



(12) 发明专利

(10) 授权公告号 CN 109618235 B

(45) 授权公告日 2021. 03. 16

(21) 申请号 201811096243.4

西里尔·康克拉托 简·勒菲弗

(22) 申请日 2014.01.17

(74) 专利代理机构 北京魏启学律师事务所

11398

(65) 同一申请的已公布的文献号

代理人 魏启学

申请公布号 CN 109618235 A

(43) 申请公布日 2019.04.12

(51) Int.Cl.

(30) 优先权数据

H04N 21/4728 (2011.01)

1300953.5 2013.01.18 GB

H04N 21/845 (2011.01)

1306451.4 2013.04.09 GB

H04N 21/854 (2011.01)

H04L 29/06 (2006.01)

(62) 分案原申请数据

审查员 姜玉静

201480016610.9 2014.01.17

(73) 专利权人 佳能株式会社

地址 日本东京都大田区下丸子3丁目30番2号

(72) 发明人 弗雷德里克·梅兹

赫夫·勒弗洛奇 弗兰克·德诺奥

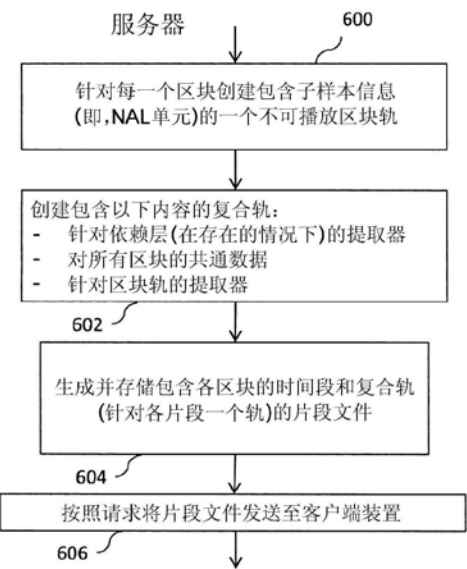
权利要求书3页 说明书32页 附图17页

(54) 发明名称

生成设备和方法、处理设备和方法以及存储介质

(57) 摘要

本发明涉及一种生成设备和方法、处理设备和方法以及存储介质。本发明还涉及在服务器中对区块化定时媒体数据进行封装并且在客户端装置中从封装后的区块化定时媒体提供定时媒体数据位流。该区块化定时媒体数据包括定时样本,并且各定时样本包括多个空间子样本。在从定时样本其中之一的多个空间子样本中选择了至少一个空间子样本之后,针对所选择的各空间子样本,创建包括所选择的空间子样本和其它定时样本各自的一个相应空间子样本的一个区块轨。创建包括识别所创建的区块轨中的至少一个区块轨的至少一个提取器的一个复合轨。将所创建的各轨独立封装在至少一个媒体片段文件中。



1. 一种用于基于视频数据来生成至少一个媒体文件的方法,所述方法包括以下步骤:
获得所述视频数据;
生成SampleToGroup框,所述SampleToGroup框包含与所述视频数据的多个样本的分组有关的信息;
生成SampleGroupDescription框,所述SampleGroupDescription框包含(i)要应用于属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的描述以及(ii)要应用于不属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的默认描述;
以及
基于所述SampleToGroup框、所述SampleGroupDescription框和所述多个样本来生成所述至少一个媒体文件。
2. 根据权利要求1所述的方法,其中,所述SampleToGroup框表示按照与组数相对应的条目计数信息和与属于各组的样本数相对应的样本计数信息来对所述多个样本进行分组。
3. 根据权利要求1所述的方法,其中,所述SampleToGroup框表示按照与组数相对应的条目计数信息、与属于各组的样本数相对应的样本计数信息和在所述SampleGroupDescription框中参考的组描述索引而要应用于属于各组的样本的描述。
4. 根据权利要求1所述的方法,其中,所述默认描述要应用于没有提供样本到组映射的所有样本。
5. 根据权利要求1所述的方法,其中,所述至少一个媒体文件符合ISO基媒体文件格式(ISO/IEC 14496-12)。
6. 根据权利要求1所述的方法,其中,所述描述和所述默认描述各自包含各样本的空间区块的水平偏移和垂直偏移。
7. 根据权利要求1所述的方法,其中,所述描述和所述默认描述各自包含各样本的空间区块的宽度和高度。
8. 一种用于基于视频数据来处理至少一个媒体文件的方法,所述方法包括以下步骤:
获得所述至少一个媒体文件,所述至少一个媒体文件包括:
SampleToGroup框,其包含与所述视频数据的多个样本的分组有关的信息,
SampleGroupDescription框,其包含(i)要应用于属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的描述以及(ii)要应用于不属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的默认描述,以及
所述多个样本;以及
通过使用所述多个样本至少之一以及所述描述和所述默认描述至少之一,来再现基于所述视频数据的视频。
9. 根据权利要求8所述的方法,其中,所述描述和所述默认描述各自包含各样本的空间区块的水平偏移和垂直偏移。
10. 根据权利要求8所述的方法,其中,所述描述和所述默认描述各自包含各样本的空间区块的宽度和高度。
11. 根据权利要求8所述的方法,其中,所述默认描述要应用于没有提供样本到组映射的所有样本。
12. 一种生成设备,用于基于视频数据来生成至少一个媒体文件,所述生成设备包括:

获得部件,用于获得所述视频数据;以及

生成部件,用于:

生成SampleToGroup框,其包含与所述视频数据的多个样本的分组有关的信息,

生成SampleGroupDescription框,其包含 (i) 要应用于属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的描述以及 (ii) 要应用于不属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的默认描述,以及

基于所述SampleToGroup框、所述SampleGroupDescription框和所述多个样本生成所述至少一个媒体文件。

13. 根据权利要求12所述的生成设备,其中,所述SampleToGroup框表示按照与组数相对应的条目计数信息和与属于各组的样本数相对应的样本计数信息来对所述多个样本进行分组。

14. 根据权利要求12所述的生成设备,还包括输出部件,所述输出部件用于输出所生成的所述至少一个媒体文件。

15. 根据权利要求12所述的生成设备,其中,所述描述和所述默认描述各自包含各样本的空间区块的水平偏移和垂直偏移。

16. 根据权利要求12所述的生成设备,其中,所述描述和所述默认描述各自包含各样本的空间区块的宽度和高度。

17. 一种处理设备,用于基于视频数据来处理至少一个媒体文件,所述处理设备包括:

获得部件,用于获得所述至少一个媒体文件,所述至少一个媒体文件包括:

SampleToGroup框,其包含与所述视频数据的多个样本的分组有关的信息,

SampleGroupDescription框,其包含 (i) 要应用于属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的描述以及 (ii) 要应用于不属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的默认描述,以及

所述多个样本;以及

再现部件,用于通过使用所述多个样本至少之一以及所述描述和所述默认描述至少之一,来再现基于所述视频数据的视频。

18. 根据权利要求17所述的处理设备,其中,所述描述和所述默认描述各自包含各样本的空间区块的水平偏移和垂直偏移。

19. 根据权利要求17所述的处理设备,其中,所述描述和所述默认描述各自包含各样本的空间区块的宽度和高度。

20. 根据权利要求17所述的处理设备,其中,所述默认描述要应用于没有提供样本到组映射的所有样本。

21. 一种计算机可读存储介质,其存储有程序,所述程序用于使计算机执行包括以下步骤的方法:

获得视频数据;

生成SampleToGroup框,所述SampleToGroup框包含与所述视频数据的多个样本的分组有关的信息;

生成SampleGroupDescription框,所述SampleGroupDescription框包含 (i) 要应用于属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的描述以及 (ii) 要应

用于不属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的默认描述；
以及

基于所述SampleToGroup框、所述SampleGroupDescription框和所述多个样本来生成至少一个媒体文件。

22. 一种计算机可读存储介质，其存储有程序，所述程序用于使计算机执行包括以下步骤的方法：

获得至少一个媒体文件，所述至少一个媒体文件包括：

SampleToGroup框，其包含与视频数据的多个样本的分组有关的信息，

SampleGroupDescription框，其包含 (i) 要应用于属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的描述以及 (ii) 要应用于不属于所述SampleToGroup框所定义的一个或多个组中的任一组的样本的默认描述，以及

所述多个样本；以及

通过使用所述多个样本至少之一以及所述描述和所述默认描述至少之一，来再现基于所述视频数据的视频。

生成设备和方法、处理设备和方法以及存储介质

[0001] (本申请是申请日为2014年1月17日、申请号为201480016610.9、发明名称为“用于封装分区定时媒体数据的方法、装置和计算机可读存储介质”的申请的分案申请。)

技术领域

[0002] 本发明通常涉及如下领域：例如根据如MPEG标准化组织所定义的基媒体文件格式(Base Media File Format)来对定时媒体数据进行封装，从而改善特别是与压缩视频流中的用户所选择的关注区域的HTTP(HyperText Transfer Protocol,超文本传输协议)流传输有关的流传递。更特别地，本发明涉及用于对包含诸如空间区块(tile)等的分区数据的基本流进行封装从而使得能够进行数据特别是一个或多个区块的高效流传输的方法、装置和计算机程序。

背景技术

[0003] 视频编码是将一系列视频图像变换成紧凑型数字化位流以使得可以传输或存储视频图像的方式。编码装置用于对视频图像进行编码，而关联的解码装置可用于重建位流以供显示和观看。一般目的是形成大小小于原始视频信息的大小的位流。这样有利地减少了针对用以传输或存储位流码的传输网络或存储装置所要求的容量。为了进行传输，通常根据典型地添加头部和校验位的传输协议来对视频位流进行封装。

[0004] 近来，运动图片专家组(MPEG)发布了用以统一并取代现有的经由HTTP(超文本传输协议)的流传输解决方案的新标准。被称为“经由HTTP的动态自适应流传输(DASH)”的该新标准意图支持基于标准web服务器的经由HTTP的媒体流传输模型，其中在这些标准web服务器中，智能(即，用以进行流传输的媒体数据的选择、以及位流针对用户选择、网络条件和客户端性能的动态适应)排他地依赖于客户端选择和装置。

[0005] 在该模型中，以数据片段并且以表示要呈现的定时媒体数据的组织的被称为“媒体呈现描述(MPD)”的清单来组织媒体呈现。特别地，清单包括供下载数据片段所使用的资源标识符，并且提供用以选择并组合这些数据片段以获得有效媒体呈现的上下文。资源标识符通常是有可能与字节范围相组合的HTTP-URL(统一资源定位符)。基于清单，客户端装置根据其需求、其性能(例如，所支持的编解码器、显示大小、帧频、质量等级等)并且依赖于网络条件(例如，可用带宽)，来随时确定要从媒体数据服务器下载哪些媒体片段。

[0006] 另外，视频分辨率不断提高、即从标准清晰度(SD)提高至高清晰度(HD)再提高至超高清晰度(例如，4K2K或8K4K，也就是说视频包括 $4,096 \times 2,400$ 个像素或 $7,680 \times 4,320$ 个像素的图像)。然而，特别是在视频具有超高清晰度的情况下，并非所有的接收和视频解码装置都具有用以以全分辨率访问视频的资源(例如，网络访问带宽或CPU(中央处理单元))，并且并非所有的用户都需访问这种视频。在这种上下文中，提供用以仅访问一些关注区域(ROI)、也就是说仅访问整个视频序列的一些空间子部分的能力，这是特别有利的。

[0007] 用以访问属于视频的帧的空间子部分的已知机制包括：将视频的各帧组织为通常被称为区块的独立可解码的空间区域的排列。诸如SVC(Scalable Video Coding,可分级视

频编码)或HEVC(High Efficiency Video Coding,高效率视频编码)等的一些视频格式提供了针对区块定义的支持。用户定义的ROI可以覆盖一个或多个连续区块。

[0008] 因此,为了根据HTTP协议对用户选择的ROI进行流传输,以能够对一个或多个区块进行空间访问并且能够组合所访问区块的方式来提供编码视频位流的定时媒体数据的封装,这很重要。

[0009] 需要重申,编码视频位流通常被构造为与完整帧相对应的一组连续时间样本,其中这些时间样本是根据解码顺序进行组织的。使用文件格式来封装并描述这些编码位流。

[0010] 为了便于例示,国际标准化组织基媒体文件格式(ISO BMFF)是描述编码定时媒体数据位流以供本地存储或者供经由网络或经由另一位流传递机构进行传输的众所周知的灵活且可扩展的格式。该文件格式是面向对象的。该文件格式包括被称为框(box)的构建块,其中这些框是顺次或层级组织的,并且定义诸如定时参数和结构参数等的编码定时媒体数据位流的参数。根据该文件格式,定时媒体数据位流包含在被称为mdat框的数据结构中,其中该mdat框是在被称为轨框的另一数据结构中所定义的。轨表示样本的定时间序列,其中样本对应于与一个时间戳相关联的所有数据、也就是说与一个帧相关联的所有数据或与共用同一时间戳的多个帧相关联的所有数据。

[0011] 对于诸如SVC格式的视频等的可分级视频,可以通过使用多个依赖轨来高效地表示分层媒体数据组织,其中各轨表示可分级性的特定级别的视频。为了避免轨之间的数据重复,可以使用提取器。根据标准文件格式,提取器(extractor)是直接包括在位流中的、使得能够从其它位流高效地提取网络抽象层(NAL)单元的数据结构。例如,增强层轨的位流可以包括参考来自基本层轨的NAL单元的提取器。然后,在从文件格式中提取这种增强层轨的情况下,必须利用提取器正参考的数据来替换这些提取器。

[0012] 在使用嵌入有这些机制的ISO BMFF来描述子信息并且容易地访问该子信息或高效地将位流组织成多个片段的情况下,可以采用若干策略。

[0013] 例如,在作者为Kofler等人的标题为“Implications of the ISO Base Media File Format on Adaptive HTTP Streaming of H.264/SVC”的文章中,提出用于考虑到ISO BMFF的可能性以及限制来对可分级视频位流(H264/SVC)进行组织以供HTTP流传输的三个不同策略:

[0014] a) 包含特定文件头部的一个文件,该特定文件头部包括文件类型框“ftyp”和包含所有的ISO BMFF元数据(包括轨定义)的动画框“moov”,其中该一个文件还包括包含整个编码位流的一个mdat框。该组织适合本地存储,但不适用于客户端可能仅需要整个位流的一部分的HTTP流传输;

[0015] b) 包含多个适于分割的moof/mdat框的单一文件。该格式使得能够进行渐进式下载。Moof框等同于分割级别的moov框。根据该方案,使用分割媒体文件,将可分级位流分割成以不同可分级性级别来表示视频的多个依赖轨。使用提取器来参考来自其它轨的NAL单元。在使用针对各区块的轨的情况下,必须预先准备所有可寻址轨并且无法独立选择轨。如果要显示多个区块,则必须解码多个位流并且分数次对基本层进行解码;

[0016] c) 多个片段文件,其中各文件是利用其自身的URL可访问并且是独立可下载的。各片段通常包括用作一种文件头部的片段类型框(styp)、可选片段索引框(sidx)以及一个或多个片段。再次地,各片段包括moof和mdat框。根据该方案,使用片断媒体文件,将各轨存储

在其自身的片段中,其中关联位流与可分级性的一个级别相关联。在需要的情况下,使用提取器来参考来自依赖轨的所需位流。这种编码方案特别适合独立对轨进行流传输。这很好地适用于DASH标准,但由于要对多个位流进行解码并由此针对各轨需要一个解码器,因而不适合区块流传输。此外,在选择一个以上的区块的情况下,存在基本层的位流的潜在重复。

[0017] 在应用于空间区块的情况下,这些策略均不允许在HTTP流传输的上下文中对特定区块进行高效访问。实际上,利用现有的文件格式定义,仍将需要访问编码位流中的多个非连续字节范围,或者,为了显示与给定时间间隔相对应的多个帧的空间区块,这将导致位流重复。

[0018] 为了解决这些问题,提供了适合空间区块的高效数据组织和轨描述方案,从而无论利用客户端应用程序选择何种轨组合,都确保了ISO BMFF解析的结果始终得到针对视频解码器的有效视频基本位流。

发明内容

[0019] 面对这些制约,本发明人提供一种用于在服务器中对区块化定时媒体数据进行封装并且用于从封装在多个媒体片段文件中的区块化定时媒体数据提供定时媒体数据的方法和装置。

[0020] 本发明的广义目的是改善如上所述的现有技术的不足之处。

[0021] 根据本发明的第一方面,提供一种用于在服务器中对分区定时媒体数据进行封装的方法,所述分区定时媒体数据包括定时样本,各定时样本包括多个子样本,所述方法包括以下步骤:

[0022] 从所述定时样本其中之一的所述多个子样本中选择至少一个子样本;

[0023] 针对所选择的各子样本,创建包括所选择的子样本和其它定时样本各自的一个相应子样本的一个分区轨;

[0024] 创建包括用于识别所创建的分区轨中的至少一个分区轨的至少一个提取器的至少一个参考轨;以及

[0025] 将所创建的各个轨独立封装在至少一个媒体片段文件中。

[0026] 因此,本发明使得无论客户端装置所下载的部分或区块的所选择集合如何,都能够组合不同的部分、特别是不同的区块,并且能够在解析时创建有效文件格式和基本流。

[0027] 因此,本发明的方法适合独立部分或区块的高效流传输,仅需要要发送至客户端装置的有用数据,适合一个以上的部分或区块选择(即,任意的ROI),降低索引化开销(相比字节范围请求),并且可以归并到MPEG标准中。

[0028] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是至少一个复合轨。所述至少一个参考轨可以包含提取器并且可选地包含初始化数据。

[0029] 在实施例中,所述至少一个提取器还识别所识别出的至少一个区块轨的至少一个空间子样本。

[0030] 在实施例中,所述方法还包括以下步骤:创建包括使得能够根据所创建的区块轨和所创建的至少一个复合轨来描述所述区块化定时媒体数据的空间部分的参数的初始化

片段文件。所述初始化片段文件可以包括向所创建的区块轨的参考、与所创建的至少一个复合轨相关联的所参考的区块轨的列表、与所创建的各个区块轨相关联的表示该轨包含空间子样本信息的区块处理程序、以及/或者与所创建的各个区块轨相关联的一般呈现信息。

[0031] 在实施例中,在针对所选择的各空间子样本创建一个区块轨的步骤中,创建至少两个区块轨,其中所述至少一个复合轨包括对同一定时样本内所包括的两个空间子样本中的任意空间子样本进行解码所要使用的至少一个数据项。

[0032] 在实施例中,所述区块化定时媒体数据是各个所述定时样本包括参考层和至少一个增强层所依据的可分级区块化定时媒体数据,其中所述参考层包括至少一个参考空间子样本并且所述至少一个增强层包括多个增强空间子样本;以及在针对所选择的各空间子样本创建一个区块轨的步骤中所创建的区块轨是包括增强空间子样本的增强区块轨,所述方法还包括创建包括参考空间子样本的至少一个参考轨的步骤。

[0033] 在实施例中,所述参考层包括多个参考空间子样本,以及创建多个参考轨,其中所述多个参考轨中的各参考轨形成参考区块轨。

[0034] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件包括用于存储包括相应轨的定义的元数据的数据结构。

[0035] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件中的至少一个媒体片段文件包括用于存储元数据的包含相应轨的至少一组子样本的至少一个定义的数据结构。

[0036] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件中的至少一个媒体片段文件包括用于存储元数据的包含相应轨的不同组子样本的不同定义的数据结构。

[0037] 在实施例中,存储元数据的所述数据结构包括至少一个定义和表示所述至少一个定义应用于相应轨的所有子样本的版本类型。

[0038] 在实施例中,与子样本解码依赖性有关的、用于存储元数据的包括相应轨的至少一组子样本的至少一个定义的数据结构的至少一个参数表示能够在无需使用另一组子样本的情况下对该组的各子样本进行解码。

[0039] 在实施例中,组的子样本是根据子样本所属的分组类型所识别出的。

[0040] 在实施例中,组的子样本是根据子样本所属的分组类型所识别出的,其中所述分组类型提供针对该组的各子样本的定义。

[0041] 在实施例中,所述初始化片段文件包括对不与通过封装至少一个区块轨所得到的媒体片段文件中所存储的任何特定定义相关联的该至少一个区块轨的子样本进行处理而默认使用的一个定义。

[0042] 在实施例中,所述方法还包括以下步骤:将初始化数据添加在通过封装所创建的各个轨所得到的媒体片段文件中的至少一个媒体片段文件中,以使得能够修改默认要使用的初始化片段文件中所存储的定义。

[0043] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件包括用于存储空间子样本的数据结构。

[0044] 通过对所创建的各个轨进行封装所得到的媒体片段文件包括用于存储提取器的数据结构。

[0045] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件中的至少一个媒体片段文件包括所述参考轨中所存储的初始化数据,以使得能够将媒体片段文件中的所述至少一个媒体片段文件作为标准媒体片段文件独立进行处理。

[0046] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件中的至少一个媒体片段文件包括用于存储识别所述参考轨中所存储的初始化数据的至少一个提取器的数据结构。

[0047] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件中的至少一个媒体片段文件包括用于存储识别所述复合轨的初始化数据的至少一个提取器的数据结构,以使得能够将媒体片段文件中的所述至少一个媒体片段文件作为标准媒体片段文件独立进行处理。

[0048] 在实施例中,所述方法还包括以下步骤:设置表示能够将媒体片段文件中的所述至少一个媒体片段文件作为标准媒体片段文件独立进行处理的标志。

[0049] 在实施例中,所述方法还包括以下步骤:创建包括初始化数据的至少一个初始化数据轨,其中所述至少一个初始化数据轨被独立封装在至少一个媒体片段文件中。

[0050] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件中的至少一个媒体片段文件包括用于存储识别所述初始化数据轨中所存储的初始化数据的至少一个提取器的数据结构,以使得能够将媒体片段文件中的所述至少一个媒体片段文件作为标准媒体片段文件独立进行处理。

[0051] 在实施例中,所述方法还包括以下步骤:设置表示能够将媒体片段文件中的所述至少一个媒体片段文件作为标准媒体片段文件独立进行处理的标志。

[0052] 在实施例中,所述服务器符合超文本传输协议即HTTP。

[0053] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件符合如国际标准化组织所定义的基媒体文件格式和经由超文本传输协议即HTTP的动态自适应流传输格式。

[0054] 在实施例中,一种用于在服务器中对分区定时媒体数据进行封装的方法,所述分区定时媒体数据包括定时样本,各定时样本包括多个空间子样本,所述方法包括以下步骤:

[0055] 从所述定时样本其中之一的所述多个子样本中选择至少一个子样本;

[0056] 针对所选择的各空间子样本,创建包括所选择的子样本和其它定时样本各自的一个相应子样本的一个分区轨;

[0057] 创建至少一个参考轨,其中所述至少一个参考轨包括使第一分区轨中所包括的第一空间子样本和第二分区轨中所包括的第二空间子样本相关联的信息;以及

[0058] 将所创建的各轨独立封装在至少一个媒体片段文件中。

[0059] 可以将这种实施例与以上通过参考本发明的第一方面所公开的前述实施例中的一个或多个相组合。

[0060] 本发明的第二方面提供一种用于在客户端装置中从封装在多个媒体片段文件中的分区定时媒体数据提供定时媒体数据位流的方法,所述分区定时媒体数据包括定时样本,各定时样本包括多个子样本,所述多个媒体片段文件包括至少一个参考轨和至少一个分区轨,所述至少一个分区轨各自包括从所述定时样本其中之一的所述多个子样本中所选择的一个子样本并且包括其它定时样本各自的一个相应空间子样本,所述至少一个参考轨

包括识别至少一个分区轨并且识别所识别出的分区轨的至少一个子样本的至少一个提取器,所述方法包括以下步骤:

[0061] 选择代表子样本的信息项;

[0062] 请求包括所选择的子样本的所述至少一个分区轨,其中所选择的子样本与所选择的信息项相对应;

[0063] 接收所述多个媒体片段文件;以及

[0064] 利用所识别出的相应子样本来替换所述至少一个参考轨的提取器,以生成所述定时媒体数据位流。

[0065] 因此,本发明使得无论客户端装置所下载的部分或区块的所选择集合如何,都能够组合不同的部分、特别是不同的区块,并且能够在解析时创建有效文件格式和基本流。

[0066] 因此,本发明的方法适合独立部分或区块的高效流传输,仅需要要发送至客户端装置的有用数据,适合一个以上的部分或区块选择(即,任意的ROI),降低索引化开销(相比字节范围请求),并且可以归并到MPEG标准中。

[0067] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是至少一个复合轨。

[0068] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是包含提取器并且可选地包含初始化数据的至少一个复合轨。

[0069] 在实施例中,所述方法还包括以下步骤:

[0070] 从所述至少一个复合轨的提取器获得轨参考;

[0071] 检查是否接收到与所获得的轨参考相对应的轨;以及

[0072] 如果没有接收到与所获得的轨参考相对应的轨,则去除相应的提取器以生成所述定时媒体数据位流。

[0073] 在实施例中,所述方法还包括以下步骤:验证为与所获得的轨参考相对应的轨是区块类型的轨。

[0074] 在实施例中,所述方法还包括以下步骤:

[0075] 从所述至少一个复合轨的提取器获得轨参考;

[0076] 检查是否接收到与所获得的轨参考相对应的轨;以及

[0077] 如果没有接收到与所获得的轨参考相对应的轨,则通过填充来替换相应的提取器以生成所述定时媒体数据位流。

[0078] 在实施例中,所述方法还包括以下步骤:验证为与所获得的轨参考相对应的轨是区块类型的轨。

[0079] 在实施例中,所述方法还包括以下步骤:接收包括使得能够根据所创建的区块轨和所创建的至少一个复合轨来形成所述区块化定时媒体数据的空间部分的参数的初始化片段文件。

[0080] 在实施例中,接收到代表至少两个区块轨的至少两个媒体片段文件,其中所述至少一个复合轨包括对属于相同集合的空间子样本的两个空间子样本中的任意空间子样本进行解码所要使用的至少一个数据项。

[0081] 在实施例中,所述区块化定时媒体数据是各个定时样本包括参考层和至少一个增强层所依据的可分级区块化定时媒体数据,其中所述参考层包括至少一个参考空间子样本并且所述至少一个增强层包括多个增强空间子样本;以及接收到代表至少两个轨的至少两个媒体片段文件,其中所接收到的至少两个轨中的一个轨是包括参考空间子样本的参考轨,并且所接收到的至少两个轨中的另一个轨是包括增强空间子样本的增强区块轨。

[0082] 在实施例中,所述参考层包括多个参考空间子样本,以及接收到各自形成参考区块轨的多个参考轨。

[0083] 在实施例中,所述媒体片段文件中的至少一个媒体片段文件包括用于存储元数据的包括相应轨的至少一组子样本的至少一个定义的数据结构,所述方法还包括以下步骤:获得用于对相应轨的所述至少一组子样本进行处理的所述至少一个定义。

[0084] 在实施例中,所述媒体片段文件中的至少一个媒体片段文件包括所述参考轨中所存储的初始化数据,所述方法还包括以下步骤:将所述媒体片段文件中的所述至少一个媒体片段文件作为标准媒体片段文件独立进行处理。

[0085] 在实施例中,所述媒体片段文件中的至少一个媒体片段文件包括用于存储识别所述参考轨中所存储的初始化数据的至少一个提取器的数据结构,所述方法还包括以下步骤:在处理所述媒体片段文件中的至少一个媒体片段文件的情况下,访问所述参考轨中所存储的数据。

[0086] 在实施例中,所述媒体片段文件中的至少一个媒体片段文件包括用于存储识别所述复合轨的初始化数据的至少一个提取器的数据结构,所述方法还包括以下步骤:访问所述参考轨中所存储的数据,以将所述媒体片段文件中的至少一个媒体片段文件作为标准媒体片段文件独立进行处理。

[0087] 在实施例中,所述方法还包括以下步骤:获得包括初始化数据的至少一个初始化数据轨,其中所述至少一个初始化数据轨独立封装在至少一个媒体片段文件中。

[0088] 在实施例中,所述媒体片段文件中的至少一个媒体片段文件包括用于存储识别所述初始化数据轨中所存储的初始化数据的至少一个提取器的数据结构,所述方法还包括以下步骤:访问所述参考轨中所存储的数据,以独立处理所述媒体片段文件中的至少一个媒体片段文件。

[0089] 在实施例中,所述客户端装置符合超文本传输协议即HTTP。

[0090] 在实施例中,所接收到的媒体片段文件符合如国际标准化组织所定义的基媒体文件格式和经由超文本传输协议即HTTP的动态自适应流传输格式。

[0091] 根据本发明的第三方面,提供一种装置,包括被配置为执行上述方法的各步骤的部件。

[0092] 根据本发明的第四方面,提供一种用于在服务器中对分区定时媒体数据进行封装的装置,所述分区定时媒体数据包括定时样本,各定时样本包括多个子样本,所述装置包括至少一个微处理器,所述至少一个微处理器被配置为执行以下步骤:

[0093] 从所述定时样本其中之一的所述多个子样本中选择至少一个子样本;

[0094] 针对所选择的各子样本,创建包括所选择的子样本和其它定时样本各自的一个相应子样本的一个分区轨;

[0095] 创建包括用于识别所创建的分区轨中的至少一个分区轨的至少一个提取器的至

少一个参考轨;以及

[0096] 将所创建的各个轨独立封装在至少一个媒体片段文件中。

[0097] 因此,本发明使得无论客户端装置所下载的部分或区块的所选择集合如何,都能够组合不同的部分、特别是不同的区块,并且能够在解析时创建有效文件格式和基本流。

[0098] 因此,本发明的方法适合独立部分或区块的高效流传输,仅需要要发送至客户端装置的有用数据,适合一个以上的部分或区块选择(即,任意的ROI),降低索引化开销(相比字节范围请求),并且可以归并到MPEG标准中。

[0099] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是至少一个复合轨。

[0100] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是包含提取器并且可选地包含初始化数据的至少一个复合轨。

[0101] 在实施例中,所述微处理器还被配置为执行以下步骤:创建包括使得能够根据所创建的区块轨和所创建的至少一个复合轨形成所述区块化定时媒体数据的空间部分的参数的初始化片段文件,其中所述初始化片段文件包括向所创建的区块轨的参考。

[0102] 在实施例中,所述初始化片段文件还包括与所创建的至少一个复合轨相关联的所参考的区块轨的列表。

[0103] 在实施例中,所述初始化片段文件还包括与所创建的各个区块轨相关联的表示该轨包含空间子样本信息的区块处理程序。

[0104] 在实施例中,所述初始化片段文件还包括与所创建的各个区块轨相关联的一般呈现信息。

[0105] 在实施例中,所述微处理器还被配置成:在针对所选择的各空间子样本创建一个区块轨的步骤中,创建至少两个区块轨,其中所述至少一个复合轨包括对同一定时样本内所包括的两个空间子样本中的任意空间子样本进行解码所要使用的至少一个数据项。

[0106] 在实施例中,所述区块化定时媒体数据是各个所述定时样本包括参考层和至少一个增强层所依据的可分级区块化定时媒体数据,其中所述参考层包括至少一个参考空间子样本并且所述至少一个增强层包括多个增强空间子样本,所述微处理器还被配置成针对所选择的各空间子样本创建一个区块轨的步骤中所创建的区块轨是包括增强空间子样本的增强区块轨,所述微处理器还被配置为执行用于创建包括参考空间子样本的至少一个参考轨的步骤。

[0107] 在实施例中,所述参考层包括多个参考空间子样本,以及所述微处理器还被配置成创建多个参考轨,其中所述多个参考轨中的各参考轨形成参考区块轨。

[0108] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件包括用于存储包括相应轨的定义的元数据的数据结构。

[0109] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件包括用于存储空间子样本的数据结构。

[0110] 在实施例中,通过对所创建的各个轨进行封装所得到的媒体片段文件包括用于存储提取器的数据结构。

[0111] 在实施例中,所述服务器符合超文本传输协议即HTTP,以及通过对所创建的各个轨进行封装所得到的媒体片段文件符合如国际标准化组织所定义的基媒体文件格式和经由超文本传输协议即HTTP的动态自适应流传输格式。

[0112] 在实施例中,一种用于在服务器中对分区定时媒体数据进行封装的装置,所述分区定时媒体数据包括定时样本,各定时样本包括多个空间子样本,所述装置包括:

[0113] 选择单元,用于从所述定时样本其中之一的所述多个子样本中选择至少一个子样本;

[0114] 第一创建单元,用于针对所选择的各空间子样本,创建包括所选择的子样本和其它定时样本各自的一个相应子样本的一个分区轨;

[0115] 第二创建单元,用于创建至少一个参考轨,其中所述至少一个参考轨包括使第一分区轨中所包括的第一空间子样本和第二分区轨中所包括的第二空间子样本相关联的信息;以及

[0116] 封装单元,用于将所创建的各轨独立封装在至少一个媒体片段文件中。

[0117] 可以将这种实施例与以上通过参考本发明的第四方面所公开的前述实施例中的一个或多个相组合。

[0118] 根据本发明的第五方面,提供一种视频编码器,其包括如上所述的装置。

[0119] 根据本发明的第六方面,提供一种用于在客户端装置中从封装在多个媒体片段文件中的分区定时媒体数据提供定时媒体数据位流的装置,所述分区定时媒体数据包括定时样本,各定时样本包括多个子样本,所述多个媒体片段文件包括至少一个参考轨和至少一个分区轨,所述至少一个分区轨各自包括从所述定时样本其中之一的所述多个子样本中所选择的一个子样本并且包括其它定时样本各自的一个相应子样本,所述至少一个参考轨包括识别至少一个分区轨并且识别所识别出的分区轨的至少一个子样本的至少一个提取器,所述装置包括至少一个微处理器,所述至少一个微处理器被配置为执行以下步骤:

[0120] 选择代表子样本的信息项;

[0121] 请求包括所选择的子样本的所述至少一个分区轨,其中所选择的子样本与所选择的信息项相对应;

[0122] 接收所述多个媒体片段文件;以及

[0123] 利用所识别出的相应子样本来替换所述至少一个参考轨的提取器,以生成所述定时媒体数据位流。

[0124] 因此,本发明使得无论客户端装置所下载的部分或区块的所选择集合如何,都能够组合不同的部分、特别是不同的区块,并且能够在解析时创建有效文件格式和基本流。

[0125] 因此,本发明的方法适合独立部分或区块的高效流传输,仅需要要发送至客户端装置的有用数据,适合一个以上的部分或区块选择(即,任意的ROI),降低索引化开销(相比字节范围请求),并且可以归并到MPEG标准中。

[0126] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是至少一个复合轨。

[0127] 在实施例中,所述分区定时媒体数据是区块化定时媒体数据,所述子样本是空间子样本,所述至少一个分区轨是至少一个区块轨,并且所述至少一个参考轨是包含提取器

并且可选地包含初始化数据的至少一个复合轨。

[0128] 在实施例中,所述微处理器还被配置为执行包括以下的步骤:

[0129] 从所述至少一个复合轨的提取器获得轨参考;

[0130] 检查是否接收到与所获得的轨参考相对应的轨;以及

[0131] 如果没有接收到与所获得的轨参考相对应的轨,则去除相应的提取器以生成所述定时媒体数据位流。

[0132] 在实施例中,所述微处理器还被配置为执行以下步骤:验证为与所获得的轨参考相对应的轨是区块类型的轨。

[0133] 在实施例中,所述微处理器还被配置为执行包括以下的步骤:

[0134] 从所述至少一个复合轨的提取器获得轨参考;

[0135] 检查是否接收到与所获得的轨参考相对应的轨;以及

[0136] 如果没有接收到与所获得的轨参考相对应的轨,则通过填充来替换相应的提取器以生成所述定时媒体数据位流。

[0137] 在实施例中,所述微处理器还被配置为执行以下步骤:验证为与所获得的轨参考相对应的轨是区块类型的轨。

[0138] 在实施例中,所述微处理器还被配置为执行以下步骤:接收包括使得能够根据所创建的区块轨和所创建的至少一个复合轨来形成所述区块化定时媒体数据的空间部分的参数的初始化片段文件。

[0139] 在实施例中,所述微处理器还被配置为:接收到代表至少两个区块轨的至少两个媒体片段文件,其中所述至少一个复合轨包括对属于相同集合的空间子样本的两个空间子样本中的任意空间子样本进行解码所要使用的至少一个数据项。

[0140] 在实施例中,所述区块化定时媒体数据是各个定时样本包括参考层和至少一个增强层所依据的可分级区块化定时媒体数据,其中所述参考层包括至少一个参考空间子样本并且所述至少一个增强层包括多个增强空间子样本;以及所述微处理器还被配置成:接收到代表至少两个轨的至少两个媒体片段文件,其中所接收到的至少两个轨中的一个轨是包括参考空间子样本的参考轨,并且所接收到的至少两个轨中的另一个轨是包括增强空间子样本的增强区块轨。

[0141] 在实施例中,所述客户端装置符合超文本传输协议即HTTP,以及所接收到的媒体片段文件符合如国际标准化组织所定义的基媒体文件格式和经由超文本传输协议即HTTP的动态自适应流传输格式。

[0142] 根据本发明的第七方面,提供一种视频解码器,其包括如上所述的装置。

[0143] 由于本发明可以以软件来实现,因此本发明可以体现为计算机可读代码以提供至可编程设备的任何适当载体介质上。有形载体介质可以包括诸如软盘、CD-ROM、硬盘驱动器、磁带装置或固态存储器装置等的存储介质。瞬态载体介质可以包括诸如电气信号、电子信号、光学信号、声学信号、磁信号或者例如微波或RF信号的电磁信号等的信号。

附图说明

[0144] 在研究附图和具体实施方式时,本领域技术人员将明白本发明的更多优点。意图是这里还包含了任何附加优点。

- [0145] 现在将仅通过示例方式并且参考以下附图来说明本发明的实施例,其中:
- [0146] 图1a和图1b示出编码视频位流中的编码区块的示例;
- [0147] 图2示出用户为了进行显示所选择的区块的时间管;
- [0148] 图3示出根据特定实施例的使用可分级媒体数据轨的初始化片段文件的框图的示例;
- [0149] 图4示出根据特定实施例的、与诸如图3所示等的初始化片段文件中所声明的轨相对应的包括区块轨和一个复合轨的媒体片段文件的框图的示例;
- [0150] 图5示出连接所下载的媒体数据片段以构建表示给定时间段内的连续视频帧的空间部分的有效可解码定时媒体数据位流的示例;
- [0151] 图6a和图6b是示出根据特定实施例的用于在服务器和客户端装置之间传输定时媒体数据的步骤的流程图;
- [0152] 图7是示出根据客户端装置所接收到的所连接的媒体片段来生成有效定时媒体数据位流的流程图;
- [0153] 图8表示可以实现一个或多个实施例的步骤的服务器或客户端装置的框图;
- [0154] 图9a、图9b和图9c示出HEVC位流中的区块和条带片段的示例;
- [0155] 图10示出根据本发明的实施例的、将HEVC位流封装为包括复合轨和独立区块轨的一组轨的示例;
- [0156] 图11示出根据第一实施例的使用媒体数据轨来对HEVC位流进行封装的初始化片段文件的框图的示例;
- [0157] 图12a和图12b示出适用于处理各种区块化结构的、子样本级别的区块在完整视频中的位置、该区块的大小、以及可以在没有任何伪影 (artifact) 的情况下对区块轨进行解码的指示的信号通知。
- [0158] 图13示出根据本发明的实施例的、将HEVC位流封装为包括复合轨和可作为标准视频轨进行播放的独立区块轨的一组轨的示例;以及
- [0159] 图14示出根据本发明的另一实施例的、将HEVC位流封装为包括复合轨、初始化数据轨以及可作为标准视频轨进行播放的独立区块轨的一组轨的示例。

具体实施方式

[0160] 根据特定实施例,将诸如包括定时样本(例如,图像)的区块化定时媒体数据(例如,视频数据)等的分区定时媒体数据作为多个定时媒体数据轨的集合进行传输,其中这多个定时媒体数据轨典型为基本层轨和多个区块轨、以及包括到定时媒体数据轨的参考的参考或复合轨。各区块轨包括多个定时样本的一个空间子样本(例如,多个NAL单元)。定义扩展提取器类型以从复合轨参考定时媒体数据轨。定时媒体数据轨被标记为不可显示并且输送并描述区块的定时媒体数据。定时媒体数据轨和复合轨的这种集合使得能够进行空间视频区块的选择、组成和高效流传输。可以将各轨作为媒体片段文件的集合从服务器装置传输至客户端装置。可以使用初始化片段文件来传输对媒体片段文件进行解码所需的元数据。

[0161] 图1a和图1b示出编码视频位流中的编码区块的示例。

[0162] 为了进行例示,在以下说明中,考虑为各视频帧(定时样本)包括与该视频帧的空

间子部分(空间子样本)相对应的独立可解码区块。优选该视频是可分级的并且是按可分级的不同级别进行组织的。如图1a所示,视频帧100可以包括HD基本层(102)和4K2K增强层(104)。仍然为了进行例示,可以将增强层104分割成表示为a、b、c和d的四个规则区块。要注意,也可以处理不同形状的区块。同样,可以将基本层102分割成多个区块。在这种情况下,可以使用多个复合轨,例如一个复合轨用于基本层并且一个复合轨用于增强层或用于各个增强层。

[0163] 还要注意,本发明不限于可分级视频格式。还可应用于使得能够独立对区块进行解码的所有视频格式。因此,可以结合本发明的实施例使用如同MPEG4、AVC、HEVC、SVC或将来的SHVC那样的任何视频压缩算法。

[0164] 图1b表示按解码顺序的典型编码视频位流。如图所示,这里编码视频位流包括按时间顺序进行编码的三个视频帧(110、112和114)。各视频帧包括基本层(BL)的所有网络抽象层(NAL)单元及其之后的增强层的NAL单元。例如,在第一个视频帧(110)的基本层(102-1)的NAL单元(1BL,116)之后是第一个视频帧的增强层(104-1)的NAL单元(1common(共通)、1a、1b、1c、1d、118)。

[0165] 与具有空间区块的增强层相对应的视频位流的一部分包括各区块的NAL单元。可选地,该部分还可以包含所有区块共通的并且是对任意区块进行解码所需的NAL单元。给定帧的所有区块共通的NAL单元可以位于视频位流的相应部分中的任何位置(即,视频帧的区块的NAL单元之前、之间或之后)。

[0166] 如图所示,包括空间区块a、b、c和d的与第一个视频帧(110)的增强层相对应的视频位流的一部分包括各区块的NAL单元(1a、1b、1c和1d)以及所有区块a、b、c和d共通的NAL单元(1common)。

[0167] 图2示出用户为了进行显示所选择的区块的时间管。更精确地,图2表示第一个视频帧n和第二个视频帧n+m(其中,n和m是整数值),其中第一个视频帧和第二个视频帧各自包括编号为1~12的十二个区块。在这十二个区块中,仅显示第三个区块和第七个区块(如粗线所示)。视频帧n和n+m属于与给定时间段相对应的一系列连续帧。因此,连续显示帧n~帧n+m中的各帧的第三个区块和第七个区块。

[0168] 如图1a和图1b所示,将视频位流的数据组织为与所有帧相对应的时间样本。因此,在如以上通过参考图2所述在给定时间段内要访问这些帧的特定空间区域的情况下,要求针对各帧访问多个小的字节范围。从所生成的请求的数量和数据开销方面而言,这在HTTP流传输中效率低。

[0169] 因此,为了在ROI流传输所用的压缩视频中提供更为高效的访问,要重新组织定时媒体数据位流,以使得将特定区块的数据组织为给定时间段(即,连续帧的集合)内的连续字节范围(形成管)。

[0170] 因此,在仅要显示视频帧的空间子部分的情况下,必须针对各管并且针对各时间段使用一个HTTP请求来仅下载与所选择的时空区域相对应的区块的管(例如,图2中的区块3和7)。

[0171] 图3示出根据特定实施例的使用媒体数据轨的初始化片段文件的框图的示例。

[0172] 图3所示的初始化片段文件和图4所示的媒体片段文件示出本发明的用以在独立流传输多个轨的情况下以ISO基媒体文件格式重新组织并封装视频位流的实施例。为了进

行例示,图3和4的内容基于图1a和图1b所示的视频位流。

[0173] 为了在独立流传输多个轨的情况下以ISO BMFF重新组织并封装视频位流,定义被称为区块轨的新类型的轨。按照定义,区块轨是作为其中的样本表示与单个时间戳相关联的所有数据的相关样本的定时序的轨。与样本通常是个体视频帧的已知视频媒体轨相反,区块轨的样本定义了整个视频帧的空间限定的子部分。因此,区块轨仅包含与给定区块有关的NAL单元。这样,可以通过将各轨存储在独立片段文件中来创建具有连续字节范围的区块的管。

[0174] 然而,由于区块轨不包含要进行解码并播放的所有所需NAL单元和信息,因此使用被称为“复合轨”的附加轨。根据特定实施例,复合轨是表示完整区块化帧(即,所有区块的组成)的媒体轨。复合轨是如通过参考图1a和图1b所述的典型组织中的编码定时媒体数据位流的表示(按解码顺序以帧为单位)。复合轨使用提取器对象来参考各自的区块轨中的NAL单元。另外,复合轨可以包括对于所有区块共通的NAL单元(在存在这些NAL单元的情况下)。

[0175] 根据特定实施例,使用初始化片段文件来发送定义封装在其它媒体片段文件中的定时媒体数据位流所需的所有元数据。如图3所示,初始化片段文件300包含文件类型框“ftyp”302和动画框“moov”304。文件类型框302优选识别片段文件符合哪些ISO BMF规格并且表示该规格的版本编号。动画框“moov”304提供描述媒体片段文件中所存储的呈现的所有元数据并且特别是呈现中可利用的所有轨。

[0176] 动画框“moov”304包含针对与作为图1a和图1b的示例所提供的可分级视频位流相对应的各个轨(“轨”框306-1~306-6)的定义。

[0177] 轨框306-1表示基本层(track_ID=1),四个轨框306-2~306-5(未示出轨框306-3和306-4)表示增强层(track_ID=2~5)的四个区块a、b、c和d,并且轨框306-6表示描述增强层(track_ID=6)的复合轨。

[0178] 各轨框至少包含一般标记为308的轨头部框“tkhd”和一般标记为310的轨媒体框“mdia”。如果轨依赖于来自其它轨的数据,则还存在轨参考框“tref”。如图所示,具有标识符track_ID=6的复合轨包括轨参考框“tref”312,这表示轨依赖于来自具有标识符track_ID=1~6的轨的数据。

[0179] 要注意,根据用于对定时媒体数据位流进行封装的ISO BMFF规格,其它框可以是强制或可选的。然而,由于本发明的实施例不依赖于这些框可应用,因此这里没有呈现这些框。

[0180] 轨头部框“tkhd”308指定轨的特性。在信息的多个项中,轨头部框“tkhd”308提供轨的标识符(track_ID)、轨的持续时间以及/或者轨的视觉呈现大小(即,显示区域的宽度和高度)。轨头部框“tkhd”308还包括表示轨是否可播放的标志参数。

[0181] 根据实施例,针对区块轨的轨头部标志的默认值为0(track_enabled=0、track_in_movie=0、track_in_preview=0),这意味着针对利用客户端装置的本地重放和预览,区块轨被忽略。在另一实施例中,可以创建新的轨头部标志以用信号通知轨是区块轨。

[0182] 轨媒体框“mdia”310可被视为包含用于声明轨内的定时媒体数据的参数的所有对象的容器。轨媒体框“mdia”310至少包含一般标记为314的媒体头部框“mdhd”、一般标记为316的处理程序参考框“hdlr”和一般标记为318的媒体信息框“minf”。

[0183] 处理程序参考框“hdlr”316声明要呈现轨的定时媒体数据的处理并由此声明轨中的定时媒体数据的性质。例如,将利用视频处理程序(注意,处理程序类型属性等于“vide”)来处理视频轨。可以通过使用VisualSampleEntry()类型的对象来描述视频样本。根据特定实施例,定义被称为区块处理程序(注意,处理程序类型属性等于“tile”)的新的处理程序类型以表示轨包含空间子样本信息。根据编码格式,如果VisualSampleEntry()类型的对象无法描述区块轨中的样本,则可以定义TileSampleEntry()类型的特定对象以描述该样本。

[0184] 媒体信息框“minf”318包含定义轨中的定时媒体数据的特性信息的所有对象。例如,对于复合轨中所定义的基本层和增强层,“minf”框可以包含标准视频媒体头部框“vmhd”(320)。

[0185] 关于区块轨,使用与新的区块处理程序相对应的被称为区块媒体头部框(tmhd, 322)的特定框,独立于编码来针对各区块定义一般呈现信息。特别地,该特定框可以包含区块相对于复合视频轨所表示的视频分辨率而覆盖的空间区域的几何信息,其中可以将该几何信息定义为如下:

```
[0186] aligned(8) class TileMediaHeaderBox
[0187] extends FullBox('tmhd',version=0,0) {
[0188]   unsigned int(16) horizontal_offset;
[0189]   unsigned int(16) vertical_offset;
[0190] }
```

[0191] 如前面所述,复合轨包括用于在呈现中向另一轨提供类型化参考的特定轨参考框“tref”312。根据特定实施例,这些类型化参考可以包括:“tile”参考(324),其可用于建立从复合轨到其所参考的区块轨的链接;以及“scal”参考(326),其可用于建立从包含该参考的轨到其所依赖的定时媒体数据轨(例如,基本层轨(track_ID=1))的链接。

[0192] 图4示出根据特定实施例的、与诸如图3所示等的初始化片段文件中所声明的轨相对应的包括区块轨和一个复合轨的媒体片段文件的框图的示例。如前面所述,图3所示的初始化片段文件和图4所示的媒体片段文件示出本发明的用以在独立对多个轨进行流传输的情况下以ISO基媒体文件格式来重新组织并封装视频位流的实施例。

[0193] 如图4所示,如在DASH标准中所规定的,媒体片段文件400-1~400-6(未示出媒体片段文件400-3~400-5)各自包括一般标记为402的片段类型框“styp”、一般标记为404的至少一个动画片断框“moof”和一般标记为406的至少一个媒体数据框“mdat”。媒体片段文件与HTTP-URL相关联。

[0194] 要注意,根据用于封装定时媒体数据位流的ISO BMFF规格,其它框可以是强制或可选的。然而,由于本发明的实施例不依赖于这些框可应用,因此这里没有呈现这些框。

[0195] 片段类型框“styp”402的格式与图3中的文件类型框“ftyp”302的格式相同,然而,片段类型框“styp”402的参考表示文件是媒体片段文件。

[0196] 动画片断框404提供通常存储在动画框“moov”内的信息。其头部(“mfhd”)包含针对各动画片断而增加的序列号(在图4中标记为seq_num)。这种序列号使得客户端装置能够按不断增大的顺序连接所接收到的片段文件并且(在需要的情况下)验证该序列的完整性。动画片断框404包含针对具有关联的媒体数据框(“mdat”,406)中的数据的各轨的轨片断框

“traf” (一般标记为408)。轨片断框408包括一般标记为410的轨片断头部框“tfhd”,其中该轨片断头部框“tfhd”用于存储存在于相应媒体数据框 (“mdat”, 406) 中的轨的位流的标识符 (track_ID)。

[0197] 媒体数据框通常包含定时媒体数据。在标准视频轨中,该媒体数据框包含视频帧。在区块轨中,媒体数据框406包含完整视频帧的空间限定的子部分。为了便于例示,与轨标识符track_ID=2相关联的媒体数据框包含与增强层的区块相对应的所有NAL单元。

[0198] 在复合轨 (图4中的track_ID=6) 中,媒体数据框406针对各区块并且针对各依赖层包含提取器 (在图4中标记为E), 并且包含对于所有区块共通的NAL单元 (在存在的情况下)。

[0199] 如图4所示,与复合轨相关联的媒体片段文件400-6的媒体数据框406特别包括以下内容:

[0200] -第一提取器412-1,用于提供向在与基本层轨相关联的媒体片段文件400-1的媒体数据框406中所存储的基本层轨内进行编码的基本层数据 (NAL单元1BL) 的链接;

[0201] -NAL单元412-2,其对于多个区块是共通的;

[0202] -第二提取器412-3,其提供向在与增强层的第一区块轨相关联的媒体片段文件400-2的媒体数据框406内进行编码的第一区块的增强层数据 (NAL单元1a) 的链接;

[0203] -第三提取器412-4,其提供向在与增强层的第二区块轨相关联的媒体片段文件400-3 (未示出) 的媒体数据框406内进行编码的第二区块的增强层数据 (NAL单元1b) 的链接;

[0204] -第四提取器412-5,其提供向在与增强层的第三区块轨相关联的媒体片段文件400-4 (未示出) 的媒体数据框406内进行编码的第三区块的增强层数据 (NAL单元1c) 的链接;以及

[0205] -第五提取器412-6,其提供向在与增强层的第四区块轨相关联的媒体片段文件400-5 (未示出) 的媒体数据框406内进行编码的第四区块的增强层数据 (NAL单元1d) 的链接。

[0206] 由于提取器412-1而可以获得的NAL单元使得能够对如下的帧的基本层进行解码,其中使用NAL单元412-2以及由于提取器412-3~412-6而可以获得的NAL单元可以完全解码该帧的增强层。如从图4可以观察到,如果仅要解码帧的空间部分,则不必下载所有的媒体片段文件400-2~400-5 (即,与区块轨相对应的位流)。

[0207] 根据特定实施例,提取器是具有以下句法的文件格式内部结构:

```

class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
[0208]    unsigned int((lengthSizeMinusOne+1)*8)
        data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_length;
}

```

[0209] 其中, NALUnitHeader() 表示符合用于对视频位流进行编码的编码格式的NAL单元的开头四个字节。这四个字节标识NAL单元作为提取器(例如, 在SVC中, 将属性nal_unit_type设置到提取器NAL单元类型(类型31))。

[0210] 值track_ref_index在复合轨的类型“scal”或“tile”的轨参考框“tref”中指定要使用的索引, 以寻找要提取数据的轨。值sample_offset给出要用作信息的源的所链接的轨中的样本的相对索引。值data_offset和data_length分别是要复制的参考样本内的第一字节的偏移和要复制的字节的数量。

[0211] 为了进行例示并且参考图3, 如果给定提取器的track_ref_index的值等于2, 则这意味着提取器参考利用tref框中的第二个条目所标识的轨(即, 具有标识符track_ID=2的轨, 也就是说针对区块a的区块轨, 其中第一个索引表示参考轨(例如, 基本层))。

[0212] 图5示出连接所下载的媒体数据片段以构建表示给定时间段内的连续视频帧(这里, 与两个连续帧相对应)的空间部分的有效可解码定时媒体数据位流的示例。针对其它时间段, 可以重复该图。

[0213] 如参考图3和4所述, 优选将定时区块化媒体数据位流作为包括一个初始化片段文件和多个媒体片段文件的一组数据进行发送, 其中多个媒体片段文件包括多个区块轨和一个复合轨。

[0214] 初始化片段文件包括提供与各轨有关的一般信息的动画框(“moov”), 特别是轨的类型(例如, 媒体轨(音频或视频)或区块轨)、编码格式、帧分辨率以及(在轨参考框“tref”中所给出的)轨之间的依赖性。使用这些数据来处理所下载的媒体片段文件。参考关于图1a、图1b、图3和图4所述的示例, 初始化片段文件的动画框的内容特别可以包括以下内容:

```

[0215]  MOOV
[0216]  -track 1:base layer
[0217]  -track 2:tile a
[0218]  -track 3:tile b
[0219]  -track 4:tile c
[0220]  -track 5:tile d
[0221]  -track 6:enhancement layer
[0222]  o tref (scal) : track_ID=1
[0223]  o tref (tile) : track_ID=2

```

[0224] track_ID=3

[0225] track_ID=4

[0226] track_ID=5

[0227] 图5粗略示出在从服务器仅下载所需的媒体片段文件(这里与基本层以及增强层中的区块a和c相对应)的情况下通过连接媒体片段所获得的文件格式。要注意,这种机制不仅使得能够仅下载所需的媒体片段文件,而且还防止了下载重复数据。

[0228] 如图所示,复合轨500使得能够通过参考来自基本层轨504(在可分级性的情况下)和来自不可播放的区块轨(506和508)的数据并且通过适当处理参考丢失数据的提取器(如参考图7所述),来使得能够构建有效的可解码定时媒体数据位流502。

[0229] 所获得的文件格式符合可分级文件格式定义。例如,客户端装置可以决定通过选择基本层轨来仅播放基本层,或者通过选择复合轨来选中具有区块a和c的增强层。客户端装置在继续播放复合轨的同时,还可以通过在稍后的时间段中下载不同的“区块轨”(即,媒体片段文件)来改变要显示的区块。

[0230] 图6a和图6b是示出根据特定实施例的用于在服务器和客户端装置之间传输定时媒体数据的步骤的流程图。在服务器内实现图6a所示的步骤,以通过根据区块化定时媒体数据位流创建适用于ROI流传输的片段区块来准备媒体呈现,而图6b所示的步骤是在客户端装置内实现的。

[0231] 在第一步骤(步骤600)中,服务器识别与区块相关联的所有NAL单元,并且针对各区块,创建包含子样本的区块轨,其中这些子样本包括与给定区块相对应的所有NAL单元。例如,服务器可以依赖于子图片级别SEI消息以识别NAL单元与不同区域的关联,并且由于在HEVC标准(提案JCTVC-K0128)中提出了ROI,因此服务器可以依赖于序列级别SEI消息以识别各ROI的位置和大小。因此,服务器可以针对给定时间段创建区块的管。

[0232] 在下一步骤(步骤602)中,服务器创建复合轨,其中该复合轨包含链接至下级可分级性轨(在存在的情况下)的提取器、对于所有区块而言共通的NAL单元、以及链接至各区块轨的提取器。有利地对提取器和共通的NAL单元进行排序,以使得利用所参考的数据替换提取器得到包括按解码顺序的完整样本的有效定时媒体数据位流(如参考图1a和图1b所述)。

[0233] 接着,在步骤604中,如参考图3和4所述,服务器根据ISO BMFF表示来生成并存储初始化片段文件和包含时间段的媒体片段文件。将所有的定时媒体数据轨(例如,视频轨)、复合轨和区块轨存储在单独的媒体片段文件中。

[0234] 然后,服务器根据请求将初始化片段文件和媒体片段文件提供至客户端装置(步骤606)。服务器可以是对HTTP请求作出响应的传统的HTTP服务器。

[0235] 在HTTP流传输的上下文中并且在优选实施例中,假定客户端装置有权访问描述服务器可利用的媒体呈现的清单文件。该清单文件向客户端装置提供充足的信息(媒体性质和片段的列表),以通过从服务器首先请求初始化片段然后请求媒体片段文件来流传输媒体呈现。

[0236] 在客户端装置端处(通常在具有诸如指示装置等的选择部件的显示器上)选择ROI时,在区块化视频的流传输期间,确定与所选择的ROI相对应的区块(步骤608)。

[0237] 接着,针对各时间段,在可分级媒体数据的情况下,客户端装置向服务器发送请求以下载与依赖层相对应的片段文件(步骤610)。根据特定实施例,在从属于依赖层的层之

前,下载这些依赖层。例如,在增强层片段文件之前,下载基本层片段文件。

[0238] 在下一步骤中,客户端装置向服务器发送请求以下载与所选择的区块相对应的媒体片段文件(步骤612),并且向服务器发送请求以下载与复合轨相对应的媒体片段文件(步骤614)。

[0239] 接着,利用客户端装置连接所下载的片段文件以构建与所选择的ROI相对应的符合ISO BMFF标准的有效(可解码)定时媒体数据位流(步骤616)。

[0240] 参考图7来详细说明步骤616。

[0241] 图7是示出根据客户端装置所接收到的所连接的媒体片段来生成有效定时媒体数据位流、也就是说在利用客户端装置播放复合轨的情况下根据所选择的区块和一个复合轨生成可解码位流的流程图。

[0242] 在第一步骤(步骤700)中,客户端装置接收先前已请求的媒体片段文件(例如,图6b中的步骤612、614和616),并且进行测试以判断是否接收到至少一个媒体片段文件(步骤702)。如果没有接收到媒体片段文件,则处理结束。

[0243] 如果接收到至少一个媒体片段文件并且如果所接收到的媒体片段文件不包含复合轨(即,这些媒体片段文件包含依赖轨、例如低级层或区块轨),则可以缓冲该至少一个媒体片段文件以供随后使用。

[0244] 相反,如果接收到至少一个媒体片段并且如果所接收到的媒体片段文件包含复合轨,则对与该复合轨相对应的位流(通常为NAL单元)进行解析,以从所接收到的媒体片段的媒体数据框“mdat”中提取第一数据项(或者在已处理了所接收到的媒体片段的至少一个数据项的情况下为下一数据项、通常为NAL单元)(步骤704)。如果在媒体数据框中不存在更多的数据项,则处理返回至步骤700以接收新的媒体片段文件(步骤706)。

[0245] 接着,进行测试以判断所提取的数据项(例如,所提取的NAL单元)是否与提取器相对应(步骤708)。如果所提取的数据项不与提取器相对应,则返回该数据项以由视频解码器进行进一步解码(步骤710)。相反,如果所提取的数据项是提取器,则必须利用正参考的数据项来替换该数据项。为此,从提取器的结构获取提取器的参数的值(步骤712)。如上所述,提取器包括从另一轨中提取数据所需的所有参数值(例如,track_ref_index、sample_offset、data_offset和data_length)。

[0246] 一旦识别出所参考的轨的标识符,则进行测试(步骤714),以判断在步骤700期间所缓冲的媒体片段文件的集合中是否可利用所参考的轨。需要重申,由于客户端装置仅下载与所选择的关注区域相对应的媒体片段文件,因此一些区块轨缺失。

[0247] 如果在步骤700期间所缓冲的媒体片段文件的集合中可利用所参考的轨,则利用正参考的数据来替换提取器(步骤716),并且将位流发送至视频解码器以进行解码(步骤710)。

[0248] 如果在步骤700期间所缓冲的媒体片段文件的集合中不可利用所参考的轨,则由于根据ISO BMF标准、提取器中所参考的数据的不存在会导致发生严重错误,因此必须进行特定步骤。进行测试,以判断所参考的轨是否是区块轨(所参考的轨可以对应于依赖分级性层)以及提取器是否具有区块类型(步骤718)。

[0249] 如果所参考的轨不是区块轨或者如果提取器不具有区块类型,则检测到标准严重错误。相反,如果所参考的轨是区块轨或者如果提取器具有区块类型,则根据用于对定时媒

体数据位流进行编码的编码格式(步骤720),去除提取器(步骤722)、或者通过来自替代的包含针对缺失区块的“跳过”数据的“填充轨”或“填充框”的填充来替换提取器(步骤724)。这里,“跳过”数据表示利用从属于同一可分级层或属于另一可分级层的先前解码的图像所获得的其它像素数据进行替换的当前图像中缺失的像素数据。通常利用至少一个标志来表示“跳过”数据。例如,在考虑HEVC视频压缩格式的情况下,填充数据可以是排他地包含利用设置为1的跳过标志进行编码的编码单位的一个或多个NALU。

[0250] 接着,将位流传输至视频解码器以进行解码(步骤710)并进行显示,并且处理循环至步骤704以处理下一数据项。

[0251] 图8表示可以实现一个或多个实施例的步骤的服务器或客户端装置800的框图。

[0252] 优选地,装置800包括:通信总线802;中央处理单元(CPU)804,其能够执行在对装置通电时来自程序ROM 806的指令、以及与通电之后来自主存储器808的软件应用程序有关的指令。主存储器808例如是经由通信总线802用作CPU 804的工作区域的随机存取存储器(RAM)型,并且其存储容量可以利用连接至扩展端口(未示出)的可选RAM来扩展。例如可以从硬盘(HD)810或程序ROM 806将与软件应用程序有关的指令加载至主存储器808。这种软件应用程序在由CPU 804执行的情况下,使得在服务器中进行参考图6a所述的步骤并且在客户端装置中进行参考图6b和7所述的步骤。

[0253] 附图标记812是网络接口,其中该网络接口812使得能够将装置800连接至通信网络814。软件应用程序在由CPU 804执行的情况下,被配置为对经由网络接口所接收到的请求作出反映,并且将数据流和请求经由网络提供至其它装置。

[0254] 附图标记816表示用以向用户显示信息和/或从用户接收输入的用户接口。

[0255] 这里,应当指出,作为变形,用于管理多媒体位流的接收或发送的装置800可以包括能够实现如参考图6a、6b和7所述的方法的一个或多个专用集成电路(ASIC)。这些集成电路例如且非限制性地集成于用于生成或显示视频序列以及/或者用于收听音频序列的设备中。

[0256] 如上所述,本发明的实施例特别可应用于已知为HEVC的视频格式。

[0257] 根据HEVC标准,可以将图像从空间上分割成区块、条带和条带片段。在该标准中,区块与利用水平边界和垂直边界(即,行和列)所定义的图像的矩形区域相对应。区块包含整数个编码树单元(CTU)。因此,区块可以高效地用于例如通过定义关注区域的位置和大小来识别关注区域。然而,HEVC位流的结构及其作为网络抽象层(NAL)单元的封装并非是考虑到区块而是基于条带进行组织的。

[0258] 在HEVC标准中,条带是条带片段的集合,其中条带片段的集合中的第一个条带片段是独立条带片段、即头部内所存储的一般信息未参考另一条带信息中的信息的条带片段。条带片段的集合中的其它条带片段(在存在的情况下)是依赖条带片段(即,头部内所存储的一般信息参考独立条带片段中的信息的条带片段)。

[0259] 条带片段包含整数个连续的(按光栅扫描顺序)编码树单元。因此,条带片段可以具有矩形形状或不具有矩形形状,因此不适合表示关注区域。根据在条带片段头部之后为条带片段数据的形式,将条带片段编码在HEVC位流中。独立条带片段和依赖条带片段在头部方面不同:由于依赖条带片段依赖于独立条带片段,因此其头部的信息量小于独立条带片段的头部的信息量。独立条带片段和依赖条带片段这两者都包含相应位流中的用于定义

区块或用作熵解码同步点的进入点的列表。

[0260] 图9a、图9b和图9c示出区块和条带片段的示例。更精确地,图9a示出利用垂直边界905-1和905-2以及水平边界910-1和910-2分割成九个部分的图像(900)。标记为915-1~915-9的这九个部分各自表示特定区块。

[0261] 图9b示出包含利用垂直边界905'定界的两个垂直区块的图像(900')。图像900'包括单一条带(无附图标记),其中该一个条带包含五个条带片段、一个独立条带片段920-1(利用阴影线表示)和四个依赖条带片段920-2~920-5。

[0262] 图9c示出利用垂直边界905''定界的两个垂直区块的图像(900'')。左侧区块包括两个条带:第一个条带包含一个独立条带片段(920'-1)和一个依赖条带片段(920'-2),并且第二条带也包含一个独立条带片段(920'-3)和一个依赖条带片段(920'-4)。右侧区块包括包含一个独立条带片段(920'-5)和一个依赖条带片段(920'-6)的一个条带。

[0263] 根据HEVC标准,按照可以总结为如下的规则(必须满足一个或者两个条件),条带片段链接至区块。

[0264] -条带片段中的所有CTU都属于同一区块(即,条带片段不能属于多个区块);以及

[0265] -区块中的所有CTU都属于同一条带片段(即,可以将区块分割成多个条带片段,只要这些条带片段各自仅属于该区块即可)。

[0266] 为了明确,以下考虑一个区块包含仅具有一个独立条带片段的一个条带。然而,可以利用如图9b和9c所示的结构那样的其它结构来执行本发明的实施例。

[0267] 如上所述,尽管可以将区块视为针对关注区域的适当支持,但条带片段是实际放置在NAL单元中以经由通信网络进行传输并聚合以形成访问单位(即,文件格式级别的编码图片或样本)的实体。

[0268] 需要重申的是,根据HEVC标准,将NAL单元的类型编码在可以如下进行定义的NAL单元头部的两个字节中:

```
[0269]  nal_unit_header() {  
[0270]  forbidden_zero_bit  
[0271]  nal_unit_type  
[0272]  nuh_layer_id  
[0273]  nuh_temporal_id_plus1  
[0274]  }
```

[0275] 用于对条带片段进行编码的NAL单元由于条带片段地址句法元素而包括表示第一个CTU在条带片段中的地址的条带片段头部。可以按照如下定义这些条带片段头部:


```

    slice_segment_header () {
        first_slice_segment_in_pic_flag
        if(nal_unit_type >= BLA_W_LP && nal_unit_type <= RSV_IRAP_VCL23)
            no_output_of_prior_pics_flag
        slice_pic_parameter_set_id
[0276]    if(!first_slice_segment_in_pic_flag){
        if(dependent_slice_segments_enabled_flag)
            dependent_slice_segment_flag
            slice_segment_address
    }
    If(!dependent_slice_segment_flag){
        [...]

```

[0277] 在PPS (Picture Parameter Set, 图片参数集) NAL单元中设置区块化信息。然后, 可以根据这些参数来推导条带片段和区块之间的关系。

[0278] 在(按照定义)在区块边界上重置空间预测的情况下, 没有什么防止区块使用来自参考帧中的不同区块的时间预测结果。因此, 为了构建独立区块, 在编码期间有利地将针对预测单位的运动矢量限制在区块内部, 以保持参考帧中的同位置区块中。另外, 优选在区块边界上禁用环路滤波器(去块和样本自适应偏移(SAO)滤波器), 以使得在对仅一个区块进行解码的情况下不引入误差漂移。要注意, 在HEVC标准中可利用环路滤波器的这种控制。将该控制设置在具有已知为loop_filter_across_tiles_enabled_flag的标志的条带片段头部中。通过明确地将该标志设置为零, 区块边界处的像素无法依赖于落在相邻区块的边界上的像素。在满足与运动矢量和环路滤波器有关的这两个条件的情况下, 可以将区块视为“独立可解码区块”或“独立区块”。

[0279] 在将视频位流编码作为一组独立区块的情况下, 该视频位流使得能够在不存在丢失参考数据或传播重建误差的风险的情况下, 针对各帧进行基于区块的解码。然后, 该结构使得能够仅重建例如可以与原始视频的如图2所示的关注区域(包括区块3和7)相对应的空间部分。可以将这种结构表示为视频位流中的补充信息, 从而表示基于区块的解码是可靠的。

[0280] 根据本发明的实施例, 通过使用应用于HEVC标准的ISO BMFF文件格式来提供HTTP流传输的上下文中的向区块的高效访问。因此, 如以下参考图10所述, 利用被称为“区块轨”的特定轨来表示要进行编码的各个独立区块(例如, 图2所示的十二个区块中的每一个区块)。

[0281] 如图10所示, 在与完整帧中的HEVC位流相对应的复合轨中(经由包含针对各个轨的定义的动画框“moov”的轨参考框“tref”来)参考这些区块轨。各区块轨包含以NAL单元为单位进行封装的压缩视频数据。复合轨包含与初始化数据相对应的各种参数集(例如, 视频参数集、序列参数集和/或图片参数集)。复合轨还包含作为特定类型的NAL单元的提取器。

[0282] 如上所述, 提取器可以是具有以下句法的文件格式内部结构:

```

class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
[0283]    unsigned int((lengthSizeMinusOne+1)*8)
        data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_length;
}

```

[0284] 提取器用作向来自其它轨的数据的指针或参考,并且代替两个轨中的数据重复,使得能够参考依赖轨来构建紧凑轨。提取器优选使用NAL单元句法。因此,提取器包括结构与NAL单元头部相同的头部,其中该头部特别包括与NAL单元类型有关的信息。例如,将该NAL单元类型设置为当前与HEVC中的预留NAL单元类型相对应的值“47”。在该头部之后是轨参考框(tref)中的索引(表示为track_ref_index),其中该索引使得能够检索包含与提取器所参考的轨相对应的轨的标识符(track_id)的tref框的条目。第三个参数是提取器所参考的样本相比当前样本的时间偏移(sample_offset)。第四个参数和第五个参数(表示为data_offset和data_length)分别提供复制源自于的位置(优选以字节为单位)以及要复制的数据量(预留值0以表示复制整个所参考的NAL单元)。

[0285] 图10示出根据本发明的实施例的将HEVC位流封装为包括复合轨和独立区块轨的一组轨的示例。为了进行例示,封装后的位流与图2示意性示出的视频序列相对应。

[0286] 如图所示,封装后的位流1000包括:初始化片段文件1005,其包含提供针对轨的定义的动画框(“moov”);以及媒体片段文件1010,其表示复合轨1015和12个区块轨1020-1~1020-12(区块轨1020-1~1020-12各自与视频序列中的一个区块相关联)。

[0287] 如在DASH标准中所指定的,复合轨1015包括:片段类型框“styp”(未示出);至少一个动画片断框“moof”1025,其包括诸如轨片段类型和标识符等的元数据;以及至少一个媒体数据框“mdat”1030,其针对各视频数据样本包括PPS和向视频数据的参考。

[0288] 同样,区块轨1020-1~1020-12各自包括:片段类型框“styp”(未示出);至少一个动画片断框“moof”,其包括诸如轨片段类型和标识符等的元数据;以及至少一个媒体数据框“mdat”,其包括以NAL单元(NALU)为单位进行封装的压缩视频数据。

[0289] 在初始化片段文件1005的轨参考框“tref”1035(在具有标识符id=1的复合轨的定义中,更精确为初始化片段文件1005的动画框“moov”)中参考具有标识符2~13的区块轨1020-1~1020-12。

[0290] 如图所示,复合轨1015包括用作向来自其它轨的数据的指针或参考的提取器。为了进行例示,示出多个参数,其中在这些参数中,区块轨的索引(track_ref_index)、数据偏移(data_offset)和数据长度(data_length)与复合轨1015的提取器1035-1和1035-p相对应。

[0291] 仍然为了进行例示,在处理复合轨1015的NAL单元1035-1的情况下,判断为NAL单元1035-1表示提取器类型的NAL单元(NALUnitHeader等于十六进制值5E00)。因此,处理NAL单元1035-1,以恢复相应的压缩视频数据。为此,获得其区块轨索引(即,track_ref_index

=1)。根据该索引,可以根据初始化片段文件1005中所存储的区块轨定义来恢复区块轨标识符。在给定示例中,由于索引等于1,因此选择tref框的第一个区块轨标识符(id=2)。接着,使用该标识符来访问相应的区块轨,然后使用提取器1035-1的数据偏移(即,要用作信息源的所识别出的轨中的样本的相对索引)和数据长度(即,要复制的字节的数量、例如在data_length=0的情况下的整个NALU)参数,来从区块轨1020-1中提取压缩视频数据(即,在给定示例中为编码条带片段NALU 1040)。

[0292] 在进行了处理之后,利用所参考的数据来替换提取器。根据图10所示的示例,提取器1035-1的解析和处理使得能够利用编码条带片段NALU 1040来进行替换,由此形成HEVC兼容位流。

[0293] 要注意,用于存储HEVC提取器的参数的语义可以接近SVC标准中所定义的语义。因此,对于参考HEVC NAL单元的提取器,可以应用以下:

[0294] -如在ISO/IEC 23008-2所规定的,设置已知为forbidden_zero_bit的参数;

[0295] -将已知为nal_unit_type的参数设置为47(当前FDIS中的预留码);

[0296] -从提取器所参考的第一个NALU复制已知为nuh_layer_id和nuh_temporal_id_plus1的参数(参考HEVC NAL单元的HEVC轨中的提取器不参考具有不同的nuh_layer_id值和nuh_temporal_id_plus1值的多个NAL单元);以及

[0297] -将已知为sample_offset的参数设置为0。

[0298] 根据本发明的实施例,区块轨要被视为用于支持空间排列的数据(例如,HEVC区块)的索引化的特定轨。因此,针对各轨需要特定信号,以特别表示区块在整个视频中的位置(水平坐标和垂直坐标)以及区块的大小(宽度和高度),并且提供用以表示是否可以在没有任何伪影的情况下对区块轨进行解码的信息。可以根据多个实施例来实现这种信号通知。

[0299] 根据图11所示的特定实施例,与上述实施例相同,可以在轨级别进行信号通知。

[0300] 图11示出根据第一实施例的使用媒体数据轨来对HEVC位流进行封装的初始化片段文件的框图的示例。这基于参考图10所述的示例。

[0301] 使用初始化片段文件1100和媒体片段文件(未示出)来以ISO基媒体文件格式重新组织并封装符合HEVC标准的视频位流。单独进行轨的流传输。

[0302] 如上所述,使用初始化片段文件来传输定义封装在其它媒体片段文件中的定时媒体数据位流所需的所有元数据。如图11所示,初始化片段文件1100包含文件类型框“ftyp”1105和动画框“moov”1110。文件类型框1105优选识别片段文件符合哪些ISO BMF规格并且表示该规格的版本编号。动画框“moov”1110提供描述媒体片段文件中所存储的呈现的所有元数据以及特别是呈现时可利用的所有轨。

[0303] 动画框1110包含针对各个轨(“track”框1115-1~1115-13)的定义,其中在给定示例中,这些轨包括一个复合轨(1115-1)和12个区块轨(1115-2~1115-13)。

[0304] 各轨框至少包含一般标记为1120的轨头部框“tkhd”和一般标记为1125的轨媒体框“mdia”。如果轨依赖于来自其它轨的数据,则还存在轨参考框“tref”。如图所示,具有标识符track_ID=1的复合轨包括轨参考框“tref”1130,这表示轨依赖于来自具有标识符track_ID=2~13的区块轨的数据。

[0305] 如上所述,要注意,根据用于对定时媒体数据位流进行封装的ISO BMFF规格,其它

框可以是强制或可选的。然而,由于本发明的实施例不依赖于这些框可应用,因此这里没有呈现这些框。

[0306] 根据参考图11所述的实施例,使用轨头部框“tkhd”(1120)和媒体信息框“mdia”的框(1125),在各轨定义中,在“moov”框(1110)中,针对要进行封装的整个HEVC位流进行一次如下内容的信号通知:区块在整个视频中的位置、区块的大小、以及可以在没有任何伪影的情况下对区块轨进行解码的表示。

[0307] 将区块的位置放置在被称为定义水平偏移和垂直偏移(horizontal_offset和vertical_offset)的TileMediaHandlerEntry框或“tmhd”(1140)的新类型的媒体头部信息框1135中。

[0308] 将所考虑的区块的视觉呈现的大小放置在现有的轨头部的宽度参数和高度参数中。在需要的情况下,可以(经由已知为SampleDescriptionBox的框中的已知为VisualSampleEntry()的参数来)在样本描述中记录所考虑的区块的真实像素尺寸。

[0309] 最后,通过在处理程序框“hdlr”(1145)中定义新类型的轨(“tile”)来明确地用信号通知特定轨(区块轨)。

[0310] 初始化片段文件1100的动画框“moov”1110还包括“mvex”框1150。该框用于通知客户端访问存在动画片断的封装文件。该框使得能够在初始化片段文件中指定呈现时的最长轨的持续时间。该框使呈现持续时间的计算更为简单,从而避免了各动画片断持续时间的检查。如图所示,mvex框1150针对各轨包含轨扩展框,以避免对于各轨的所有片断(即,区块轨和复合轨)共通的信息的重复、例如轨标识符和轨中的样本的默认大小。

[0311] 要注意,用信号通知区块轨的这种实施例相比通常的视频轨信号通知没有引入大的开销。此外,该实施例仍然独立于呈现被分割得到的动画片断的数量。然而,其不适用于沿着视频序列改变的区块化结构。

[0312] 根据适用于处理沿着视频序列的区块化结构的变化的另一特定实施例,使用来自ISO BMFF标准的样本分组机制来以样本级别进行信号通知。

[0313] 这种样本分组机制用于表示轨中的样本的分区。这些样本分组机制依赖于以下两个框的使用:SampleToGroup框(“sbgp”),用于描述样本向样本组的分配;以及SampleGroupDescription框(“sgpd”),用于描述特定样本组内的样本的共通特性。经由类型字段(“grouping_type”)利用一个SampleToGroup框和一个SampleGroupDescription框的组合来定义特定类型的样本分组。多个样本分组实例(即,SampleToGroup框和SampleGroupDescription框的对)可以基于不同的分组标准而共存。

[0314] 根据本发明的实施例,定义与样本的区块化有关的新的分组标准。被称为“区块”的该新的grouping_type描述区块的性质并且是根据标准VisualSampleGroupEntry所推导出的。该grouping_type可被称为TileRegionSampleGroupEntry或HEVCSpatialEntry,并且按照如下进行定义:

[0315] class HEVCSpatialEntry() extends VisualSampleGroupEntry('trsg') {

[0316] unsigned int(32) tileID;

[0317] unsigned int(16) horizontal_offset;

[0318] unsigned int(16) vertical_offset;

[0319] unsigned int(16) region_width;

```
[0320] unsigned int (16) region_height;  
[0321] unsigned int (2) independent;  
[0322] unsigned int (6) reserved=0;  
[0323] }
```

[0324] 根据该新类型的组条目, tileID参数是针对该组所描述的区块的唯一标识符。在基区域的亮度样本中, 使用horizontal_offset参数和vertical_offset参数来分别设置利用区块所表示的矩形区域的左上角像素相对于HEVC帧的左上角像素的水平偏移和垂直偏移。在HEVC帧的亮度样本中, 使用region_width参数和region_height参数来分别设置利用区块所表示的矩形区域的宽度和高度。independent参数是2位字, 其中该2位字如以上参考独立区块的定义所述指定如下内容: 区块包括与仅属于相同区块的样本有关的解码依赖性。为了进行例示并且参考用于描述区块组织的SEI消息的标准使用, 可以使用已知为tile_section_exact_match_flag的标志来设置独立标志的值。可以按照如下设置后者的含义:

[0325] -如果independent参数等于0, 则该区块与相同帧或前一帧中的其它区块之间的编码依赖性是可知的;

[0326] -如果independent参数等于1, 则在该区块和同一帧中的其它区块之间不存在空间编码依赖性, 但在该区块和前一帧中的具有相同tileID的区块之间可以存在编码依赖性; 以及

[0327] -如果independent参数等于2, 则在该区块和同一帧或前一帧中的具有相同tileID的其它区块之间不存在编码依赖性;

[0328] 预留independent参数值3。

[0329] 可选地, 将描述针对各区块的平均位率的参数设置在区块描述符中, 从而提供至流传输客户端以进行基于带宽的适应性改变。

[0330] 根据本实施例, 通过针对各区块轨定义具有“区块”grouping_type和HEVCSpatialEntry的一个SampleGroupDescription框(“sgpd”)来在动画头部(“moov”框)中一次给出各区块的性质。然后, 根据ISO BMFF标准, 由于预先并不知晓样本的数量, 因此在各区块轨片断中定义SampleToGroup框以使区块轨片断的各样本与其性质相关联。

[0331] 在区块的网格随时间的经过而改变的情况下, 可以在轨片断框(“traf”)中定义具有新的HEVCSpatialEntry的新的SampleGroupDescription框(“sgpd”)并且可以利用SampleToGroup框(“sbgp”)参考该SampleGroupDescription框(“sgpd”)。因此, 在网格随时间的经过呈静态的情况下, 针对各区块轨和区块轨片断定义至少一个SampleToGroup框。该框在描述方面表示至少28个字节。假定具有持续时间为2秒的片断的16个区块, 仅针对SampleToGroup框, 这将表示1792位/秒以用信号通知随时间经过的区块化结构。在网格随时间的经过正改变的情况下, (在数据量方面的)成本将变得更高。如以下所述, 可以减少补充的初始化数据量。

[0332] 要注意, 样本组框提供用以向各个样本添加元数据的高效且可扩展的工具。然而, 常见的是, 所给定的元数据针对给定轨的所有样本均为有效, 利用具有各AU中的固定NALU模式的区块化描述符(即, 利用沿着时间的静态区块化结构)可以是这种情况。

[0333] 这可以通过相应地使用SampleToGroup框标示组的各样本来进行表示。

[0334] 可选地,为了减少每区块轨的初始化数据量,可以将一些样本组标记为“默认”(即,针对所有样本均有效)。由于针对静态区块化结构在“moov”框级别进行定义(针对整个序列的各轨进行一次),因此该默认分组在字节方面限制了描述成本。

[0335] 为此,可以在已知为SampleGroupDescriptionBox的框中使用组描述类型的第二个版本(版本=2)(针对各traf/stbl框可以存在多个SampleGroupDescriptionBox),这(经由已知为grouping_type的参数)表示所参考的样本组适用于当前轨或当前轨片断中的所有样本。

[0336] 可以按照如下定义新版本的样本组描述框:

```

        aligned(8) class SampleGroupDescriptionBox (unsigned int (32)
handler_type) extends FullBox('sgpd', version, 0) {
    unsigned int(32) grouping_type;
    if (version == 1) || (version == 2) {unsigned int (32) default_length;}
    unsigned int (32) entry_count;
    int i;
    for ( i = 1; i <= entry_count; i++ ) {
        if (version != 0) {
            if (default_length == 0) {
                unsigned int(32) description_length;
            }
        }
    }
    switch(handler_type) {
        case 'vide': // for video tracks
            VisualSampleGroupEntry(grouping_type);
            break;
        case 'soun': // for audio tracks
            AudioSampleGroupEntry(grouping_type);
            break;
        case 'hint': // for hint tracks
            HintSampleGroupEntry(grouping_type);
            break;
    }
}

```

[0337]

[0338] 根据该扩展类型的SampleGroupDescription框,在版本等于2的情况下,该组描述和任何SampleGroupEntry均适用于当前轨或轨片断的所有样本。grouping_type参数是识别与该样本组描述相关联的SampleToGroup框的整数。要注意,在轨或轨片断在其初始化数据中包括版本=2的SampleGroupDescription框的情况下,无需在这些初始化数据中放置描述相应的grouping_type的SampleToGroup框(这是因为描述按照定义适用于所有的样本)。

[0339] 在新版本的样本组描述框的该定义中,entry_count是下表中给出条目的数量的

整数,default_length表示每个组条目的长度(在长度恒定的情况下)或者在可变的情况下为零,并且description_length在针对各条目改变并且因此default_length为零的情况下表示各个组条目的长度。

[0340] 根据本实施例,grouping_type可以具有用以指示空间/区块分组的特定值:该值例如可以是与针对“tile”的ASCII码(0x74696C65)相对应的十六进制值。针对一个轨,应存在针对grouping_type具有相同值的该框的至多一次发生。

[0341] 要注意,在自适应网格随时间的经过而移动的情况下,样本组框保持相同(即,“区块”grouping_type)并且继续应用于所有的样本。如此,在区块化结构相对于moov/trak/mdia/minf/stbl框中用信号通知的默认结构发生改变的区块轨的轨片断中,仅需更新样本组描述框。这样减少了自适应区块的信号通知成本。

[0342] 可选地并且仍为了减少各区块轨的初始化数据量(以避免在各区块轨片断中重复SampleToGroup框),将被称为“dsdp”的新的DefaultSampleToGroups框(或具有相同语义的另一相似框,而与名称无关)定义为作为初始化信息的一部分而仅包括在来自各moov/trak框的SampleTable框(“stbl”)中。该新框使将适用于轨内的所有样本的一组样本组描述关联至所有样本。

[0343] 可以按照如下所述定义新的DefaultSampleToGroup框:

```

                                aligned(8) class DefaultSampleToGroups extends FullBox('dsdp', version,
[0344]                                0) {
                                    unsigned int(32) entry_count;
                                    for (i=1; i <= entry_count; i++) {
                                        unsigned int(32) grouping_type;
                                        if (version == 1) {
                                            unsigned int(32) grouping_type_parameter;
                                        }
                                    }
                                    unsigned int(32) group_description_index;
[0345]                                }
                                }

```

[0346] 其中,entry_count参数给出要关联至各样本的组的列表中的条目的数量,并且grouping_type参数是在SampleGroupDescription框中所参考的针对分组的类型的标识符。例如,在特定实施例中,分组类型可以具有表示空间/区块分组的特定值。例如,该值可以是与针对“tile”的ASCII码(0x74696C65)相对应的十六进制值。group_description_index参数是给出描述该组中的样本的样本组条目的索引的整数。索引在1~SampleGroupDescription框中的样本组条目的数量的范围内,或者取值0以表示该样本不是该类型的任何组的成员。最后,grouping_type_parameter参数是针对分组的子类型(在由分组类型使用的情况下)的指示。

[0347] 这样使得能够用信号通知以下内容:在仅区块分组处于使用中(entry_count=1)的情况下,无论动画片断的数量如何,使用最多32字节/区块,来自轨的所有样本针对给定分组类型跟随同一组描述。在自适应网格随时间的经过而移动的情况下,可以在轨片断中定义新的DefaultSampleToGroups框和新的SampleGroupDescription框。新的DefaultSampleToGroups框将替换先前的定义并且参考新的SampleGroupDescription框中

的新区块描述。因而，不是针对各轨片断而是仅在区块网格定义改变的情况下定义SampleToGroup框。

[0348] 此外，根据被配置为处理区块化结构沿着视频序列的变化的特定实施例，利用新的样本映射框以子样本级别进行信号通知。

[0349] 针对与区块化结构有关的样本描述的该实施例适用于一个轨封装型区块化HEVC位流。在解析了复合轨的提取器以将所下载的空间部分保存在符合MPEG-4标准的一个视频轨中之后，可以发现针对区块化HEVC位流的MPEG-4封装的应用或在客户端侧的应用。在涉及将复合轨变换成符合MPEG-4的标准单轨的情况下，另一应用用于不支持提取器的MPEG-4解析器。

[0350] 图12a和图12b示出区块在完整视频中的位置、该区块的大小的信号通知，以及指示适于处理各种区块化结构、可以在没有任何伪影的情况下以子样本级别对该区块进行解码的信号通知。

[0351] 图12a示出客户端装置（例如，视频播放器）所执行的步骤。在第一步骤（步骤1200）中，客户端装置下载初始化数据、或者在文件是本地文件的情况下读取初始化数据、例如符合MPEG-4标准的封装后的位流的初始化数据（通常为moov框的内容）。

[0352] 根据这些初始化数据，客户端装置可以解析编码有区块信息的轨头部信息（步骤1205）。利用该区块信息，用户可以经由客户端装置的图形界面选择与一个或多个区块相对应的关注区域（步骤1210）。

[0353] 利用客户端装置来下载或读取相应的区块轨以及复合轨（步骤1215和1220）。接着，使用区块轨来解析复合轨的提取器以获得一个视频轨（步骤1225）。最后，客户端装置在所获得的视频轨中例如以SampleTableBox构建并添加区块化描述（步骤1230）。

[0354] 在图12b中示出区块化描述的示例。如图所示，区块化描述1250包括动画框“moof”1255和数据框“mdat”1260。“moof”框1255针对各轨包含一个SampleTable框，其中该SampleTable框包含：SampleToGroup框1265，其描述不同的样本组；样本组描述框1270，其描述各样本的NAL单元和区块之间的映射；以及样本组描述框1275，其包含区块描述。样本组框1265表示针对组条目TileSampleMapEntry的“tsgm”分组类型。

[0355] TileNALUMapEntry组条目1270定义样本的NAL单元和区块之间的映射（这是这种实施例提及子样本级别信号通知的原因）。grouping_type参数等于“tsgm”的该框包含针对各样本的NAL单元的数量。

[0356] 可以按照如下所述定义TileNALUMapEntry框（如图12b所示）：

```
class TileNALUMapEntry() extends VisualSampleGroupEntry('tsgm') {
    unsigned int(8) reserved = 0;
    unsigned int(8) entry_count;
[0357]    for (i=1; i<= entry_count; i++)
        unsigned int(32) tileID;
    }
}
```

[0358] 其中，entry_count表示轨样本中的NALU的数量，并且tileID提供针对当前轨所描

述的空间区块的唯一标识符。

[0359] 还可以按照如下所述从大小方面优化TileNALUMapEntry框：

```
class TileNALUMapEntry() extends VisualSampleGroupEntry ('tsgm') {  
    unsigned int(6) reserved = 0;  
    unsigned int(1) large_size;  
    unsigned int(1) mode;  
    if (large_size) {  
        unsigned int(16) entry_count;  
    } else {  
        unsigned int(8) entry_count;  
    }  
    for (i=1; i<= entry_count; i++)  
[0360]     if (mode) {  
        if (large_size) {  
            unsigned int(16) NALU_start_number;  
        } else {  
            unsigned int(8) NALU_start_number;  
        }  
    }  
    unsigned int(16) tileID;  
}
```

[0361] 其中，large_size参数表示轨样本中的NAL单元条目的数量是否是以8或16位表示的，并且mode参数表示是描述每个NAL单元条目（在没有设置mode的情况下）还是仅与tileID相对应的NAL单元条目发生改变（在设置了mode的情况下）。在后者情况下，根据large_size参数的值来将NAL单元数量以16或8位来编码。

[0362] 指定这些NAL单元各自和特定区块之间的映射。如这里以上所述，在样本组描述框1275中设置区块描述，其中逐个描述各区块。

[0363] 要注意，给定示例是一个NAL单元包含针对一个区块的数据并且沿着序列持续时间都如此的特定情况。在区块数据分割在多个NAL单元中的情况下，使用多个聚合器来描述与区块相对应的NAL单元的集合。在NAL单元的数量随时间的经过而发生变化的情况下，可以定义多个tileSampleMap条目，并且样本组可选地针对各片断在轨片断头部中经由分组类型而参考适当的区块映射条目。

[0364] 在如参考图10所述的数据封装方案中，将HEVC位流封装为指向实际包含压缩视频数据的区块轨1020-1~1020-2的复合轨1015。复合轨包含来自于不同的HEVC参数集NAL单元（在图10中表示为PS）的结构数据。复合轨的其它元素主要包括针对各样本并且针对各区

块轨均为一个的提取器的列表,从而(经由初始化片段文件1005的moov框中所包含的轨参考框(“tref”))指向封装在区块轨中的压缩视频数据。

[0365] ISO BMFF标准(该标准的第15部分)中的当前依赖性信号通知机制位于作为初始化片段文件1005的moov框中的轨框的一部分的轨参考框“tref”中。“tref”框提供从包含轨向呈现时的另一轨的参考。包含轨可以参考呈现时的多个其它轨。在当前标准中,利用可以取“scal”或“sbas”这两个值的reference_type参数来指定轨之间的依赖性的类型。“sbas”值代表可分级基础。其表示所参考的轨在可分级呈现时是当前轨的可分级基础轨。“scal”值代表可分级性。其指示代表可分级表示的不同层的轨之间的关系。这意味着包含轨依赖于所参考的轨。

[0366] 在参考图10所述的实施例中,不存在特定的与可分级性相关的依赖性。即使可以考虑到可分级视频,这里也关注复合轨和区块轨之间的空间依赖性。例如,可以利用新的“tile”值来明确表示这些依赖性,如在与复合轨1015(id=1)相对应的初始化片段文件1005的moov框的tref框中所进行的那样。

[0367] 尽管针对各轨的依赖性的例示示例涉及区块(即,从一个复合轨向区块轨的依赖性),但可以同样处理从一个轨向一个或多个轨的其它类型的依赖性。因此,reference_type参数例如可以表示子层依赖性(例如,“subl”值)以表示轨依赖于一个或多个轨。

[0368] 因此,可以将子层轨定义为包含HEVC基本位流的在不会损害其它HEVC NAL单元的解码处理的情况下可丢弃的部分的轨。如上所述,这种定义特别应用于可分级HEVC位流中的时间层并且应用于区块轨。可以在HEVCConfiguration记录中(即,在SampleTableBox中)使用如下的位(或标志)来标记与子层轨相对应的各轨,其中该位(或标志)在被设置为预定值的情况下,表示该HEVC轨是子层轨并且仅包含从其它轨(例如,从复合轨)所参考的NAL单元(即,该HEVC轨不可显示)。在该位或标志的值具有相反值的情况下,这表示该HEVC轨是还包含初始化数据的子层轨(即,该HEVC轨可显示)。例如,可以将所预留的位用在当前HEVCDecoderConfigurationRecord框中。

[0369] 根据参考图12a和图12b所述的特定实施例,各区块轨可作为标准视频轨进行播放。

[0370] 图13示出根据本发明的实施例的将HEVC位流封装为包括可作为标准视频轨进行播放的复合轨和独立区块轨的一组轨的示例。为了进行例示,封装后的位流与图2中示意性示出的视频序列相对应。

[0371] 图13所示的HEVC位流封装与图10所示的HEVC位流封装的主要不同之处在于:各区块轨包括使得能够恢复初始化和配置数据的特定提取器。

[0372] 如图所示,区块轨1300-1~1300-12各自包括指向复合轨1310的表示初始化和配置数据的HEVC参数集NAL单元(表示为PS)上的提取器1305-1~1305-12,其中要重申的是,根据HEVC标准,这些初始化和配置数据通常与HEVC位流的各种参数集相对应。因此,这种初始化和配置数据使各区块轨可作为正常视频轨进行播放。

[0373] 在视频数据样本之前,添加至各区块轨的提取器有利地位于各区块轨的媒体数据框“mdat”的开头。

[0374] 例如必须在关联至区块轨的“tref”框1320-1~1320-12的reference_type参数中(在初始化片段文件1325的动画框“moov”中)用信号通知表示为1315-1~1315-12的从区块

轨(1300-1~1300-12)向复合轨(1310)的这些依赖性。根据本实施例,将包含参数集的轨视为HEVC基本轨“hbas”(这接近SVC的情况,其中包含可分级呈现时的最低工作点的轨被视为“可分级基本轨”“sbas”)。如图所示,依赖于基本轨的轨(即,依赖于具有标识符id=1的复合轨1310的具有标识符id=2~12的区块轨1300-1~1300-12)在它们的轨参考框(1320-1~1320-12)中具有值“hbas”。

[0375] 再次,区块化信号通知可以处于轨级别、样本级别或者轨级别和样本级别。

[0376] 要注意,区块轨默认被视为不可播放。然而,符合MPEG-4标准的高级解析器可以例如通过查看“tref”框在流传输表单文件中检测可显示的区块轨并列出这些区块轨(如果区块轨包含“hbas”类型的参考类型,则该区块轨可被视为可显示)。这意味着,该区块轨即使在处理程序框中被标记有“tile”值,也可被视为标准视频轨。在区块化信号通知是基于样本的情况下,由于将区块化信息放置在已知为SampleTableBox的框中,因此将区块轨或子层轨在它们的处理程序框中加标签为“vide”。

[0377] 图14示出根据本发明的另一实施例的将HEVC位流封装为包括复合轨、初始化数据轨和可作为标准视频轨进行播放的独立区块轨的一组轨的示例。为了进行例示,封装后的位流与图2中示意性示出的视频序列相对应。

[0378] 图14所示的HEVC位流封装与图13所示的HEVC位流封装的不同之处在于:将初始化数据放置在专用初始化数据轨1400中(而不是放置在复合轨1310中)。

[0379] 这种实施例相比参考图13所述的实施例而提供的优点其中之一涉及在要独立播放区块轨的情况下要传输的数据量。由于在专用轨中传输初始化数据,因此没有请求传输复合轨。

[0380] 需要重申的是,根据HEVC文件格式的当前规格,存在用以以文件格式(仅在已知为Sample Entry的框或者在Sample Entry框并且在数据样本中)传送参数集(PS)的两个可能性。利用已知为Sample Table的框中的“hvc1”框和“hev1”框来分别用信号通知这两个配置。在将参数存储在样本中更加复杂的情况下,允许在参数集更新的情况下的更多动态。因此,在优选实施例中,(利用Sample Table框中的HEVC Sample Entries参数中的“hev1”值)在Sample Entry框并且在数据样本中传送参数集,从而特别是针对区块化配置变化来处理图片参数集(PPS)。

[0381] 因此,专用初始化数据轨1400仅包含如同分别与视频参数集、序列参数集或图片参数集相对应的类型等于32、33或34那样的非VCL HEVC NAL单元作为数据。

[0382] 如图14所示,位于区块轨1410-1~1410-12的媒体数据框“mdat”的开头的提取器1415-1~1415-12指向专用初始化数据轨1400的数据。同样,复合轨1405的第一个提取器(1420)指向专用初始化数据轨1400的数据。因此,初始化数据轨1400是没有参考任何其它轨的封装后的HEVC位流的唯一轨。如此,由于不存在关联至初始化数据轨1400(id=2)的tref框中所示的依赖性(在tref框中无“hbas”依赖性),因此将后者视为不可独立显示。

[0383] 在视频位流中修改一些初始化数据的情况下(即,在HEVC位流中发生图片参数集的情况下),在发生变化的时间位置处,如附图标记1425所示,将这些初始化数据放置在样本数据中。将附图标记为1430和1435-1~1435-12的相应提取器分别插入在复合轨1405以及区块轨1410-1~1410-12各自中、即参考这些新的PPS的各区块轨中。

[0384] 在封装后的HEVC位流的各轨中,按时间顺序组织样本(以及关联的NALU)。同样,在

专用初始化数据轨1400中按时间顺序组织图片参数集。“trun”框(图14中未示出)使得能够针对各样本提供正确的解码时间。

[0385] 当然,为了满足本地和特定要求,本领域技术人员可以向上述解决方案应用许多修改和改变,然而所有这些修改和改变均包括在如所附权利要求书所定义的本发明的保护范围内。

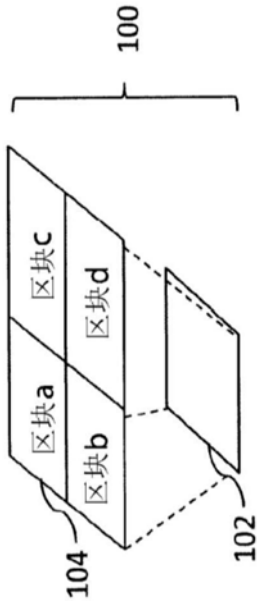


图1a

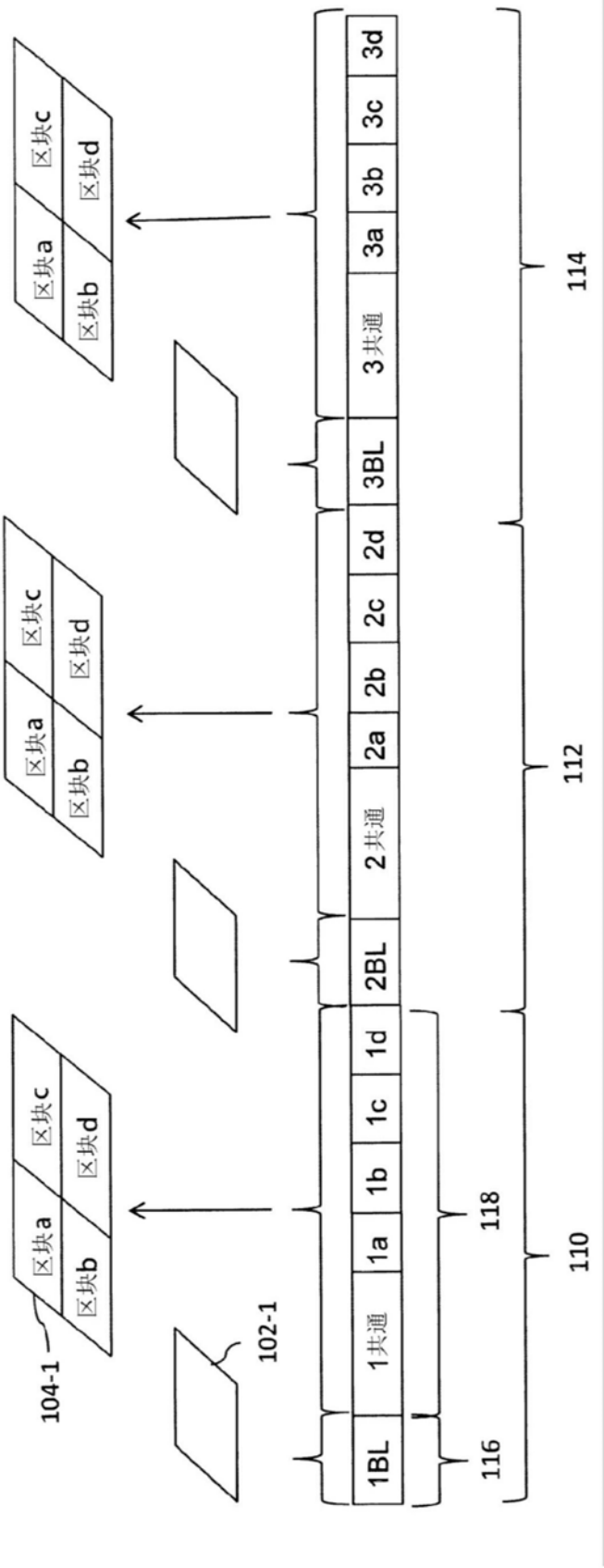


图1b

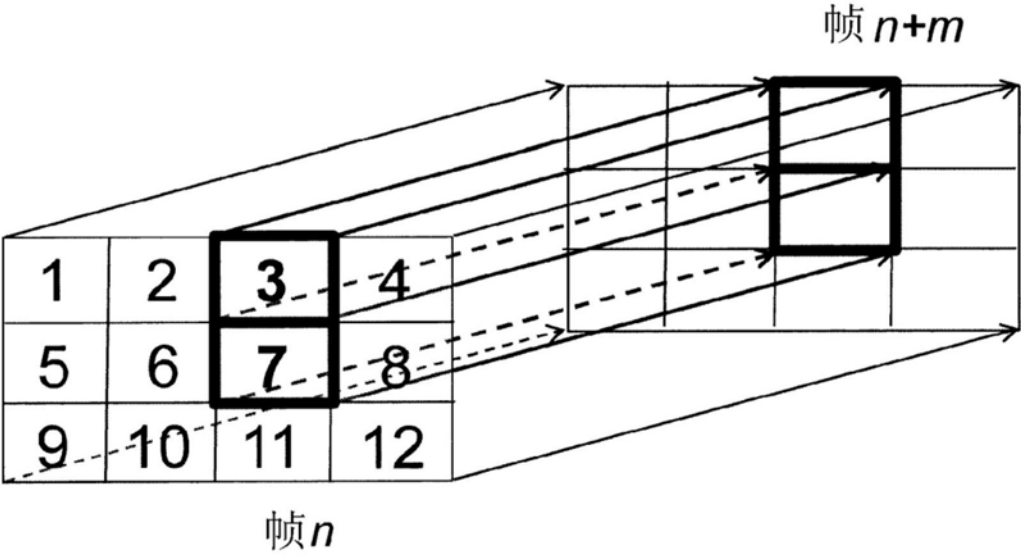


图2

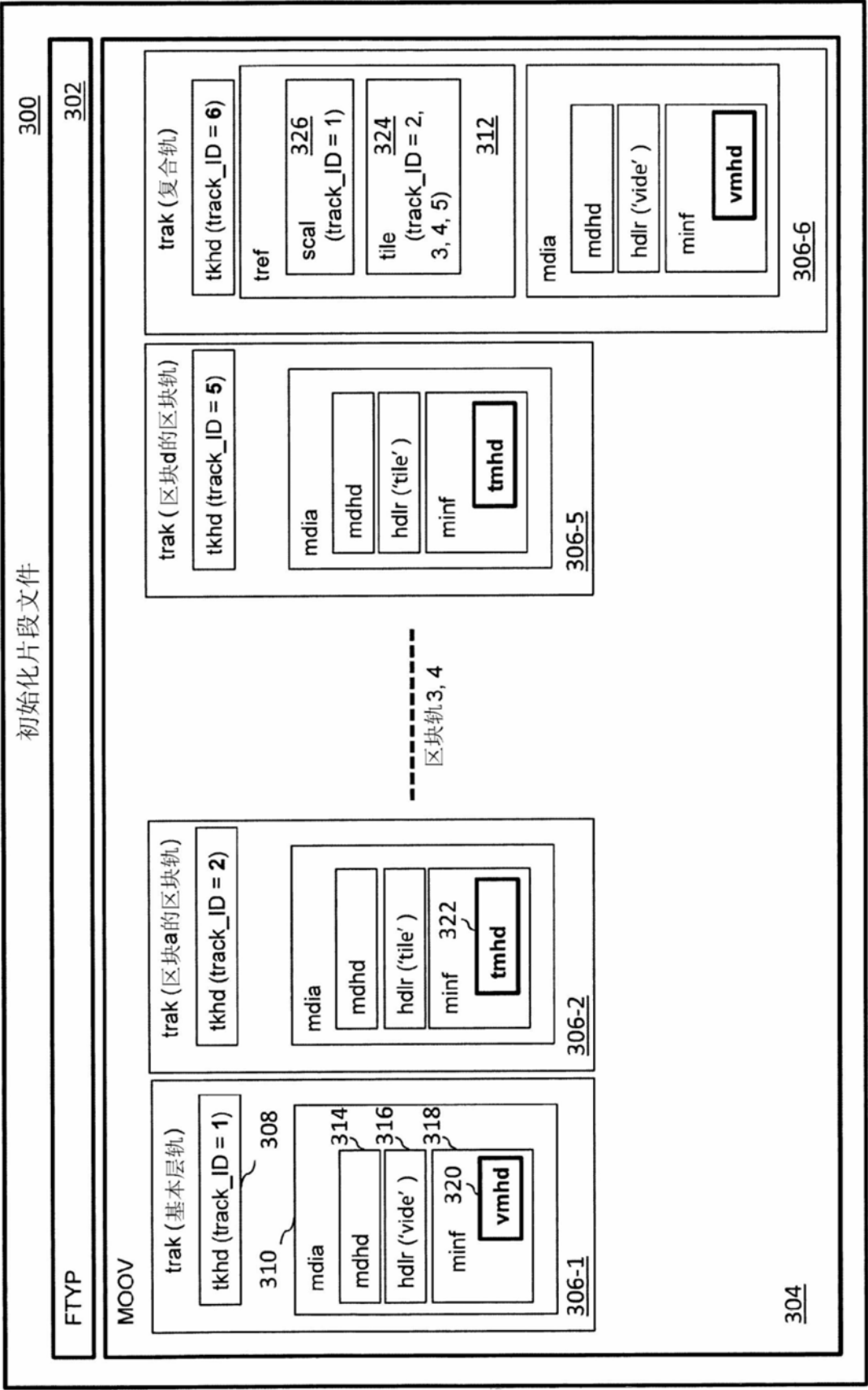


图3

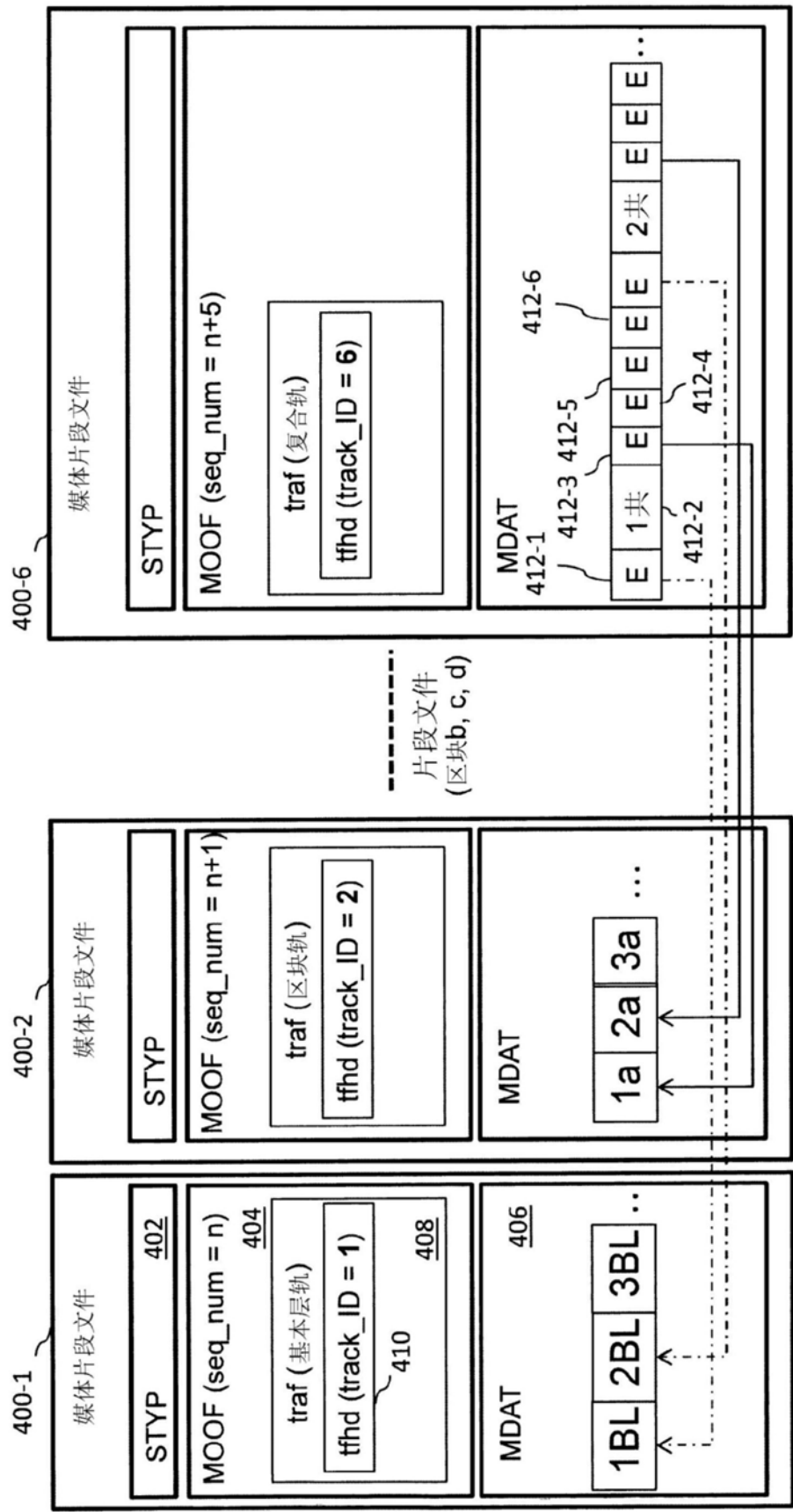


图4

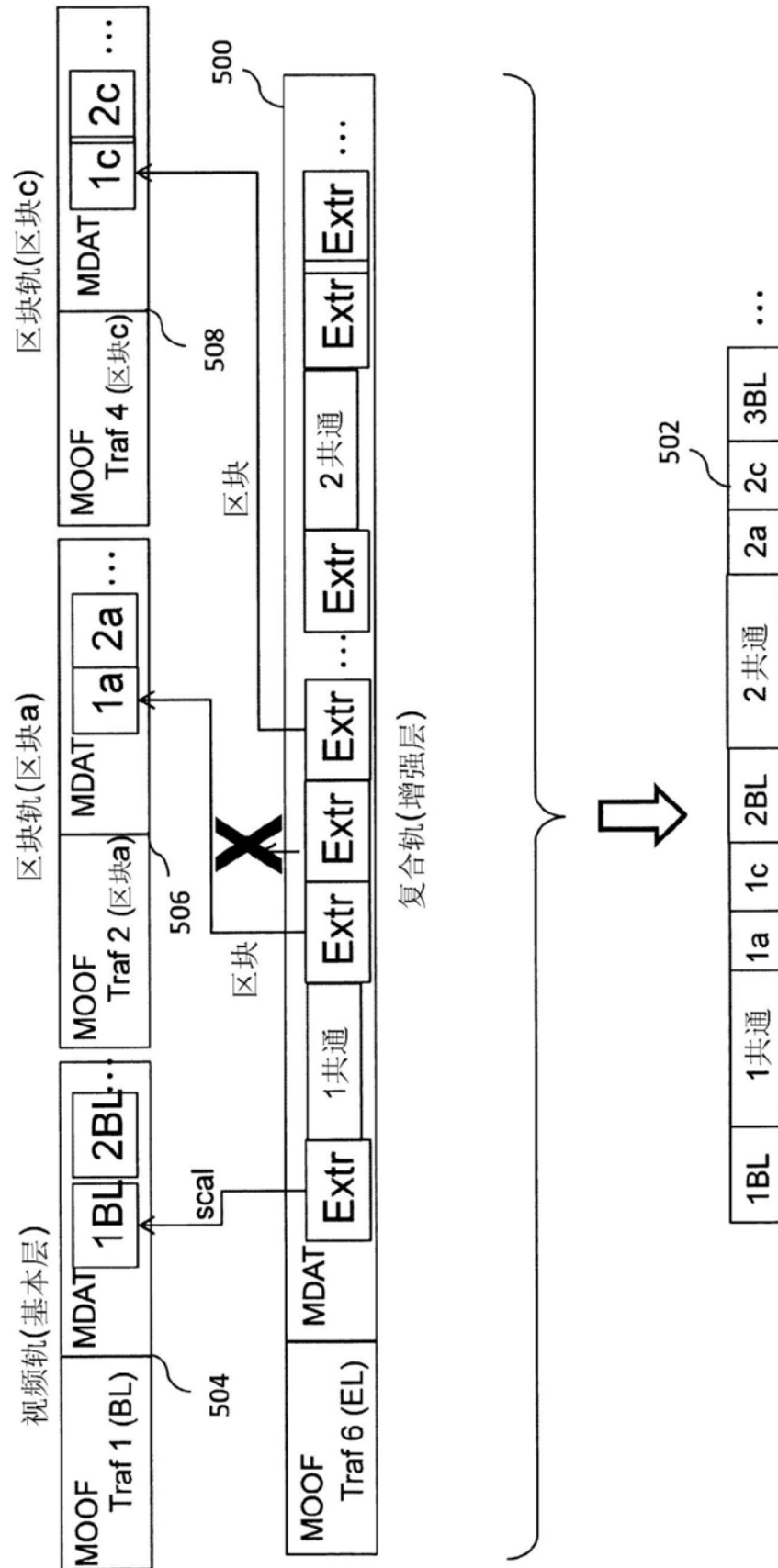


图5

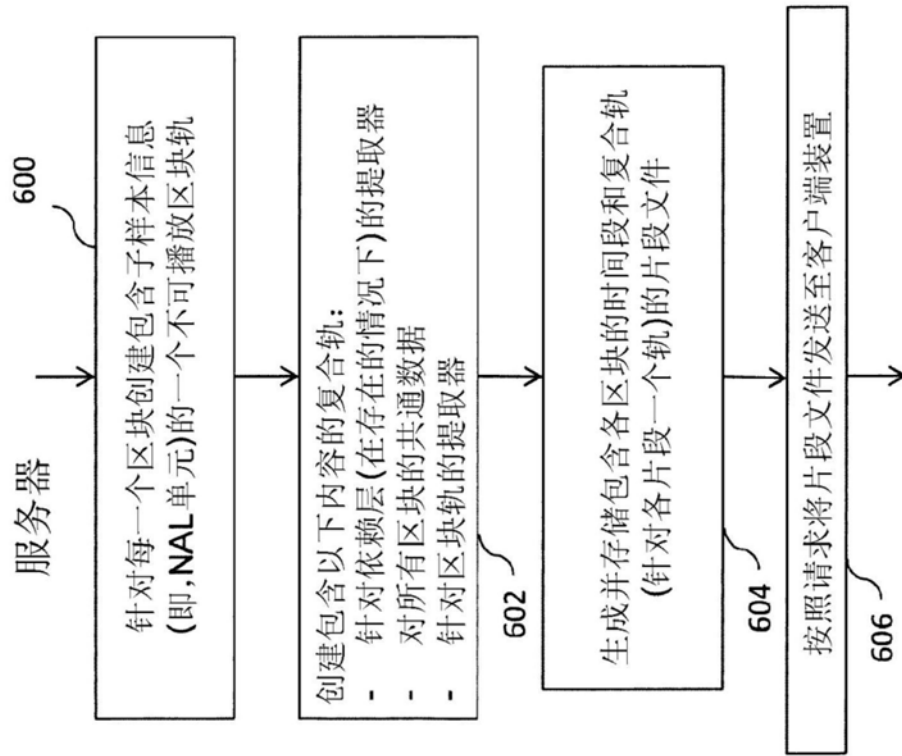


图6a

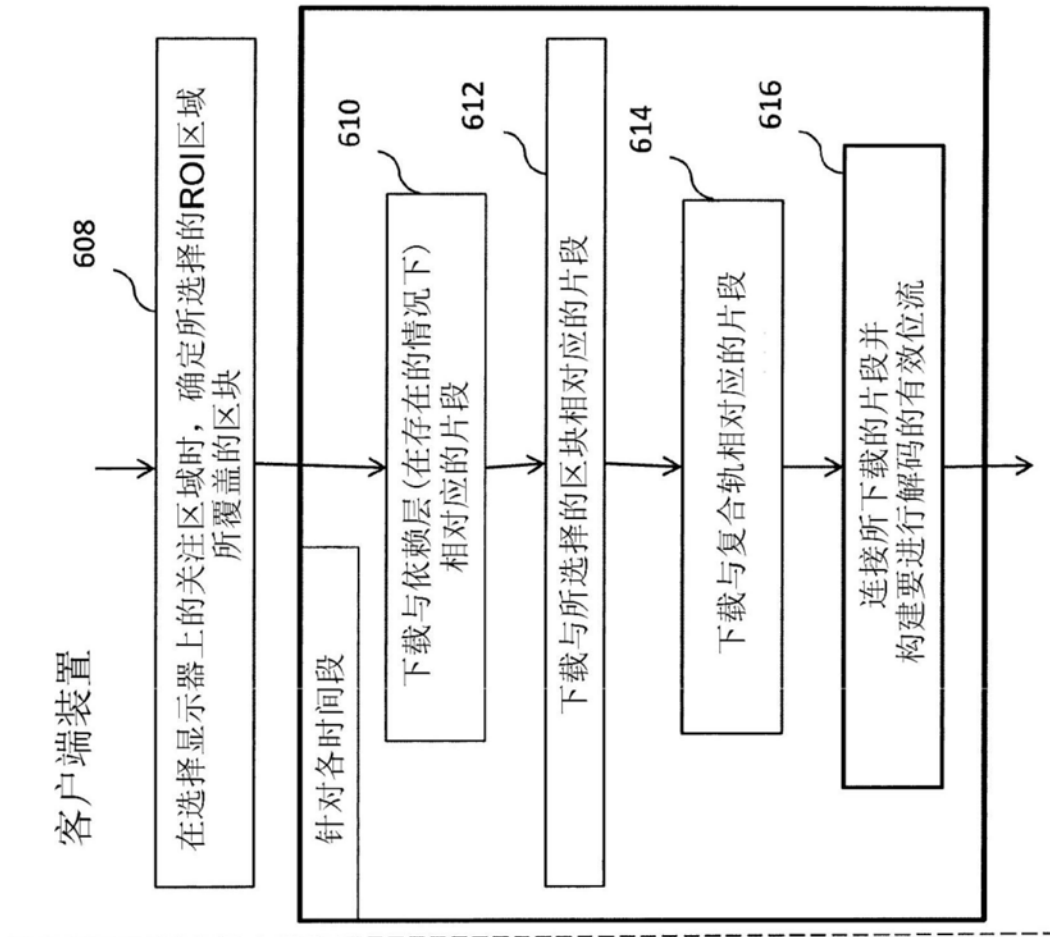


图6b

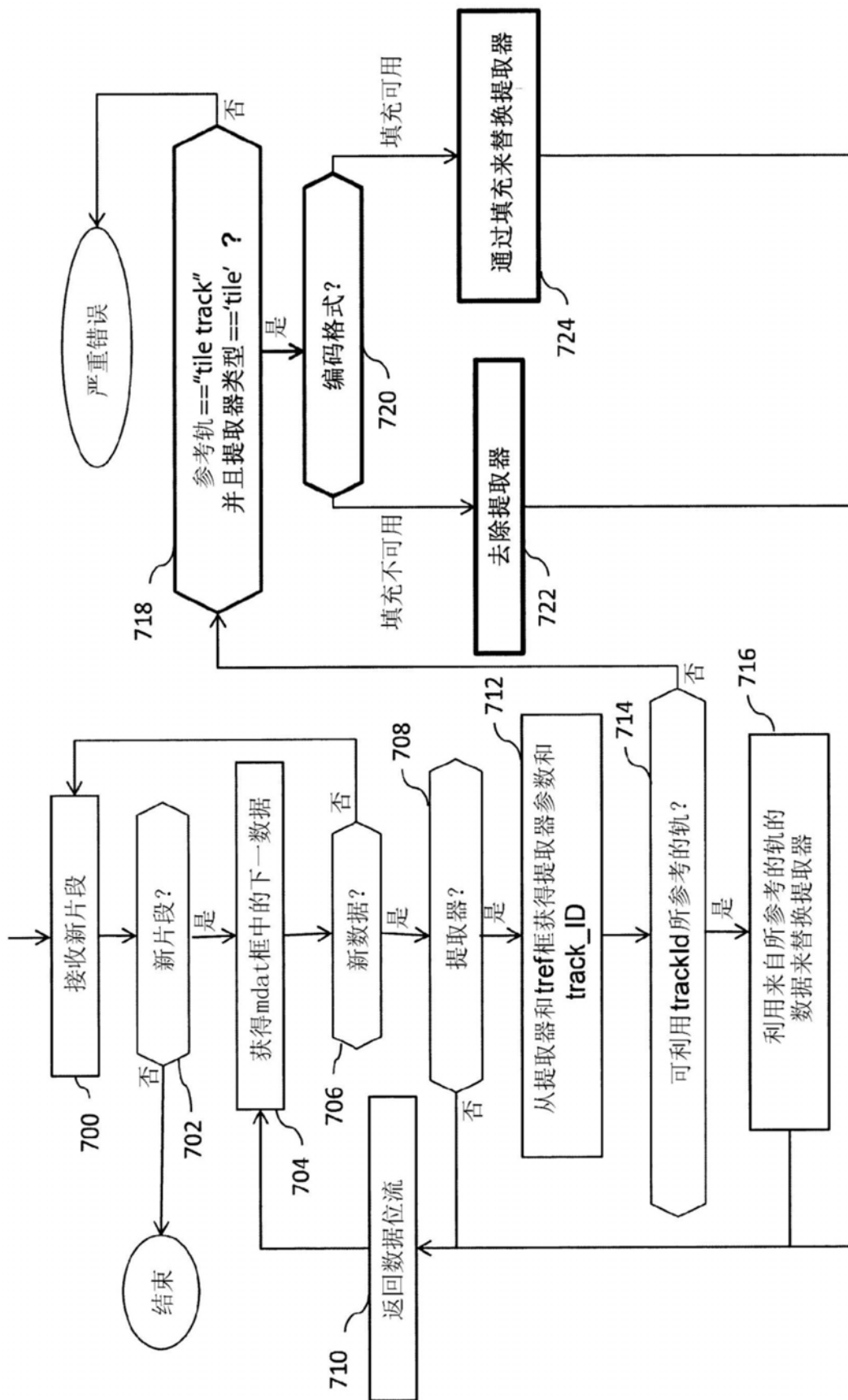


图7

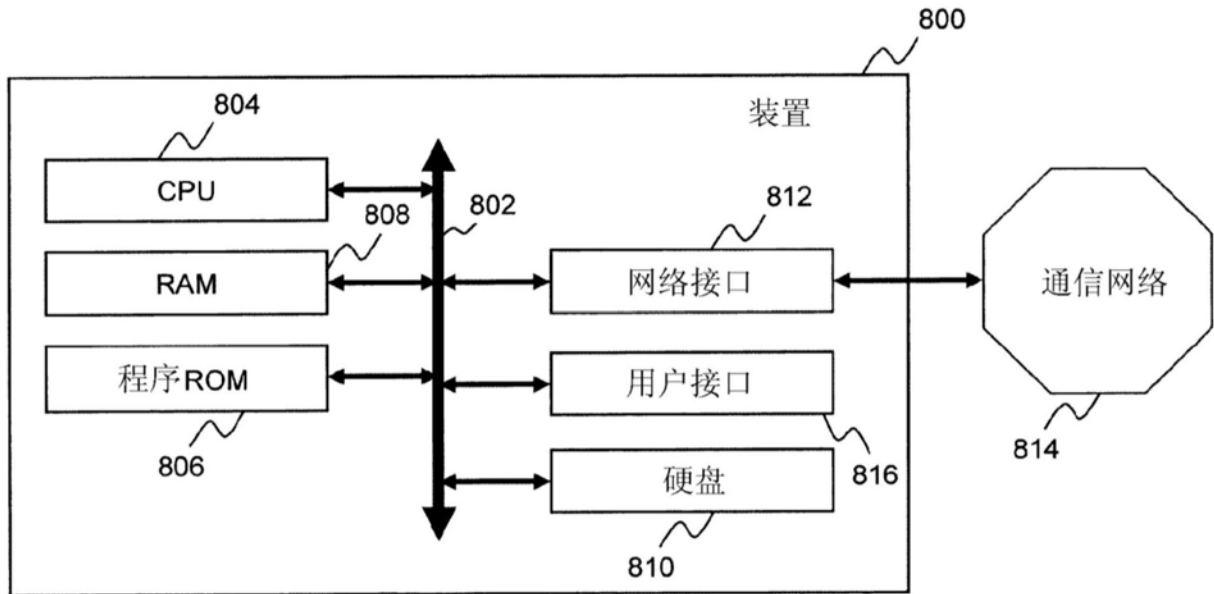


图8

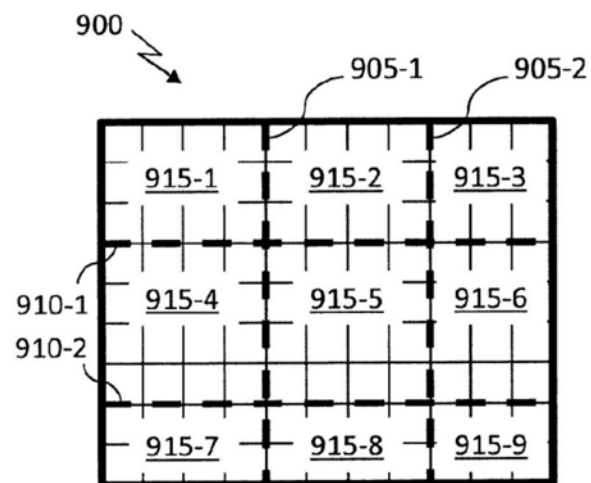


图9a

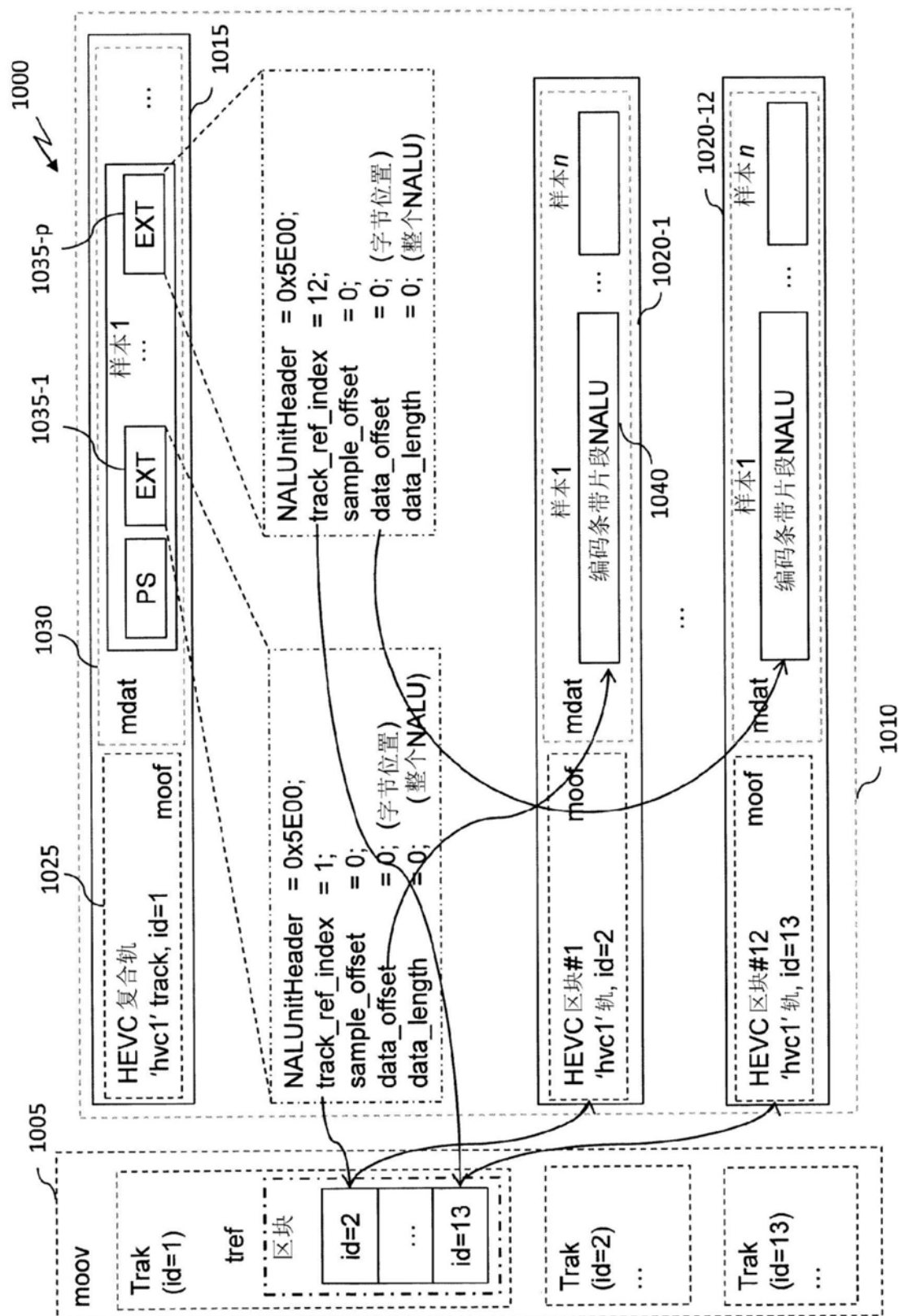


图10

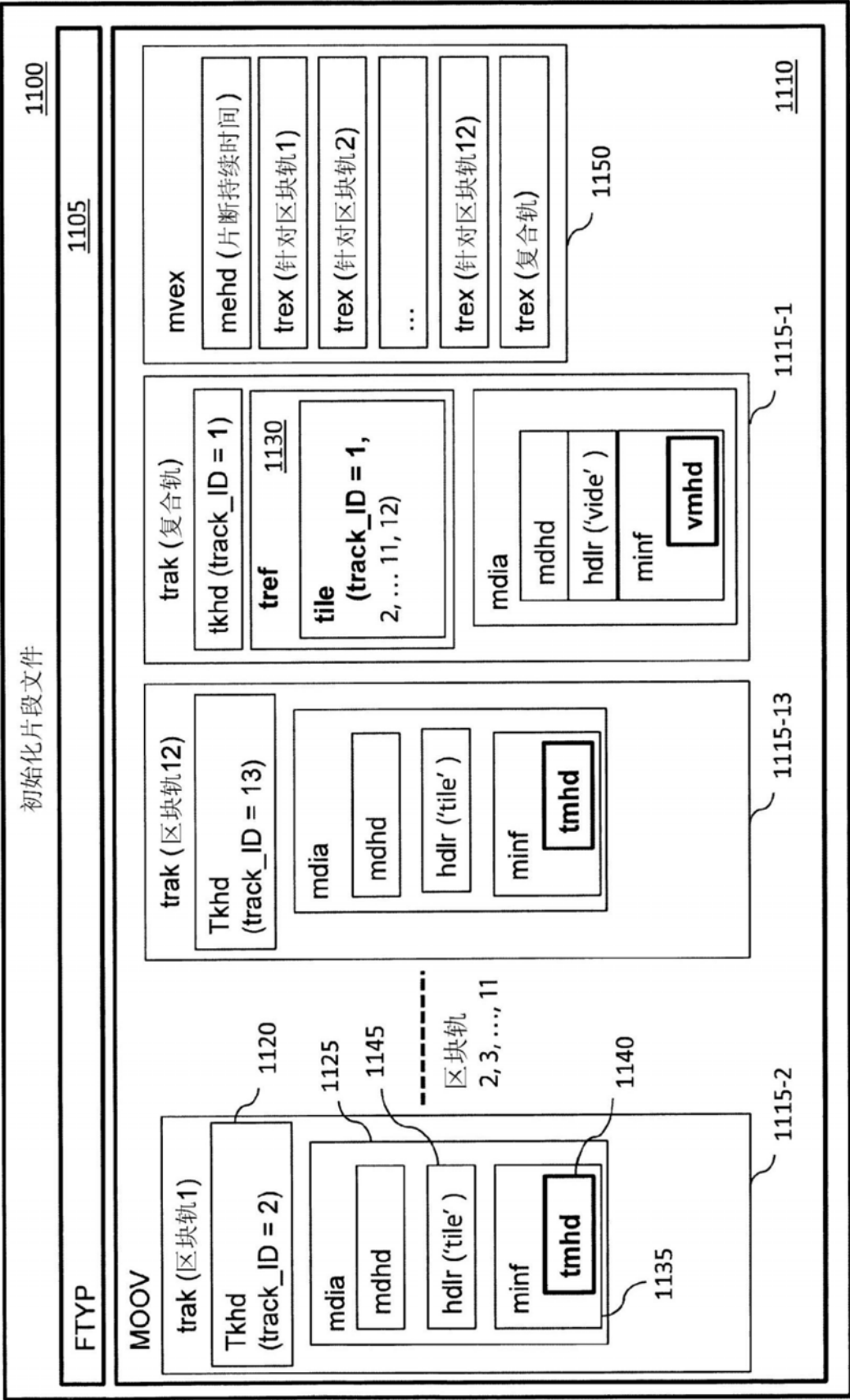


图11

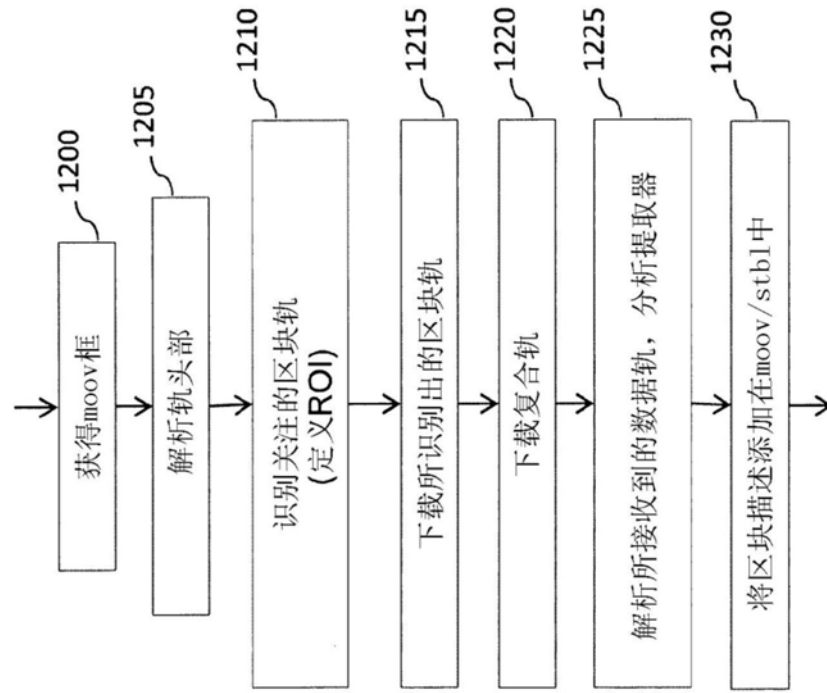


图12a

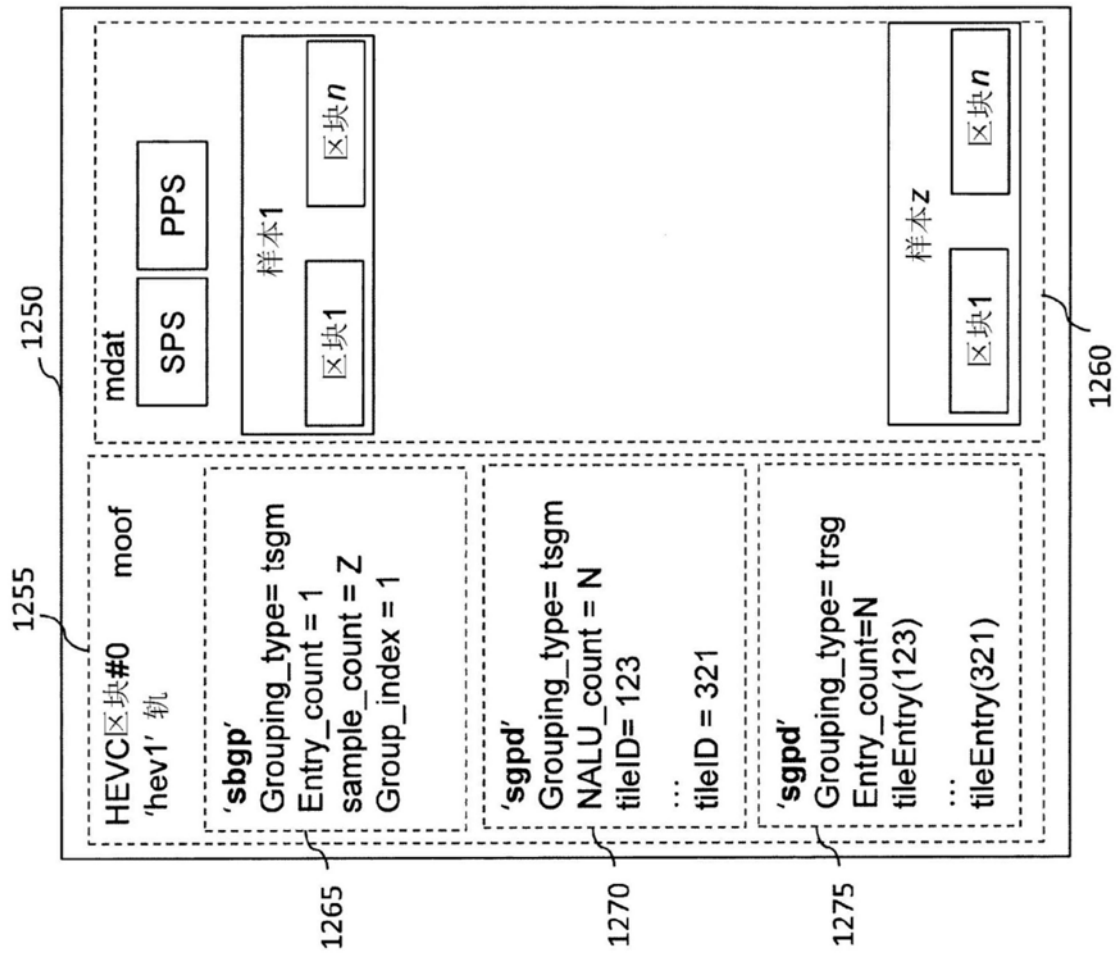


图126b

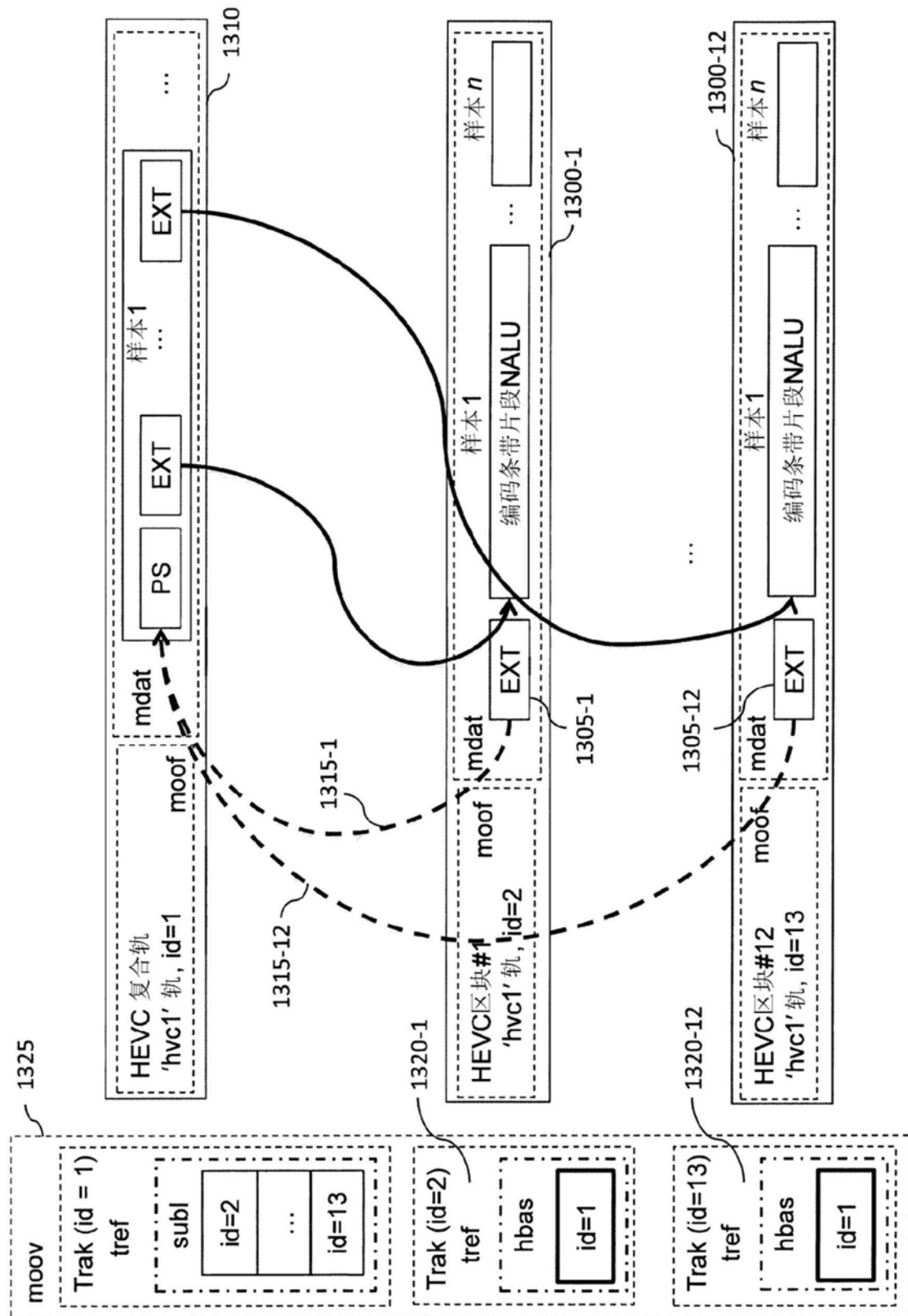


图13

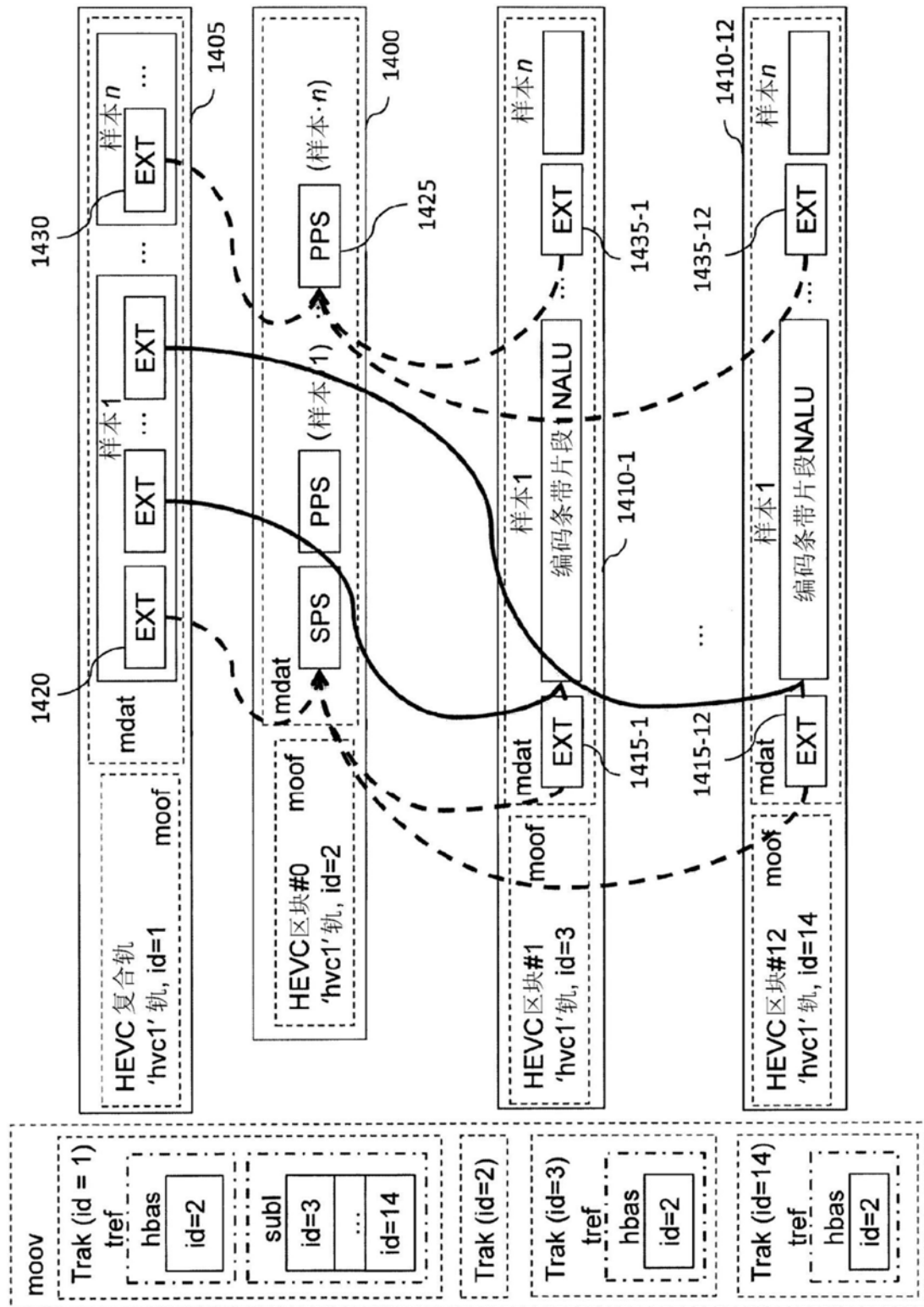


图14