



(12) 发明专利

(10) 授权公告号 CN 110099278 B

(45) 授权公告日 2022.10.14

(21) 申请号 201910433376.4

(51) Int.CI.

(22) 申请日 2014.12.19

H04N 19/109 (2014.01)

(65) 同一申请的已公布的文献号

H04N 19/136 (2014.01)

申请公布号 CN 110099278 A

H04N 19/139 (2014.01)

(43) 申请公布日 2019.08.06

H04N 19/147 (2014.01)

(30) 优先权数据

H04N 19/17 (2014.01)

61/925,090 2014.01.08 US

H04N 19/177 (2014.01)

61/934,574 2014.01.31 US

H04N 19/179 (2014.01)

14/513,132 2014.10.13 US

H04N 19/523 (2014.01)

(62) 分案原申请数据

201480072763.5 2014.12.19

(56) 对比文件

US 2010166079 A1, 2010.07.01

(73) 专利权人 微软技术许可有限责任公司

US 2012008686 A1, 2012.01.12

地址 美国华盛顿州

KR 20110017302 A, 2011.02.21

(72) 发明人 G·J·沙利文 Y·周 李明杰

CN 101658044 A, 2010.02.24

C-L·林

US 2011251431 A1, 2011.10.13

(74) 专利代理机构 永新专利商标代理有限公司

WO 2013002716 A2, 2013.01.03

72002

CN 102007769 A, 2011.04.06

专利代理人 张立达 王英

US 2012051431 A1, 2012.03.01

(续)

US 2013057646 A1, 2013.03.07

审查员 李景芳

权利要求书3页 说明书24页 附图8页

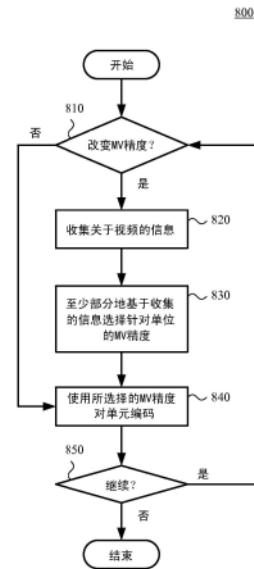
(54) 发明名称

选择运动向量精度

(57) 摘要

提出了在视频编码期间选择运动向量 (“MV”) 精度的方法。这些方法可以便于在率失真性能和/或计算效率方面有效的压缩。例如，视频编码器从多个MV精度中确定针对视频的单位的MV精度，其包括一个或多个分数样本MV精度和整数样本MV精度。视频编码器可以识别具有分数样本MV精度的一组MV值，然后至少部分地基于分数部分为零的MV值 (该组内) 的流行，选择针对单位的MV精度。或者，视频编码器可以执行率失真分析，其中率失真分析朝向整数样本MV精度偏置。或者，视频编码器可以收集关于视频的信息，并至少部分地基于收集到的信息选择针对所述单位的MV精度。

B
CN 110099278 B



[转续页]

[接上页]

(56) 对比文件

CN 102187668 A, 2011.09.14

CN 103202016 A, 2013.07.10

US 2009257668 A1, 2009.10.15

Jicheng An et.al. Progressive MV
Resolution.《JC-VC MEETING, Torino》.2011,
陈文洪等.H.264视频编码在数码监控系统
中的应用研究.《安防科技》.2007, (第06期),

1. 一种计算机系统,包括一个或多个处理单元和存储器,其中,所述计算机系统实现视频编码器,所述视频编码器被配置为执行包括以下各项的操作:

对视频序列的帧进行编码以产生编码数据,所述帧中的每一个帧包括一个或多个区域,其中,所述编码包括:

在报头中设置指示符的值,所述报头应用于所述视频序列的所述帧中的至少一个帧,所述指示符指示:

针对所述帧中的所述至少一个帧的运动向量(“MV”)精度是否是在逐区域的基础上被控制的;以及

如果针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,则针对所述帧中的所述至少一个帧的所述MV精度是四分之一样本精度还是整数样本精度;以及

如果针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的,则针对所述帧中的所述至少一个帧的所述一个或多个区域中的每个区域,在针对所述区域的报头中设置标记的值,所述标记指示针对所述区域的MV精度是四分之一样本精度还是整数样本精度;以及

输出所述编码数据作为比特流的一部分,所述比特流包括所述指示符,并且如果针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的,则所述比特流还包括针对所述帧中的所述至少一个帧的所述一个或多个区域中的每个区域的、指示针对所述区域的所述MV精度的标记。

2. 如权利要求1所述的计算机系统,其中,应用于所述帧中的所述至少一个帧的所述报头是序列参数集或图片参数集,其中,所述区域是切片,并且其中,所述区域的所述报头是切片报头。

3. 如权利要求1所述的计算机系统,其中,针对所述指示符:

第一可能值指示针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,并且还指示针对所述帧中的所述至少一个帧的所述MV精度是四分之一样本精度;

第二可能值指示针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,并且还指示针对所述帧中的所述至少一个帧的所述MV精度是整数样本精度;以及

第三可能值指示针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的。

4. 如权利要求1所述的计算机系统,其中,针对所述帧中的所述至少一个帧的所述一个或多个区域中的给定区域,如果针对所述给定区域的所述标记不存在于所述比特流中,则针对所述给定区域的所述标记被推断为具有等于所述指示符的值。

5. 如权利要求1所述的计算机系统,其中,所述指示符是二比特值。

6. 如权利要求1所述的计算机系统,其中,对MV值的解释取决于所述MV精度是四分之一样本精度还是整数样本精度而变化,如果所述MV精度是整数样本精度,则所述MV值表示整数样本偏移,而如果所述MV精度是四分之一样本精度,则所述MV值表示四分之一样本偏移。

7. 如权利要求1所述的计算机系统,其中,所述设置所述指示符的所述值是基于所述视

频序列的所述帧的源的。

8. 如权利要求1所述的计算机系统,其中,所述设置所述指示符的所述值是基于对性能启发式或历史数据的测量的。

9. 一种应用于计算机系统中的方法,包括:

接收针对视频序列的帧的编码数据作为比特流的一部分,所述帧中的每一个帧包括一个或多个区域,所述比特流包括:

报头中的指示符,所述报头应用于所述视频序列的所述帧中的至少一个帧,所述指示符指示:

针对所述帧中的所述至少一个帧的运动向量(“MV”)精度是否是在逐区域的基础上被控制的;以及

如果针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,则针对所述帧中的所述至少一个帧的所述MV精度是四分之一样本精度还是整数样本精度;以及

如果针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的,则针对所述帧中的所述至少一个帧的所述一个或多个区域中的每个区域,所述比特流还包括针对所述区域的报头中的标记,所述标记指示针对所述区域的MV精度是四分之一样本精度还是整数样本精度;以及

对所述编码数据进行解码以重构所述视频序列的所述帧。

10. 如权利要求9所述的方法,其中,应用于所述帧中的所述至少一个帧的所述报头是序列参数集或图片参数集,其中,所述区域是切片,并且其中,所述区域的所述报头是切片报头。

11. 如权利要求9所述的方法,其中,针对所述指示符:

第一可能值指示针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,并且还指示针对所述帧中的所述至少一个帧的所述MV精度是四分之一样本精度;

第二可能值指示针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,并且还指示针对所述帧中的所述至少一个帧的所述MV精度是整数样本精度;以及

第三可能值指示针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的。

12. 如权利要求9所述的方法,其中,针对所述帧中的所述至少一个帧的所述一个或多个区域中的给定区域,如果针对所述给定区域的所述标记不存在于所述比特流中,则针对所述给定区域的所述标记被推断为具有等于所述指示符的值。

13. 如权利要求9所述的方法,其中,所述指示符是二比特值。

14. 如权利要求9所述的方法,其中,对MV值的解释取决于所述MV精度是四分之一样本精度还是整数样本精度而变化,如果所述MV精度是整数样本精度,则所述MV值表示整数样本偏移,而如果所述MV精度是四分之一样本精度,则所述MV值表示四分之一样本偏移。

15. 一个或多个计算机可读介质,具有存储于其上的计算机可执行指令,所述计算机可执行指令用于在被编码时促使一个或多个处理单元执行包括以下各项的操作:

接收针对视频序列的帧的编码数据作为比特流的一部分,所述帧中的每一个帧包括一个或多个区域,所述比特流包括:

报头中的指示符,所述报头应用于所述视频序列的所述帧中的至少一个帧,所述指示符指示:

针对所述帧中的所述至少一个帧的运动向量(“MV”)精度是否是在逐区域的基础上被控制的;以及

如果针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,则针对所述帧中的所述至少一个帧的所述MV精度是四分之一样本精度还是整数样本精度;以及

如果针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的,则针对所述帧中的所述至少一个帧的所述一个或多个区域中的每个区域,所述比特流还包括针对所述区域的报头中的标记,所述标记指示针对所述区域的MV精度是四分之一样本精度还是整数样本精度;以及

对所述编码数据进行解码以重构所述视频序列的所述帧。

16. 如权利要求15所述的一个或多个计算机可读介质,其中,应用于所述帧中的所述至少一个帧的所述报头是序列参数集或图片参数集,其中,所述区域是切片,并且其中,所述区域的所述报头是切片报头。

17. 如权利要求15所述的一个或多个计算机可读介质,其中,针对所述指示符:

第一可能值指示针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,并且还指示针对所述帧中的所述至少一个帧的所述MV精度是四分之一样本精度;

第二可能值指示针对所述帧中的所述至少一个帧的所述MV精度不是在逐区域的基础上被控制的,并且还指示针对所述帧中的所述至少一个帧的所述MV精度是整数样本精度;以及

第三可能值指示针对所述帧中的所述至少一个帧的所述MV精度是在逐区域的基础上被控制的。

18. 如权利要求15所述的一个或多个计算机可读介质,其中,针对所述帧中的所述至少一个帧的所述一个或多个区域中的给定区域,如果针对所述给定区域的所述标记不存在于所述比特流中,则针对所述给定区域的所述标记被推断为具有等于所述指示符的值。

19. 如权利要求15所述的一个或多个计算机可读介质,其中,所述指示符是二比特值。

20. 如权利要求15所述的一个或多个计算机可读介质,其中,对MV值的解释取决于所述MV精度是四分之一样本精度还是整数样本精度而变化,如果所述MV精度是整数样本精度,则所述MV值表示整数样本偏移,而如果所述MV精度是四分之一样本精度,则所述MV值表示四分之一样本偏移。

选择运动向量精度

[0001] 本申请是申请日为2014年12月19日、申请号为201480072763.5、发明名称为“选择运动向量精度”的中国专利申请的分案申请。

背景技术

[0002] 工程师使用压缩(也称为源编码或源译码)来减少数字视频的比特率。压缩通过将信息转换为较低比特率形式而降低了存储和传输视频信息的成本。解压缩(也称作解码)根据压缩形式重构原始信息的版本。“编码解码器”是编码器/解码器系统。

[0003] 在过去的二十年里,采用了各种视频编码解码器标准,包括ITU-T H.261、H.262(MPEG-2或ISO/IEC 13818-2)、H.263和H.264(MPEG-4AVC或ISO/IEC 14496-10)标准、MPEG-1(ISO/IEC 11172-2)和MPEG-4视觉(ISO/IEC 14496-2)标准、以及SMPTE 421M(VC-1)标准。最近,批准了HEVC标准(ITU-T H.265或ISO/IEC 23008-2)。目前正在开发HEVC标准的扩展(例如,用于可扩展视频编码/解码,用于对具有较高保真度的视频在样本位深度或色度采样率方面进行编码/解码,或用于多视图编码/解码)。视频编码解码器标准通常定义编码视频比特流的语法的选项,详述在编码和解码中使用特定特征时在比特流中的参数。在多种情况下,视频编码解码器标准还提供关于解码器应该执行以实现解码的一致性结果的解码操作的细节。除了编码解码器标准外,各种属性编码解码器格式定义了用于编码视频比特流和对应的解码操作的语法的其他选项。

[0004] 通常,视频压缩技术包括“图片内”压缩和“图片间”压缩。图片内压缩技术压缩个体图片,而图片间压缩技术参考先前的和/或后续的一张或多张图片(通常称作参考图片或锚定图片)来压缩图片。

[0005] 图片间压缩技术通常使用运动估计和运动补偿,通过利用视频序列中的时间冗余来减少比特率。运动估计是用于估计图片之间的运动的过程。在一种常见技术中,使用运动估计的编码器试图将当前图片中的当前样本值的块与在另一图片(参考图片)中搜索区域内的相同大小的候选块进行匹配。当编码器在参考图片的搜索区域中发现精确的或“足够靠近”的匹配时,编码器将在当前和候选块之间位置的变化参数化为运动数据(例如,运动向量(“MV”))。MV通常是二维值,其具有表示左右空间位移的水平MV分量和表示上下空间位移的垂直MV分量。通常,运动补偿是根据参考图片利用运动数据重构图片的过程。

[0006] MV可以表示当前块在参考图片的共同定位的位置处开始的整数数量的样本网格位置方面的空间位移。例如,对于当前图片中位置(32,16)处的当前块,MV(-3,1)表示在参考图片中的位置(29,17)。或者,MV可以表示当前块在参考图片的共同定位的位置处起的分数量(fractional number)的样本网格位置方面的空间位移。例如,对于在当前图片的位置(32,16)处的当前块,MV(-3.5,1.25)表示在参考图片中的位置(28.5,17.25)。为了确定在参考图片中分数偏移处的样本值,编码器通常在整数样本位置处的样本值之间进行插值。这种插值可能是计算密集的。在运动补偿期间,编码器还执行如在参考图片中的分数偏移处计算样本值所需的插值。

[0007] 不同视频编码解码器标准和格式已经使用了具有不同MV精度的MV。对于整数样本

MV精度, MV分量表示空间位移的整数数量的样本网格位置。对于分数样本MV精度, 例如1/2样本MV精度或1/4样本MV精度, MV分量可以表示空间位移的整数数量的样本网格位置或分数量的样本网格位置。例如, 如果MV精度是1/4样本MV精度, 则MV分量可以表示0样本、0.25样本、0.5样本、0.75样本、1.0样本、1.25样本等的空间位移。一些视频编码解码器标准和格式支持在编码期间切换MV精度。然而, 在特定编码场景中, 不能有效进行关于使用哪个MV精度的编码器侧决策。

发明内容

[0008] 在发明内容部分, 详细描述表示用于选择运动向量(“MV”)精度的在编码器侧操作中的创新。例如, 当视频编码器编码视频时, 视频编码器确定针对视频的单位的MV精度。

[0009] 根据本文描述的创新的一个方面, 当确定针对单位的MV精度时, 视频编码器可以识别具有分数样本MV精度的一组MV值。视频编码器可以至少部分地基于在所述一组MV值中分数部分为零的MV值的流行, 而选择针对所述单位的MV精度。

[0010] 根据本文描述的创新的另一方面, 当为单位确定MV精度时, 视频编码器可以执行率失真分析以在多个MV精度之间进行决定, 其包括一个或多个分数样本MV精度和整数样本MV精度。通过(a)缩放失真成本, (b)向失真成本添加惩罚, (c)缩放比特率成本, (d)向比特率成本添加惩罚, 和/或(e)调整拉格朗日乘数, 率失真分析朝向整数样本MV精度偏置。

[0011] 根据本文描述的创新的另一方面, 当确定针对单位的MV精度时, 视频编码器可以收集关于视频的信息, 并至少部分地基于所收集到的信息从多个MV精度中选择针对单位的MV精度。多个MV精度包括一个或多个分数样本MV精度和整数样本MV精度。

[0012] 用于选择MV精度的编码器侧的选项的创新可以实现为方法的一部分、适于执行方法的计算设备的一部分, 或者存储用于使得计算设备执行方法的计算机可执行指令的有形计算机可读介质的一部分。可以组合或分开使用各种创新。

[0013] 根据以下结合附图进行的详细描述, 本发明的前述和其它目的、特征和优点将变得更清晰。

附图说明

[0014] 图1是可以实现一些所述实施例的示例性计算系统的图。

[0015] 图2a和2b是可以实现一些所述实施例的示例性网络环境的图。

[0016] 图3是可以结合其实现一些所述实施例的示例性编码器系统的图。

[0017] 图4a和4b是示出可以结合其实现一些所述实施例的示例性视频编码器的图。

[0018] 图5是示出具有可以提供用于屏幕捕捉的输入的内容的计算机桌面环境的图。

[0019] 图6是示出具有自然视频内容和人工视频内容的混合内容视频的图。

[0020] 图7a和7b是分别示出带有具有整数样本空间位移和分数样本空间位移的MV值的运动补偿的图。

[0021] 图8是示出在编码期间用于适应MV精度的广义技术的流程图。

[0022] 图9是示出在编码期间利用低复杂度方法用于适应MV精度的示例性技术的流程图。

[0023] 图10是示出根据低复杂度方法的一些变型的图片的不同区域的图。

具体实施方式

[0024] 详细描述呈现了在编码期间选择运动向量(“MV”)精度的创新。这些方法可以便于压缩,其在率失真性能和/或计算效率方面是有效的。例如,视频编码器从多个MV精度确定针对视频的单位的MV精度,所述多个MV精度包括一个或多个分数样本MV精度和整数样本MV精度。视频编码器可以识别具有分数样本MV精度的一组MV值,然后至少部分地基于具有零的分数部分的MV值(组内)的流行,选择针对单位的MV精度。或者,视频编码解码器可以执行率失真分析,其中率失真分析朝向整数样本MV精度偏置。或者,视频编码解码器可以收集关于视频的信息,并至少部分地基于所收集的信息选择针对单位的MV精度。或者,视频编码解码器可以以一些其它方式确定针对视频的单位的MV精度。

[0025] 虽然本文描述的操作被适当地描述为由视频编码器执行,但是在许多情况下,所述操作可以由其它类型的媒体处理工具执行。

[0026] 结合HEVC标准专用的语法元素和操作示出了本文所描述的一些创新。本文描述的创新还可以被实现用于其它标准或格式。

[0027] 更一般地,对本文描述的例子的各种变型都是可以的。例如,可以通过改变所描述的方法动作的顺序、通过分离、重复或省略某些方法动作等来改变一些本文描述的方法。可以组合或分开地使用所公开技术的各个方面。不同实施例使用所描述创新中的一个或多个。本文所描述的一些创新解决了背景中提到的一个或多个问题。通常,给定的技术/工具并不能解决所有这种问题。

[0028] I.示例性计算系统

[0029] 图1示出了可以实现所描述的若干创新的适当计算系统100的广义例子。计算系统100并不打算暗示对使用或功能的范围的任何限制,因为创新可以实现于各种计算系统中,包括适应视频编码的专用计算系统。

[0030] 参考图1,计算系统100包括一个或多个处理单元(110、115)和存储器(120、125)。处理单元(110、115)执行计算机可执行指令。处理单元可以是中央处理单元(“CPU”)、在专用集成电路(“ASIC”)中的处理器或者任意其它类型的处理器。在多处理系统中,多个处理单元执行计算机可执行指令以增加处理能力。例如,图1示出了中央处理单元(110)以及图形处理单元或协处理单元(115)。有形存储器(120、125)可以是通过处理单元可访问的易失性存储器(例如,寄存器、高速缓存、RAM)、非易失性存储器(例如,ROM、EEPROM、闪存等)或者两者的某组合。存储器(120、125)存储实现在编码期间用于选择MV精度的一个或多个创新的软件(180),该软件是适于由处理单元执行的计算机可执行指令的形式。

[0031] 计算系统可以具有额外的特征。例如,计算系统(100)包括存储设备(140)、一个或多个输入设备(150)、一个或多个输出设备(160)、以及一个或多个通信连接(170)。例如总线、控制器或网络的互连机构(未示出)使得计算系统(100)的部件互连。通常,操作系统软件(未示出)提供在计算系统(100)内执行的其它软件的操作环境,并协调计算系统(100)的部件的活动。

[0032] 有形存储设备(140)可以是可移除的或不可移除的,并包括磁盘、磁带或卡带、CD-ROM、DVD或能够用于存储信息并在计算系统(100)内被访问的任意其它介质。存储设备(140)存储实现在编码期间选择MV精度的一个或多个创新的软件(180)的指令。

[0033] 输入设备(150)可以是触摸输入设备,例如,键盘、鼠标、笔或轨迹球、语音输入设

备、扫描设备、或向计算系统(100)提供输入的其它设备。对于视频,输入设备(150)可以是摄像机、视频卡、TV调谐卡、屏幕捕捉模块、或接受模拟或数字形式的视频输入的类似设备、或将视频输入读入到计算系统(100)中的CD-ROM或CD-RW。输出设备(160)可以是显示器、打印机、扬声器、CD刻录机、或从计算系统(100)提供输出的另一设备。

[0034] 通信连接(170)使得能够在通信介质上与另一计算实体进行通信。通信介质传达信息,例如,计算机可执行指令、音频或视频输入或输出、或调制数据信号中的其它数据。调制数据信号是具有一个或多个其特性集或以对信号中的信息进行编码的方式而改变的信号。通过示例而非限制的方式,通信介质可以使用电、光、RF或其它载体。

[0035] 可以在计算机可读介质的一般上下文中描述创新。计算机可读介质是能够在计算环境中被访问的任意可用的有形介质。通过示例而非限制的方式,通过计算系统(100),计算机可读介质包括存储器(120、125)、存储设备(140)和上述任意的组合。

[0036] 可以在计算机可执行指令的一般上下文中描述创新,例如,包含于程序模块中的指令,其在目标真实或虚拟处理器上在计算系统中执行。一般而言,程序模块包括例程、程序、库、对象、类别、部件、数据结构等,其执行特定任务或实现特定抽象数据类型。程序模块的功能可以按照各个实施例所期望地在程序模块之间组合或分离。可以在局部或分布式计算系统中执行用于程序模块的计算机可执行指令。

[0037] 术语“系统”和“设备”在本文中可交换使用。除非上下文中另有明确表示,否则任一术语都不暗示对计算系统或计算设备的类型的任何限制。一般而言,计算系统或计算设备可以是局部的或分布式的,并可以包括专用硬件和/或硬件与实现本文描述的功能的软件的任意组合。

[0038] 还可以利用配置为执行任意所公开方法的专用计算硬件来实现所公开的方法。例如,可以通过专门设计或配置为实现任意所公开方法的集成电路(例如,ASIC,如ASIC数字信号处理器(“DSP”),图形处理单元(“GPU”),或可编程逻辑设备(“PLD”),如现场可编程门阵列(“FPGA”))来实现所公开的方法。

[0039] 出于描述的原因,详细的描述使用类似“确定”和“使用”的术语来描述计算系统中的计算机操作。这些术语是计算机执行的操作的高度抽象,并不应该与人类执行的动作混淆。对应于这些术语的实际计算机操作依赖于实现方式而变化。如本文所使用的,术语“优化(optimiz*)”(包括变形(如optimization和optimizing))指的是在给定决策范围内选项中的选择,并不暗示优化的选择是决策的扩展范围的“最佳”或“最优”选择。

[0040] II.示例性网络环境

[0041] 图2a和2b示出了示例性网络环境(201、202),其包括视频编码器(220)和视频解码器(270)。编码器(220)和解码器(270)利用适当的通信协议通过网络(250)连接。网络(250)可以包括互联网或其它计算机网络。

[0042] 在图2a所示的网络环境(201)中,每个实时通信(“RTC”)工具(210)包括编码器(220)和解码器(270)用于双向通信。给定的编码器(220)可以产生与HEVC标准(还已知为H.265)、SMPTE 421M标准、ISO/IEC 14496-10标准(还已知为H.264或AVC)、其它标准或属性格式的变型或扩展兼容的输出,对应的解码器(270)接受来自编码器(220)的编码数据。双向通信可以是以下项目的一部分:视频会议、视频电话呼叫、或其它两方或多方通信场景。虽然在图2a中的网络环境(201)包括两个实时通信工具(210),但网络环境(201)可以替代

地包括参与多方通信的三个或更多实时通信工具(210)。

[0043] 实时通信工具(210)通过编码器(220)管理编码。图3示出了可以包含于实时通信协议(210)中的示例性编码器系统(300)。替代地,实时通信工具(210)使用另一编码器系统。实时通信工具(210)还通过解码器(270)管理解码。

[0044] 在图2b中所示的网络环境(202)中,编码工具(212)包括编码器(220),其编码视频用于输送到多个回放工具(214)(其包括解码器(270))。可以为视频监控系统、网络摄像机监视系统、屏幕捕捉模块、远程桌面会议呈现或对视频进行编码并将其从一个位置发送到一个或多个其它位置的其它场景提供单向通信。虽然图2b中的网络环境(202)包括两个回放工具(214),但网络环境(202)可以包括更多或更少的回放工具(214)。一般而言,回放工具(214)与编码工具(212)通信以确定回放工具(214)接收的视频流。回放工具(214)接收流,在适当周期内缓冲接收到的编码数据,并开始解码并进行回放。

[0045] 图3示出了可以包含于编码工具(212)中的示例性编码器系统(300)。替代地,编码工具(212)使用另一编码器系统。编码工具(212)还可以包括服务器侧控制器逻辑,用于管理与一个或多个回放工具(214)的连接。回放工具(214)还可以包括客户端侧控制器逻辑,用于管理与编码工具(212)的连接。

[0046] III.示例性编码器系统

[0047] 图3是结合其可以实现一些所述实施例的示例性编码器系统(300)的框图。编码器系统(300)可以是通用编码工具,其能够以任意多个编码模式进行操作,所述编码模式例如是用于实时通信的低延迟编码模式、转码模式、以及用于产生用于从文件或流回放的媒体的较高延迟编码模式;或者编码器系统(300)可以是专用编码工具,其适于一种这样的编码模式。编码器系统(300)可以实现为操作系统模块,作为应用库的一部分或者作为独立的应用。总之,编码器系统(300)从视频源(310)接收源视频帧序列(311),并产生经编码的数据作为到信道(390)的输出。输出到信道的经编码的数据可以包括利用所选的MV精度而编码的内容。

[0048] 视频源(310)可以是摄像机、调谐卡、存储介质、屏幕捕捉模块或其它数字视频源。视频源(310)以例如每秒30帧的帧率产生视频帧序列。如本文所使用的,术语“帧”通常指的是源、编码的或重构的图像数据。对于逐行扫描视频,帧是逐行扫描视频帧。对于交错视频,在示例性实施例中,交错视频帧可以是在编码之前被去交错的。替代地,将两个补充的交错视频字段编码到一起作为单个视频帧或编码为两个分开编码的字段。除了表示逐行扫描视频帧或交错扫描视频帧,术语“帧”或“图片”可以表示单个非配对视频字段、补充视频字段对、表示给定时间的视频对象的视频对象平面、或在较大图像中的感兴趣区域。视频对象平面或区域可以是包括多个对象或场景区域的较大图像的一部分。

[0049] 到达源帧(311)存储于源帧临时存储器存储区域(320)中,其包括多个帧缓冲器存储区域(321、322、…、32n)。帧缓冲器(321、322等)在源帧存储区域(320)中保存一个源帧。在一个或多个源帧(311)已经存储于帧缓冲器(321、322等)之后,帧选择器(330)从源帧存储区域(320)中选择个体源帧。通过帧选择器(330)选择用于输入到编码器(340)中的帧的顺序可能不同于视频源(310)产生帧的顺序,例如,对一些帧的编码可能在顺序上延迟,从而允许一些后续帧首先被编码并因此便于暂时反向预测。在编码器(340)之前,编码器系统(300)可以包括预处理器(未示出),其在编码之前执行所选帧(331)的预处理(例如,过滤)。

预处理可以包括颜色空间转换为主要(例如,亮度)和次要(例如,朝向红色和朝向蓝色的色度差)分量,并对处理进行重采样(例如,减少色度分量的空间分辨率)以进行编码。一般而言,在编码之前,视频已经转换为颜色空间,例如YUV,其中亮度(Y)分量的样本值表示亮度或强度值,并且色度(U,V)分量的样本值表示颜色差值。可以对色度采样值进行子采样到较低色度采样率(例如,用于YUV 4:2:0格式或YUV 4:2:2),或者色度样本值可以具有与亮度样本值相同的分辨率(例如,用于YUV 4:4:4格式)。在YUV 4:2:0格式中,水平地通过两个因素且垂直地通过两个因素对色度分量进行降低采样。在YUV 4:2:2格式中,水平地通过两个因素对色度分量进行降低采样。或者可以以另一格式(例如,RGB 4:4:4格式)对视频进行编码。

[0050] 编码器(340)对所选帧(331)进行编码以产生经编码的帧(341),并且还产生存储器管理控制操作(“MMCO”)信号(342)或参考图片集(“RPS”)信息。如果当前帧不是已经编码的第一帧,则当执行其编码过程时,编码器(340)可以使用一个或多个先前编码/解码的帧(369),该帧(369)已经存储于解码帧临时存储器存储区域(360)内。这种存储的解码帧(369)用作当前源帧(331)的内容的帧间预测的参考帧。MMCO/RPS信息(342)向编码器表明哪个重构帧可以用作参考帧,并因此应该存储于帧存储区域内。

[0051] 一般而言,编码器(340)包括执行编码任务的多个编码模块,所述任务例如分割成瓦片(tile),帧内预测估计和预测,运动估计和补偿,频率转换,量化和熵编码。编码器(340)执行的准确操作可以依赖于压缩格式而变化。输出编码数据的格式可以是HEVC格式(H.265)、Windows媒体视频格式、VC-1格式、MPEG-x格式(例如,MPEG-1、MPEG-2或MPEG-4)、H.26x格式(例如,H.261、H.262、H.263、H.264)或其它格式的变型或扩展。

[0052] 编码器(340)可以将帧分割为相同尺寸或不同尺寸的多个瓦片。例如,编码器(340)将帧沿着瓦片行和瓦片列进行分割,借助帧边界,所述瓦片行和瓦片列定义在帧内的瓦片的水平和垂直边界,其中每个瓦片是矩形区域。瓦片通常用于提供用于并行处理的选项。可以将帧组织为一个或多个切片,其中切片可以是整个帧或帧的区域。可以对切片独立于帧中其它切片进行解码,这改善了误差恢复。为了进行编码和解码,切片或瓦片的内容还被分隔为块或样本的其它集合。

[0053] 对于根据HEVC标准的语法,编码器将帧(或切片或瓦片)的内容分割成编码树单元。编码树单元(“CTU”)包括组织为亮度编码树块(“CTB”)的亮度样本值以及组织为两个色度CTB的对应的色度样本值。CTU(及其CTB)的大小由编码器进行选择。亮度CTB例如可以包含64x64、32x32、16x16的亮度样本值。CTU包括一个或多个编码单元。编码单元(“CU”)具有亮度编码块(“CB”)和两个对应的色度CB。例如,具有64x64亮度CTB和两个64x64的色度CTB(YUV 4:4:4格式)的CTU可以被分割成四个CU,每个CU包括32x32的亮度CB和两个32x32的色度CB,每个CU还可能被分割成更小的CU。或者,作为另一例子,具有64x64亮度CTB和两个32x32色度CTB(YUV 4:2:0格式)的CTU可以被分割为四个CU,每个CU包括32x32亮度CB和两个16x16色度CB,每个CU还可以被分割成更小的CU。CU的最小可允许尺寸(例如,8x8、16x16)可以在比特流中用信号传递。

[0054] 一般而言,CU具有例如帧间或帧内的预测模式。CU包括一个或多个预测单元,用于对预测信息(例如,预测模式细节、位移值等)和/或预测处理进行信号传递的目的。预测单元(“PU”)具有亮度预测块(“PB”)和两个色度PB。对于帧内预测的CU,PU具有与CU相同的尺

寸,除非CU具有最小尺寸(例如,8x8)。在这种情况下,CU可以分割成四个较小的PU(例如,当最小CU尺寸是8x8时,每个是4x4),或者PU可以具有最小CU尺寸,如由CU的语法元素指示的。CU还具有一个或多个变换单元,用于残差编码/解码的目的,其中变换单元(“TU”)具有亮度变换块(“TB”)和两个色度TB。在帧内预测CU中的PU可以包含单个TU(在尺寸上等于PU)或多个TU。编码器决定如何将视频分割为CTU、CU、PU、TU等。在H.264/AVC标准的上下文中,术语“宏块”指示类似于H.265/HEVC标准的CTU的块状区域的块状区域,并且术语“子宏块分割”指示类似于CU或PU的块状区域的块状区域。如本文所使用的,术语“块”可以表示CB、PB、TB、CTU、CU、PU、TU、宏块、子宏块分割或其它样本值集合,这取决于上下文。

[0055] 返回图3,编码器以根据帧(331)中其它先前重构的样本值的预测的方式呈现源帧(331)的内部编码块。对于块内复制(“BC”)预测,图片内估计器估计块相对其它先前重构样本值的位移。帧内预测参考区域(或简称帧内预测区域)是用于生成块的BC预测值的帧中的样本的区域。帧内预测区域可以用块向量(“BV”)值(在BV估计中确定的)来指示。对于块的帧内空间预测,图片内估计器估计相邻的重构样本值到块的外推。图片内估计器可以输出预测信息(例如,BC内预测的BV值或帧内空间预测的预测模式(方向)),其是熵编码的。帧内预测预测器应用预测信息来确定帧内预测值。

[0056] 编码器(340)表示按照根据参考帧的预测的方式的源帧(331)的帧间编码的预测块。运动估计器估计块相对于一个或多个参考帧(369)的运动。运动估计器可以选择如文本所述的运动向量(“MV”)精度(例如,整数样本MV精度、1/2样本MV精度或1/4样本MV精度),然后在运动估计期间使用所选的MV精度。当使用多个参考帧时,多个参考帧可以来自不同的时间方向或相同的时间方向。运动补偿预测参考区域是在用于生成当前帧的样本块的运动补偿的预测值的参考帧中样本的区域。运动估计器输出运动信息,例如,MV信息,其是熵编码的。运动补偿器将具有所选MV精度的MV值应用于参考帧(369),以确定用于帧间预测的运动补偿的预测值。

[0057] 编码器可以确定块的(帧内或帧间)预测值和对应的原始值之间的差(如果有的话)。利用频率变换(如果频率变换不被省略)、量化和熵编码对这些预测残差值进一步编码。例如,编码器(340)设置用于图片、瓦片、切片和/或视频的其它部分的量化参数(“QP”)的值,并因此量化变换系数。编码器(340)的熵编码器压缩量化变换系数值以及某些边带信息(例如,MV信息、所选MV精度、BV值、QP值、模式决策、参数选择)。典型的熵编码技术包括指数哥伦布编码、Golomb-Rice编码、算数编码、差分编码、霍夫曼编码、行程编码、可变长度到可变长度(“V2V”)编码、可变长度到固定长度(“V2F”)编码、Lempel-Ziv(“LZ”)编码、词典编码、概率区间分割熵编码(“PIPE”)以及上述组合。熵编码器可以针对不同种类的信息使用不同的编码技术,可以应用组合的多个技术(例如,通过应用Golomb-Rice编码之后是算数编码),并可以从特定编码技术中的多个代码表中进行选择。在一些实现方式中,频率变换可以被省略。在这种情况下,可以量化和熵编码预测残差值。

[0058] 在编码器(340)中的运动补偿循环(也就是,“循环中”过滤)内包括自适应去块滤波来平滑跨解码帧中的块边界行和/或列的不连续性。其它过滤(例如,未示出的去环过滤、自适应循环过滤(“ALF”)或样本适应偏移(“SAO”)过滤)可以替代的或额外地应用为循环内过滤操作。

[0059] 通过解码过程模拟器(350)处理编码帧(341)和MMCO/RPS信息(342)(或者等价于

MMCO/RPS信息(342)的信息,因为在编码器(340)处已知帧的依赖性和顺序结构)。解码过程模拟器(350)实现解码器的一些功能,例如,解码任务以重构参考帧。以与MMCO/RPS信息(342)一致的方式,解码过程模拟器(350)确定给定的编码帧(341)是否需要被重构,并存储用作将被编码的后续帧的帧间预测中的参考帧。如果编码帧(341)需要被存储,则解码过程模拟器(350)对接收编码帧(341)且产生对应解码帧(351)的解码器所要进行的解码过程进行建模。这样,当编码器(340)已经使用存储于解码帧存储区域(360)内的解码帧(369)时,解码过程模拟器(350)还使用来自存储区域(360)的解码帧(369)作为解码过程的一部分。

[0060] 解码帧临时存储器存储区域(360)包括多个帧缓冲器存储区域(361、362、…、36n)。以与MMCO/RPS信息(342)一致的方式,解码过程模拟器(350)管理存储区域(360)的内容,以便识别具有编码器(340)不再需要用作参考帧的帧的任意帧缓冲器(361、362等)。在对解码过程建模之后,解码过程模拟器(350)在帧缓冲器(361、362等)中存储以这种方式识别出的新解码的帧(351)。

[0061] 编码帧(341)和MMCO/RPS信息(342)在临时编码数据区域(370)中缓冲。在编码数据区域(370)内聚集的编码数据包含用于一个或多个图片的编码数据作为元素编码视频比特流的语法的一部分。在编码数据区域(370)中聚集的编码数据还可能包括与编码视频数据相关的媒体元数据(例如,作为一个或多个补充增强信息(“SEI”)消息或视频可用信息(“VUI”)消息中的一个或多个参数)。

[0062] 通过信道编码器(380)对来自临时编码数据区域(370)的聚集数据(371)进行处理。信道编码器(380)可以分封化和/或多路复用聚集的数据以进行传输或存储为媒体流(例如,根据媒体程序流或传输流格式,例如,ITU-T H.222.0|ISO/IEC 13818-1或互联网实时传输协议格式,例如IETF RFC 3550),在该情况下,信道编码器(380)可以添加语法元素作为媒体传输流的语法的一部分。或者,信道编码器(380)可以组织聚集数据用于存储为文件(例如,根据媒体容器格式,例如,ISO/IEC 14496-12),在该情况下,信道编码器(380)可以添加语法元素作为媒体存储文件的语法的一部分。或者更一般地,信道编码器(380)可以实现一个或多个媒体系统多路复用协议或传输协议,在该情况下,信道编码器(380)可以添加语法元素作为协议的语法的一部分。信道编码器(380)向信道(390)提供输出,其表示存储、通信连接或用于输出的另一信道。信道编码器(380)或信道(390)还可以包括其它元件(未示出),例如,用于前向纠错(“FEC”)编码和模拟信号调制。

[0063] IV.示例性视频编码器

[0064] 图4a和4b是可以结合其实现一些所述实施例的广义视频编码器(400)的框图。编码器(400)接收包括当前图片的视频图片序列作为输入视频信号(405),并在编码视频比特流(495)中产生编码数据作为输出。

[0065] 编码器(400)是基于块的,且使用依赖于实现方式的块格式。块还可以在不同阶段被细分,例如,在预测、频率变换和/或熵编码阶段。例如,图片可以被划分为64x64块、32x32块或16x16块,其反过来可以划分为较小块的样本值用于编码和解码。在针对HEVC标准的编码的实现方式中,编码器将图片分割为CTU(CTB)、CU(CB)、PU(PB)和TU(TB)。

[0066] 编码器(400)利用图片内编码和/或图片间编码来压缩图片。编码器(400)的许多部件用于图片内编码和图片间编码。通过这些部件执行的准确操作可以依赖于被压缩的信息类型而变化。

[0067] 瓦片模块(410)可选地将图片分割为相同尺寸或不同尺寸的多个瓦片。例如,瓦片模块(410)沿着瓦片行和瓦片列分割图片,借助图片边界,所述瓦片行和瓦片列定义图片内瓦片的水平和垂直边界,其中每个瓦片是矩形区域。

[0068] 通用编码控制件(420)从编码器(400)的各个模块接收用于输入视频信号(405)的图片以及反馈(未示出)。总之,通用编码控制件(420)向其它模块(例如,瓦片模块(410)、变换器/缩放器/量化器(430)、缩放器/逆变换器(435)、图片内估计器(440)、运动估计器(450)和内/间切换)提供控制信号(未示出),以在编码期间设置和改变编码参数。特别地,结合运动估计器(450),通用编码控制件(420)可以在编码期间确定MV精度。通用编码控制件(420)还可以在编码期间评估中间结果,例如,执行率失真分析。通用编码控制件(420)产生指示在编码期间做出的决策的通用控制数据(422),从而对应的解码器可以做出一致决策。向报头格式化器/熵编码器(490)提供通用控制数据(422)。

[0069] 如果利用图片间预测来预测当前图片,则运动估计器(450)估计输入视频信号(405)的当前图片(相对一个或多个参考图片)的样本值的块运动。如本文所述,运动估计器(450)可以选择运动向量(“MV”)精度(例如,整数样本MV精度、1/2样本MV精度或1/4样本MV精度),然后在运动估计期间使用所选的MV精度。解码图片缓冲器(470)缓冲一个或多个重构的先前编码的图片以用作参考图片。当使用多个参考图片时,多个参考图片可以来自于不同时间方向或相同时间方向。运动估计器(450)产生边带信息运动数据(452),例如,MV数据、合并模式索引值和参考图片选择数据、以及表示所选的MV精度的边带信息。将包括运动数据(452)的边带信息提供给报头格式化器/熵编码器(490)以及运动补偿器(455)。

[0070] 运动补偿器(455)将具有所选的MV精度的MV值应用到来自解码图片缓冲器(470)的重构的参考图片。当图片的色度数据具有与亮度数据相同的分辨率时(例如,当格式是YUV 4:4:4或RGB 4:4:4格式时),应用于色度块的MV值可以与应用于亮度块的MV值相同。在另一方面,当图片的色度数据具有相对于亮度数据减少的分辨率时(例如,当格式是YUV 4:2:0格式或YUV 4:2:2格式时),应用于色度块上的MV值可以是按比例缩小的MV值,并可以对其进行化整,以调节色度分辨率的差(例如,对于YUV 4:2:0格式,通过用2除MV值的垂直和水平分量,以及对其进行舍位或化整到用于色度运动补偿处理的精度;对于YUV 4:2:2格式,通过用2除MV值的水平分量,以及对其进行舍位或化整到用于色度运动补偿处理的精度)。运动补偿器(455)针对当前图片产生运动补偿预测。

[0071] 在编码器(400)的分开的路径中,图片内估计器(440)确定如何执行输入视频信号(405)的当前图片的样本值的块的图片内预测。可以利用图片内编码对当前图片进行整体或局部编码。利用当前图片的重构(438)的值,对于空间内预测,图片内估计器(440)确定如何根据当前图片的邻近的先前重构的样本值来空间上预测当前图片的当前块的样本值。或者,对于利用BV值的BC内预测,图片内估计器(440)估计当前块的样本值到当前图片内的不同候选区域的位移。

[0072] 图片内估计器(440)产生边带信息内预测数据(442),例如,指示内预测是否使用空间预测或BC内预测(例如,每个块内标记值)、预测模式方向(对于空间内预测)和BV值(对于BC内预测)的信息。将内预测数据(442)提供到报头格式化器/熵编码器(490)以及图片内预测器(445)。

[0073] 根据内预测数据(442),图片内预测器(445)根据当前图片的邻近的先前重构的样

本值在空间上预测当前图片的当前块的样本值。或者,对于BC内预测,图片内预测器(445)利用内预测区域的先前重构的样本值预测当前块的样本值,其由当前块的BV值指示。

[0074] 内/间切换选择运动补偿预测或图片内预测的值用作给定块的预测(458)。当没有略过残差编码时,在预测(458)的块与输入视频信号(405)的原始当前图片的对应部分之间的差(如果有的话)提供残差值(418)。在重构当前图片期间,当已经对残差值进行编码/信号传递时,组合重构的残差值与预测(458)以从视频信号(405)产生原始内容的重构(438)。然而,在有损压缩中,仍会从视频信号(405)中丢失一些信息。

[0075] 在变换器/缩放器/量化器(430)中,当没有略过频率变换时,频率变换器将空间域视频数据转换为频域(即,频谱、变换)数据。对于基于块的视频编码,频率变换器对预测残差数据的块(或当预测(458)是空时对样本值数据)应用离散余弦变换(“DCT”、其整数近似、或另一类型的前向块变换(例如,离散正弦变换或其整数近似),产生频率变换系数的块。编码器(400)还能够指示略过这种变换步骤。缩放器/量化器缩放且量化变换系数。例如,量化器将死区标量量化应用到具有基于逐帧、逐瓦片、逐切片、逐块、专用频率或其它而变化的量化步长尺寸的频域数据上。将量化变换系数数据(432)提供给报头格式化器/熵编码器(490)。如果略过频率变换,则缩放器/量化器可以缩放和量化预测残余数据的块(或者当预测(458)为空时的样本值数据),产生提供给报头格式化器/熵编码器(490)的量化值。

[0076] 在缩放器/逆变换器(435)中,缩放器/逆量化器对量化的变换系数执行逆缩放和逆量化。逆频率变换器执行逆频率变换,产生重构预测残差值或样本值的块。如果已经略过了变换阶段,则也略过逆频率变换。在这种情况下,缩放器/逆量化器可以对预测残差数据(或样本值数据)的块执行逆缩放和逆量化,产生重构值。当已经对残差值编码/信号传递时,编码器(400)组合重构的残差值与预测值(458)(例如,运动补偿预测值,图片内预测值)以形成重构(438)。当尚未对残差值编码/信号传递时,编码器(400)使用预测值(458)作为重构(438)。

[0077] 对于图片内预测,可以将重构(438)的值馈送回图片内估计器(440)和图片内预测器(445)。另外,重构(438)的值可以用于后续图片的运动补偿预测。重构(438)的值可以进一步被过滤。针对给定的视频信号(405)的图片,过滤控制件(460)确定如何在重构(438)的值上执行去块过滤和SAO过滤。过滤控制件(460)产生过滤控制数据(462),该数据被提供给报头格式化器/熵编码器(490)和合并器/过滤器(465)。

[0078] 在合并器/过滤器(465)中,编码器(400)将来自不同瓦片的内容合并到重构版本的图片中。编码器(400)根据过滤控制数据(462)选择性执行去块过滤和SAO过滤,从而适应地平滑跨帧中的边界的不连续性。可以替代地或额外的应用其它过滤(例如,未示出的去环过滤或ALF)。瓦片边界可以取决于编码器(400)的设置而选择性过滤或根本不过滤,并且编码器(400)可以提供编码比特流中的语法以指示是否应用这种过滤。解码图片缓冲器(470)缓冲重构的当前图片以用于后续的运动补偿预测。

[0079] 报头格式化器/熵编码器(490)对通用控制数据(422)、量化变换系数数据(432)、内预测数据(422)、运动数据(452),以及过滤控制数据(462)格式化和/或熵编码。可以对MV值进行预测性编码。例如,报头格式化器/熵编码器(490)使用指数哥伦布编码用于在MV预测之后对各种语法元素进行熵编码,例如,差分MV值的语法元素。

[0080] 报头格式化器/熵编码器(490)提供编码视频比特流(495)中的编码数据。编码视

频比特流(495)的格式可以是HEVC格式、Windows媒体视频格式、VC-1格式、MPEG-x格式(例如,MPEG-1、MPEG-2或MPEG-4)、H.26x格式(例如,H.261、H.262、H.263、H.264)或其它格式的变型或扩展。

[0081] 取决于期望压缩的实现方式和类型,可以添加、省略编码器的模块,将其分割成多个模块,使其与其它模块组合和/或以类似模块进行替代。在替代实施例中,具有不同模块的编码器和/或模块的其它配置执行一个或多个所描述的技术。编码器的具体实施例通常使用变型或补充版本的编码器(400)。在编码器(400)内的模块之间示出的关系指示在编码器中的信息的一般流;为了简便的原因未示出其它关系。

[0082] V. 在编码期间选择MV精度

[0083] 该部分表示用于在编码期间选择运动向量(“MV”)精度的各种方法。这些方法可以便于压缩,其在编码和解码的率失真性能和/或计算效率方面是有效的。

[0084] 当对任意类型的视频编码时可以应用本文所描述的用于选择MV精度的方法。然而,特别地,当编码某些人工创建的视频内容(例如,屏幕捕捉内容)时,如本文所描述的选择MV精度可以改善性能。

[0085] A. 视频类型

[0086] 通常,屏幕捕捉视频(也称作屏幕内容视频或屏幕捕捉内容)表示呈现生成计算机屏幕或其它显示器的内容的处理的图形的输出。这与自然视频形成对比,自然视频指的是从真实世界对象的摄像机传感器视图捕捉到的视频图像,或具有类似特性的视频。屏幕捕捉视频通常包含所呈现的文本、计算机图形、动画生成内容或从计算机显示器的呈现过程的输出捕捉到的其它类似类型的内容,这与(或除了)仅摄像机捕捉视频内容相反。编码/解码屏幕捕捉内容的共同场景包括远程桌面会议和在自然视频或其它“混合内容”视频上的图片或文本覆盖的编码/解码。本文描述的若干创新适应于编码屏幕捕捉视频或其它人工创造视频。这些创新还可以用于自然视频,但是可能不一样有效。本文描述的其它创新在编码自然视频或人工创造视频方面是有效的。

[0087] 图5示出了具有可以为屏幕捕捉提供输入的内容的计算机桌面环境(510)。例如,屏幕捕捉视频可以表示整个计算机桌面(511)的一系列图像。或者,屏幕捕捉视频可以表示计算机桌面环境的窗口之一的一些列图像,例如,包括游戏内容的app窗口(513),具有网页内容的浏览器窗口(512)或者具有文字处理器内容的窗口(514)。

[0088] 与利用视频摄像机捕捉到的自然视频内容相比,如计算机生成的、人工创造的视频内容一样,屏幕捕捉内容趋向于具有相对少的离散样本值。例如,屏幕捕捉内容的区域通常包括单一均匀颜色,而自然视频内容中的区域更可能包括逐渐变换的颜色。另外,屏幕捕捉内容通常包括独特结构(例如,图形、文本特性),其逐帧准确地重复,即使内容在空间上可能被替代(例如,由于滚动)。屏幕捕捉内容通常以具有高色度采样分辨率的格式(例如,YUV 4:4:4或RGB 4:4:4)编码,但是其还可以以具有较低色度采样分辨率的格式(例如,YUV 4:2:0,YUV 4:2:2)编码。

[0089] 图6示出了混合内容视频(620),其包括一些自然视频(621)和一些人工创造的视频内容。人工创造的视频内容包括在自然视频(621)和在自然视频(621)下运行的收报机(ticker)(623)旁边的图形(622)。类似于图5所示的屏幕捕捉内容,在图6中示出的人工创造的视频内容倾向于具有相对少的离散样本值。其还倾向于具有(例如,由于滚动)逐帧准

确重复的独特结构(例如,图形、文本特性)。

[0090] 可以从显示设备的输出缓冲器中或从一个或多个存储帧的缓冲器中周期性读取屏幕捕捉视频或混合内容视频。或者,可以从屏幕捕捉模块(其可以从显示设备的输出缓冲器中周期性地读取值,从操作系统模块拦截显示命令,或者以其它方式捕捉待显示的样本值)提供屏幕捕捉视频。屏幕捕捉视频或混合内容视频可以来自“现场”流或来自在存储设备中先前记录的流。

[0091] B.不同的MV精度

[0092] 在多个编码场景中,当编码屏幕捕捉视频或其它人工创造的视频内容时,大部分MV值表示整数样本空间位移,并且非常少的MV值表示分数样本空间位移。这提供了机会来减少MV精度以改善整体性能。

[0093] 图7a示出了带有具有整数样本空间位移的MV(720)的运动补偿。MV(720)指示向左的四个样本以及一个向上样本的空间位移,相对于当前块的参考图片中的共同定位位置(710)。例如,对于在当前图片中位置(64,96)处的4x4当前块,MV(720)表示4x4预测区域(730),其在参考图片上的位置是(60,95)。预测区域(730)包括在参考图片中整数样本位置处的重构的样本值。编码器或解码器不需要执行插值来确定预测区域(730)的值。

[0094] 图7b示出了带有具有分数样本空间位移的MV(721)的运动补偿。MV(721)指示向左的3.75个样本和向上的0.5个样本的空间位移,相对于当前块在参考图片中共同定位的位置(710)。例如,对于在当前图片中位置(64,96)处的4x4当前块,MV(721)指示4x4预测区域(731),其在参考图片上的位置是(60.25,95.5)。预测区域(731)包括在参考图片中分数样本位置处的插值样本值。编码器或解码器执行插值来确定预测区域(731)的样本值。当允许分数样本空间位移时,存在可以与当前块匹配的多个候选预测区域,并因此运动补偿预测的质量通常会改善至少一些类型的视频内容(例如,自然视频)。

[0095] 当MV精度对于视频的单位是整数样本精度时,在该单位中的块的所有的MV值表示整数样本空间位移。当MV精度对于视频的单位是分数样本精度时,在该单位中的块的MV值可以表示分数样本空间位移或整数样本空间位移。也就是,当MV精度对视频的单位是分数样本精度时,在该单位中块的一些MV值可以表示分数样本空间位移,而在该单位中块的其它MV值指示整数样本空间位移。

[0096] 当使用运动估计和运动补偿对块进行编码时,编码器通常计算在块的样本值及其运动补偿预测之间的逐样本差(也称作残差值或误差值)。然后对残差值进行编码。对于残差值,编码效率取决于残差值的复杂度以及引入多少损失或失真作为压缩过程的一部分。通常,良好的运动补偿预测密切近似块,从而残差值是可以有效编码的小幅度差。在另一方面,不良的运动补偿预测通常产生包括较大幅度值的残差值,其更难以有效编码。编码器通常花费大比例的编码时间执行运动估计,试图找到良好的匹配,并因此改善率失真性能。

[0097] 当编码解码器使用具有整数样本MV精度的MV值时,编码器和解码器不需要针对运动补偿在参考图片的样本值之间执行插值操作,因为MV值指示整数样本空间位移。当编码解码器使用具有分数样本MV精度的MV值时,编码器和解码器可以针对运动补偿在参考图片的样本值之间执行插值操作(至少对于指示分数样本空间位移的MV值增加了计算复杂度),但是相对于整数样本MV精度,运动补偿预测趋向于更密切近似块(导致具有较小有效值的残差值)。

[0098] C.MV值的表示

[0099] 通常利用含义取决于相关联的MV精度的整数值来表示MV值。对于整数样本MV精度,例如,整数值1指示1个样本的空间位移,整数值2表示2个样本的空间位移,等等。对于1/4样本MV精度,例如,整数值1指示0.25个样本的空间位移。整数值2、3、4和5分别指示0.5、0.75、1.0和1.25个样本的空间位移。不管MV精度,整数值可以指示空间位移的幅度,并且分开的标志值可以指示位移是正的还是负的。可以利用两个整数值表示给定的MV值的水平MV分量和垂直MV分量。因此,表示MV值的两个整数值的含义取决于MV精度。例如,对于具有2样本水平位移以及没有垂直位移的MV值,如果MV精度是1/4样本MV精度,则MV值表示为(8,0)。然而,如果MV精度是整数样本MV精度,则MV值表示为(2,0)。

[0100] 在编码视频数据的比特流中的MV值通常是熵编码的(例如,基于MV分量方式)。还可以相对于预测MV值不同地对MV值进行编码(例如,基于MV分量方式)。在许多情况下,MV值等于预测的MV值,从而差分MV值为零,其可以被非常有效地编码。可以利用指数哥伦布编码、上下文自适应二进制算术编码或熵编码的另一形式,对差分MV值(或在不使用MV预测时的MV值)进行熵编码。虽然在MV值(或差分MV值)和编码比特之间的密切关系取决于所使用的熵编码形式,但一般而言,较小的值会被更有效地编码(也就是,利用较少的比特),这是因为它们更普通,而较大的值会被效率较低地编码(也就是,利用更多的比特),这是因为它们较不常见。

[0101] D.自适应MV精度一介绍

[0102] 为了总结先前的三个部分,使用具有整数样本MV精度的MV值趋向于降低与用信号传递MV值相关联的比特率,并降低编码和解码的计算复杂度(通过避免在参考图片中分数样本位置处样本值的插值),但是可能降低运动补偿预测的质量,并因此至少对于一些类型的视频内容,增加残差值的幅度。在另一方面,使用具有分数样本MV精度的MV值趋向于增加与用信号传递MV值相关联的比特率,并增加编码和解码的计算复杂度(通过包括在参考图片中分数样本位置处的样本值的插值),但是可能改善运动补偿预测的质量,并至少对于一些类型的视频内容,减小残差值的幅度。通常,计算复杂度、用信号传递MV值的比特率以及运动补偿预测的质量随着MV精度增加而增加(例如,从整数样本到1/2样本,或从1/2样本到1/4样本),直到收益递减点。同时,虽然增加的MV精度趋向于增加用信号传递MV值所需的比特率,但在编码自然内容时,运动补偿预测的质量中的相关联改善可能减少发送残差值的足够近似所需的比特率,并因此减少用足够图片质量编码视频内容所需的总比特率。

[0103] 当对屏幕捕捉视频或其它人工创造的视频内容进行编码时,分数样本MV精度的添加成本(在比特率和计算复杂度方面)可能是不公平的。例如,如果大部分MV值表示整数样本空间位移,而只有非常少的MV值表示分数样本空间位移,则不能保证分数样本MV精度的添加成本。编码器可以在运动估计期间略过搜索分数样本位置(并略过插值操作以确定分数样本位置处的样本值)。对于这种内容,通过使用具有整数样本MV精度的MV值,可以减少比特率和计算复杂度,而没有对运动补偿预测质量的显著惩罚。

[0104] 由于分数样本MV精度对于其它类型的视频内容(例如,摄像机捕捉到的自然视频)可能仍是有用的,所以编码器和解码器可以适于在MV精度之间切换。例如,编码器和解码器可以对于屏幕捕捉视频使用整数样本MV精度,但是对于自然视频使用分数样本MV精度(例如,1/4样本MV精度)。在下一段中描述在选择MV精度时编码器可以遵从的方法。编码器可以

利用比特流中的一个或多个语法元素向解码器用信号传递所选择的MV精度。

[0105] 在用信号传递MV精度的一种方法中,当启用自适应选择MV精度时,编码器逐切片选择MV精度。在序列参数集(“SPS”)、图片参数集(“PPS”)或其它语法结构中的标记值指示是否启用MV精度的自适应选择。如果是,则在给定切片的切片报头中的一个或多个语法元素指示对所述切片的块的所选择的MV精度。例如,标记值0指示1/4样本MV精度,标记值1指示整数样本MV精度。

[0106] 在用信号传递MV精度的另一方法中,编码器逐图片或逐切片选择MV精度。PPS中的语法元素指示三种MV精度模式之一:(0)对于与PPS相关联的图片的切片的MV值的1/4样本MV精度,(1)对于与PPS相关联的图片的切片的MV值的整数样本MV精度,或(2)取决于每切片报头用信号传递的标记值的切片自适应MV精度,其中在切片的切片报头中的标记值可以针对切片的MV值指示1/4样本MV精度或整数样本MV精度。对于在一个实现方式中关于该方法的额外细节,参见JCTVC-P0277。

[0107] 在用信号传递MV精度的另一方法中,当启用MV精度的自适应选择时,编码器逐CU选择MV精度。对于给定CU的结构中的一个或多个语法元素指示所述CU的块的所选择的MV精度。例如,在CU的CU语法结构中的标记值表示对于与CU相关联的所有PU的MV值是具有整数样本MV精度还是1/4样本MV精度。对于在一个实现方式中关于该方法的额外细节,参见JCTVC-P0283。

[0108] 在这些方法的任一个中,编码器和解码器可以针对水平和垂直MV分量使用不同的MV精度。当对已经水平或垂直缩放的屏幕捕捉视频编码(例如,利用在未缩放维度中的整数样本MV精度,以及利用在缩放维度中的分数样本MV精度)时,这可能是有用的。在一些实现方式中,如果不能单独通过调整QP值而实现率控制,则编码器可以水平或垂直调整屏幕捕捉视频的大小以减小比特率,然后对调整大小后的视频进行编码。在解码器侧,在解码后将视频缩放回其原始维度。编码器可以用信号传递水平MV分量的MV精度(例如,利用第一标记值或语法元素)并还可以用信号传递垂直MV分量的MV精度(例如,利用第二标记值或语法元素)到解码器。

[0109] 更一般地,当启用MV精度的自适应选择时,编码器以某种方式选择MV精度并用信号传递所选的MV精度。例如,在SPS、PPS或其它语法结构中的标记值可以表示是否启用自适应选择MV精度。当启用自适应MV精度时,在序列-层语法、图片组-层语法(“GOP-层语法”)、图片-层语法、切片-层语法、瓦片-层语法、块-层语法或其它语法结构中的一个或多个语法元素可以指示MV值的所选的MV精度。或者,在序列-层语法、GOP-层语法、图片-层语法、切片-报头-层语法、切片-数据-层语法、瓦片-层语法、块-层语法或其它语法结构中的一个或多个语法元素可以指示不同MV分量的MV精度。当存在两个可用MV精度时,标记值可以指示在两个MV精度之间的选择。当存在更多可用MV精度时,整数值可以指示在这些MV精度之间的选择。

[0110] 除了修改用信号传递/解析指示所选MV精度的语法元素外,可以修改解码以改变取决于所选MV精度如何解释用信号传递的MV值。如何编码和重构MV值的细节可以取决于MV精度而变化。例如,当MV精度是整数样本精度时,可以对预测到的MV值化整为最接近的整数,并且差分MV值可以指示整数样本偏移。或者,当MV精度是1/4样本精度时,可以对预测的MV值化整到最接近的1/4样本偏移,并且差分MV值可以指示1/4样本偏移。或者,可以用某种

其它方式用信号传递MV值。当MV值具有整数样本MV精度且视频使用4:2:2或4:2:0色度采样时,可以通过缩放等导出色度MV值,这可能导致色度的1/2样本位移。或者,可以将色度MV值化整为整数值。

[0111] E. 选择MV精度的方法

[0112] 当在视频编码期间自适应MV精度时,编码器针对视频的单位选择MV精度。编码器可以基于来自适配源的线索(hint)选择MV精度来使用(参见以下的方法1)。例如,视频源可以指示视频是屏幕捕捉内容或自然视频(由摄像机捕捉到的)。或者编码器可以基于对各种MV精度的详尽评估来选择MV精度(参见以下的方法2)。或者,编码器可以基于分析来自先前单元的统计数据和/或正在编码的当前单位的统计数据来选择MV精度(参见下文的方法3-4)。

[0113] 对选择MV精度的一些方法进行自适应于屏幕捕捉编码场景。当编码任意类型的视频内容时更一般地应用其它方法。

[0114] 在本段描述的一些例子中,编码器在使用1/4样本MV精度和整数样本MV精度之间进行选择。更一般地,编码器在多个可用MV精度之间进行选择,其可以包括整数样本MV精度、1/2样本MV精度、1/4样本MV精度和/或另一MV精度。

[0115] 当编码器针对视频的单位选择MV精度时,视频的单位可以是序列、GOP、图片、切片、瓦片、CU、PU、其它块或其它类型的视频的单位。取决于在复杂度和灵活度之间期望的权衡,以高度局部基础(例如,逐CU)、较大的逐区域(例如,逐瓦片或逐切片)、整个图片基础、或者更全局的基础(例如,每个编码会话、每个序列、每个GOP、或检测到的场景变化之间的每个图片序列)选择MV精度可能是合适的。

[0116] 1. 使用来自应用、操作系统或视频源的线索的方法

[0117] 编码器可以基于由应用、操作系统或视频源用信号传递的线索来选择MV精度。例如,线索可以指示将被编码的视频内容由特定应用呈现,所述应用例如是文字处理器、电子表单应用或网络浏览器(不具有嵌入式视频区域,其可以是自然视频内容)。利用这种应用呈现可以趋向于仅产生内容的整数样本空间位移。基于这种线索,编码器可以选择整数样本MV精度。对于以文字处理器、电子表单应用、网络浏览器或不经常呈现自然视频内容的其它应用呈现的内容,整数样本MV精度可能优于分数样本MV精度。(但是当已经对视频调整大小时,分数样本MV精度可能是更可取的。)

[0118] 或者,线索可以指示通过屏幕捕捉模块或者通常输送人工创造的视频内容的其它视频源输送视频内容。对于这种内容,整数样本MV精度可能优于分数样本MV精度,从而编码器选择整数样本MV精度。(但是当已经对视频调整大小时,分数样本MV精度可能是更可取的。)

[0119] 在另一方面,如果线索指示视频内容是由摄像机、DVD或其它磁盘或调谐卡输送的或者由视频播放器呈现的,则编码器可以选择分数样本MV精度。对于这种内容,分数样本MV精度可能优于整数样本MV精度。

[0120] 线索可以应用于编码会话、一系列帧、单个视频帧或部分视频帧(例如,对应于与应用相关联的窗口的区域)。

[0121] 在一些情况下,编码器可能不接收或可能不能解释视频源、操作系统或应用提供的关于视频内容本质的线索。或者,线索可能是不正确的或令人误解的(例如,对于包括自

然视频内容和人工创造的视频内容的混合内容视频,或对于已经调整大小的视频)。在这种情况下,编码器可以使用另一方法确定应该选择哪个MV精度。

[0122] 2. 强力编码方法

[0123] 在选择MV精度的另一组方法中,编码器利用不同的MV精度对视频的单位多次进行编码(例如,一次利用整数采样MV精度,一次利用1/4样本MV精度)。编码器选择提供最佳性能的MV精度,并在对所述单位进行编码用于输出时使用所选的MV精度。视频的单位可以是块、PU、CU、切片、瓦片、图片、GOP、序列或其它类型的视频的单位。通常,编码器以这种方法执行多次编码。

[0124] 为了评估哪个MV精度提供了最佳性能,编码器可以在编码单位期间使用不同MV精度时,确定率失真成本,并选择具有最低率失真成本的选项。率失真成本具有失真成本D和比特率成本R,其具有因子 λ (通常称作拉格朗日乘子),该因子相对失真成本,对比特率成本加权($D+\lambda R$)或反之亦然($R+\lambda D$)。比特率成本可以是估计的或实际的比特率成本。通常,失真成本基于对原始样本和重构样本的比较。可以将失真成本测量为绝对差的和("SAD")、绝对Hadamard变换差的和("SAHD")或绝对变换差的其它和("SATD")、平方误差的和("SSE")、均方误差("MSE")、均值方差或另一失真基准。因子 λ 可以在编码期间变化(例如,当量化步长较大时增加比特率成本的相对权重)。率失真成本通常提供对不同MV精度选项性能的最准确的评估,但还具有最高的计算复杂度。

[0125] 编码器可以将率失真成本函数的一个或多个项改变为使率失真分析朝向整数样本MV精度选项偏置。例如,当利用率失真分析在多个MV精度之间进行决策以确定视频的单位的MV精度时,通过缩放失真成本,对失真成本增加惩罚,缩放比特率成本,向比特率成本增加惩罚和/或调整拉格朗日乘子因子,使率失真分析朝向整数样本MV精度偏置。当评估分数样本MV精度时,编码器可以按比例放大失真成本(通过大于1的因子),按比例放大比特率成本(通过大于1的因子),增加失真惩罚,增加比特率惩罚和/或使用较大的拉格朗日乘子因子。或者,当评估整数样本MV精度时,编码器可以按比例缩小失真成本(通过小于1的因子),按比例缩小比特率成本(通过小于1的因子),和/或使用较小的拉格朗日乘子因子。

[0126] 编码器可以在编码期间改变朝向或背离整数样本MV精度的偏置程度。例如,编码器可以取决于整数样本MV值可能更适合编码视频内容(例如,当视频内容可能是人工创造的内容时增加朝向整数样本MV精度的偏置)的置信度而朝向整数样本MV精度调整偏置。或者,编码器可以取决于编码和/或解码的计算能力而朝向整数样本MV精度调整偏置(例如,当可用计算能力较低时增加朝向整数样本MV精度的偏置)。

[0127] 替代地,编码器可以使用另一方法来评估哪个MV精度提供最佳性能。例如,编码器针对给定的量化步长测量哪个MV精度导致最低比特的编码数据。或者,编码器仅对使用不同MV精度的编码的失真进行评估。或者,编码器使用采样器测量,例如,与整数样本MV精度相比分数样本MV精度的失真减少益处,其可以最够简单以确定单次编码。例如,当使用分数样本MV精度时,与使用整数样本MV精度时相比,编码器检查失真减少的量(按照SAD、SATD、TSE、MSE或其它失真基准)。

[0128] 强力编码方法可以是计算密集的。与使用固定MV精度编码相比,它们可能涉及大量额外计算、额外存储器存储以及额外的存储器读取和写入操作。

[0129] 3. 使用内容分析的方法

[0130] 在选择MV精度的另一方法集中,编码器基于对输入视频内容和/或编码视频内容的分析针对视频的单位选择MV精度。视频的单位可以是块、PB、PU、CU、CTU、子宏块分割、宏块、切片、瓦片、图片、GOP、序列或其它类型的视频的单位。

[0131] 图8示出了在编码期间自适应MV精度的技术(800)。可以通过编码器执行技术(800),所述编码器例如参考图3或图4a和4b所描述的或者通过其它编码器。根据技术(800),在对视频编码期间,编码器从视频的单位的多个MV精度确定MV精度。多个MV精度可以包括一个或多个分数样本MV精度以及整数样本MV精度。例如,多个MV精度可以包括整数样本MV精度和1/4样本MV精度。或者,多个MV精度可以包括整数样本MV精度、1/2样本MV精度和1/4样本MV精度。

[0132] 具体地,当对视频的单位进行编码时,编码器确定(810)是否改变MV精度。在编码开始时,编码器可以根据默认值初始设置MV精度,或者就像改变MV精度一样进行。对于视频的后面的单位,编码器可以使用当前MV精度(其用于一个或多个先前编码的单位)或改变MV精度。例如,编码器可以在发生定义事件时(例如,在对阈值数量的单位进行编码之后,在场景变化之后,在确定已经改变了视频类型之后)决定改变MV精度。

[0133] 为了改变MV精度,编码器收集(820)关于视频的信息。通常,所收集到的信息可以是输入视频的特性或经编码视频的特性。所收集到的信息可以涉及视频的正在被编码的当前单位和/或涉及先前编码的单位。(当所收集到的信息涉及视频的一个或多个先前编码的单位时,可以在编码先前单位之前、期间或之后收集(820)这种信息。该收集(820)不同于图8所示的定时,而且不管关于改变MV精度的决策(810)而发生)。编码器然后至少部分地基于所收集到的信息选择(830)针对视频的单位的MV精度。

[0134] 作为一个例子,编码器可以收集针对当前单位的样本值。少量离散样本值的存在趋向于表示屏幕捕捉内容,并因此暗示应该选择整数样本MV精度。在另一方面,大量离散样本值的存在倾向于指示自然视频,并因此暗示应该选择分数样本MV精度。可以将样本值组织为直方图。可以从YUV颜色空间中的仅亮度(Y)样本、从YUV颜色空间中的亮度和色度(U,V)样本、从RGB颜色空间中的R、G和B样本或从RGB颜色空间中的仅G(或R或B)样本收集样本值。例如,当选择MV精度时,编码器在所收集到的样本值中确定不同的样本值的计数。编码器比较该计数与阈值。如果计数低于阈值,则编码器选择整数样本MV精度。如果计数高于阈值,则编码器选择分数样本MV精度。可以利用任一选项处理边界条件(计数等于阈值),取决于实现方式。或者,编码器考虑来自所收集到的样本值的统计。例如,编码器确定最常见的收集样本值x是否占样本值的多于y%。如果是,则编码器选择整数样本MV精度;否则,编码器选择分数样本MV精度。x和y的值取决于实现方式。x的值可以是10或者一些其它计数。y的值可以是80、90或一些其它小于100的百分比。

[0135] 作为另一例子,编码器可以针对用各自的MV精度编码的当前单位的块收集失真测量。例如,当使用分数样本MV精度时,与整数样本MV精度相比,编码器记录失真中的改善(减小)。当选择MV精度时,编码器确定失真的减小是否证明MV精度的增加。

[0136] 再如,编码器可以收集针对一个或多个先前单位的MV值(具有分数样本MV精度)。可以根据其分数部分的值组织所收集的MV值,例如,对于1/4样本MV精度MV值,在直方图中对于分数部分为0的MV值的二进数(bin),对于分数部分为0.25的MV值的二进数,对于分数部分为0.5的MV值的二进数,以及对于分数部分为0.75的MV值的二进数。在下一段中将描述

该方法的低复杂度变型。

[0137] 例如,编码器可以收集关于使用分数样本MV精度编码的块的MV数据(差分MV值)的编码比特的计数的信息。差分MV值的低平均数量的比特指示常规(可预测)运动,并且在整数样本MV精度合适时更常见。用于差分MV值的高平均数比特在分数样本MV精度合适时更常见。当选择MV精度时,编码器在差分MV值的编码比特的计数中测量比特的平均(或中间)数。编码器比较测量与阈值。如果测量低于阈值,则编码器选择整数样本MV精度。如果测量高于阈值,则编码器选择分数样本MV精度。可以使用任意选项处理边界条件(测量等于阈值),取决于实现方式。

[0138] 再如,当编码单位时,编码器评估单位的每个块(例如,PU)的多个MV精度,并且收集每个块指示哪个MV精度为所述块提供最佳性能的信息。当使用整数样本MV精度对块进行编码时,编码器可以确定率失真成本(例如, $D+\lambda R$);并还在使用分数样本MV精度对块进行编码时,确定率失真成本(例如, $D+\lambda R$)。编码器确定有多少次多个MV精度中的每个对单位中的各个块是最佳的,并选择具有最大计数的MV精度。例如,对于图片中的每个块,当使用整数样本MV精度对块进行编码时,编码器确定率失真成本;并还在使用1/4样本MV精度对块进行编码时,确定率失真成本。编码器对次数整数样本MV精度计数更好,并且对次数1/4样本MV精度计数更好,然后挑选两者中较高者。替代地,编码器确定对多少次整数样本MV精度计数对于单位的块最佳,然后仅在计数高于单位中块的数量的阈值百分比时选择整数样本MV精度。在一些实现方式中,编码器考虑具有MV任意值的块。在其它实现方式中,编码器仅考虑具有非零值MV的块。可以对给定单位的块执行多个MV精度的该块方式的评估,以便选择针对一个或多个后续单位的MV精度,而不管用于给定单位的MV精度模式。或者,可以针对给定单位执行多个MV精度的块方式评估,以便为给定单位选择MV精度。

[0139] 替代地,编码器使用另一方法来收集信息并至少部分地基于所选的信息选择MV精度。

[0140] 返回图8,不管MV精度是否变化,编码器使用所选的MV精度对单位进行编码(840)。在视频的单位内的块(例如,PU、宏块或其它块)的MV值具有所选的MV精度。编码器例如以比特流输出用于当前单位的编码数据。编码数据可以包括指示所选的MV精度的语法元素。

[0141] 编码器决定(850)是否继续下一单位。如果是,则编码器决定(810)是否改变针对下一单位的MV精度。因此,可以选择针对每个单位(例如,每段、每GOP、每图片、每切片、每CTU、每CU、每PU、每PB、每宏块、每子宏块分割)的MV精度。或者,为了减小复杂度,单位的MV精度可以不时变化(例如,周期性或当发生预定事件时),然后对于一个或多个后续单位重复。

[0142] 当编码器逐图片使用相同的瓦片模式时,编码器可以逐图片重复每瓦片MV精度。逐图片共同定位的瓦片可以使用相同的MV精度。类似地,逐图片共同定位的切片可以使用相同的MV精度。例如,假设视频描绘了计算机桌面,并且部分桌面具有显示自然视频内容的窗口。可以逐图片在桌面的区域内使用分数样本MV精度,而显示文本或其它呈现内容的其它区域使用整数样本MV精度编码。

[0143] 在该方法集中,编码器可以使用单次编码。对于正在编码的视频的当前单位,用于当前单位的所选的MV精度至少部分地取决于从视频的一个或多个先前单位所收集到的信息(以编码顺序,其还被称作解码顺序或比特流顺序,而不是输入顺序,其还被称作时间顺

序、输出顺序或显示顺序)。

[0144] 替代地,在该方法集中,编码器可以利用短的预见性窗口使用多次编码或解码(有时称作1.5次编码)。对于正在编码的视频的当前单位,所选的MV精度至少部分地取决于从当前单位收集到的信息。选择的针对当前单位的MV精度还可以至少部分地取决于从视频的一个或多个先前单位收集到的信息(以编码顺序而非输入顺序)。

[0145] 在该方法集中,编码器可以至少部分地基于整数样本MV精度是合适的置信度,调整朝向或背离整数样本MV精度的偏置量。编码器还可以至少部分地基于编码和/或解码的计算能力(如果较少计算能力可用,则优选整数样本MV精度来减少计算复杂度),调整朝向或背离整数样本MV精度的偏置量。例如,为了优选整数样本MV精度,编码器可以调整在比较操作中使用的阈值,以使得更可能选择整数样本MV精度。

[0146] 在该方法集中,所选的MV精度可以用于视频的单位内的块的MV值的水平MV分量和/或垂直MV分量,其中允许水平MV分量和垂直MV分量具有不同的MV精度。或者,所选的MV精度可以用于视频的单位内的块的MV值的水平MV分量和垂直MV分量两者,其中水平MV分量和垂直MV分量具有相同的MV精度。

[0147] 在该方法集中,(例如,比特流中的)编码视频包括指示单位的所选MV精度的一个或多个语法元素。替代地,编码视频可能缺乏指示该单位的所选MV精度的任何语法元素(参见下文,在关于非规范方法的段中)。例如,即使比特流支持用新号传递具有分数样本MV精度的MV值,编码器可以将视频的单位的运动估计约束到仅使用具有分数部分为0的MV值。这可以通过避免插值操作而产生编码和解码的计算复杂度。

[0148] 4. 使用低复杂度内容分析的方法

[0149] 为了简化决策制定过程,编码器可以在选择MV精度之前考虑较小的数据集,或者当选择MV精度时使用较简单的决策逻辑,避免多次编码。

[0150] 图9示出了利用低复杂度方法在编码期间自适应MV精度的技术(900)。可以通过编码器执行技术(900),所述编码器例如是参考图3或图4a和4b所描述的,或者通过其它编码器。技术(900)详述了一种收集关于视频的信息并至少部分地基于收集到的信息选择MV精度的方法,如参考图8所描述的。

[0151] 根据技术(900),在对视频编码期间,编码器确定视频的单位的MV精度。当确定针对所述单位的MV精度时,编码器识别(910)具有分数样本MV精度的MV值集合。可以允许该MV值集合包括零值MV和非零值MV。或者,可以将MV值集合约束为仅包括非零值MV。或者,还可以将MV值集合约束为仅包括来自具有特定块尺寸或更大块尺寸的块的非零值MV。

[0152] 编码器至少部分地基于在MV值集合中分数部分为零的MV值的流行而选择(920)针对单位的MV精度。可以根据分数部分为零的MV值集合的分数来测量所述流行。例如,对于图片,编码器可以确定分数部分为零的MV值的百分比。或者,对于使用MV值集合的区域或区域集,可以根据分数部分为零的区域或区域集的分数测量流行。如果分数超过阈值,则选择的针对单位的MV精度是整数样本MV精度。如果分数没有超过阈值,则选择的针对单位的MV精度是分数样本MV精度。可以使用任意选项处理边界条件(分数等于阈值),取决于实现方式。

[0153] 选择(920)针对单位的MV精度还可以至少部分地基于非零值MV的流行,从而当存在阈值量的非零值MV时允许切换到整数样本MV精度。可以以非零值MV的MV值的分数的形式,以使用非零值MV的块的计数的形式,或者以使用非零值MV的区域或区域集的分数的形式

式,来测量非零值MV的流行。在这种情况下,可以从区域或区域集的非零值MV中识别出具有分数样本MV精度的MV值集合。因此,编码器可以考虑在作为非零值MV的MV集合中分数部分为零的非零值MV的流行。例如,当满足以下两个条件时编码器切换到整数样本MV精度: (1) 检测到足够大量的非零值MV,以及(2)在非零值MV集合中,存在足够多的分数部分为零的非零值MV(或者替代地,足够少的具有非零分数部分的非零值MV)。通过对MV值进行计数(不管其相关联的块大小),或通过考虑MV值的相关联的块大小(例如,因为一些MV值被应用于比其它块大的块中),来确定非零值MV的流行和具有分数部分为零的MV值的流行。

[0154] 编码器使用选择的针对单位的MV精度对该单位进行编码。在视频的单位中的块(例如,PU、宏块或其它块)的MV值具有该单位的选择的MV精度。编码器例如以比特流输出用于当前单位的经编码的数据。经编码的数据可以包括指示选择的针对单位的MV精度的语法元素。

[0155] 为了减少编码器设置MV精度所花费的时间量,在选择了针对单位的整数样本MV精度之后,选择的MV精度可以用于后续视频单位,直到事件使得MV精度切换回分数样本MV精度。例如,事件可以是对定义数量的单位进行编码、场景变化、或基于在编码期间的观察确定切换回分数样本MV精度是有利的。

[0156] 在一个示例性实现方式中,编码器对视频的单位(例如,图片、瓦片、切片或CU)仅编码一次。在开始时,编码器使用1/4样本MV精度对单位编码。在编码期间,编码器确定MV值的分数部分是否是零。例如,编码器测量MV值的什么分数具有非零分数部分。或者,由于一些MV值比其它影响较大的图片区域,所以编码器测量图片间预测区域的什么分数使用具有非零分数部分的MV值(测量区域,而不是MV值的计数)。如果分数超过阈值(其取决于实现方式,例如是75%),则编码器对于视频的一个或多个后续单位切换到整数样本MV精度。

[0157] 在该示例性实现方式中,在编码器切换到整数样本的MV精度之后,编码器可以无限期地保持整数样本MV精度,或者直到至少临时地定义的事件触发切换回分数样本MV精度。例如,所述事件可以是编码特定数量的单位(例如,100个单位)。或者,所述事件可以是场景变化。或者,所述事件可以是基于在编码时收集到的统计,确定切换回分数样本MV精度可能是有利的。(可以在对一些有限量的区域编码期间收集这种统计,以决定分数样本MV精度对该区域更好,然后应用于对于一个或多个单位切换MV精度)。

[0158] 不管视频内容是自然视频内容还是人工创造的视频内容,大部分视频可能是静止的。例如,静止部分可能是自然视频中的固定背景或屏幕捕捉内容中的固定内容。视频的静止部分具有零值MV,当MV精度是分数样本MV精度时其分数部分为零。存在大量零值MV将使得考虑具有非零分数部分的MV值的分数的决策逻辑混乱。

[0159] 因此,编码器可以从考虑中排除零值MV。图10示出了包括(大部分)具有零值MV的非移动部分(1001)和(大部分)具有非零值MV的两个移动部分(1002、1003)的图片1000。编码器考虑移动部分(1002、1003)中的非零值MV,但是不考虑非移动部分(1001)的MV值。当分数部分为零的非零值MV的分数(在移动部分(1002、1003)中)超过阈值时(或者当使用分数部分为零的非零MV的图片的分数(以区域的形式)超过阈值时),编码器可以切换到整数样本MV精度。

[0160] 编码器还可以检查被评估超过阈值量的非零值MV的数量,从而基于少量MV值而不做出决策。这可以使得决策制定过程更健壮。

[0161] 在另一示例性实现方式中,编码器使用1/4样本MV精度对给定的视频的单位(例如,图片、瓦片、切片或CU)进行编码。如果(1)多于x%的单位使用具有非零值MV的图片间预测,以及(2)多于y%的使用非零MV的单位部分具有整数值MV(分数部分为零),则编码器对于视频的一个或多个后续单位切换到整数样本MV精度。x和y的值取决于实现方式,并且例如可以分别是5和75。

[0162] 在类似的示例性实现方式中,编码器使用1/4样本MV精度对给定的视频的单位(例如,图片、瓦片、切片或CU)进行编码。如果(1)单位中多于z个PU具有非零值MV,以及(2)那些PU中多于y%具有整数值MV(分数部分为零),则编码器对于视频的一个或多个后续单位切换到整数样本MV精度。z和y的值取决于实现方式,并且例如可以分别是100和75。

[0163] 较大区域的MV值可能比较小区域的MV值更可靠。编码器可以限定评估那些MV值。例如,编码器可以仅为特定块尺寸或更大(例如,16x16或更大)的块评估MV值。

[0164] 在另一示例性实现方式中,编码器使用1/4样本MV精度对给定的视频的单位(例如,图片、瓦片、切片或CU)进行编码。如果(1)单位中多于z个PU是w x w或更大并具有非零值MV,以及(2)那些PU中多于y%具有整数值MV(分数部分为零),则编码器对于视频的一个或多个后续单位切换到整数样本MV精度。w、z和y的值取决于实现方式,并且例如可以分别是16、100和75。

5. 非规范方法

[0165] 在大部分先前的例子中,编码器在比特流中用信号传递在编码数据中指示选择的MV精度的一个或多个语法元素。解码器解析指示选择的MV精度的语法元素,并根据选择的MV精度解释MV值。

[0166] 替代地,在非规范方法中,编码器并不用信号传递指示编码器所选的MV精度的任何语法元素。例如,编码器在整数样本MV精度和分数样本MV精度之间进行选择,但是总是在分数样本MV精度处对MV值进行编码。编码器在分数样本MV精度处重构并应用MV值。

[0167] 当选择整数样本MV精度时,编码器可以通过避免分数样本偏移处的样本值的插值,以及通过评估仅在整数样本偏移处的候选预测区域,来简化运动估计。另外,如果MV预测产生分数值(例如,使用时间MV预测),则编码器可以仅考虑这些MV差,其可能导致在向分数值MV预测添加MV差时的整数值(例如,来自临时MV预测)。在解码期间,可以通过避免在分数样本偏移处的样本值的插值而简化运动补偿。

[0168] 在前述段中描述的某些方法(例如,利用通过缩放失真成本和/或比特率成本的缩放率失真成本,或添加失真成本惩罚或比特率成本惩罚,或调整权重因子)还可以适应于非规范方法。编码器可以在编码期间改变朝向或背离整数样本MV精度的偏置程度。通过缩放、惩罚和/或权重因子,编码器可以取决于整数样本MV值可能更适于编码视频内容的置信度,或取决于编码或解码的计算能力,而朝向整数样本MV精度调整偏置。

6. 备选和变型

[0169] 在一些使用场景中,图片的编码顺序(也称作解码顺序或解码的顺序)不同于输入/摄像机捕捉和显示的时间顺序(也称作显示顺序)。当选择MV精度时,编码器可以考虑这种重排序。例如,编码器可以基于图片的时间顺序而不是图片的编码顺序来选择MV精度。

[0170] 在本文描述的多个例子中,以分开的部件或过程实现BC内预测和运动补偿,并且以分开的部件或过程实现BV估计和运动估计。替代地,BC内预测可以实现为运动补偿的特

殊情况，并BV估计可以实现为运动估计的特殊情况，对此，将当前图片用作参考图片。在这种实现方式中，BV值可以作为MV值用信号传递，但是用于BC内预测（在当前图片内）而不是图片间预测。如本文使用的术语，“BC内预测”表示在当前图片内的预测，不管是使用图片内预测模块、运动补偿模块还是些其它模块提供预测。类似地，可以使用MV值或使用不同类型的参数或语法元素来呈现BV值，并且可以使用图片内估计模块、运动估计模块或一些其它模块来提供BV估计。本文所描述的用于选择MV精度的方法可以应用于确定将用作BC内预测的BV值的MV值的精度（也就是，当前图片作为参考图片）。

[0173] VI. 创新特征

[0174] 除了下文呈现的权利要求外，本文描述的创新特征包括但不限于以下：

#	特征
A1	<p>一种计算设备，包括：</p> <p>用于对视频进行编码的单元，包括用于确定针对视频的单位的运动向量（“MV”）精度的单元，其中在所述视频的单位中的块的 MV 值具有针对所述单位的 MV 精度，并且其中用于确定针对所述单位的 MV 精度的单元包括：</p> <p>用于识别具有分数样本 MV 精度的一组 MV 值的单元；以及</p> <p>用于至少部分地基于在一组 MV 值内分数部分为零的 MV 值的流行来选择用于所述单位的 MV 精度的单元；以及</p> <p>用于输出经编码的视频的单元。</p>
B1 [0175]	<p>一种计算设备，包括：</p> <p>用于对视频进行编码的单元，包括用于确定针对视频的单位的运动向量（“MV”）精度的单元，其中在视频的单位内的块的 MV 值具有所述单位的 MV 精度，其中用于确定的单元包括：用于执行率失真分析以在多个 MV 精度之间进行决定的单元，所述多个 MV 精度包括一个或多个分数样本 MV 精度和整数样本 MV 精度，并且其中通过 (a) 缩放失真成本，(b) 向失真成本添加惩罚，(c) 缩放比特率成本，(d) 向比特率成本添加惩罚，和/或 (e) 调整拉格朗日乘子因子；以及</p> <p>用于输出经编码的视频的单元。</p>
C1	<p>一种计算设备，包括：</p> <p>用于对视频进行编码的单元，包括用于从多个 MV 精度中确定针对视频的单位的运动向量（“MV”）精度的单元，所述多个 MV 精度包括一个或多个分数样本 MV 精度和整数样本 MV 精度，其中在视频的单位内的块的 MV 值具有所述单位的 MV 精度，并且其中用于确定的单元包括：</p>
[0176]	<p>用于收集关于视频的信息的单元；以及</p> <p>用于至少部分地基于收集的信息选择针对所述单位的 MV 精度的单元；以及</p> <p>用于输出经编码的视频的单元。</p>

[0177] 考虑到可以应用所公开的本发明的原理的多种可能实施例，可以认识到的是，图示实施例仅是本发明的优选示例，而不应该认为是限制本发明的范围。而是，本发明的范围

由后续权利要求所定义。因此,我们要求在这些权利要求的范围和精神内的所有内容为我们的发明。

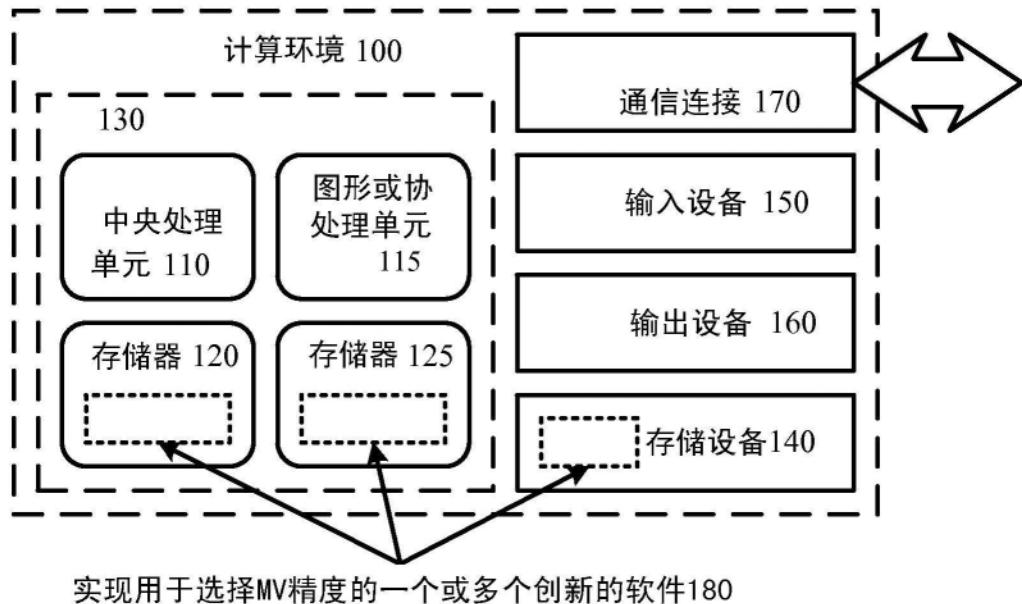


图1

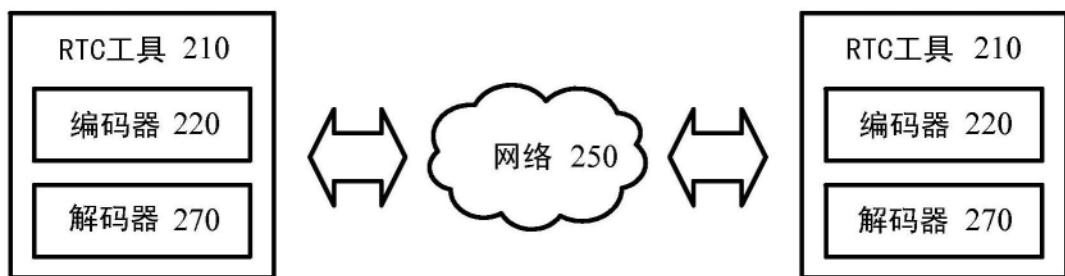
201

图2a

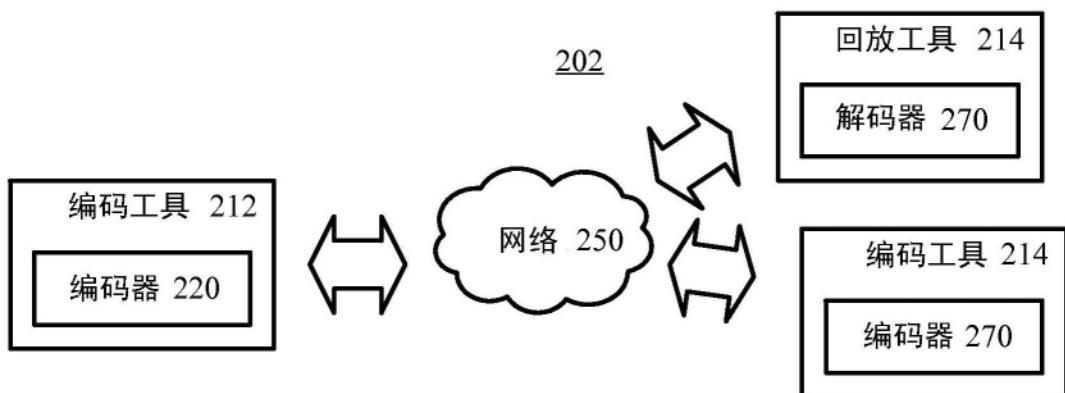


图2b

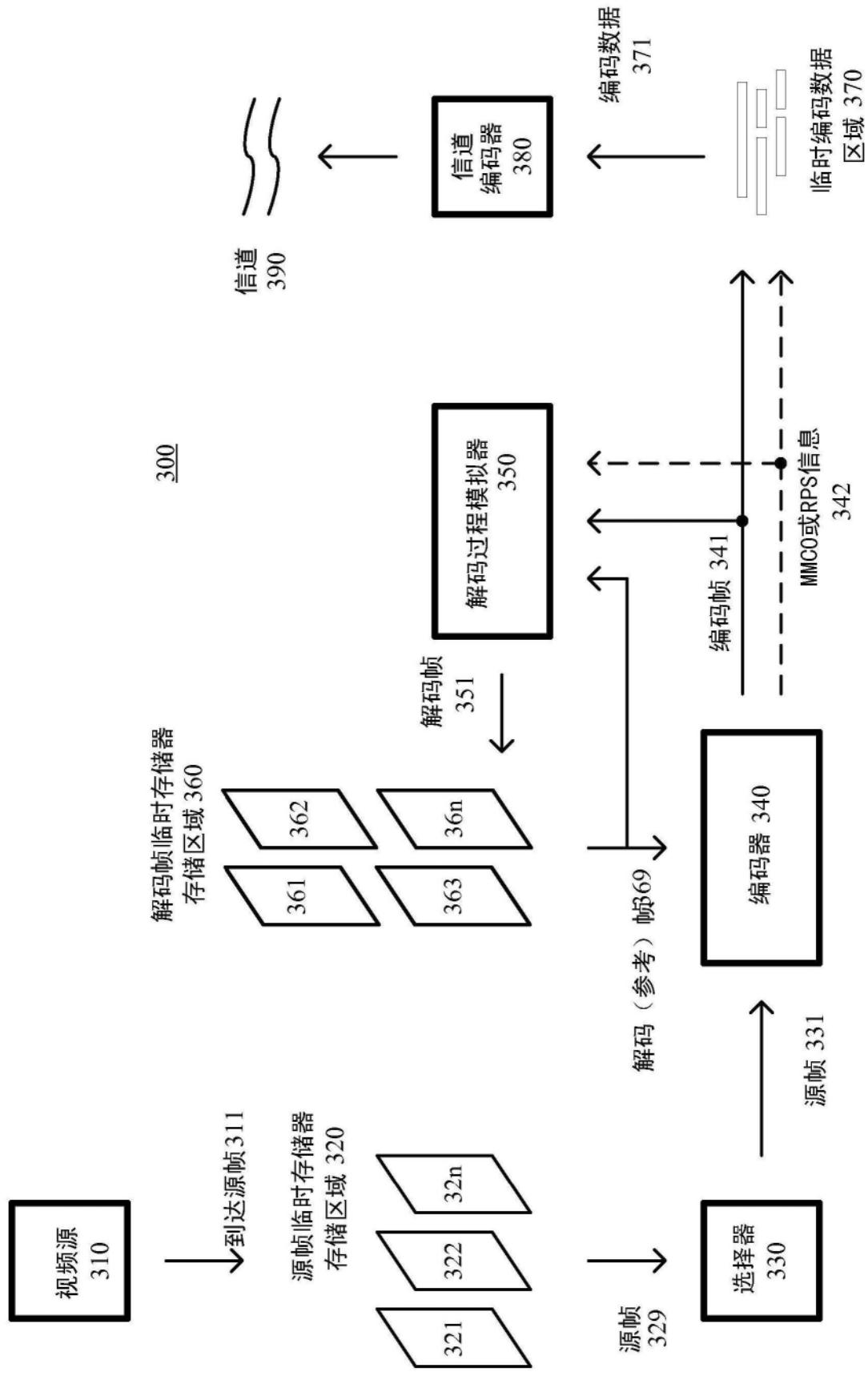


图3

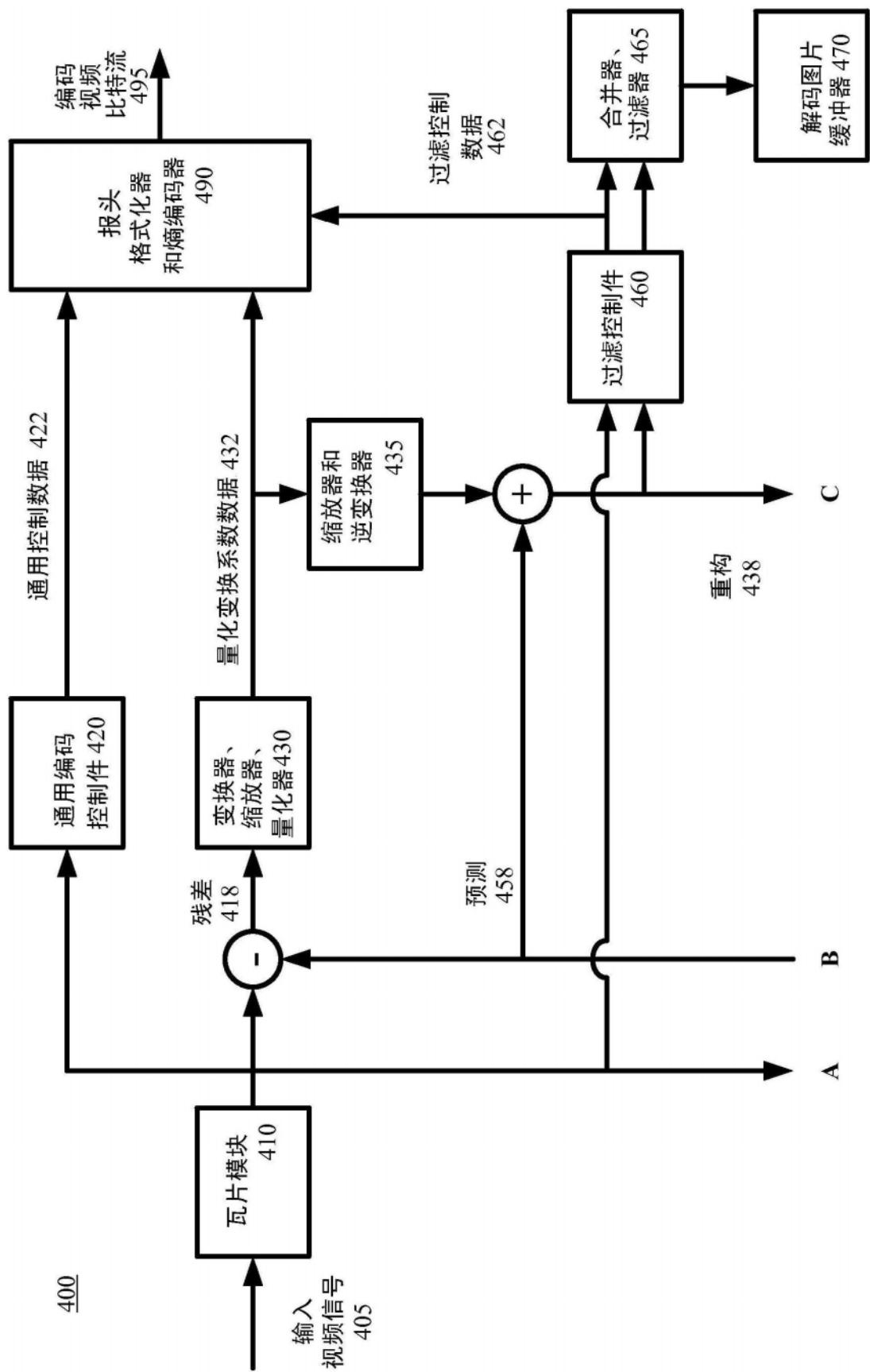


图 4a

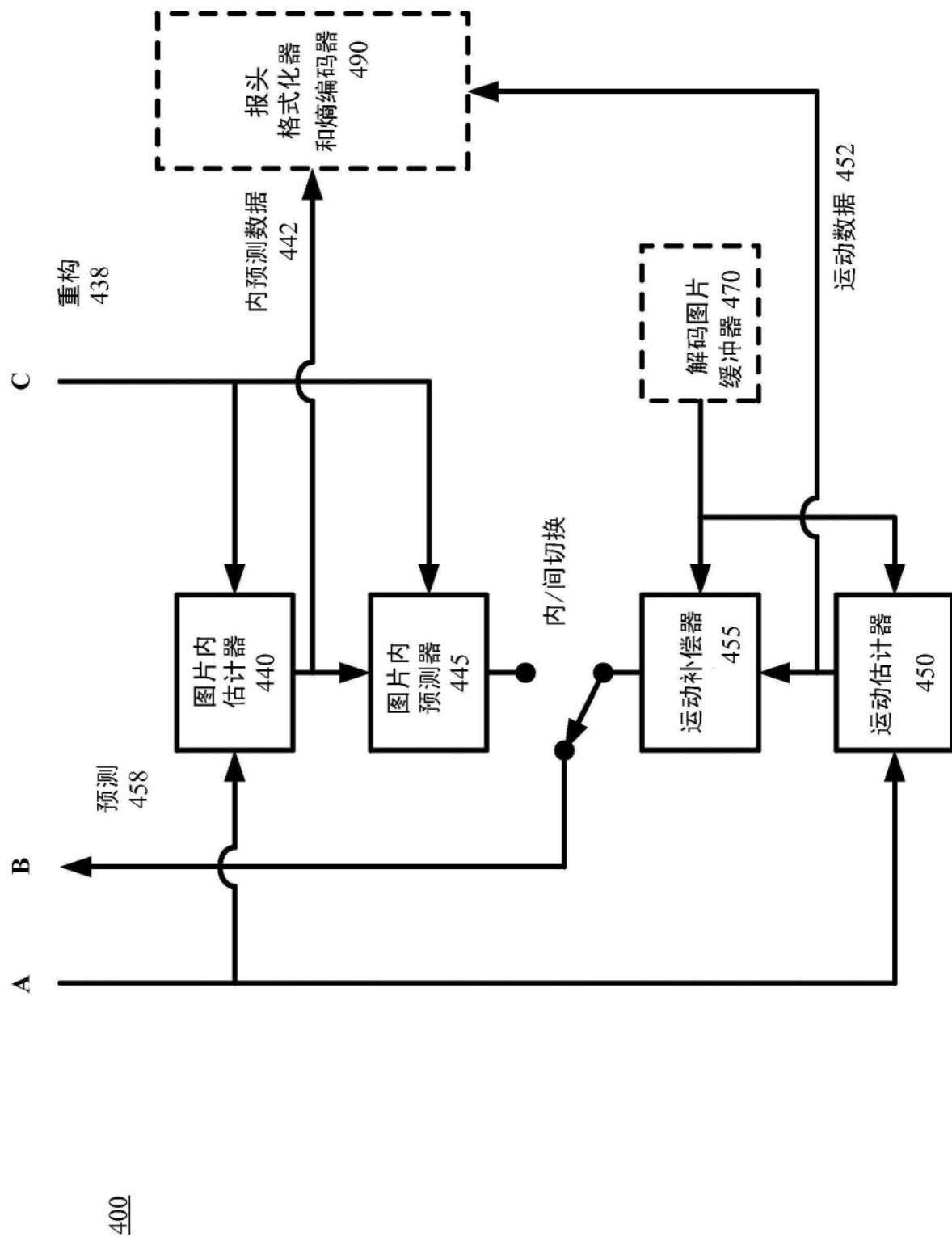
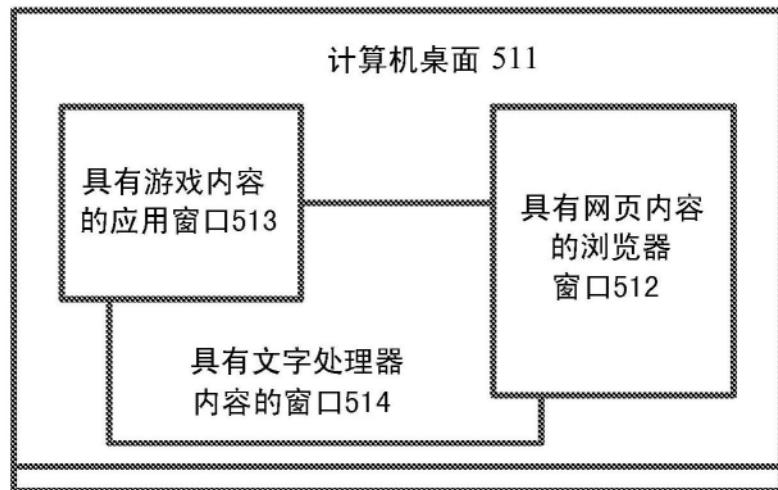
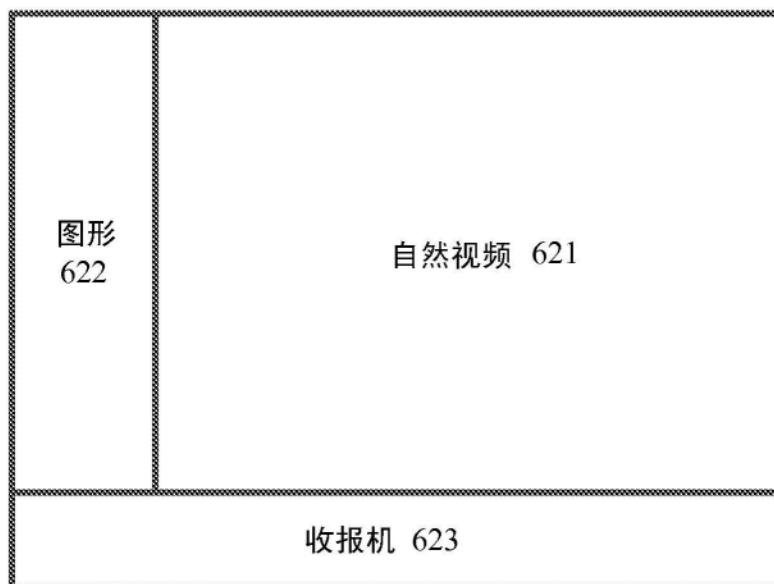


图4b



具有可能为屏幕捕捉提供输入的内容的计算机桌面环境 (510)

图5



具有自然视频内容和基于调色板内容的复合视频 (620)

图6

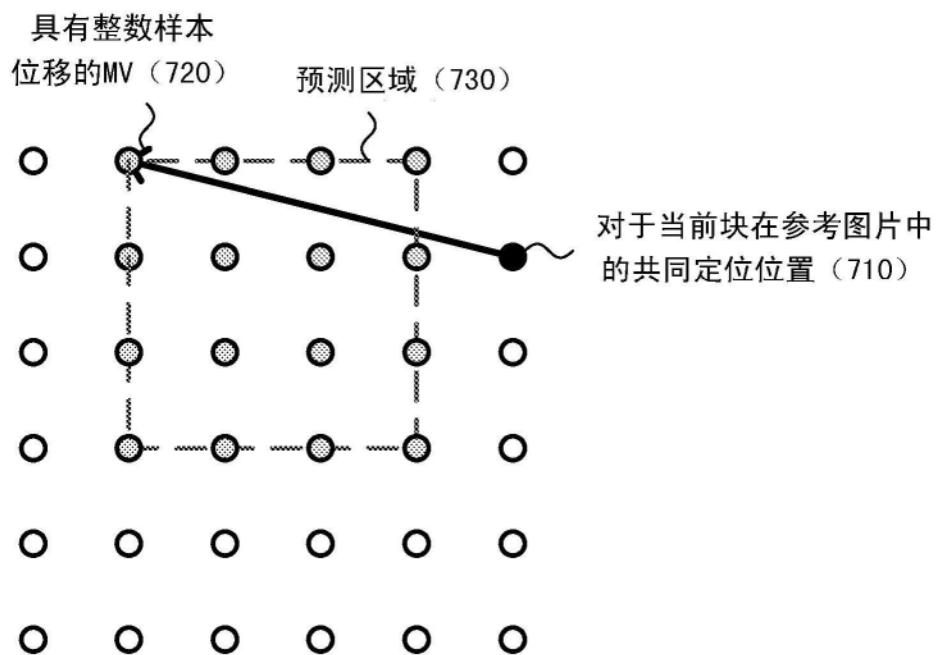
701

图7a

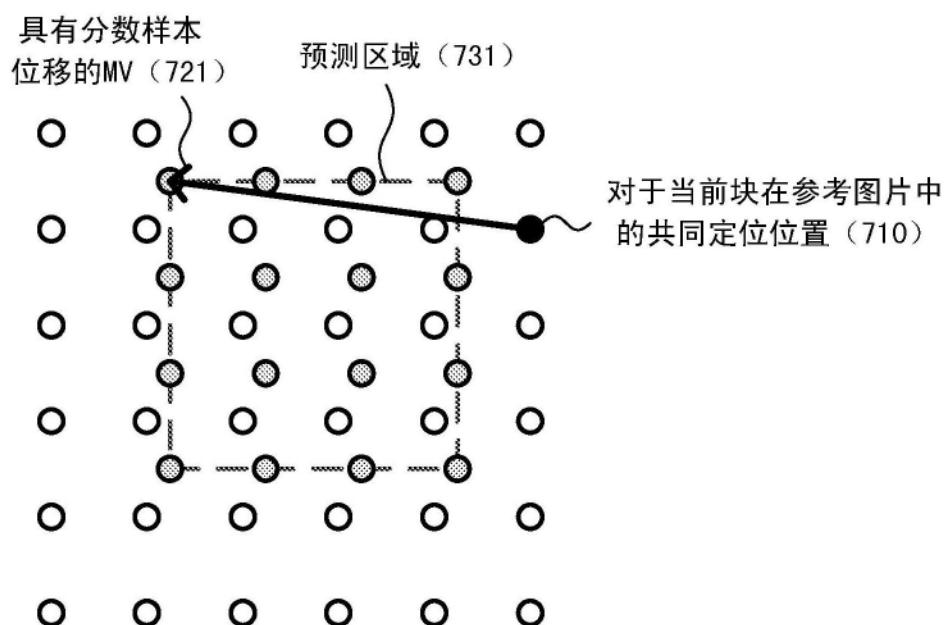
702

图7b

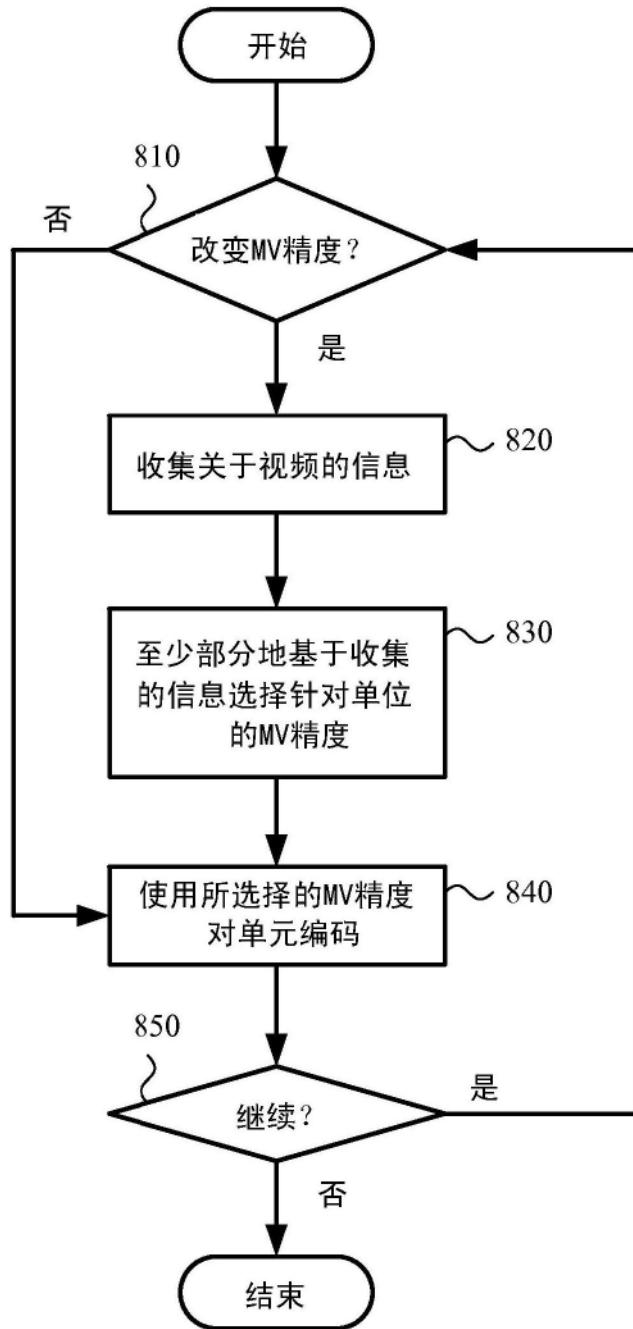
800

图8

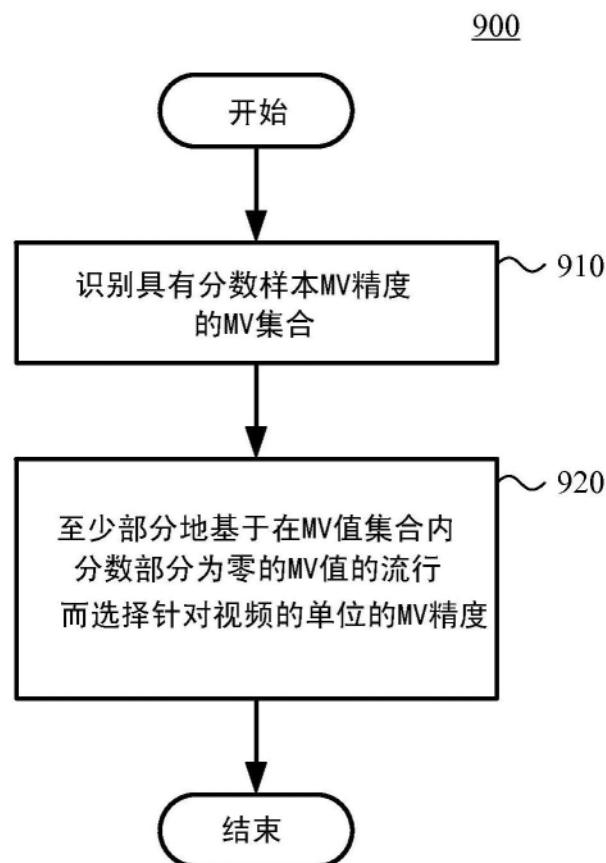
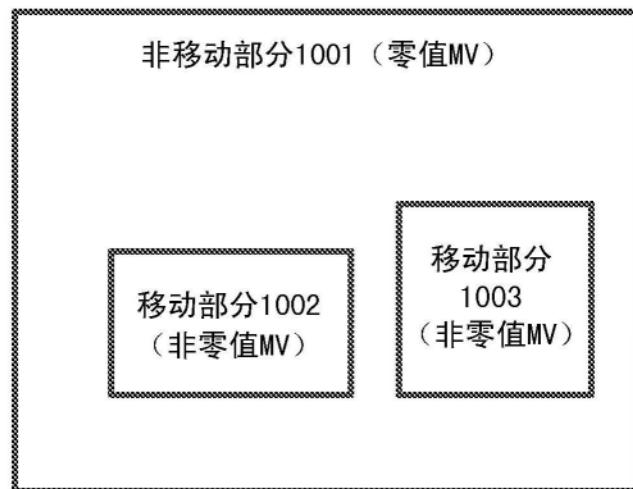


图9



通过视频编码器分析的图片1000

图10