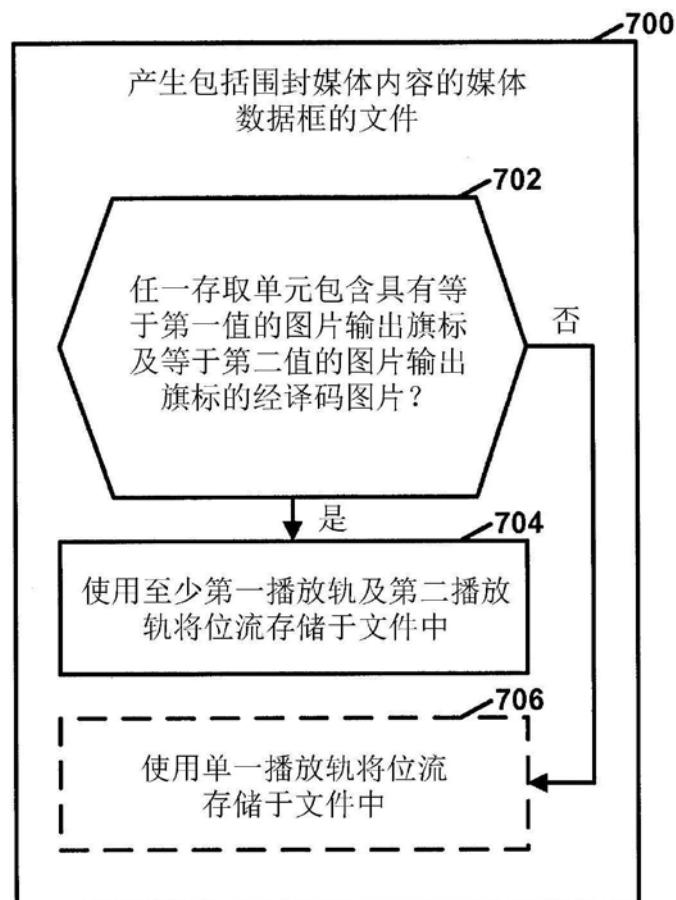


图11

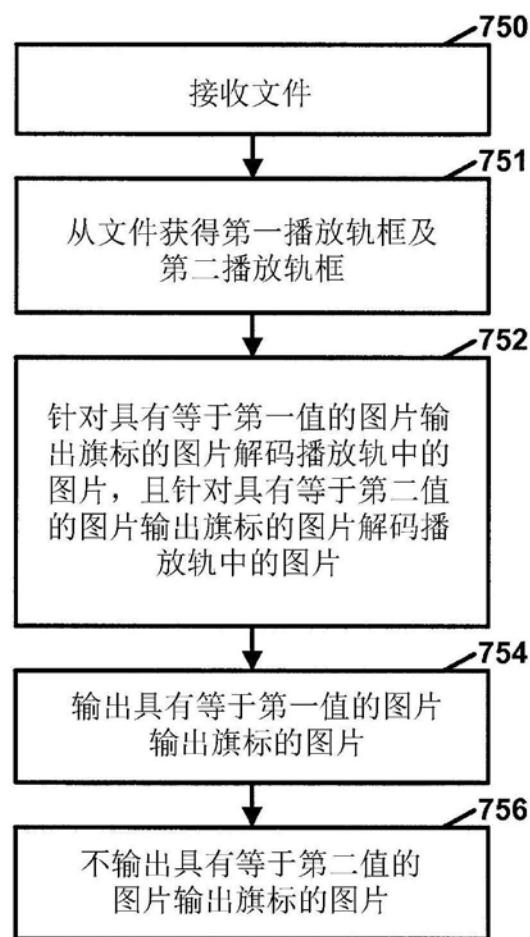


图12