



(12) 发明专利申请



(10) 申请公布号 CN 118632011 A

(43) 申请公布日 2024. 09. 10

(21) 申请号 202410754883.9

(22) 申请日 2020.12.16

(30) 优先权数据

62/956,254 2020.01.01 US

17/077,035 2020.10.22 US

(62) 分案原申请数据

202080028087.7 2020.12.16

(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 崔秉斗 史蒂芬·文格尔 刘杉

(74) 专利代理机构 北京派特恩知识产权代理有
限公司 11270

专利代理师 崔晓岚 浦彩华

(51) Int.Cl.

H04N 19/46 (2014.01)

H04N 19/136 (2014.01)

H04N 19/70 (2014.01)

H04N 19/174 (2014.01)

H04N 19/91 (2014.01)

H04N 21/854 (2011.01)

H04N 21/845 (2011.01)

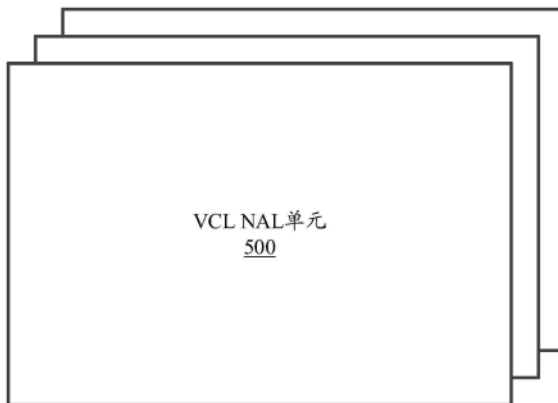
权利要求书3页 说明书18页 附图6页

(54) 发明名称

视频编解码的方法、装置、系统、介质及存储
码流的方法

(57) 摘要

本申请提供了一种视频编解码的方法、装置、系统、介质及存储码流的方法。方法包括：接收已编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和已编码图片的第二切片的第二VCL NAL单元，第一VCL NAL单元具有第一VCL NAL单元类型，第二VCL NAL单元具有不同于第一VCL NAL单元类型的第二VCL NAL单元类型；以及基于第一VCL NAL单元的第一VCL NAL单元类型和第二VCL NAL单元的第二VCL NAL单元类型，确定已编码图片的图片类型，包括：基于第一VCL NAL单元类型，指示第一VCL NAL单元包含逐渐解码刷新GDR图片已编码切片；以及基于第二VCL NAL单元类型，指示第二VCL NAL单元不包含瞬时解码刷新IDR图片已编码切片或干净随机存取CRA图片已编码切片，确定已编码图片是拖尾图片。



1. 一种视频解码的方法,其特征在于,用于对已编码视频流进行解码,所述方法包括:

接收已编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和所述已编码图片的第二切片的第二VCL NAL单元,所述第一VCL NAL单元具有第一VCL NAL单元类型,所述第二VCL NAL单元具有不同于所述第一VCL NAL单元类型的第二VCL NAL单元类型;以及

基于所述第一VCL NAL单元的所述第一VCL NAL单元类型和所述第二VCL NAL单元的所述第二VCL NAL单元类型,确定所述已编码图片的图片类型,包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐渐解码刷新GDR图片已编码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时解码刷新IDR图片已编码切片或干净随机存取CRA图片已编码切片,确定所述已编码图片是拖尾图片。

2. 根据权利要求1所述的方法,其特征在于,所述确定所述已编码图片的图片类型还包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含拖尾图片已编码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元包含瞬时解码刷新IDR图片已编码切片或干净随机存取CRA图片已编码切片,确定所述已编码图片是拖尾图片。

3. 根据权利要求1所述的方法,其特征在于,所述确定所述已编码图片的图片类型还包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含随机存取解码前置RADL图片已编码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元包含瞬时解码刷新IDR图片已编码切片或干净随机存取CRA图片已编码切片,确定所述已编码图片是RADL图片。

4. 根据权利要求1所述的方法,其特征在于,所述确定所述已编码图片的图片类型还包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐步式时间子层存取STSA图片已编码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时解码刷新IDR图片已编码切片,确定所述已编码图片是STSA图片。

5. 根据权利要求1所述的方法,其特征在于,所述确定所述已编码图片的图片类型还包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐步式时间子层存取STSA图片已编码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含干净随机存取CRA图片已编码切片,确定所述已编码图片是拖尾图片。

6. 根据权利要求1至5任一项所述的方法,其特征在于,所述方法还包括:

基于指示所述已编码图片包括混合VCL NAL单元类型的指示符,确定所述已编码图片的图片类型;

当所述已编码图片位于第一层中时,接收所述指示符;以及

基于所述指示符指示所述已编码图片包括混合VCL NAL单元类型,确定位于第二层中

的附加已编码图片包括混合VCL NAL单元类型,其中,所述第二层为所述第一层的参考层。

7.根据权利要求1至5中任一项所述的方法,其特征在于,

所述确定所述已编码图片的图片类型还包括:基于指示符指示所述已编码图片包括混合VCL NAL单元类型,确定所述已编码图片是拖尾图片。

8.根据权利要求1至5中任一项所述的方法,其特征在于,所述方法还包括:

基于指示符指示所述已编码图片包括混合VCL NAL单元类型,确定所述已编码图片的时间ID为0。

9.一种视频解码系统,其特征在于,包括:

存储器,配置成存储计算机程序代码;以及

至少一个处理器,配置成接收至少一个已编码视频流、访问所述计算机程序代码并根据所述计算机程序代码的指令操作以实现根据权利要求1至8中任一项所述的视频解码方法,以对所述已编码视频流进行解码。

10.一种视频编码方法,其特征在于,用于产生已编码视频流,所述方法包括:

接收待编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和所述待编码图片的第二切片的第二VCL NAL单元,所述第一VCL NAL单元具有第一VCL NAL单元类型,所述第二VCL NAL单元具有不同于所述第一VCL NAL单元类型的第二VCL NAL单元类型;以及

基于所述第一VCL NAL单元的所述第一VCL NAL单元类型和所述第二VCL NAL单元的所述第二VCL NAL单元类型,来设置所述待编码图片的图片类型,包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐渐编码刷新GDR图片的编码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时编码刷新IDR图片的编码切片或干净随机存取CRA图片的编码切片,确定所述待编码图片是拖尾图片。

11.一种计算机可读介质,其特征在于,所述计算机可读介质存储计算机指令,所述计算机指令在由至少一个处理器执行时,使得所述至少一个处理器实现根据权利要求1至8所述的视频解码方法,以对已编码流视频进行解码,或者权利要求10中所述的视频编码方法,以生成已编码视频流。

12.一种视频解码装置,其特征在于,用于对已编码视频流进行解码,所述装置包括:

接收模块,用于接收已编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和所述已编码图片的第二切片的第二VCL NAL单元,所述第一VCL NAL单元具有第一VCL NAL单元类型,所述第二VCL NAL单元具有不同于所述第一VCL NAL单元类型的第二VCL NAL单元类型;以及

解码模块,用于基于所述第一VCL NAL单元的所述第一VCL NAL单元类型和所述第二VCL NAL单元的所述第二VCL NAL单元类型,确定所述已编码图片的图片类型,包括:

基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐渐编码刷新GDR图片已解码切片;以及

基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时编码刷新IDR图片已解码切片或干净随机存取CRA图片已解码切片,确定所述已编码图片是拖尾图片。

13.一种视频编码装置,其特征在于,所述装置包括:

存储器,配置成存储计算机程序代码;以及

至少一个处理器,配置成接收视频数据、访问所述计算机程序代码并根据所述计算机程序代码的指令操作以实现根据权利要求10中所述的视频编码方法,以对所述视频数据进行编码。

14.一种存储码流的方法,其特征在于,在非暂时性计算机可读存储介质上存储视频码流,所述视频码流基于权利要求1至8任一项所述的视频解码方法进行解码,或者根据权利要求10所述的视频编码方法产生。

视频编解码的方法、装置、系统、介质及存储码流的方法

[0001] 本申请是申请日为2020年12月16日、中国专利申请号为202080028087.7、发明名称为“视频编解码的方法、装置、系统以及计算机可读介质”的专利申请的主动修改版本的分案申请。

技术领域

[0002] 本申请实施例的实施例涉及视频编码和解码,更具体地涉及用于使已编码图片支持混合网络抽象(NAL)单元类型。

背景技术

[0003] 在下一代视频编码(Versatile Video Coding,VVC)规范草案JVET-P2001(其全部内容并入本文中)(由JVET-Q0041编辑更新)中,支持混合网络抽象层(NAL)单元类型特征,该特征使得所具有的一个或多个切片NAL单元的NAL单元类型能够等于帧内随机存取点(IRAP)或干净随机存取(CRA),以及使得一个或多个切片NAL单元的NAL单元类型能够等于非IRAP。该特征可用于将两个不同的比特流合并成一个比特流,或者对每个局部区域(子图片)支持不同的随机存取周期。目前,定义如下语法和语义以支持该功能:

[0004] 在下面的表1中提供了示例性图片参数集原始字节序列载荷(RBSP)语法。

[0005] 表1

[0006]	pic_parameter_set_rbsp () {	描述符
	...	
	mixed_nalu_types_in_pic_flag	U (1)
	...	
	}	

[0007] 语法元素mixed_nalu_types_in_pic_flag等于1,则表示参考图片参数集(PPS)的每个图片具有多于一个视频编码层(VCL)NAL单元,且VCL NAL单元不具有相同的nal_unit_type值,且该图片不是IRAP图片。语法元素mixed_nalu_types_in_pic_flag等于0,则表示参考PPS的每个图片具有一个或多个VCL NAL单元,且参考PPS的每个图片的VCL NAL单元具有相同的nal_unit_type值。

[0008] 当语法元素no_mixed_nalu_types_in_pic_constraint_flag等于1时,语法元素mixed_nalu_types_in_pic_flag的值应等于0。

[0009] 根据当前VVC规范,NAL单元类型代码和NAL单元类型类别如下面的表2所示的定义。

[0010] 表2

[0011]

nal_unit_type	nal_unit_type的名称	NAL单元的内容和RBSP语法结构	NAL单元类型类别
0	TRAIL_NUT	拖尾图片的已编码切片 slice_layer_rbsp ()	VCL
1	STSA_NUT	STSA图片的已编码切片 slice_layer_rbsp ()	VCL
2	RASL_NUT	RASL图片的已编码切片 slice_layer_rbsp ()	VCL
3	RADL_NUT	RADL图片的已编码切片 slice_layer_rbsp ()	VCL
4...7	RSV_VCL_4... RSV_VCL_7	预留的非IRAP VCL NAL单元类型	VCL
8 9	IDR_W_RADL IDR_N_LP	IDR图片的已编码切片 slice_layer_rbsp ()	VCL
10	CRA_NUT	CRA图片的已编码切片 silce_layer_rbsp ()	VCL
11	GDR_NUT	GDR图片的已编码切片 slice_layer_rbsp ()	VCL
12 13	RSV_IRAP_VCL12 RSV_IRAP_VCL13	预留的IRAP VCL NAL单元类型	VCL
14...15	RSV_VCL14... RSV_VCL15	预留的非IRAP VCL NAL单元类型	VCL

[0012]

16	SPS_NUT	序列参数集 seq_parameter_set_rbsp ()	非VCL
17	PPS_NUT	图片参数集 pic_parameter_set_rbsp ()	非VCL
18	APS_NUT	自适应参数集 adaptation_parameter_set_rbsp ()	非VCL
19	AUD_NUT	存取单元分隔符 access_unit_delimiter_rbsp ()	非VCL
20	EOS_NUT	序列结束end_of_seq_rbsp ()	非VCL
21	EOB_NUT	比特流结束 end_of_bitstream_rbsp ()	非VCL
22, 23	PREFIX_SEI_NUT SUFFIX_SEI_NUT	补充增强信息sei_rbsp ()	非VCL
24	DPS_NUT	解码参数集 decoding_parameter_set_rbsp ()	非VCL
25...27	RSV_NVCL25... RSV_NVCL27	预留的非VCL NAL单元类型	非VCL
28...31	UNSPEC28... UNSPEC31	未指定的非VCL NAL单元类型	非VCL

[0013] 对于图片picA中的、其nal_unit_type值nalUnitTypeA在IDR_W_RADL至CRA_NUT (包括端点) 范围内的每个切片 (图片picA还包含具有另一个nal_unit_type值的一个或多个切片 (即, 图片picA的mixed_nalu_types_in_pic_flag的值等于1)), 应用如下:

[0014] (A) 切片应属于子图片subpicA, 其中, 对于子图片subpicA, 对应语法元素subpic_treated_as_pic_flag[i]的值等于1。

[0015] (B) 切片不应属于包含VCL NAL单元的picA的子图片, 其中切片的语法元素nal_unit_type不等于nalUnitTypeA。

[0016] (C) 对于编码层视频序列 (CLVS) 中按解码次序在后面的所有PU, subpicA中的切片的RefPicList[0]和RefPicList[1]均不应包括在有效条目中按解码次序位于picA之前的任何图片。

[0017] 对于任何特定图片的VCL NAL单元, 应用如下:

[0018] 如果语法元素mixed_nalu_types_in_pic_flag等于0, 则对于图片的所有已编码切片NAL单元, 语法元素nal_unit_type的值应该相同。图片或PU称为与图片或PU的已编码切片NAL单元具有相同的NAL单元类型。

[0019] 否则 (语法元素mixed_nalu_types_in_pic_flag等于1), 一个或多个VCL NAL单元应全部具有在IDR_W_RADL至CRA_NUT (包括端点) 范围内的特定nal_unit_type值, 而其他VCL NAL单元应全部具有在TRAIL_NUT至RSV_VCL_6 (包括端点) 范围内的特定nal_unit_type值。

type值,或者其nal_unit_type值等于GDR_NUT。

发明内容

[0020] 在上述背景技术部分中描述的混合VCL NAL单元类型的当前设计可能存在一些问题。

[0021] 在一些情况下,当图片由混合VCL NAL单元类型组成时,图片的图片类型可能是模棱两可的。

[0022] 在一些情况下,当NAL单元类型在同一个PU(图片)中混合时,时间标识符(例如,TemporalId)约束可能发生冲突。

[0023] 例如,当前VVC规范对TemporalId的约束如下:当语法元素nal_unit_type在IDR_W_RADL至RSV_IRAP_12(包括端点)的范围内时,语法元素TemporalId应等于0。当语法元素nal_unit_type等于STSA_NUT时,语法元素TemporalId不应等于0。

[0024] 在一些情况下,如果语法元素mixed_nalu_types_in_pic_flag在PPS中用信号表示,则至少两个PPS NAL单元应由CLVS中的切片NAL单元引用。此外,当提取子图片时,应通过改变语法元素mixed_nalu_types_in_pic_flag的值来重写相关联的PPS。

[0025] 在一些情况下,当前设计可能不支持随机存取解码前置(RADL)/随机存取跳过前置(RASL)NAL单元与图片(PU)中的拖尾图片共存。

[0026] 在一些情况下,当一层中的图片引用不同层中的另一图片时,语法元素mixed_nalu_types_in_pic_flag可能不对齐。

[0027] 本申请实施例的实施例可解决上文描述的一个或多个问题和/或其他问题。

[0028] 本申请实施例提供一种视频解码的方法,用于对已编码视频流进行解码,所述方法包括:

[0029] 接收已编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和所述已编码图片的第二切片的第二VCL NAL单元,所述第一VCL NAL单元具有第一VCL NAL单元类型,所述第二VCL NAL单元具有不同于所述第一VCL NAL单元类型的第二VCL NAL单元类型;以及

[0030] 基于所述第一VCL NAL单元的所述第一VCL NAL单元类型和所述第二VCL NAL单元的所述第二VCL NAL单元类型,确定所述已编码图片的图片类型,包括:

[0031] 基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐渐解码刷新GDR图片已编码切片;以及

[0032] 基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时解码刷新IDR图片已编码切片或干净随机存取CRA图片已编码切片,确定所述已编码图片是拖尾图片。

[0033] 在一些实施例中,本申请实施例还提供一种视频解码系统,包括:

[0034] 存储器,配置成存储计算机程序代码;以及

[0035] 至少一个处理器,配置成接收至少一个已编码视频流、访问所述计算机程序代码并根据所述计算机程序代码的指令操作以实现根据本申请实施例的视频解码方法,以对所述已编码视频流进行解码。

[0036] 在一些实施例中,本申请实施例还提供一种视频编码方法,用于产生已编码视频

流,所述方法包括:

[0037] 接收待编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和所述待编码图片的第二切片的第二VCL NAL单元,所述第一VCL NAL单元具有第一VCL NAL单元类型,所述第二VCL NAL单元具有不同于所述第一VCL NAL单元类型的第二VCL NAL单元类型;以及

[0038] 基于所述第一VCL NAL单元的所述第一VCL NAL单元类型和所述第二VCL NAL单元的所述第二VCL NAL单元类型,来设置所述待编码图片的图片类型,包括:

[0039] 基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐渐编码刷新GDR图片的编码切片;以及

[0040] 基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时编码刷新IDR图片的编码切片或干净随机存取CRA图片的编码切片,确定所述待编码图片是拖尾图片。

[0041] 在一些实施例中,本申请实施例还提供一种计算机可读介质,所述计算机可读介质存储计算机指令,所述计算机指令在由至少一个处理器执行时,使得所述至少一个处理器实现根据本申请实施例的视频解码方法,以对已编码视频流进行解码,或者本申请实施例的视频编码方法,以生成已编码视频流。

[0042] 在一些实施例中,本申请实施例还提供一种视频解码装置,用于对已编码视频流进行解码,所述装置包括:

[0043] 接收模块,用于接收已编码图片的第一切片的第一视频编码层VCL网络抽象层NAL单元和所述已编码图片的第二切片的第二VCL NAL单元,所述第一VCL NAL单元具有第一VCL NAL单元类型,所述第二VCL NAL单元具有不同于所述第一VCL NAL单元类型的第二VCL NAL单元类型;以及

[0044] 解码模块,用于基于所述第一VCL NAL单元的所述第一VCL NAL单元类型和所述第二VCL NAL单元的所述第二VCL NAL单元类型,确定所述已编码图片的图片类型,包括:

[0045] 基于所述第一VCL NAL单元类型,指示所述第一VCL NAL单元包含逐渐编码刷新GDR图片已解码切片;以及

[0046] 基于所述第二VCL NAL单元类型,指示所述第二VCL NAL单元不包含瞬时编码刷新IDR图片已解码切片或干净随机存取CRA图片已解码切片,确定所述已编码图片是拖尾图片。

[0047] 在一些实施例中,本申请实施例还提供一种视频编码装置,所述装置包括:

[0048] 存储器,配置成存储计算机程序代码;以及

[0049] 至少一个处理器,配置成接收视频数据、访问所述计算机程序代码并根据所述计算机程序代码的指令操作以实现根据本申请实施例的视频编码方法,以对所述视频数据进行编码。

[0050] 在一些实施例中,本申请实施例还提供一种存储码流的方法,在非暂时性计算机可读存储介质上存储视频码流,所述视频码流基于本申请实施例的视频解码方法进行解码,或者根据本申请实施例的视频编码方法产生。

附图说明

[0051] 通过以下详细描述和附图,所公开的主题的进一步特征、性质和各种优点将更加明显,在附图中:

[0052] 图1是根据一个实施例的通信系统的简化框图的示意图。

[0053] 图2是根据一个实施例的通信系统的简化框图的示意图。

[0054] 图3是根据一个实施例的解码器的简化框图的示意图。

[0055] 图4是根据一个实施例的编码器的简化框图的示意图。

[0056] 图5是根据一个实施例的NAL单元的框图。

[0057] 图6是根据一个实施例的解码器的框图。

[0058] 图7是适合于实现实施例的计算机系统的图。

具体实施方式

[0059] 图1示出了根据本申请实施例的实施例的通信系统(100)的简化框图。系统(100)可包括通过网络(150)互连的至少两个终端(110,120)。为了进行单向数据传输,第一终端(110)可在本地位置对视频数据进行编码以通过网络(150)传输到另一终端(120)。第二终端(120)可从网络(150)接收另一终端的已编码视频数据,对已编码数据进行解码,并显示所恢复的视频数据。单向数据传输在媒体服务等应用中可能是较常见的。

[0060] 图1示出了第二终端对(130,140),第二终端对(130,140)设置成支持例如可在视频会议期间发生的已编码视频的双向传输。对于双向数据传输,每个终端(130,140)可在本地位置对所采集的视频数据进行编码,以通过网络(150)传输到另一终端。每个终端(130,140)还可接收由另一终端传输的已编码视频数据,可对该已编码数据进行解码,且可在本地显示装置上显示所恢复的视频数据。

[0061] 在图1中,终端(110-140)可被示出为服务器、个人计算机和智能电话,和/或任何其他类型的终端。例如,终端(110-140)可以是膝上型计算机、平板电脑、媒体播放器和/或专用视频会议设备。网络(150)表示在终端(110-140)之间传送已编码视频数据的任何数量的网络,包括例如有线和/或无线通信网络。通信网络(150)可在电路交换和/或分组交换信道中交换数据。代表性的网络包括电信网络、局域网、广域网和/或互联网。出于本讨论的目的,除非在下文中有所解释,否则网络(150)的架构和拓扑对于本申请实施例的操作来说可能是无关紧要的。

[0062] 作为所公开的主题的应用示例,图2示出了视频编码器和视频解码器在流式传输环境中的放置方式。所公开的主题可同等地适用于其他支持视频的应用,包括例如视频会议、数字TV、在包括CD、DVD、存储棒等的数字介质上存储压缩视频等等。

[0063] 如图2所示,流式传输系统(200)可包括采集子系统(213),该采集子系统可包括视频源(201)和编码器(203)。视频源(201)可以是例如数码相机,并可配置成创建未压缩的视频样本流(202)。相较于已编码视频比特流,未压缩的视频样本流(202)可提供高数据量,并可由耦接到相机(201)的编码器(203)处理。编码器(203)可包括硬件、软件或软硬件组合以实现或实施如下文更详细地描述的所公开的主题的各方面。相较于样本流,已编码视频比特流(204)可提供较低数据量,并可存储在流式传输服务器(205)上以供将来使用。一个或多个流式传输客户端(206)可访问流式传输服务器(205)以检索视频比特流(209),视频比

特流(209)可以是已编码视频比特流(204)的副本。

[0064] 在实施例中,流式传输服务器(205)还可用作媒体感知网络单元(MANE)。例如,流式传输服务器(205)可配置成给一个或多个流式传输客户端(206)修剪已编码视频比特流(204),以裁剪可能不同的比特流。在实施例中,MANE可与流式传输系统(200)中的流式传输服务器(205)分开提供。

[0065] 流式传输客户端(206)可包括视频解码器(210)和显示器(212)。例如,视频解码器(210)可对作为已编码视频比特流(204)的传入副本的视频比特流(209)进行解码,并创建可在显示器(212)或另一呈现设备(未描绘)上呈现的输出视频样本流(211)。在一些流式传输系统中,可根据某些视频编码/压缩标准对视频比特流(204,209)进行编码。这些标准的示例包括但不限于ITU-T建议书H.265。正在开发的视频编码标准非正式地称为下一代视频编码(VVC),所公开的主题可用于VVC的上下文中。

[0066] 图3示出了根据本申请实施例的一个实施例的附接到显示器(212)的视频解码器(210)的示例性功能框图。

[0067] 视频解码器(210)可包括信道(312)、接收器(310)、缓冲存储器(315)、熵解码器/解析器(320)、缩放器/逆变换单元(351)、帧内预测单元(352)、运动补偿预测单元(353)、聚合器(355)、环路滤波器单元(356)、参考图片存储器(357)和当前图片存储器(358)。在至少一个实施例中,视频解码器(210)可包括集成电路、一系列集成电路和/或其他电子电路。视频解码器(210)还可部分地或完全地在软件中实现,该软件在具有相关联存储器的一个或多个CPU上运行。

[0068] 在本实施例和其他实施例中,接收器(310)可接收将由解码器(210)解码的一个或多个已编码视频序列,可一次接收一个已编码视频序列,其中每个已编码视频序列的解码独立于其他已编码视频序列的解码。可从信道(312)接收已编码视频序列,该信道可以是通向存储已编码视频数据的存储装置的硬件/软件链路。接收器(310)可接收可转发到它们各自的使用实体(未描绘)的已编码视频数据以及其他数据,例如,已编码音频数据和/或辅助数据流。接收器(310)可将已编码视频序列与其他数据分开。为了防止网络抖动,缓冲存储器(315)可耦接在接收器(310)与熵解码器/解析器(320)(此后称为“解析器”)之间。当接收器(310)从具有足够带宽和可控性的存储/转发装置或从等时同步网络接收数据时,可以不使用缓冲器(315),或可以将该缓冲器做得较小。为了在互联网等业务分组网络上使用,可能需要缓冲器(315),该缓冲器可相对较大且可有利地具有自适应性大小。

[0069] 视频解码器(210)可包括解析器(320)以根据已熵编码的视频序列重构符号(321)。例如,这些符号的类别包括用于管理解码器(210)的操作的信息,以及用于控制显示装置例如显示器(212)的潜在信息,该显示装置可耦接到解码器,如图2中所示。例如,用于显示装置的控制信息可以是辅助增强信息(Supplemental Enhancement Information, SEI)消息或视频可用性信息(Video Usability Information, VUI)的参数集片段(未描绘)形式。解析器(320)可对接收到的已编码视频序列进行解析/熵解码。已编码视频序列的编码可根据视频编码技术或标准进行,且可遵循本领域技术人员公知的各种原理,包括可变长度编码、霍夫曼编码(Huffman coding)、具有或不具有上下文灵敏度的算术编码等等。解析器(320)可基于对应于群组的至少一个参数,从已编码视频序列提取用于视频解码器中的像素的子群中的至少一个子群的子群参数集。子群可包括图片群组(Group of

Pictures, GOP)、图片、图块、切片、宏块、编码单元 (Coding Unit, CU)、块、变换单元 (Transform Unit, TU)、预测单元 (Prediction Unit, PU) 等等。解析器 (320) 还可从已编码视频序列提取信息, 例如变换系数、量化器参数值、运动矢量等等。

[0070] 解析器 (320) 可对从缓冲器 (315) 接收的视频序列执行熵解码/解析操作, 从而创建符号 (321)。

[0071] 取决于已编码视频图片或一部分已编码视频图片 (例如: 帧间图片和帧内图片、帧间块和帧内块) 的类型以及其他因素, 符号 (321) 的重构可涉及多个不同单元。涉及哪些单元以及涉及方式可由解析器 (320) 从已编码视频序列解析的子群控制信息来控制。为了简洁起见, 未描述解析器 (320) 与下文的多个单元之间的此类子群控制信息流。

[0072] 除已经提及的功能块以外, 解码器 210 可在概念上细分成如下文所描述的多个功能单元。在商业约束下运行的实际实现方式中, 这些单元中的许多单元彼此紧密交互且可至少部分地彼此集成。然而, 出于描述所公开的主题的目的, 概念上细分成下文的功能单元是适当的。

[0073] 一个单元可以是缩放器/逆变换单元 (351)。缩放器/逆变换单元 (351) 可从解析器 (320) 接收作为符号 (321) 的量化变换系数以及控制信息, 包括使用哪种变换方式、块大小、量化因子、量化缩放矩阵等。缩放器/逆变换单元 (351) 可输出包括样本值的块, 该样本值可输入到聚合器 (355) 中。

[0074] 在一些情况下, 缩放器/逆变换单元 (351) 的输出样本可属于帧内编码块; 即: 不使用来自先前重构的图片的预测性信息, 但可使用来自当前图片的先前重构部分的预测性信息的块。此类预测性信息可由帧内图片预测单元 (352) 提供。在一些情况下, 帧内图片预测单元 (352) 采用从当前图片存储器 (358) 的当前 (部分重构的) 图片提取的周围已重构信息生成大小和形状与正在重构的块相同的块。在一些情况下, 聚合器 (355) 基于每个样本, 将帧内预测单元 (352) 生成的预测信息添加到由缩放器/逆变换单元 (351) 提供的输出样本信息中。

[0075] 在其他情况下, 缩放器/逆变换单元 (351) 的输出样本可属于帧间编码和潜在运动补偿块。在此情况下, 运动补偿预测单元 (353) 可访问参考图片存储器 (357) 以提取用于预测的样本。在根据属于块的符号 (321) 对提取的样本进行运动补偿之后, 这些样本可由聚合器 (355) 添加到缩放器/逆变换单元 (351) 的输出 (在这种情况下称为残差样本或残差信号), 从而生成输出样本信息。运动补偿预测单元 (353) 从参考图片存储器 (357) 内的地址获取预测样本可受到运动矢量控制。该运动矢量以符号 (321) 的形式而供运动补偿预测单元 (353) 使用, 符号 (321) 可具有例如 X、Y 和参考图片分量。运动补偿还可包括在使用子样本精确运动矢量时, 从参考图片存储器 (357) 提取的样本值的内插、运动矢量预测机制等等。

[0076] 聚合器 (355) 的输出样本可经受环路滤波器单元 (356) 中的各种环路滤波技术。视频压缩技术可包括环路内滤波器技术, 该环路内滤波器技术受控于包括在已编码视频比特流中并且作为来自解析器 (320) 的符号 (321) 可用于环路滤波器单元 (356) 的参数, 然而, 视频压缩技术还可响应于在解码已编码图片或已编码视频序列的先前 (按解码次序) 部分期间获得的元信息, 以及响应于先前重构且经过环路滤波的样本值。

[0077] 环路滤波器单元 (356) 的输出可以是样本流, 该样本流可输出到显示装置例如显示器 (212) 以及存储在参考图片存储器 (357) 中, 以用于后续的帧间图片预测。

[0078] 一旦完全重构,某些已编码图片就可用作参考图片以用于将来预测。一旦已编码图片被完全重构,且已编码图片(通过例如解析器(320))被识别为参考图片,则当前参考图片可变为参考图片存储器(357)的一部分,且可在开始重构后续已编码图片之前重新分配新的当前图片存储器。

[0079] 视频解码器(210)可根据例如ITU-T建议书H.265标准中记录的预定视频压缩技术执行解码操作。在已编码视频序列遵循视频压缩技术或标准的语法(如视频压缩技术或标准中指定的,特别是如视频压缩技术或标准中的配置文件指定的)的意义上,已编码视频序列可符合所使用的视频压缩技术或标准指定的语法。此外,对于一些视频压缩技术或标准下的合规性,还要求已编码视频序列的复杂度处于视频压缩技术或标准的层级所限定的范围内。在一些情况下,层级限制最大图片大小、最大帧率、最大重构取样率(以例如每秒兆(mega)个样本为单位进行测量)、最大参考图片大小等。在一些情况下,由层级设定的限制可通过假想参考解码器(Hypothetical Reference Decoder,HRD)规范和在已编码视频序列中用信号表示的HRD缓冲器管理的元数据来进一步限定。

[0080] 在一个实施例中,接收器(310)可连同已编码视频一起接收附加(冗余)数据。该附加数据可以被包括作为已编码视频序列的一部分。该附加数据可由视频解码器(210)使用以对数据进行适当解码和/或较准确地重构原始视频数据。附加数据可以是例如时间、空间或信噪比(signal noise ratio,SNR)增强层、冗余切片、冗余图片、前向纠错码等形式。

[0081] 图4示出了根据本申请实施例的一个实施例的与视频源(201)相关联的视频编码器(203)的示例性功能框图。

[0082] 视频编码器(203)可包括例如作为源编码器(430)的编码器、编码引擎(432)、(本地)解码器(433)、参考图片存储器(434)、预测器(435)、发送器(440)、熵编码器(445)、控制器(450)和信道(460)。

[0083] 编码器(203)可从视频源(201)(并非编码器的一部分)接收视频样本,该视频源可采集将由编码器(203)编码的视频图像。

[0084] 视频源(201)可提供将由编码器(203)编码的呈数字视频样本流形式的源视频序列,该数字视频样本流可具有任何合适位深度(例如:8位、10位、12位……)、任何色彩空间(例如BT.601Y CrCb,RGB……)和任何合适采样结构(例如Y CrCb 4:2:0、Y CrCb 4:4:4)。在媒体服务系统中,视频源(201)可以是存储先前已准备的视频的存储装置。在视频会议系统中,视频源(203)可以是采集本地图像信息作为视频序列的相机。可将视频数据提供为多个单独的图片,当按顺序观看时,这些图片被赋予运动。图片自身可构造为空间像素阵列,其中取决于所用的取样结构、色彩空间等,每个像素可包括一个或多个样本。本领域技术人员可以很容易理解像素与样本之间的关系。下文侧重于描述样本。

[0085] 根据一个实施例,编码器(203)可实时或在由应用所要求的任何其他时间约束下,将源视频序列的图片编码且压缩成已编码视频序列(443)。施行适当的编码速度是控制器(450)的一个功能。控制器(450)还可控制如下文所描述的其他功能单元且可在功能上耦接到这些单元。为了简洁起见,图中未标示耦接。由控制器(450)设置的参数可包括速率控制相关参数(图片跳过、量化器、率失真优化技术的 λ 值……)、图片大小、图片群组(GOP)布局、最大运动矢量搜索范围等。本领域技术人员可容易地识别控制器(450)的其他功能,这些功能可涉及针对某一系统设计优化的视频编码器(203)。

[0086] 一些视频编码器在本领域技术人员容易识别的“编码环路”中进行操作。作为过度简化的描述,编码环路可包括源编码器(430)(负责基于待编码的输入图片和参考图片创建符号)的编码部分和嵌入于编码器(203)中的(本地)解码器(433),解码器(433)重构符号以用类似于(远程)解码器还可创建样本数据的方式创建样本数据(因为在某些视频压缩技术中,符号与已编码视频比特流之间的压缩是无损的)。重构的样本流可输入到参考图片存储器(434)。由于符号流的解码产生与解码器位置(本地或远程)无关的位精确结果,因此参考图片存储器中的内容在本地编码器与远程编码器之间也是按比特位精确对应的。换句话说,编码器的预测部分“看到”的参考图片样本与解码器将在解码期间使用预测时所“看到”的样本值完全相同。这种参考图片同步性的基本原理(以及在例如因信道误差而无法维持同步性的情况下产生的漂移)也是本领域技术人员已知的。

[0087] “本地”解码器(433)的操作可与例如已在上文结合图3详细描述“远程”解码器(210)相同。然而,当符号可用且熵编码器(445)和解析器(320)能够无损地将符号编码/解码为已编码视频序列时,包括信道(312)、接收器(310)、缓冲器(315)和解析器(320)在内的解码器(210)的熵解码部分,可能无法完全在本地解码器(433)中实施。

[0088] 此时可以观察到,除存在于解码器中的解析和/或熵解码之外的任何解码器技术,也必定以基本上相同的功能形式存在于对应的编码器中。出于此原因,所公开的主题侧重于解码器操作。可简化编码器技术的描述,因为编码器技术与全面地描述的解码器技术互逆。仅在某些区域中需要更详细的描述,并在下文提供。

[0089] 在一部分操作期间,源编码器(430)可执行运动补偿预测编码,参考来自视频序列中被指定为“参考帧”的一个或多个先前已编码帧,该运动补偿预测编码对输入帧进行预测性编码。以此方式,编码引擎(432)对输入帧的像素块与参考帧的像素块之间的差异进行编码,该参考帧可被选作该输入帧的预测参考。

[0090] 本地视频解码器(433)可基于源编码器(430)创建的符号,对可指定为参考帧的帧的已编码视频数据进行解码。编码引擎(432)的操作可有利地为有损过程。当已编码视频数据可在视频解码器(图4中未示出)处被解码时,重构的视频序列通常可以是带有一些误差的源视频序列的副本。本地视频解码器(433)复制解码过程,该解码过程可由视频解码器对参考帧执行,且可使重构的参考帧存储在参考图片存储器(434)中。以此方式,编码器(203)可在本地存储重构的参考帧的副本,该副本与将由远端视频解码器获得的重构的参考帧具有共同内容(不存在传输误差)。

[0091] 预测器(435)可针对编码引擎(432)执行预测搜索。即,对于将要编码的新帧,预测器(435)可在参考图片存储器(434)中搜索可作为该新图片的适当预测参考的样本数据(作为候选参考像素块)或某些元数据,例如参考图片运动矢量、块形状等。预测器(435)可基于样本块逐像素块操作,以找到合适的预测参考。在一些情况下,如由预测器(435)获得的搜索结果所确定的那样,输入图片可具有从参考图片存储器(434)中存储的多个参考图片取得的预测参考。

[0092] 控制器(450)可管理视频编码器(430)的编码操作,包括例如设置用于对视频数据进行编码的参数和子群参数。

[0093] 可在熵编码器(445)中对所有上述功能单元的输出进行熵编码。熵编码器根据本领域技术人员已知的技术例如霍夫曼编码、可变长度编码、算术编码等对各种功能单元生

成的符号进行无损压缩,从而将该符号转换成已编码视频序列。

[0094] 传输器(440)可缓冲由熵编码器(445)创建的已编码视频序列,从而为通过通信信道(460)进行传输做准备,该通信信道可以是通向将存储已编码视频数据的存储装置的硬件/软件链路。传输器(440)可将来自视频编码器(430)的已编码视频数据与要传输的其他数据合并,该其他数据例如是已编码音频数据和/或辅助数据流(未示出来源)。

[0095] 控制器(450)可管理编码器(203)的操作。在编码期间,控制器(450)可以为每个已编码图片分配某一已编码图片类型,但这可能影响可应用于相应的图片的编码技术。例如,通常可将图片分配为帧内图片(I图片)、预测性图片(P图片)或双向预测性图片(B图片)。

[0096] 帧内图片(I图片)可以是不将序列中的任何其他帧用作预测源就可被编码和解码的图片。一些视频编解码器容许不同类型的帧内图片,包括例如瞬时解码刷新(IDR)图片。本领域技术人员了解I图片的变体及其相应的应用和特征。

[0097] 预测性图片(P图片)可以是可使用帧内预测或帧间预测进行编码和解码的图片,该帧内预测或帧间预测使用至多一个运动矢量和参考索引来预测每个块的样本值。

[0098] 双向预测性图片(B图片)可以是可使用帧内预测或帧间预测进行编码和解码的图片,该帧内预测或帧间预测使用至多两个运动矢量和参考索引来预测每个块的样本值。类似地,多个预测性图片可使用多于两个参考图片和相关联元数据以用于重构单个块。

[0099] 源图片通常可在空间上细分成多个样本块(例如, 4×4 、 8×8 、 4×8 或 16×16 个样本的块),且逐块进行编码。这些块可参考其他(已编码)块进行预测性编码,该其他块由应用于块的相应图片的编码分配来确定。举例来说,I图片的块可进行非预测性编码,或该块可参考同一图片的已编码块来进行预测性编码(空间预测或帧内预测)。P图片的像素块可参考一个先前编码的参考图片通过空间预测或通过时间预测进行非预测性编码。B图片的块可参考一个或两个先前编码的参考图片通过空间预测或通过时间预测进行非预测性编码。

[0100] 视频编码器(203)可根据例如ITU-T建议书H.265的预定视频编码技术或标准执行编码操作。在操作中,视频编码器(203)可执行各种压缩操作,包括利用输入视频序列中的时间和空间冗余的预测编码操作。因此,已编码视频数据可符合所使用的视频编码技术或标准指定的语法。

[0101] 在实施例中,传输器(440)可在传输已编码视频时传输附加数据。视频编码器(430)可包括此类数据作为已编码视频序列的一部分。附加数据可包括时间/空间/SNR增强层、冗余图片和切片等其他形式的冗余数据、辅助增强信息(SEI)消息、视频可用性信息(VUI)参数集片段等。

[0102] 本申请实施例的实施例可修改当前VVC规范,且可实现如上表2中定义的NAL单元类型代码和NAL单元类型类别。

[0103] “帧内随机存取点图片”(或“IRAP图片”)可以是如下图片,该图片在其解码过程中不参考除了自身之外的用于帧间预测的任何图片,且可以是干净随机存取(CRA)图片或瞬时解码刷新(IDR)图片。比特流中按解码次序第一的图片可以是IRAP或逐渐解码刷新(GDR)图片。如果在需要参考必要的参数集时必要的参数集可用,则可正确地解码IRAP图片和已编码视频序列(CVS)中按解码次序所有后续的非RASL图片,而无需执行按解码次序在IRAP图片之前的任何图片的解码处理。

[0104] “拖尾图片”可以是非IRAP图片,其按输出次序在相关联的IRAP图片之后,且不是逐步式时间子层存取(STSA)图片。

[0105] “逐步式时间子层存取图片”(或“STSA图片”)可以是如下图片,该图片不使用与用于帧间预测参考的STSA图片具有相同的TemporalId的图片。按解码次序在STSA图片之后、与STSA图片具有相同的TemporalId的图片可以不使用按解码次序在STSA图片之前、与用于帧间预测参考的STSA图片具有相同的TemporalId的图片。STSA图片可以在STSA图片处允许从紧邻的下子层向上切换到包含STSA图片的子层。STSA图片可具有大于0的TemporalId。

[0106] “随机存取跳过前置图片”(或“RASL图片”)可以是相关联的CRA图片即前置图片的图片。当相关联的CRA图片的NoIncorrectPicOutputFlag等于1时,RASL图片可能不输出且不能正确解码,因为RASL图片可包含对比特流中不存在的图片的参考。RASL图片可以不用作对RASL图片进行解码处理的参考图片。当存在时,所有RASL图片可按解码次序在同一个相关联的CRA图片的所有拖尾图片之前。

[0107] “随机存取解码前置图片”(或“RADL图片”)可以是前置图片,该前置图片不用作对同一个相关联的IRAP图片的拖尾图片进行解码处理的参考图片。当存在时,所有RADL图片可按解码次序在同一个相关联的IRAP图片的所有拖尾图片之前。

[0108] “瞬时解码刷新图片”(或“IDR图片”)可以是如下图片,该图片不具有比特流中存在的前置图片(例如,nal单元类型等于IDR_N_LP),或者不具有比特流中存在的前置图片,但是可具有比特流中相关联的RADL图片(例如,nal单元类型等于IDR_W_RADL)。

[0109] “干净随机存取图片”(或“CRA图片”)可以是如下图片,该图片在其解码过程中不参考除了自身之外的用于帧间预测的任何图片,且可以是比特流中按解码次序第一的图片,或者可稍后出现在比特流中。CRA图片可具有相关联的RADL或RASL图片。当CRA图片的NoIncorrectPicOutputFlag等于1时,相关联的RASL图片可不由解码器输出,因为RASL图片不可解码,原因是RASL图片可包含对比特流中不存在的图片的参考。

[0110] 根据一个或多个实施例,当由已编码图片引用的PPS的语法元素mixed_nalu_types_in_pic_flag等于1时,已编码图片的图片类型(例如由解码器)如下确定:

[0111] (A) 当图片的NAL单元的nal_unit_type等于TRAIL_NUT且图片的另一NAL单元的nal_unit_type在IDR_W_RADL至CRA_NUT的范围内时,将图片确定为拖尾图片。

[0112] (B) 当图片的NAL单元的nal_unit_type等于RADL_NUT且图片的另一NAL单元的nal_unit_type在IDR_W_RADL至CRA_NUT的范围内时,将图片确定为RADL图片。

[0113] (C) 当图片的NAL单元的nal_unit_type等于STSA_NUT且图片的另一NAL单元的nal_unit_type为IDR_W_RADL或IDR_N_LP时,将图片确定为STSA图片。

[0114] (D) 当图片的NAL单元的nal_unit_type等于STSA_NUT且图片的另一NAL单元的nal_unit_type为CRA_NUT时,将图片确定为拖尾图片。

[0115] (E) 当图片的NAL单元的nal_unit_type等于GDR_NUT且图片的另一NAL单元的nal_unit_type在IDR_W_RADL至CRA_NUT的范围内时,将图片确定为拖尾图片。

[0116] 根据一个或多个实施例,当由已编码图片引用的PPS的语法元素mixed_nalu_types_in_pic_flag等于1时,(例如由解码器)将已编码图片的图片类型确定为拖尾图片。

[0117] 上述方面可提供对上文在发明内容部分中描述的“问题1”的解决方案。

[0118] 根据一个或多个实施例,可以不允许STSANAL单元与IRAP NAL单元的混合物。

[0119] 例如,对于任何特定图片的VCL NAL单元,可如下实现:

[0120] 如果mixed_nalu_types_in_pic_flag等于0,则对于图片的所有已编码切片NAL单元,nal_unit_type值应该相同(例如,可确定为相同)。图片或PU称为与图片或PU的已编码切片NAL单元具有相同的NAL单元类型。

[0121] 否则(mixed_nalu_types_in_pic_flag等于1),一个或多个VCL NAL单元应全部具有(例如,可确定为全部具有)在IDR_W_RADL至CRA_NUT(包括端点)范围内的特定nal_unit_type值,而其他VCL NAL单元应全部具有(例如,可确定为全部具有)在RADL_NUT至RSV_VCL_6(包括端点)范围内的特定nal_unit_type值,或者其nal_unit_type值等于GDR_NUT或TRAIL_NUT。

[0122] 根据一个实施例,编码器可配置成应用以上内容以不允许STSANAL单元与IRAP NAL单元的混合物。根据一个实施例,解码器可配置成基于以上内容确定NAL单元类型的值。

[0123] 根据一个或多个实施例,可去除当前VVC规范草案JVET-P2001中对STSA_NUT的TemporalId约束。

[0124] 也就是说,例如,本申请实施例的实施例可以不实现如下约束:当nal_unit_type等于STSA_NUT时,TemporalId不应等于0。然而,实施例仍可实现如下约束:当nal_unit_type在IDR_W_RADL至RSV_IRAP_12(包括端点)的范围内时,TemporalId应(例如,可确定为)等于0。

[0125] 根据一个或多个实施例,可实现如下约束:其mixed_nalu_types_in_pic_flag等于1的图片的TemporalId应等于0。例如,本申请实施例的编码器或解码器可基于标志mixed_nalu_types_in_pic_flag等于1来确定图片的时间ID为0。

[0126] 上述方面可提供对上文在发明内容部分中描述的“问题2”的解决方案。

[0127] 根据一个或多个实施例,语法元素mixed_nalu_types_in_pic_flag可设置在图片报头或切片报头中,而不是设置在PPS中。在下面的表3中提供了图片报头中的语法元素mixed_nalu_types_in_pic_flag的示例。

[0128] 表3

	picture_header_rbsp () {	描述符
	...	
[0129]	mixed_nalu_types_in_pic_flag	U (1)
	...	
	}	

[0130] 语法元素mixed_nalu_types_in_pic_flag等于1,可表示与图片报头PH相关联的每个图片具有多于一个VCL NAL单元,且VCL NAL单元不具有相同的nal_unit_type值,且该图片不是IRAP图片。语法元素mixed_nalu_types_in_pic_flag等于0,可表示与PH相关联的每个图片具有一个或多个VCL NAL单元,且与PH相关联的每个图片的VCL NAL单元具有相同的nal_unit_type值。

[0131] 当语法元素no_mixed_nalu_types_in_pic_constraint_flag等于1时,mixed_nalu_types_in_pic_flag的值应(例如,可确定为)等于0。

[0132] 根据一个或多个实施例,语法元素mixed_nalu_types_in_pic_flag可设置在图片报头或切片报头中,存在标志位于SPS中。

[0133] 在下面的表4中提供了具有存在标志(sps_mixed_nalu_types_present_flag)的SPS的示例。

[0134] 表4

[0135]	seq_parameter_set_rbsp () {	描述符
	...	
	sps_mixed_nalu_types_present_flag	U (1)
	...	
	}	

[0136] 在下面的表5中提供了具有语法元素mixed_nalu_types_in_pic_flag的图片报头的示例。

[0137] 表5

[0138]	picture_header_rbsp () {	描述符
	...	
	if (sps_mixed_nalu_types_present_flag)	
	mixed_nalu_types_in_pic_flag	U (1)
	}	

[0139] 语法元素sps_mixed_nalu_types_present_flag等于1,可表示参考SPS的零个或多个图片具有多于一个VCL NAL单元,且VCL NAL单元不具有相同的nal_unit_type值,且该图片不是IRAP图片。语法元素sps_mixed_nalu_types_present_flag等于0,可表示参考SPS的每个图片具有一个或多个VCL NAL单元,且参考SPS的每个图片的VCL NAL单元具有相同的nal_unit_type值。

[0140] 当语法元素no_mixed_nalu_types_in_pic_constraint_flag等于1时,语法元素sps_mixed_nalu_types_present_flag的值应(例如,可确定为)等于0。

[0141] 语法元素mixed_nalu_types_in_pic_flag等于1,可表示与PH相关联的每个图片具有多于一个VCL NAL单元,且VCL NAL单元不具有相同的nal_unit_type值,且该图片不是IRAP图片。语法元素mixed_nalu_types_in_pic_flag等于0,可表示与PH相关联的每个图片具有一个或多个VCL NAL单元,且与PH相关联的每个图片的VCL NAL单元具有相同的nal_unit_type值。当不存在时,可(例如由解码器)推断mixed_nalu_types_in_pic_flag的值等于0。

[0142] 上述方面可提供对上文在发明内容部分中描述的“问题3”的解决方案。

[0143] 根据一个或多个实施例,语法元素标志mixed_nalu_types_in_pic_flag可以用指示符mixed_nalu_types_in_pic_idc替换。

[0144] 在下面的表6中提供了具有语法元素mixed_nalu_types_in_pic_idc的图片参数集的示例。

[0145] 表6

	pic_parameter_set_rbsp () {	描述符
	...	
[0146]	mixed_nalu_types_in_pic_idc	U (2)
	...	
	}	

[0147] 语法元素mixed_nalu_types_in_pic_idc等于1或2,可表示参考PPS的每个图片具有多于一个VCL NAL单元,且VCL NAL单元不具有相同的nal_unit_type值,且该图片不是IRAP图片。语法元素mixed_nalu_types_in_pic_idc等于0,可表示参考PPS的每个图片具有一个或多个VCL NAL单元,且参考PPS的每个图片的VCL NAL单元具有相同的nal_unit_type值。语法元素mixed_nalu_types_in_pic_idc的其他值可预留给ITU-T|ISO/IEC在将来使用。

[0148] 当语法元素no_mixed_nalu_types_in_pic_constraint_idc等于1时,mixed_nalu_types_in_pic_idc的值应(例如,由例如解码器确定为)等于0。

[0149] 对于图片picA中的、其nal_unit_type值nalUnitTypeA在IDR_W_RADL至CRA_NUT(包括端点)范围内的每个切片(图片picA还包含具有另一个nal_unit_type值的一个或多个切片(即,图片picA的mixed_nalu_types_in_pic_idc的值等于1)),可如下实现:

[0150] (A) 切片应属于(例如,可确定为属于)子图片subpicA,其中,对于子图片subpicA,对应subpic_treated_as_pic_flag[i]的值等于1。

[0151] (B) 切片不应属于(例如,可确定为不属于)包含VCL NAL单元的picA的子图片,其中切片的nal_unit_type不等于nalUnitTypeA。

[0152] (C) 对于CLVS中按解码次序在后面的所有PU,subpicA中的切片的RefPicList[0]和RefPicList[1]均不应包括在有效条目中按解码次序位于picA之前的任何图片。

[0153] RefPicList[0]可以是用于P切片的帧间预测的参考图片列表或者用于B切片的帧间预测的第一参考图片列表。RefPicList[1]可以是用于B切片的帧间预测的第二参考图片列表。

[0154] 对于任何特定图片的VCL NAL单元,可如下实现:

[0155] (A) 如果语法元素mixed_nalu_types_in_pic_idc等于1,则一个或多个VCL NAL单元应全部具有(例如,可确定为具有)在IDR_W_RADL至CRA_NUT(包括端点)范围内的特定nal_unit_type值,而其他VCL NAL单元应全部具有(例如,可确定为具有)在TRAIL_NUT至RSV_VCL_6(包括端点)范围内的特定nal_unit_type值,或者其nal_unit_type值等于GDR_NUT。

[0156] (B) 如果语法元素mixed_nalu_types_in_pic_idc等于2,则一个或多个VCL NAL单元应全部具有(例如,可确定为具有)等于RASL_NUT或RADL_NUT(包括端点)或者等于GDR_NUT的特定nal_unit_type值,而其他VCL NAL单元应全部具有(例如,可确定为具有)在TRAIL_NUT至RSV_VCL_6(包括端点)范围内的特定nal_unit_type值,或者其nal_unit_type值等于GDR_NUT,其中nal_unit_type不同于其他nal_unit_type。

[0157] (C) 否则(mixed_nalu_types_in_pic_idc等于0),对于图片的所有已编码切片NAL

单元, `nal_unit_type` 值应 (例如, 可确定为) 相同。图片或PU称为与图片或PU的已编码切片NAL单元具有相同的NAL单元类型。

[0158] 当由已编码图片引用的PPS的`mixed_nalu_types_in_pic_idc`等于1或2时, 该图片可 (例如由解码器) 确定为拖尾图片。

[0159] 上述方面可提供对上文在发明内容部分中描述的“问题4”的解决方案。

[0160] 根据一个或多个实施例, 当A层中的图片的语法元素`mixed_nalu_types_in_pic_flag`等于1时, A层的参考层即B层中的图片的`mixed_nalu_types_in_pic_flag`应 (例如, 可确定为) 等于1 (在同一个AU下)。

[0161] 上述方面可提供对上文在发明内容部分中描述的“问题5”的解决方案。

[0162] 根据一个或多个实施例, 一个或多个已编码视频数据比特流以及其中的语法结构和元素 (例如上述VCL NAL单元和参数集) 可由本申请实施例的解码器接收, 以对所接收的视频数据进行解码。根据本申请实施例的实施例, 本申请实施例的解码器可基于已编码图片的、具有混合VCL NAL单元类型的VCL NAL单元 (例如, 图5中示出的VCL NAL单元 (500)) 来对视频的已编码图片进行解码。

[0163] 例如参考图6, 解码器 (600) 可包括解码代码 (610), 解码代码 (610) 配置成使得解码器 (600) 的至少一个处理器基于VCL NAL单元对已编码图片进行解码。根据一个或多个实施例, 解码代码 (610) 可包括确定代码 (620), 确定代码 (620) 配置成使得解码器 (600) 的至少一个处理器如本申请实施例的实施例中所描述的: (a) 基于已编码图片的另一个或多个VCL NAL单元的NAL单元类型, 或基于指示符 (例如标志), 来确定或约束已编码图片的一个或多个VCL NAL单元的NAL单元类型, (b) 基于已编码图片的VCL NAL单元的一个或多个NAL单元类型, 或基于指示符 (例如标志), 来确定或约束已编码图片的图片类型, (c) 基于已编码图片的一个或多个VCL NAL单元的一个或多个NAL单元类型, 或基于指示符 (例如标志), 来确定或约束已编码图片的TemporalID, 和/或 (d) 基于所接收的或所确定的另一指示符 (例如标志), 来确定或约束如下指示符 (例如标志), 该指示符 (例如标志) 指示已编码图片是否具有多个VCL NAL单元, 而该多个VCL NAL单元具有混合VCL NAL单元类型。

[0164] 本申请实施例的实施例可以以任何顺序单独使用或组合使用。此外, 本申请实施例的方法、编码器和解码器中的每一个可通过处理电路 (例如, 一个或多个处理器或者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时性计算机可读介质中的程序。

[0165] 可以将上述技术实现为计算机软件, 该计算机软件使用计算机可读指令, 且物理地存储在一个或多个计算机可读介质中。例如, 图7示出了适合于实施所公开的主题的实施例的计算机系统 (900)。

[0166] 可使用任何合适的机器代码或计算机语言来对计算机软件进行编码, 任何合适的机器代码或计算机语言可经受汇编、编译、链接或类似的机制以创建包括指令的代码, 该指令可由一个或多个计算机中央处理单元 (CPU)、图形处理单元 (GPU) 等直接执行或通过解释、微代码等执行。

[0167] 指令可以在各种类型的计算机或其组件上执行, 例如包括个人计算机、平板计算机、服务器、智能电话、游戏装置、物联网装置等。

[0168] 图7所示的计算机系统 (900) 的组件本质上是示例性的, 并不旨在对实施本申请实

施例的实施例的计算机软件的使用范围或功能提出任何限制。组件的配置也不应被解释为具有与计算机系统(900)的示例性实施例中所示的组件中的任何一个组件或组件的组合相关的任何依赖性要求。

[0169] 计算机系统(900)可包括某些人机接口输入装置。此类人机接口输入装置可响应于一个或多个人类用户例如通过下述的输入:触觉输入(例如:击键、划动,数据手套移动)、音频输入(例如:语音、拍手)、视觉输入(例如:手势)、嗅觉输入(未描绘)。人机接口装置还可用于捕获不一定与人的意识输入直接相关的某些媒体,例如音频(例如:语音、音乐、环境声音)、图像(例如:扫描的图像、从静止图像相机获取摄影图像)、视频(例如二维视频、包括立体视频的三维视频)等。

[0170] 输入人机接口装置可包括下述中的一项或多项(每种中仅示出一个):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)和相机(908)。

[0171] 计算机系统(900)还可包括某些人机接口输出装置。此类人机接口输出装置可例如通过触觉输出、声音、光和气味/味道来刺激一个或多个人类用户的感官。此类人机接口输出装置可包括触觉输出装置(例如触摸屏(910)的触觉反馈、数据手套或操纵杆(905),但是也可以是不作为输入设备的触觉反馈装置)。例如,此类装置可以是音频输出装置(例如:扬声器(909)、耳机(未示出))、视觉输出装置(例如包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕的屏幕(910),每种屏幕有或没有触摸屏输入功能,每种屏幕都有或没有触觉反馈功能-其中的一些屏幕能够通过诸如立体图像输出之类的装置、虚拟现实眼镜(未描绘)、全息显示器和烟箱(未描绘)以及打印机(未描绘)来输出二维视觉输出或超过三维输出。

[0172] 计算机系统(900)还可包括人类可访问存储装置及其关联介质:例如包括具有CD/DVD等介质(921)的CD/DVD ROM/RW(920)的光学介质、指状驱动器(922)、可拆卸硬盘驱动器或固态驱动器(923)、诸如磁带和软盘之类的传统磁性介质(未描绘)、诸如安全软件狗之类的基于专用ROM/ASIC/PLD的装置(未描绘)等。

[0173] 本领域技术人员还应该理解,结合当前公开的主题使用的术语“计算机可读介质”不涵盖传输介质、载波或其他瞬时信号。

[0174] 计算机系统(900)还可包括到达一个或多个通信网络的接口。网络可例如是无线网络、有线网络、光网络。网络可进一步是本地网络、广域网络、城域网络、车辆和工业网络、实时网络、耐延迟网络等。网络的示例包括诸如以太网之类的局域网、无线LAN、包括GSM、3G、4G、5G、LTE等的蜂窝网络、包括有线电视、卫星电视和地面广播电视的电视有线或无线广域数字网络、包括CANBus的车辆和工业用电视等等。某些网络通常需要连接到某些通用数据端口或外围总线(949)的外部网络接口适配器(例如计算机系统(900)的USB端口);如下所述,其他网络接口通常通过连接到系统总线而集成到计算机系统9200的内核中(例如,连接到PC计算机系统以太网接口或连接到智能手机计算机系统蜂窝网络接口)。计算机系统(900)可使用这些网络中的任何一个与其他实体通信。此类通信可以是仅单向接收的(例如,广播电视)、仅单向发送的(例如,连接到某些CANbus装置的CANbus)或双向的,例如,使用局域网或广域网数字网络连接到其他计算机系统。此类通信可包括与云计算环境(955)的通信。如上所述,可以在那些网络和网络接口中的每一个上使用某些协议和协议栈。

[0175] 上述人机接口装置、人机可访问的存储装置和网络接口 (954) 可附接到计算机系统 (900) 的内核 (940)。

[0176] 内核 (940) 可包括一个或多个中央处理单元 (CPU) (941)、图形处理单元 (GPU) (942)、现场可编程门区域 (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。这些装置以及只读存储器 (ROM) (945)、随机存取存储器 (946)、诸如内部非用户可访问的硬盘驱动器、SSD 等之类的内部大容量存储器 (947) 可通过系统总线 (948) 连接。在一些计算机系统中, 可以以一个或多个物理插头的形式访问系统总线 (948), 以能够通过附加的 CPU、GPU 等进行扩展。外围装置可直接连接到内核的系统总线 (948) 或通过外围总线 (949) 连接到内核的系统总线 (948)。外围总线的体系结构包括 PCI、USB 等。图形适配器 950 可包括在内核 940 中。

[0177] CPU (941)、GPU (942)、FPGA (943) 和加速器 (944) 可执行某些指令, 这些指令可以组合来构成上述计算机代码。该计算机代码可存储在 ROM (945) 或 RAM (946) 中。过渡数据也可以存储在 RAM (946) 中, 而永久数据可例如存储在内部大容量存储器 (947) 中。可通过使用高速缓存来进行到任何存储装置的快速存储及检索, 该高速缓存可与下述紧密关联: 一个或多个 CPU (941)、GPU (942)、大容量存储器 (947)、ROM (945)、RAM (946) 等。

[0178] 计算机可读介质可以在其上具有用于执行各种由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出于本申请实施例的目的而专门设计和构造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用的类型。

[0179] 作为非限制性示例, 可由于一个或多个处理器 (包括 CPU、GPU、FPGA、加速器等) 执行包含在一种或多种有形的计算机可读介质中的软件而使得具有架构 (900), 特别是内核 (940) 的计算机系统提供功能。此类计算机可读介质可以是与如上所述的用户可访问的大容量存储器相关联的介质, 以及某些非暂时性的内核 (940) 的存储器, 例如内核内部大容量存储器 (947) 或 ROM (945)。可以将实施本申请实施例的各种实施例的软件存储在此类装置中并由内核 (940) 执行。根据特定需要, 计算机可读介质可包括一个或多个存储装置或芯片。软件可引起内核 (940), 特别是其中的处理器 (包括 CPU、GPU、FPGA 等) 执行本文所描述的特定过程或特定过程的特定部分, 包括定义存储在 RAM (946) 中的数据结构以及根据由软件定义的过程来修改此类数据结构。附加地或替换地, 可由于硬连线或以其他方式体现在电路 (例如, 加速器 (944)) 中的逻辑而使得计算机系统提供功能, 该电路可替换软件或与软件一起运行以执行本文描述的特定过程或特定过程的特定部分。在适当的情况下, 提及软件的部分可包含逻辑, 反之亦然。在适当的情况下, 提及计算机可读介质的部分可包括存储用于执行的软件的电路 (例如集成电路 (IC))、体现用于执行的逻辑的电路或两者都包括。本申请实施例包括硬件和软件的任何合适的组合。

[0180] 虽然本申请实施例已经描述了多个非限制性示例性实施例, 但是存在落入本申请实施例的范围内的修改、置换和各种替换等效物。因此, 应当理解, 本领域技术人员将能够设计出许多系统和方法, 这些系统和方法虽然未在本文中明确示出或描述, 但是其体现了本申请实施例的原理, 因此落入本申请实施例的精神和范围内。

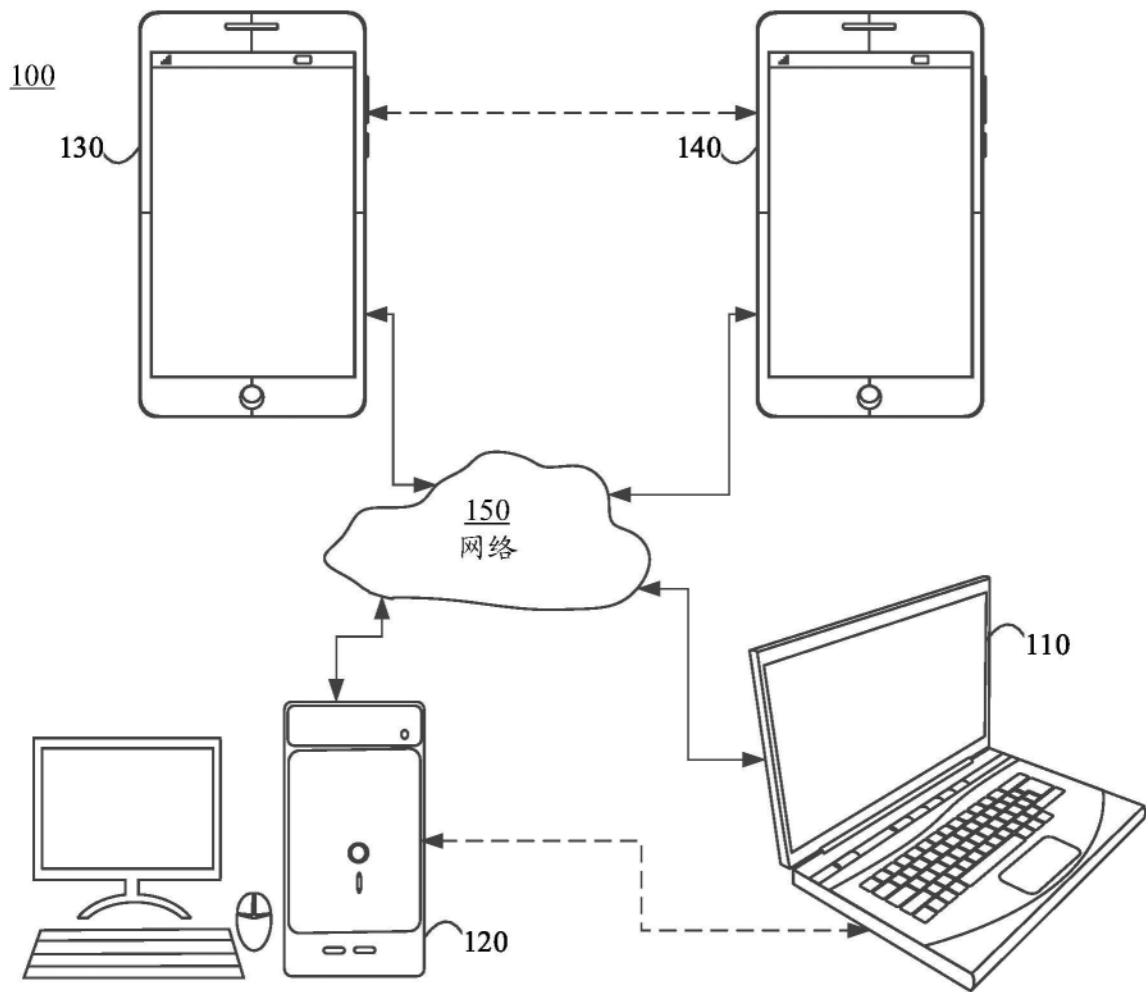


图1

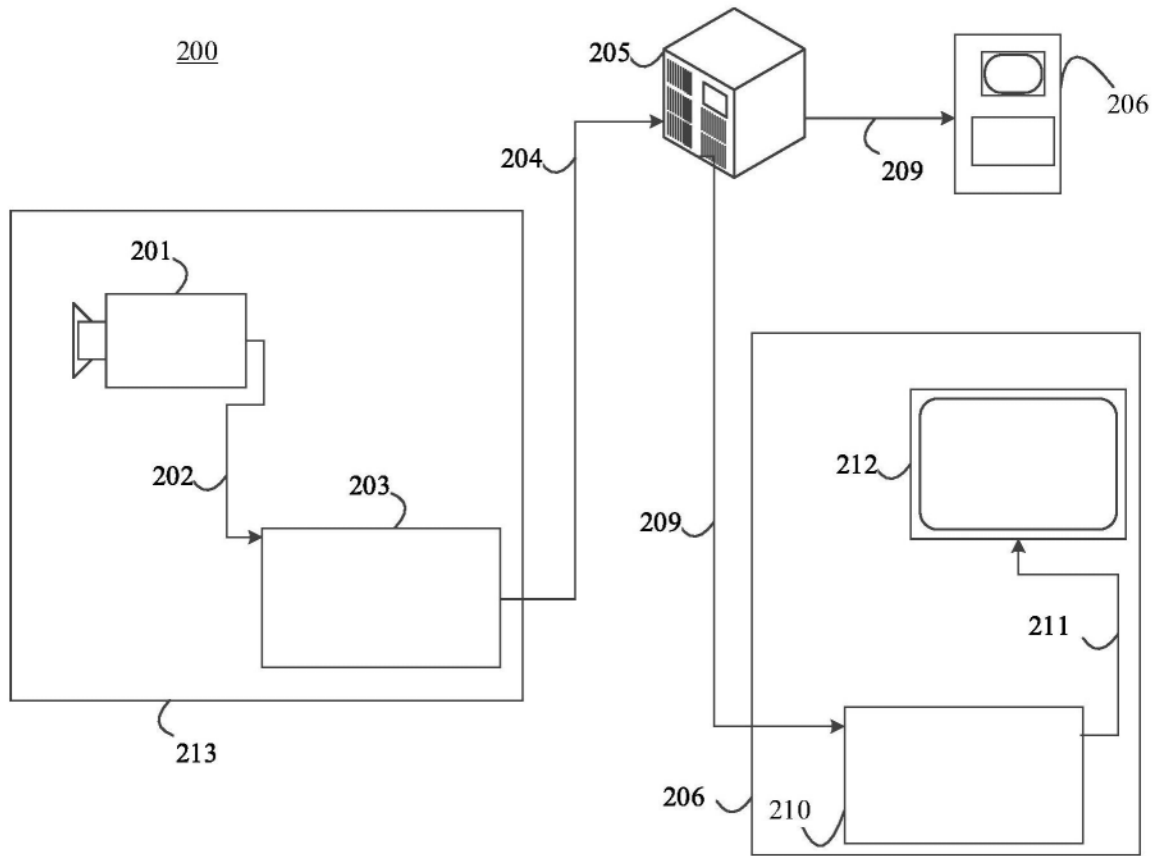


图2

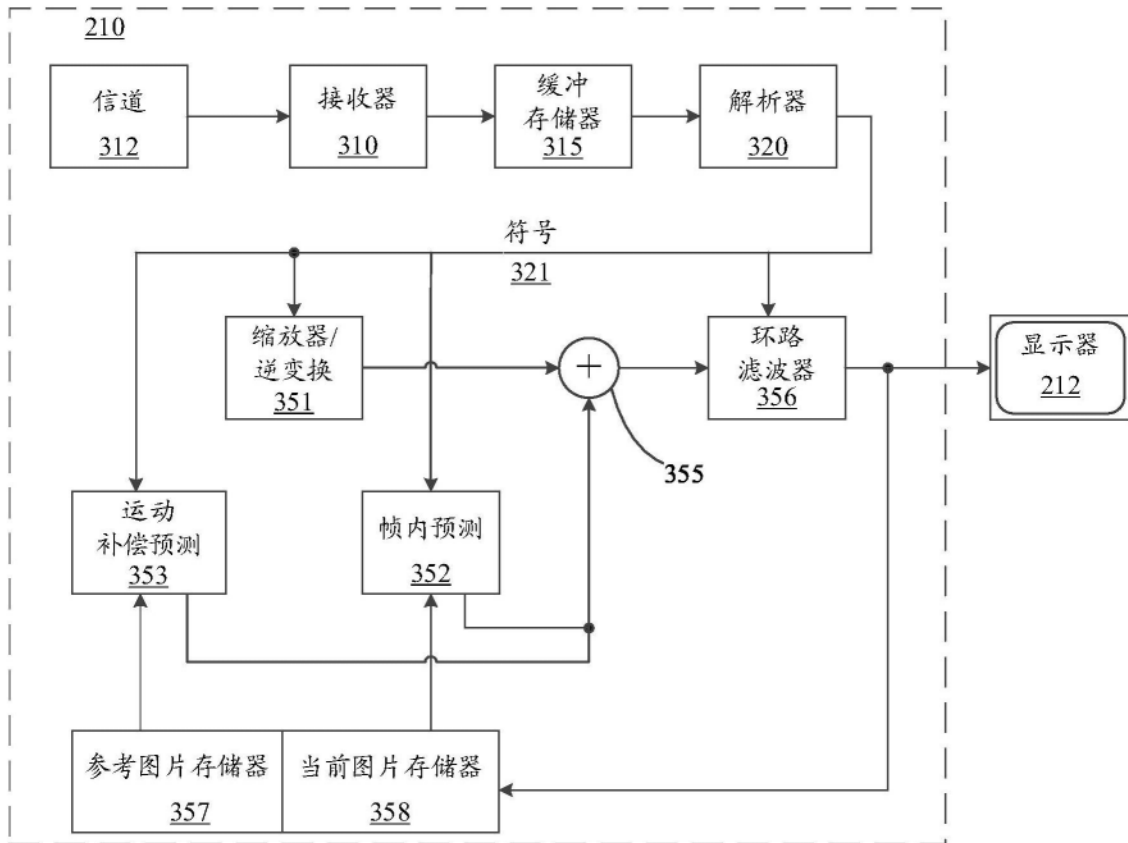


图3

203

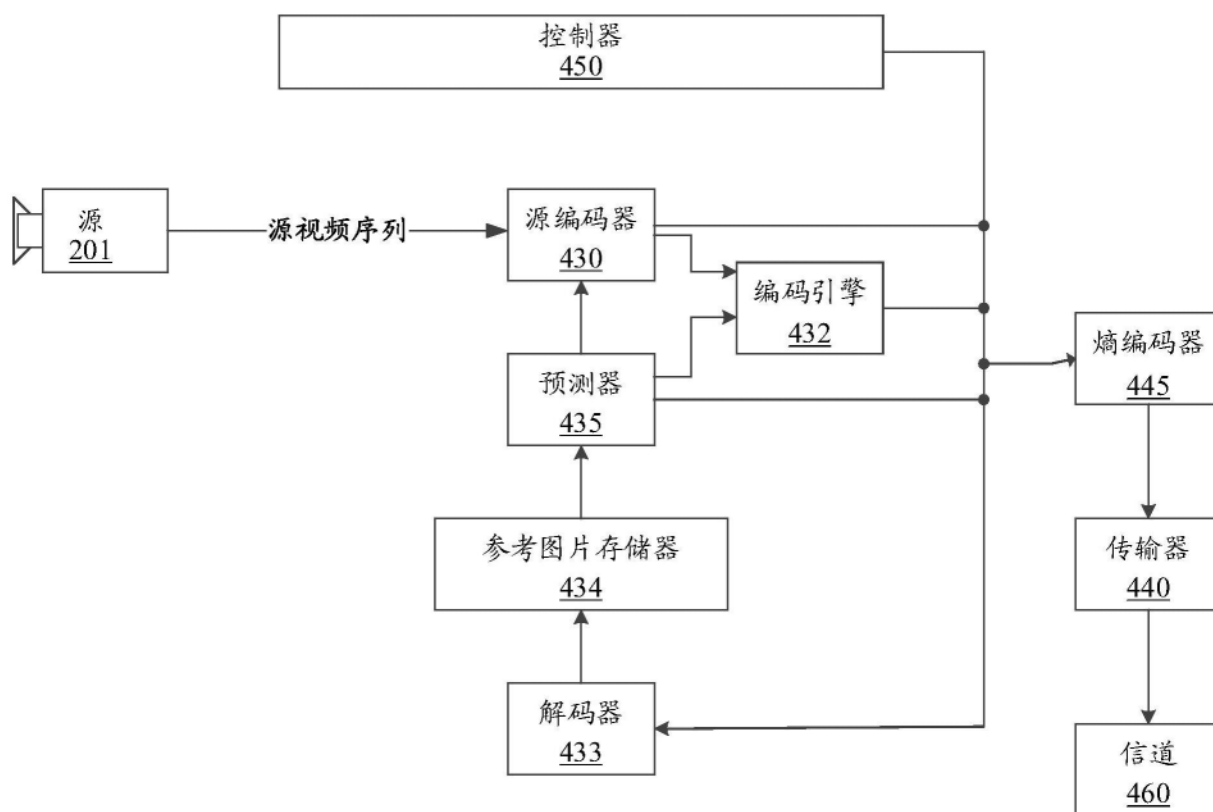


图4

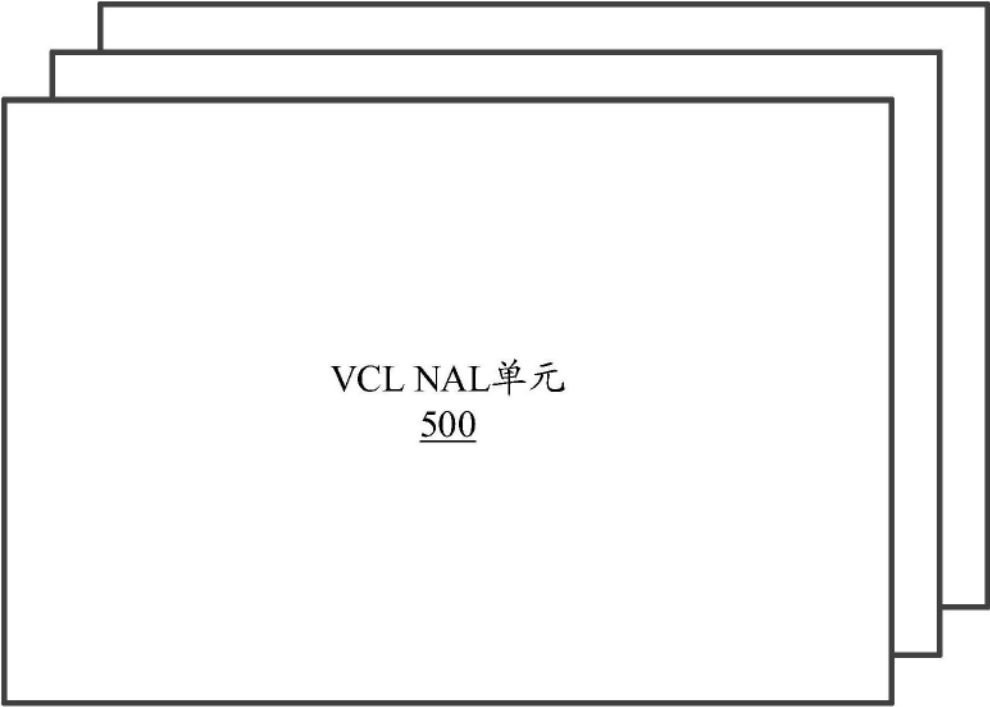


图5

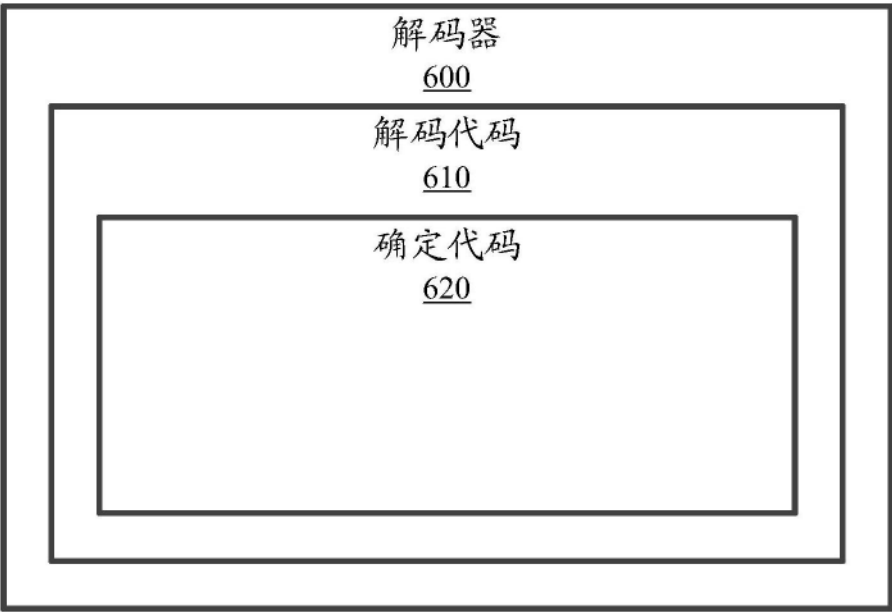


图6

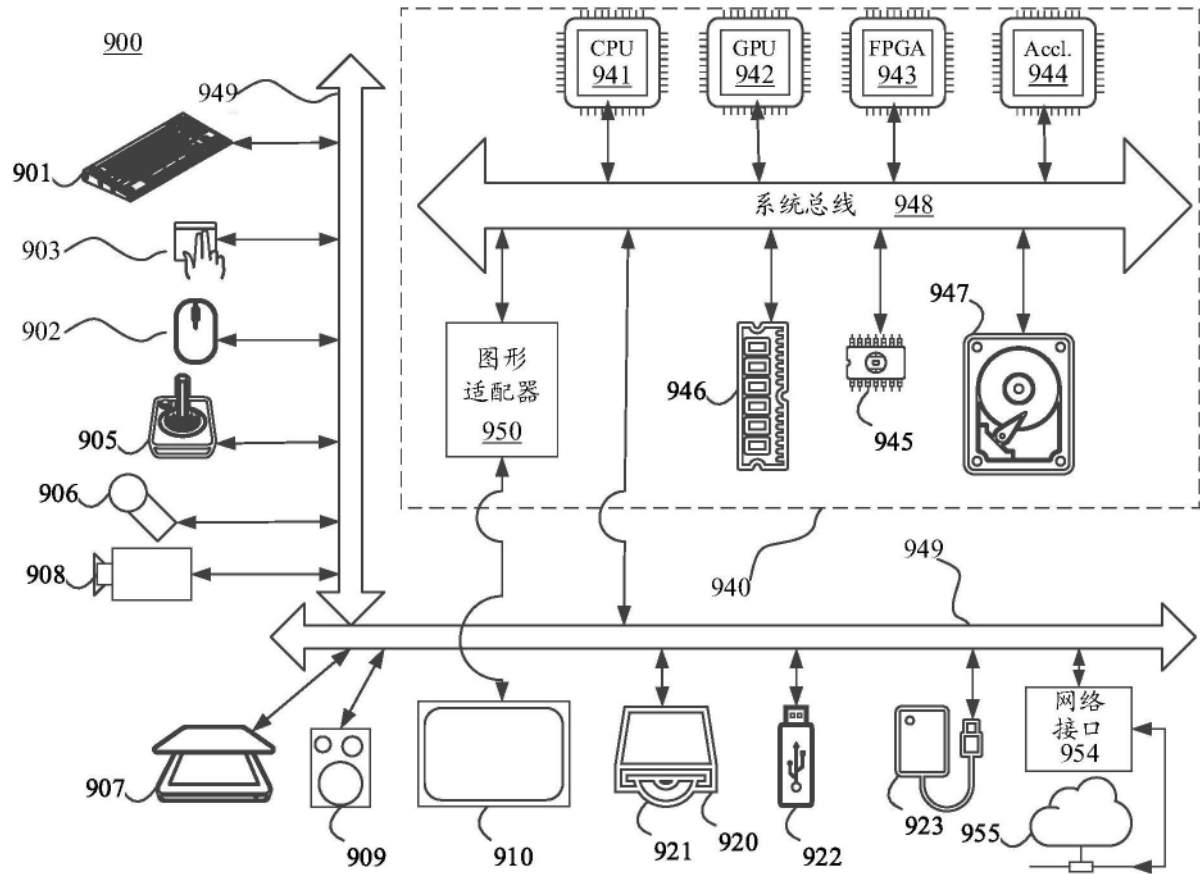


图7