



(10) 授权公告号 CN 112154663 B

(45) 授权公告日 2022. 07. 12

(21) 申请号 201980033757.1

(22) 申请日 2019.05.30

(65) 同一申请的已公布的文献号
申请公布号 CN 112154663 A

(43) 申请公布日 2020.12.29

(30) 优先权数据
62/680,468 2018.06.04 US
16/210,635 2018.12.05 US

(85) PCT国际申请进入国家阶段日
2020.11.19

(86) PCT国际申请的申请数据
PCT/US2019/034485 2019.05.30

(87) PCT国际申请的公布数据
W02019/236362 EN 2019.12.12

(73) 专利权人 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 许晓中 李翔 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 王兆林

(51) Int.Cl.
H04N 19/119 (2006.01)
H04N 19/513 (2006.01)
H04N 19/105 (2006.01)
H04N 19/159 (2006.01)
H04N 19/176 (2006.01)
H04N 19/139 (2006.01)

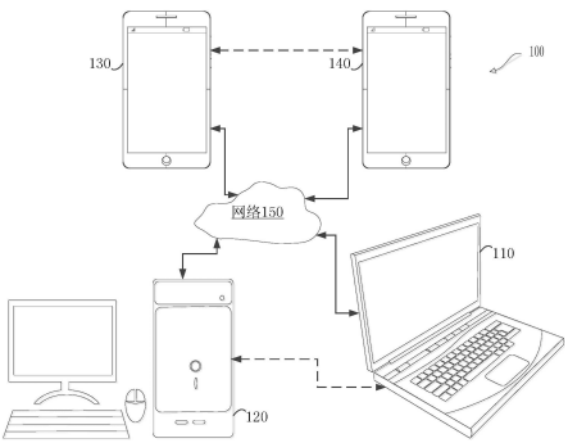
(56) 对比文件
WO 2016123081 A1,2016.08.04
CN 107005708 A,2017.08.01
CN 107211156 A,2017.09.26
US 2013058394 A1,2013.03.07

审查员 尚琴

权利要求书2页 说明书19页 附图10页

(54) 发明名称
视频解码方法和相关装置

(57) 摘要
本公开提供一种视频解码的方法,包括:获取当前图片,并且对于当前图片中包括的当前块,识别与当前图片不同的参考图片中包括的参考块,其中,当前块被划分为多个子块(CBSB),并且参考块具有多个子块(RBSB)。该方法包括:确定RBSB的参考图片是否是当前图片,并且响应于确定RBSB的参考图片是当前图片,将RBSB的编码模式确定为帧内模式。该方法还包括:响应于确定RBSB的参考图片不是当前图片,基于对应的RBSB的编码模式是否是帧内模式和帧间模式之一来确定CBSB中的一个CBSB的运动矢量预测值。



1. 一种视频解码的方法,其特征在于,所述方法包括:

从已编码视频比特流中获取当前图片;

对于所述当前图片中包括的当前块,识别参考图片中包括的参考块,所述当前块被划分为多个子块(CBSB),所述参考块包括多个子参考块(RBSB),所述多个RBSB分别对应于所述多个CBSB;

确定所述RBSB的所述参考图片是否是所述当前图片;

响应于确定所述RBSB的所述参考图片是所述当前图片,将所述RBSB的编码模式确定为帧内模式,将所述RBSB对应的CBSB的运动矢量预测值设置为默认运动矢量;

响应于确定所述RBSB的所述参考图片不是所述当前图片,

对于每个CBSB,响应于对应的RBSB以帧内模式或帧内块复制模式编码,将所述CBSB的运动矢量预测值设置为默认运动矢量,所述默认运动矢量是零运动矢量;

对于每个CBSB,响应于对应的RBSB以帧间模式编码,将所述CBSB的运动矢量预测值确定为等于所述对应的RBSB的运动矢量预测值;

基于每个CBSB的运动矢量预测值对每个CBSB执行基于子块的时间运动矢量预测,对每个CBSB进行解码。

2. 根据权利要求1所述的方法,其特征在于,所述CBSB的运动矢量预测值是与所述对应的RBSB相关联的所述运动矢量预测值的缩放版本。

3. 根据权利要求1-2任一项所述的方法,其特征在于,所述识别参考图片中包括的参考块,包括:

根据与相邻于所述当前块的块相关联的运动矢量预测值来识别所述参考块。

4. 根据权利要求1-2任一项所述的方法,其特征在于,所述参考图片是来自参考图片序列中的第一参考图片,所述参考图片序列与所述当前图片相关联。

5. 一种视频解码的装置,其特征在于,所述装置包括获取单元、识别单元、第一确定单元和第二确定单元:

所述获取单元,用于从已编码视频比特流中获取当前图片;

所述识别单元,用于对于所述当前图片中包括的当前块,识别参考图片中包括的参考块,所述当前块被划分为多个子块(CBSB),所述参考块包括多个子块(RBSB),所述多个RBSB分别对应于所述多个CBSB;

所述第一确定单元,用于确定所述RBSB的所述参考图片是否是所述当前图片;

所述第二确定单元,用于响应于确定所述RBSB的所述参考图片是所述当前图片,将所述RBSB的编码模式确定为帧内模式,将所述RBSB对应的CBSB的运动矢量预测值设置为默认运动矢量;

响应于确定所述RBSB的所述参考图片不是所述当前图片,

对于每个CBSB,响应于对应的RBSB以帧内模式或帧内块复制模式编码,将所述CBSB的运动矢量预测值设置为默认运动矢量,所述默认运动矢量是零运动矢量;对于每个CBSB,响应于对应的RBSB以帧间模式编码,将所述CBSB的运动矢量预测值确定为等于所述对应的RBSB的运动矢量预测值;

所述装置还用于,基于每个CBSB的运动矢量预测值对每个CBSB执行基于子块的时间运动矢量预测,对每个CBSB进行解码。

6. 根据权利要求5所述的装置,其特征在于,所述CBSB的运动矢量预测值是与所述对应的RBSB相关联的所述运动矢量预测值的缩放版本。

7. 根据权利要求5所述的装置,其特征在于,所述识别单元,用于:

根据与相邻于所述当前块的块相关联的运动矢量预测值来识别所述参考块。

8. 根据权利要求5所述的装置,其特征在于,所述参考图片是来自参考图片序列中的第一参考图片,所述参考图片序列与所述当前图片相关联。

9. 一种存储有指令的非暂态计算机可读介质,其特征在于,所述指令在由视频解码器中的处理器执行时使所述处理器执行权利要求1-4任意一项所述的方法。

10. 一种计算机设备,其特征在于,所述设备包括处理器以及存储器:

所述存储器用于存储程序代码,并将所述程序代码传输给所述处理器;

所述处理器用于根据所述程序代码中的指令执行权利要求1-4中任意一项所述的方法。

视频解码方法和相关装置

[0001] 交叉引用

[0002] 本公开内容要求于2018年6月4日提交的美国临时申请No. 62/680,468“METHODS FOR SUB-BLOCK BASED TEMPORAL MOTION VECTOR PREDICTION”的优先权的权益,该申请的全部内容通过引用合并于本文中。

技术领域

[0003] 本公开涉及视频处理技术,尤其涉及基于子块的时间运动矢量预测的方法和相关装置。

背景技术

[0004] 本文中提供的背景描述是出于总体上呈现本公开内容的背景的目的。就本背景技术部分中描述的工作的程度而言,目前署名的发明人的工作以及在提交时可以不另外被限定作为现有技术的描述的方面既没有明确地也没有隐含地被承认为针对本公开内容的现有技术。

[0005] 近几十年来,已经已知使用具有运动补偿的图片间预测来执行视频编码和解码。未压缩的数字视频可以包括一系列图片,每个图片具有空间维度,例如 1920×1080 亮度样本和相关联的色度样本。该一系列图片可以具有例如每秒60个图片或60Hz的固定的或可变的图片速率(也被非正式地称为帧速率)。未压缩的视频具有高的码率要求。例如,每样本8位的1080p60 4:2:0视频(60Hz帧速率下 1920×1080 亮度样本分辨率)需要接近1.5Gbit/s的带宽。一小时这样的视频需要大于600GB的存储空间。

[0006] 视频编码和解码的一个目的可以是通过压缩来减少输入视频信号中的冗余。压缩可以帮助减少以上提及的带宽或存储空间需求,在一些情况下减少两个数量级或更多。可以采用无损压缩、有损压缩及其组合。无损压缩是指可以从已压缩的原始信号重建原始信号的精确副本的技术。当使用有损压缩时,重建的信号可能与原始信号不同,但是原始信号与重建的信号之间的失真足够小,以使重建的信号对于预期应用是有用的。在视频的情况下,有损压缩应用广泛。容忍的失真量取决于应用;例如,某些消费者流媒体应用的用户可能比电视贡献应用的用户容忍更高的失真。可达到的压缩比可以反映出:更高的可允许/可容忍的失真可以产生更高的压缩比。

[0007] 运动补偿可以是有损压缩技术,并且可以涉及下述技术:在由运动矢量(此后被称为MV)指示的方向上进行空间移位之后,使用来自先前重建的图片或其一部分(参考图片)的样本数据的块来预测新重建的图片或图片部分。在一些情况下,参考图片可以与当前重建中的图片相同。MV可以具有两个维度X和Y,也可以具有三个维度,其中第三维度是使用中的参考图片的指示(后者间接地可以是时间维度)。

[0008] 在一些视频压缩技术中,可以从其他MV预测适用于样本数据的某些区域的MV,例如从与样本数据的另一区域有关并且在解码顺序上前于该MV的其他MV来进行预测,样本数据的另一区域在空间上与被重建的区域相邻。这样做可以大幅减少对MV进行编码所需的数

据量,从而消除冗余并且增加压缩。MV预测可以有效地工作,例如,这是由于在对从摄像装置导出的输入视频信号(被称为自然视频)进行编码时,存在统计上的可能性——即比单个MV适用的区域大的区域在相似的方向上移动,因此可以在一些情况下,使用从相邻区域的MV导出的相似运动矢量来进行预测。这导致针对给定区域找到的MV与从周围MV预测的MV相似或相同,并且又可以在熵编码之后以比直接对MV进行编码的情况下使用的位数小的位数来表示。在一些情况下,MV预测可以是原始信号(即样本流)导出的信号(即MV)的无损压缩的示例。在其他情况下,MV预测本身可以是有损的,例如这是由于从若干周围MV计算预测值时的舍入误差。

[0009] 在H.265/HEVC(ITU-T H.265建议书,“高效视频编解码(High Efficiency Video Coding)”,2016年12月)中描述了各种MV预测机制。在H.265提供的多种MV预测机制中,本申请描述的是此后被称为“空间合并”的技术。

[0010] 一些形式的帧间预测是在子块级执行的。然而,基于子块的时间运动矢量预测模式,例如替选时间运动矢量预测(ATMVP)和空间时间运动矢量预测(STMVP),要求以帧间模式对对应的子块进行编码。然而,这些时间运动矢量预测模式不能够处理以诸如帧内块复制模式之类的帧内模式编码的子块。

发明内容

[0011] 本公开内容的示例性实施例包括一种用于解码器的视频解码的方法。该方法包括:从已编码视频比特流中获取当前图片。该方法还包括:对于当前图片中包括的当前块,识别与当前图片不同的参考图片中包括的参考块,其中,当前块被划分为多个子块(CBSB),并且参考块具有多个子块(RBSB),多个子块(RBSB)分别对应于多个CBSB中的不同CBSB。该方法还包括:确定RBSB的参考图片是否是当前图片;以及响应于确定RBSB的参考图片是当前图片,将RBSB的编码模式确定为帧内模式。该方法还包括:响应于确定RBSB的参考图片不是当前图片,(i)对于CBSB中的一个CBSB,确定RBSB的编码模式是否是帧内模式和帧间模式之一,以及(ii)基于对应的RBSB的编码模式是否是帧内模式和帧间模式之一来确定CBSB中的所述一个CBSB的运动矢量预测值。

[0012] 本公开内容的示例性实施例包括一种用于视频解码的装置。所述装置包括获取单元、识别单元、第一确定单元和第二确定单元。所述获取单元,用于从已编码视频比特流中获取当前图片;所述识别单元,用于对于所述当前图片中包括的当前块,识别与所述当前图片不同的参考图片中包括的参考块,所述当前块被划分为多个子块(CBSB),所述参考块具有多个子块(RBSB),所述多个RBSB分别对应于所述多个CBSB中的不同CBSB;所述第一确定单元,用于确定所述RBSB的所述参考图片是否是所述当前图片;所述第二确定单元,用于响应于确定所述RBSB的所述参考图片是所述当前图片,将所述RBSB的编码模式确定为帧内模式;响应于确定所述RBSB的所述参考图片不是所述当前图片,(i)对于所述CBSB中的一个CBSB,确定所述RBSB的编码模式是否是帧内模式和帧间模式之一,(ii)基于对应的RBSB的编码模式是否是所述帧内模式和所述帧间模式之一来确定所述CBSB中的所述一个CBSB的运动矢量预测值。

[0013] 本公开内容的示例性实施例包括一种存储有指令的非暂态计算机可读介质,该指令在由视频解码器中的处理器执行时使处理器执行一个方法。该方法包括:从已编码视频

比特流中获取当前图片。该方法还包括：对于当前图片中包括的当前块，识别与当前图片不同的参考图片中包括的参考块，其中，当前块被划分为多个子块 (CBSB)，并且参考块具有多个子块 (RBSB)，多个RBSB分别对应于多个CBSB中的不同CBSB。该方法还包括：确定RBSB的参考图片是否是当前图片；以及响应于确定RBSB的参考图片是当前图片，将RBSB的编码模式确定为帧内模式。该方法还包括：响应于确定RBSB的参考图片不是当前图片，(i) 对于CBSB中的一个CBSB，确定RBSB的编码模式是否是帧内模式和帧间模式之一，以及(ii) 基于对应的RBSB的编码模式是否是帧内模式和帧间模式之一来确定CBSB中的所述一个CBSB的运动矢量预测值。

[0014] 本公开内容的示例性实施例包括一种计算机设备，所述设备包括处理器以及存储器；所述存储器用于存储程序代码，并将所述程序代码传输给所述处理器；所述处理器用于根据所述程序代码中的指令执行任一项所述的用于解码器的视频解码方法。

[0015] 由上述方案可以看出，本公开的一些实现方式可以实现基于时间运动矢量预测模式处理以诸如帧内块复制模式之类的帧内模式编码的子块。

附图说明

[0016] 根据以下详细描述和附图，所公开的主题的另外的特征、性质和各种优点将变得更加明显，在附图中：

[0017] 图1是根据一个实施例的通信系统的简化框图的示意图；

[0018] 图2是根据另一个实施例的通信系统的简化框图的示意图；

[0019] 图3是根据一个实施例的解码器的简化框图的示意图；

[0020] 图4是根据一个实施例的编码器的简化框图的示意图；

[0021] 图5示出了根据另一实施例的编码器的框图；

[0022] 图6示出了根据另一实施例的解码器的框图；

[0023] 图7是帧内图片块补偿的示意图；

[0024] 图8是当前块以及当前块的周围空间合并候选的示意图；

[0025] 图9是当前块的子块以及参考块的对应子块的示意图；

[0026] 图10示出了由编码器或解码器执行的处理的实施例；

[0027] 图11是根据一个实施例的计算机系统的示意图。

具体实施方式

[0028] 本公开中涉及的英文缩写的中文释义和英文全称请参见说明书后附的附录A。

[0029] 图1示出了根据本公开内容的一个实施例的通信系统(100)的简化框图。通信系统(100)包括可以经由例如网络(150)彼此通信的多个终端设备。例如，通信系统(100)包括经由网络(150)互连的第一对终端设备(110)和(120)。在图1示例中，第一对终端设备(110)和(120)执行单向数据传输。例如，终端设备(110)可以对视频数据(例如，由终端设备(110)捕获的视频图片流)进行编码以用于经由网络(150)传输到另一终端设备(120)。经编码的视频数据可以以一个或多个已编码视频比特流的形式传输。终端设备(120)可以从网络(150)接收已编码视频数据，对已编码视频数据进行解码以恢复所述视频图片，并且根据恢复的视频数据显示视频图片。单向数据传输在媒体服务应用等中是较常见的。

[0030] 在另一示例中,通信系统(100)包括执行例如在视频会议期间发生的已编码视频数据的双向传输的第二对终端设备(130)和(140)。对于双向数据传输,在示例中,终端设备(130)和(140)中的每个终端设备可以对视频数据(例如,由终端设备捕获的视频图片流)进行编码以用于经由网络(150)传输到终端设备(130)和(140)中的另一终端设备。终端设备(130)和(140)中的每个终端设备还可以接收由终端设备(130)和(140)中的另一终端设备传输的已编码视频数据,并且可以对已编码视频数据进行解码以恢复所述视频图片,并且可以根据恢复的视频数据在可访问的显示设备上显示视频图片。

[0031] 在图1示例中,终端设备(110)、(120)、(130)和(140)可以被示出为个人计算机和智能电话,但本公开内容的原理可以不限于此。本公开内容的实施例适用于膝上型计算机、平板电脑、媒体播放器和/或专用视频会议设备。网络(150)表示在终端设备(110)、终端设备(120)、终端设备(130)和终端设备(140)之间传送已编码视频数据的任何数目的网络,包括例如有线连接(有线)和/或无线通信网络。网络(150)可以在电路交换和/或分组交换信道中交换数据。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,除非在本文中另有说明,否则网络(150)的架构和拓扑对于本公开内容的操作来说可以是不重要的。

[0032] 作为所公开的主题的应用的示例,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。所公开的主题可以同等地适用于其他支持视频的应用,包括例如视频会议,数字电视,在包括CD、DVD、存储棒等的数字介质上存储压缩视频等。

[0033] 流式传输系统可以包括捕获子系统(213),该捕获子系统(213)可以包括创建例如未压缩的视频图片流(202)的视频源(201),例如数字摄像装置。在示例中,视频图片流(202)包括由数字摄像装置拍摄的样本。当与经编码的视频数据(204)(或已编码视频比特流)进行比较时被描绘为强调高数据量的粗线的视频图片流(202)可以由包括耦接至视频源(201)的视频编码器(203)的电子设备(220)处理。视频编码器(203)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。当与视频图片流(202)进行比较时,经编码的视频数据(204)(或已编码视频比特流(204))可以存储在流服务器(205)上以供将来使用,经编码的视频数据(204)被描绘为强调较低数据量的细线。诸如图2中的客户端子系统(206)和(208)的一个或多个流客户端子系统可以访问流服务器(205)以检索经编码的视频数据(204)的副本(207)和(209)。客户端子系统(206)可以包括例如电子设备(230)中的视频解码器(210)。视频解码器(210)对经编码的视频数据(204)的传入副本(207)进行解码,并且创建可以在显示器(212)(例如,显示屏)或另一呈现设备(未描绘)上呈现的传出视频图片流(211)。在一些流式传输系统中,可以根据某些视频编码/压缩标准对经编码的视频数据(204)、其副本(207)和(209)(例如,视频比特流)进行编码。这些标准的示例包括ITU-T H.265建议书。在示例中,正在开发的视频编码标准被非正式地称为通用视频编码或VVC。所公开的主题可以在VVC的上下文中使用。

[0034] 应注意,电子设备(220)和电子设备(230)可以包括其他部件(未示出)。例如,电子设备(220)可以包括视频解码器(未示出),并且电子设备(230)可以包括视频编码器(未示出)。

[0035] 图3示出了根据本公开内容的一个实施例的视频解码器(310)的框图。视频解码器(310)可以被包括在电子设备(330)中。电子设备(330)可以包括接收器(331)(例如,接收电

路)。视频解码器(310)可以用于代替图2示例中的视频解码器(210)。

[0036] 接收器(331)可以接收要由视频解码器(310)解码的一个或多个已编码视频序列;在同一实施例或另一实施例中,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。可以从信道(301)接收已编码视频序列,信道(301)可以是到存储经编码的视频数据的存储设备的硬件/软件链路。接收器(331)可以接收经编码的视频数据以及其他数据,例如,可以转发到其各自的使用实体(未描绘)的已编码音频数据和/或辅助数据流。接收器(331)可以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)可以耦接在接收器(331)与熵解码器/解析器(320)(此后被称为“解析器(320)”)之间。在一些应用中,缓冲存储器(315)是视频解码器(310)的一部分。在其他应用中,缓冲存储器(315)可以在视频解码器(310)的外部(未描绘)。在又一些其他应用中,在视频解码器(310)的外部可以有缓冲存储器(未描绘)以例如防止网络抖动,并且在视频解码器(310)的内部可以有另外的缓冲存储器(315)以例如处理播出定时。当接收器(331)从具有足够带宽和可控性的存储/转发设备或从等时同步网络接收数据时,可能不需要缓冲存储器(315),或者缓冲存储器(315)可以是小的。为了在诸如因特网的最优分组网络上使用,可能需要缓冲存储器(315),缓冲存储器(315)可以相对较大并且可以有利地具有自适应大小,并且可以至少部分在操作系统或视频解码器(310)的外部的类似元件(未描绘)中实现。

[0037] 视频解码器(310)可以包括根据已编码视频序列重建符号(321)的解析器(320)。这些符号(321)的类别包括用于管理视频解码器(310)的操作的信息,以及控制诸如呈现设备(312)(例如,显示屏)的呈现设备的潜在信息,该呈现设备(312)不是电子设备(330)的一体部分,而是可以耦接至电子设备(330),如图3所示。用于(一个或多个)呈现设备的控制信息可以是辅助增强信息(SEI消息)或视频可用性信息(VUI)参数集片段(未描绘)的形式。解析器(320)可以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以根据视频编码技术或标准,并且可以遵循各种原理,包括可变长度编码、霍夫曼编码(Huffman coding)、具有或不具有背景灵敏度的算术编码等。解析器(320)可以基于与群组对应的至少一个参数,从已编码视频序列提取用于视频解码器中的像素的子群组中的至少一个子群组的子群组参数集。子群组可以包括图片群组(GOP)、图片、图块、切片、宏块、编码单元(CU)、块、变换单元(TU)、预测单元(PU)等。解析器(320)还可以从已编码视频序列中提取信息,例如变换系数、量化器参数值、运动矢量等。

[0038] 解析器(320)可以对从缓冲存储器(315)接收的视频序列执行熵解码/解析操作,从而创建符号(321)。

[0039] 取决于已编码视频图片或其部分(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)的重建可以涉及多个不同的单元。涉及哪些单元以及涉及方式可以由解析器(320)从已编码视频序列解析的子群组控制信息来控制。出于简洁起见,未描述解析器(320)与下面的多个单元之间的这样的子群组控制信息流。

[0040] 除了已经提及的功能块之外,视频解码器(310)可以在概念上被细分为如下所述的许多功能单元。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并且可以至少部分地彼此集成。然而,出于描述所公开的主题的目的,概念上细分为下面的功能单元是适当的。

[0041] 第一单元是缩放器/逆变换单元(351)。缩放器/逆变换单元(351)从解析器(320)接收作为(一个或多个)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)可以输出包括可以输入到聚合器(355)中的样本值的块。

[0042] 在一些情况下,缩放器/逆变换单元(351)的输出样本可以属于帧内编码块;即:不使用来自先前重建的图像的预测性信息,但可以使用来自当前图像的先前重建部分的预测性信息的块。这样的预测性信息可以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)使用从当前图片缓冲器(358)提取的周围已重建信息生成大小和形状与重建中的块相同的块。例如,当前图片缓冲器(358)缓冲部分重建的当前图片和/或完全重建的当前图片。在一些情况下,聚合器(355)基于每个样本将帧内图片预测单元(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。

[0043] 在其他情况下,缩放器/逆变换单元(351)的输出样本可以属于帧间编码和潜在运动补偿块。在这样的情况下,运动补偿预测单元(353)可以访问参考图片存储器(357)以提取用于预测的样本。在根据属于该块的符号(321)对提取的样本进行运动补偿之后,这些样本可以由聚合器(355)添加到缩放器/逆变换单元(351)的输出(在该情况下被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)从其提取预测样本的参考图片存储器(357)内的地址可以由运动矢量控制,运动矢量以符号(321)的形式供运动补偿预测单元(353)使用,符号(321)可以具有例如X、Y和参考图片分量。运动补偿还可以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。

[0044] 聚合器(355)的输出样本可以在环路滤波器单元(356)中采用各种环路滤波技术。视频压缩技术可以包括环路内滤波器技术,环路内滤波器技术由被包括在已编码视频序列(也被称为已编码视频比特流)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)的参数控制,然而视频压缩技术还可响应于在对已编码图片或已编码视频序列的先前(在解码顺序上)部分进行解码期间获得的元信息,以及响应于先前重建并经环路滤波的样本值。

[0045] 环路滤波器单元(356)的输出可以是样本流,样本流可以输出至呈现设备(312)以及存储在参考图片存储器(357),以用于将来的帧间图片预测。

[0046] 一旦完全重建,某些已编码图片就可以用作参考图片以用于将来预测。例如,一旦与当前图片对应的已编码图片被完全重建,并且已编码图片(通过例如解析器(320))被识别为参考图片,则当前图片缓冲器(358)就可以变为参考图片存储器(357)的一部分,并且可以在开始重建随后的已编码图片之前重新分配新的当前图片缓冲器(358)。

[0047] 视频解码器(310)可以根据诸如ITU-T H.265建议书的标准中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以及视频压缩技术或标准中记录的配置文件两者的意义上,已编码视频序列可以符合正在使用的视频压缩技术或标准指定的语法。具体地,配置文件可以从视频压缩技术或标准中可用的所有工具中选择某些工具作为在配置文件下可供使用的仅有工具。对于合规性,还要求已编码视频序列的复杂度在由视频压缩技术或标准的层级限定的范围内。在一些情况下,层级限制最大图片大小、最大帧率、最大重建取样率(以例如每秒兆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,由层级设置的限制可以通过假设参考解码器(HRD)规范以及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。

[0048] 在一个实施例中,接收器(331)可以接收附加(冗余)数据和经编码的视频。附加数据可以被包括为(一个或更多个)已编码视频序列的一部分。附加数据可以由视频解码器(310)使用来正确地对数据进行解码和/或更准确地重建原始视频数据。附加数据可以是例如时间、空间或信噪比(SNR)增强层、冗余切片、冗余图片、前向纠错码等形式。

[0049] 图4示出了根据本公开内容的一个实施例的视频编码器(403)的框图。视频编码器(403)被包括在电子设备(420)中。电子设备(420)包括传输器(440)(例如,传输电路)。视频编码器(403)可以用于代替图2示例中的视频编码器(203)。

[0050] 视频编码器(403)可以从视频源(401)(并非图4示例中的电子设备(420)的一部分)接收视频样本,视频源(401)可以捕获要由视频编码器(403)编码的视频图像。在另一示例中,视频源(401)是电子设备(420)的一部分。

[0051] 视频源(401)可以提供要由视频编码器(403)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以具有任何合适的位深度(例如:8位、10位、12位…)、任何色彩空间(例如,BT.601Y CrCb、RGB…)以及任何合适的采样结构(例如,Y CrCb 4:2:0、Y CrCb 4:4:4)。在媒体服务系统中,视频源(401)可以是存储先前已准备的视频的存储设备。在视频会议系统中,视频源(401)可以是捕获本地图像信息作为视频序列的摄像装置。可以将视频数据提供为在按次序观看时被赋予运动的多个单独的图片。图片自身可以被组织为空间像素阵列,其中,取决于使用中的采样结构、色彩空间等,每个像素可以包括一个或更多个样本。所属领域技术人员可以很容易理解像素与样本之间的关系。以下描述着眼于样本。

[0052] 根据一个实施例,视频编码器(403)可以实时地或在应用所需的任何其他时间约束下,将源视频序列的图片编码且压缩为已编码视频序列(443)。施行适当的编码速度是控制器(450)的一个功能。在一些实施例中,控制器(450)控制如下所述的其他功能单元并且在功能上耦接至这些其他功能单元。出于简洁起见,未描绘该耦接。由控制器(450)设置的参数可以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 值等)、图片大小、图片群组(GOP)布局、最大运动矢量搜索范围等。控制器(450)可以被配置成具有其他合适的功能,这些功能属于针对某些系统设计优化的视频编码器(403)。

[0053] 在一些实施例中,视频编码器(403)被配置成在编码环路中操作。作为非常简单的描述,在一个示例中,编码环路可以包括源编码器(430)(例如,负责基于要编码的输入图片和(一个或更多个)参考图片创建诸如符号流的符号)以及嵌入于视频编码器(403)中的(本地)解码器(433)。(本地)解码器(433)以类似于(远程)解码器将创建样本数据的方式重建符号以创建样本数据(因为在所公开的主题中考虑的视频压缩技术中,符号与已编码视频比特流之间的任何压缩是无损的)。将重建的样本流(样本数据)输入至参考图片存储器(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(434)中的内容在本地编码器与远程编码器之间也是位精确的。换言之,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“看到”的样本值完全相同。该参考图片同步性基本原理(以及在例如由于信道误差而无法维持同步性的情况下产生的漂移)也用于一些相关领域。

[0054] (本地) 解码器 (433) 的操作可以与例如已经在上面结合图3详细描述的视频解码器 (310) 的“远程”解码器的操作相同。然而, 还简要参照图3, 当符号可用并且熵编码器 (445) 和解析器 (320) 可以无损地将符号编码/解码为已编码视频序列时, 包括缓冲存储器 (315) 和解析器 (320) 的视频解码器 (310) 的熵解码部分可能无法完全在本地解码器 (433) 中实现。

[0055] 此时可以观察到, 除了存在于解码器中的解析/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。出于该原因, 所公开的主题着眼于解码器操作。由于编码器技术与全面地描述的解码器技术互逆, 因此可以简化对编码器技术的描述。仅在某些区域中需要并且在下面提供更详细的描述。

[0056] 在操作期间, 在一些示例中, 源编码器 (430) 可以执行运动补偿预测编码, 该运动补偿预测编码参考来自视频序列中被指定为“参考图片”的一个或更多个先前编码的图片对输入图片进行预测性编码。以此方式, 编码引擎 (432) 对输入图片的像素块与可以被选作输入图片的 (一个或更多个) 预测参考的 (一个或更多个) 参考图片的像素块之间的差异进行编码。

[0057] (本地) 解码器 (433) 可以基于由源编码器 (430) 创建的符号对可以被指定为参考图片的图片的已编码视频数据进行解码。编码引擎 (432) 的操作可以有利地是有损处理。当已编码视频数据可以在视频解码器 (图4中未示出) 处被解码时, 重建的视频序列通常可以是具有一些误差的源视频序列的副本。(本地) 解码器 (433) 复制可以由视频解码器对参考图片执行的解码处理, 并且可以使重建的参考图片存储在参考图片存储器 (434) 中。以此方式, 视频编码器 (403) 可以本地存储重建的参考图片的副本, 该副本与将由远端视频解码器获得的重建的参考图片具有共同内容 (不存在传输误差)。

[0058] 预测器 (435) 可以针对编码引擎 (432) 执行预测搜索。即, 对于要编码的新图片, 预测器 (435) 可以在参考图片存储器 (434) 中搜索可以用作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) 可以基于样本块逐像素块操作, 以找到适当的预测参考。在一些情况下, 根据由预测器 (435) 获得的搜索结果所确定的, 输入图片可以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。

[0059] 控制器 (450) 可以管理源编码器 (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组参数。

[0060] 可以在熵编码器 (445) 中对所有以上提及的功能单元的输出进行熵编码。熵编码器 (445) 通过根据本领域技术人员已知的例如诸如霍夫曼编码、可变长度编码、算术编码等的技术对由各种功能单元生成的符号进行无损压缩, 将符号转换成已编码视频序列。

[0061] 传输器 (440) 可以缓冲由熵编码器 (445) 创建的 (一个或更多个) 已编码视频序列, 以为经由通信信道 (460) 进行传输做准备, 该通信信道可以是到将存储经编码的视频数据的存储设备的硬件/软件链路。传输器 (440) 可以将来自视频编码器 (403) 的已编码视频数据与要传输的其他数据例如已编码音频数据和/或辅助数据流 (未示出源) 合并。

[0062] 控制器 (450) 可以管理视频编码器 (403) 的操作。在编码期间, 控制器 (450) 可以为每个已编码图片分配可能影响可以应用于相应的图片的编码技术的某些编码图片类型。例如, 通常可以将图片分配为以下图片类型中之一:

[0063] 帧内图片(I图片),其可以是在不将序列中的任何其他图片用作预测源的情况下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,包括例如独立解码器刷新(“IDR”)图片。所属领域技术人员了解I图片的变体及其相应的应用和特征。

[0064] 预测性图片(P图片),其可以是在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的情况下,使用帧内预测或帧间预测进行编码和解码的图片。

[0065] 双向预测性图片(B图片),其可以是在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的情况下使用帧内预测或帧间预测进行编码和解码的图片。类似地,多个预测性图片可以使用多于两个参考图片和相关联的元数据以用于单个块的重建。

[0066] 源图片通常可以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),并且逐块进行编码。这些块可以参考其他(已经编码的)块进行预测性编码,所述其他块通过应用于块的相应的图片的编码分配来确定。例如,I图片的块可以进行非预测性编码,或I图片的块可以参照同一图片的已经编码的块来进行预测性编码(空间预测或帧内预测)。P图片的像素块可以参照一个先前已编码的参考图片经由空间预测或经由时间预测进行预测性编码。B图片的块可以参照一个或两个先前已编码的参考图片经由空间预测或经由时间预测进行预测性编码。

[0067] 视频编码器(403)可以根据例如ITU-T H.265建议书的预定视频编码技术或标准执行编码操作。在其操作中,视频编码器(403)可以执行各种压缩操作,包括利用输入视频序列中的时间和空间冗余的预测性编码操作。因此,已编码视频数据可以符合由使用的视频编码技术或标准指定的语法。

[0068] 在一个实施例中,传输器(440)可以传输附加数据和已编码视频。源编码器(430)可以包括这样的数据作为已编码视频序列的一部分。附加数据可以包括时间/空间/SNR增强层、诸如冗余图片和切片的其他形式的冗余数据、补充增强信息(SEI)消息、视觉可用性信息(VUI)参数集片段等。

[0069] 捕获到的视频可以作为呈时间序列的多个源图片(视频图片)。帧内图片预测(通常被简化为帧内预测)利用给定图片中的空间相关性,而帧间图片预测利用图片之间的(时间或其他)相关性。在示例中,被称为当前图片的编码/解码中的特定图片被分割成块。在当前图片中的块类似于视频中先前已编码且仍被缓冲的参考图片中的参考块时,可以通过被称为运动矢量的矢量对当前图片中的块进行编码。所述运动矢量指向参考图片中的参考块,并且可以在使用多个参考图片的情况下具有识别参考图片的第三维度。

[0070] 在一些实施例中,双向预测技术可以用于帧间图片预测。根据双向预测技术,使用两个参考图片,例如在解码顺序上均在视频中的当前图片之前(但在显示顺序上可能分别是过去和将来)的第一参考图片和第二参考图片。可以通过指向第一参考图片中的第一参考块的第一运动矢量以及指向第二参考图片中的第二参考块的第二运动矢量来对当前图片中的块进行编码。可以通过第一参考块和第二参考块的组合来预测块。

[0071] 此外,合并模式技术可以用于帧间图片预测以提高编码效率。

[0072] 根据本公开内容的一些实施例,诸如帧间图片预测和帧内图片预测的预测以块为单位来执行。例如,根据HEVC标准,视频图片序列中的图片被分割成编码树单元(CTU)以用于压缩,图片中的CTU具有相同大小,例如 64×64 像素、 32×32 像素或 16×16 像素。一般来说,CTU包括三个编码树块(CTB),即一个亮度CTB和两个色度CTB。每个CTU可以被递归地以

四叉树拆分成一个或多个编码单元 (CU)。例如, 可以将 64×64 像素的CTU拆分成一个 64×64 像素的CU, 或4个 32×32 像素的CU, 或16个 16×16 像素的CU。在示例中, 分析每个CU以确定用于CU的预测类型, 例如帧间预测类型或帧内预测类型。取决于时间和/或空间可预测性, CU被拆分成一个或更多个预测单元 (PU)。通常, 每个PU包括亮度预测块 (PB) 和两个色度PB。在一个实施例中, 编码 (编码/解码) 中的预测操作以预测块为单位来执行。使用亮度预测块作为预测块的示例, 预测块包括像素值 (例如, 亮度值) 的矩阵, 例如 8×8 像素、 16×16 像素、 8×16 像素、 16×8 像素等。

[0073] 图5示出了根据本公开内容的另一实施例的视频解码器 (503) 的图。该视频编码器 (503) 被配置成接收视频图片序列中的当前视频图片内的样本值的处理块 (例如, 预测块), 并且将处理块编码到作为已编码视频序列的一部分的已编码图片中。在示例中, 视频编码器 (503) 用于代替图2示例中的视频编码器 (203)。

[0074] 在HEVC示例中, 视频编码器 (503) 接收用于处理块的样本值的矩阵, 处理块例如 8×8 样本的预测块等。视频编码器 (503) 使用例如率失真优化来确定是否使用帧内模式、帧间模式或双向预测模式来最佳地对处理块进行编码。当要在帧内模式下编码处理块时, 视频编码器 (503) 可以使用帧内预测技术以将处理块编码到编码图片中; 并且当要在帧间模式或双向预测模式下编码处理块时, 视频编码器 (503) 可以分别使用帧间预测或双向预测技术以将处理块编码到编码图片中。在某些视频编码技术中, 合并模式可以是帧间图片预测子模式, 其中, 在不借助于预测值外部的已编码运动矢量分量的情况下, 从一个或更多个运动矢量预测值导出运动矢量。在某些其他视频编码技术中, 可以存在适用于主题块的运动矢量分量。在示例中, 视频编码器 (503) 包括其他部件, 例如确定处理块的模式的模式决策模块 (未示出)。

[0075] 在图5示例中, 视频编码器 (503) 包括如图5所示耦接在一起的帧间编码器 (530)、帧内编码器 (522)、残差计算器 (523)、开关 (526)、残差编码器 (524)、通用控制器 (521) 和熵编码器 (525)。

[0076] 帧间编码器 (530) 被配置成接收当前块 (例如, 处理块) 的样本、将该块与参考图片中的一个或更多个参考块 (例如, 先前图片和后来图片中的块) 进行比较、生成帧间预测信息 (例如, 根据帧间编码技术的冗余信息描述、运动矢量、合并模式信息)、以及使用任何合适的技术基于帧间预测信息来计算帧间预测结果 (例如, 预测的块)。

[0077] 帧内编码器 (522) 被配置成接收当前块 (例如, 处理块) 的样本、在一些情况下将该块与同一图片中已经编码的块进行比较、在变换之后生成量化系数、以及在一些情况下还生成帧内预测信息 (例如, 根据一个或更多个帧内编码技术的帧内预测方向信息)。

[0078] 通用控制器 (521) 被配置成确定通用控制数据, 并且基于通用控制数据控制视频编码器 (503) 的其他部件。在示例中, 通用控制器 (521) 确定块的模式, 并且基于该模式将控制信号提供到开关 (526)。例如, 当模式是帧内模式时, 通用控制器 (521) 控制开关 (526) 来选择供残差计算器 (523) 使用的帧内模式结果, 并且控制熵编码器 (525) 来选择帧内预测信息并将帧内预测信息包括在比特流中; 以及当模式是帧间模式时, 通用控制器 (521) 控制开关 (526) 来选择供残差计算器 (523) 使用的帧间预测结果, 并且控制熵编码器 (525) 来选择帧间预测信息并将帧间预测信息包括在比特流中。

[0079] 残差计算器 (523) 被配置成计算接收的块与从帧内编码器 (522) 或帧间编码器

(530) 选择的预测结果之间的差(残差数据)。残差编码器(524)被配置成基于残差数据操作,以对残差数据进行编码以生成变换系数。在示例中,残差编码器(524)被配置成转换频域中的残差数据,并且生成变换系数。变换系数然后经受量化处理以获得量化的变换系数。

[0080] 熵编码器(525)被配置成将比特流格式化以包括经编码的块。熵编码器(525)被配置成根据诸如HEVC标准的合适标准包括各种信息。在示例中,熵编码器(525)被配置成包括通用控制数据、选择的预测信息(例如,帧内预测信息或帧间预测信息)、残差信息和比特流中的其他合适的信息。应注意,根据公开的主题,当在帧间模式或双向预测模式的合并子模式中对块进行编码时,不存在残差信息。

[0081] 图6示出了根据本公开内容的另一实施例的视频解码器(610)的图。视频解码器(610)被配置成接收作为已编码视频序列的一部分的已编码图片,并且对已编码图片进行解码以生成重建的图片。在示例中,视频解码器(610)用于代替图2示例中的视频解码器(210)。

[0082] 在图6示例中,视频解码器(610)包括如图6所示耦接在一起的熵解码器(671)、帧间解码器(680)、残差解码器(673)、重建模块(674)和帧内解码器(672)。

[0083] 熵解码器(671)可以被配置成根据已编码图片重建表示构成已编码图片的语法元素的某些符号。这样的符号可以包括例如用于对块进行编码的模式(例如,帧内模式、帧间模式、双向预测模式、后两者的合并子模式或另一子模式)、可以分别识别供帧内解码器(672)或帧间解码器(680)使用以进行预测的某些样本或元数据的预测信息(例如,帧内预测信息或帧间预测信息)、呈例如量化的变换系数的形式的残差信息等。在示例中,当预测模式是帧间或双向预测模式时,将帧间预测信息提供给帧间解码器(680);以及当预测类型是帧内预测类型时,将帧内预测信息提供给帧内解码器(672)。残差信息可以经受逆量化并且被提供给残差解码器(673)。

[0084] 帧间解码器(680)被配置成接收帧间预测信息,并且基于帧间预测信息生成帧间预测结果。

[0085] 帧内解码器(672)被配置成接收帧内预测信息,并且基于帧内预测信息生成预测结果。

[0086] 残差解码器(673)被配置成执行逆量化以提取解量化的变换系数,并且处理解量化的变换系数以将残差从频域转换到空间域。残差解码器(673)还可能某些控制信息(以包括量化器参数(QP)),并且信息可以由熵解码器(671)提供(由于这仅是低量控制信息,因此未描绘数据路径)。

[0087] 重建模块(674)被配置成在空间域中组合由残差解码器(673)输出的残差与预测结果(可以是由帧间预测模块或帧内预测模块输出的情况)以形成重建的块,该重建的块可以是重建的图片的一部分,重建的图片又可以是重建的视频的一部分。应注意,可以执行诸如解块操作等的其他合适的操作来提高视觉质量。

[0088] 应注意,可以使用任何合适的技术来实现视频编码器(203)、视频编码器(403)和视频编码器(503)以及视频解码器(210)、视频解码器(310)和视频解码器(610)。在一个实施例中,可以使用一个或更多个集成电路来实现视频编码器(203)、视频编码器(403)和视频编码器(503)以及视频解码器(210)、视频解码器(310)和视频解码器(610)。在另一实施例中,视频编码器(203)、(403)和(503)以及视频解码器(210)、(310)和(610)可以使用执行

软件指令的一个或更多个处理器来实现。

[0089] 来自不同图片的基于块的补偿可以被称为运动补偿。块的补偿还可以根据同一图片内的先前重建的区域进行,这可以被称为帧内图片块补偿或者帧内块复制。例如,将指示当前块与参考块之间的偏移的位移矢量称为块矢量。根据一些实施例,块矢量指向已经被重建并且可用于参考的参考块。同样,出于并行处理的考虑,超出图块/切片边界或波前梯形边界的参考区域也可以从块矢量的参考中排除。由于这些约束,块矢量可以与运动补偿中的运动矢量(MV)不同,在运动补偿中,运动矢量可以为任意值(在x或y方向为正或负)。

[0090] 图7示出了帧内图片块补偿(例如,帧内块复制模式)的实施例。在图7中,当前图片700包括已经被编码/解码的一组块(即,灰色块)以及尚未被编码/解码的一组块(即,白色块)。尚未被编码/解码的块中的一个块的子块702可以与指向先前已经被编码/解码的另一子块706的块矢量704相关联。因此,与子块706相关联的任何运动信息可以用于子块702的编码/解码。

[0091] 根据一些实施例,块矢量的编码是显式的。在其他实施例中,块矢量的编码是隐式的。在显式模式下,用信号表示块矢量与其预测值之间的差异,而在隐式模式下,以与合并模式中的运动矢量预测相似的方式从块矢量的预测值中恢复出块矢量。在一些实施例中,块矢量的分辨率限于整数位置。在其他实施例中,块矢量指向分数位置。

[0092] 根据一些实施例,使用参考索引来用信号表示使用块级的帧内图片块补偿(即,帧内块复制模式),其中,将当前解码的图片视为参考图片,将该参考图片放置在参考图片列表的最后位置。该参考图片还可以与其他时间参考图片一起在解码图片缓冲器(DPB)中进行管理。

[0093] 根据一些实施例,参考块在用于预测当前块之前被水平或垂直地翻转(例如,翻转的帧内块副本)。在一些实施例中, $M \times N$ 编码块内部的每个补偿单元是 $M \times 1$ 或 $1 \times N$ 行(例如,基于行的帧内块复制)。

[0094] 根据一些实施例,执行块级的运动补偿,其中,当前块是用于使用相同的运动信息来执行运动补偿的处理单元。就这一点而言,若给定块的大小,则该块中的所有像素将使用相同的运动信息来形成其预测块。块级运动补偿的示例包括使用空间合并候选、时间候选以及在双向预测中的来自现有合并候选的运动矢量的组合。

[0095] 参照图8,当前块(801)包括在运动搜索处理期间已由编码器/解码器发现的样本,该样本能够根据已经在空间偏移的相同大小的先前块进行预测。在一些实施例中,取代于直接对MV编码,可以使用与标记为A0、A1和B0、B1、B2(分别对应802到806)的五个周围样本中的任意一个样本相关联的MV,从与一个或更多个参考图片,例如最近(按解码次序)的参考图片,相关联的元数据中直接得到所述MV。块A0、A1、B0、B1和B2可被称为空间合并候选。

[0096] 根据一些实施例,运动补偿块内部的不同位置处的像素(例如,子块)可以具有不同的运动信息。与块级运动信息的这些差异可以被导出而不是用信号表示。这种类型的运动补偿可以被称为子块级运动补偿,其允许块的运动补偿小于块本身。就这一点而言,每个块可以具有多个子块,每个子块可以包含不同的运动信息。

[0097] 子块级运动补偿的示例包括基于子块的时间运动矢量预测,其中,当前块的子块具有不同的运动矢量。子块级运动补偿的另一个示例是ATMVP,其为下述方法:允许每个编码块从并置的参考图片的多个块中提取多组运动信息,所述多个块小于当前编码块。

[0098] 子块级运动补偿的另一个示例包括具有子块调整的空间/时间融合,其中,对于当前块中的每个子块,根据该子块的空间/时间相邻者的运动矢量来调整该子块的运动矢量。在这种模式下,对于某些子块,可能需要来自时间参考图片中的对应子块的运动信息。

[0099] 子块级运动补偿的另一个示例是仿射编码运动补偿块,其中,根据相邻块的运动矢量,首先导出当前块的四个角处的运动矢量。随后,使用仿射模型来导出(例如,子块或像素级的)当前块的其余运动矢量,使得与各自子块的相邻者相比,每个子块可以具有不同的运动矢量。

[0100] 子块级运动补偿的另一个示例是使用解码器侧运动矢量导出的合并候选细化。在这种模式下,在获得当前块或者当前块的子块的一个或多个运动矢量预测值之后,可以使用诸如模板匹配或双边匹配的方法来进行进一步细化给定的(一个或多个)运动矢量预测值。细化的运动矢量预测值可以用于执行运动补偿。可以在编码器侧和解码器侧二者处都执行相同的细化操作,使得解码器不需要关于细化如何从原始预测值移位的附加信息。此外,可以将跳过模式视为特殊的合并模式,其中,除了从当前块的相邻者中导出当前块的运动信息之外,当前块的预测残差也为零。

[0101] 根据一些实施例,在子块时间运动矢量预测中,当前块的子块可以具有从时间参考图片导出的不同的运动矢量预测值。例如,识别出一组运动信息,该一组运动信息包括当前块的运动矢量和相关联的参考索引。可以根据第一可用空间合并候选确定运动信息。使用该运动信息,确定用于当前块的参考图片中的参考块。参考块也被划分为子块。在一些实施例中,对于当前图片中的每个当前块子块(CBSB),在参考图片中存在对应的参考块子块(RBSB)。

[0102] 在一些实施例中,对于每个CBSB,如果对应的RBSB使用一组运动信息以帧间模式被编码,则该运动信息被转换(例如,使用诸如时间运动矢量预测中的运动矢量缩放等方法)并且被用作该CBSB的运动矢量预测值。下面进一步详细描述用于处理以帧内模式(例如,帧内块复制模式)编码的RBSB的方法。

[0103] 根据一些实施例,当使用基于子块的时间运动矢量预测模式时,不允许每个CBSB以诸如帧内块复制模式的帧内模式进行编码。这可以通过将帧内块复制编码的RBSB视为帧内模式来实现。特别地,无论如何考虑帧内块复制模式(例如,被视为帧间模式、帧内模式或第三模式),对于CBSB,当其对应的RBSB以帧内块复制模式编码时,该RBSB被认为是基于子块的时间运动矢量预测中的帧内模式。因此,在一些实施例中,根据在基于子块的时间运动矢量预测中的默认设置来处理以帧内块复制模式编码的RBSB。例如,当以帧内块复制模式对CBSB的对应的RBSB进行编码时,诸如零运动矢量的默认运动矢量可以用作CBSB的运动矢量预测值。在该示例中,用于CBSB的参考图片将不再是当前图片,而是时间参考图片。例如,时间参考图片可以是由当前块的所有子块共享的图片、参考图片列表中的第一参考图片、用于TMVP目的的同位图片等。在另一个示例中,对于CBSB,当其对应的RBSB以帧间模式编码但是参考图片是当前图片时,例如零运动矢量的默认运动矢量被分配给该CBSB。就这一点而言,即使RBSB以帧间模式进行编码,由于当前图片与参考图片相同,因此如同RBSB以帧内模式编码来处理该RBSB。

[0104] 图9示出了执行基于子块的时间运动矢量预测的示例。图9示出了当前图片900,当前图片900具有包括当前块900A的九个块。当前块900A被划分为四个子块1至4。当前图片

900可以与参考图片902相关联,参考图片902包括九个先前编码/解码的块。此外,如图9所示,当前块900A具有指向参考块902A的运动矢量904。可以通过使用当前块900A的一个或多个相邻块的运动矢量(例如,空间合并候选)来确定运动矢量904。参考块902A被划分为四个子块1至4。参考块902A的子块1至4分别对应于当前块900A的子块1至4。如果参考图片902与当前图片900相同,则块902A中的每个RBSB被视为如同这些块是以帧内模式编码的。在这方面,例如,如果块902A的子块1以帧间模式编码,但是参考图片902与当前图片900相同,则将块902A的子块1视为如同该子块以帧内模式编码,其中,默认运动矢量被分配给块900A的子块1。

[0105] 如果参考图片902与当前图片900不同,则参考块902A的子块1至4可分别用于针对块900A的子块1至4执行基于子块的时间运动矢量预测。例如,基于是以帧间模式还是帧内模式(例如,帧内块复制模式)对参考块902A的子块1进行编码,来确定当前块900A的子块1的运动矢量预测值。如果参考块902A的子块1以帧间模式编码,则参考块902A的子块1的运动矢量用于确定当前块900A的子块1的运动矢量。然而,如果参考块902A的子块1以帧内模式编码,则当前块900A的子块1的运动矢量被设置为零运动矢量。

[0106] 图10示出了可以由诸如帧内编码器522或帧内解码器672的编码器或解码器执行的处理的实施例。该处理可以开始于步骤S1000,在步骤S1000处,从已编码视频比特流获取当前图片。例如,参照图9,可以从已编码视频比特流中获取当前图片900。处理进行到步骤S1002,在步骤S1002处,对于当前图片中的当前块,识别来自参考图片的参考块。例如,参照图9,可以从与当前块900A相关联的参考图片列表中检索参考图片902。当对当前块900A执行子块时间运动矢量预测时,运动矢量904可以用于识别参考图片902的参考块902A。

[0107] 处理进行到步骤S1004,在步骤S1004处,确定参考图片是否与当前图片相同。如果参考图片与当前图片不同,则处理进行到步骤S1006,在步骤S1006处,确定与CBSB对应的RBSB的编码模式。例如,参照图9,确定参考块902A的RBSB 1的编码模式,RBSB 1与当前块900A的CBSB 1相对应。处理进行到步骤S1008,在步骤S1008处,确定所述RBSB的编码模式是否是帧间模式。如果RBSB的编码模式是帧间模式,则处理进行到步骤S1010,在步骤S1010处,基于RBSB的运动矢量预测值来确定CBSB的运动矢量预测值。例如,如果参考块902A的RBSB 1的编码模式是帧间模式,则基于参考块902A的RBSB 1或运动矢量预测值来确定当前块900A的CBSB 1的运动矢量预测值。例如,参考块902A的RBSB 1的运动矢量预测值被转换(例如,使用诸如时间运动矢量预测中的运动矢量缩放等的方法)并且被用作当前块900A的CBSB 1的运动矢量预测值。

[0108] 返回到步骤S1008,如果RBSB的编码模式不是帧间模式(例如,RBSB的编码模式是帧内模式),则处理进行到步骤S1012,在步骤S1012处,将CBSB的运动矢量预测值设置为默认运动矢量。例如,如果参考块902A的RBSB 1以帧内模式编码,则当前块900A的CBSB 1的运动矢量预测值被设置为例如零运动矢量的默认运动矢量。

[0109] 返回到步骤S1004,如果参考图片与当前图片相同,则处理进行到步骤S1012,在步骤S1012处,将CBSB的运动矢量预测值设置为默认运动矢量。就这一点而言,当参考图片与当前图片相同时,将RBSB的编码模式确定为帧内模式,其中,用于对应的CBSB的运动矢量预测值被设置为默认运动矢量。就这一点而言,即使RBSB是帧间模式,也可以通过将CBSB的运动矢量预测值设置为默认运动矢量来将RBSB视为以帧内模式编码。可以针对当前块900A中

的每个子块重复步骤S1004至S1012。

[0110] 上述技术可以使用计算机可读指令被实现为计算机软件,并且被物理地存储在一个或多个计算机可读介质中。例如,图11示出了适于实现所公开的主题的某些实施例的计算机系统(1100)。

[0111] 计算机软件可以使用任何合适的机器代码或计算机语言来编码,机器代码或计算机语言可以经受汇编、编译、链接等机制以创建包括指令的代码,该指令可以由一个或多个计算机中央处理单元(CPU)、图形处理单元(GPU)等直接执行或者通过解译、微代码执行等来执行。

[0112] 所述指令可以在各种类型的计算机或其部件上执行,包括例如个人计算机、平板计算机、服务器、智能电话、游戏设备、物联网设备等。

[0113] 图11中示出的用于计算机系统(1100)的部件本质上是示例性的,并且不旨在对实现本公开内容的实施例的计算机软件的使用范围或功能提出任何限制。部件的配置也不应当被解释为具有与计算机系统(1100)的示例性实施例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。

[0114] 计算机系统(1100)可以包括某些人机接口输入设备。这样的人机接口输入设备可以响应于由一个或多个人类用户通过例如触觉输入(诸如:击键、滑动、数据手套移动)、音频输入(诸如:语音、拍打)、视觉输入(诸如:姿势)、嗅觉输入(未示出)的输入。人机接口设备还可以用于捕获不一定与人的意识输入直接有关的某些介质,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)。

[0115] 输入人机接口设备可以包括以下中的一个或多个(描述的每个中的仅一个):键盘(1101)、鼠标(1102)、触控板(1103)、触摸屏(1110)、数据手套(未示出)、操纵杆(1105)、麦克风(1106)、扫描仪(1107)、摄像装置(1108)。

[0116] 计算机系统(1100)还可以包括某些人机接口输出设备。这样的人机接口输出设备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人类用户的感受。这样的人机接口输出设备可以包括触觉输出设备(例如,通过触摸屏(1110)、数据手套(未示出)或操纵杆(1105)的触觉反馈,但是也可以有不用作输入设备的触觉反馈设备)、音频输出设备(例如:扬声器(1109)、头戴式耳机(未示出))、视觉输出设备(例如,屏幕(1110),包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕,每个屏幕具有或不具有触摸屏输入能力,每个具有或不具有触觉反馈能力——其中的一些可能能够通过诸如立体图像输出的方式输出二维视觉输出或多于三维输出;虚拟现实眼镜(未描绘)、全息显示器和烟罐(未描绘))和打印机(未描绘)。

[0117] 计算机系统(1100)还可以包括人类可访问存储设备及其相关联的介质,例如,包括具有CD/DVD等介质(1121)的CD/DVD ROM/RW(1120)的光学介质、拇指驱动器(1122)、可移除硬盘驱动器或固态驱动器(1123)、例如磁带和软盘(未描绘)的传统磁性介质、基于专用ROM/ASIC/PLD的设备,例如安全加密狗(未描绘)等。

[0118] 本领域技术人员还应当理解,结合当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、载波或其他瞬时信号。

[0119] 计算机系统(1100)还可以包括到一个或多个通信网络的网络接口(1154)。网络

(1155) 可以是例如无线的、有线的、光学的。网络 (1155) 还可以是本地的、广域的、城市的、车载的和工业的、实时的、延迟容忍的等。网络 (1155) 的示例包括：诸如以太网的局域网，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的车载和工业网络等。某些网络通常需要附接至某些通用数据端口或外围总线 (1149) (诸如，例如计算机系统 (1100) 的USB端口) 的外部网络接口适配器；其它网络通常通过附接至如下所述的系统总线 (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 (1100) 的核中。使用这些网络中的任何网络，计算机系统 (1100) 可以与其它实体进行通信。这样的通信可以是单向的、仅接收的 (例如，广播电视)、单向仅发送的 (例如，到某些CANbus设备的CANbus)、或双向的，例如使用局域或广域数字网络到其它计算机系统)。可以在如上所述的这些网络 (1155) 和网络接口 (1154) 中的每个网络和网络接口上使用特定的协议和协议栈。

[0120] 上述人机接口设备、人类可访问存储设备和网络接口 (1154) 可以附接至计算机系统 (1100) 的核 (1140)。

[0121] 核 (1140) 可以包括一个或更多个中央处理单元 (CPU) (1141)、图形处理单元 (GPU) (1142)、现场可编程门区 (FPGA) (1143) 形式的专用可编程处理单元、用于某些任务的硬件加速器 (1144) 等。这些设备连同只读存储器 (ROM) (1145)、随机存取存储器 (1146)、诸如内部非用户可访问硬盘驱动器、SSD等的内部大容量存储装置 (1147) 一起可以通过系统总线 (1148) 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 (1148)，以使得能够通过另外的CPU、GPU等进行扩展。外围设备可以直接地或通过外围总线 (1149) 附接至核的系统总线 (1148)。外围总线的架构包括PCI、USB等。

[0122] CPU (1141)、GPU (1142)、FPGA (1143) 和加速器 (1144) 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 (1145) 或RAM (1146) 中。过渡数据也可以存储在RAM (1146) 中，而永久数据可以存储在例如内部大容量存储装置 (1147) 中。可以通过使用缓存存储器来实现存储器设备中的任何存储器设备的快速存储和检索，该缓存存储器可以与一个或更多个CPU (1141)、GPU (1142)、大容量存储装置 (1147)、ROM (1145)、RAM (1146) 等紧密相关联。

[0123] 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出于本公开内容的目的而专门设计和构造的介质和计算机代码，或者可以是计算机软件领域的技术人员公知且可用的类型。

[0124] 作为示例而非限制，具有架构 (1100) 的计算机系统，特别是核 (1140)，可以由于 (一个或更多个) 处理器 (包括CPU、GPU、FPGA、加速器等) 执行体现在一个或更多个有形计算机可读介质中的软件的结果而提供功能。这样的计算机可读介质可以是与如以上所介绍的用户可访问的大容量存储装置相关联的介质，以及具有非暂态性的核 (1140) 的某些存储装置，例如，核心内大容量存储装置 (1147) 或ROM (1145)。可以将实现本公开内容的各种实施例的软件存储在这样的设备中并且由核 (1140) 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设备或芯片。软件可以使核 (1140)，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 (1146) 中的数据结构以及根据由软件限定的处理修改这样的数据结构。另外地或替代地，计算机系统可以由于逻辑硬连线或以其他方式体现在电路 (例如，加速器 (1144)) 中而提供

功能,该电路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。在适当的情况下,提及软件可以包含逻辑,反之提及逻辑也可以包含软件。在适当的情况下,提及计算机可读介质可以包含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上述两者。本公开内容包含硬件和软件的任何合适的组合。

- [0125] 附录A:首字母缩写
- [0126] 运动矢量 (MotionVector,MV)
- [0127] 高效视频编码 (High Efficiency Video Coding,HEVC)
- [0128] 辅助增强信息 (Supplementary Enhancement Information,SEI)
- [0129] 视频可用性信息 (Video Usability Information,VUI)
- [0130] 图片群组 (group of pictures,GOP)
- [0131] 变换单元 (Transform Unit,TU)
- [0132] 预测单元 (Prediction Unit,PU)
- [0133] 编码树单位 (coding tree unit,CTU)
- [0134] 编码树块 (coding tree block,CTB)
- [0135] 预测块 (Prediction Block,PB)
- [0136] 假设参考解码器 (Hypothetical Reference Decoder,HRD)
- [0137] 信噪比 ((Signal-noise ratio,SNR)
- [0138] 中央处理单元 (Central Processing Unit,CPU)
- [0139] 图形处理单元 (Graphics Processing Unit,GPU)
- [0140] 阴极射线管 (Cathode Ray Tube,CRT)
- [0141] 液晶显示器 (Liquid-Crystal Display,LCD)
- [0142] 有机发光二极管 (Organic Light-Emitting Diode,OLED)
- [0143] 致密盘 (Compact Disc,CD)
- [0144] 数字视频光盘 (Digital Video Disc,DVD)
- [0145] 只读存储器 (Read-Only Memory,ROM)
- [0146] 随机存取存储器 (Random Access Memory,RAM)
- [0147] 专用集成电路 (Application-Specific Integrated Circuit,ASIC)
- [0148] 可编程逻辑设备 (Programmable Logic Device,PLD)
- [0149] 局域网 (Local Area Network,LAN)
- [0150] 全球移动通信系统 (Global System for Mobile communications,GSM)
- [0151] 长期演进 (Long-Term Evolution,LTE)
- [0152] 控制器局域网总线 (Controller Area Network Bus,CANBus)
- [0153] 通用串行总线 (Universal Serial Bus,USB)
- [0154] 外围部件互连 (Peripheral Component Interconnect,PCI)
- [0155] 现场可编程门区 (Field Programmable Gate Areas,FPGA)
- [0156] 固态驱动器 (solid-state drive,SSD)
- [0157] 集成电路 (Integrated Circuit,IC)
- [0158] 编码单元 (Coding Unit,CU)

[0159] 尽管本公开内容已经描述了若干示例性实施例,但是存在落入本公开内容的范围内的改变、置换和各种替换等效物。因此将认识到,本领域技术人员能够设想虽然本文中并没有明确示出或描述但是体现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。

[0160] 本公开提供一种用于解码器的视频解码的方法,包括:从已编码视频比特流中获取当前图片;对于所述当前图片中包括的当前块,识别与所述当前图片不同的参考图片中包括的参考块,所述当前块被划分为多个子块(CBSB),所述参考块具有多个子块(RBSB),多个RBSB分别对应于所述多个CBSB中的不同CBSB;确定所述RBSB的参考图片是否是所述当前图片;响应于确定所述RBSB的参考图片是所述当前图片,将所述RBSB的编码模式确定为帧内模式;以及响应于确定所述RBSB的参考图片不是所述当前图片,(i)对于所述CBSB中的一个CBSB,确定所述RBSB的编码模式是否是帧内模式和帧间模式之一,以及(ii)基于对应的RBSB的编码模式是否是所述帧内模式和所述帧间模式之一来确定所述CBSB中的所述一个CBSB的运动矢量预测值。

[0161] 在一种可能的实现方式中,若确定所述RBSB的编码模式是帧内模式,所述方法还包括,响应于确定所述对应的RBSB的编码模式是所述帧内模式,将所述一个CBSB的运动矢量预测值设置为默认运动矢量,默认运动矢量使参考图片改变为时间上的参考图片。

[0162] 在一种可能的实现方式中,所述默认运动矢量是下述中的一个:(i)零运动矢量;以及(ii)所述CBSB与所述RBSB之间的偏移。

[0163] 在一种可能的实现方式中,响应于确定所述对应的RBSB的编码模式是所述帧间模式,所述一个CBSB的运动矢量预测值是基于与所述对应的RBSB相关联的运动矢量预测值得到的。

[0164] 在一种可能的实现方式中,所述一个CBSB的运动矢量预测值是与所述对应的RBSB相关联的所述运动矢量预测值的缩放版本。

[0165] 在一种可能的实现方式中,所述方法还包括:响应于确定所述RBSB的所述参考图片是所述当前图片且所述RBSB的编码模式为帧间模式,将所述一个CBSB的运动矢量预测值设置为默认运动矢量以将所述RBSB的编码模式确定为帧内模式,默认运动矢量使参考图片改变为时间上的参考图片。

[0166] 在一种可能的实现方式中,所述方法还包括:通过基于每个CBSB的运动矢量预测值对每个CBSB执行基于子块的时间运动矢量预测,对每个CBSB进行解码。

[0167] 在一种可能的实现方式中,根据与相邻于所述当前块的块相关联的运动矢量预测值来识别所述参考块。

[0168] 在一种可能的实现方式中,所述参考图片是来自参考图片序列中的第一参考图片,所述参考图片序列与所述当前图片相关联。

[0169] 本公开内容的示例性实施例包括一种用于视频解码的装置。所述装置包括获取单元、识别单元、第一确定单元和第二确定单元。所述获取单元,用于从已编码视频比特流中获取当前图片;所述识别单元,用于对于所述当前图片中包括的当前块,识别与所述当前图片不同的参考图片中包括的参考块,所述当前块被划分为多个子块(CBSB),所述参考块具有多个子块(RBSB),所述多个RBSB分别对应于所述多个CBSB中的不同CBSB;所述第一确定单元,用于确定所述RBSB的所述参考图片是否是所述当前图片;所述第二确定单元,用于响

应于确定所述RBSB的所述参考图片是所述当前图片,将所述RBSB的编码模式确定为帧内模式;响应于确定所述RBSB的所述参考图片不是所述当前图片,(i)对于所述CBSB中的一个CBSB,确定所述RBSB的编码模式是否是帧内模式和帧间模式之一,(ii)基于对应的RBSB的编码模式是否是所述帧内模式和所述帧间模式之一来确定所述CBSB中的所述一个CBSB的运动矢量预测值。

[0170] 本公开内容提供一种存储有指令的非暂态计算机可读介质,所述指令在由视频解码器中的处理器执行时使所述处理器执行前述提供的任一项用于解码器的视频解码的方法。

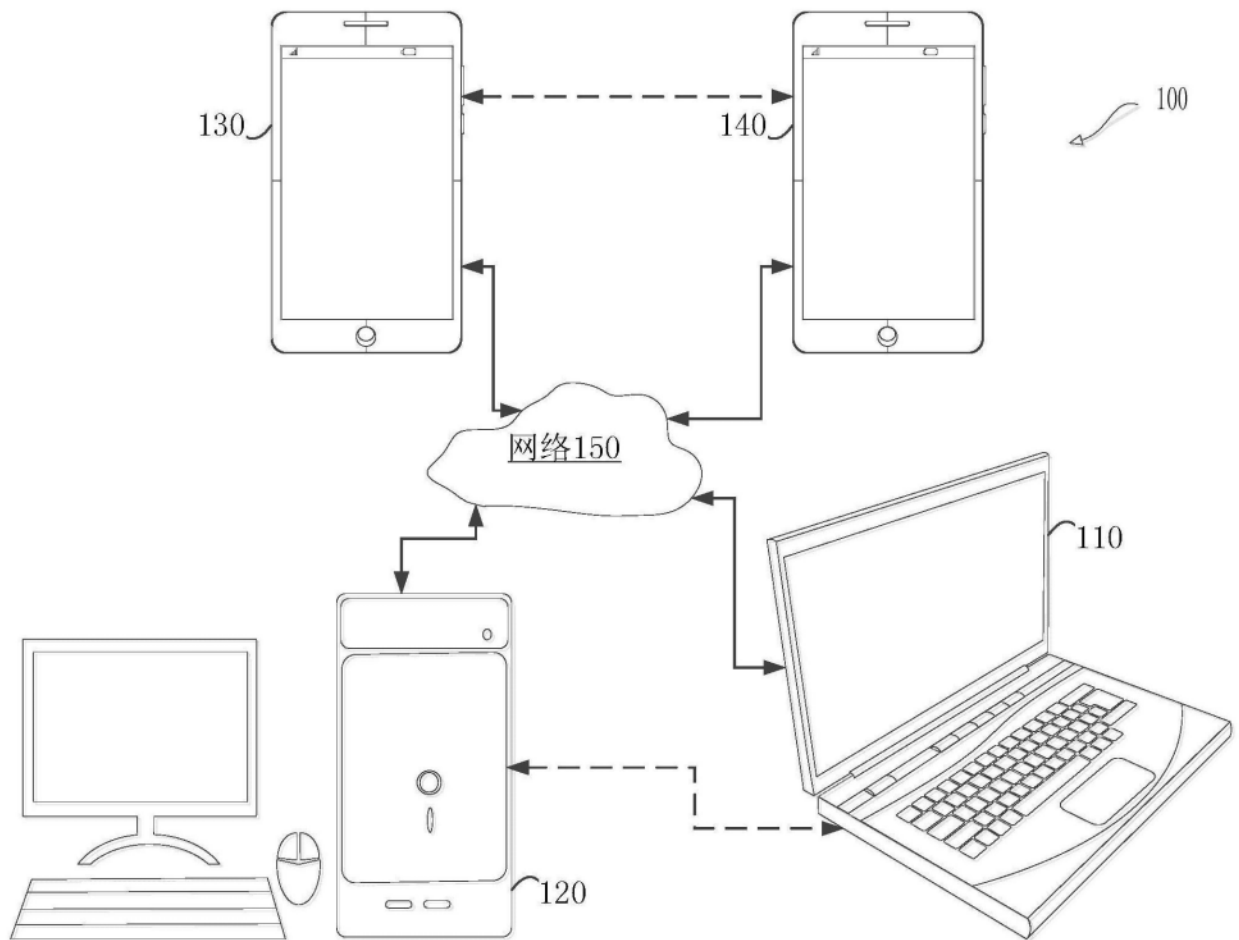


图1

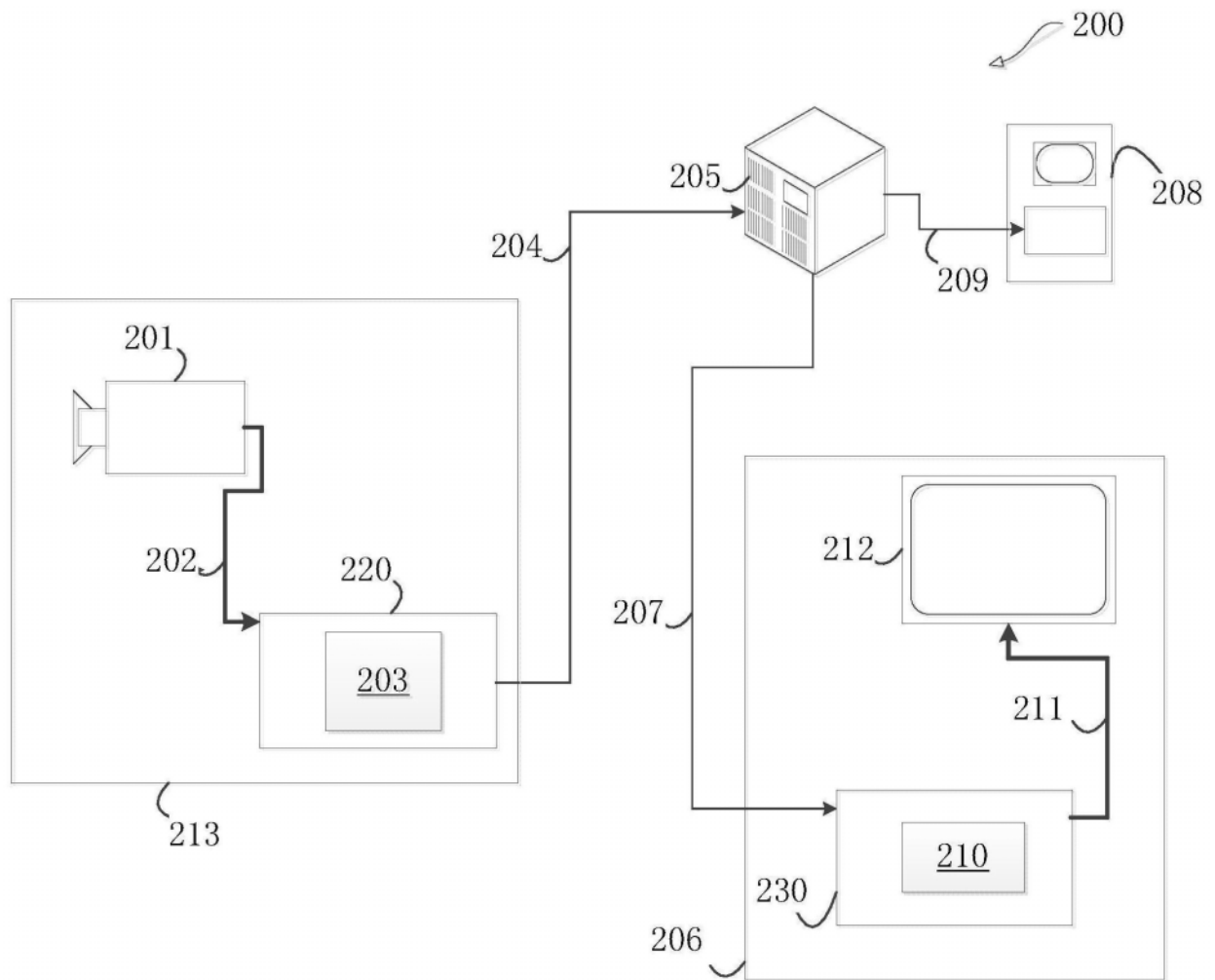


图2

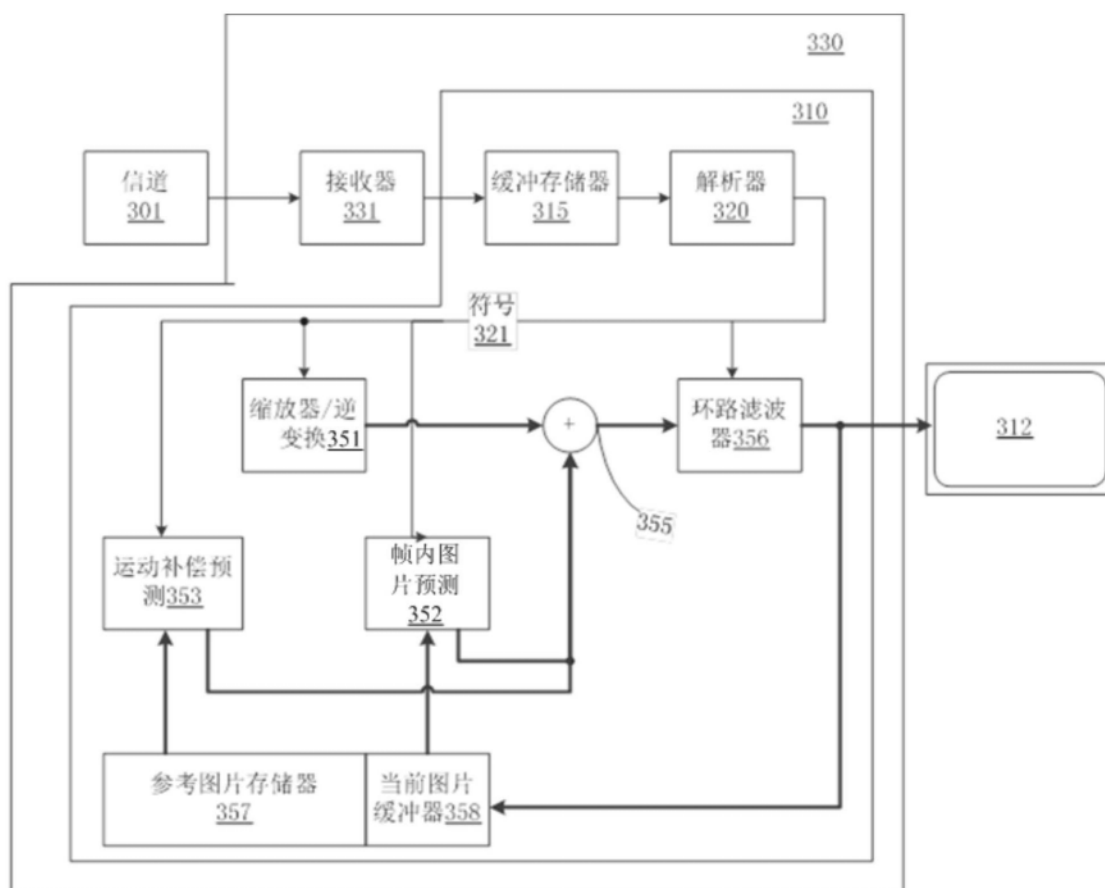


图3

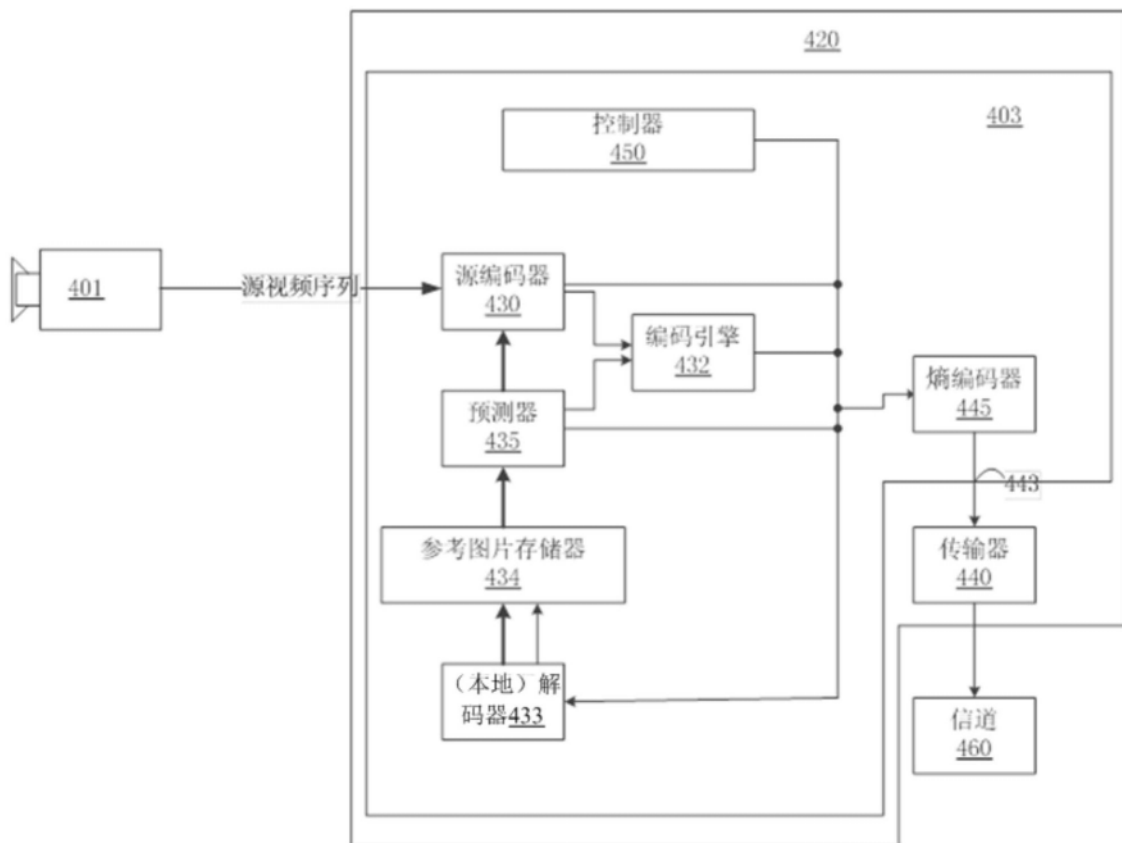


图4

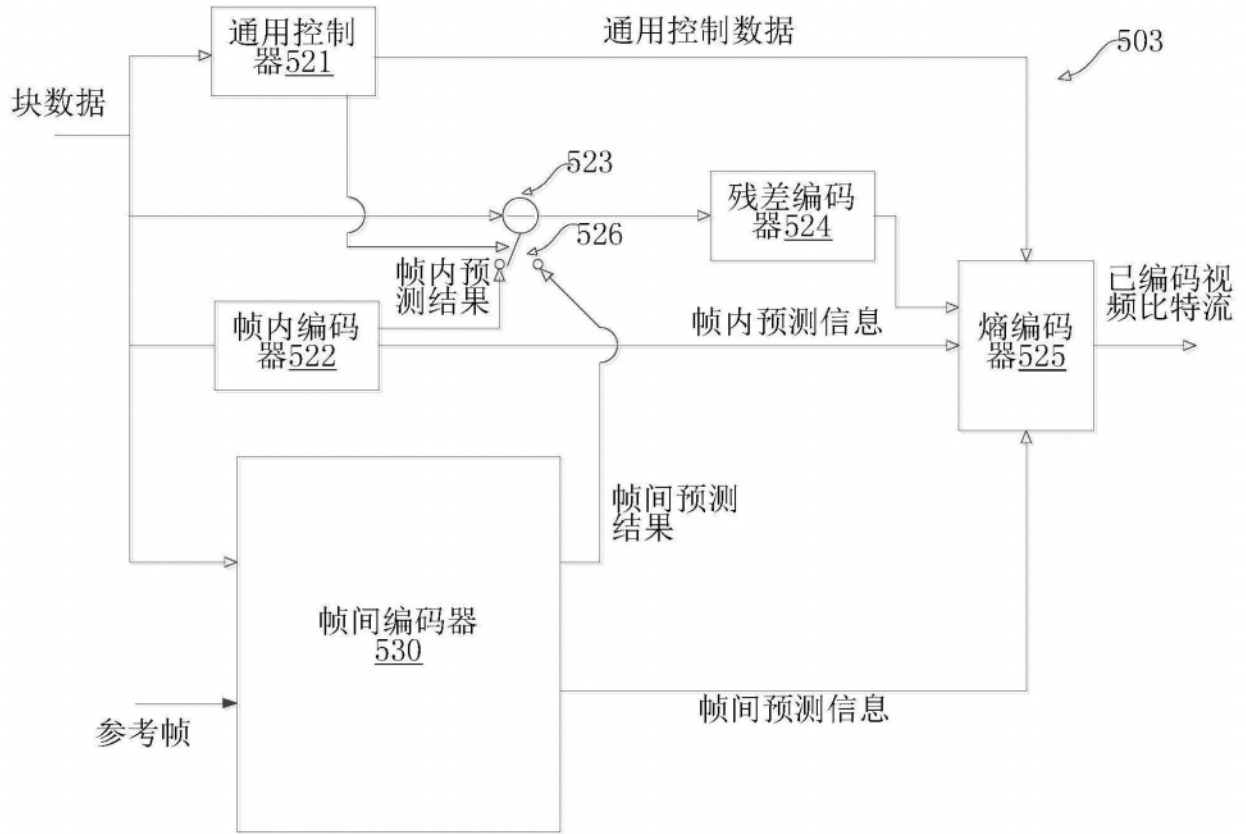


图5

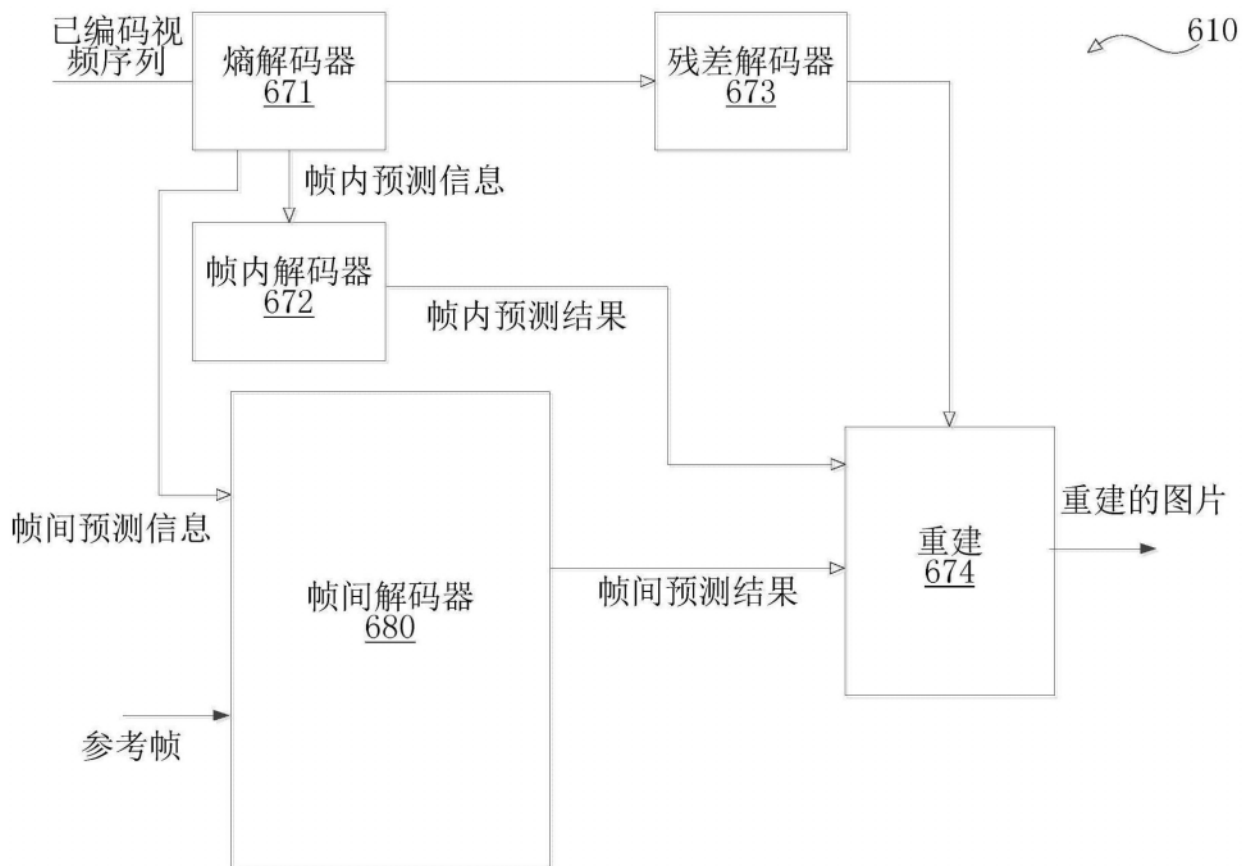


图6

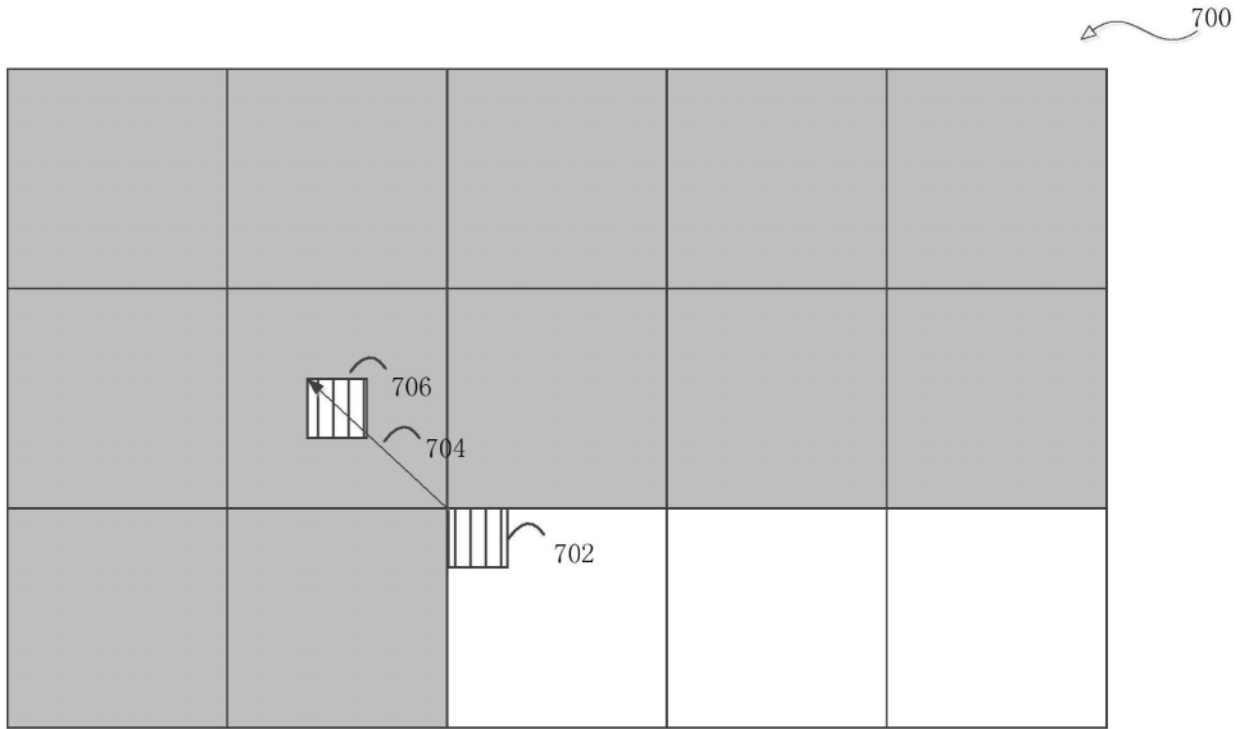


图7

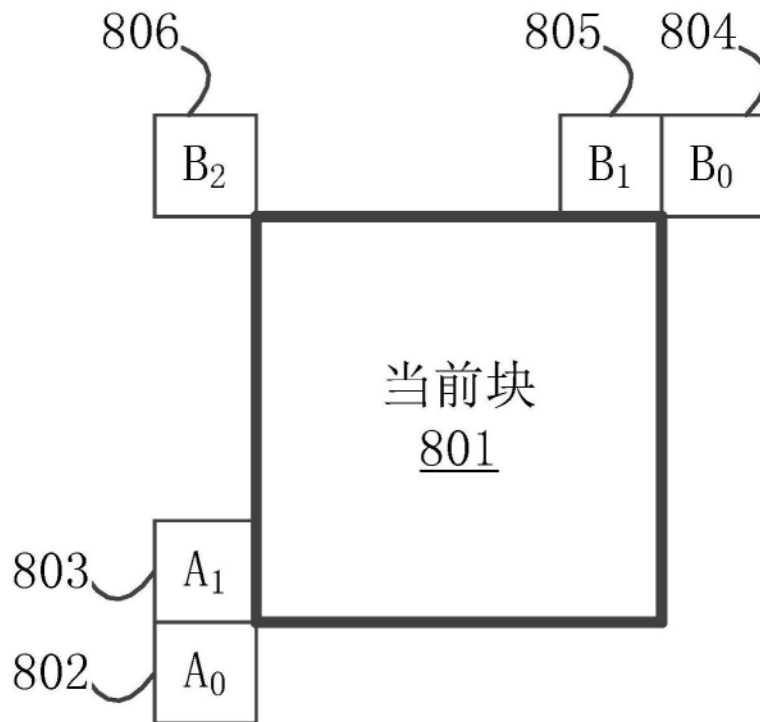


图8

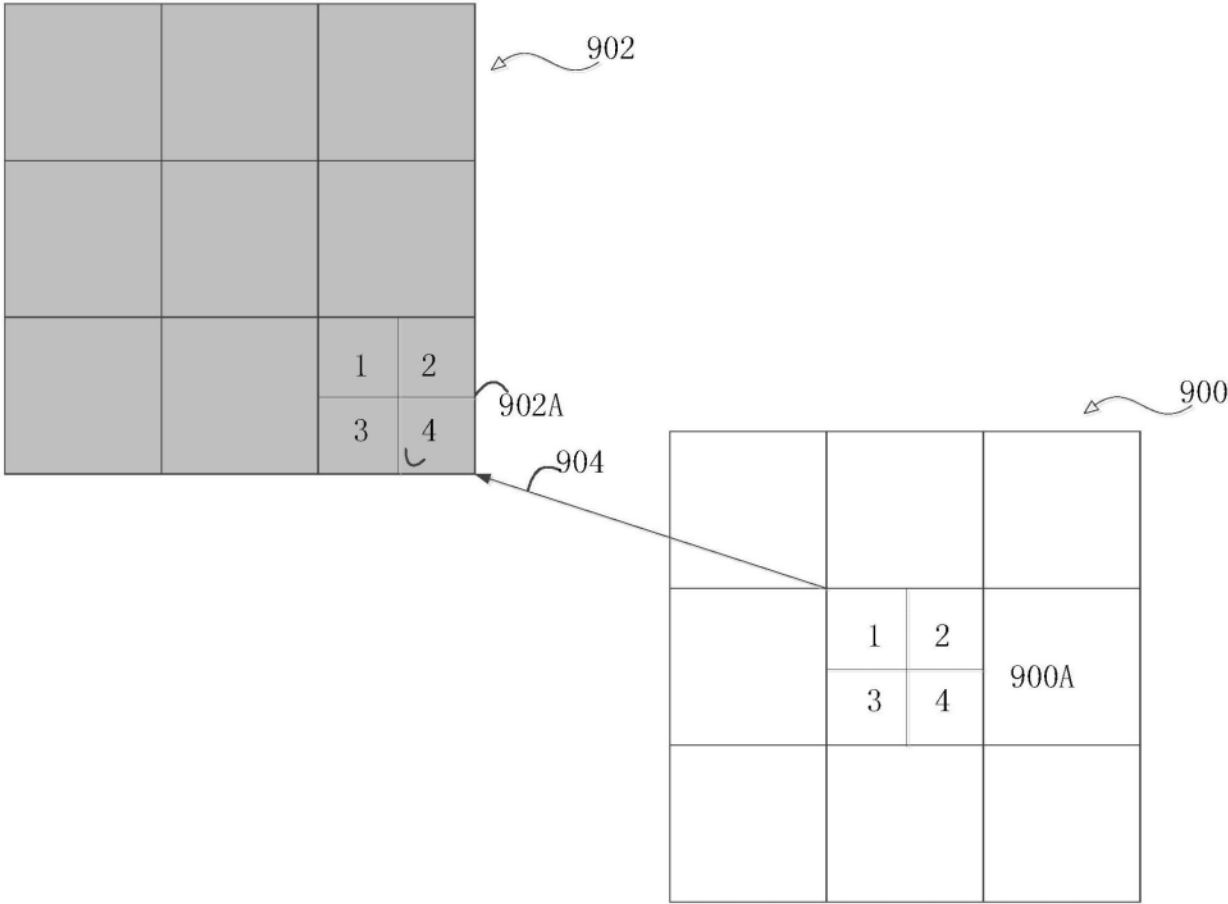


图9

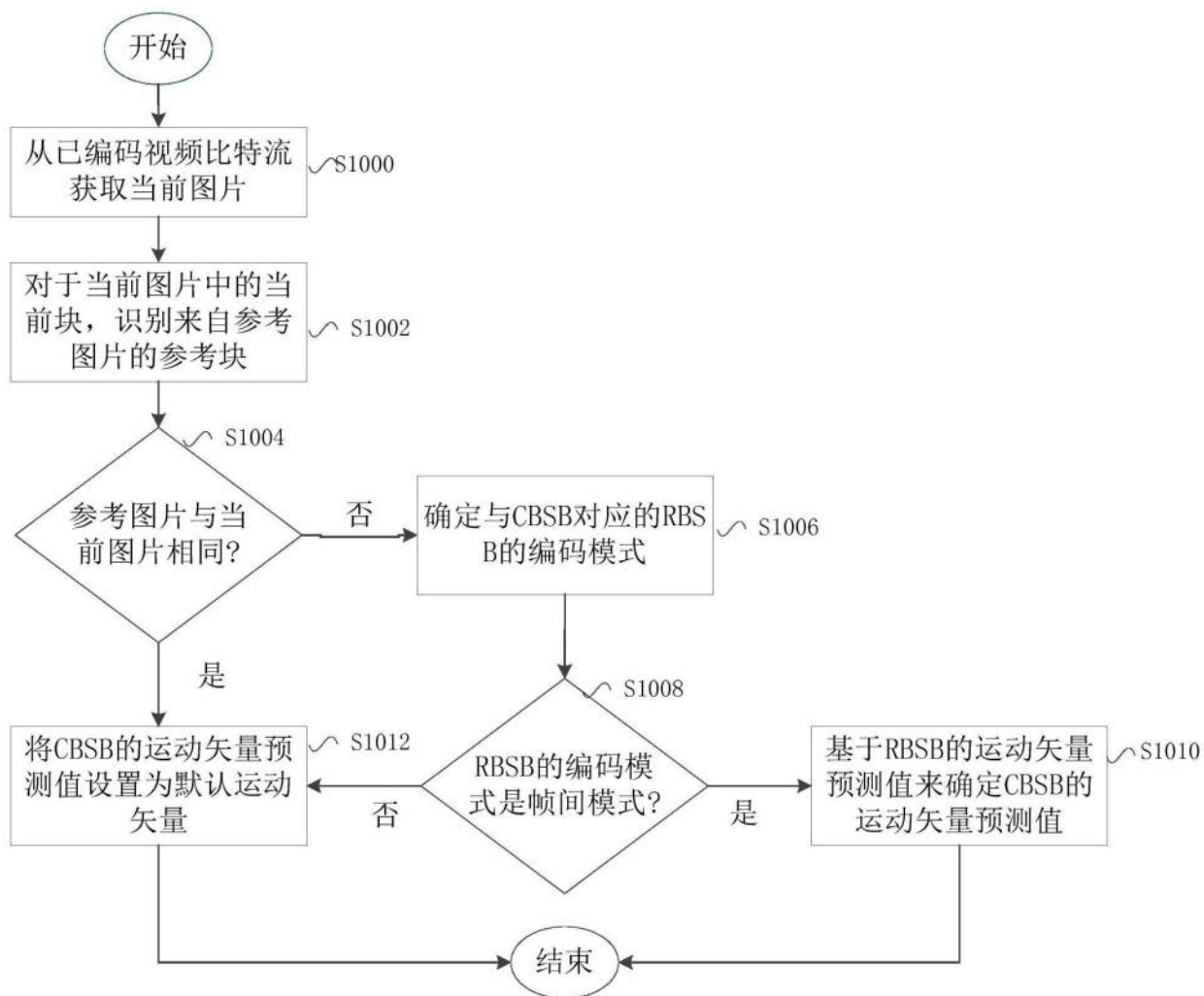


图10

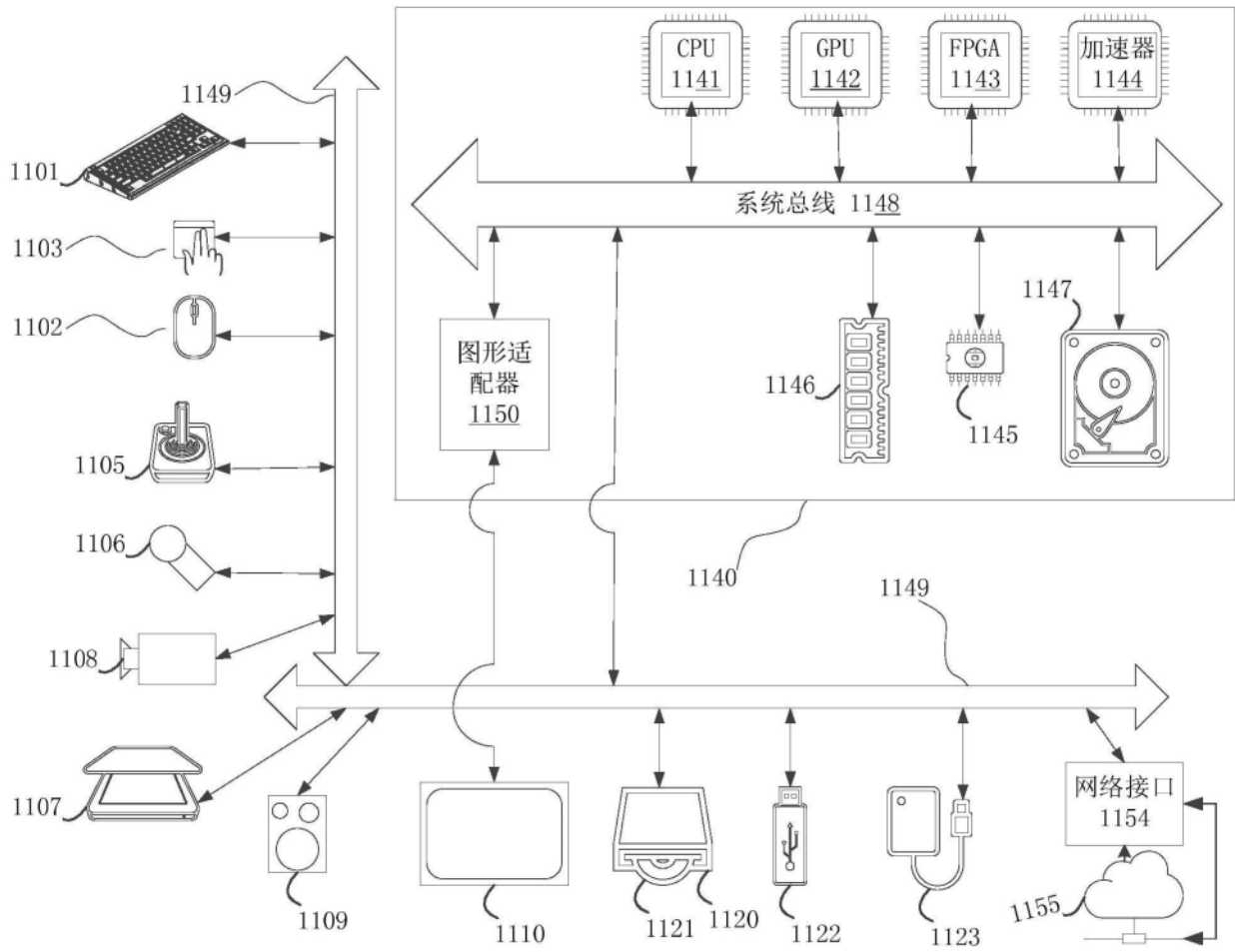


图11