



(12) 发明专利申请

(10) 申请公布号 CN 116458163 A

(43) 申请公布日 2023. 07. 18

(21) 申请号 202180075086.2

(22) 申请日 2021.10.15

(30) 优先权数据

2016457.0 2020.10.16 GB

(85) PCT国际申请进入国家阶段日

2023.05.06

(86) PCT国际申请的申请数据

PCT/GB2021/052685 2021.10.15

(87) PCT国际申请的公布数据

W02022/079450 EN 2022.04.21

(71) 申请人 威诺瓦国际有限公司

地址 英国伦敦

(72) 发明人 圭多·梅迪 L·温尼尔

(74) 专利代理机构 北京聿华联合知识产权代理

有限公司 11611

专利代理师 刘华联

(51) Int.Cl.

H04N 21/44 (2006.01)

权利要求书3页 说明书29页 附图15页

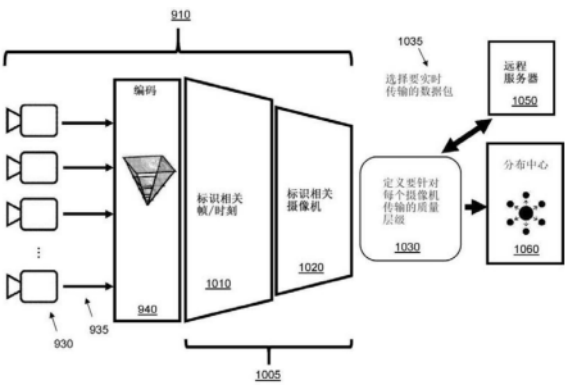
(54) 发明名称

多层信号编码的分布式分析

(57) 摘要

本公开涉及一种分析多个视频摄像机源的方法,所述方法包括:在第一位置使用基于层的编码对所述多个视频摄像机源进行编码,包括生成针对在所述基于层的编码内的多个层中的每个层的编码数据流,其中在所述多个层中的不同层对应于不同的空间分辨率,较高层表示较高空间分辨率;向相对于所述第一位置远程的第二位置传输针对所述多个视频摄像机源的一个或多个最低层的编码数据流;在所述第二位置对编码数据流进行解码来生成所述多个视频摄像机源的处于第一空间分辨率的一组重构;将一个或多个视频分析功能应用于所述一组重构来标识用于进一步分析的一个或多个视频摄像机源;向针对所标识的用于进一步分析的一个或多个视频摄像机源的所述第一位置发送针对在所述一个或多个最低层之上的一个或多个层的进一步编码数据流的请求;响应于所述请求,向所述第二位置传输针对在所述一个或多个最低层之上

的一个或多个层的所述进一步编码数据流;在第二位置处,对所述进一步编码数据流进行解码来生成针对所标识的一个或多个视频摄像机源的处于第二空间分辨率的一组重构;并将一个或多个视频分析功能应用于处于所述第二空间分辨率的所述一组重构。



1. 一种分析多个视频摄像机源的方法,所述方法包括:

在第一位置处,使用基于层的编码对所述多个视频摄像机源进行编码,包括生成针对在所述基于层的编码内的多个层中的每个层的编码数据流,其中在所述多个层中的不同层对应于不同的空间分辨率,较高层表示较高空间分辨率;

向相对于所述第一位置远程的第二位置传输针对所述多个视频摄像机源的一个或多个最低层的编码数据流;

在所述第二位置处,对所述编码数据流进行解码来生成所述多个视频摄像机源的处于第一空间分辨率的一组重构;

将一个或多个视频分析功能应用于所述一组重构来标识用于进一步分析的一个或多个视频摄像机源;

向针对所标识的用于进一步分析的一个或多个视频摄像机源的所述第一位置发送针对在所述一个或多个最低层之上的一个或多个层的进一步编码数据流的请求;

响应于所述请求,向所述第二位置传输针对在所述一个或多个最低层之上的一个或多个层的所述进一步编码数据流;

在所述第二位置处,对所述进一步编码数据流进行解码来生成针对所标识的一个或多个视频摄像机源的处于第二空间分辨率的一组重构;并且

将一个或多个视频分析功能应用于处于所述第二空间分辨率的所述一组重构。

2. 一种远程信号分析方法,其包括:

在计算设备处,通过至少一个网络接收针对对应于多个输入信号的多个编码信号流的编码数据,

其中在相对于所述计算设备远程的编码位置处执行对所述多个输入信号的编码,

其中使用基于层的编码对所述编码信号流进行编码,

其中在所述基于层的编码中的不同层对应于不同的质量层级,较高层表示较高质量层级,并且

其中所接收到的编码数据包括与在所述基于层的编码内的一个或多个最低层相关联的编码数据;

对所接收到的编码数据进行解码来获得处于所述一个或多个最低层的所述多个输入信号的多个重构;

处理所述多个输入信号的重构中的至少一个重构来确定用于进一步处理的所述编码信号流的部分;并且

从所述编码位置请求针对在所述一个或多个最低层之上的与所述编码信号流的所确定的部分相关联的一个或多个层的编码数据。

3. 一种编码方法,其包括:

在第一位置处,生成针对对应于多个输入信号的多个编码信号流的编码数据,

其中使用基于层的编码对所述编码信号流进行编码,

其中在所述基于层的编码中的不同层对应于不同的质量层级,较高层表示较高质量层级,并且

其中所接收到的编码数据包括与在所述基于层的编码内的一个或多个最低层相关联的编码数据;

向相对于所述第一位置远程的第二位置传输所述编码数据用于解码；

接收对针对在所述一个或多个最低层之上的一个或多个层的进一步编码数据的请求，所述请求与所述多个输入信号中的一个或多个输入信号相关；并且

向所述第二位置传输所述进一步编码数据。

4. 根据权利要求2或权利要求3所述的方法，其中不同的质量层级对应于不同的空间分辨率，较高层表示较高空间分辨率。

5. 根据权利要求1或权利要求4所述的方法，其中所述基于层的编码包括处于最低空间分辨率的基础层和在所述基础层之上的处于较高空间分辨率的多个残差层，在所述多个残差层中的给定残差层是基于从处于针对所述层的分辨率的输入信号导出的数据与从较低层导出的数据之间的比较来确定的。

6. 根据前述权利要求中任一项所述的方法，其包括：

确定在处于所述一个或多个最低层的所述多个输入信号的重构中的至少一个重构中的感兴趣区域；并且

请求与所述感兴趣区域相关的进一步编码数据。

7. 根据权利要求6所述的方法，其进一步包括：

使用所述进一步编码数据来生成针对处于所述一个或多个最低层之上的一个或多个层的所述感兴趣区域的一组重构。

8. 根据权利要求6或权利要求7所述的方法，其包括：

确定生成处于最高质量层级的重构所需的差分编码数据，所述差分编码数据不包括所接收到的编码数据和所接收到的进一步编码数据；并且

从所述第一位置请求所述差分编码数据。

9. 根据前述权利要求中任一项所述的方法，其包括：

在处于所述一个或多个最低层的所述多个输入信号的重构中的至少一个重构内执行对象检测。

10. 根据前述权利要求中任一项所述的方法，其包括：

在处于所述一个或多个最低层之上的一个或多个层的所述多个输入信号的重构中的至少一个重构内执行对象检测。

11. 根据前述权利要求中任一项所述的方法，所述方法包括以下项中的一项或多项：

标识针对所述多个输入信号中的一个或多个输入信号的一组或多组帧用于进一步分析，以及

标识一个或多个输入源用于进一步分析；

其中在第一编码位置和第二分析位置中的一者或多者处执行所述标识。

12. 根据权利要求11所述的方法，其包括：

传输针对在所述一个或多个最低层之上的层的编码数据用于所标识的帧或所标识的输入源。

13. 根据前述权利要求中任一项所述的方法，其包括：

将编码数据传输到远程服务器用于信号分析；并且

将编码数据传输到多个用户。

14. 根据权利要求13所述的方法，其中将编码数据传输到多个用户包括将编码数据传

输到分布中心用于分布给所述多个用户。

15. 根据权利要求14所述的方法,其包括:

在所述远程服务器处对一组重构进行分析来确定将传输到所述分布中心的用于一个或多个输入源的编码数据的一组层。

16. 根据权利要求15所述的方法,其中所述多个输入信号或多个视频输入源与用于捕获现场事件的多个视频摄像机相关联,其中所述远程服务器确定将选择所述多个视频摄像机中的哪一个视频摄像机用于传输到所述多个用户,并且确定针对将传输到所述多个用户的所述编码数据的质量层级用于所选择的视频摄像机。

17. 根据权利要求2或权利要求3所述的方法,其中每个质量层级具有不同的编码位率。

18. 根据前述权利要求中任一项所述的方法,其中从所述第一编码位置传输的所述编码数据包括用于编码信号分量的子集的编码数据。

19. 根据权利要求18所述的方法,其中所述编码信号分量包括所述视频信号的不同颜色分量,并且编码分量的所述子集至少包括明度或亮度分量。

20. 根据前述权利要求中任一项所述的方法,其包括:

在所述第二远程位置处接收针对在所述一个或多个最低层之上的一个或多个层的编码数据用于存档。

21. 根据前述权利要求中任一项所述的方法,其包括:

对所接收到的编码数据进行解码来获得处于所述一个或多个最低层的所述多个输入信号的多个重构;

放大所述多个重构;

对所放大的重构进行分析。

22. 根据权利要求21所述的方法,其中所述放大使用在训练集上训练的神经网络架构,所述训练集包括在处于两个空间分辨率的图像上的版本,以便学习参数值来将处于较低分辨率的重构映射到处于较高分辨率的重构。

23. 根据权利要求22所述的方法,其中响应于针对处于所述较高分辨率以及所述较高分辨率之上的一个或多个层的编码数据是不可用的而执行所述放大。

24. 根据前述权利要求中任一项所述的方法,其中多层编码包括基于SMPTE VC-62117的编码。

25. 根据权利要求1至24中任一项所述的方法,其中所述多层编码包括基于ISO/IEC MPEG-5第2部分LCEVC的编码。

26. 一种实施当从属于权利要求3时的方法权利要求4至25中任一项的编码系统。

27. 一种实施当从属于权利要求2时的方法权利要求4至25中任一项的远程分析系统。

多层信号编码的分布式分析

技术领域

[0001] 本公开涉及用于执行数字信号的自动分析的系统,例如用于执行图像序列中的对象分析和/或场景分类的人工智能(“AI”)系统。此外,本公开涉及用于操作上述系统以对在分析与执行数据分析的系统的位置不同的位置捕获、编码和/或存档的数据执行分析的方法。此外,本公开涉及一种计算机程序产品,其包括上面存储有计算机可读指令的非暂时性计算机可读存储介质,这些计算机可读指令可由包括处理硬件的计算机化设备执行以执行上述方法。本文描述的某些实例采用自动深度学习分析来执行复杂信号(诸如,作为非限制性实例,音频、图像、视频和沉浸式媒体信号)的场景和/或对象检测、分类、识别和标记。为简单起见,本公开主要针对视频信号的情况,但是本领域的技术人员能够容易地推导出相同的概念如何可以应用于除视频信号之外的其他类型的信号的自动分析。

背景技术

[0002] 应当理解,时间演变信号诸如视频信号典型地通过时间上的样本序列(例如,速率为60Hz或每秒60个样本的时间上的样本序列)来表示,其中每个样本(例如,视频帧)包含多个颜色平面(例如RGB、YUV、HSV等),其中每个颜色平面包含大量像素(例如,3840x2160像素、或大约800万像素用于超高清全分辨率平面),并且其中每个像素值用给定位深度(例如,8位、10位等)的值表示。总而言之,应当理解,未压缩的视频信号需要每秒极高的位数才能从一个给定位置传送到另一个位置,诸如12千兆位(即1.5GB)仅用于4:2:2 UHD 10位视频信号的一秒。

[0003] 虽然人工智能和神经网络算法的力量令人兴奋,但信号分析仍然存在基本的分布问题。许多组织拥有大型数据中心,其中包含许多图形或张量处理单元,能够使用复杂的神经网络架构执行推理。然而,这些复杂的神经网络架构被配置成接收和处理高分辨率信号,诸如高分辨率图像或视频源。遇到的问题之一是如何有效地将构成分析输入的高分辨率图像或视频源提供给这些数据中心;特别是在这些源的来源在地理上偏远的情况下。例如,体育赛事或音乐事件可能有许多摄像机以高分辨率捕获视频数据;然而,需要与任何数据中心的高速连接来处理混合此视频数据。为此,需要在现场执行复杂的视频分析或混合;然而,用于高级神经网络分析的处理资源可能无法移植到赛事或事件场地(因为其可能包括足球场大小的高级冷却处理设备)。类似的问题适用于来自远程事件和/或边缘传感器网络的高质量音频信号。

[0004] 更详细地说,每当待分析的信号位于与意图分析该信号的计算资源的位置不同的位置时,传统方法对图像帧执行AI分析的一个显著局限性就会随之而来。此情况对于视频信号来说非常频繁,因为视频信号通常是通过轻型摄像机和/或边缘设备“在现场”捕获的,而复杂的AI分析需要在数据中心(其中很多时候由于数据来自多个位置而定期重新训练神经网络算法)中通常可用的处理资源。为了应对此挑战,最常见的方法是将AI处理能力转移到边缘设备(由于处理资源的限制以及仅能获得本地数据用于神经网络训练和机器学习的目的,这会产生成本低效并且牺牲分析的价值)或以某种方式将待分析的所有视频数据从

边缘传输到数据中心。

[0005] 已经讨论的解决方案侧重于传输用于人工智能分析的特征。然而,通常没有关于这些“特征”应该是什么的指导。典型地,建议的解决方案具有某种形式的现场特定本地特征提取,并且将此数据传输到具有用于更复杂的分析的处理资源的数据中心。然而,此方法是不灵活的,例如很难改变特征提取。通常需要本地实施神经网络架构的一部分,这会产生另外的问题。此外,在分析之前很难知道哪些特征可能对分析有用。

发明内容

[0006] 本发明的各个方面和变化在所附权利要求中阐述。

[0007] 在下面的详细描述中进一步阐述了某些未要求保护的方面。

附图说明

[0008] 图1示出了多层或分级编码和解码过程的高层级示意图;

[0009] 图2示出了多层解构过程的高层级示意图;

[0010] 图3示出了多层解构过程的替代性高层级示意图;

[0011] 图4示出了适用于对分层输出的残差进行编码的编码过程的高层级示意图;

[0012] 图5示出了适用于对来自图4的每个输出层级进行解码的多层解码过程的高层级示意图;

[0013] 图6示出了多层编码技术的编码过程的高层级示意图;

[0014] 图7示出了适用于对图6的输出进行解码的解码过程的高层级示意图;

[0015] 图8示出了多层编码格式的高层级示意图;

[0016] 图9示出了用于远程分析本地捕获的信号的第一示例性过程的高层级示意图;

[0017] 图10示出了用于远程分析本地捕获的信号的第二示例性过程的高层级示意图;

[0018] 图11示出了用于远程分析本地捕获的信号的第二示例性过程的高层级示意图;

[0019] 图12示出了用于实施本文描述的实例的示例性装置;

[0020] 图13A至图13D示出了用于一个示例性摄像机的对视频数据进行分层传输的示例性过程;并且

[0021] 图14示出了指示与比较操作相比的改进的表。

具体实施方式

[0022] 本文描述的某些实例涉及用于编码信号的方法。处理数据可包含但不限于获取、导出、输出、接收和重构数据。

[0023] 本文描述的某些实例涉及信号的分布式处理。具体地,实例涉及使用多层编码(也称为基于分层的和/或分级编码)对信号进行编码,其中不同的层表示不同的质量层级。这些质量层级可以包括不同的空间分辨率、不同的传输位率、不同的量化层级、不同的采样层级和/或不同的位深度。

[0024] 在基于分层的编码格式中,诸如ISO/IEC MPEG-5第2部分LCEVC(以下简称“LCEVC”)或SMPTE VC-6 2117(以下简称“VC-6”),信号被分解成多个数据“梯队”(也称为“分级分层”),每个梯队对应于信号的“质量层级”(“LoQ”),从原始信号的采样率的最高梯

队到通常具有比原始信号更低的采样率的最低梯队。在非限制性实例中,当信号是视频流的帧时,最低梯队可以是原始帧的缩略图,或甚至仅是单个像素。其它梯队含有有关应用于重构再现以便产生最终输出的校正的信息。梯队可以基于残差信息,例如特定质量层级的原始信号的版本和相同质量层级的信号的重构版本之间的差异。最低梯队可以不包含残差信息,但可以包含原始信号的最低采样。给定质量层级的解码信号通过以下方式进行重构:首先对最低梯队(因此以第一最低质量层级重构该信号)进行解码,然后预测第二质量层级——下一较高质量层级信号的再现,然后对相应的第二梯队重构数据(也称为第二质量层级“残差数据”)进行解码,然后将预测数据与重构数据组合,以便重构第二较高质量层级信号的再现,等等,直到重构给定质量层级。重构该信号可以包含对残差数据进行解码,并使用该残差数据来校正特定质量层级的版本,该版本导出自较低质量层级的信号版本。不同梯队的队数据可以使用不同的编码格式进行编码,并且不同的质量层级可以具有不同的采样率(例如,针对图像或视频信号的情况的分辨率)。后续梯队可以指信号的相同信号分辨率(即,采样率),或指越来越高的信号分辨率。

[0025] 在优选的实例中,多层编码方案包括视频编码标准VC-6:SMPTE VC-6 ST-2117,如PCT/GB2018/053552(和相关联的已公布的标准文件)中所述的,所有这些文件都通过引用并入本文。某些实例还可以使用其他多层编码方案,诸如LCEVC:MPEG-5第2部分LCEVC(“低复杂度增强视频编码”),如PCT/GB2020/050695(和相关联的已公布的标准文件)中所述的,这些文件也通过引用并入本文。然而,本文所示的概念不必限于这些特定的多层编码方案。在其他实例中,多层编码方案可以包括(可缩放的)基于小波的编码方案。

[0026] 在某些实例中,在网络边缘的信号源(诸如视频摄像机和/或麦克风)处输出一组全质量信号,然后使用多层编码格式(诸如VC-6)在边缘对该组全质量信号进行编码。通过至少一个网络将与在多层编码分级或锥体中的较低层级相关的编码数据流传输到远程计算设备进行分析。此远程计算设备可以包括所谓的“云”服务器或数据处理中心。对在多层编码分级或锥体中的针对较低层级的编码数据进行解码并且将其提供给远程计算设备,该远程计算设备然后能够对从针对较低层级的编码数据获得的重构执行第一层级分析。例如,这些重构可以是低分辨率的“缩略图”帧或帧组。远程计算设备然后可以基于第一层级分析的输出对信号进行进一步分析。此操作可以包括请求表示在多层编码分级或锥体中的较高层级的进一步编码数据。从网络边缘接收到此编码数据后,远程计算设备可以能够解码进一步编码数据,并且基于较高质量层级的重构(例如,较高空间分辨率的重构)对信号执行进一步分析。此方法可以应用于多个信号源,诸如记录体育或音乐事件的一大组视频摄像机,其中每个摄像机从不同的地点、角度和/或位置捕捉动作。因此,本方法可以用于对多个较小的编码数据流(例如,来自所有捕获设备)执行快速分析,可以通过网络很容易地传输这些数据流,然后仅请求用于帧和/或感兴趣的捕获设备的另外的编码数据的子集。例如,可以基于对来自多个摄像机的一组重构的分析来确定足球的位置,然后可以仅向根据分析被认为含有足球的摄像机请求进一步数据。以此方式,可以提供人工智能混合应用。因此,可以对有效传送和重构的信号数据远程执行高级和复杂的数据中心处理。

[0027] 例如,通常应用卷积神经网络作为在远程计算设备处的视频分析功能的一部分(例如所谓的“基于云的对象识别”)。多层编码格式使此视频分析能够多倍加速,并且使基于云的边缘视频分析成为可能。多层编码格式(诸如VC-6)本身对图像识别算法“友好”,并

且可以将处理速度提高5-10倍。

[0028] 本文描述的某些实例可以解决具有多个摄像机的问题,诸如20个视频摄像机以每秒60帧的超高清(UHD)(UHDp60)流式传输数据,并且还需要对该流式传输的数据执行视频分析。例如,在捕获点(诸如足球场或音乐会会场)与用于高级视频分析的场地(诸如实施卷积神经网络架构以识别对象的数据中心)之间通常没有可用带宽。将数据中心带到足球场也是不可能的,即使偶尔使用适度的处理基础设施也会成本高昂,和/或需要移动大量固定式处理硬件。然而,使用本实例,可以在现场执行多层格式的编码,并且仅传输层的子集用于分析(通常是多个源设备的较低层或最低层)。如果传输较低层或最低层,则该位置的编码流可以包括多个低位率编码流,例如低分辨率流,可以使用现有带宽将其轻松容纳。然后,可以按需通过网络将较高层的不同组编码数据传输到数据中心用于进一步处理。

[0029] 与比较方法相比,当前描述的实例具有优势以应对传输约束和处理功率约束两者。例如,一种比较方法是在分析之前抽取视频帧。然而,此方法有一个缺点,即在有机会分析其内容之前可能会修剪有用的帧。由于抽取,在连续多组分析数据之间的大的离散变化也可能“混淆”神经网络架构和/或使这些数据对不同的采样率不灵活(例如,经过1/10修剪训练的系统可能无法适用1/5修剪)。相反,本实例允许远程位置接收较高频率的信号流(例如,避免需要按时大量抽取),因为较低层信号很小并且可以通过受限带宽连接传输。因此,远程处理位置可以接收接近连续的多组帧。另一种比较方法是将每个分析的帧缩小到较低分辨率,例如在传输到远程处理位置之前。然而,此方法意味着在远程位置处的处理仅限于较低分辨率的重构,例如可能会丢失可以提高分类准确度的细节。这最终会限制系统的整体分类性能。在本实例中,以低分辨率提供第一编码数据流用于初始分析,但是可以有效地请求和传输进一步细节,例如无需适配可用于全无损编码的多层编码。例如,远程处理位置可以请求进一步编码数据,并且该进一步编码数据可以包括用于增强在远程处理位置处的现有低分辨率重构的较高分辨率的残差数据。在使用多个层对源信号进行编码的实现方式中,可以请求和传输不同的层来选择性地提高准确度,有效地允许AI处理算法“按需”请求最少量的数据以适当地分析边缘数据信号。

[0030] 引言

[0031] 作为进一步背景,本文描述的实例涉及信号处理。信号可以被认为是样本序列(即二维图像、视频帧、视频场、声音帧等)。在说明书中,术语“图像”、“图片”或“平面”(意指“超平面”的最广泛含义,即具有任意维数和给定采样网格的元素阵列)将通常用于标识沿样本序列的信号样本的数字渲染,其中,每个平面对于其每个维度(例如X和Y)具有给定的分辨率,并且包括一组平面元素(或“元素”或“画素”或通常称为“像素”的用于二维图像的显示元素,通常称为“体素”的用于体积图像的显示元素等),其以一个或多个“值”或“设置”为特征(例如,通过非限制性实例的方式,合适颜色空间中的颜色设置、指示密度等级的设置、指示温度等级的设置、指示音频音高的设置、指示幅度的设置、指示深度的设置、指示阿尔法通道透明度等级的设置等)。每个平面元素均由一组合适的坐标标识,指示所述元素在图像的采样网格中的整数位置。信号维度可以仅包含空间维度(例如,在图像的情况下)或时间维度(例如,在信号随时间演变的情况下,诸如视频信号)。

[0032] 作为非限制性实例,信号可以是图像、音频信号、多通道音频信号、遥测信号、视频信号、3DoF/6DoF视频信号、体积信号(例如,医学成像、科学成像、全息成像等)、体积视频信

号,甚至是具有超过四个维度的信号。

[0033] 为简单起见,本文描述的实例通常是指显示为设置的2D平面的信号(例如,合适颜色空间中的2D图像),诸如例如视频信号。术语“帧”或“场”将与术语“图像”可互换地使用,以指示视频信号的时间上的样本:针对由帧构成的视频信号(逐行视频信号)示出的任何概念和方法也可以容易地应用于由场构成的视频信号(隔行视频信号),反之亦然。尽管本文说明的实施例的重点是图像和视频信号,但本领域技术人员可以容易地理解相同的概念和方法也适用于任何其他类型的多维信号(例如,音频信号、体积信号、立体视频信号、3DoF/6DoF视频信号、全光信号、点云等)。

[0034] 本文描述的某些基于分层的分级格式使用变化的校正量(例如,也以“残差数据”的形式,或简称为“残差”),以便在给定的质量层级上生成与原始信号最相似(或甚至无损重构)的信号重构。校正量可以基于给定质量层级的预测再现的保真度。

[0035] 基于分层的分级编码方案或格式的实例

[0036] 在优选实例中,编码器或解码器是基于分层的分级编码方案或格式的一部分。基于分层的分级编码方案的实例包含LCEVC:MPEG-5第2部分LCEVC(“低复杂度增强视频编码”)和VC-6:SMPTE VC-6 ST-2117,前者描述在PCT/GB2020/050695(和相关联的标准文献)中,后者描述在PCT/GB2018/053552(和相关联的标准文献)中,所有这些文献通过引用并入本文。然而,本文所示的概念不必限于这些特定的分级编码方案。本文描述的某些实例最佳地与VC-6编码一起工作。

[0037] 图1至图7提供了不同示例性基于分层的分级编码格式的概览。这些是作为用于添加进一步信号处理操作的上下文提供的,这些操作在图7之后的图中列出。图1至图5提供了类似于SMPTE VC-6 ST-2117的实现方式的实例,而图6和图7提供了类似于MPEG-5第2部分LCEVC的实现方式的实例。可以看出,两组实例利用共同的底层操作(例如,下采样、上采样和残差生成),并且可以共享模块化实现技术。

[0038] 图1非常一般地示出了分级编码方案。分级编码器102检索待编码的数据101并且输出编码数据103。随后,分级解码器104接收编码数据103,将该数据解码并且输出解码数据105。

[0039] 典型地,本文实例中使用的分级编码方案创建一个基础或核心层级以及一个或多个层级的残差,该基础或核心层级是处于较低质量层级的原始数据的表示,该一个或多个层级的残差可用于使用基础层级数据的解码版本来重建处于较高质量层级的原始数据。一般来说,如本文所使用的术语“残差”指代参考阵列或参考帧的值与数据的实际阵列或帧之间的差。阵列可以是表示编码单元的一维或二维阵列。举例来说,编码单元可以是对应于输入视频帧的类似大小的区域的 2×2 或 4×4 组的残差值。

[0040] 应注意,该一般化实例对于所输入信号的性质为不可知的。对如本文中所使用的“残差数据”的提及指代从一组残差导出的数据,例如残差集合本身或对残差集合执行的一组数据处理操作的输出。在整个本说明书中,一般来说,一组残差包含多个残差或残差要素,每一残差或残差要素对应于信号要素,即信号或原始数据的要素。

[0041] 在特定实例中,该数据可以是图像或视频。在这些实例中,一组残差对应于视频的图像或帧,其中每一残差与信号的像素相关联,该像素为信号要素。

[0042] 本文描述的方法可以应用于反映视频信号的不同颜色分量的所谓的数据平面。例

如,本文描述的方法可以应用于反映不同颜色通道的YUV或RGB数据的不同平面。可并行地处理不同颜色通道。可以按照任何逻辑顺序整理每个流的分量。

[0043] 现在将描述分级编码方案,其中可以部署本发明的概念。该方案在图2至图5中概念性地示出,并且通常对应于上文描述的VC-6。在此类编码技术中,残差数据被用于逐步提高的质量层级。在此提出的技术中,核心层以第一分辨率表示图像,并且分层分级中的后续层是解码侧以较高分辨率重构图像所需的残差数据或调整层。每一层或层级可以被称为梯队指数,使得残差数据是校正较低梯队指数中存在的低质量信息所需的数据。此分级技术中的每一层或梯队指数,尤其是每一残差层,通常是具有许多零值元素的相对稀疏的数据组。当提到梯队指数时,它是指该层级的所有梯队或分量组,例如,在该质量层级执行的变换步骤中产生的所有子组。

[0044] 以此特定的分级方式中,所描述的数据结构消除了对先前或后续质量层级的任何要求或依赖。可以单独对质量层级进行编码和解码,而无需参考任何其他层。因此,与许多已知的其他分级编码方案相反,其中需要解码最低质量层级以便解码任何更高质量层级,所描述的方法不需要解码任何其他层。然而,下面描述的交换信息的原理也可以适用于其他分级编码方案。

[0045] 如图2中所示,编码数据表示一组层或层级,这里通常称为梯队指数。基础或核心层级表示原始数据帧210,尽管处于最低质量层级或分辨率,并且后续残差数据梯队可以与核心梯队指数处的数据相组合,以逐步更高的分辨率重建原始图像。

[0046] 为了创建核心-梯队指数,可以使用与要在分级编码操作中使用的层级或梯队指数的数量对应的多个下采样操作201对输入数据帧210进行下采样。所需的下采样操作201比分级中的层级数少一个。在本文说明的所有实例中,存在输出编码数据的4个层级或梯队指数以及相应的3个下采样操作,但是当然应理解,这些仅是为了说明。其中 n 指示层级的数量,下采样器的数量是 $n-1$ 。核心层级 R_{1-n} 是第三下采样操作的输出。如上所述,核心层级 R_{1-n} 对应于处于最低质量层级的输入数据帧的指示。

[0047] 为了区分下采样操作201,将按照对输入数据210执行操作的顺序或通过其输出所表示的数据来引用每一个操作。例如,实例中的第三下采样操作 201_{1-n} 也可以被称为核心下采样器,因为其输出生成核心梯队指数或梯队 $_{1-n}$,也就是说,在此层级的所有梯队的指数是 $1-n$ 。因此,在此实例中,第一下采样操作 201_1 对应于 R_1 下采样器,第二下采样操作 201_2 对应于 R_2 下采样器,并且第三下采样操作 201_{1-n} 对应于核心或 R_3 下采样器。

[0048] 如图2中所示,表示核心质量层级 R_{1-n} 的数据经历上采样操作 202_{1-n} ,这里称为核心上采样器。第二下采样操作 201_2 的输出(R_2 下采样器的输出,即核心下采样器的输入)和核心上采样器 202_{1-n} 的输出之间的差异 203_2 被输出作为第一残差数据 R_2 。此第一残差数据 R_2 相应地表示核心层级 R_3 和用于创建该层级的信号之间的误差。因为在此实例中,该信号本身已经经历了两个下采样操作,所以第一残差数据 R_2 是调整层,其可以用于以比核心质量层级更高的质量层级但比输入数据帧210更低的质量层级重建原始信号。

[0049] 图2和图3概念性地示出了如何创建表示更高质量层级的残差数据的变化。

[0050] 在图2中,第二下采样操作 201_2 (或 R_2 下采样器,即用于创建第一残差数据 R_2 的信号)的输出被上采样 202_2 ,并且第二下采样操作 201_2 (或 R_2 下采样器,即 R_1 下采样器的输出)的输入之间的差异 203_1 以与创建第一残差数据 R_2 几乎相同的方式进行计算。此差异相

应地是第二残差数据 R_{-1} 并且表示调整层,该调整层可以用于使用来自较低层的数据以较高质量层级重建原始信号。

[0051] 然而,在图3的变型中,第二下采样操作 201_{-2} (或 R_{-2} 下采样器)的输出与第一残差数据 R_{-2} 组合或求和 304_{-2} ,以重建核心上采样器 202_{1-n} 的输出。在此变型中,正是此重建的数据被上采样 202_{-2} ,而不是下采样的数据。类似地,将上采样数据与第二下采样操作的输入(或 R_{-2} 下采样器,即 R_{-1} 下采样器的输出)进行比较 203_{-1} ,以创建第二残差数据 R_{-1} 。

[0052] 图2和图3的实现方式之间的变化导致两个实现方式之间的残差数据的轻微变化。图2受益于更大的并行化潜力。

[0053] 重复所述过程或循环以创建第三残差 R_0 。在图2和图3的实例中,输出残差数据 R_0 (即,第三残差数据)对应于最高层级,并且在解码器处被用于重建输入数据帧。在此层级处,差分操作基于与第一下采样操作的输入相同的输入数据帧。

[0054] 图4示出了实例编码过程401,用于对数据的层级或梯队指数中的每一个进行编码,以产生一组具有梯队指数的编码数据梯队。此编码过程仅用作用于对层级中的每一个进行编码的合适编码过程的实例,但是应理解,可以使用任何合适的编码过程。过程的输入是从图2或图3输出的相应层级的残差数据,并且输出是一组编码残差数据的梯队,编码残差数据的梯队一起分级地表示编码数据。

[0055] 在第一步中,执行变换402。变换可以是如W02013/171173中描述的方向分解变换或小波或离散余弦变换。如果使用方向分解变换,则可以输出一组四个分量(也称为变换系数)。当提到梯队指数时,它统称为所有方向(A、H、V、D),即4个梯队。然后,在熵编码之前,将分量组量化403。在此实例中,熵编码操作404被耦合到稀疏化步骤405,该稀疏化步骤利用残差数据的稀疏性来减小总的文件大小,并且涉及将数据元素映射到有序二叉树。在W02019/111004中进一步描述了熵编码和稀疏化的此类耦合,但是此类过程的精确细节与对本发明的理解无关。每个残差阵列都可以被认为是一个梯队。

[0056] 上面阐述的过程对应于适用于根据SMPTE ST 2117,VC-6多平面图片格式对用于重构的数据进行编码的编码过程。VC-6是一种灵活的、多分辨率的、仅帧内位流格式,能够压缩任何有序的整数元素网格组,网格中的每一个都具有独立的大小,但也是为图片压缩而设计的。它采用数据不可知技术(data agnostic technique)进行压缩,并且能够压缩低位或高位深度的图片。位流的标头可以含有关于图片的多种元数据。

[0057] 如将理解的,可以使用单独的编码器或编码操作来实现每个梯队或梯队指数。类似地,编码模块可以被分成下采样和比较的步骤,以产生残差数据,并且随后对残差进行编码,或替代地,梯队的步骤中的每一个都可以在组合的编码模块中实现。因此,过程可以例如使用4个编码器来实现,每个梯队指数一个编码器,1个编码器和多个编码模块并行或串行操作,或一个编码器对不同的数据组重复操作。

[0058] 下面给出了重构原始数据帧的实例,该数据帧已经使用上述示例性过程进行了编码。此重构过程可以被称为锥体重构。有利的是,该方法提供了一种高效的技术,用于重构在所接收的数据组中编码的图像,所接收的数据组可以通过数据流的方式接收,例如,通过单独解码对应于不同图像大小或分辨率层级的不同分量组,并且将来自一个解码分量组的图像细节与来自较低分辨率分量组的放大的解码图像进行组合。因此,通过对两个或更多个分量组执行此过程,在结构处的数字图像或细节可以针对逐步更高的分辨率或更多数量

的像素来重构,而不需要接收最高分辨率分量组的全部或完整的图像细节。相反,该方法有助于逐步增加分辨率越来越高的细节,同时以分阶段的方式从较低分辨率的分量组重构图像。

[0059] 此外,单独解码每个分量组便于并行处理所接收的分量组,从而在多个处理可用的实现方式中提高重构速度和效率。

[0060] 每个分辨率层级对应于一个质量层级或梯队指数。这是一个集合术语,与描述所有新的输入或所接收的分量组的平面(在此实例中是整数值元素的网格表示)以及指数 m 循环的输出重构图像相关联。例如,梯队指数为零的重构图像是锥体重构的最后循环的输出。

[0061] 锥体重构可以是这样一种过程:从初始梯队指数开始重构倒锥体,并且使用新残差的循环来导出更高的梯队指数,直到在梯队指数为零处的最大质量,质量为零。循环可以被认为是此类锥体重构中的步骤,该步骤由指数 m 识别。该步骤通常包括对来自可能的先前步骤的输出数据进行上采样,例如,对解码的第一分量组进行放大,并且采取新的残差数据作为进一步输入,以便获取将在可能的后续步骤中进行上采样的输出数据。在仅接收到第一分量组和第二分量组的情况下,梯队指数的数量将是两个,并且不存在可能的后续步骤。然而,在分量组或梯队指数的数量为三个或更多的实例中,输出数据可以在下面的步骤中被逐步上采样。

[0062] 第一分量组通常对应于初始梯队指数,所述初始梯队指数可以由梯队指数 $1-N$ 表示,其中 N 是平面中梯队指数的数量。

[0063] 通常,对解码的第一分量组的放大包括对初始梯队指数的解码过程的输出应用上采样器。在实例中,这涉及使从初始梯队指数分量组的解码输出的重构图片的分辨率与对应于 $2-N$ 的第二分量组的分辨率一致。典型地,来自较低梯队指数分量组的放大输出对应于较高梯队指数分辨率的预测图像。由于较低分辨率的初始梯队指数图像和上采样过程,预测图片通常对应于平滑或模糊的图片。

[0064] 将来自上述梯队指数的较高分辨率细节添加到此预测图片,提供了组合的、重构的图像组。有利的是,在一个或多个较高梯队指数分量组的所接收的分量组包括残差图像数据或指示放大的预测图片和原始的、未压缩的或预编码的图像之间的像素值差异的数据的情况下,为了重构给定分辨率或质量的图像或数据组所需的接收数据量可以显著小于使用其他技术接收相同质量的图像所需的数据量或数据率。因此,根据该方法,通过将以较低分辨率所接收的低细节图像数据与以越来越高的分辨率所接收的越来越多细节的图像数据相组合,降低了数据速率要求。

[0065] 典型地,编码数据组包括一个或多个进一步分量组,其中一个或多个进一步分量组中的每一个分量组对应于比第二分量组更高的图像分辨率,并且其中一个或多个进一步分量组中的每一个分量组对应于逐步更高的图像分辨率,该方法包括,对于一个或多个进一步分量组中的每一个分量组,解码分量组以便获得解码组,该方法进一步包括,对于一个或多个进一步分量组中的每一个分量组,按照对应图像分辨率的升序:对具有最高对应图像分辨率的重构组进行放大,以便将重构组的对应图像分辨率增加到等于进一步分量组的对应图像分辨率,并且将重构组和进一步分量组组合在一起,以便产生进一步重构组。

[0066] 以此方式,该方法可以涉及采取给定分量组层级或梯队指数的重构图像输出,放大该重构组,并且将其与上述分量组或梯队指数的解码输出相组合,以产生新的、较高分辨

率的重构图片。应理解,对于逐步更高的梯队指数,这可以根据所接收的组中的分量组的总数重复执行。

[0067] 在典型的实例中,分量组中的每一个分量组对应于逐步更高的图像分辨率,其中每个逐步更高的图像分辨率对应于在对应图像中的像素数量的四倍增加。因此,典型地,对应于给定分量组的图像大小是对应于下面的分量组的图像的大小或像素数的四倍,或是两倍的高度和两倍的宽度,该下面的分量组是具有比所讨论的梯队指数小一的梯队指数的分量组。例如,其中每个对应图像的线性大小相对于下面的图像大小是两倍的分量组中的所接收的一组可以使放大操作更简单。

[0068] 在所示的实例中,进一步分量组的数量是两个。因此,所接收的组中的分量组的总数为四个。这对应于初始梯队指数为梯队-3。

[0069] 第一分量组可以对应于图像数据,并且第二分量组和任何进一步分量组对应于残差图像数据。如上所述,在最低梯队指数(即第一分量组)含有正在传输的图像的低分辨率或下采样版本的情况下,该方法为给定的图像大小提供了尤其有利的数据速率需求降低。以此方式,对于重构的每个循环,从低分辨率图像开始,该图像被放大以便产生高分辨率但平滑的版本,然后通过增加该放大的预测图片和将以该分辨率传输的实际图像之间的差异来改善该图像,并且此另外的改善可以在每个循环重复。因此,在初始梯队指数之上的每个分量组仅需要含有残差数据,以便重新引入在将原始图像向下采样到最低梯队指数时可能已经丢失的信息。

[0070] 该方法提供了一种获得图像数据的方式,例如,在接收到含有已经(例如通过分解、量化、熵编码和稀疏化)被压缩的数据的组时,所述图像数据可以是残差数据。当结合原始或传输前数据稀疏的组使用时,稀疏化步骤尤其有利,该原始或传输前数据稀疏的组可以通常对应于残差图像数据。残差可以是第一图像的元素和第二图像的元素之间的差异,通常位于同一位置。此类残差图像数据通常可能具有高度的稀疏性。这可以被认为是对应于图像,其中细节区域稀疏地分布在细节最少、可忽略或不存在的区域中。此类稀疏数据可以被描述为数据阵列,其中数据以至少二维结构(例如,网格)来组织,并且其中如此组织的大部分数据为零(逻辑上或数字上)或被认为低于某个阈值。残差数据只是一个实例。此外,元数据可能是稀疏的,因此通过此过程在很大程度上减少了大小。发送已经稀疏化的数据允许通过省略发送此类稀疏区域,而是在解码器处的所接收的字节组内的适当位置重新引入它们,来实现所需数据速率的显著降低。

[0071] 通常,熵解码、去量化和方向合成变换(directional composition transform)步骤是根据编码器或发送所接收的编码数据组的节点所限定的参数来执行的。对于每个梯队指数或分量组,这些步骤用于解码图像数据,以便得到可以按照上面公开的技术与不同梯队指数组合的组,同时允许以数据高效的方式传输每个层级的组。

[0072] 还可以提供一种根据上面公开的方法来重构编码数据组的方法,其中根据上面公开的方法来执行第一分量组和第二分量组中的每一个的解码。因此,本公开的有利解码方法可以用于所接收的图像数据组中的每个分量组或梯队指数,并且相应地进行重构。

[0073] 参考图5,现在描述解码实例。接收一组编码数据501,其中该组包括四个梯队指数,每个梯队指数包括四个梯队:从梯队₀(最高分辨率或质量层级)到梯队₃(初始梯队)。梯队₃分量组中携带的图像数据对应于图像数据,并且其他分量组含有该传输的图像的残差

数据。尽管层级中的每一个层级可以输出可被认为是残差的数据,但是初始梯队层级(即梯队₋₃)中的残差有效地对应于实际的重构图像。在阶段503处,并行处理分量组中的每一个,以便对该编码组进行解码。

[0074] 参考初始梯队指数或核心梯队指数,对每一个分量组梯队₋₃到梯队₀执行以下解码步骤。

[0075] 在步骤507处,分量组被去稀疏化。去稀疏化可以是在其他基于分层的分级格式中不执行的任选步骤。在此实例中,去稀疏化导致由在每个梯队处所接收的编码字节组重建稀疏二维阵列。在二维阵列内未被接收的位置处分组的零值(由于为了减少传输的数据量而从传输的字节分组中省略了)由此过程重新填充。阵列中的非零值在重建的二维阵列中保持其正确的值和位置,去稀疏化步骤在适当的位置或其间的的位置的分组重新填充传输的零值。

[0076] 在步骤509处,将范围解码器应用于每个梯队的去稀疏化组,以便使用像素值替换阵列内的编码符号,所述范围解码器的配置参数对应于在传输之前对传输数据进行编码所使用的参数。根据图像的像素值分布的近似值,用像素值代替所接收的组中的编码符号。使用分布的近似值(即图像中所有像素值的每个值的相对频率,而不是真实分布),准许减少解码该组所需的数据量,这是因为范围解码器需要分布信息以便执行此步骤。如本公开中所描述的,去稀疏化和范围解码的步骤是相互依赖的,而不是顺序的。这由流程图中的箭头形成的回路指示。

[0077] 在步骤511处,值阵列被去量化。根据在传输之前对分解的图像进行量化的参数再次执行此过程。

[0078] 在去量化之后,在步骤513通过合成变换来变换该组,该合成变换包括对去量化的阵列应用逆方向分解操作。这使得根据包括平均、水平、竖直和对角算子的算子组的方向滤波被反转,使得所得阵列是梯队₋₃的图像数据和梯队₋₂到梯队₀的残差数据。

[0079] 阶段505示出了在利用梯队分量组501中的每一个的合成变换的输出进行重构所涉及的若干循环。阶段515指示从解码器503输出的用于初始梯队的重构图像数据。在实例中,重构图片515具有64x64的分辨率。在516处,对此重构图片进行上采样,以便将其像素的组成数量增加四倍,从而产生具有128x128分辨率的预测图片517。在阶段520处,将预测图片517添加到来自梯队₋₂的解码器的输出的解码残差518。这两个128x128大小的图像相加产生一个128x128大小的重构图像,该重构图像含有来自初始梯队的平滑图像细节,该平滑图像细节由来自梯队₋₂的残差的较高分辨率细节增强。如果所需的输出分辨率是对应于梯队₋₂的分辨率,则可以输出或显示此所得的重构图片519。在本实例中,重构图片519用于进一步循环。在步骤512处,以与步骤516相同的方式对重构图像519进行上采样,以便产生256x256大小的预测图片524。然后在步骤528处将其与解码的梯队₋₁输出526组合,从而产生256x256大小的重构图片527,该重构图片是用残差526的较高分辨率细节增强的预测519的放大版本。在530处,最后一次重复此过程,并且将重构图片527放大到512x512的分辨率,用于在阶段532处与梯队0残差组合。从而获得512x512的重构图片531。

[0080] 图6和图7示出了可以利用本发明原理的其他的分级编码技术。此技术是一种灵活的、适应性强的、高效的且计算成本低廉的编码格式,其将不同的视频编码格式、基础编解码器(例如,AVC、HEVC或任何其他当前或未来的编解码器)与编码数据的至少两个增强层级

相组合。

[0081] 编码方案的一般结构使用以基础编解码器进行编码的经下采样源信号,将第一层级的校正数据添加到基础编解码器的经解码输出以产生经校正图片,且接着将另一层级的增强数据添加到经校正图片的经上采样版本。因此,流被认为是基础流和增强流,它们可以被进一步复用或以其他方式组合以生成编码数据流。在某些情况下,基础流和增强流可以单独传输。如本文描述的,对编码数据的引用可以指增强流或基础流和增强流的组合。基础流可以由硬件解码器解码,而增强流可以适用于具有适当功耗的软件处理实现。此通用编码结构创建了多个自由度,这些自由度允许对许多情况的极大灵活性和适应性,从而使得编码格式适用于许多用例,包含OTT传输、直播流、直播超高清UHD广播等。尽管基础编解码器的经解码输出并非意图用于检视,但其为较低分辨率下的完全经解码视频,从而使得输出与现有解码器兼容,并且在认为合适的情况下也可用作较低分辨率输出。

[0082] 在某些实例中,可以使用一组网络抽象层单元(NALU)将每个或两个增强流封装到一个或多个增强位流中。NALU意图封装增强位流以便将增强应用于正确的基础重构帧。NALU可例如含有到NALU的参考索引,其含有增强必须应用于的基础解码器经重构帧位流。以此方式,增强可同步到基础流和组合的每一位流的帧以产生经解码输出视频(即,增强层级的每一帧的残差与基础解码流的帧组合)。图片的群组可表示多个NALU。

[0083] 返回到上文所描述的初始过程,其中基础流连同增强流内的两个层级(或子层级)的增强一起提供,一般化编码过程的实例描绘于图6的框图中。初始分辨率的输入视频600经处理以生成各种编码流601、602、603。通过向基础编解码器(例如,AVC、HEVC或任何其它编解码器)馈送输入视频的下取样版本来产生第一编码流(经编码基础流)。编码的基础流可被称为基础层或基础层级。通过处理通过取经重构基础编解码器视频与输入视频的下取样版本之间的差而获得的残差来产生第二编码流(经编码层级1流)。通过处理通过取经重构基础编码视频的经校正版本的上取样版本与输入视频之间的差而获得的残差来产生第三编码流(经编码层级2流)。在某些情况下,图6的组件可提供一般低复杂性编码器。在某些情况下,可通过形成低复杂性编码器的一部分的编码过程来生成增强流,且低复杂性编码器可被配置成控制独立的基础编码器和解码器(例如,封装为基础编解码器)。在其他情况下,基础编码器和解码器可供应为低复杂性编码器的一部分。在一种情况下,图6的低复杂性编码器可被视为用于基础编解码器的一种形式的封套(wrapper),其中基础编解码器的功能性可对于实施低复杂性编码器的实体隐藏。

[0084] 由下采样组件105示出的下采样操作可以应用于输入视频以产生待由基础编解码器的基础编码器613编码的下采样视频。下采样可在竖直和水平两个方向上进行,或替代地仅在水平方向上进行。基础编码器613和基础解码器614可以由基础编解码器实现(例如,作为公共编解码器的不同功能)。基础编解码器和/或基础编码器613和基础解码器614中的一个或多个可以包括适当配置的电子电路系统(例如,硬件编码器/解码器)和/或由处理器执行的计算机程序代码。

[0085] 每个增强流编码过程可能不一定包含上采样步骤。例如,在图6中,第一增强流在概念上是校正流,而第二增强流被上采样以提供增强层级。

[0086] 更详细地参见生成增强流的过程,为了生成编码的层级1流,编码的基础流由基础解码器614解码(即,解码操作应用于编码的基础流以生成解码的基础流)。解码可由基础编

解码器的解码功能或模式执行。接着在层级1比较器610处创建经解码基础流与下采样的输入视频之间的差(即,减法运算应用于下采样的输入视频和经解码基础流以生成第一组残差)。比较器610的输出可以被称为第一组残差,例如残差数据的表面或帧,其中在基础编码器613、基础解码器614和下采样块605的输出的分辨率下针对每个像素确定残差值。

[0087] 该差接着由第一编码器615(即,层级1编码器)编码以生成编码的层级1流602(即,将编码操作应用于第一组残差以生成第一增强流)。

[0088] 如上所述,增强流可包括第一增强层级602和第二增强层级603。第一增强层级602可被视为经校正流,例如以比输入视频600低的分辨率向基础编码/解码视频信号提供校正层级的流。第二增强层级603可被视为将经校正流转换为原始输入视频600的另一增强层级,例如其将增强层级或校正应用于从经校正流重构的信号。

[0089] 在图6的实例中,通过对另一组残差进行编码来创建第二增强层级603。该另一组残差由层级2比较器619生成。层级2比较器619确定解码层级1流的上采样版本(例如上采样组件617的输出)与输入视频600之间的差异。上采样组件617的输入是通过将第一解码器(即层级1解码器)应用于第一编码器615的输出而生成的。这生成一组解码的层级1残差。这些残差接着在求和组件620处与基础解码器614的输出组合。这有效地将层级1残差应用于基础解码器614的输出。其允许层级1编码和解码过程中的损失由层级2残差校正。求和组件620的输出可被视为表示对解码器处的编码的基础流601和编码的层级1流602应用层级1处理的输出的模拟信号。

[0090] 如所提到的,将上采样的流与输入视频进行比较,这创建了另一组残差(即,将差分操作应用于上采样的重建流以生成另一组残差)。另一组残差接着由第二编码器621(即,层级2编码器)编码为编码的层级2增强流(即,编码操作接着应用于另一组残差以生成另一编码的增强流)。

[0091] 因此,如图6中所示出和上文所描述,编码过程的输出为基础流601和一个或多个增强流602、603,所述一个或多个增强流优选地包括第一增强层级和另一增强层级。三个流601、602和603可在具有或不具有诸如控制标头等额外信息的情况下组合以生成用于表示输入视频600的视频编码框架结构的组合流。应注意,图6中展示的组件可对数据的块或编码单元进行操作,这些块或译码单元例如对应于处于特定分辨率水平的帧的 2×2 或 4×4 部分。这些组件在无任何块间相依性的情况下操作,因此其可并行地应用于帧内的多个块或编码单元。这不同于比较的视频编码方案,在比较的视频编码方案中,块之间存在相依性(例如空间相依性或时间相依性)。比较的视频编码方案的相依性限制并行水平且需要高得多的复杂性。

[0092] 在图7的框图中描绘相应的一般化解码过程。据称,图7可展示对应于图6的低复杂性编码器的低复杂性解码器。低复杂性解码器接收由低复杂性编码器生成的三个流601、602、603连同含有另外的解码信息的标头704。编码的基础流601由对应于在低复杂性编码器中使用的基础编解码器的基础解码器710解码。编码的层级1流602由第一解码器711(即,层级1解码器)接收,所述第一解码器对如由图1的第一编码器615编码的第一组残差进行解码。在第一求和组件712处,将基础解码器710的输出与从第一解码器711获取的解码的残差组合。通过上采样组件713对可称为层级1经重构视频信号的组合视频进行上采样。编码的层级2流103由第二解码器714(即,层级2解码器)接收。第二解码器714对如由图1的第二编

码器621编码的第二组残差进行解码。尽管标头704在图7中展示为由第二解码器714使用，但其也可由第一解码器711以及基础解码器710使用。第二解码器714的输出是第二组解码的残差。这些可处于比第一组残差和到上采样组件713的输入更高的分辨率。在第二求和组件715处，来自第二解码器714的第二组残差与上采样组件713的输出（即，上采样的重构的层级1信号）组合，以重构经解码视频750。

[0093] 按照低复杂性编码器，图7的低复杂性解码器可在视频信号的给定帧的不同块或编码单元上并行地操作。另外，可并行地执行由基础解码器710、第一解码器711和第二解码器714中的两个或更多个进行的解码。这是可能的，因为不存在块间相依性。

[0094] 在解码过程中，解码器可解析标头704（其可含有全局配置信息、图片或帧配置信息和数据块配置信息）且基于那些标头而配置低复杂性解码器。为了重新创建输入视频，低复杂性解码器可对基础流、第一增强流和另一或第二增强流中的每一个进行解码。所述流的各帧可以经同步并且然后经组合以导出经解码视频750。取决于低复杂性编码器和解码器的配置，经解码视频750可为原始输入视频100的有损或无损重构。在许多情况下，经解码视频750可为原始输入视频600的有损重构，其中所述损失对经解码视频750的感知具有减小的影响或最小影响。

[0095] 在图6和图7的每一个中，层级2和层级1编码操作可以包含变换、量化和熵编码的步骤（例如，以该顺序）。这些步骤可以以类似于图4和图5中示出的操作的方式实现。编码操作还可以包含残差分级、加权和滤波。类似地，在解码阶段，可将残差传递通过熵解码器、解量化器（de-quantizer）和逆变换模块（例如，按该次序）。可使用任何合适的编码和相应的解码操作。然而，优选地，层级2和层级1编码步骤可在软件中执行（例如，如由编码设备中的一个或多个中央或图形处理单元执行）。

[0096] 如本文中所描述的变换可使用定向分解变换，诸如基于哈达玛的变换（Hadamard-based transform）。这两者可包括应用于残差的扁平编码单元（即， 2×2 或 4×4 残差块）的小内核或矩阵。关于变换的更多细节可例如查阅以引用的方式并入本文中的专利申请PCT/EP2013/059847或PCT/GB2017/052632。编码器可在待使用的不同变换之间选择，例如在待应用的内核的大小之间选择。

[0097] 变换可将残差信息变换到四个表面。举例来说，变换可产生以下分量或变换系数：平均、竖直、水平和对角。特定表面可以包括特定分量的所有值，例如第一表面可以包括所有平均值，第二表面可以包括所有竖直值等等。如在本公开中较早提及，由变换输出的这些分量可在这些实施例中采用为待根据所描述的方法量化的系数。量化方案可用于将残差信号创建为量，使得特定变量可采用仅特定离散量值。在此实例中，熵编码可以包括游程编码（RLE），然后使用霍夫曼编码器处理编码输出。在某些情况下，当需要熵编码时，可使用这些方案中的仅一个。

[0098] 综上所述，本文中的方法和装置基于一种总体方法，其经由现有编码和/或解码算法（例如MPEG标准，诸如AVC/H.264、HEVC/H.265等；以及非标准算法，诸如VP9、AV1等）构建，所述现有编码和/或解码算法用作相应地用于不同编码和/或解码方法的增强层的基线。实例的总体方法背后的想法是以阶层方式对视频帧进行编码/解码，与使用MPEG系列算法中所使用的基于块的方法形成对比。以阶层方式对帧进行编码包含针对全帧生成残差，且接着针对抽取帧生成残差，等等。

[0099] 如上所述,由于不存在块间相依性,这些过程可以并行地应用于帧的颜色分量的编码单元或块。还可以并行地执行对一组颜色分量内的每一个颜色分量的编码(例如,使得根据(帧数)*(颜色分量数)*(每帧的编码单元数)来复制该操作)。还应注意,不同颜色分量可以具有每帧不同数量的编码单元,例如可以在人类视觉可以检测到明度改变大于颜色改变时在比一组色度(例如,U或V)分量更高的分辨率下处理亮度(例如,Y)分量。

[0100] 因此,如上文所示出且描述,解码过程的输出是(任选的)基础重构,以及在较高级别处的原始信号重构。此实例尤其非常适合于在不同帧分辨率下创建经编码及经解码视频。举例来说,输入信号30可以是包括 1920×1080 分辨率下的帧的HD视频信号。在某些情况下,基础重构和层级2重构两者可由显示设备使用。例如,在网络流量的情况下,层级2流可能比层级1流和基础流更多地中断(因为该层级2流可以含有多达4倍的数据量,其中下采样将每个方向上的维度减少2)。在此情况下,当流量发生时,显示设备可以恢复显示基础重构,而层级2流被中断(例如,当层级2重构不可用时),并且然后当网络条件改进时恢复显示层级2重构。当解码设备遭受资源约束时可应用类似方法,例如执行系统更新的机顶盒可具有操作基础解码器220以输出基础重构,但可不具有处理容量来计算层级2重建。

[0101] 编码布置还使得视频分布器能够将视频分布到非均质设备集合;仅具有基础解码器720的那些设备检视基础重构,而具有增强层级的那些设备可检视较高质量层级2重构。在比较案例中,需要单独分辨率下的两个完整视频流来服务于两个设备集合。由于层级2和层级1增强流对残差数据进行编码,可更有效地对层级2和层级1增强流进行编码,例如残差数据的分布通常大部分质量都在0左右(即,不存在差)且通常采取约0的小范围值。量化之后的情况尤其如此。相比而言,处于不同分辨率的完整视频流将具有非零均值或中值的不同分布,该值需要较高位速率以供传输到解码器。在本文中所描述的实例中,残差由编码管道进行编码。这可包含变换、量化和熵编码操作。其还可包含残差分级、加权和滤波。随后将残差传输到解码器,例如作为L-1和L-2增强流,该增强流可与基础流组合作为混合流(或单独传输)。在一种情况下,设定用于包括基础流和两个增强流的混合数据流的位速率,且接着基于正处理的数据将不同自适应位速率应用于个别流以满足设定的位速率(例如,以低假影水平所感知的高质量视频可通过自适应地将位速率指派给不同个别流(即使在逐帧层级处)而构造,以使得经约束数据可由感知上最有影响的个别流使用,该个别流可随着图像数据改变而改变)。

[0102] 如本文中所描述的残差集合可被视为稀疏数据,例如在许多情况下对于给定像素或区域不存在差,且所得残差值为零。当查看残差的分布时,将许多概率质量分配到接近零定位的小残差值,例如对于-2、-1、0、1、2等的某些视频值发生得最频繁。在某些情况下,残差值的分布关于0对称或近似对称。在某些测试视频情况下,发现残差值的分布关于0类似于对数分布或指数分布的形状(例如,对称或近似对称)。残差值的准确分布可取决于输入视频流的内容。

[0103] 残差可自身被处理为二维图像,例如差的差量图像。以此方式,可以看到数据的稀疏性涉及在残差图像中可见的比如“点”、“小”、“线”、“边缘”、“拐角”等特征。已发现这些特征通常不完全相关(例如,在空间上和/或在时间上)。这些特征具有不同于其来源于的图像数据的特性(例如,原始视频信号的像素特性)的特性。

[0104] 由于残差的特性不同于所述残差来源于的图像数据的特性,因此通常不可能应用

标准编码方法,诸如传统动态图片专家组(MPEG)编码和解码标准中发现的那些方法。例如,许多比较方案使用大的变换(例如,在正常视频帧中的大的像素区域的变换)。归因于例如如上文所描述的残差的特性,对残差图像使用这些比较大变换将是极低效的。举例来说,使用经设计用于正常图像的区域的大块对残差图像中的小点进行编码将是非常困难的。

[0105] 本文描述的某些实例通过代替地使用小的和简单的变换内核(例如,如本文所呈现的 2×2 或 4×4 内核——定向分解和定向分解平方)来解决这些问题。可以使用哈达玛矩阵(例如,用于扁平 2×2 编码块的 4×4 矩阵,或用于扁平 4×4 编码块的 16×16 矩阵)来应用本文描述的变换。这在与比较的视频编码方法不同的方向上移动。将这些新方法应用于残差块会生成压缩效率。例如,某些变换生成不相关的经变换的系数(例如,在空间上),这些经变换的系数可以有效地压缩。虽然可以(例如针对在残差图像中的线)利用经变换的系数之间的相关,但这些相关可能导致编码复杂性,使得难以在传统设备和低资源设备上实施,并且这些相关通常生成需要校正的其他复杂假影。通过将某些残差值设定为0(即,不转发这些残差值以供处理)来预处理残差可提供可控制且灵活的方式来管理位速率和流带宽,以及资源使用。

[0106] 与对多层信号进行远程分析相关的实例

[0107] 图8示出了处于不同的质量层级的一组重构。在此实例中,不同的质量层级对应于不同的空间分辨率。在图8中示出了四个层:低于标清或缩略图分辨率层801;小分辨率层802;中分辨率层803;以及大分辨率层804。针对大分辨率层804的重构可以包括处于最高分辨率(例如UHD)和最高帧率(每秒60帧)的全帧的数据。针对中分辨率层803的重构可以处于较低分辨率,诸如高清(HD)。这些重建也可以具有较低帧率,诸如每秒30帧。在小层802处的重构可以处于标清(SD)分辨率并且可以处于与中层803相同或更低的帧率。针对最低层801的重构可以处于低于HD层级,例如,nHD或VGA类分辨率,并且还可以处于与小层802相同或更低的帧率。应注意,这些分辨率仅作为实例提供,并且不同的实现方式可以根据需要具有不同分层的分辨率。

[0108] 可以通过对信号进行编码和解码来生成图8的各种分层或层,如参考上文的图1至图5所描述的。在远程计算设备处,仅与缩略图层801相关的编码数据或与缩略图层801和小层802相关的编码数据才可以初始地从一组摄像机接收、解码并且用于重构在图中虚线805下面的表示。远程计算设备可以对与一个或多个最低层相关的重构执行一个或多个视频分析功能。该一个或多个视频分析功能可以包括对象检测和定位,如图8中所示。

[0109] 在实例中,如果在远程计算设备处执行的一个或多个视频分析功能指示特定信号源是感兴趣的,例如,如果在初始重构中检测到和/或定位到对象,则可以向在其中信号被编码的网络边缘发出请求以传输与另外的分层或层相关联的编码数据。例如,远程计算设备可以请求与中层803相关联的编码数据以便在中层的分辨率处构建重构。在上文描述的编码实例中,可以看出在最低层之上的层可以包括残差数据,使得针对中层803的编码数据仅包括在中分辨率处的编码残差数据,其可以被解码并添加到在远程计算设备处的上采样的或放大的预先存在的重构。因此,针对中层803的编码数据不包括在中层分辨率处的全信号流,并且因此可以被有效地传输和解码。

[0110] 在图8中,基于在中层803处构建的进一步编码数据的重构用于对在信号中的一个或多个感兴趣区域进行标识。这可以在中层803或一个或较低层处执行。感兴趣区域可以包

括在检测到的对象周围的区域。在此实例中,远程计算机设备可以被配置成在中层803处标识在重构中的感兴趣区域,并且使用此感兴趣区域来有条件地发出对与在较高层(例如高层804)处的感兴趣区域相关的编码数据的进一步请求。编码格式诸如VC-6允许仅对感兴趣区域进行部分解码,并且仅可以从编码场地传输针对感兴趣区域的编码数据。

[0111] 图9示出了在两个位置上的分布式信号处理的实例。第一位置910可以被称为“边缘”位置,因为此位置最接近一组信号源。第二位置920相对于第一位置910是远程的,例如通过一个或多个数据通信网络915通信耦合到第一位置910。第二位置920可以被称为“远程”位置。在图9中,第二位置920包括云数据中心。在其他实例中,第二位置920可以包括具有视频分析功能的任何其他场地。典型地,在数据通信网络915上的数据通信带宽比在第一位置910和第二位置920中的每一个位置处的数据通信带宽更加受限;例如,数据通信网络915可以包括因特网等。

[0112] 第一位置910包括一组信号源。在此实例中,这些信号源包括提供多个视频摄像机源935的一组视频摄像机930。在其他实例中,信号源可以包括麦克风、传感器等。第一位置910可以例如包括体育场和/或艺术(音乐、戏剧或电影)会场。第一位置910还包括编码平台940。这可以包括一个或多个计算设备(或专用硬件编码器),其通信耦合到多个视频摄像机源935。来自该一组信号源的流可以包括在高分辨率处的原始的和/或未压缩的信号,例如通过有线耦合的多个UHDp60流。编码平台940被配置成在基于层的编码中针对多个层中的每一个层生成编码数据流,其中在多个层中的不同层对应于不同的质量层级(例如,不同的空间分辨率),其中较高层表示较高质量层级。例如,在图1至图5的上下文中,较低层可以包括第⁰梯队下面的较低梯队,其中最低层是 R_{1-n} 层。这些层也对应于图8中所示的层。如果标准诸如VC-6用于在编码平台940处进行的编码,则将每个视频源编码为VC-6数据流。

[0113] 在比较实例中,在编码平台940处的编码之后,将编码数据流传输到另一个位置进行解码。在希望重构全质量信号流的其他位置,则将针对每个层的编码数据传输到该其他位置,并且通过解码针对每个层的编码数据来重构信号(例如,如在图5和图7中所描述的)。

[0114] 然而,在本实例中,可能没有合适的网络或处理带宽来将针对所有层的全组编码数据流传输到第二位置。另外,即使合适的带宽是可用的,视频分析功能也可能无法同时(例如实时)处理所有多个信号源的信号。

[0115] 因此,在本实例中,仅将多个信号源的针对一个或多个最低层的编码数据流从第一位置910传输到第二位置920。例如,至少可以将针对在分级中的最低层的编码数据流传输到第二位置920以供解码平台950接收和解码。由于多层编码(例如,如上文描述的),这可以在不调整编码或解码过程的情况下实现——第一位置简单地传输针对所需的较低层的编码数据流,而(至少初始地)不传输针对较高层(例如,与较高质量层级相关联的层)的编码数据流。

[0116] 在图9中,如框942所示,初始地,仅将每个多层锥体的“尖端”传输给该一组视频摄像机930。这可以包括仅传输的针对缩略图分辨率层810的编码数据。在其他情况下,其可以包括传输最低层诸如缩略图分辨率层801和小层802两者(例如,线805下面的层)的定义数量。所选择的最低层的数量可以取决于静态用户定义参数和/或动态参数,诸如在数据通信网络915上的可用带宽和/或在第二位置930处的处理资源。在此情况下,处理资源可以涉及对来自解码平台950的输出执行一个或多个视频分析功能的一个或多个处理器。解码平台

950通常实施与编码平台940的编码过程相对应的解码过程。例如,解码平台950可以实施VC-6解码过程。

[0117] 在本实例中,由解码平台950对通过框942传输的针对最低层的编码数据进行解码以生成处于第一质量层级的多个视频摄像机源的一组重构。例如,此操作可以包括为每个视频摄像机源生成重构,诸如在图5中的520。在第二位置处,一个或多个视频分析功能应用于重构。如上文描述的,这些功能可以包括场景分析、人脸检测、姿态检测、对象检测等。视频分析功能可以基于AI或神经网络方法。

[0118] 在图9的实例中,一个或多个视频分析功能的输出是对重构的一个或多个部分的标识以用于进一步分析。这可以包括对一个或多个编码数据流的标识,即对一个或多个特定视频摄像机源的标识。其还可以或替代性地包括对较多个帧(例如时间上的图像)中的一个帧和/或一个或多个感兴趣区域(例如,如由在帧内的空间区域和/或针对帧的特定的一组像素和/或编码单元所定义)的标识。此标识可以在第二位置920处(例如,通过在解码平台950的输出上运行的远程处理设备)执行。因此,第二位置920可以请求用于所标识的视频摄像机源、帧和/或感兴趣区域的不同数据。这由在图9中的双向箭头和框944示出。

[0119] 在接收到标识视频摄像机源、帧和/或感兴趣区域的请求时,编码平台940(或在第一位置910处耦合到编码平台940的设备)传输与所标识的视频摄像机源、帧和/或感兴趣区域相对应的另外的或进一步编码数据。这可以包括从在第一位置910中的编码平台940(或耦合的设备)执行部分文件召回操作。例如,可以需要仅编码数据以获得在图5中的残差数据526,并且这可以与较低层重构的上采样版本(例如在图5中来自520的524)组合。可以使用比处于相同层质量(例如分辨率)的全视频源更少的带宽来有效地传送此编码数据。例如,编码平台940可以传输针对中层803的数据。在一种情况下,编码平台940可以仅传输针对在帧内的位置的子集的编码残差数据,例如仅传输针对落入标识的感兴趣区域内的像素和/或编码单元的残差数据。编码方案诸如VC-6允许对视频帧的部分部分进行有效解码,例如由于独特的S树编码结构。例如,S树与算术上简单的变换(诸如上文描述的哈达玛变换)相组合,允许使用具有多数据的相同指令(SIMD)计算元素,并且因此该编码和/或解码尤其适合于具有SIMD和向量单元(诸如图形处理单元(GPU)和现场可编程门阵列(FPGA))的现代硬件。另外,本文描述的多层编码不需要比较JPEG/MPEG编解码器通常需要的应用特定加速。解码也比在整个数据帧上操作的比较残差数据编码方案更快。在一种情况下,进一步编码数据可以涉及一组分量的子集(例如,仅一个颜色分量诸如Y或一个主/中心音频通道)。

[0120] 因此,第二位置950接收针对标识的信号源、帧和/或感兴趣区域的进一步编码数据,例如与先前接收到的编码数据的较高质量层级相关联。然后由解码平台950对此进一步编码数据进行解码以生成第二组重构。这些重构例如可以包括在图5中示出为528的重构(或这些重构的一部分)。第二位置920然后能够指示对来自第二组重构的数据的一个或多个视频分析功能,其可以包括在较高空间分辨率处对至少感兴趣区域的进一步分析,例如在图8的中层803处的分析。

[0121] 此增量分析过程可以继续用于与在编码中的层一样多的层,或者直到来自视频分析功能的度量超过预定义的阈值。例如,可以迭代地请求、传输、解码和分析与增加的质量层级相关联的编码数据,直到超过准确度阈值和/或直到达到最高质量层级。虽然实例是关于从最低质量层级上升来说明的,但是本领域技术人员将理解类似的方法可以应用于在多

层编码中的任何两层,其中第一层处于比第二层更低的质量层级。

[0122] 在图9中,可以向视频摄像机源的子集(例如向与当前视频分析相关的一个或多个信号源)发出对用于全质量重构(例如在分级中的所有层)的差分或残差数据的请求,如在框946中所示。对视频摄像机源的子集的此选择可以根据作为先前框(例如,942和944)的一部分执行的分析来做出。例如,可能需要全质量重构用于当前正在捕获来自现场事件(诸如足球的位置、乐队中的主唱和/或舞台上的演员)的动作的视频摄像机。用于视频摄像机源的编码数据流当前是对视频分析无用的(例如,因为该编码数据流没有人和/或动作),不需要将该编码数据流传输到第二位置920。整个过程可以在现场事件期间动态地应用,使得被选为框946的一部分的视频摄像机可以在事件的过程期间不断地改变。

[0123] 在图13A至图13D中针对一个信号源示出了增量数据交换和分析的过程。图13A示出了信号源1330、信号流1335(例如,处于全质量)、编码器1340、通信网络1315、解码器1350和视频分析模块1308。编码器1340接收信号流1335。信号流1335然后被编码以构成编码的小层802、编码的中层803和编码的大层804。还示出了通过通信网络1315将编码的小层802发送(1309)到解码器1350(例如在远程位置处)。线1309指示将编码的小层发送到解码器。解码器1350对编码的小层802进行解码以产生解码的小层802',其被传输到视频分析模块1308。解码的小层802'也可能受到一定量的失真,这可以由在编码器1340中执行的信号流1335的量化来支配。视频分析模块1308对解码的小层802'执行视频分析。在一些情况下,视频分析模块1308可以对编码的小层802执行分析(即,作为对解码的小层802'的分析的补充或替代方案)。视频分析模块1308可以能够确定编码的信息的某些属性。在此情况下,编码的小层802和解码的小层802'都可以被发送到视频分析模块1308,但是编码的小层802不通过解码器1350传递。

[0124] 如图13B中所示,一旦视频分析模块1308接收到解码的小层802',该视频分析模块就通过通信网络1315向编码器1340发送请求1310,请求从编码器1340待传输到解码器1350的编码信号流1335的另一层。可以在对解码的小层802'进行分析之后发送请求1310。具体地,可以响应于对解码的小层802'的分析的结果来发送请求1310。例如,响应于对解码的小层802'的分析,视频分析模块1308可以确定对另一层(例如,解码的中层803')执行分析将是有益的,在此类实例中,视频分析模块1308向编码器1340发送请求1310,请求从编码器1340待传输到解码器1350的编码信号流1335的另一层(例如编码的中层803)。

[0125] 如图13C中所示,在编码器1340接收到请求1310(如图13B中所示)之后,通过通信网络1315(如图13C中所示)将编码的中层803发送(1311)到解码器1350。线1311指示将编码的中层803发送到解码器。解码器1350对编码的中层803进行解码以生成解码的中层803'。将解码的中层803'从解码器1350发送到视频分析模块1308。也可能下一帧的小层是信号流1335与编码的中层803并行地行进通过编码器和解码器,但这未示出。

[0126] 如图13D中所示,编码器1340通过通信网络1315将编码的大层804发送到解码器1350。虚线1312指示不总是将编码的大层804发送到解码器。在下文对不总是发送编码的大层804的有益效果进行说明。

[0127] 本文描述的方法和系统可以应用于一个信号源或应用于多个信号源。图9的配置可以被认为是AI辅助的传输漏斗,其允许对本地数据进行基于云的AI分析。在此实例中,针对所有信号源的全质量信号流可以是在边缘位置处可用的,并且可以由第二AI分析位置分

阶段获得。如图13D中所示,对于其中未传输针对该全组层的编码数据的信号源,可以稍后(例如,在非实时设置中)传输此编码数据用于存档目的。例如,可以针对来自现场事件的另外的编码数据整夜执行此操作,使得全质量视频摄像机源在现场事件之后在第二位置920处是可用的(例如,存档和/或使用/编辑以供分布)。以此方式,对于视频源,可以执行增强的基于AI的视频数据传送并且所有视频数据可以在远程是可用的而不需要在现场事件期间和在现场事件之后两者的额外带宽。在本文描述的某些情况下,可能以低于全质量或分辨率存储存档,例如在中层803处或处于HD而不是UHD的质量层级。在此情况下,可以不执行图13D的操作,并且可以仅需要将编码数据的中间层级的层传输到远程位置用于视频摄像机源。

[0128] 图10示出了分布式信号处理系统的另一个实例。此实例可以被视为图9的实例的替代性变型。图10是混合方法的实例,由此在第一位置910处而不是在第二位置处执行编码数据流的标识。因此,关于在何处执行一组编码信号的特定处理,本方法提供了灵活性。

[0129] 图10示出了如图9中所示的一组信号源930、一组信号流935(例如,处于全质量)和编码平台940。在此实例中,对编码平台940的输出执行了两个另外的处理操作。由预处理计算设备1005执行这些另外的处理操作。第一操作包括帧标识1010,并且第二操作包括摄像机标识1020。虽然在图10中以一种顺序示出,但是在其他的实现方式中这些操作也可以与操作1020至操作1010一起执行。虽然预处理计算设备1005是在第一位置910处示出的,但是该预处理计算设备可以替代性地在远程位置(例如图9的第二位置)处或在中间位置处。可以操作帧标识1010和摄像机标识1020以确定针对每个信号源930的特定高质量层级。可以使用类似于图9中所示的过程的过程来执行此操作。替代性地,如果在第一位置910处的计算设备上执行帧标识1010操作和摄像机标识1020操作,则这些操作可以访问层的全堆栈,例如编码信号流的所有编码信号流或大部分编码信号流。例如,可以给该组信号源分配一个优先级。由帧标识1010和摄像机标识1020标识的低优先级信号源可以与在多层格式中的一个或多个最低层(例如图8中的缩略图和/或小分辨率层801和802)相关联;由帧标识1010和摄像机标识1020标识的高优先级信号源可以与在多层格式中的全组或接近全组的层相关联。在图10中的框1030处,帧标识1010和摄像机标识1020的输出(可以是分配给每个视频摄像机源的优先级)用于定义针对每个信号源930传输的一组质量层级。

[0130] 在图10的实例中,边缘位置910与远程服务器1050和分布中心1060两者通信。远程服务器1050可以执行参考图9和图13A至图13D所描述的远程解码和视频分析。如通信1035所示,远程服务器1050可以提供远程视频分析以便分配优先级或以其他方式指示编码数据流的待传输到分布中心1060的数据包(例如用于与现场事件相关的视频和/或音频的实时分布)。因此,分布中心1060可以针对信号源的子集接收AI选择的处于全质量层级的视频源,其中在基于自动视频分析的现场事件期间发送到分布中心1060的编码数据和针对主要(现场)分布选择的信号源改变。这可以允许例如自动的足球或演员跟踪、自动的全人脸跟踪、自动的动作跟踪、自动的隐藏字幕摄像机审查等。分布中心1060可以将编码数据分布给一个或多个终端用户设备用于对解码信号进行渲染,例如观看视频源。图11是可以如何将本文描述的方法和系统应用在现场体育事件的环境中的进一步实例。图11中所示的配置类似于图9和图10中所示的配置。体育场1110位于第一位置,并且分布中心1120位于不同的远程位置。多个视频摄像机1130布置在体育场1110周围以从不同位置、角度和地点捕获体育事

件。视频摄像机1130向边缘机箱1140提供一组全质量视频流。边缘机箱1140包括计算设备以实施先前实例的编码平台。计算设备可以被配置成如图12所示的并且可以包括服务器计算设备。

[0131] 体育场1110和分布中心1120通过智能互连1115通信耦合,可以使用图9的云数据中心920、图10的远程服务器1050和一组数据通信网络中的一者或多者来实施该智能互连。服务器计算设备可以构成智能混合器1150的一部分,用于控制发送到分布中心1120以供实时分布的编码视频流。

[0132] 在此实例中,在构成第一(边缘)位置的体育场1110处将所有视频流编码(例如压缩)。边缘机箱1140可以实施编码平台,该编码平台使用多层编码诸如VC-6来对一组视频流1135进行编码。可以在最高的仅帧内质量处将该组视频流1135编码。

[0133] 实时地(即,使得可以在具有最低延迟的终端设备适当地渲染和观看帧),在全质量(即,所有层都在编码中)处传输用于来自该组视频摄像机1130的关键摄像机的编码数据,并且仅使用在编码中的较低层(例如仅使用较低分辨率的数据,诸如与图8中的层801和层802中的一个或多个层相关联的数据)向在该组视频摄像机1130中的其他摄像机发送。

[0134] 在图11中,代理源1152由分布中心1120(和任何远程处理服务器设备)接收,该分布中心包括针对在多层编码中的至少最低层的编码数据。这可以包括使得用于解码在图5中的重构515或在图8中的缩略图分辨率层801的这样的编码数据。分布中心1120还接收以主要源1154的形式的针对所有层的编码数据;例如,这可以包括编码数据以重构表示当前分布的视频源(例如,当前选择的用于正在由观看设备接收的足球比赛的摄像机)的全(例如4K)质量信号。也可以发送其他关键源1156以允许全质量重构(例如,针对当前一组动作的替代性摄像机地点)。可以基于由智能混合器1150进行的智能混合来选择主要源1154和关键源1156。例如,如参考图9和图13A所描述的,远程处理设备可以接收代理源1152,并且使用此操作来选择针对一个或多个较高层的编码数据,这些层通过次要源1158传输。远程处理设备可以使用代理和/或次要源1152和1158来标识相关时刻(例如,发生动作的时间段,诸如30分钟-50分钟)和/或相关摄像机,并且使用此标识来指示另外的层由分布中心1120传输和接收以构成全高质量编码流。

[0135] 在一种情况下,如果远程处理设备位于分布中心1120,则远程处理设备和分布中心1120两者都可以接收相同的代理源1152,并且由远程处理设备请求的任何次要源1158可以是可访问的(例如传递到)分布中心1120。在此情况下,如果远程处理设备指示切换到另一个摄像机源,并且已经接收到代理源1152和次要源1158,则接收全质量源1154或1156所需的全部是针对在次要源1158中携带的层之上的层的残差数据。这可以允许在源之间高效和快速的切换,因为仅需要针对源的一组增量的或差分的层。这允许实时的混合和摄像机/帧切换。如图13D中所示,在事件结束时,然后将针对未用作现场事件的分布的一部分但已被编码并且存储在边缘机箱1140处的较高层的剩余数据包传输到分布中心1120(例如,在非实时的基础上)用于存档。

[0136] 在特定实现方式中,在视频捕获位置处,信号源930或1130可以包括多个视频摄像机,该多个视频摄像机的全质量视频流(即,视频摄像机的默认输出)被传送到在第一位置910或1110处的服务器计算设备,其中编码平台940或1140以贡献标准压缩所有接收到的视频源。压缩数据或编码数据可以作为一组不断增加的文件本地保存到磁盘。在图10的情况

下,可以执行现场处理来索引编码数据;在图11的情况下,此操作可以在本地或远程执行。索引可以允许选择现场视频源并且允许选择每个摄像机的质量层级。可以手动地执行选择,例如由用户操作本地的或远程的混合应用程序或桌面执行,和/或通过远程的或本地的视频分析功能自动执行。在某些情况下,可以发送与最低或非常低的分辨率(例如图8中的801或802)相关联的编码数据,以便使用高级的视频分析功能(例如AI应用程序)进行自动索引。在某些情况下,可以仅传输一个颜色分量诸如Y或亮度分量用于远程分析和/或可以降低帧率(例如,可以传输N帧中的1帧,可以使用时分复用或帧分复用对其进行复用)。未立即传输的编码数据可以在本地存储为一组文件。此编码数据可以稍后传输,例如在现场事件之后,和/或保存到存储介质用于物理传输或本地存档。

[0137] 根据本文描述的实例,使用比较编码和分析方法发现的延迟优势可以显著降低,例如,可以远程执行视频分析功能,与本地处理相比,几乎没有另外的延迟。可以执行对特定的视频摄像机源的索引来选择视频摄像机源以供未来分布,例如与回放和/或视频助理裁判(VAR)相关的片段。视频分析功能可以用于标识特定的视频摄像机源以供人工审查和/或优先观看。例如,视频分析功能可以标识视频的一个或多个部分(例如,在存储的编码文件和/或特定摄像机中的特定帧)以供人类裁判审查。此外,可以标识感兴趣区域(诸如足球是否越过球门线)并且可以比完整帧更快地将针对这些区域的高质量重构的编码数据传输到远程位置。

[0138] 可以应用本文描述的某些实例作为在事件期间的不同模式。

[0139] 在第一事件模式中,在事件位置与远程演播室和/或数据中心之间的数据通信链路可以被配置成按优先级排列人类或AI选择的视频摄像机源(例如,图11中的源1154和1156)和请求的随机访问编码数据梯队或与感兴趣区域相关联的编码数据。如图11中所示,可以针对一个或多个摄像机连续传输低分辨率代理源1152来允许远程分析。可以基于对用于在远程位置处分析的编码数据的自动选择来使用针对数据通信链路的剩余备用容量,诸如次要源1158。这可以被视为一种自动“填隙”的形式。它可以被执行来按优先级排列梯队、平面、颜色分量和/或帧以供传输。可以在稍后的时刻对优先的编码数据进行解码来产生事件的精彩集锦。因此,这可以在事件期间有效地将编码数据从事件位置传输到远程位置,使得在远程位置(例如,分布中心1120)处针对视频的有用部分的编码数据更有可能是可用的。在一种情况下,可以仅传输产生HD可视视频所需的编码数据梯队来产生进一步视频,诸如新闻和精彩集锦。

[0140] 第二事件模式可以与事件的开始、结束和/或中间相关联,诸如在体育赛事中的半场或全场。在此模式中,用于编码的压缩比可以较高(因为在户外,不需要高质量的动作源)并且所需的源(即活动摄像机)的数量可以较少。例如,可能只有一个或两个摄像机来自本地演播室。在此第二事件模式中,数据通信链路可以显著较少地填充并且因此可以基于如本文描述的自动优先级排列来增加上文描述的填隙。

[0141] 最后,可以有针对事件何时结束的后事件模式。在此情况下,可以没有正在主动分布给观看者的摄像机源。在此情况下,数据通信链路的所有带宽都可以专用于上述“填隙”优先级排列。这可以继续直到在远程位置处接收到所有编码数据(例如,用于如图13D中所示的存档)和/或直到本地的一组计算设备关闭。

[0142] 在某些情况下,在第一位置处的计算设备可以位于在事件之前被驱动到第一位置

的移动单元(诸如生产卡车)内。在此情况下,未通过数据通信链路传输的任何编码数据都可以存储在位于移动单元内的存储介质中,并且移动单元可以被物理地驱动到远程位置,在该位置可以上传此移动单元存储的内容。

[0143] 本实例允许更快的远程处理。这可以通过高级神经网络架构实现实时推理。图14示出了在比较方法(1405)中应用于JPEG图像的各种处理操作与应用于使用本文描述的多层编码生成的重构(1410)的各种处理操作之间的比较。有三个处理操作:解码、检测和识别。解码可以包括对编码流进行解码来获得对在重构内的像素数据的访问。检测可以包括人脸检测,并且识别可以包括人脸识别。这些数字指示每HD帧用于各种活动的时间(以毫秒计)。可以看出,在解码时间上存在显著减少(减少到1/26)。检测时间和识别时间也分别减少到1/2和1/3。该比较是基于针对普通电视频道的视频源。还记录了4%的准确度增加(例如,F1分数增加0.02,F1分数是精度和召回率指标的组合)。

[0144] 改进的原因是多种多样的。首先,多层编码方案诸如VC-6允许非常有效和快速的解码(例如,如上文参考图5的实例所讨论的)。这意味着编码数据可以由远程处理设备非常快速地解码,并且可以获得对重构的不同分层的访问。可以通过在必要时仅使用较低层级的分层来实现在检测和识别中的加速。这可以允许检测算法和识别算法在低分辨率帧上更稳健地运行,这些低分辨率帧也没有高频细节,从神经网络架构的角度来看,这可能是“噪声”。还有一个有益效果是,此类处理更对操纵高分辨率像素内容的对抗性攻击免疫。

[0145] 更详细地,比较视频源可以包括全分辨率UHDp60视频源等。当使用比较视频编解码器对此视频源进行编码时,在全分辨率处生成重构的帧,并且需要将重构的帧缩小到适合视频分析功能的分辨率。此操作需要另外的时间,而使用多层编码方法诸如VC-6,由于编码的性质,可以直接访问较低数据梯队(即较低层)。当检测和识别是由在不同分辨率输入上运行的视频分析功能执行时,比较方法也可能受到影响,因为分辨率转换需要应用于视频或图像数据的解码帧,而使用多层方法,可以访问不同的分辨率。此外,在解码过程内使用放大的分层很容易允许放大到用户定义的分辨率和超分辨率放大(例如使用预测放大器)。因此,很容易支持不同的模块化视频分析功能。

[0146] 在某些实例中,可以在最高公因数分辨率处执行检测和/或识别。使用多层方法诸如VC-6,可以选择最接近此公因数分辨率的梯队。在某些情况下,对由视频分析的检测阶段输出的感兴趣区域执行识别,诸如对象或人脸识别。例如,人脸或对象检测功能可以输出定义边界框的参数,其可以用于标识感兴趣区域。可以通过后续的对象或人脸识别功能对此感兴趣区域进行分析(例如,可以识别在边界框内的对象或人脸)。在本实例中,可以仅需要对来自较低层重构的感兴趣区域进行上采样(例如,根据图5),并且仅需要接收和解码覆盖上采样的感兴趣区域的残差数据。此操作可以加速处理,因为相比较来说可能需要全图像重构,这通常是通过在最高质量下重构全帧并且将该全帧下采样到识别所需的分辨率来获得的。此外,使用多层方法诸如VC-6允许请求和传输针对感兴趣区域的进一步数据梯队,以便根据需要(例如,如果识别准确度低于阈值)添加另外的细节和分辨率。因此,使用本方法,可以在较低分辨率层上执行检测,但可以对使用一个或多个较高层重构的较高分辨率数据执行识别。通过对一个分量(诸如数据的Y平面或亮度平面)执行一个或多个视频分析功能(诸如此处讨论的检测和识别)可以进一步增强这些有益效果。同样,如果在Y或亮度平面数据上的准确度低,则可以请求并且模块化地传输和接收针对进一步颜色分量(诸如U或

V)的数据。

[0147] 实例方面

[0148] 根据本文描述的发明的一个方面,本公开提供了一种使用系统对输入信号内的一个或多个元素进行分类的方法。此方法包括第一步,接收在第一位置中的输入信号,并且使用基于分层的分级编码方法(诸如,通过非限制性实例的方式,编码标准SMPTE VC-6 ST-2117或MPEG-5第2部分LCEVC ISO/IEC 23094-2)对该信号进行压缩。该基于分层或层的编码方法包括生成在分级中的至少两组压缩数据,从而使得在解码时能够将该信号重构到相应的质量层级。在该方法中,将对应于该输入信号的第一组压缩数据传输到在第二位置中的分析系统。该分析系统通过对第一组压缩数据进行解码来将该信号的压缩版本解码到第一质量层级以便生成第一重构信号并且对该第一重构信号执行第一分类操作。至少部分地基于所述第一分类操作,可以标识该信号的相关区域。例如,通过非限制性实例的方式,这些相关区域可以包括在所述帧内的帧和感兴趣区域。这些相关区域可以指示在何处执行第二分类操作。响应于此,该分析系统执行从在第一位置中的系统的部分文件召回操作。这可以包括对另外的压缩数据(例如另外的数据分层或数据层)的请求。然后,在第一位置中的系统可以向在第二位置中的分析系统传输对应于该信号的所述相关区域的第二组压缩数据。然后,在第二位置中的分析系统可以通过对所述第二组压缩数据进行解码来将该信号的压缩版本解码到第二质量层级以便生成第二重构信号。然后,该分析系统可以对第二重构信号执行一个或多个第二分类操作。在本文描述的方面中,术语分类可以被视为更一般的信号分析的一个实例,并且术语“压缩”和“编码”可互换地使用,其中可以通过使用量化和/或对信号进行压缩的熵编码来对信号进行编码。

[0149] 根据第二方面,一种分析多个视频摄像机源的方法包括:

[0150] 在第一位置处,使用基于层的编码对该多个视频摄像机源进行编码,包括生成针对在该基于层的编码内的多个层中的每个层的编码数据流,其中在该多个层中的不同层对应于不同的空间分辨率,较高层表示较高空间分辨率;

[0151] 向相对于第一位置远程的第二位置传输针对该多个视频摄像机源的一个或多个最低层的编码数据流;

[0152] 在第二位置处,对该编码数据流进行解码来生成多个视频摄像机源的处于第一空间分辨率的一组重构;

[0153] 将一个或多个视频分析功能应用于该一组重构来标识用于进一步分析的一个或多个视频摄像机源;

[0154] 向针对所标识的用于进一步分析的一个或多个视频摄像机源的第一位置发送针对在该一个或多个最低层之上的一个或多个层的进一步编码数据流的请求;

[0155] 响应于该请求,向第二位置传输针对在该一个或多个最低层之上的一个或多个层的该进一步编码数据流;

[0156] 在第二位置处,对该进一步编码数据流进行解码来生成针对所标识的一个或多个视频摄像机源的处于第二空间分辨率的一组重构;并且

[0157] 将一个或多个视频分析功能应用于处于第二空间分辨率的该一组重构。

[0158] 根据第三方面,一种远程信号分析方法,包括:

[0159] 在计算设备处,通过至少一个网络接收针对对应于多个输入信号的多个编码信号

流的编码数据,其中在相对于该计算设备远程的编码位置处执行对该多个输入信号的编码,其中使用基于层的编码对该编码信号流进行编码,其中在该基于层的编码中的不同层对应于不同的质量层级,较高层表示较高质量层级,并且其中所接收到的编码数据包括与在该基于层的编码内的一个或多个最低层相关联的编码数据;

[0160] 对所接收到的编码数据进行解码来获得处于该一个或多个最低层的该多个输入信号的多个重构;

[0161] 处理该多个输入信号的重构中的至少一个重构来确定用于进一步处理的该编码信号流的部分;并且

[0162] 从该编码位置请求针对在该一个或多个最低层之上的与该编码信号流的所确定的部分相关联的一个或多个层的编码数据。

[0163] 可以由图9的云数据中心920、图10的远程服务器1060或图1120的分布中心1120以及其它实现方式中的一者或多者应用此方法。

[0164] 根据第四方面,一种编码方法,例如在第一位置处执行,该方法包括:

[0165] 在第一位置处,生成针对对应于多个输入信号的多个编码信号流的编码数据,其中使用基于层的编码对该编码信号流进行编码,其中在该基于层的编码中的不同层对应于不同的质量层级,较高层表示较高质量层级,并且其中所接收到的编码数据包括与在该基于层的编码内的一个或多个最低层相关联的编码数据;

[0166] 向相对于第一位置远程的第二位置传输该编码数据用于解码;

[0167] 接收针对在该一个或多个最低层之上的一个或多个层的进一步编码数据的请求,该请求与该多个输入信号中的一个或多个输入信号相关;并且

[0168] 向第二位置传输该进一步编码数据。

[0169] 例如,可以由图9和图10的编码平台940或图11的边缘机箱应用此方法。

[0170] 可选地,在上述方法中的一种或多种方法中,当在第一位置中的系统中检测到用于向在第二位置中的系统的数据传输的可用带宽不足以传输所有第一组压缩数据时,该方法包括传输第一组压缩数据的子集,其中在第二位置中的系统基于所接收到的第一组压缩数据的子集来重构处于第一质量层级的信号。在非限制性实施例中,当该系统接收第一组压缩数据的子集时产生的处于第一质量层级的信号的渲染的分辨率和颜色平面是与当该系统接收第一全组压缩数据时产生的渲染的分辨率和颜色平面相同的。

[0171] 可选地,在上述方法中的一种或多种方法中,至少部分地基于在第二位置中的系统中的视频分析任务的输出,该方法包括向在第一位置中的系统发信号来传输针对与处于第三质量层级的重构相对应的信号的压缩数据,该第三质量层级高于第一质量层级。

[0172] 可选地,在上述方法中的一种或多种方法中,至少部分地基于在第二位置中的系统中的视频分析任务的输出,该方法包括向在第一位置中的系统发信号来传输针对与处于第四质量层级的重构相对应的信号的压缩数据,该第四质量层级低于第一质量层级。

[0173] 可选地,在上述方法中的一种或多种方法中,在第二位置中的系统对在第一位置中可用的多个信号执行根据本文描述的方法的分析和/或分类,并且标识将在第三质量层级处传输的一个或多个信号,该第三质量层级高于第一质量层级,并且标识将在第四质量层级处传输的一个或多个信号,该第四质量层级低于第一质量层级。此操作是有利的,因为该分析系统可以实时控制哪些信号(基于该分类的输出,在任何一个时间判断为较高相关

性)将在较高保真度下传输到第二位置以及哪些信号(基于该分类的输出,在任何一时间判断为较低相关性)将在代理质量下传输,从而优化可用网络容量的使用。

[0174] 可选地,在上述方法中的一种或多种方法中,当额外的网络容量可用时,该方法包括向第二位置传输与尚未出于视频分析任务的目的而传输的信号有关的压缩数据组。此操作有利于存档目的和/或未来的分类任务,这将能够受益于在现场可用的所有压缩数据。

[0175] 可选地,在上述方法中的一种或多种方法中,使用相同的处理单元和本地存储器(例如,通过非限制性实例的方式,一组一个或多个图形处理单元)例如在远程处理位置处的处理单元和一组存储器来执行解码和执行信号分析的步骤。

[0176] 可选地,在上述方法中的一种或多种方法中,视频分析任务包括分类并且以分级方式(例如,通过非限制性实例的方式,利用贝叶斯条件概率分析方法)组织第一分类、一个或多个第二分类和一个或多个第三分类,使得第一分类是在该输入信号中的某些元素的粗分类,并且该一个或多个第二分类或该一个或多个第三分类进一步细化在该输入信号中的元素的粗分类。更可选地,该方法可以进一步包括,基于该粗分类,仅选择、检索并且解码第二组压缩数据的被确定需要进一步分类的那些部分,并且基于那些解码的部分来生成第二重构信号。

[0177] 可选地,在上述方法中的一种或多种方法中,使用一个或多个神经网络检测器(例如,多个神经网络检测器)执行一个或多个分类。可选地,神经检测器包括数据归一化阶段和后续的数据比较阶段。可选地,神经网络检测器是以数字硬件(例如,现场可编程门阵列(FPGA)、被配置成执行软件产品以实施神经网络功能的计算硬件)或两者的组合来实施的。

[0178] 根据第五方面,提供了一种对在输入信号内的一个或多个元素进行分类的方法(即,一种用于对在输入信号内的一个或多个元素进行分类的方法),该方法包括:

[0179] 对该输入信号的第一版本执行分类任务的第一分类子任务;

[0180] 至少部分地基于第一分类子任务的输出,标识第二组压缩数据并且从第一位置召回第二组压缩数据;

[0181] 至少部分地基于所述第二组压缩数据,重构该输入信号的第二版本;

[0182] 对该输入信号的第二版本执行该分类任务的一个或多个第二分类子任务,

[0183] 其中该信号的第一版本和第二版本是从远程位置递增获取的。

[0184] 可选地,上述方法进一步包括对该输入信号的第三版本执行该分类任务的一个或多个第三分类子任务,所述第三版本是基于从第一位置召回的第三组压缩数据生成的,所述第三组压缩数据是基于一个或多个第二分类子任务的输出标识的。可选地,在上述方法中,基于第一分类子任务的输出,第二版本仅包括该输入信号的一部分。

[0185] 可选地,在上述方法中的一种或多种方法中,检索到的第二组压缩数据包括的数据仅为针对构成该输入信号的颜色平面的子集的数据。在另一个实例中,该信号的第二版本仍包括构成该信号的第一版本的所有颜色平面。

[0186] 可选地,在上述方法中的一种或多种方法中,对不同质量层级的提及对应于不同的空间分辨率,较高层表示较高空间分辨率。

[0187] 可选地,在上述方法中的一种或多种方法中,该基于层的编码包括处于最低空间分辨率的基础层和在该基础层之上的处于较高空间分辨率的多个残差层,在该多个残差层中的给定残差层是基于对从处于针对该层的分辨率的输入信号导出的数据与从较低层导

出的数据之间的比较来确定的。

[0188] 可选地,在上述方法中的一种或多种方法中,存在以下进一步的步骤:确定在处于一个或多个最低层的多个输入信号的重构的至少一个重构中的感兴趣区域并且请求与该感兴趣区域相关的进一步编码数据。这可以包括使用该进一步编码数据来生成针对处于在该一个或多个最低层之上的一个或多个层的感兴趣区域的一组重构。在一种情况下,这可以包括确定生成处于最高质量层级的重构所需的差分编码数据,该差分编码数据不包括所接收到的编码数据和所接收到的进一步编码数据,并且从第一位置请求差分编码数据。

[0189] 可选地,在上述方法中的一种或多种方法中,该方法包括在处于该一个或多个最低层的该多个输入信号的重构的至少一个重构内执行对象检测。在某些情况下,该方法可以包括在处于该一个或多个最低层之上的一个或多个层的该多个输入信号的重构的至少一个重构内执行对象检测。

[0190] 可选地,在上述方法中的一种或多种方法中,该方法进一步包括标识针对该多个输入信号中的一个或多个输入信号的一组或多组帧用于进一步分析,以及标识一个或多个输入源用于进一步分析,其中在第一编码位置和第二分析位置中的一者或多者处执行所述标识。在此情况下,该方法可以进一步包括传输针对在该一个或多个最低层之上的层的编码数据用于所标识的帧或所标识的输入源。

[0191] 可选地,在上述方法中的一种或多种方法中,该方法进一步包括将编码数据传输到远程服务器用于信号分析,并且将编码数据传输到多个用户。将编码数据传输到多个用户可以包括将编码数据传输到分布中心用于分布给该多个用户。这可以包括在该远程服务器处对一组重构进行分析来确定将传输到该分布中心的针对一个或多个输入源的编码数据的一组层。

[0192] 可选地,在上述方法中的一种或多种方法中,该多个输入信号或多个视频输入源与用于捕获现场事件的多个视频摄像机相关联,其中该远程服务器确定将选择该多个视频摄像机中的哪一个视频摄像机用于传输到该多个用户,并且确定针对将传输到该多个用户的编码数据的质量层级用于所选择的视频摄像机。

[0193] 可选地,在上述方法中的一种或多种方法中,每个质量层级具有不同的编码位率。

[0194] 可选地,在上述方法中的一种或多种方法中,从第一编码位置传输的编码数据包括用于编码信号分量的子集的编码数据。这些编码信号分量可以包括该视频信号的不同颜色分量,并且编码分量的子集可以至少包括明度或亮度分量。

[0195] 可选地,在上述方法中的一种或多种方法中,该方法进一步包括在第二远程位置处接收针对在该一个或多个最低层之上的一个或多个层的编码数据用于存档。

[0196] 可选地,在上述方法中的一种或多种方法中,所述方法进一步包括对所接收到的编码数据进行解码来获得处于该一个或多个最低层的该多个输入信号的多个重构,放大该多个重构,并且对所放大的重构执行分析。例如,图5中的上采样器526、522和530可以包括用于预测上采样的重构520、528和531的内容的神经网络上采样器。在某些情况下,即使在没有针对每个层的另外的残差数据(518、526等)的情况下,上采样的重构517、524和/或532也可以提供改进的视频分析。在此情况下,图5的表示多层信号的解码的布置可以有利于容纳具有用于推理的固定输入分辨率的不同模块化视频分析功能(例如,用于分类的神经网络架构)。当存在网络拥塞或带宽以其他方式受限时,这可以是有利的,即响应于针对处于

该较高分辨率以及该较高分辨率之上的一个或多个层的编码数据是不可用的而执行该放大。在这些情况下,可以仍在固定的质量层级处应用视频分析功能;当带宽可用时,可以再次对针对一个或多个层的残差数据进行编码、接收和解码,并且从而改进用于视频分析功能的输出。在这些情况下,放大可选地使用在训练集上训练的神经网络架构,该训练集包括在图像上处于两个空间分辨率的版本,以便学习参数值来将处于较低分辨率的重构映射到处于较高分辨率的重构。

[0197] 可选地,在上述方法中的一种或多种方法中,该多层编码包括基于SMPTE VC-6 2117的编码。

[0198] 可选地,在上述方法中的一种或多种方法中,该多层编码包括基于ISO/IEC MPEG-5第2部分LCEVC的编码。

[0199] 上文列出的适当方法可以由在第一位置处的编码系统实施。这些方法中的其他一些方法可以由在第二位置处的远程分析系统实施。

[0200] 本公开的实施例基本上消除或至少部分地解决了现有技术中的上述问题,并且使分布式深度学习分类系统能够对在不同位置中在全分辨率和质量下可用的信号执行准确和有效的分析,而所述分类不需要将所有这些信号传输到分类系统的位置。

[0201] 本文描述的某些实例允许非常有效地编码一组信号源,例如在事件场地处于接近无损UHDp60的全质量的多个所有摄像机。使用这些实例,可以仅向远程场地实时传输与达到给定LOQ的质量相对应的字节。在VC-6实例中,这可以是LOQ-4(负4),并且可能仅传输一个颜色分量诸如亮度。在远程场地(诸如在远程处理设备和/或在云数据中心或分布中心处的处理设备)处,可以对这些较低层级“代理视频”(即来自较低层级视频的重构)执行有效的检测传递,并且可以基于自动分析算法标识感兴趣的帧和区域。任何标识的感兴趣部分都可以用于从事件场地获取一组残差或差分字节,这些字节是具有较高质量和/或分辨率重构所必需的,来允许对处于高质量层级(例如,达到原始全质量或全分辨率信号源)的重构的部分执行进一步视频分析。

[0202] 此处描述的某些实例可以具有针对远程/家庭生产应用、媒体档案分析和/或分布、工业应用、安全/监视应用和/或物联网应用等的应用。这些实例避免了在需要“在现场”对质量非常差的源执行高级视频分析还是在云中对质量非常差的源执行高级视频分析之间做出错误选择。相反,本实例提供了一种部分文件召回的机制,允许以可能实时交互的方式“从远程”增量地执行分析,这可以很容易地并入并行化的AI处理管道中。

[0203] 用于实施边缘或远程处理设备的示例性装置

[0204] 参考图12,示出了装置1200的实例的示意性框图。装置1200可以用于实施执行如本文描述的编码操作的边缘计算设备和执行如本文描述的远程信号分析的远程计算设备中的一者或多者。在一种情况下,在边缘和远程(即“云”)位置中的一者或多者处可以存在多个装置1200。

[0205] 装置1200的实例包括但不限于移动计算机、个人计算机系统、无线设备、基站、电话设备、台式计算机、膝上型计算机、笔记本电脑、上网本计算机、大型计算机系统、手持式计算机、工作站、网络计算机、应用服务器、存储设备、消费型电子设备诸如摄像机、摄影机、移动设备、视频游戏机、手持式视频游戏设备、外围设备诸如交换机、调制解调器、路由器、车辆等或通常任何类型的计算或电子设备。在优选实例中,该装置包括服务器计算设备。

[0206] 在此实例中,装置1200包括被配置成处理信息和/或指令的一个或多个处理器1201。一个或多个处理器1201可以包含中央处理单元(CPU)。一个或多个处理器1201与总线1202耦合。由一个或多个处理器1201执行的操作可以由硬件和/或软件实行。一个或多个处理器1201可以包括多个协同定位的处理器或多个全异地定位的处理器。

[0207] 在此实例中,装置1200包括被配置成存储用于一个或多个处理器1201的信息和/或指令的计算机可用易失性存储器1203。计算机可用易失性存储器1203与总线1202耦合。计算机可用易失性存储器1203可以包括随机存取存储器(RAM)。

[0208] 在此实例中,装置1200包括被配置成存储用于一个或多个处理器501的信息和/或指令的计算机可用非易失性存储器1204。计算机可用非易失性存储器1204与总线1202耦合。计算机可用非易失性存储器1204可以包括只读存储器(ROM)。

[0209] 在此实例中,装置1200包括被配置成存储信息和/或指令的一个或多个数据存储单元1205。一个或多个数据存储单元1205与总线1202耦合。一个或多个数据存储单元1205可以例如包含磁盘或光盘和盘驱动器或固态驱动器(SSD)。

[0210] 在此实例中,装置1200包括被配置成向和/或从一个或多个处理器1201传送信息的一个或多个输入/输出(I/O)设备1206。一个或多个I/O设备506与总线1202耦合。一个或多个I/O设备506包括至少一个网络接口。该至少一个网络接口使装置1200能够通过一个或多个数据通信网络进行通信。数据通信网络的实例包括但不限于广域网(WAN诸如因特网)和局域网(LAN)。许多数据通信链路可以包括通过如本领域已知的多个互连网络的通信。一个或多个I/O设备1206可以使用户能够通过一个或多个输入设备(未示出)向装置1200提供输入。一个或多个I/O设备1206可以使信息能够通过一个或多个输出设备(未示出)提供给用户。

[0211] 针对装置1200示出了各种其他实体。例如,当存在时,操作系统1207、数据处理模块1208、一个或多个进一步模块1209和数据1210被示为驻留在计算机可用易失性存储器1203、计算机可用非易失性存储器1204和一个或多个数据存储单元1205中的一者或它们的组合中。信号处理模块1208可以通过存储在计算机可用非易失性存储器1204内的存储器位置中、一个或多个数据存储单元1205内的计算机可读存储介质中和/或其他有形计算机可读存储介质中的计算机程序代码来实施。有形计算机可读存储介质的实例包括但不限于:光学介质(例如,CD-ROM、DVD-ROM或蓝光)、闪存卡、软盘或硬盘或能够在至少一个ROM或RAM或可编程ROM(PROM)芯片中存储诸如固件或微代码等计算机可读指令的或作为专用集成电路(ASIC)的任何其他介质。

[0212] 因此,装置1200可以包括可以由一个或多个处理器1201执行的数据处理模块1208。数据处理模块1208可以被配置成包括指令来实施本文描述的操作中的至少一些操作。在操作期间,一个或多个处理器1201启动、运行、执行、解释或以其他方式执行在数据处理模块1208中的指令。

[0213] 尽管本文参考附图描述的实例的至少一些方面包含在处理系统或处理器中执行的计算机过程,但是本文描述的实例还扩展到计算机程序,例如载体上或载体中的计算机程序,其适于将实例付诸实践。载体可以是任何能够承载程序的实体或设备。

[0214] 应当理解,装置1200可以包括更多、更少和/或与图12中所示的组件不同的组件。在边缘位置和云位置中的每一个位置处,装置1200可以被实施为单一实体或分布式计算平

台。例如，远程处理设备可以包括与高速数据通信链路（例如，与在边缘和远程处理设备之间的更有限的链路相比）互连的多个处理设备。类似地，在边缘位置处的处理可以由多个互连的设备（例如，通过LAN耦合）执行。

[0215] 本文描述的技术可以通过软件或硬件实现，或者可以使用软件和硬件的组合来实现。它们可以包含配置装置以实施和/或支持本文描述的任何或所有技术。

[0216] 以上实施例应理解为说明性实例。设想另外的实施例。所描述的选项和特征的不同组合，包括某些特征的省略，是明确允许的，并且可以在不同的实现方式中提供。例如，应当理解，关于任何一个实例描述的任何特征可单独使用，或者与描述的其他特征结合使用，并且还可以与这些实例中的任何其他实例的一个或多个特征结合使用，或者与这些实例中的任何其他实例的任何组合结合使用。此外，在不脱离由所附权利要求书限定的本发明的范围的情况下，还可采用上文未描述的等效者和修改。

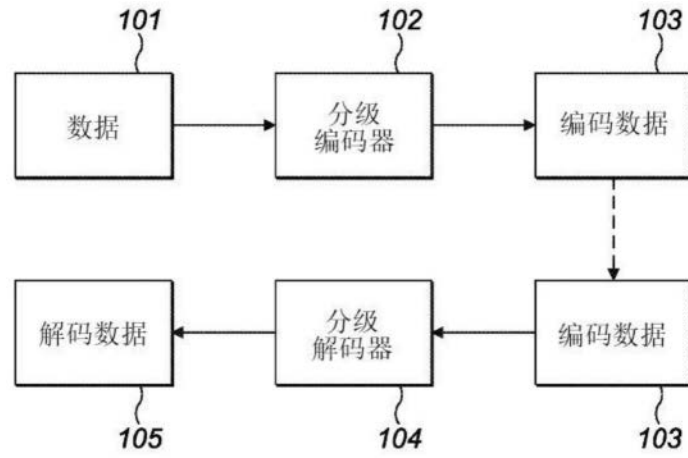


图1

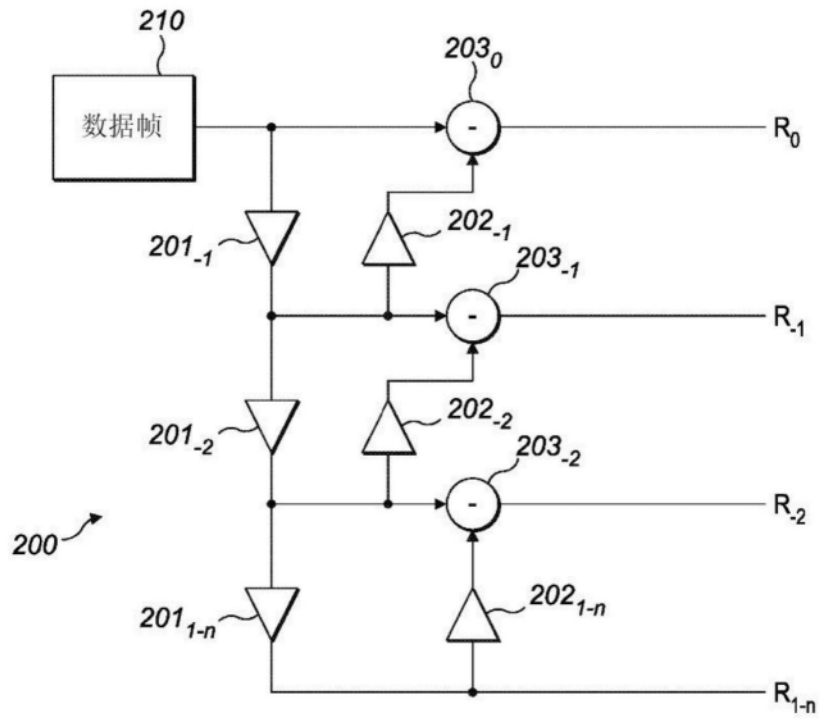


图2

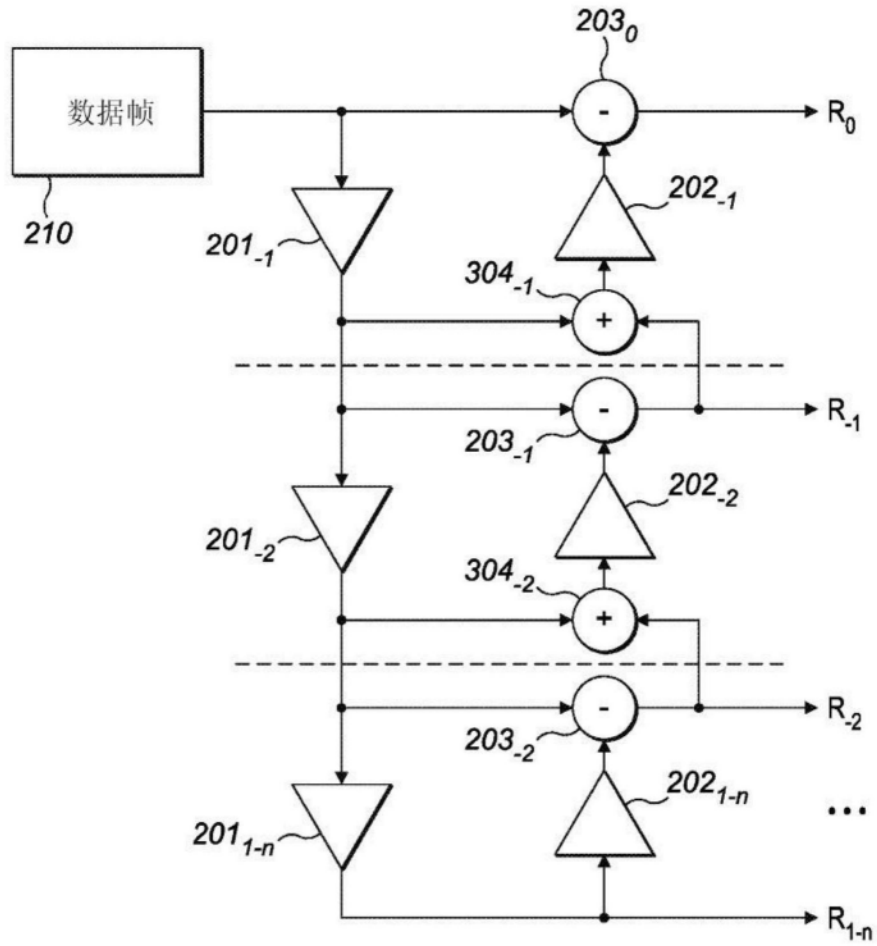


图3

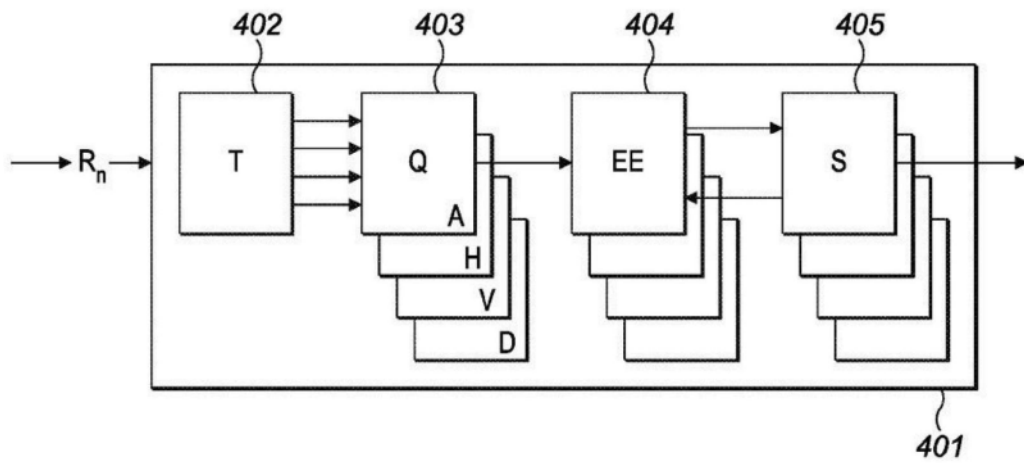


图4

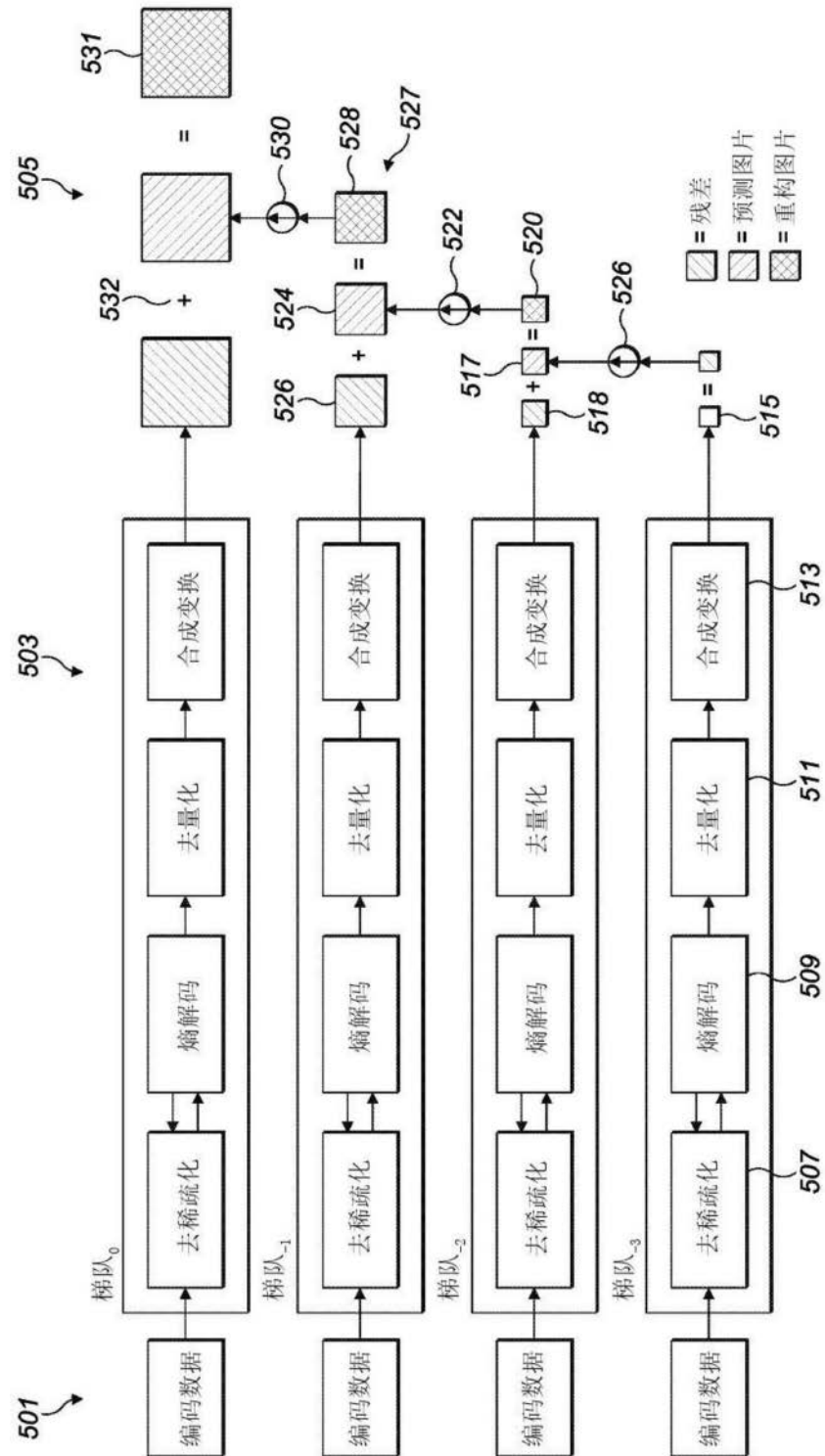


图5

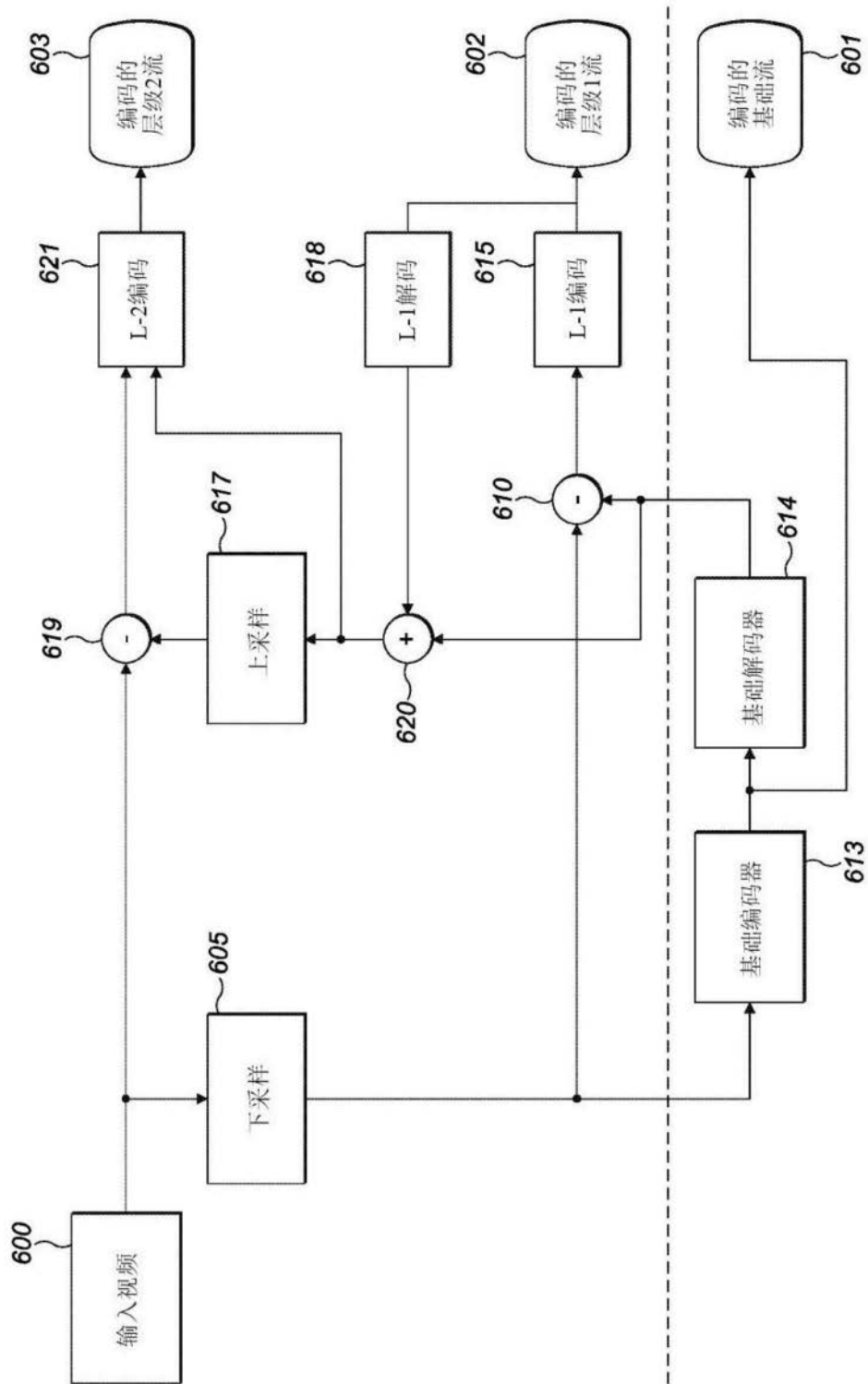


图6

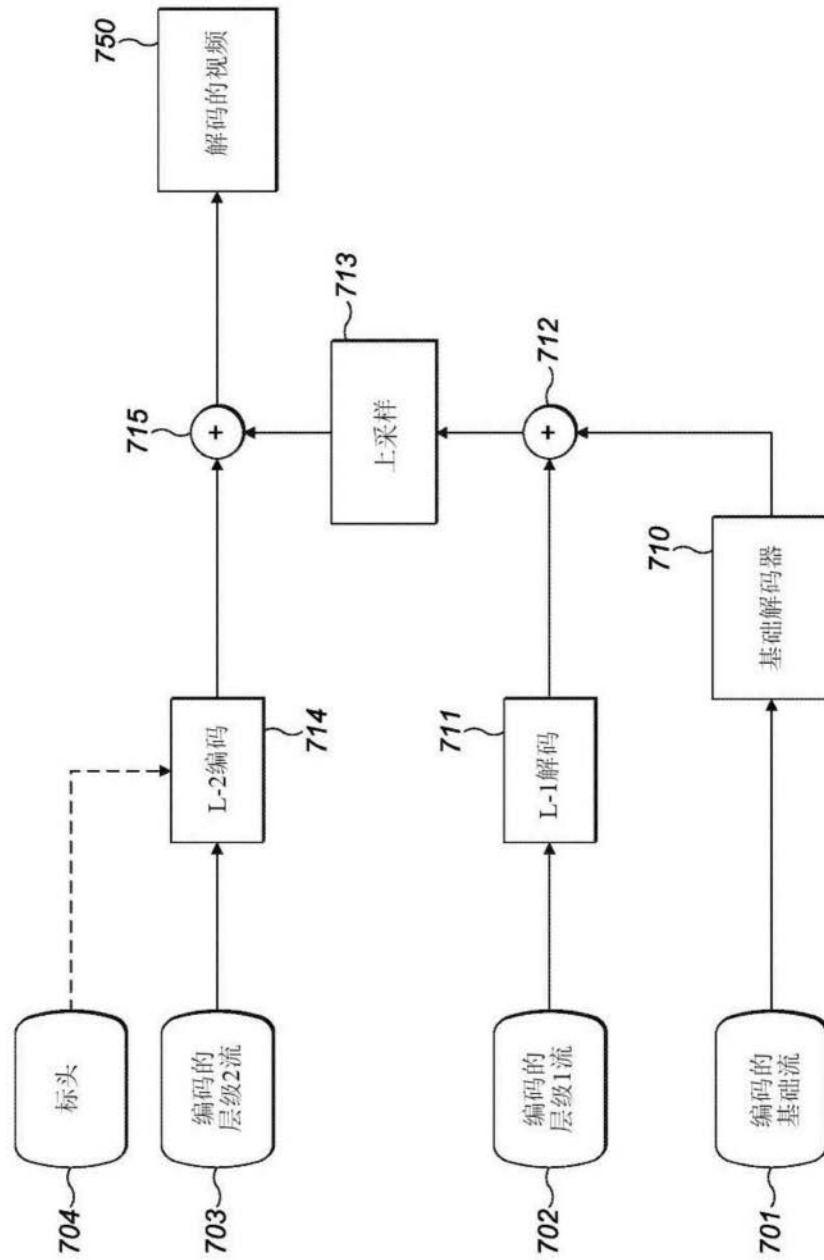


图7

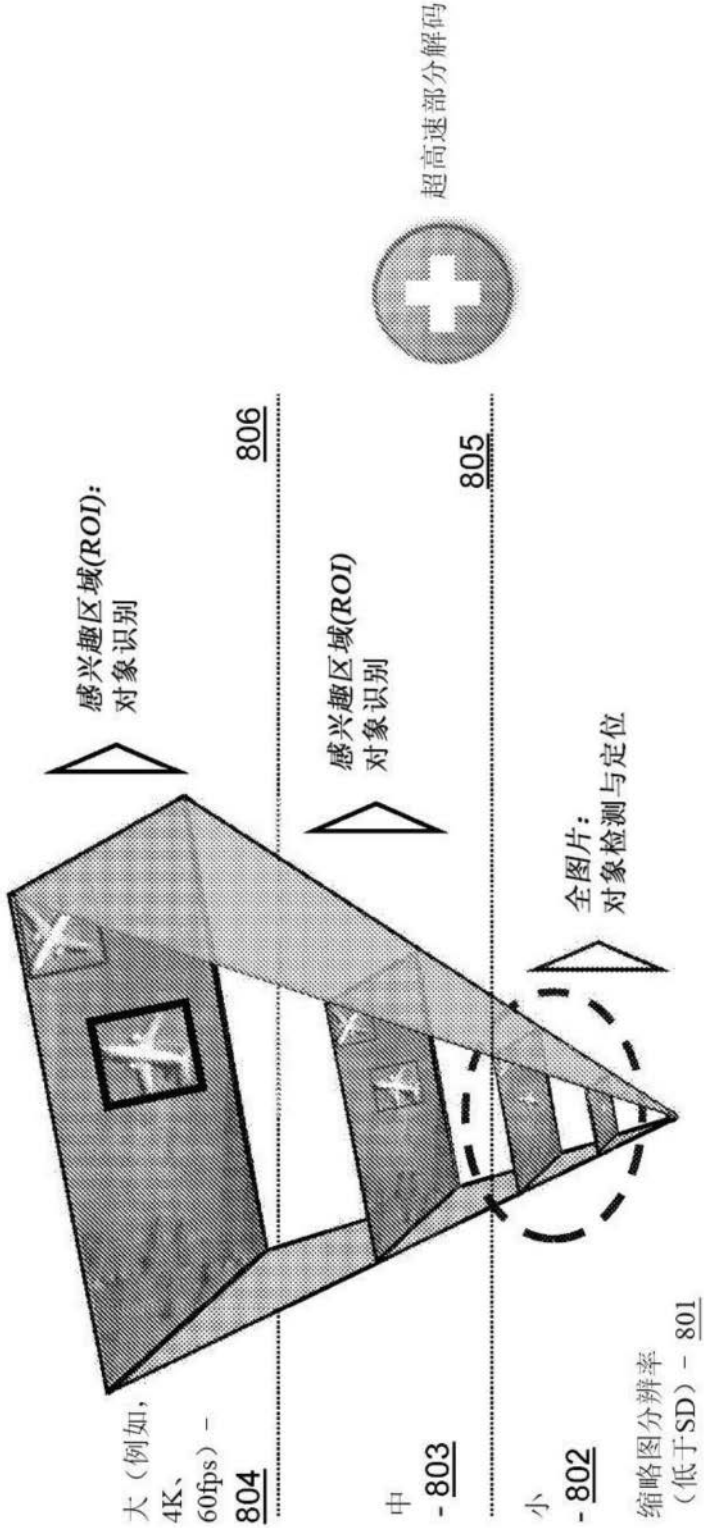


图8

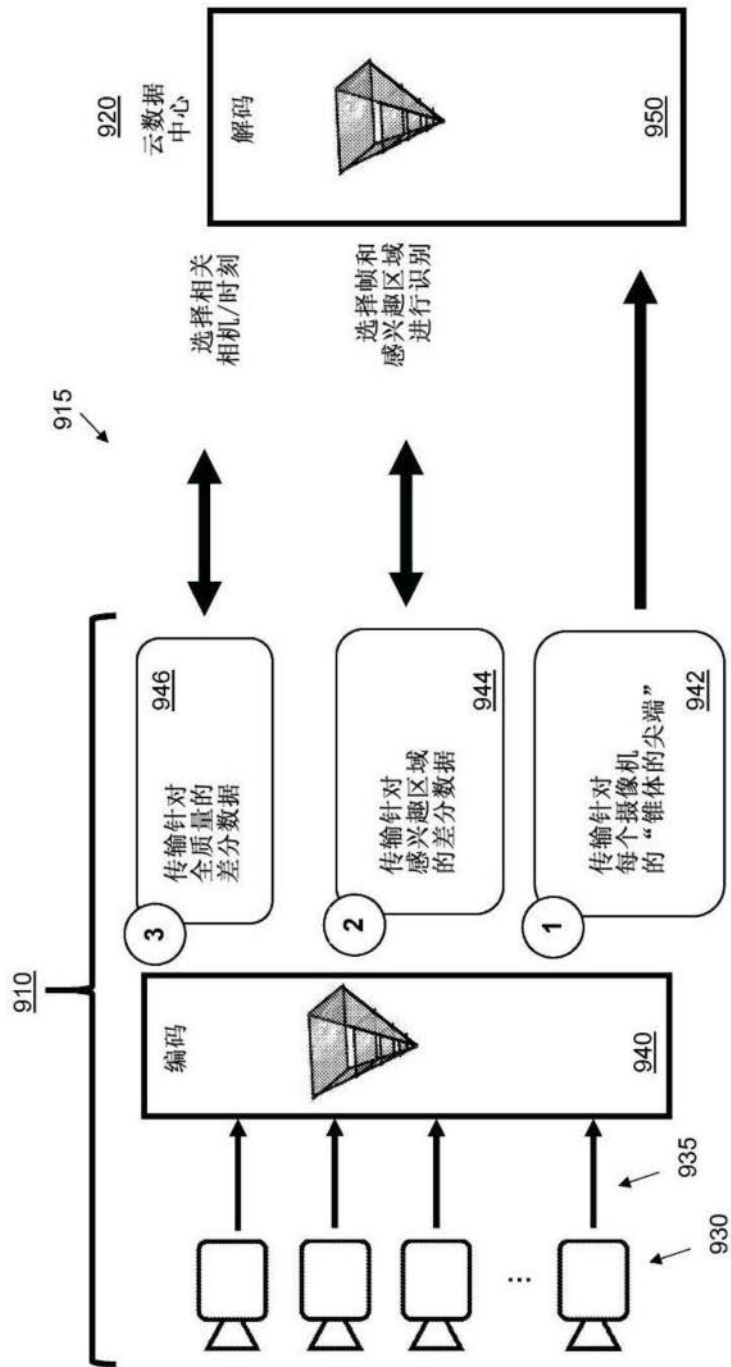


图9

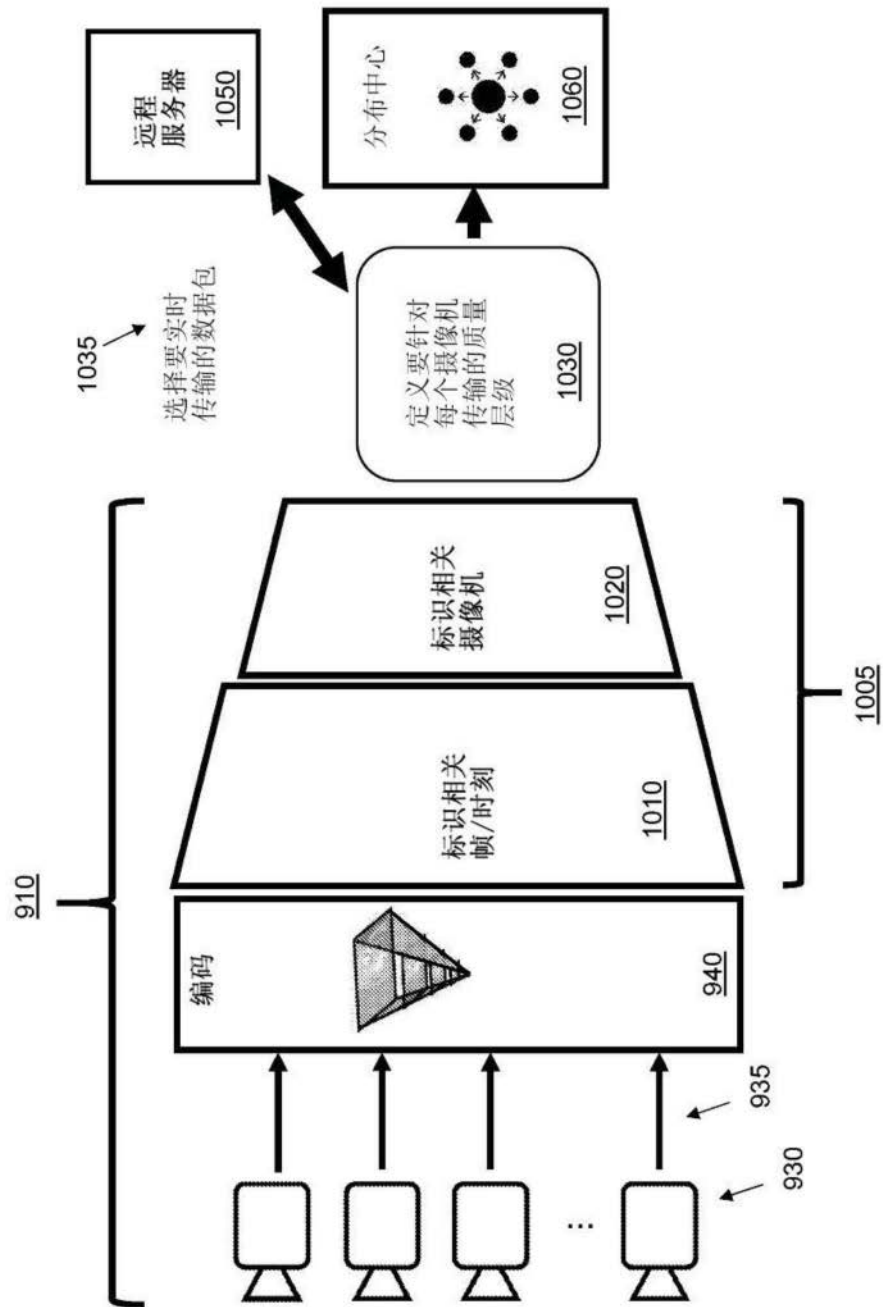


图10

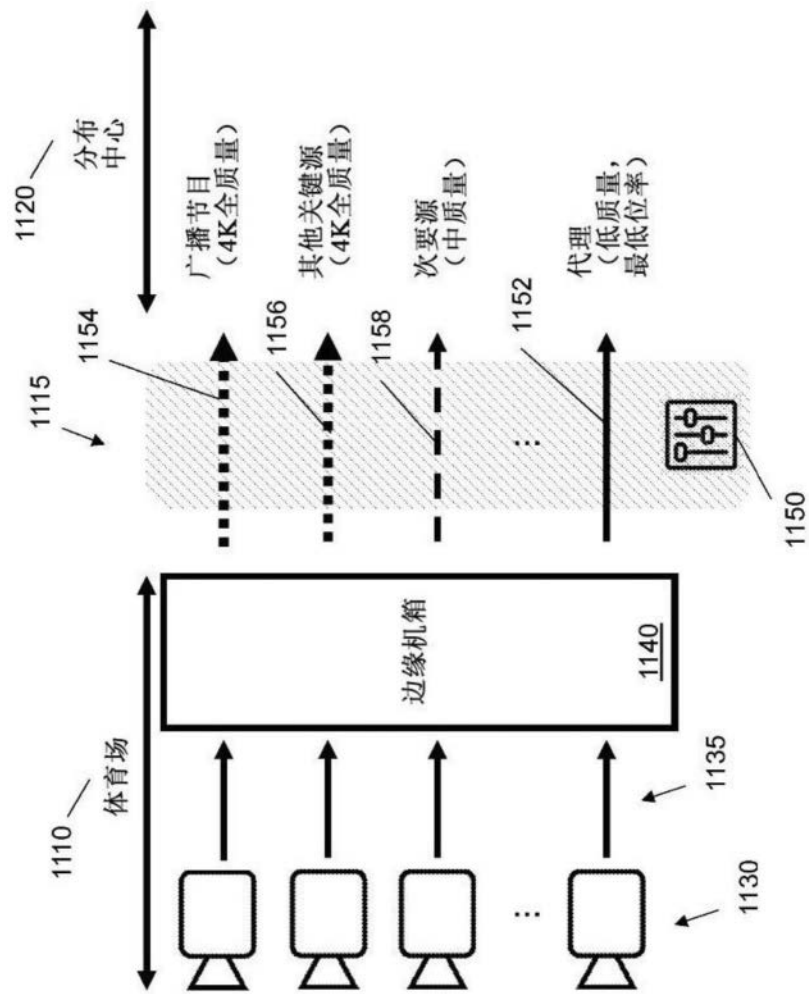


图11

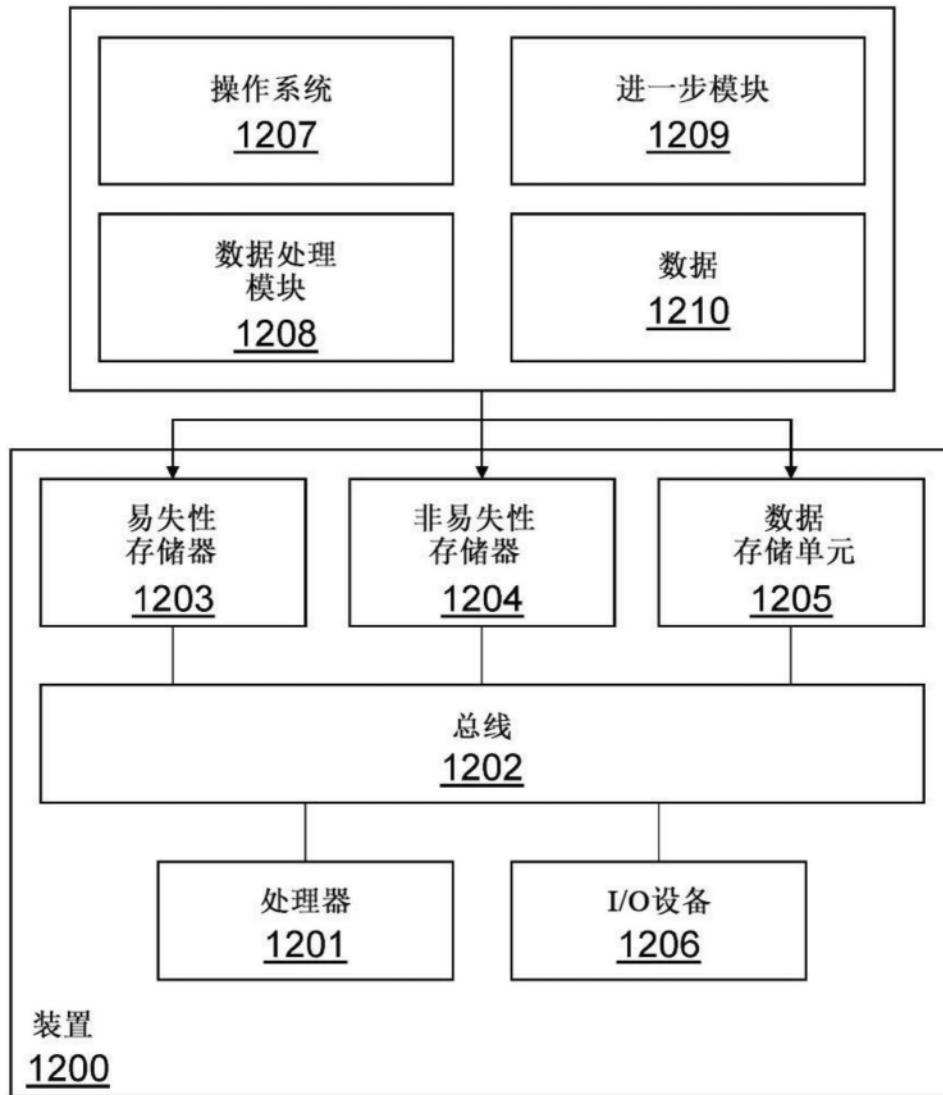


图12

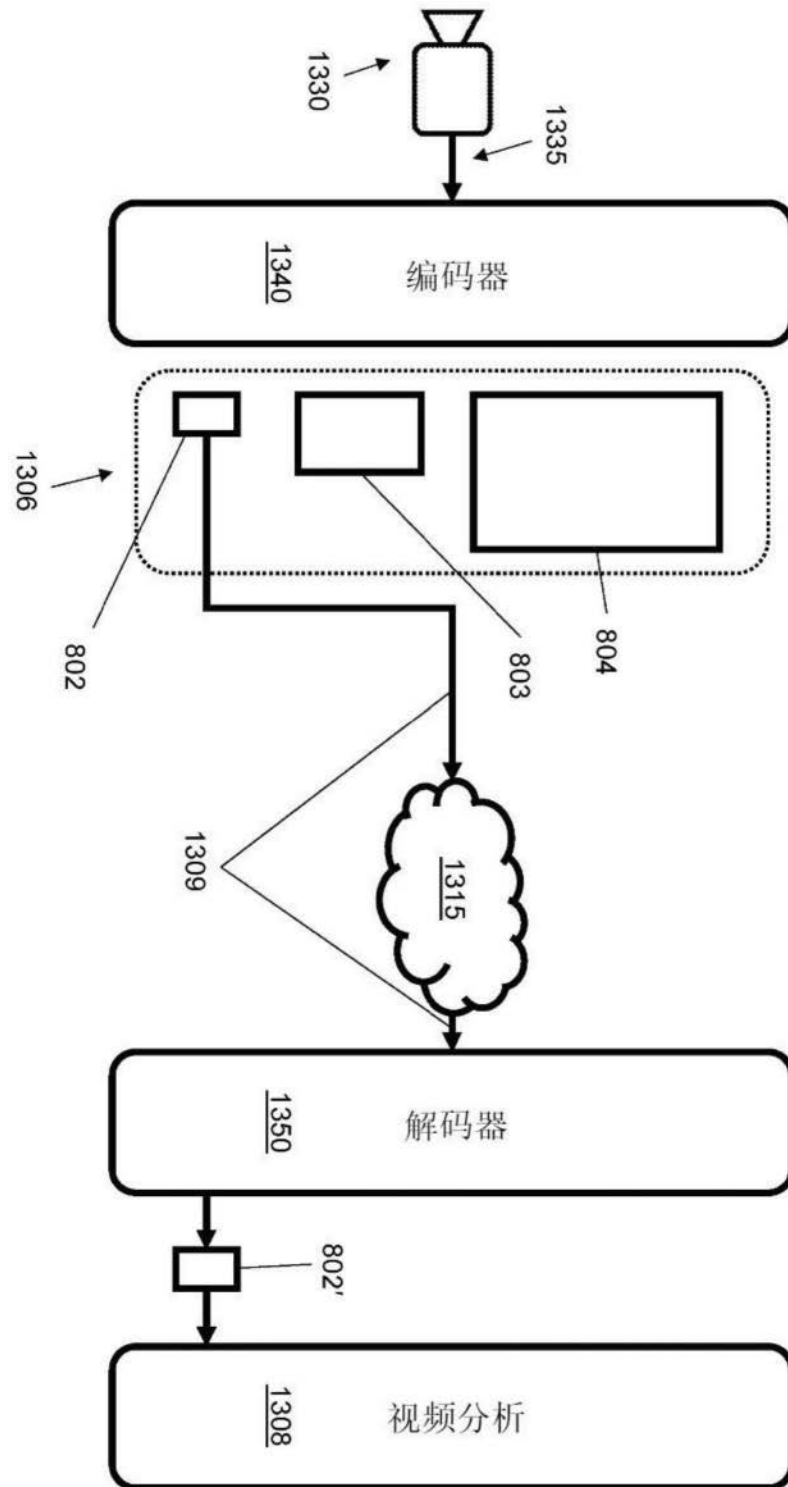


图13A

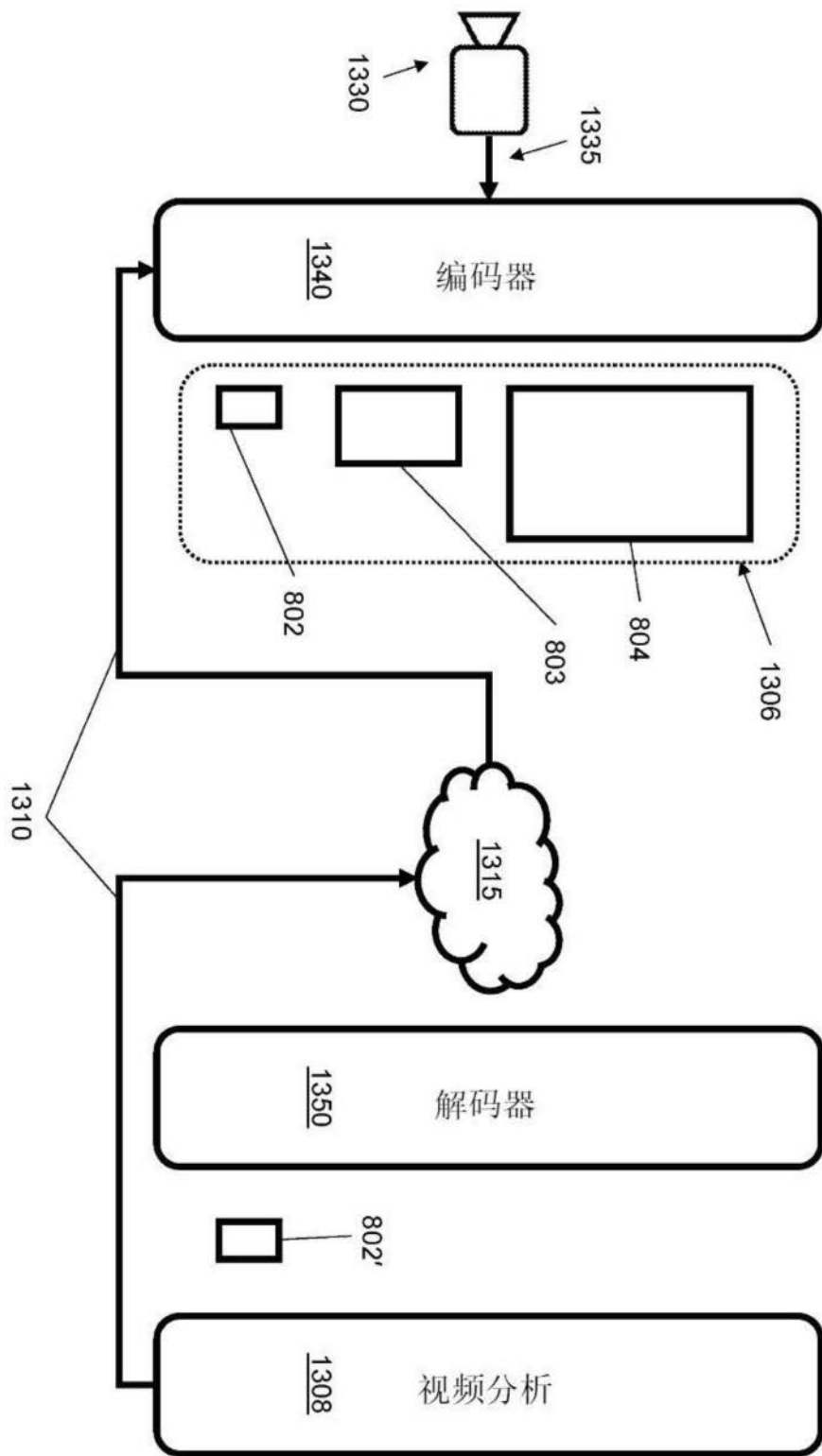


图13B

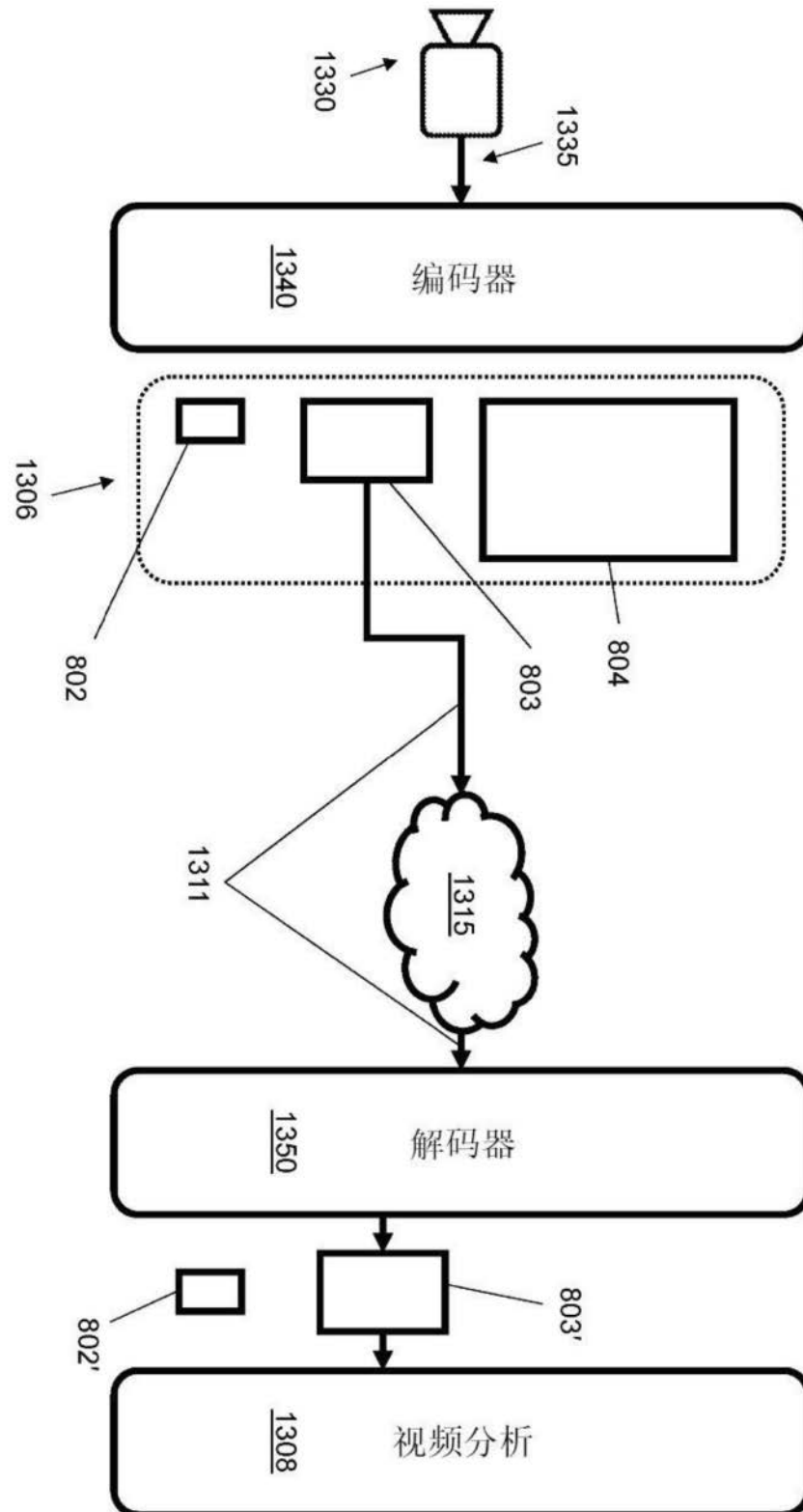


图13C

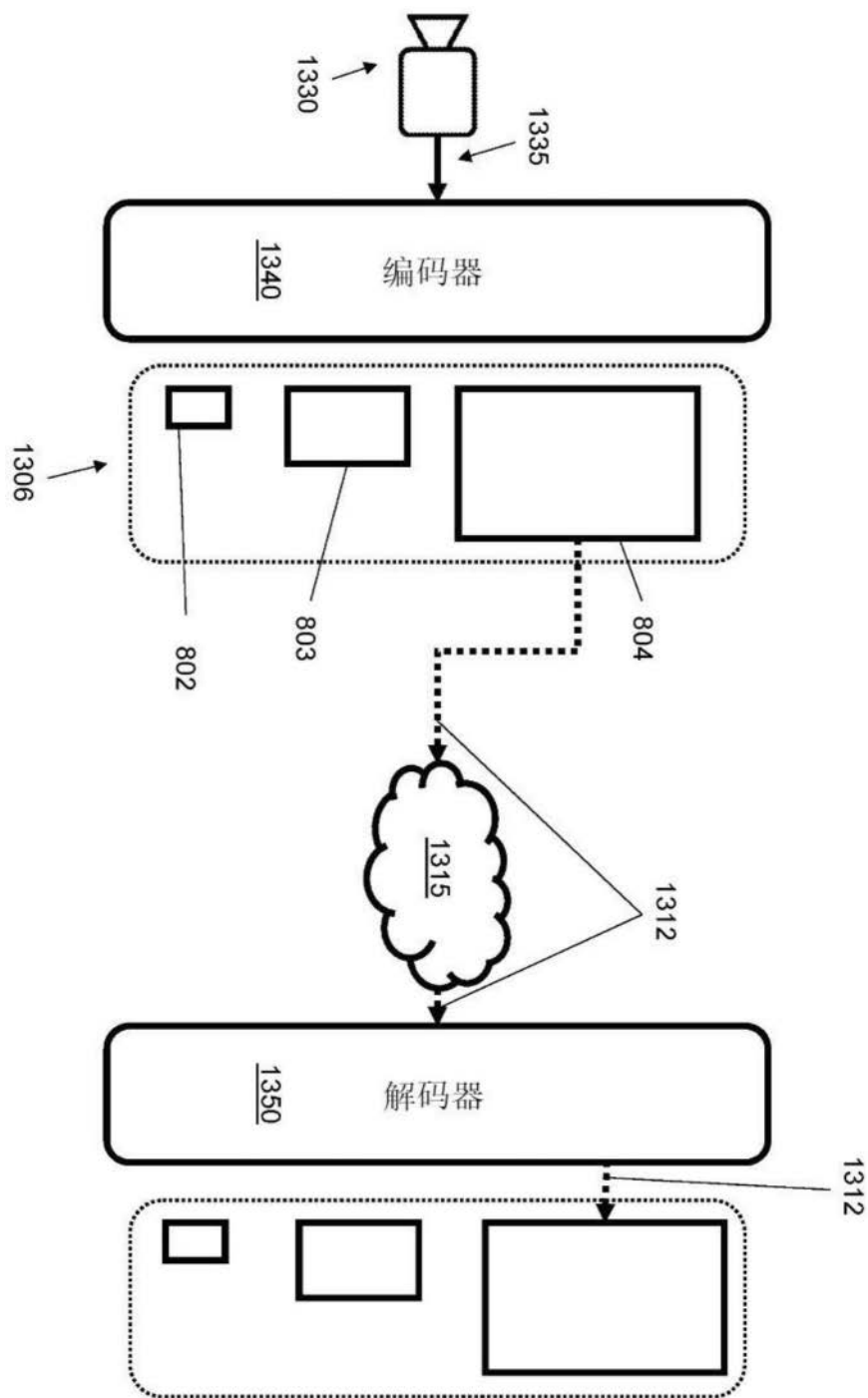


图13D

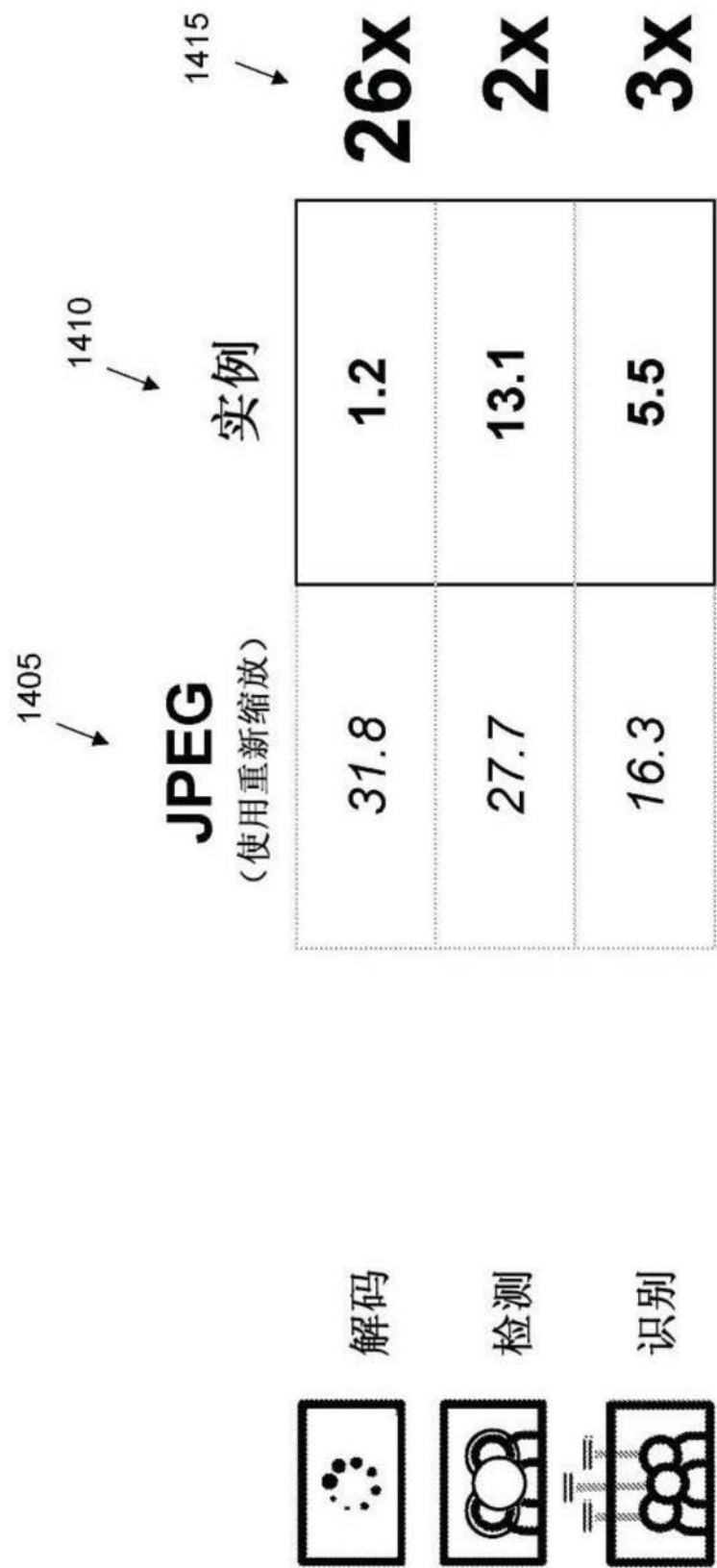


图14