



(12) 发明专利申请

(10) 申请公布号 CN 118870028 A

(43) 申请公布日 2024. 10. 29

(21) 申请号 202411050677.6

(22) 申请日 2020.11.02

(30) 优先权数据

63/020,280 2020.05.05 US

17/072,606 2020.10.16 US

(62) 分案原申请数据

202080062531.7 2020.11.02

(71) 申请人 腾讯美国有限责任公司

地址 美国加利福尼亚州帕洛阿尔托公园大道2747号

(72) 发明人 赵欣 芮世薰 刘杉

(74) 专利代理机构 深圳市深佳知识产权代理事

务所(普通合伙) 44285

专利代理师 张晓

(51) Int.Cl.

H04N 19/61 (2014.01)

H04N 19/12 (2014.01)

H04N 19/44 (2014.01)

H04N 19/124 (2014.01)

H04N 19/186 (2014.01)

H04N 19/70 (2014.01)

H04N 19/91 (2014.01)

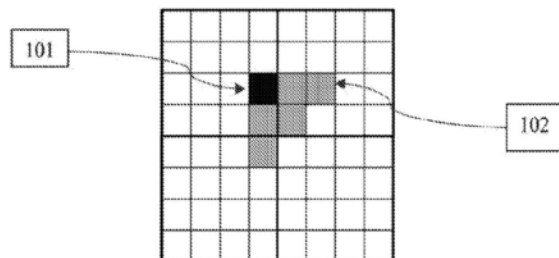
权利要求书2页 说明书12页 附图6页

(54) 发明名称

视频编解码的方法和装置

(57) 摘要

一种使用至少一个处理器对编码后的视频比特流进行解码的方法包括:获得编码后的视频比特流,该编码后的视频比特流包括编码后的颜色分量;对编码后的颜色分量进行熵解析;对颜色分量进行去量化并获得颜色分量的变换系数;对颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出;对JCST输出执行反向变换,从而获得颜色分量的残差分量;以及基于颜色分量的残差分量对编码后的视频比特流进行解码。



1. 一种对编码后的视频比特流进行解码的方法,其特征在于,所述方法包括:
获得编码后的视频比特流,所述编码后的视频比特流包括编码后的颜色分量;
对所述编码后的颜色分量进行熵解析;
对所述熵解析后的颜色分量进行去量化并获得所述颜色分量的变换系数;
对所述颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出;
对所述JCST输出执行反向变换,从而获得所述颜色分量的残差分量;以及
基于所述颜色分量的残差分量对所述编码后的视频比特流进行解码。
2. 根据权利要求1所述的方法,其特征在于,所述变换系数包括Cb和Cr变换系数。
3. 根据权利要求1所述的方法,其特征在于,所述变换系数包括Y、Cb和Cr变换系数。
4. 根据权利要求1-3中任一项所述的方法,其特征在于,逐元素地应用所述JCST。
5. 根据权利要求1所述的方法,其特征在于,所述颜色分量包括位于不同坐标处的多于一对的不同颜色分量。
6. 根据权利要求1-3或5中任一项所述的方法,其特征在于,针对有限范围的块大小应用所述JCST。
7. 根据权利要求1-3或5中任一项所述的方法,其特征在于,还包括:
从所述编码后的视频比特流中获得包括与编码后的图片对应的图片单元的编码后的视频序列(CVS);
获得包括在所述图片单元中的图片标头(PH)网络抽象层(NAL)单元;
获得包括在所述图片单元中的至少一个视频编码层(VCL)NAL单元;
对JCST标志进行解析,所述JCST标志用于在变换块级别处用信号通知何时将应用所述JCST。
8. 根据权利要求1-3或5中任一项所述的方法,其特征在于,还包括:
从所述编码后的视频比特流中获得包括与编码后的图片对应的图片单元的编码后的视频序列(CVS);
获得包括在所述图片单元中的图片标头(PH)网络抽象层(NAL)单元;
获得包括在所述图片单元中的至少一个视频编码层(VCL)NAL单元;
对JCST标志进行解析,所述JCST标志用于在CU或CB级别处用信号通知何时将应用所述JCST。
9. 根据权利要求1-3或5中任一项所述的方法,其特征在于,还包括:
从所述编码后的视频比特流中获得包括与编码后的图片对应的图片单元的编码后的视频序列(CVS);
获得包括在所述图片单元中的图片标头(PH)网络抽象层(NAL)单元;
获得包括在所述图片单元中的至少一个视频编码层(VCL)NAL单元;
对JCST标志进行解析,所述JCST标志用于用信号通知何时经由高级语法应用所述JCST。
10. 根据权利要求1-3或5中任一项所述的方法,其特征在于,所述JCST包括经由编码信息确定的第二变换。
11. 一种生成视频比特流的视频编码方法,其特征在于,所述方法包括:
获得待编码的视频数据,所述视频数据包括颜色分量;

对所述颜色分量的残差分量进行变换,获得所述颜色分量的变换系数;
对所述颜色分量的变换系数应用联合分量二次变换(JCST),生成JCST输出;
对所述JCST输出进行量化;
对量化后的所述颜色分量进行熵编码;以及
输出所述熵编码后的视频比特流。

12. 一种对编码后的视频比特流进行解码的装置,其特征在于,所述装置包括:
存储器,所述存储器存储指令;以及
处理器,所述处理器与所述存储器通信,并且被配置成在执行指令时使所述装置执行
根据权利要求1至10中任一项所述的方法。

13. 一种生成视频比特流的编码装置,其特征在于,所述装置包括:
存储器,所述存储器存储指令;以及
处理器,所述处理器与所述存储器通信,并且被配置成在执行指令时使所述装置执行
根据权利要求1至11中任一项所述的方法。

14. 一种非暂态计算机可读介质,其特征在于,所述非暂态计算机可读介质存储有指令,所述指令在由计算机执行以使所述计算机执行根据权利要求1至11中任一项所述的方法。

15. 一种处理视频比特流的方法,包括:
在非暂态计算机可读介质上存储视频比特流,所述非暂态计算机可读介质存储有指令,所述指令在由计算机执行以使所述计算机执行根据权利要求1至11中任一项所述的方法。

视频编解码的方法和装置

相关申请的交叉引用

[0001] 本申请要求于2020年5月5日提交的美国临时专利申请第63/020,280号以及于2020年10月16日提交的美国专利申请第17/072,606号的优先权,上述专利申请的全部内容并入本文中。

技术领域

[0002] 本公开内容总体上涉及数据处理领域,并且更具体地涉及一种对编码后的视频比特流进行解码的方法和装置。

[0003]

背景技术

[0004] AOMedia视频1(AOMedia Video 1,AV1)是设计用于通过因特网进行视频传输的开放视频编码格式。作为VP9的后继者,AV1由开放媒体联盟(Alliance for Open Media,AOMedia)开发,该联盟成立于2015年,其包括半导体公司、视频点播提供商、视频内容生产商、软件开发公司和网络浏览器供应商。AV1项目的许多分项源自AOMedia成员先前的研究工作。个人贡献者在多年前启动了实验性技术平台。例如,Xiph's/Mozilla的Daala于2010年发布了代码;谷歌的实验性VP9进化项目VP10于2014年9月12日公布;以及思科的Thor于2015年8月11日发布。

[0005] 在VP9代码库的基础上,AV1结合了另外的技术,这些技术中的一些是以这些实验格式开发的。AV1参考编解码器的第一版本(0.1.0)于2016年4月7日发布。AOMedia于2018年3月28日宣布发布AV1比特流规范,以及基于参考软件的编码器和解码器。2018年6月25日,发布了AV1规范的验证版本1.0.0。2019年1月8日,AV1规范的验证版本1.0.0与勘误表1一起发布。AV1比特流规范包括参考视频编解码器。

发明内容

[0006] 在AV1中,为色度通道生成的预测残差信号例如Cb和Cr彼此高度相关。因此,现有的残差编码方式有待改进。

[0007] 本公开内容的实施方式提供了用于对编码后的视频比特流进行解码的方法和装置,其可以通过减少Cb与Cr预测残差之间的统计冗余来进一步改进残差编码。

[0008] 本公开内容的一个实施方式提供了一种对编码后的视频比特流进行解码的方法,包括:获得编码后的视频比特流,该编码后的视频比特流包括编码后的颜色分量;对编码后的颜色分量进行熵解析;对熵解析后的颜色分量进行去量化并获得颜色分量的变换系数;对颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出;对JCST输出执行反向变换,从而获得颜色分量的残差分量;以及基于颜色分量的残差分量对所述编码后的视频比特流进行解码。

[0009] 本公开内容的一个实施方式提供了一种对编码后的视频比特流进行解码的装置,

包括:至少一个存储器,被配置成存储计算机程序代码;以及至少一个处理器,被配置成访问所述至少一个存储器并根据所述计算机程序代码进行操作,所述计算机程序代码包括:第一获得代码,被配置成使所述至少一个处理器获得编码后的视频比特流,该编码后的视频比特流包括编码后的颜色分量;第一解析代码,被配置成使所述至少一个处理器对编码后的颜色分量进行熵解析;去量化代码,被配置成使所述至少一个处理器对熵解析后的颜色分量进行去量化并获得颜色分量的变换系数;联合分量二次变换(JCST)应用代码,被配置成使所述至少一个处理器对颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出;反向变换代码,被配置成使所述至少一个处理器对JCST输出执行反向变换,从而获得颜色分量的残差分量;以及解码代码,被配置成使至所述少一个处理器基于颜色分量的残差分量对所述编码后的视频比特流进行解码。

[0010] 本公开内容的一个实施方式提供了一种非暂态计算机可读存储介质,其存储有指令,所述指令使至少一个处理器进行以下操作:获得编码后的视频比特流,该编码后的视频比特流包括编码后的颜色分量;对编码后的颜色分量进行熵解析;对熵解析后的颜色分量进行去量化并获得颜色分量的变换系数;对颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出;对JCST输出执行反向变换,从而获得颜色分量的残差分量;以及基于颜色分量的残差分量对所述编码后的视频比特流进行解码。

[0011] 本公开内容的实施方式提供了一种用于对编码后的视频比特流进行解码的方法和装置。首先获得编码后的视频比特流,该编码后的视频比特流包括编码后的颜色分量。对编码后的颜色分量进行熵解析。对熵解析后的颜色分量进行去量化并获得颜色分量的变换系数。对颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出。对JCST输出执行反向变换,从而获得颜色分量的残差分量。基于颜色分量的残差分量对所述编码后的视频比特流进行解码。以上述方式,通过减少为色度通道生成的预测残差信号例如Cb和Cr之间的统计冗余来改进残差编码。

附图说明

[0012] 图1是由局部模板覆盖的编码系数的示意性图示。

[0013] 图2是根据实施方式的通信系统的框图。

[0014] 图3是根据实施方式的基于图的点云压缩(graph-based point cloud compression,G-PCC)压缩器和G-PCC解压缩器在环境中的放置方式的图。

[0015] 图4是根据实施方式的编码器/解码器方案的示意性图示。

[0016] 图5是根据实施方式的编码器/解码器方案的示意性图示。

[0017] 图6是根据实施方式的来自两个 4×2 块的Cb和Cr变换系数对的示意性图示。

[0018] 图7是根据实施方式的应用于两个 4×2 Cb和Cr块上的JCST的示意性图示。

[0019] 图8是根据实施方式的使用四点变换的JCST的示意性图示。

[0020] 图9是示出根据实施方式的解码方法的流程图。

[0021] 图10是适于实现各实施方式的计算机系统的图。

具体实施方式

[0022] 本文中描述的实施方式提供了用于对图像数据进行编码和/或解码的方法和装

置。本公开内容的实施方式涉及用于对来自多个颜色分量的残差例如来自两个色度分量的残差进行编码的JCST方案。

[以AV1进行残差编码]

[0023] 对于每个给定的变换单元,AV1系数编码器从对跳过符号进行编码开始,接下来是对变换核类型和所有非零系数的块结束(end-of-block,E0B)位置进行编码(在不会跳过变换编码时)。然后每个系数值可以映射至多个级别图和符号,其中符号平面覆盖了系数的符号和三个级别平面,每个系数值可以对应于不同的系数幅度范围,即较低级别平面、中级别平面和较高级别平面。较低级别平面对应于0至2的范围,中级别平面对应于3至14的范围,以及较高级别平面对应于15及大于15的范围。

[0024] 在对E0B位置进行编码之后,将较低级别平面和中级别平面按反向扫描顺序一起编码,较低级别平面指示系数幅度是否在0与2之间,以及中级别平面指示系数幅度是否在3与14之间。然后将符号平面和较高级别平面按前向扫描顺序一起编码,并且较高级别平面指示具有幅度大于14的残差值。剩余部分使用Exp-Golomb码进行熵编码。AV1采用传统的锯齿形扫描顺序。

[0025] 这样的分离允许将丰富上下文模型分配给较低级别平面,其考虑了包括双向、水平和垂直的变换方向,变换大小,以及在适度的上下文模型大小下提高压缩效率的多达五个相邻系数。中级别平面使用类似于较低级别平面的上下文模型,其中上下文邻居系数的数量从5减少至2。较高级别平面使用Exp-Golomb码进行编码,而不使用上下文模型。在符号平面中,使用相邻变换单元的DC符号作为上下文信息对除DC符号外的符号进行编码。其他符号位直接编码而不使用上下文模型。

[0026] 在通用视频编码(Versatile Video Coding,VVC)中,首先将编码块分割为 4×4 子块,并且根据预定义的扫描顺序对编码块内部的子块以及子块内的变换系数进行编码。对于具有至少一个非零变换系数的子块,变换系数的编码被分成四个扫描阶段(pass)。

[0027] 例如,假设absLevel是当前变换系数的绝对值。在第一阶段中,对语法元素sig_coeff_flag(指示absLevel大于0)、par_level_flag(指示absLevel的奇偶校验)、以及rem_abs_gt1_flag(指示 $(\text{absLevel}-1) \gg 1$ 大于0)进行编码;在第二阶段中,对语法元素rem_abs_gt2_flag(指示absLevel大于4)进行编码;在第三阶段中,调用系数级别的剩余值(被称为abs_remainder);并且如有必要,在第四阶段中,对符号信息进行编码。

[0028] 为了利用变换系数之间的相关性,在当前系数的上下文选择中,使用图1中所示的由局部模板覆盖的先前编码的系数,其中以黑色显示的位置(101)指示当前变换系数的位置,并且以阴影显示的位置(102)指示它的五个邻居。此处,absLevel1[x][y]表示第一阶段后系数在其位置(x,y)处的部分重构绝对水平,d表示当前系数的对角线位置($d=x+y$),numSig表示局部模板中非零系数的数量,以及sumAbs1表示由局部模板覆盖的系数的部分重构绝对水平absLevel1[x][y]的总和。

[0029] 在对当前系数的sig_coeff_flag进行编码时,根据sumAbs1和对角线位置d选择上下文模型索引。更具体地说,对于Luma分量,根据如下式1确定上下文模型索引:ctxSig= $18 * \max(0, \text{状态}-1) + \min(\text{sumAbs1}, 5) + (d < 2 ? 12 : (d < 5 ? 6 : 0))$ 。这相当于下面的式2与式3。式2:ctxIdBase= $18 * \max(0, \text{状态}-1) + (d < 2 ? 12 : (d < 5 ? 6 : 0))$ 。式3:ctxSig=ctxIdSigTable[$\min(\text{sumAbs1}, 5)$]+ctxIdBase。

[0030] 对于色度,根据如下式4确定上下文模型索引: $ctxSig = 12 * \max(0, 状态 - 1) + \min(\text{sumAbs1}, 5) + (d < 2 ? 6 : 0)$,这相当于下面的式5与式6。式5: $ctxIdBase = 12 * \max(0, 状态 - 1) + (d < 2 ? 6 : 0)$ 。式6: $ctxSig = ctxIdSigTable[\min(\text{sumAbs1}, 5)] + ctxIdBase$ 。

[0031] 此处,如果启用相关量化并且使用状态转换过程导出状态,则使用标量量化器;表 $ctxIdSigTable$ 存储了上下文模型索引偏移量, $ctxIdSigTable[0 \sim 5] = \{0, 1, 2, 3, 4, 5\}$ 。

[0032] 在对当前系数的 par_level_flag 进行编码时,根据 sumAbs1 、 numSig 和对角线位置 d 来选择上下文模型索引。更具体地说,对于Luma分量,根据如下式7确定上下文模型索引: $ctxPar = 1 + \min(\text{sumAbs1} - \text{numSig}, 4) + (d == 0 ? 15 : (d < 3 ? 10 : (d < 10 ? 5 : 0)))$,这相当于下面的式8与式9。式8: $ctxIdBase = (d == 0 ? 15 : (d < 3 ? 10 : (d < 10 ? 5 : 0)))$ 。式9: $ctxPar = 1 + ctxIdTable[\min(\text{sumAbs1} - \text{numSig}, 4)] + ctxIdBase$ 。对于色度,根据如下式10确定上下文模型索引: $ctxPar = 1 + \min(\text{sumAbs1} - \text{numSig}, 4) + (d == 0 ? 5 : 0)$,这相当于下面的式11与式12。式11: $ctxIdBase = (d == 0 ? 5 : 0)$ 。式12: $ctxPar = 1 + ctxIdTable[\min(\text{sumAbs1} - \text{numSig}, 4)] + ctxIdBase$ 。

[0033] 此处,表 $ctxIdTable$ 存储了上下文模型索引偏移量, $ctxIdTable[0 \sim 4] = \{0, 1, 2, 3, 4\}$ 。在对当前系数的 $rem_abs_gt1_flag$ 和 $rem_abs_gt2_flag$ 进行编码时,以与 par_level_flag 相同方式确定上下文模型索引: $ctxGt1 = ctxPar$ 和 $ctxGt2 = ctxPar$ (式13)。

[0034] 对于 $rem_abs_gt1_flag$ 和 $rem_abs_gt2_flag$ 使用不同的上下文模型集。因此,即使 $ctxGt1$ 等于 $ctxGt2$,用于 $rem_abs_gt1_flag$ 的上下文模型与用于 $rem_abs_gt2_flag$ 的上下文模型也不同。

[用于变换跳过模式(Transform Skip Mode, TSM)和差分脉冲编码调制(Differential Pulse-Code Modulation, DPCM)的残差编码]

[0035] 为了使残差编码适于表示量化预测残差(空间域)的变换跳过和块差分脉冲编码调制(Block Differential Pulse-Code Modulation, BDPCM)残差级别的统计和信号特征,提出在以上AV 1的残差编码部分中描述的残差编码方案之上修改以下残差编码过程,并应用于TSM和BDPCM模式。

[0036] 现在描述三个编码阶段。在第一编码阶段中,首先在一个阶段中对 sig_coeff_flag 、 $coeff_sign_flag$ 、 $abs_level_gt1_flag$ 、 par_level_flag 进行编码。在第二编码阶段中,对 $abs_level_gtX_flag$ 进行编码,其中 X 可以是3、5、7、……。在第三阶段中,对系数级别的其余部分进行编码。在系数组(Coefficient Group, CG)级别操作所述编码阶段,即对于每个CG,执行三个编码阶段。

[0037] 没有有效扫描位置。由于残差信号反映了预测后的空间残差,并且没有对TS进行由变换引起的能量压缩,因此没有赋予变换块的右下角处拖尾零或无意义级别的高概率。因此,在这种情况下,省略了最后的有效扫描位置信令。相反,待处理的第一个子块是变换块内最右下角的子块。

[0038] 现在描述子块编码块标志(Coding Block Flag, CBF)。由于最后有效扫描位置信令的缺失,需要对TS的带有 $coded_sub_block_flag$ 的子块CBF信令进行如下修改:

[0039] 由于量化,前述的无意义序列可能仍会局部地出现在变换块内。因此,如前所述移除最后的有效扫描位置,并且针对所有子块对 $coded_sub_block_flag$ 进行编码。

[0040] 覆盖DC频率位置的子块(左上子块)的 $coded_sub_block_flag$ 呈现特殊情况。在

VVC草案3中,该子块的coded_sub_block_flag从不用信号通知,并被推断为等于1。当最后有效扫描位置位于另一子块中时,在DC子块之外存在至少一个有效级别。因此,尽管该子块的coded_sub_block_flag被推断为等于1,但DC子块可能仅包含零/非有效级别。由于TS中没有最后扫描位置信息,会用信号通知每个子块的coded_sub_block_flag。这也包括DC子块的coded_sub_block_flag,除了在所有其他coded_sub_block_flag语法元素已经等于0的情况下。这样,DC coded_sub_block_flag被推断为等于1($\text{inferDcSbCbf}=1$)。由于在该DC子块中存在至少一个有效级别,因此没有用信号通知在(0,0)处的第一位置的sig_coeff_flag语法元素,并且与在该DC子块中的所有其他sig_coeff_flag语法元素等于0的情况下相反,得出为等于1($\text{inferSbDcSigCoeffFlag}=1$)。

[0041] 可以改变coded_sub_block_flag的上下文建模。上下文模型索引可以计算为右侧的coded_sub_block_flag和当前子块下方的coded_sub_block_flag的总和,而不是两者的逻辑析取。

[0042] 下面是对sig_coeff_flag上下文建模的描述:sig_coeff_flag上下文建模中的局部模板可以修改为仅包括当前扫描位置的右边邻居(NB_0)和下面邻居(NB_1)。上下文模型偏移量表示有效相邻位置的数量 $\text{sig_coeff_flag}[\text{NB}_0]+\text{sig_coeff_flag}[\text{NB}_1]$ 。因此,不同上下文的选择是根据当前变换块内的对角线d来设置的(d被去除)。这产生用于编码sig_coeff_flag标志的三个上下文模型和单个上下文模型集。

[0043] 下面是对abs_level_gt1_flag和par_level_flag上下文建模的描述:针对abs_level_gt1_flag和par_level_flag采用单个上下文模型。

[0044] 下面是对abs_remainder编码的描述:尽管变换跳过残差绝对水平的经验分布通常可以仍符合拉普拉斯分布或几何分布,但与变换系数绝对水平相比,可能存在较大的不平稳性。特别地,对于残差绝对水平,连续实现窗口内的方差较高。这激发了对abs_remainder语法二值化和上下文建模的以下修改:

[0045] 在二值化中使用较高的截止值,即从利用sig_coeff_flag、abs_level_gt1_flag、par_level_flag和abs_level_gt3_flag的编码到用于abs_remainder的莱斯码(Rice code)的转换点,以及针对每个二进制数(bin)位置的专用上下文模型产生较高的压缩效率。增加截止值将产生更多的“大于X”标志,例如引入abs_level_gt5_flag、abs_level_gt7_flag等,直到达到截止值。截止值本身固定为5($\text{numGtFlags}=5$)。

[0046] 可以修改莱斯参数(rice parameter)得出的模板,即仅考虑当前扫描位置左边的邻居和下面的邻居,类似于sig_coeff_flag上下文建模的局部模板。

[0047] 下面是对coeff_sign_flag上下文建模的描述:由于符号序列内部的不平稳性以及预测残差经常有偏差的事实,即使在全局经验分布为几乎均匀分布的情况下,也可以使用上下文模型对符号进行编码。单个专用上下文模型可以用于符号的编码,并且可以在sig_coeff_flag之后解析符号以将所有上下文编码的二进制数保持在一起。

[0048] 下面是对上下文编码的二进制数的限制的描述:每个变换单元(Transform Unit, TU)的上下文编码的二进制数的总数被限制为TU区域大小乘以2,例如对于 16×8 TU的上下文编码的二进制数的最大数量是 $16 \times 8 \times 2 = 256$ 。上下文编码二进制数的预算是在TU级别处消耗的,即不是每个CG的上下文编码二进制数的单独预算,而是当前TU内的所有CG共享一个上下文编码二进制数的预算。

[色度残差的联合编码]

[0049] VVC草案6支持其中联合编码色度残差的模式。联合色度编码模式的使用(激活)由TU级别标志`tu_joint_cbr_residual_flag`指示,并且所选模式由色度CBF隐式指示。如果TU的一个或两个色度CBF等于1,则标志`tu_joint_cbr_residual_flag`存在。在PPS和切片标头中,为联合色度残差编码模式用信号通知色度量化的参数(quantization parameter, QP)偏移值,以区别于为常规色度残差编码模式用信号通知的色度QP偏移值。这些色度QP偏移值用于为使用联合色度残差编码模式编码的这些块得出色度QP值。当相应的联合色度编码模式(下面表1中的模式2)在TU中处于活动时,色度QP偏移在该TU的量化和解码期间被添加至应用的亮度得出色度QP。对于其他模式(表1中的模式1和3),以与针对传统Cb或Cr块相同的方式得出色度QP。来自传输的变换块的色度残差(`resCb`和`resCr`)的重构过程在表1中进行了描绘。当该模式被激活时,一个单一的联合色度残差块(表1中的`resJointC[x][y]`)被用信号通知,考虑诸如`tu_cbf_cb`、`tu_cbf_cr`和`CSign`(在切片标头中指定的符号值)的信息,得出针对Cb的残差块(`resCb`)和针对Cr的残差块(`resCr`)。

[0050] 仅在帧内编码的编码单元(Coding Unit, CU)中支持上述三种联合色度编码模式。在帧间编码的CU中,仅支持模式2。因此,对于帧间编码的CU,语法元素`tu_joint_cbr_residual_flag`仅在两个色度cbf都等于1的情况下才存在。

表1:色度残差的重构

<code>tu_cbf_cb</code>	<code>tu_cbf_cr</code>	Cb 和 Cr 残差的重构	模式
1	0	$\text{resCb}[x][y] = \text{resJointC}[x][y]$ $\text{resCr}[x][y] = (\text{CSign} * \text{resJointC}[x][y]) \gg 1$	1
1	1	$\text{resCb}[x][y] = \text{resJointC}[x][y]$ $\text{resCr}[x][y] = \text{CSign} * \text{resJointC}[x][y]$	2
0	1	$\text{resCb}[x][y] = (\text{CSign} * \text{resJointC}[x][y]) \gg 1$ $\text{resCr}[x][y] = \text{resJointC}[x][y]$	3

[0051] 此处,值`CSign`是符号值(+1或-1),其在切片标头中指定,并且`resJointC[][]`是传输残差。

[0052] 现在参照图2,图2是根据实施方式的通信系统200的框图。通信系统200可以包括经由网络250互连的至少两个终端210和220。对于数据的单向传输,第一终端210可以在本地位置处对数据进行编码,以经由网络250传输至第二终端220。第二终端220可以从网络250接收第一终端210的编码后的数据,对编码后的数据进行解码并且显示解码后的数据。单向数据传输在媒体服务应用等中可以是常见的。

[0053] 图2还示出了第二对终端230和240,第二对终端230和240被提供用于支持例如在视频会议期间可能发生的编码后的数据的双向传输。对于数据的双向传输,每个终端230或240可以对在本地位置处捕获的数据进行编码,以经由网络250传输至其他终端。每个终端230或240还可以接收由其他终端传输的编码后的数据,可以对编码后的数据进行解码并且可以在本地显示装置处显示解码后的数据。

[0054] 在图2中,终端210至240可以被示为服务器、个人计算机和智能电话,但是实施方式的原理不限于此。实施方式适用于膝上型计算机、平板电脑、媒体播放器和/或专用视频会议装备。网络250表示在终端210至240之间传送编码后的数据的任何数目的网络,包括例如有线和/或无线通信网络。通信网络250可以在电路交换和/或分组交换通道中交换数据。代表性网络包括电信网络、局域网、广域网和/或因特网。出于本论述的目的,除非在下文中有所说明,否则网络250的架构和拓扑对于实施方式的操作来说可以是不重要的。

[0055] 图3是根据实施方式的G-PCC压缩器303和G-PCC解压缩器310在环境中的放置方式的图。所公开的主题可以同等地适用于其他支持的应用,包括例如:视频会议、数字电视、在包括CD、DVD、记忆棒等的数字介质上存储压缩数据等。

[0056] 流式传输系统300可以包括捕获子系统313,该捕获子系统313可以包括源301例如数码摄像装置,该源301创建例如未压缩的数据302。可以通过耦接至源301的G-PCC压缩器303来处理具有较高数据量的数据302。G-PCC压缩器303可以包括硬件、软件或其组合,以实现或实施如下更详细地描述的所公开的主题的各方面。可以将具有较低数据量的编码后的数据304存储在流式传输服务器305上,以备将来使用。一个或多个流式传输客户端306和308可以访问流式传输服务器305以检索编码后的数据304的副本307和309。客户端306可以包括G-PCC解压缩器310,该G-PCC解压缩器310对编码后的数据的传入副本307进行解码,并且创建可以在显示器312或其他呈现装置(未描绘)上呈现的传出(outgoing)数据311。在一些流式传输系统中,可以根据视频编码/压缩标准对编码后的数据304、307和309进行编码。这些标准的示例包括由MPEG针对G-PCC开发的那些标准。

[0057] 本公开内容的实施方式可以对多个颜色分量的变换系数联合应用二次变换。所提出的联合变换方案可以被称为联合分量二次变换(JCST)。图4中示出了所提出的对两个颜色分量使用JCST的编码器方案,其中JCST在前向变换之后以及量化之前执行。

[0058] 本公开内容的实施方式也可以在去量化变换之后以及反向(逆)变换之前执行JCST,如图5所示。

[0059] 图9示出了根据实施方式的解码方法的流程图。参照图9,LRRX在第一框901中,方法900包括获得编码后的视频比特流,该编码后的视频比特流包括编码后的颜色分量。

[0060] 在第二框902中,方法900包括对编码后的颜色分量进行熵解析。

[0061] 在第三框903中,方法900包括对颜色分量进行去量化并获得颜色分量的变换系数。

[0062] 在第四框904中,方法900包括对颜色分量的变换系数应用联合分量二次变换(JCST),从而生成JCST输出。

[0063] 根据实施方式,可以提供第五框905。在第五框905中,方法900可以包括对JCST输出执行反向变换,从而获得颜色分量的残差分量。

[0064] 根据实施方式,可以提供第六框906。在第六框906中,方法900可以包括基于颜色分量的残差分量对编码后的视频比特流进行解码。

[0065] 根据实施方式,该方法可以作为编码方法反向执行。实际上,虽然本文中的描述可能涉及特定的编码或解码方案,但这些描述不限于特定的编码或解码方案。也就是说,它们同样适于编码方案和解码方案两者。

[0066] 在一个实施方式中,JCST的输入可以是Cb和Cr变换系数。

- [0067] 在另一实施方式中,JCST的输入可以是Y、Cb和Cr变换系数。
- [0068] 在一个实施方式中,可以逐元素地执行JCST,使得对位于相同坐标处的每对Cb和Cr变换系数执行JCST。图6中示出了来自两个 4×2 块的成对Cb和Cr变换系数的示例。
- [0069] 在一个实施方式中,JCST可以是两点变换,并且输入可以是位于相同坐标处的一对Cb和Cr变换系数。
- [0070] 在一个实施方式中,JCST可以是两点变换,并且输出可以是替换一对Cb和Cr变换系数的一对二次变换系数。
- [0071] 在一个实施方式中,Cb和Cr变换系数的输出对可以位于与用作JCST的输入的Cb和Cr变换系数对相同的位置处。在图7中示出应用在两个 4×2 Cb和Cr块上的JCST的示例,并且构建另外两个 4×2 Cb和Cr块的JCST的输出将在编码器/解码器处进一步量化/去量化以及熵编码/解析。
- [0072] 在一个实施方式中,JCST的输出可以小于输入。例如,输入可以是一对Cb和Cr系数,而输出可以只是一个二次变换系数。
- [0073] 在一个实施方式中,在JCST中应用的变换可以包括但不限于Hadamard变换、离散余弦/正弦变换和数据驱动变换例如KLT或LGT(线图变换)。
- [0074] 在一个实施方式中,JCST的输入可以来自位于不同坐标处的多于一对(例如,三对)不同颜色分量。
- [0075] 在一个实施方式中,JCST可以是四点变换,并且输入可以是位于相同坐标处的两对Cb和Cr变换系数。示例示于图8中。
- [0076] 在一个实施方式中,输出可以是多于一对二次变换系数,所述多于一对二次变换系数替代多于一对Cb和Cr变换系数。
- [0077] 在一个实施方式中,JCST的输出可以小于输入。例如,输入可以是两对Cb和Cr系数,而输出可以只是两个二次变换系数。
- [0078] 在一个实施方式中,在JCST中应用的变换可以包括但不限于Hadamard变换、离散余弦/正弦变换或数据驱动变换例如KLT、LGT(线图变换)。
- [0079] 在一个实施方式中,可以针对有限范围的块大小应用JCST。
- [0080] 在一个示例中,可以针对小于或等于给定阈值的块大小应用JCST,其中,块大小可以指块宽度、块高度、块区域大小、块宽度和高度,以及块宽度和高度的最大值(或最小值)。
- [0081] 在一个示例中,可以针对大于或等于给定阈值的块大小应用JCST,其中,块大小可以指块宽度、块高度、块区域大小、块宽度和高度,以及块宽度和高度的最大值(或最小值)。
- [0082] 在一个实施方式中,可以通过变换块级别处的JCST标志用信号通知是否应用JCST。
- [0083] 在一个实施方式中,可以在变换系数之后用信号通知JCST标志。
- [0084] 在一个实施方式中,可以仅应用JCST的至少一个颜色分量具有至少一个非零系数时才用信号通知JCST标志。
- [0085] 在一个实施方式中,仅应用JCST的颜色分量中的每一个具有至少一个非零系数时才用信号通知JCST标志。
- [0086] 在一个实施方式中,仅应用JCST的颜色分量的非零系数的总数目大于或等于给定阈值例如0、1、2、3或4的情况下,才用信号通知JCST标志。

[0087] 在一个实施方式中,仅在应用JCST的颜色分量的最后非零系数位于沿扫描顺序大于给定阈值例如0、1、2、3或4的位置处的情况下,才用信号通知JCST标志。

[0088] 在一个实施方式中,通过CU级别(或CB级别)处的JCST标志用信号通知是否应用JCST。

[0089] 在一个实施方式中,是否可以使用JCST是使用高级句法用信号通知的,高级句法包括但不限于视频参数集(Video Parameter Set,VPS)、序列参数设置(Sequence Parameter Setting,SPS)、图片参数集(Picture Parameter Set,PPS)、切片标头、图片标头、Tile、Tile组或编码树单元(Coding Tree Unit,CTU)级别。

[0090] 在一个实施方式中,可以在获得包括在图片单元中的图片标头(PH)网络抽象层(NAL)单元和包括在所述图片单元中的至少一个视频编码层(VCL)NAL单元之后用信号通知JCST标志。

[0091] 在一个实施方式中,在应用JCST时,初级变换是给定的变换类型。

[0092] 在一个示例中,给定的变换类型是DCT-2。在另一示例中,给定的变换类型是DST-7。

[0093] 在一个实施方式中,在JCST中应用的变换的选择取决于编码信息,编码信息包括但不限于:初级变换类型、帧内预测模式,例如帧内预测方向/角度、帧间预测模式、是否应用块内复制(Intra Block Copy,IntraBC)、是否应用调色板模式、是否应用DPCM模式、运动矢量信息(方向,幅度)、是否应用子块运动以及是否应用扭曲运动(仿射运动)。

[0094] 在一个实施方式中,用于对指示是否应用JCST的标志进行熵编码的上下文取决于来自相邻块的相邻块信息,相邻块信息包括但不限于上面列出的信息。

[0095] 上述技术可以在适于压缩/解压缩的视频编码器和/或解码器中实现。编码器和/或解码器可以以硬件、软件或其任何组合来实现,并且软件(如果有的话)可以存储在一个或多个非暂态计算机可读介质中。例如,方法(或实施方式)、编码器和解码器中的每一个可以通过处理电路系统(例如,一个或多个处理器或一个或多个集成电路)来实现。在一个示例中,一个或多个处理器执行存储在非暂态计算机可读介质中的程序。

[0096] 可以使用计算机可读指令将上述技术实现为计算机软件,并且将其物理地存储在一个或多个计算机可读介质中。例如,图10示出了适于实现本公开内容的某些实施方式的计算机系统900。

[0097] 可以使用任何合适的机器代码或计算机语言对计算机软件进行编码,可以对任何合适的机器代码或计算机语言进行汇编、编译、链接等机制以创建包括指令的代码,这些指令可以由计算机中央处理单元(CPU)、图形处理单元(GPU)等直接执行或者通过解译、微代码执行等来执行。

[0098] 指令可以在各种类型的计算机或其部件——包括例如个人计算机、平板计算机、服务器、智能电话、游戏装置、物联网装置等上执行。

[0099] 图10中示出的用于计算机系统900的部件是示例性的,并且不旨在对实现本公开内容的实施方式的计算机软件的使用范围或功能范围提出任何限制。部件的配置也不应当被解释为具有与计算机系统900的非限制性实施方式中示出的部件中的任何一个部件或部件的组合有关的任何依赖性要求。

[0100] 计算机系统900可以包括某些人机接口输入装置。这样的人机接口输入装置可以

响应于由一个或多个个人用户通过例如触觉输入(例如:击键、滑动、数据手套移动)、音频输入(例如:语音、拍打)、视觉输入(例如:姿势)、嗅觉输入(未描绘)的输入。人机接口装置还可以用于捕获不一定与人类有意识的输入直接相关的某些媒体,例如,音频(诸如:语音、音乐、环境声音)、图像(诸如:扫描图像、从静态图像摄像装置获得的摄影图像)、视频(诸如:二维视频、包括立体视频的三维视频)。

[0101] 输入人机接口装置可以包括以下中的一个或多个(每者仅描绘了一个):键盘901、鼠标902、触控板903、触摸屏910、数据手套、操纵杆905、麦克风906、扫描仪907、摄像装置908。

[0102] 计算机系统900还可以包括某些人机接口输出装置。这样的人机接口输出装置可以通过例如触觉输出、声音、光和气味/味道来刺激一个或多个个人用户的感受。这样的人机接口输出装置可以包括:触觉输出装置(例如,通过触摸屏910、数据手套或操纵杆905的触觉反馈,但是也可以存在不用作输入装置的触觉反馈装置)。例如,这样的装置可以是音频输出装置(诸如:扬声器909、头戴式耳机(未描绘))、视觉输出装置(诸如屏幕910,包括阴极射线管(cathode ray tube,CRT)屏幕、液晶显示器(liquid crystal display,LCD)屏幕、等离子屏幕、有机发光二极管(organic light-emitting diode,OLED)屏幕,各自具有或不具有触摸屏输入能力,各自具有或不具有触觉反馈能力——其中的一些屏幕可以能够通过诸如立体图形输出的方式输出二维视觉输出或三维以上输出、虚拟现实眼镜(未描绘)、全息显示器和烟罐(未描绘)、以及打印机(未描绘)。

[0103] 计算机系统900还可以包括人类可访问的存储装置及其相关联的介质,例如包括具有CD/DVD等介质921的CD/DVD只读存储器(read only memory,ROM)/读写(read write,RW)920的光学介质、拇指驱动器922、可移动硬盘驱动器或固态驱动器923、例如磁带和软盘(未描绘)的遗留磁性介质、基于专用ROM/专用集成电路(application specific integrated circuit,ASIC)/可编程逻辑器件(programmable logic device,PLD)的装置例如安全加密狗(未描绘)等。

[0104] 本领域技术人员还应当理解,结合目前公开的主题使用的术语“计算机可读介质”不包含传输介质、载波或其他瞬态信号。

[0105] 计算机系统900还可以包括到一个或多个通信网络的接口。网络可以例如是无线的、有线的、光的。网络还可以是局域网、广域网、城域网、车载和工业网络、实时网络、延迟容忍网络等等。网络的示例包括诸如以太网的局域网、无线局域网(local area network,LAN)、蜂窝网络(包括全球移动通信系统(global system for mobile communications,GSM)、第三代(the third generation,3G)、第四代(the fourth, generation,4G)、第五代(the fifth generation,5G)、长期演进(long term evolution,LTE)等)、电视有线或无线广域数字网络(包括有线电视、卫星电视和地面广播电视)、车载的和工业的(包括控制器局域网(controller area network,CAN)总线)等。某些网络通常需要附接至某些通用数据端口或外围总线949(诸如,例如计算机系统900的通用串行总线(universal serial bus,USB)端口)的外部网络接口适配器;其他网络通常通过附接至如下描述的系统总线(例如,到PC计算机系统的以太网接口或到智能电话计算机系统的蜂窝网络接口)而集成到计算机系统900的核中。使用这些网络中的任何网络,计算机系统900可以与其他实体进行通信。这样的通信可以是单向仅接收的(例如,广播电视)、单向仅发送的

(例如,到某些CAN总线装置的CAN总线)、或双向的(例如使用局域或广域数字网络到其他计算机系统)。这样的通信可以包括到云计算环境955的通信。可以在如上所述的这些网络和网络接口中的每一个上使用某些协议和协议栈。

[0106] 以上提及的人机接口装置、人类可访问存储装置和网络接口954可以附接至计算机系统900的核940。

[0107] 核940可以包括一个或更多个中央处理单元(Central Processing Unit,CPU) 941、图形处理单元(Graphics Processing Unit,GPU) 942、现场可编程门阵列(Field Programmable Gate Area,FPGA) 943形式的专用可编程处理单元、用于某些任务的硬件加速器944等。这些装置,连同只读存储器(Read-only Memory,ROM) 945、随机存取存储器946、例如内部非用户可访问硬盘驱动器、固态驱动器(solid-状态drive,SSD)等的内部大容量存储装置947,可以通过系统总线948连接。在一些计算机系统中,可以以一个或更多个物理插头的形式访问系统总线948,以使得能够通过附加的CPU、GPU等进行扩展。外围装置可以直接地或通过外围总线949附连至核的系统总线948。外围总线的架构包括外围部件互连(peripheral component interconnect,PCI)、USB等。图形适配器950可以包括在核940中。

[0108] CPU 941、GPU 942、FPGA 943和加速器944可以执行某些指令,这些指令可以组合起来构成以上提及的计算机代码。该计算机代码可以存储在ROM 945或RAM 946中。瞬态数据也可以存储在RAM 946中,而永久数据可以存储在例如内部大容量存储装置947中。可以通过使用缓存存储器来实现对存储器装置中的任何存储装置的快速存储和检索,该缓存存储器可以与一个或更多个CPU 941、GPU 942、大容量存储装置947、ROM 945、RAM 946等紧密相关联。

[0109] 计算机可读介质可以在其上具有用于执行各种计算机实现的操作的计算机代码。介质和计算机代码可以是针对本公开内容的目的而专门设计和构建的介质和计算机代码,或者介质和计算机代码可以是计算机软件领域的技术人员公知且可获得的类型。

[0110] 作为示例而通过非限制的方式,具有架构的计算机系统900,并且特别是核940可以由处理器(包括CPU、GPU、FPGA、加速器等)执行实施在一个或更多个有形计算机可读介质中的软件而提供功能。这样的计算机可读介质可以是与如以上所介绍的以下项相关联的介质:用户可访问的大容量存储装置,以及核940的具有非暂态性的某些存储装置,例如,核内部大容量存储装置947或ROM 945。实现本公开内容的各种实施方式的软件可以存储在这样的装置中并且由核940执行。根据特定需要,计算机可读介质可以包括一个或更多个存储装置或芯片。软件可以使核940并且特别是其中的处理器(包括CPU、GPU、FPGA等)执行本文中描述的特定处理或特定处理的特定部分,包括限定存储在RAM 946中的数据结构以及根据由软件限定的处理来修改这些数据结构。另外地或作为替选,计算机系统可以由硬连线或以其他方式体现在电路(例如:加速器944)中的逻辑而提供功能,该逻辑可以代替软件或与软件一起操作以执行本文中描述的特定处理或特定处理的特定部分。在合适的情况下,所涉及的软件可以涵盖逻辑,反之,所涉及的逻辑也可以涵盖软件。在合适的情况下,所涉及的计算机可读介质可以涵盖存储用于执行的软件的电路(例如,集成电路(integrated circuit,IC))、实施用于执行的逻辑的电路、或上述两者。本公开内容包含硬件和软件的任何合适的组合。

[0111] 尽管本公开内容已经描述了若干非限制性实施方式,但是存在落入本公开内容的

范围内的改变、置换和各种替换等同物。因此将理解,虽然本文中并没有明确示出或描述,但是本领域技术人员能够设想实施本公开内容的原理并且因此在本公开内容的精神和范围内的许多系统和方法。

[0112] 缩写:

ALF:自适应环路滤波器,
APS:适应参数集,
AV1:AOMedia视频1,
AV2:AOMedia视频2,
CB:编码块,
CC-ALF:跨分量自适应环路滤波器,
CDEF:受约束方向增强滤波器,
CU:编码单元,
CTU:编码树单元,
DPCM:差分脉冲编码调制,
DPS:解码参数集,
HDR:高动态范围,
HEVC:高效视频编码,
ISP:帧内子分区,
JCCT:联合色度分量变换,
JVET:联合视频探索组,
LR:环路恢复滤波器,
PDPC:位置相关预测组合,
PPS:图片参数集,
PU:预测单元,
SDR:标准动态范围,
SPS:序列参数设置,
TSM:变换跳过模式,
TU:变换单元,
VVC:通用视频编码,
WAIP:广角帧内预测,以及
VPS:视频参数集。

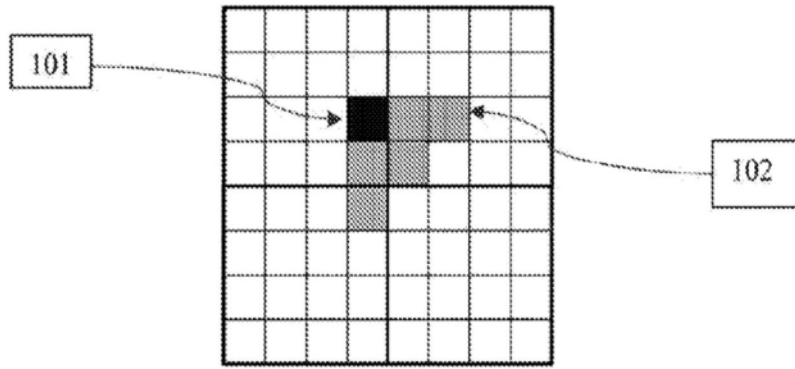


图1

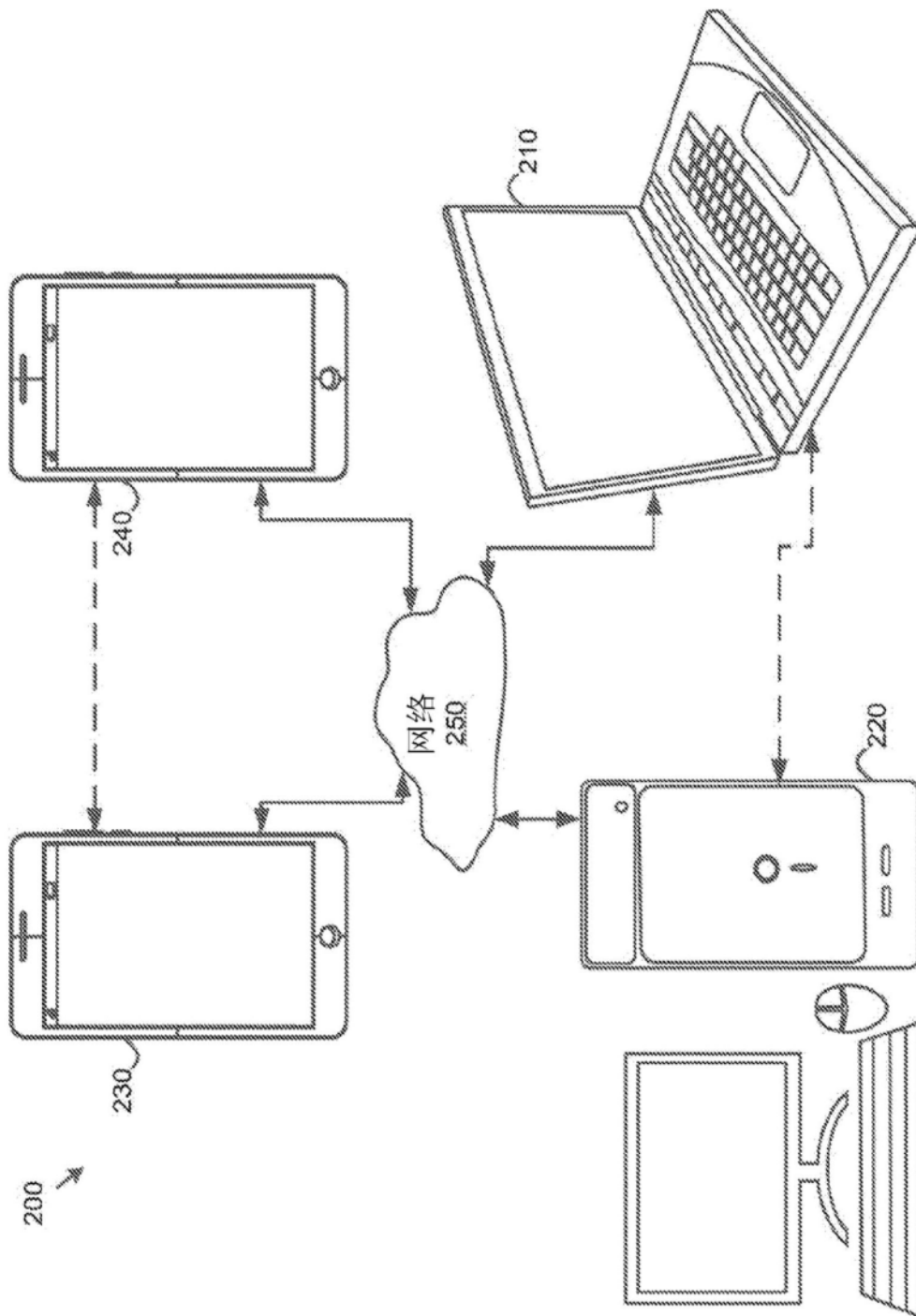


图2

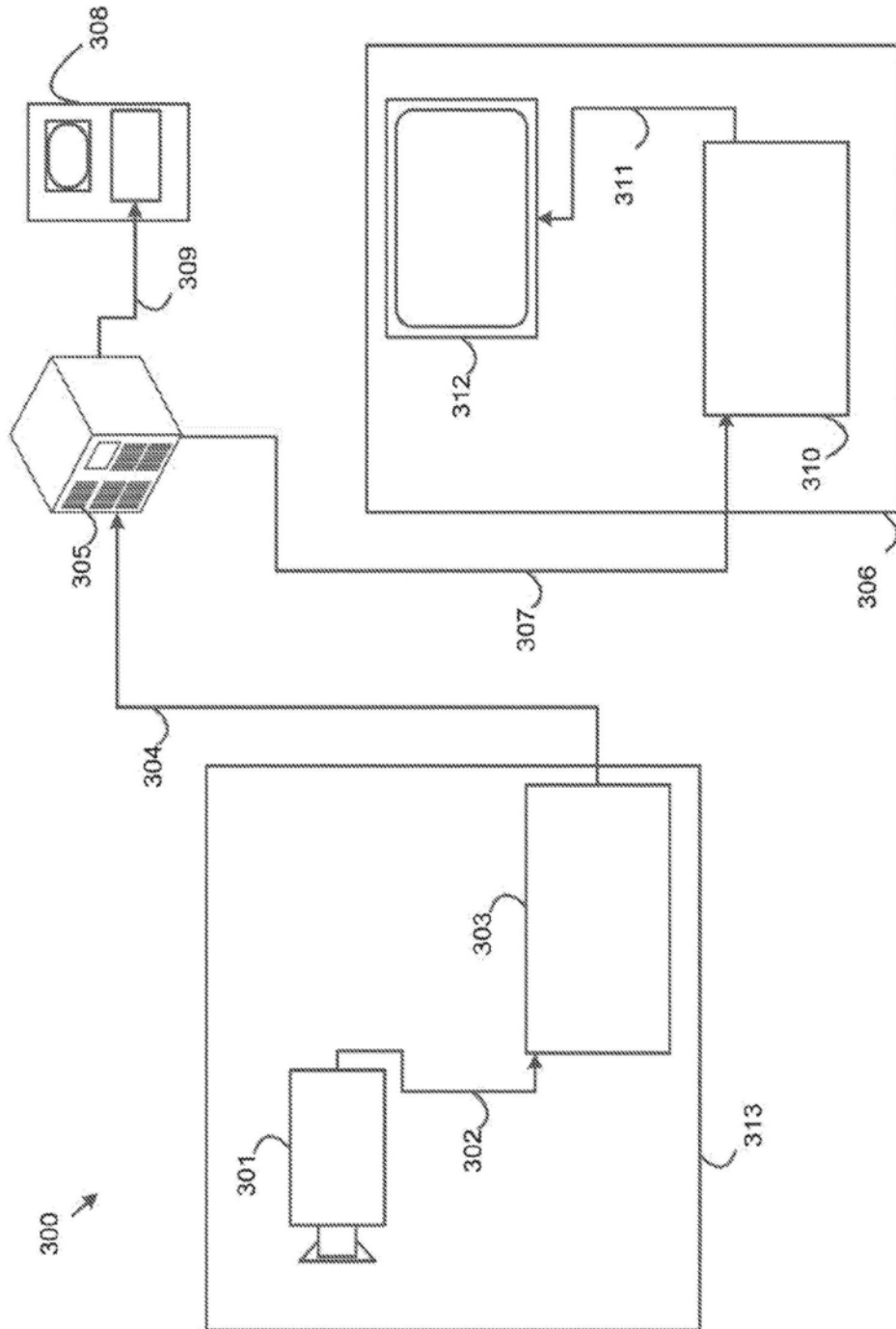


图3

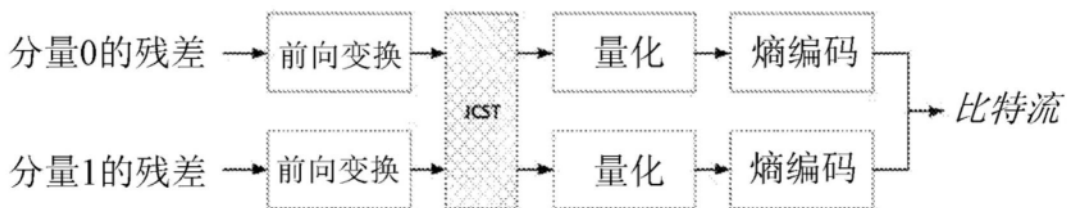


图4

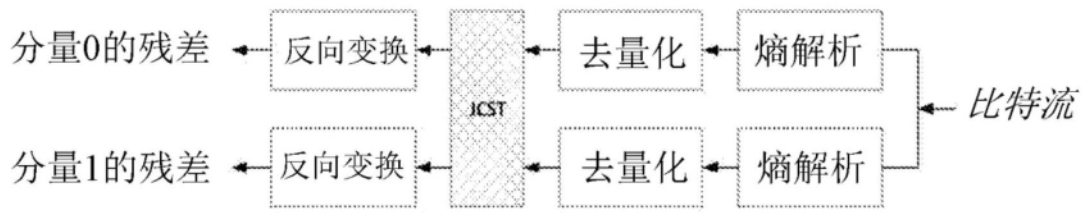


图5

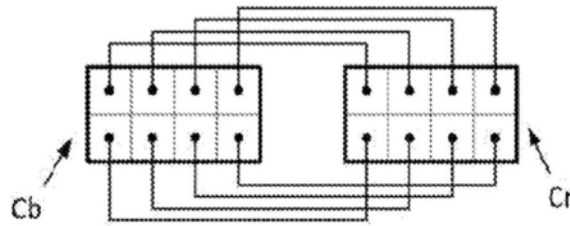


图6

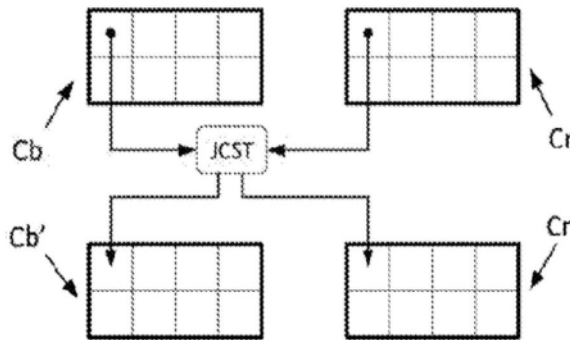


图7

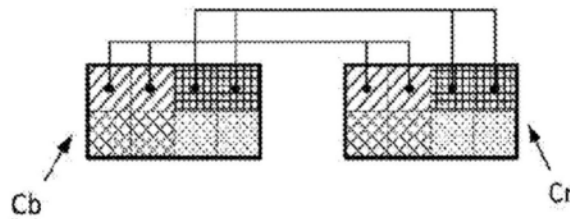


图8

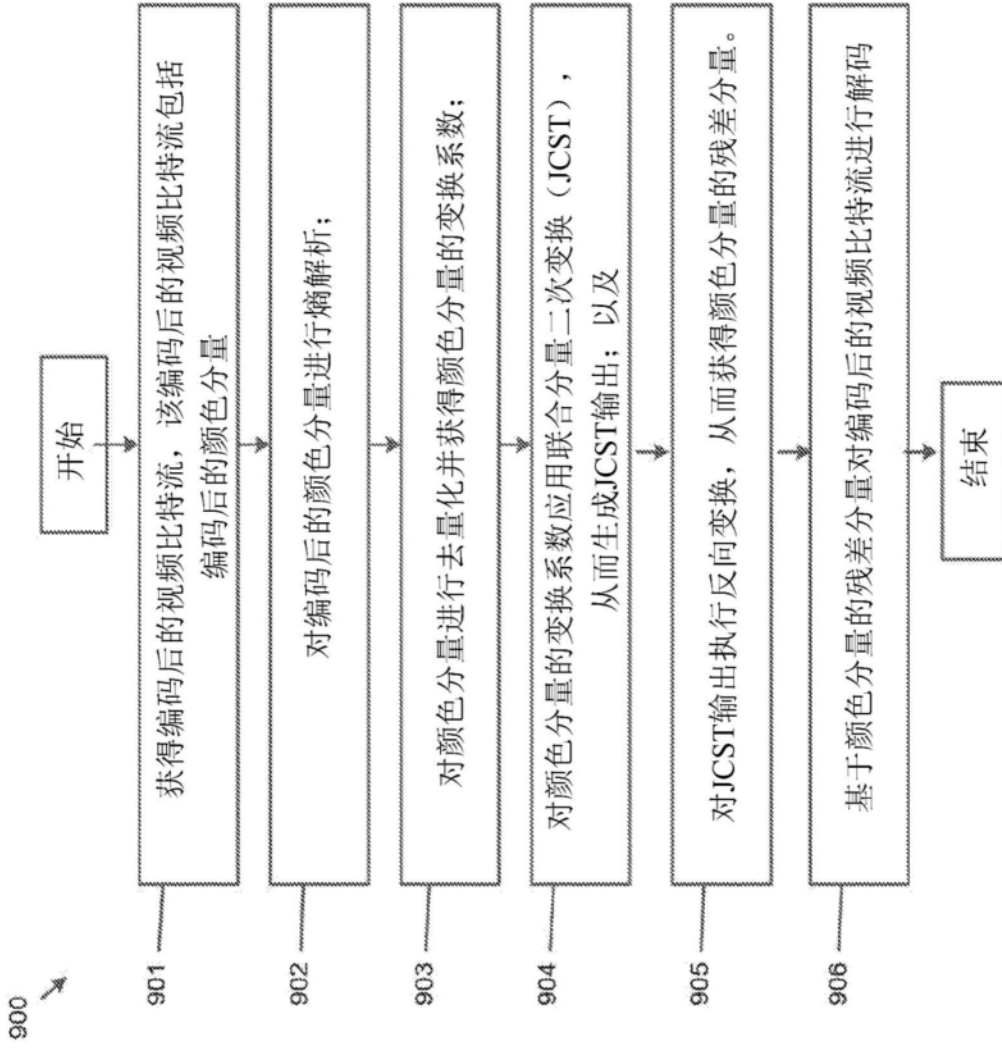


图9

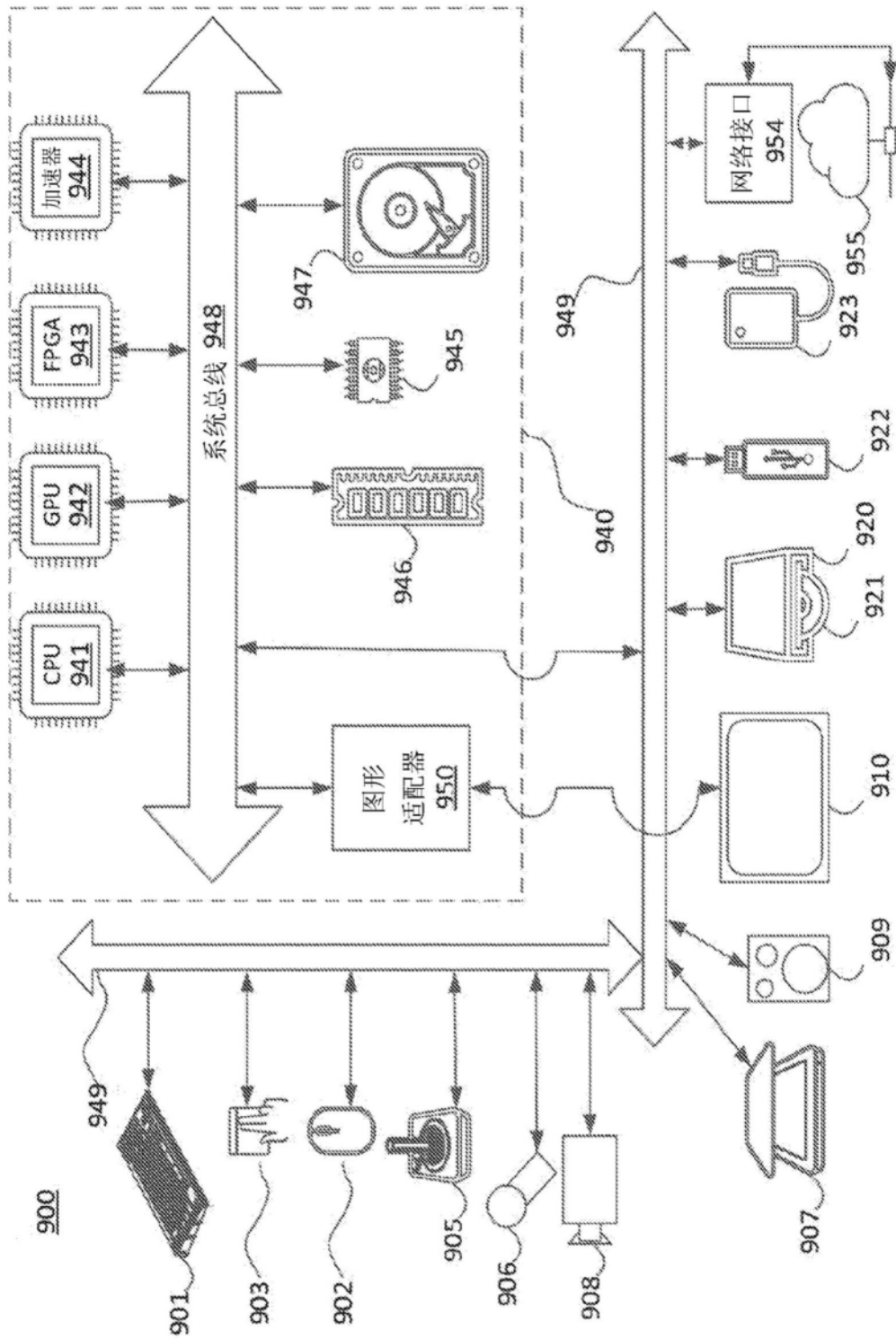


图10