



(12) 发明专利申请

(10) 申请公布号 CN 117413513 A

(43) 申请公布日 2024. 01. 16

(21) 申请号 202280020308.5

(22) 申请日 2022.11.08

(30) 优先权数据

63/342,441 2022.05.16 US

17/982,139 2022.11.07 US

(85) PCT国际申请进入国家阶段日

2023.09.08

(86) PCT国际申请的申请数据

PCT/US2022/049243 2022.11.08

(87) PCT国际申请的公布数据

W02023/224655 EN 2023.11.23

(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 赵亮 高晗 赵欣 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 万欣慰

(51) Int.Cl.

H04N 19/103 (2006.01)

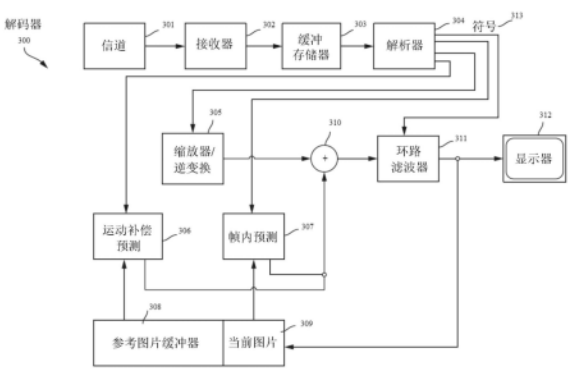
权利要求书3页 说明书22页 附图15页

(54) 发明名称

用于联合MVD编码的灵活缩放因子

(57) 摘要

一种用于视频编码的方法包括：获得视频比特流的编码块；确定是否使用运动矢量差的联合编码(JMVD)来预测编码块；基于确定选择使用JMVD来预测编码块，获得缩放因子；以及基于将缩放因子应用于JMVD的沿一个或多个预定义方向的一个或多个分量来导出针对一个或多个参考帧列表的运动矢量差(MVD)，以及至少基于所导出的MVD来重建编码块。



1. 一种由至少一个处理器执行的用于视频编码的方法,所述方法包括:
获取视频比特流的编码块;
确定是否使用运动矢量差的联合编码(JMVD)来预测所述编码块;
基于确定使用所述JMVD来预测所述编码块,从所述视频比特流获得多个缩放因子和所述JMVD;
基于将所述多个缩放因子应用于所述联合运动矢量差的沿一个或更多个预定义方向的一个或更多个分量来导出针对至少一个参考帧列表的运动矢量差(MVD);以及
至少基于所导出的MVD重建所述编码块。
2. 根据权利要求1所述的用于视频编码的方法,
其中,还基于参考帧与当前帧之间的距离中的任何距离来导出所述MVD。
3. 根据权利要求2所述的用于视频编码的方法,
其中,导出所述MVD包括确定标志是否指示所述缩放因子中的至少一个缩放因子不等于第一预定义默认值,
其中,所述缩放因子中的所述至少一个缩放因子用于针对所述参考帧中的一个参考帧从所述JMVD导出MVD,以及
其中,用于针对所述参考帧中的第二参考帧从所述JMVD导出MVD的所述缩放因子中的另一个缩放因子被设置为第二预定义默认值。
4. 根据权利要求1所述的用于视频编码的方法,
其中,获得所述缩放因子基于获得在所述视频比特流中用信号发送的至少一个标志,以及
所述至少一个标志指示沿所述一个或更多个预定义方向的所述分量中的至少一个分量的所述缩放因子。
5. 根据权利要求1所述的用于视频编码的方法,
其中,导出所述MVD包括:
确定第一标志是否指示所述缩放因子中的至少一个缩放因子不等于第一预定义默认值,以及
基于确定所述第一标志指示所述缩放因子中的所述至少一个缩放因子不等于所述第一预定义默认值,基于第二标志的值确定是否将所述缩放因子应用于所述MVD的所述预定义方向中的至少一个方向。
6. 根据权利要求5所述的用于视频编码的方法,
其中,导出所述MVD包括:基于确定所述第二标志指示两个所述预定义方向,将所述一个或更多个缩放因子等同地应用于两个所述预定义方向。
7. 根据权利要求1所述的用于视频编码的方法,
其中,获得所述缩放因子包括:在查找表中获得所述缩放因子的索引,
其中,所述查找表指示所述缩放因子中的在所述索引中的第一索引处的至少一对缩放因子在两个所述预定义方向上具有相同缩放因子值,并且
其中,所述查找表指示所述缩放因子中的在所述索引中的第二索引处的至少第二对缩放因子在多个所述预定义方向之间具有不同缩放因子值。
8. 根据权利要求7所述的用于视频编码的方法,

其中,所述相同缩放因子值和所述不同缩放因子值中的至少一者是分数缩放因子值,以及

其中,所述相同缩放因子值和所述不同缩放因子值中的至少另一者是 m/M ,其中 M 是 n 的2次方,并且 m 和 n 是整数。

9.根据权利要求1所述的用于视频编码的方法,

其中,基于量化步长、量化参数、块大小、当前块的运动矢量预测块之间的差、MVD类别、参考图片和与所述编码块相邻的相邻块的MVD缩放因子中的至少一者的编码信息来导出所述缩放因子。

10.根据权利要求1所述的用于视频编码的方法,

其中,帧标头、切片标头和序列标头中的至少一者指示是否在所述视频比特流中针对所述缩放因子来发送信号。

11.一种用于视频编码的设备,所述设备包括:

至少一个存储器,其被配置成存储计算机程序代码;

至少一个处理器,其被配置成访问所述计算机程序代码并按照所述计算机程序代码所指示进行操作,所述计算机程序代码包括:

获得代码,其被配置成使所述至少一个处理器获得视频比特流的编码块;

确定代码,其被配置成使所述至少一个处理器确定是否使用运动矢量差的联合编码(JMVD)来预测所述编码块;

另外的获得代码,其被配置成基于确定使用所述JMVD来预测所述编码块,使所述至少一个处理器从所述视频比特流获得多个缩放因子和所述JMVD;

导出代码,其被配置成使所述至少一个处理器基于将所述多个缩放因子应用于所述JMVD的沿一个或更多个预定义方向的一个或更多个分量来导出针对至少一个参考帧列表的运动矢量差(MVD);以及

重建代码,其被配置成使所述至少一个硬件处理器至少基于所导出的MVD重建所述编码块。

12.根据权利要求11所述的用于视频编码的设备,

其中,还基于参考帧与当前帧之间的距离中的任何距离来导出所述MVD。

13.根据权利要求12所述的用于视频编码的设备,

其中,导出所述MVD包括确定标志是否指示所述缩放因子中的至少一个缩放因子不等于第一预定义默认值,

其中,所述缩放因子中的所述至少一个缩放因子用于针对所述参考帧中的一个参考帧从所述JMVD导出MVD,以及

其中,用于针对所述参考帧中的第二参考帧从所述JMVD导出MVD的所述缩放因子中的另一个缩放因子被设置为第二预定义默认值。

14.根据权利要求11所述的用于视频编码的设备,

其中,获得所述缩放因子基于获得在所述编码块的比特流中用信号发送的至少一个标志,以及

所述至少一个标志指示沿所述一个或更多个预定义方向的所述分量中的至少一个分量的所述缩放因子。

15. 根据权利要求11所述的用于视频编码的设备,

其中, 导出所述MVD包括:

确定第一标志是否指示所述缩放因子中的至少一个缩放因子不等于第一预定义默认值, 以及

基于确定所述第一标志指示所述缩放因子中的所述至少一个缩放因子不等于所述第一预定义默认值, 基于第二标志的值确定是否将所述缩放因子应用于所述MVD的所述预定义方向中的至少一个轴。

16. 根据权利要求15所述的用于视频编码的设备,

其中, 导出所述MVD包括: 基于确定所述第二标志指示两个所述预定义方向, 将所述一个或更多个缩放因子等同地应用于两个所述预定义方向。

17. 根据权利要求12所述的用于视频编码的设备,

其中, 获得所述缩放因子包括: 在查找表中获得所述缩放因子的索引,

其中, 所述查找表指示所述缩放因子中的在所述索引中的第一索引处的至少一对缩放因子在两个所述预定义方向上具有相同缩放因子值, 并且

其中, 所述查找表指示所述缩放因子中的在所述索引中的第二索引处的至少第二对缩放因子在多个所述预定义方向之间具有不同缩放因子值。

18. 根据权利要求17所述的用于视频编码的设备,

其中, 所述相同缩放因子值和所述不同缩放因子值中的至少一者是分数缩放因子值, 以及

其中, 所述相同缩放因子值和所述不同缩放因子值中的至少另一者是 m/M , 其中 M 是 n 的 2 次方, 并且 m 和 n 是整数。

19. 根据权利要求11所述的用于视频编码的设备,

其中, 基于量化步长、量化参数、块大小、当前块的运动矢量预测块之间的差、MVD类别、参考图片和与所述编码块相邻的相邻块的MVD缩放因子中的至少一者的编码信息来导出所述缩放因子。

20. 一种存储计算机程序代码的非暂态计算机可读介质, 所述计算机程序代码在由计算机执行时, 使所述计算机至少执行以下操作:

获取视频比特流的编码块;

确定是否使用运动矢量差的联合编码(JMVD)来预测所述编码块;

基于确定使用所述JMVD来预测所述编码块, 从所述视频比特流获得多个缩放因子和所述JMVD;

基于将所述多个缩放因子应用于所述JMVD的沿一个或更多个预定义方向的一个或更多个分量来导出针对至少一个参考帧列表的运动矢量差(MVD); 以及

至少基于所导出的MVD重建所述编码块。

用于联合MVD编码的灵活缩放因子

[0001] 相关申请的交叉引用

[0002] 本申请基于2022年5月16日提交的美国临时申请第63/342,441号和2022年11月7日提交的美国申请第17/982,139号,并要求上述美国申请的优先权,上述美国申请在此通过引用将其全部内容明确地并入本申请中。

背景技术

1. 技术领域

[0003] 本公开内容涉及一组高级图像和视频编码技术,更具体地,涉及用于运动矢量差的联合编码(joint coding of motion vector difference, JMVD)的改进方案。

[0004] 2. 背景技术

[0005] 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。AOMedia Video 1是由开放媒体联盟(AOMedia,于2015年成立的联盟,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web浏览器供应商)作为VP9的继承者开发的。AV1项目的许多组成部分源于联盟成员先前的研究工作。个体贡献者在若干年前就启动了实验性技术平台:2010年发布了Xiph/Mozilla的Daala代码、2014年9月12日公布了Google的实验性VP9演进项目VP10、以及2015年8月11日发布了Cisco的Thor。在VP9代码库的基础上,AV1结合了其他技术,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个版本0.1.0于2016年4月7日发布。该联盟于2018年3月28日宣布发布AV1比特流规范,以及基于软件的参考编码器和解码器。于2018年6月25日发布了该规范的验证版本1.0.0。于2019年1月8日发布了该规范的具有勘误表1的验证版本1.0.0。AV1比特流规范包括参考视频编解码器。

[0006] ITU-T VCEG (Q6/16) 和ISO/IEC MPEG (JTC 1/SC 29/WG 11) 在2013年(版本1)、2014年(版本2)、2015年(版本3)和2016年(版本4)发布了H.265/HEVC (High Efficiency Video Coding, 高效视频编码) 标准。从那时起,他们一直在研究未来视频编码技术标准化的潜在需求,这种技术在压缩能力上可以明显优于HEVC。2017年10月,他们发布了具有超越HEVC的能力的视频压缩的联合建议书(CfP)。到2018年2月15日,分别提交了关于标准动态范围(standard dynamic range, SDR)的总共22个CfP响应、关于高动态范围(high dynamic range, HDR)的12个CfP响应以及关于360个视频类别的12个CfP响应。在2018年4月,在122MPEG/第10次JVET (Joint Video Exploration Team-Joint Video Expert Team, 联合视频探索组-联合视频专家组) 会议中评估了所有接收到的CfP响应。通过仔细的评估, JVET正式地发起了超越HEVC的下一代视频编码的标准化,即所谓的通用视频编码(Versatile Video Coding, VVC)。

[0007] 并且对于JMVD,在JMVD编码模式下假设线性运动存在技术问题,因为两个参考帧之间的运动可能并非总是线性运动,并且例如,运动从后向参考帧到前向参考帧可能变得更慢或更快。因此,需要解决这些问题的技术方案。

发明内容

[0008] 根据一些实施方式的一方面,提供了一种由至少一个处理器执行的用于视频编码的方法。该方法包括:获得视频数据的编码块;确定是否使用运动矢量差的联合编码(JMVD)来预测编码块;基于确定使用JMVD来预测编码块,获得缩放因子;以及基于将缩放因子应用于JMVD的沿一个或多个预定义方向的一个或更多个分量来导出针对至少一个参考帧列表的运动矢量差(motion vector difference,MVD),以及至少基于所导出的MVD来重建编码块。

[0009] 根据一些实施方式的其他方面,还提供了与该方法一致的设备 and 计算机可读介质。

附图说明

[0010] 所公开的主题的其他特征、性质和各种优点将根据以下详细描述和附图更加明显,在附图中:

- [0011] 图1是根据一些实施方式的图的简化图示;
- [0012] 图2是根据一些实施方式的图的简化图示;
- [0013] 图3是根据一些实施方式的图的简化图示;
- [0014] 图4是根据一些实施方式的图的简化图示;
- [0015] 图5是根据一些实施方式的图的简化图示;
- [0016] 图6是根据一些实施方式的图的简化图示;
- [0017] 图7是根据一些实施方式的图的简化图示;
- [0018] 图8是根据一些实施方式的图的简化图示;
- [0019] 图9A是根据一些实施方式的图的简化图示;
- [0020] 图9B是根据一些实施方式的图的简化图示;
- [0021] 图10A是根据一些实施方式的图的简化图示;
- [0022] 图10B是根据一些实施方式的图的简化图示;
- [0023] 图10C是根据一些实施方式的图的简化图示;
- [0024] 图11是根据一些实施方式的图的简化图示;
- [0025] 图12是根据一些实施方式的图的简化图示;
- [0026] 图13是根据一些实施方式的简化流程图;以及
- [0027] 图14是根据一些实施方式的图的示意性图示。

具体实施方式

[0028] 可以单独使用或以任何顺序组合使用下面讨论的所提出的特征。此外,可以通过处理电路系统(例如,一个或更多个处理器或一个或更多个集成电路)来实现实施方式。在一个示例中,一个或更多个处理器执行被存储在非暂态计算机可读介质中的程序。

[0029] 根据一些实施方式,存在被配置成存储计算机程序代码的至少一个存储器,以及被配置成访问计算机程序代码并按照计算机程序代码的指示操作的至少一个处理器,该计算机程序代码包括:获得代码,其被配置成使至少一个处理器获得视频数据的编码块;确定代码,其被配置成使至少一个处理器确定是否使用运动矢量差的联合编码(JMVD)来预测编

码块;另外的获得代码,其被配置成使至少一个处理器响应于确定使用JMVD来预测编码块来获得缩放因子;导出代码,其被配置成使至少一个处理器基于将缩放因子应用于JMVD的沿一个或更多个预定义方向的一个或更多个分量来导出针对至少一个参考帧列表的运动矢量差(motion vector difference,MVD);以及重建代码,其被配置成使至少一个硬件处理器至少基于所导出的MVD重建编码块。

[0030] 根据一些示例性实施方式,还基于参考帧与当前帧之间的距离中的任何距离来导出MVD。

[0031] 根据一些示例性实施方式,导出MVD包括:确定标志是否指示缩放因子中的至少一个缩放因子不等于第一预定义默认值,缩放因子中的至少一个缩放因子用于针对参考帧中的一个参考帧从JMVD导出MVD,并且用于针对参考帧中的第二参考帧从JMVD导出MVD的缩放因子中的另一个缩放因子被设置为第二预定义默认值。

[0032] 根据一些示例性实施方式,获得缩放因子基于获得在编码块的比特流中用信号发送的至少一个标志,并且该至少一个标志指示沿一个或更多个预定义方向的分量中的至少一个分量的缩放因子。

[0033] 根据一些示例性实施方式,导出MVD包括:确定第一标志是否指示缩放因子中的至少一个缩放因子不等于第一预定义默认值,以及响应于确定第一标志指示缩放因子中的至少一个缩放因子不等于第一预定义默认值,基于第二标志的值确定是否将缩放因子应用于MVD的预定义方向中的至少一个方向。

[0034] 根据一些示例性实施方式,导出MVD包括:基于确定第二标志指示两个预定义方向,将一个或更多个缩放因子等同地应用于两个预定义方向。

[0035] 根据一些示例性实施方式,获得缩放因子包括:在查找表中获得缩放因子的索引,该查找表指示缩放因子中的在索引中的第一索引处的至少一对缩放因子在两个预定义方向上具有相同缩放因子值,并且该查找表指示缩放因子中的在索引中的第二索引处的至少第二对缩放因子在多个预定义方向之间具有不同的缩放因子值。

[0036] 根据一些示例性实施方式,相同缩放因子值和不同缩放因子值中的至少一者是分数缩放因子值,并且相同缩放因子值和不同缩放因子值中的至少另一者是 m/M ,其中 M 是 n 的2次方,并且 m 和 n 是整数。

[0037] 根据一些示例性实施方式,基于量化步长、量化参数、块大小、当前块的运动矢量预测块之间的差、MVD类别、参考图片和与编码块相邻的相邻块的MVD缩放因子中的至少一者的编码信息来导出缩放因子。

[0038] 根据一些示例性实施方式,帧标头、切片标头和序列标头中的至少一者指示是否在编码块的比特流中针对缩放因子发送信号。

[0039] 图1示出了根据本公开内容的实施方式的通信系统100的简化框图。通信系统100可以包括经由网络105互连的至少两个终端102和103。为了数据的单向传输,第一终端103可以在本地位置处对视频数据进行编码,以经由网络105传输至另一终端102。第二终端102可以从网络105接收另一终端的经编码的视频数据,对经编码的数据进行解码并且显示恢复的视频数据。单向数据传输可能在媒体服务应用等中是常见的。

[0040] 图1示出了第二对终端101和104,第二对终端101和104被提供用于支持例如在视频会议期间可能发生的经编码的视频的双向传输。对于数据的双向传输,每个终端101和

104可以对在本地位置处捕获的视频数据进行编码,以经由网络105传输至另一终端。每个终端101和104还可以接收由另一终端传输的经编码的视频数据,可以对经编码的数据进行解码,并且可以在本地显示装置上显示恢复的视频数据。

[0041] 在图1中,终端101、102、103和104可以被示出为服务器、个人计算机和智能电话,但本公开内容的原理不限于此。本公开内容的实施方式适用于膝上型计算机、平板计算机、媒体播放器和/或专用视频会议装备的应用。网络105表示在终端101、102、103和104之间传送经编码的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105可以在电路交换信道和/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。出于本讨论的目的,除非在下文中有所说明,否则网络105的架构和拓扑对于本公开内容的操作而言可以是无关紧要的。

[0042] 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,作为针对所公开的主题的应用的示例。所公开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TV、在包括CD、DVD、存储棒等的数字介质上存储压缩视频等等。

[0043] 流式传输系统可以包括捕获子系统203,捕获子系统203可以包括创建例如未压缩的视频样本流213的视频源201,例如数字摄像装置。样本流213可以被强调为与经编码的视频比特流相比的较高的数据量,并且可以由耦接至摄像装置201的编码器202进行处理。编码器202可以包括硬件、软件或其组合以实现或实施如下文更详细地描述的所公开主题的各项。经编码的视频比特流204可以被强调为与样本流相比的较低的数据量,经编码的视频比特流204可以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212和207可以访问流式传输服务器205以检索经编码的视频比特流204的副本208和副本206。客户端212可以包括视频解码器211,视频解码器211对经编码的视频比特流208的传入副本进行解码,并且创建可以在显示器209或另一呈现装置(未描绘)上呈现的传出视频样本流210。在一些流式传输系统中,可以根据某些视频编码/压缩标准来对视频比特流204、206和208进行编码。这些标准的示例在上面已经提到并且在本文中进一步描述。

[0044] 图3可以是根据本发明的实施方式的视频解码器300的功能框图。

[0045] 接收器302可以接收要由解码器300解码的一个或多个编码视频序列;在相同实施方式或另一实施方式中,一次接收一个经编码的视频序列,其中,每个经编码的视频序列的解码独立于其他经编码的视频序列。可以从信道301接收经编码的视频序列,信道301可以是到存储经编码的视频数据的存储装置的硬件/软件链路。接收器302可以接收经编码的视频数据以及其他数据,例如,经编码的音频数据和/或辅助数据流,这些数据可以被转发至其相应的使用实体(未描绘)。接收器302可以将经编码的视频序列与其他数据分开。为了防止网络抖动,可以将缓冲存储器303耦接在接收器302与熵解码器/解析器304(此后被称为“解析器”)之间。在接收器302正在从具有足够带宽和可控性的存储/转发装置或者从等时同步网络接收数据时,可以不需要缓冲器303或者缓冲器303可以是小的。为了在诸如因特网的尽力而为(best effort)型分组网络上使用,可能需要缓冲器303,缓冲器303可以相对较大并且可以有利地具有自适应大小。

[0046] 视频解码器300可以包括解析器304,以根据经熵编码的视频序列重建符号313。这些符号的类别包括:用于管理解码器300的操作的信息;以及潜在地用于控制诸如显示器312的呈现装置的信息,呈现装置不是解码器的组成部分但是可以耦接至解码器。用于呈现

装置的控制信息可以是补充增强信息 (SEI (Supplementary Enhancement Information, SEI) 消息) 或视频可用性信息参数集片段 (未描绘) 的形式。解析器304可以对接收到的经编码的视频序列进行解析/熵解码。经编码的视频序列的编码可以根据视频编码技术或标准进行, 并且可以遵循本领域技术人员公知的原理, 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。解析器304可以针对视频解码器中的像素的至少一个子群组基于与该群组对应的至少一个参数来从经编码的视频序列中提取子群组参数集。子群组可以包括图片群组 (Group of Picture, 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。熵解码器/解析器还可以从经编码的视频序列中提取诸如变换系数、量化器参数值、运动矢量等的信息。

[0047] 解析器304可以对从缓冲器303接收到的视频序列进行熵解码/解析操作, 以创建符号313。解析器304可以接收经编码的数据, 并且选择性地对特定符号313进行解码。此外, 解析器304可以确定是否将特定符号313提供给运动补偿预测单元306、缩放器/逆变换单元305、帧内预测单元307或环路滤波器311。

[0048] 根据经编码的视频图片或一部分经编码的视频图片的类型 (诸如: 帧间图片和帧内图片、帧间块和帧内块) 以及其他因素, 符号313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通过由解析器304从经编码的视频序列解析的子群组控制信息来控制。出于简洁起见, 未描绘解析器304与下面的多个单元之间的这样的子群组控制信息流。

[0049] 除了已经提及的功能块之外, 解码器300可以在概念上细分成如下面描述的多个功能单元。在商业限制下的实际实现操作中, 这些单元中的许多单元彼此紧密地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, 在概念上细分为以下功能单元是适当的。

[0050] 第一单元是缩放器/逆变换单元305。缩放器/逆变换单元305从解析器304接收量化变换系数以及控制信息——包括使用哪种变换、块大小、量化因子、量化缩放矩阵等——作为符号313。缩放器/逆变换单元305可以输出包括样本值的块, 块可以被输入至聚合器310中。

[0051] 在一些情况下, 缩放器/逆变换305的输出样本可以属于帧内编码块; 也就是说: 不使用来自先前重建的图像的预测信息, 但是可以使用来自当前图像的先前重建部分的预测信息的块。这样的预测信息可以由帧内图片预测单元307提供。在一些情况下, 帧内图片预测单元307使用从当前 (部分重建的) 图像309提取的周围已经重建的信息来生成与重建的块相同大小和形状的块。在一些情况下, 聚合器310基于每个样本将帧内预测单元307已经生成的预测信息添加至由缩放器/逆变换单元305提供的输出样本信息。

[0052] 在其他情况下, 缩放器/逆变换单元305的输出样本可以属于经帧间编码并且可能经运动补偿的块。在这种情况下, 运动补偿预测单元306可以访问参考图像存储器308以提取用于预测的样本。在根据属于块的符号313对所提取的样本进行运动补偿之后, 这些样本可以由聚合器310添加至缩放器/逆变换单元的输出 (在这种情况下被称为残差样本或残差信号), 以生成输出样本信息。运动补偿单元从其提取预测样本的参考图像存储器内的地址可以由运动矢量控制, 运动矢量能够以符号313的形式可用于运动补偿单元, 符号313可以

具有例如X分量、Y分量和参考图片分量。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提取的样本值的内插、运动矢量预测机制等等。

[0053] 聚合器310的输出样本可以在环路滤波器单元311中经受各种环路滤波技术。视频压缩技术可以包括环路内滤波技术,环路内滤波技术由被包括在经编码的视频比特流中并且作为来自解析器304的符号313可用于环路滤波器单元311的参数来控制,但是视频压缩技术还可以响应于在对经编码的图片或经编码的视频序列的先前(按解码顺序)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。

[0054] 环路滤波器单元311的输出可以是样本流,该样本流可以输出至呈现装置312以及存储在参考图片存储器557中,以用于将来的帧间图片预测。

[0055] 一旦完全重建,某些经编码的图片就可以用作参考图片以用于将来预测。一旦经编码的图片被完全重建并且该经编码的图片已经被识别为参考图片(通过例如解析器304),则当前参考图片309就可以变为参考图片缓冲器308的一部分,并且可以在开始重建后续经编码的图片之前重新分配新的当前图片存储器。

[0056] 视频解码器300可以根据可以记录在诸如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。在经编码的视频序列遵循视频压缩技术或标准的语法的意义上,经编码的视频序列可以符合由正使用的视频压缩技术或标准指定的语法,如在视频压缩技术文档或标准中并且具体地在其中的配置文档中指定的语法。对于合规性,还要求经编码的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、最大帧速率、最大重建样本速率(以例如每秒兆个样本为单位进行度量)、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(Hypothetical Reference Decoder,HRD)规范以及在经编码的视频序列中用信号发送的HRD缓冲器管理的元数据来进一步限定。

[0057] 在实施方式中,接收器302可以连同经编码的视频一起接收附加(冗余)数据。附加数据可以被包括为经编码的视频序列的一部分。附加数据可以由视频解码器300使用以对数据进行适当解码和/或更准确地重建原始视频数据。附加数据可以呈例如时间、空间或信噪比(signal-to-noise ratio,SNR)增强层、冗余切片、冗余图片、前向纠错码等形式。

[0058] 图4可以是根据本公开内容的实施方式的视频编码器400的功能框图。

[0059] 编码器400可以从视频源401(其并非编码器的一部分)接收视频样本,视频源401可以捕获要由编码器400进行编码的视频图像。

[0060] 视频源401可以提供要由编码器303进行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比特深度(例如:8比特、10比特、12比特,……)、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401可以是存储先前准备的视频的存储装置。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像素空间阵列,其中,取决于所使用的采样结构、色彩空间等,每个像素可以包括一个或更多个样本。本领域技术人员可以容易地理解像素与样本之间的关系。以下描述侧重于样本。

[0061] 根据实施方式,编码器400可以实时地或者按照应用所要求的任何其他时间约束,将源视频序列的图片编码并压缩成经编码的视频序列410。执行适当的编码速度是控制器

402的一个功能。控制器控制如下所述的其他功能单元并且在功能上耦接至这些单元。为了清楚起见,未描绘耦接。由控制器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、……)、图片大小、图片群组(GOP)布局、最大运动矢量搜索范围等。本领域技术人员可以容易地识别控制器402的其他功能,这是因为这些功能可能属于针对某个系统设计优化的视频编码器400。

[0062] 一些视频编码器以本领域技术人员所容易了解的“编码环路”进行操作。作为过于简化的描述,编码环路可以包括:编码器402(此后被称为“源编码器”)的编码部分(负责基于要被编码的输入图片和参考图片创建符号)以及嵌入在编码器400中的(本地)解码器406,(本地)解码器406重建符号以创建(远程)解码器也将创建的样本数据(因为在所公开的主题中考虑的视频压缩技术中,符号与经编码的视频比特流之间的任何压缩是无损的)。将重建的样本流输入至参考图片存储器405。由于对符号流的解码导致独立于解码器位置(本地或远程)的比特精确结果,因此参考图片缓冲器内容在本地编码器与远程编码器之间也是比特精确的。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在解码期间使用预测时将“看到”的样本值完全相同。该参考图片同步性的基本原理(以及在例如由于信道误差而不能保持同步性情况下而导致的漂移)是本领域技术人员公知的。

[0063] “本地”解码器406的操作可以与已经在上面结合图3详细描述“远程”解码器300的操作相同。然而,还简要参考图4,当符号可用并且熵编码器408和解析器304可以无损地将符号编码/解码为经编解码的视频序列时,在本地解码器406中可以不完全实现包括信道301、接收器302、缓冲器303和解析器304的解码器300的熵解码部分。

[0064] 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需要以基本上相同的功能形式存在于对应的编码器中。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。更详细的描述仅在某些区域中需要并且在下面被提供。

[0065] 作为其操作的一部分,源编码器403可以进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“参考帧”的一个或多个先前经编码的帧对输入帧进行预测性编码。以这样的方式,编码引擎407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,参考帧可以被选作输入帧的预测参考。

[0066] 本地视频解码器406可以基于由源编码器403创建的符号对可以指定为参考帧的帧的经编码的视频数据进行解码。编码引擎407的操作可以有利地是有损处理。当经编码的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。本地视频解码器406复制可以由视频解码器对参考帧进行的解码处理,并且可以使重建的参考帧存储在参考图片缓存405中。以这样的方式,编码器400可以在本地存储重建的参考帧的副本,该副本与将由远端视频解码器获得的重建的参考帧具有共同内容(不存在传输误差)。

[0067] 预测器404可以针对编码引擎407执行预测搜索。也就是说,对于要编码的新帧,预测器404可以在参考图片存储器405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,诸如,参考图片运动矢量、块形状等。预测器404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,如由预测器404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405中存储的多个参考图片提取的预

测参考。

[0068] 控制器402可以管理视频编码器403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。

[0069] 可以在熵编码器408中对所有上述功能单元的输出进行熵编码。熵编码器通过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)对符号进行无损压缩来将由各种功能单元生成的符号转换成经编码的视频序列。

[0070] 传输器409可以缓冲由熵编码器408创建的经编码的视频序列,从而为经由通信信道411进行传输做准备,通信信道411可以是到将存储经编码的视频数据的存储装置的硬件/软件链路。传输器409可以将来自视频编码器403的经编码的视频数据与要传输的其他数据合并,其他数据例如是经编码的音频数据和/或辅助数据流(未示出源)。

[0071] 控制器402可以管理编码器400的操作。在编码期间,控制器405可以向每个经编码的图片分配特定编码图片类型,这可能影响可以应用于相应的图片的编码技术。例如,通常可以向图片分配以下帧类型之一:

[0072] 帧内图片(I图片),其可以是不将序列中的任何其他帧用作预测源就可以被编码和解码的图片。一些视频编码器允许不同类型的帧内图片,包括例如独立解码器刷新图片。本领域的技术人员了解I图片的这些变型及其相应的应用和特征。

[0073] 预测图片(P图片),其可以是可以使用利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帧内预测或帧间预测进行编码和解码的图片。

[0074] 双向预测图片(B图片),其可以是可以使用利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帧间预测或帧内预测进行编码和解码的图片。类似地,多预测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重建单个块。

[0075] 源图片通常可以在空间上细分成多个样本块(例如,分别为 4×4 、 8×8 、 4×8 或 16×16 样本的块),并且逐块进行编码。这些块可以参考根据应用于块的相应图片的编码分配确定的其他(已经编码的)块进行预测性编码。举例来说,可以对I图片的块进行非预测性编码,或可以参考相同图片的已经编码的块对I图片的块进行预测性编码(空间预测或帧内预测)。可以参考一个先前编码的参考图片经由空间预测或经由时间预测对P图片的像素块进行非预测性编码。可以参考一个或两个先前编码的参考图片经由空间预测或经由时间预测对B图片的块进行非预测性编码。

[0076] 视频编码器400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400可以进行各种压缩操作,包括利用输入视频序列中的时间冗余和空间冗余的预测编码操作。因此,编码的视频数据可以遵循由使用的视频编码技术或标准指定的语法。

[0077] 在实施方式中,传输器409可以传输附加数据和编码的视频。源编码器403可以包括这样的数据作为经编码的视频序列的一部分。附加数据可以包括时间/空间/SNR增强层、其他形式的冗余数据例如冗余图片和切片、辅助增强信息(SEI)消息、视觉可用性信息(Visual Usability Information,VUI)参数集片段等。

[0078] 图5示出了HEVC和JEM中使用的帧内预测模式。为了捕获自然视频中呈现的任意边缘方向,方向性帧内模式的数量从HEVC中使用的33个扩展到65个。在HEVC的基础上的JEM中的附加方向性模式在图1(b)中被描绘为虚线箭头,并且平面模式和DC模式保持不变。这些

更密集的方向性帧内预测模式适用于所有块大小以及亮度和色度帧内预测两者。如图5所示,与奇数帧内预测模式索引相关联的、由虚线箭头标识的方向性帧内预测模式被称为奇数帧内预测模式。与偶数帧内预测模式索引相关联的、由实线箭头标识的方向性帧内预测模式被称为偶数帧内预测模式。在本文档中,如图5中的实线或虚线箭头所指示的方向性帧内预测模式也被称为角度模式。

[0079] 在JEM中,总共67个帧内预测模式用于亮度帧内预测。为了编码帧内模式,基于相邻块的帧内模式来构建大小为6的最可能模式(Most Probable Mode,MPM)列表。如果帧内模式不是来自MPM列表,则用信号发送标志以指示帧内模式是否属于所选模式。在JEM-3.0中,存在被每四个角度模式均匀选择的16个所选模式。在JVET-D0114和JVET-G0060中,导出16个次级MPM以替换均匀选择的模式。

[0080] 图6示出了针对帧内方向性模式采用的N个参考层。存在块单元611、段A 601、段B 602、段C 603、段D 604、段E 605、段F 606、第一参考层610、第二参考层609、第三参考层608和第四参考层607。

[0081] 在HEVC和JEM两者以及诸如H.264/AVC的一些其他标准中,用于预测当前块的参考样本被限于最近参考线(行或列)。在多参考线帧内预测方法中,对于帧内方向性模式,候选参考线(行或列)的数量从一(即最近的)增加到N,其中N是大于或等于一的整数。图2采用 4×4 预测单元(PU)作为示例,示出了多线帧内方向性预测方法的概念。帧内方向性模式可以任意选择N个参考层中的一个来生成预测器。换句话说,根据参考样本S1、S2、……和SN中的一个生成预测器 $p(x, y)$ 。用信号发送标志以指示哪一参考层被选择用于帧内方向性模式。如果N被设置为1,则帧内方向性预测方法与JEM2.0中的传统方法相同。在图6中,参考线610、609、608和607由六个段601、602、603、604、605和606以及左上参考样本组成。在本文档中,参考层也被称为参考线。当前块单元内的左上像素的坐标是(0,0),并且第1参考线中的左上像素的坐标是(-1,-1)。

[0082] 在JEM中,对于亮度分量,在生成处理之前对用于帧内预测样本生成的相邻样本进行滤波。滤波通过给定帧内预测模式和变换块大小进行控制。如果帧内预测模式是DC或者变换块大小等于 4×4 ,则不对相邻样本进行滤波。如果给定帧内预测模式与竖直模式(或水平模式)之间的距离大于预定义阈值,则启用滤波处理。对于相邻样本滤波,使用[1,2,1]滤波器和双线性滤波器。

[0083] 位置相关的帧内预测组合(position dependent intra prediction combination,PDPC)方法是帧内预测方法,其调用未经滤波的边界参考样本和具有经滤波的边界参考样本的HEVC式帧内预测的组合。如下计算位于(x,y)的每个预测样本 $\text{pred}[x][y]$:

[0084] $\text{pred}[x][y] = (wL * R_{-1,y} + wT * R_{x,-1} + wTL * R_{-1,-1} + (64 - wL - wT - wTL) * \text{pred}[x][y] + 32) > 6$

[0085] (等式.2-1)

[0086] 其中 $R_{x,-1}$ 、 $R_{-1,y}$ 分别表示位于当前样本(x,y)的上方处和左侧处的未经滤波的参考样本,并且 $R_{-1,-1}$ 表示位于当前块的左上角处的未经滤波的参考样本。如下计算权重:

[0087] $wT = 32 >> ((y << 1) >> \text{移位})$ (等式2-2)

[0088] $wL = 32 >> ((x << 1) >> \text{移位})$ (等式2-3)

[0089] $wTL = -(wL \gg 4) - (wT \gg 4)$ (等式2-4)

[0090] 移位 = $(\log_2(\text{宽度}) + \log_2(\text{高度}) + 2) \gg 2$ (等式2-5)。

[0091] 图7示出了图700, 其中一个 4×4 块内的(0,0)位置和(1,0)位置的DC模式PDPC权重(wL, wT, wTL)。如果将PDPC应用于DC帧内模式、平面帧内模式、水平帧内模式和竖直帧内模式, 则不需要附加的边界滤波器, 例如HEVC DC模式边界滤波器或水平模式/竖直模式边缘滤波器。图7示出了用于应用于右上对角模式的PDPC的参考样本 $R_{x,-1}, R_{-1,y}$ 和 $R_{-1,-1}$ 的定义。预测样本 $\text{Pred}(x', y')$ 位于预测块内的 (x', y') 处。参考样本 $R_{x,-1}$ 的坐标 x 由下式给出: $x = x' + y' + 1$, 并且参考样本 $R_{-1,y}$ 的坐标 y 类似地由下式给出: $y = x' + y' + 1$ 。

[0092] 图8示出了局部照明补偿(Local Illumination Compensation, LIC)图800, 并且是基于使用缩放因子 a 和偏移 b 的用于照明变化的线性模型。并且其针对每个帧间模式编码的编码单元(CU)自适应地被启用或禁用。

[0093] 当针对CU应用LIC时, 通过使用当前CU的相邻样本及其对应的参考样本, 采用最小平方误差法来导出参数 a 及参数 b 。更具体地, 如图8所示, 使用CU的子采样(2:1子采样)的相邻样本和参考图片中的(由当前CU或子CU的运动信息识别的)对应样本。导出IC参数并且分别针对每个预测方向应用这些IC参数。

[0094] 当使用合并模式对CU进行编码时, 以与合并模式中的运动信息复制类似的方式从相邻块复制LIC标志; 否则, 针对CU用信号发送LIC标志以指示是否应用了LIC。

[0095] 图9A示出了在HEVC中使用的帧内预测模式900。在HEVC中, 总共存在35个帧内预测模式, 其中模式10是水平模式, 模式26是竖直模式, 并且模式2、模式18和模式34是对角模式。帧内预测模式由三个最可能模式(MPM)和32个剩余模式用信号发送。

[0096] 图9B示出了, 在VVC的实施方式中总共存在87个帧内预测模式, 其中模式18是水平模式, 模式50是竖直模式, 并且模式2、模式34和模式66是对角模式。模式-1至模式-10以及模式67至模式76被称为广角帧内预测(Wide-Angle Intra Prediction, WAIP)模式。

[0097] 根据PDPC表达式, 使用帧内预测模式(DC、平面、角度)和参考样本的线性组合来预测位于位置 (x, y) 处的预测样本 $\text{pred}(x, y)$:

[0098] $\text{pred}(x, y) = (wL \times R_{-1,y} + wT \times R_{x,-1} - wTL \times R_{-1,-1} + (64 - wL - wT + wTL) \times \text{pred}(x, y) + 32) \gg 6$

[0099] 其中, $R_{x,-1}, R_{-1,y}$ 分别表示位于当前样本 (x, y) 的上方处和左侧处的参考样本, 并且 $R_{-1,-1}$ 表示位于当前块的左上角处的参考样本。

[0100] 对于DC模式, 针对具有宽度和高度的尺寸的块, 如下计算权重:

[0101] $wT = 32 \gg ((y < 1) \gg nScale)$, $wL = 32 \gg ((x < 1) \gg nScale)$, $wTL = (wL \gg 4) + (wT \gg 4)$,

[0102] 其中, $nScale = (\log_2(\text{宽度}) - 2 + \log_2(\text{高度}) - 2 + 2) \gg 2$, 其中, wT 表示位于上参考线内具有相同横坐标的参考样本的加权因子, wL 表示位于左参考线内具有相同纵坐标的参考样本的加权因子, 并且 wTL 表示当前块的左上参考样本的加权因子, $nScale$ 指定加权因子沿轴减小(wL 从左到右减小或者 wT 从上到下减小)的速度, 即加权因子递减速率, 并且在当前设计中该加权因子递减速率沿 x 轴(从左到右)和 y 轴(从上到下)是相同的。并且32表示相邻样本的初始加权因子, 并且初始加权因子也是分配给当前CB中的左上样本的上(左或左上)加权, 并且PDPC处理中的相邻样本的加权因子应当等于或小于该初始加权因子。

[0103] 对于平面模式而言 $w_{TL}=0$,而对于水平模式而言 $w_{TL}=w_T$,并且对于竖直模式而言 $w_{TL}=w_L$ 。可以仅利用加法和移位来计算PDPC权重。 $\text{pred}(x,y)$ 的值可以使用等式1在单个步骤中进行计算。

[0104] 本文提出的方法可以单独使用或以任何顺序组合使用。此外,方法(或实施方式)、编码器以及解码器中的每个可以由处理电路系统(例如,一个或更多个处理器或者一个或更多个集成电路)来实现。在一个示例中,一个或更多个处理器执行被存储在非暂态计算机可读介质中的程序。在下文中,术语块可以被解释为预测块、编码块或编码单元,即CU。

[0105] 图10A示出了通过使用QTBT进行块分割的示例1000,并且图10B示出了对应的树表示1001。实线指示四叉树划分,并且虚线指示二叉树划分。在二叉树的每个划分(即,非叶)节点中,用信号发送一个标志以指示使用哪种划分类型(即,水平或竖直),其中0指示水平划分并且1指示竖直划分。对于四叉树划分,不需要指示划分类型,因为四叉树划分总是水平地并且竖直地划分块以产生具有相等大小的4个子块。

[0106] 在HEVC中,通过使用表示为编码树的四叉树结构将CTU划分成CU以适应各种局部特征。在CU级别处做出是否使用图片间(时间)或图片内(空间)预测对图片区域进行编码的决定。可以根据PU划分类型,进一步将每个CU划分成一个、两个或四个PU。在一个PU内部,应用相同的预测处理,并且逐PU将相关信息传输至解码器。在通过应用基于PU划分类型的预测处理获得残差块之后,可以根据另一四叉树结构(如CU的编码树)将CU分割为变换单元(TU)。HEVC结构的关键特征之一是其具有包括CU、PU和TU的多分割构思。

[0107] 根据实施方式,QTBT结构去除了多个分割类型的概念,即它去除了CU、PU和TU概念的划分,并且支持CU分割形状的更大灵活性。在QTBT块结构中,CU可以具有正方形或矩形形状。在图11的流程图1100处,根据示例性实施方式,在S11获得的编码树单元(coding tree unit,CTU)或CU首先在S12处通过四叉树结构分割。在S14,进一步确定是否通过二叉树结构分割四叉树叶节点,如果是,则在S15,如图10C所描述的,例如,在二叉树划分中存在两种划分类型,对称水平划分和对称竖直划分。二叉树叶节点被称为编码单元(CU),并且在没有任何其他分割的情况下使用该段进行预测和变换处理。这意味着在QTBT编码块结构中CU、PU和TU具有相同的块大小。在VVC中,有时CU由不同颜色分量的编码块(Coding Block,CB)构成,例如在4:2:0色度格式的P切片和B切片的情况下,一个CU包含一个亮度CB和两个色度CB,有时一个CU由单个分量的CB构成,例如,在I切片的情况下,一个CU仅包含一个亮度CB或者仅包含两个色度CB。

[0108] 根据实施方式,为QTBT分割方案定义以下参数:

[0109] -CTU size:四叉树的根节点大小,与HEVC中的概念相同,

[0110] -MinQTSIZE:允许的最小四叉树叶节点大小,

[0111] -MaxBTSIZE:允许的最大二叉树根节点大小,

[0112] -MaxBTDepth:允许的最大二叉树深度,以及

[0113] -MinBTSIZE:允许的最小二叉树叶节点大小。

[0114] 在QTBT分割结构的一个示例中,CTU size被设置为 128×128 亮度样本和两个对应的 64×64 色度样本块,其中QT是四叉树的MinQTSIZE被设置为 16×16 ,MaxBTSIZE被设置为 64×64 ,MinBTSIZE(对于宽度和高度两者)被设置为 4×4 ,并且MaxBTDepth被设置为4。四叉树分割首先应用于CTU,以在S12或S15生成四叉树叶节点。四叉树叶节点可以具有从 16×16

(即,MinQTSIZE)到 128×128 (即,CTU size)的大小。如果叶四叉树节点是 128×128 ,则由于如在S14检查的,大小超过了MaxBTSIZE(即 64×64),因此不会通过二叉树进一步划分。否则,叶四叉树节点可以在S15通过二叉树进一步分割。因此,四叉树叶节点也是二叉树的根节点并且其具有为0的二叉树深度。当二叉树深度达到MaxBTDepth(即,4)时,在S14不考虑进一步划分。当二叉树节点的宽度等于MinBTSIZE(即4)时,在S14不考虑进一步水平划分。类似地,当二叉树节点的高度等于MinBTSIZE时,在S14不考虑进一步竖直划分。如下面关于描述QT/TT/BT大小的语法所讨论的,在S16提供信号,用于诸如在S17通过预测和变换处理进一步处理的二叉树的叶节点的处理,并且类似地如本文关于这样的预测和变换处理所讨论的,没有任何进一步的分割。如图11所示,根据示例性实施方式,这样的信号发送也可以在S12之后的S13提供。在JEM中,最大CTU大小为 256×256 亮度样本。

[0115] 另外,根据实施方式,QTBT方案支持亮度和色度具有单独的QTBT结构的能力/灵活性。目前,对于P切片和B切片,一个CTU中的亮度和色度编码树块(coding tree block,CTB)共享相同的QTBT结构。然而,对于I切片,亮度CTB通过QTBT结构被分割为CU,而色度CTB通过另一QTBT结构被分割为色度CU。这意味着I切片中的CU由亮度分量的编码块或两个色度分量的编码块组成,并且P切片或B切片中的CU由所有三个颜色分量的编码块组成。

[0116] 在HEVC中,针对小块的帧间预测受到限制以减少运动补偿的存储器访问,使得对于 4×8 块和 8×4 块不支持双预测,并且对于 4×4 块不支持帧间预测。如在JEM-7.0中实现的QTBT中,这些限制被去除。

[0117] 图10C表示关于所包括的多类型树(Multi-type-tree,MTT)结构1002的简化框图1100VVC,其是所示的四叉树(quadtrees,QT)与嵌套二叉树(binary tree,BT)和三元树/三叉树(triple-tree/ternary tree,TT)的组合,即QT/BT/TT。首先,CTU或CU通过QT被递归地分割成正方形块。然后,每个QT叶可以由BT或TT进一步分割,其中,BT和TT划分可以递归地应用和交织,但是不能应用另外的QT分割。在所有相关提案中,TT使用1:2:1比率将矩形块竖直或水平地划分成三个块(从而避免非二次方的宽度和高度)。为了防止分割模仿(emulation),通常对MTT施加附加的划分约束,如图10C的简化图1002所示,在VVC中对于块1103(四叉树)、块1104(二叉树、JEM)和块1105(三叉树)的QT/BT/TT块分割,以避免重复分割(例如,抑制由竖直/水平三叉划分导致的中间分割上的竖直/水平二叉划分)。可以对BT和TT的最大深度设置进一步的限制。

[0118] 如上面的三叉树块1105所述,这样的三叉树分割的关键优点为:作为四叉树和二叉树分割的补充,三叉树分割能够捕获位于块中心的对象,而四叉树和二叉树总是沿着块中心分割,并且所提出的三叉树的分割的宽度和高度总是2次方,使得不需要额外的变换。

[0119] 两级树的设计主要是为了降低复杂度。理论上,遍历树的复杂度为 T^D ,其中,T表示划分类型的数量,D是树的深度。

[0120] 图11示出了VP9和AV1中的块分割的示例1100,其中VP9的示例编码树单元(coding tree unit,CTU)1111示出了VP 9使用从 64×64 级别1112开始向下到 4×4 级别1113的4路分割树,针对块 8×8 以及块 8×8 以下具有一些附加限制,如级别1113的上半部分所示。注意,指定为R的分割涉及递归,这在于以较小规模重复相同的分割树,直到达到最低的 4×4 级别。AV1的示例CTU 1104不仅将分割树扩展到10路结构1116,而且还增加最大大小(在VP9/AV1的说法中称为超级块)以从 128×128 级别1115开始。注意,这包括VP9中不存在的4:1/1:

4矩形分割。并且矩形分割都不能被进一步细分。此外,在可以在某些情况下进行 2×2 色度帧间预测的意义上,AV1为使用 8×8 级别以下的分割增加了更多的灵活性。

[0121] 并且在HEVC中,可以通过使用被表示为编码树的四叉树结构来将编码树单元(CTU)划分成编码单元(CU),以适于各种局部特性。在CU级别处做出是否使用图片间(时间)或图片内(空间)预测对图片区域进行编码的决定。可以根据PU划分类型,进一步将每个CU划分成一个、两个或四个预测单元(PU)。在一个PU内部,应用相同的预测处理,并且逐PU将相关信息传输至解码器。在通过应用基于PU划分类型的预测处理获得残差块之后,可以根据另一四叉树结构(如CU的编码树)将CU分割为变换单元(TU)。HEVC结构的关键特征之一是其具有包括CU、PU和TU的多分割构思。在HEVC中,CU或TU可以仅为正方形形状,而PU针对帧间预测块可以是正方形或矩形形状。在HEVC中,一个编码块还可以被划分成四个正方形子块,并且对每个子块即TU执行变换。每个TU可以被进一步递归地(使用四叉树划分)划分成较小的TU,其被称为残差四叉树(Residual Quad-Tree, RQT)。并且在图片边界处,HEVC采用隐式四叉树划分,使得块将保持四叉树划分,直到大小适合图片边界。

[0122] 此外,图12示出了根据示例性实施方式的与具有运动矢量差的合并模式(merge mode with motion vector difference, MMVD)相关的示例1200。例如,除了其中将隐式地导出的运动信息直接用于当前CU的预测样本生成的合并模式之外,在VVC中引入具有运动矢量差的合并模式(MMVD)。并且可以在发送跳过标志和合并标志之后立即用信号发送MMVD标志,以指定MMVD模式是否用于CU。并且在MMVD中,在选择合并候选之后,通过用信号发送的运动矢量差(MVD)信息进一步细化合并候选,使得另外的信息包括合并候选标志、用于指定运动幅度的索引和用于指示运动方向的索引。在MMVD模式下,合并列表中的前两个候选中的一个被选择用作MV基础。可以用信号发送合并候选标志以指定使用哪一个。

[0123] 距离索引指定运动幅度信息,并指示从起点的预定义偏移。图12示出了L0基准1201和L1基准1202,其中偏移被添加到起始MV的水平分量或竖直分量。表1中指定了距离索引和预定义偏移量的关系。

[0124] 表1-距离索引与预定义偏移量的关系

	距离 IDX	0	1	2	3	4	5	6	7
[0125]	偏移量（以亮度 样本为单位）	1/4	1/2	1	2	4	8	16	32

[0126] 根据示例性实施方式,方向索引表示MVD相对于起点的方向。方向索引可以表示四个方向,如下表2所示。MVD符号的含义可以根据起始MV的信息而变化。例如,当起始MV是单预测MV或双预测MV,其中两个列表都指向当前图片的同一侧时(即(两个参考的图片顺序计数(picture order count, POC)都大于当前图片的POC,或者都小于当前图片的POC),表2中的符号指定添加至起始MV的MV偏移的符号。并且/或者当起始MV是双预测MV,其中两个MV指向当前图片的不同侧(即,一个参考的POC大于当前图片的POC,而另一个参考的POC小于当前图片的POC),并且列表0中的POC差大于列表1中的POC差时,表2中的符号指定添加至起始MV的列表0MV分量的MV偏移的符号,并且列表1MV的符号具有相反的值。否则,如果列表1中的POC差大于列表0,则表2中的符号指定添加至到起始MV的列表1MV分量的MV偏移的符号,

并且列表0MV的符号具有相反的值。

[0127] 根据示例性实施方式,可以根据POC在每个方向上的差来缩放MVD。如果两个列表中POC的差相同,则不需要缩放。否则,如果列表0中的POC差大于列表1中的POC差,则缩放列表1的MVD。如果L1的POC差大于L0的POC差,则列表0的MVD以相同的方式缩放。如果起始MV是单预测的,则MVD被添加至可用MV。

[0128] 表2-方向索引指定的MV偏移的符号

方向IDX	00	01	10	11
x轴	+	-	N/A	N/A
y轴	N/A	N/A	+	-

[0130] 根据示例性实施方式,可以存在对称MVD编码,其中可以根据在每个方向上的POC差来缩放MVD。如果两个列表中POC的差相同,则不需要缩放。否则,如果列表0中的POC差大于列表1中的POC差,则缩放列表1的MVD。如果L1的POC差大于L0的POC差,则列表0的MVD以相同的方式缩放。如果起始MV是单预测的,则MVD被添加至可用MV。

[0131] 并且根据示例性实施方式,在VVC中,除了正常的单向预测和双向预测模式MVD信号发送之外,还可以应用用于双向MVD信号发送的对称MVD模式。在对称MVD模式下,包括列表-0和列表-1的参考图片索引和列表-1的MVD的运动信息不被用信号发送,而是被导出。对称MVD模式的解码处理如下:

[0132] 1.在切片级别,变量BiDirPredFlag、RefIdxSymL0和RefIdxSymL1被导出如下:

[0133] -如果mvd_l1_zero_flag为1,则BiDirPredFlag设置为等于0。

[0134] -否则,如果列表-0中最近的参考图片和列表-1中最近的参考图片形成前向和后向的参考图片对或后向和前向的参考图片对,则BiDirPredFlag被设置为1,并且列表-0和列表-1参考图片都是短期参考图片。否则BiDirPredFlag被设置为0。

[0135] 2.在CU级别,如果CU被双预测编码并且BiDirPredFlag等于1,则显式地用信号发送指示是否使用对称模式的对称模式标志。

[0136] 并且当对称模式标志为真时,仅mvp_l0_flag、mvp_l1_flag和MVD0被显式地用信号发送。将列表-0和列表-1的参考索引分别设置为等于参考图片对。MVD1被设置为等于(-MVD0)。

[0137] 根据示例性实施方式,在CWG-B018中可以存在帧间模式编码,其中在AV1中,对于帧间的每个编码块,如果当前块的模式不是跳过模式而是帧间编码模式,则用信号发送另一个标志以指示当前块是使用单参考模式还是复合参考模式,其中预测块在单参考模式中由一个运动矢量生成,而预测块在复合参考模式中通过对从两个运动矢量导出的两个预测块进行加权平均来生成。

[0138] 例如,对于单参考情况,可以用信号发送以下模式:

[0139] NEARMV-使用由动态参考列表(Dynamic Reference List,DRL)索引指示的列表中的运动矢量预测器(motion vector predictor,MVP)之一;

[0140] NEWMV-使用通过DRL索引信号发送的列表中的运动矢量预测器(MVP)之一作为参考,并对MVP应用增量;

[0141] GLOBALMV-使用基于帧级别全局运动参数的运动矢量。

[0142] 并且对于复合参考模式,可以用信号发送以下模式:

[0143] NEAR_NEARMV-使用通过DRL索引信号发送的列表中的运动矢量预测器(MVP)之一;

[0144] NEAR_NEWMV-使用通过DRL索引信号发送的列表中的运动矢量预测器(MVP)之一作为参考,并为第二MV发送增量MV;

[0145] NEW_NEARMV-使用通过DRL索引信号发送的列表中的运动矢量预测器(MVP)之一作为参考,并为第一MV发送增量MV;

[0146] NEW_NEWMV-使用通过DRL索引信号发送的列表中的运动矢量预测器(MVP)之一作为参考,并为两个MV都发送增量MV;

[0147] GLOBAL_GLOBALMV-基于其帧级别全局运动参数使用来自每个参考的MV。

[0148] 并且根据示例性实施方式,在AV1中还可以存在运动矢量差编码,其中AV1允许1/8像素运动矢量精度(或准确度),并且使用以下语法来用信号发送参考帧列表0或列表1中的运动矢量差:

[0149] mv_joint指定运动矢量差的哪些分量是非零的:

[0150] 0指示沿水平或竖直方向不存在非零MVD,

[0151] 1指示仅沿水平方向存在非零MVD,

[0152] 2指示仅沿竖直方向存在非零MVD,以及

[0153] 3指示沿水平和竖直方向均存在非零MVD,

[0154] mv_sign指定运动矢量差是正的还是负的,

[0155] mv_class指定运动矢量差的类别(如表3所示,类别越高表示运动矢量差的幅度越大)

[0156] 表3:运动矢量差的幅度类别

MV类别	MVD的幅度
MV_CLASS_0	(0,2]
MV_CLASS_1	(2,4]
MV_CLASS_2	(4,8]
MV_CLASS_3	(8,16]
MV_CLASS_4	(16,32]
MV_CLASS_5	(32,64]
MV_CLASS_6	(64,128]
MV_CLASS_7	(128,256]
MV_CLASS_8	(256,512]
MV_CLASS_9	(512,1024]
MV_CLASS_10	(1024,2048]

[0158] mv_bit指定运动矢量差与每个MV类别的起始幅度之间的偏移的整数部分,

[0159] mv_fr指定运动矢量差的前2个小数比特,并且

[0160] mv_hp指定运动矢量差的第三小数比特。

[0161] 并且根据示例性实施方式,在CWG-B092中可以存在自适应MVD分辨率,其中对于NEW_NEARMV和NEAR_NEWMV模式,MVD的精度取决于相关联的类别和MVD的幅度。

[0162] 首先,只有当MVD幅度等于或小于一个像素时,才可以允许分数MVD。其次,当相关

联的MV类别的值等于或大于MV_CLASS_1时,仅可以允许一个MVD值,并且对于MV类别1(MV_CLASS_1)、MV类别2(MV_CLASS_2)、MV类别3(MV_CLASS_3)、MV类别4(MV_CLASS_4)或MV类别5(MV_CLASS_5),每个MV类别中的MVD值被导出为4、8、16、32、64。每个MV类别中允许的MVD值如表4所示:

[0163] 表4:每个MV幅度类别中的自适应MVD

MV类别	MVD的幅度
MV_CLASS_0	(0,1], {2}
MV_CLASS_1	{4}
MV_CLASS_2	{8}
MV_CLASS_3	{16}
MV_CLASS_4	{32}
MV_CLASS_5	{64}
MV_CLASS_6	{128}
MV_CLASS_7	{256}
MV_CLASS_8	{512}
MV_CLASS_9	{1024}
MV_CLASS_10	{2048}

[0165] 另外,如果当前块被编码为NEW_NEARMV或NEAR_NEWMV模式,则使用一个上下文来用信号发送mv_joint或mv_class。否则,使用另一上下文来用信号发送mv_joint或mv_class。

[0166] 根据示例性实施方式,在CWG-B092中还可以存在联合MVD编码(JMVD),其中可以应用名为JOINT_NEWMV的新的帧间编码模式来指示两个参考列表的MVD是否被联合用信号发送。如果帧间预测模式等于JOINT_NEWMV模式,则参考列表0和参考列表1的MVD被联合用信号发送。因此,仅名为joint_mvd的一个MVD可以被用信号发送并传输至解码器,并且参考列表0和参考列表1的增量mv是从joint_mvd导出的。

[0167] JOINT_NEWMV模式可以与NEAR_NEARMV、NEAR_NEWMV、NEW_NEARMV、NEW_NEWMV和GLOBAL_GLOBALMV模式一起用信号发送。根据示例性实施方式,不添加额外的上下文。

[0168] 并且当用信号发送JOINT_NEWMV模式时,并且两个参考帧与当前帧之间的POC距离不同时,基于POC距离为参考列表0或参考列表1缩放MVD。具体地说,参考帧列表0与当前帧之间的距离记为td0,参考帧列表1与当前帧之间的距离记为td1。如果td0等于或大于td1,则joint_mvd直接用于参考列表0,并且参考列表1的mvd基于等式(1)从joint_mvd导出。

$$[0169] \quad \textit{derived_mvd} = \frac{td1}{td0} * \textit{joint_mvd} \quad \text{等式(1)}$$

[0170] 否则,如果td1等于或大于td0,则joint_mvd直接用于参考列表1,并且参考列表0的mvd基于等式(2)从joint_mvd导出。

$$[0171] \quad \textit{derived_mvd} = \frac{td0}{td1} * \textit{joint_mvd} \quad \text{等式(2)}$$

[0172] 根据示例性实施方式,CWG-C011中还存在自适应MVD分辨率的改进,其中名为AMVDMV的新帧间编码模式可以被添加至单个参考案例。当选择AMVDMV模式时,该选择指示

AMVD被应用至信号MVD。名为amvd_flag的一个标志可以在JOINT_NEWMV模式下被添加,以指示AMVD是否应用于联合MVD编码模式。并且当自适应MVD分辨率应用于联合MVD编码模式时,两个参考帧的MVD被联合用信号发送,并且MVD的精度由MVD幅度隐式地确定。否则,两个(或多于两个)参考帧的MVD被联合用信号发送,并且可以应用其他MVD编码。

[0173] 根据示例性实施方式,可以在CWG-C012和CWG-C020中使用自适应运动矢量分辨率(adaptive motion vector resolution,AMVR)。CWG-C012的AMVR总共包括7个运动矢量(motion vector,MV)精度(8,4,2,1,1/2,1/4,1/8)被支持。对于每个预测块,自适应AMVR编码器可以搜索所有支持的精度值,并将最佳精度用信号发送至解码器。为了减少编码器运行时间,支持两个精度集。每个精度集包含4个预定义的精度。基于帧的最大精度值,在帧级别自适应地选择精度集。与AV1相同,最大精度在帧标头中用信号发送。表5总结了基于帧级别最大精度的支持精度值:

[0174] 表5:两个集合中支持的MV精度

[0175]	帧级别最大精度	支持的MV精度
	1/8	1/8,1/2,1,4
	1/4	1/4,1,4,8

[0176] 在当前的AMVR软件(类似于AV1)中,存在帧级别标志以指示帧的MV是否包含子像素精度。只有当cur_frame_force_integer_mv标志的值为0时,才可以启用AMVR。在AMVR中,如果块的精度低于最大精度,则不用信号发送运动模型和插值滤波器。如果块的精度低于最大精度,则运动模式被推断为平移运动,插值滤波器被推断为常规插值滤波器。类似地,如果块的精度是4像素或8像素,则帧间模式不被用信号发送并被推断为0。

[0177] 并且可以进行另外的改进,因为在JMVD的情况下,假设在后向参考帧与前向参考帧之间存在线性运动,但是当为一个块选择JMVD编码模式时,对于两个参考帧用信号发送一个联合MVD,并且基于参考帧与当前帧之间的距离从联合MVD导出两个参考帧的MVD。然而,两个参考帧之间的运动可能不总是线性运动,因为例如,从后向参考帧到前向参考帧,运动可能变得更慢或更快。

[0178] 根据示例性实施方式,参考帧的方向通过参考帧按照显示顺序是在当前帧之前还是按照显示顺序在当前帧之后来确定。并且,术语x轴和y轴指的是2-D值的水平和垂直分量,但是这些术语可以由沿着彼此垂直的两个预定义方向的另外两个轴来代替,并且本文所描述的相同实施方式也适用。例如,x轴和y轴可以用45度轴和135度轴代替。

[0179] 图13示出了流程图1300,其中在S130获得视频数据的编码块,并且在S131可以确定为该块选择哪种模式。例如,根据示例性实施方式,其中确定为一个块选择了JMVD模式,则在S133可以使用灵活缩放因子,以在S135基于比特流中的用信号发送的缩放因子或当前块(或相邻块)的编码信息,从用信号发送的联合MVD导出参考帧列表0和/或列表1的MVD。灵活缩放因子可以联合地或单独地应用于MVD的沿一个或更多个预定义方向的一个或更多个分量。

[0180] 并且根据示例性实施方式,MVD的预定义方向指的是沿着x轴和/或y轴的MVD分量。

[0181] 根据示例性实施方式,在S131,当确定为一个块选择了JMVD模式时,则在S135,基于参考帧与当前帧之间的距离和/或JMVD模式的缩放因子,从用信号发送的联合MVD导出参考帧列表0或列表1的MVD。

[0182] 根据示例性实施方式,在S133,可以确定用信号发送/导出的标志指示缩放因子不等于第一预定义默认值(例如,1)。然后,在S135,可以使用用信号发送的标志的相关缩放因子,以针对参考帧中的一个参考帧从联合MVD导出MVD。用于导出其他参考帧的MVD的缩放因子被设置为第二预定义默认值(例如,1)。

[0183] 根据示例性实施方式,当在S131确定当前块被编码为JMVD模式时,在S136,可以在比特流中用信号发送至少一个标志,以指示沿一个或更多个预定义方向的MVD的分量的缩放因子。沿一个或更多个预定义方向的MVD的分量指的是沿着x轴和/或y轴的MVD分量。

[0184] 并且,当在S131确定当前块是用联合AMVD(或AMVR)编码模式编码时,在S133和S134,可以确定x轴和y轴的缩放因子相等。

[0185] 根据示例性实施方式,当确定当前块被编码为JMVD模式时,在S136,可以用信号发送诸如jmvd_scale_factor_flag的标志,以指示用于导出MVD的缩放因子。并且如果jmvd_scale_factor_flag指示缩放因子的值不等于1,则名为scale_factor_dir的另一个标志可以被用信号发送,以指示缩放因子是被应用于MVD的x轴,y轴,还是x轴和y轴两者。

[0186] 根据示例性实施方式,当确定缩放因子被应用于x轴和y轴两者时,在S133和S134,x轴和y轴两者的缩放因子相等。

[0187] 根据示例性实施方式,当在S131确定当前块被编码为JMVD模式时,在S136,诸如scale_factor_dir的标志可以被用信号发送以指示缩放因子是被应用于MVD的x轴,y轴,还是x轴和y轴两者。然后,名为jmvd_scale_factor_flag的另一个标志被用信号发送以指示用于导出MVD的缩放因子的值。

[0188] 根据示例性实施方式,当在S131确定当前块被编码为JMVD模式时,诸如jmvd_scale_factor_flag的标志可以被用信号发送以指示用于导出MVD的缩放因子查找表中的索引。并且该缩放因子查找表中的每个条目指定x轴或y轴的缩放因子值。查找表中条目的顺序可以是固定的、根据统计或其他规则或者根据诸如模板匹配或双边匹配(即,基于模板匹配/基于双边匹配的重新排序)的解码器侧搜索算法预定义的。表6示出了缩放因子查找表的一个示例:

[0189] 表6:缩放因子查找表

[0190]

索引	x轴的缩放因子	y轴的缩放因子
0	1	1
1	1	2
2	1	1/2
3	2	1
4	2	2
5	2	1/2
6	1/2	1
7	1/2	2
8	1/2	1/2

[0191] 根据示例性实施方式,当在S131确定当前块被编码为JMVD模式时,诸如jmvd_scale_factor_equal_to_one的标志可以被编码并且指示联合MVD的缩放因子和方向。也就是说,如果标志jmvd_scale_factor_equal_to_one等于一个值(例如,1),则在S133将缩放

因子设置为等于1,并且在S134将该因子应用于水平(x轴)和竖直(y轴)两者。否则,如果标志jmvmd_scale_factor_equal_to_one等于另一个值(例如,0),则用信号发送附加jmvmd_scaling_dir语法元素。基于查找表(例如,表7),可以获得不同方向(轴)的缩放因子。注意,查找表中条目的顺序可以是固定的、根据统计或其他规则或者根据诸如模板匹配或双边匹配(即,基于模板匹配/基于双边匹配的重新排序)的解码器侧搜索算法预定义的。

[0192] 表7:缩放因子查找表

[0193]	jmvmd_scaling_dir	x 轴的缩放因子	y 轴的缩放因子
	0	1/2	0
[0194]	1	0	1/2
	2	2	0
	3	0	2

[0195] 根据示例性实施方式,用于用信号发送jmvmd_scale_factor_flag和/或scale_factor_dir的上下文取决于当前块或相邻块的编码信息,诸如当前块的块大小,或当前块的MVP,或当前块的编码模式是否是联合AMVD(或AMVR)编码模式,或相邻块的MVD,或相邻块的编码模式。并且缩放因子的值可以被限制为n的2次方,其中n可以是0或正整数或负整数。例如,缩放因子的值可以限制为{1,2},或者缩放因子的值可以限制为{1,1/2,2}。

[0196] 根据示例性实施方式,缩放因子的值可以是m/M,其中M是n的2次方,并且m是整数。并且缩放因子的值被限制为{1/8,2/8,3/8,4/8,……,15/8,16/8}。

[0197] 根据示例性实施方式,MVD的缩放处理与全部并入本文的于2022年4月6日提交的US 63/328,062中描述的方法相同。

[0198] 根据示例性实施方式,可以在S133基于编码信息导出缩放因子,编码信息包括但不限于量化步长或量化参数、块大小、MVP0与MVP1之间的差/关系、当前块的运动矢量预测块、MVD类别、参考图片、相邻块的MVD缩放因子。并且在S136,可以以高级语法用信号发送一种语法,以指示是否需要在比特流中用信号发送jmvmd_scale_factor_flag和/或scale_factor_dir。高级语法包括但不限于序列标头、帧标头和切片标头。

[0199] 并且如果在S131没有确定特定模式,则在S132可以根据其他模式发生上述任何其他处理,并且在S137可以类似地根据本文所描述的任何处理模式发生附加处理。另外,在到达处理S137之后在S130获得编码块可以以交互方式依赖于对先前编码块的先前处理。

[0200] 在图13的上述描述中,应当理解,一个或更多个操作可以以与所示的顺序交替的顺序执行和/或与彼此并行以及与其他操作并行地执行。

[0201] 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更多个计算机可读介质中,或者可以通过特别配置的一个或更多个硬件处理器来实现上述技术。例如,图14示出了适于实现所公开的主题的某些实施方式的计算机系统1400。

[0202] 该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,该机器代码或计算机语言可以经过汇编、编译、链接等机制来创建包括指令的代码,该指令可以由计算机中央处理单元(central processing unit,CPU)、图形处理单元(Graphics Processing

Unit, GPU) 等直接执行, 或者通过解释、微代码运行等来执行。

[0203] 指令可以在各种类型的计算机或其组件上执行, 包括例如个人计算机、平板计算机、服务器、智能电话、游戏装置、物联网装置等。

[0204] 图14中所示的用于计算机系统1400的组件在本质上是示例性的, 并且不旨在暗示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的任何限制。组件的配置也不应当被解释为具有与计算机系统1400的示例性实施方式中示出的组件中的任何一个组件或组件的组合有关的任何依赖性 or 要求。

[0205] 计算机系统1400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更多个人类用户通过例如触觉输入 (诸如, 键击、轻扫、数据手套移动)、音频输入 (诸如, 语音、拍手)、视觉输入 (诸如, 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有关的某些媒体, 诸如音频 (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)。

[0206] 输入人机接口装置可以包括以下中的一个或更多个 (每个仅描绘一个): 键盘1401、鼠标1402、触控板1403、触摸屏1410、操纵杆1405、麦克风1406、扫描仪1408、摄像装置1407。

[0207] 计算机系统1400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/味道来刺激一个或更多个人类用户的感官。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1410或者操纵杆1405的触觉反馈, 但也可以存在不充当输入装置的触觉反馈装置)、音频输出装置 (诸如, 扬声器1209、头戴式耳机 (未描绘))、视觉输出装置 (诸如, 包括CRT屏幕、LCD屏幕、等离子屏幕、OLED屏幕的屏幕1410, 各自具有或者不具有触摸屏输入能力, 各自具有或者不具有触觉反馈能力, 其中一些可以输出二维视觉输出或者通过诸如立体图形输出的方式输出多于三维输出; 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)。

[0208] 计算机系统1400还可以包括人类可访问的存储装置及其相关联的介质, 诸如包括具有CD/DVD 1411等介质的CD/DVD ROM/RW 1420的光学介质、拇指驱动器1422、可移动硬盘驱动器或固态驱动器1423、传统磁介质例如磁带和软盘 (未描绘)、基于专用ROM/ASIC/PLD的装置诸如安全软件狗 (未描绘) 等。

[0209] 本领域技术人员还应当理解, 结合目前公开内容的主题使用的术语“计算机可读介质”不包含传输介质、载波或者其他暂态信号。

[0210] 计算机系统1400还可以包括至一个或更多个通信网络1498的接口1499。网络1498可以例如是无线的、有线的、光的。网络1498还可以是局域的、广域的、城市的、车载的和工业的、实时的、延迟容忍的等。网络1498的示例包括局域网, 诸如以太网、无线LAN、蜂窝网络 (包括GSM、3G、4G、5G、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CANBus) 等。某些网络1498一般需要附接至某些通用数据端口或外围总线 (1450和1451) (诸如, 例如计算机系统1400的USB端口) 的外部网络接口适配器; 其他网络一般通过附接至如下所述的系统总线集成至计算机系统1400的核中 (例如, 集成至PC计算机系统以太网接口或集成至智能电话计算机系统蜂窝网络接口)。使用这些网络1498中的任何网络, 计算机系统1400可以与其他实体进行通信。这样的

通信可以是单向仅接收(例如,广播电视)、单向仅发送(例如,至某些CANbus装置的CANbus)、或双向的,例如至使用局域数字网络或广域数字网络的其他计算机系统。可以在如上所述的这些网络和网络接口的每个上使用某些协议和协议栈。

[0211] 上述人机接口装置、人类可访问的存储装置和网络接口可以附接至计算机系统1400的核1440。

[0212] 核1440可以包括一个或更多个中央处理单元(CPU) 1441、图形处理单元(GPU) 1442、图形适配器1417、现场可编程门阵列(Field Programmable Gate Area, FPGA) 1443形式的专用可编程处理单元、用于某些任务的硬件加速器1444等。这些装置连同只读存储器(Read-only memory, ROM) 1445、随机存取存储器1446、内部大容量存储装置诸如内部非用户可访问的硬盘驱动器、SSD等1447,可以通过系统总线1448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1448,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1451附接至核的系统总线1448。用于外围总线的架构包括PCI、USB等。

[0213] CPU 1441、GPU 1442、FPGA 1443和加速器1444可以执行某些指令,这些指令可以组合构成上述的计算机代码。该计算机代码可以存储在ROM 1445或RAM 1446中。过渡数据也可以存储在RAM 1446中,而永久数据可以存储在例如内部大容量存储装置1447中。可以通过使用高速缓冲存储器来实现针对任何存储器装置的快速存储以及检索,该高速缓冲存储器可以与一个或更多个CPU 1441、GPU 1442、大容量存储装置1447、ROM 1445、RAM 1446等紧密相关联。

[0214] 计算机可读介质上可以具有计算机代码,该计算机代码用于执行各种由计算机实现的操作。介质和计算机代码可以针对本公开内容的目的而专门设计和构建,或者介质和计算机代码可以是对计算机软件领域的技术人员公知且可用的类型。

[0215] 作为示例而非限制的方式,具有架构1400特别是核1440的计算机系统可以由于处理器(包括CPU、GPU、FPGA、加速器等)执行包含在一个或更多个有形计算机可读介质中的软件来提供功能。这样的计算机可读介质可以是与如上文介绍的用户可访问型大容量存储装置相关联的介质,以及具有非暂态性质的核1440的某些存储装置,诸如核内大容量存储装置1447或ROM 1445。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由核1440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器装置或芯片。软件可以使核1440——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处理或特定处理的特定部分,包括定义存储在RAM 1446中的数据结构以及根据软件定义的处理修改这样的数据结构。另外地或作为替选,计算机系统可以通过逻辑硬连线或以其他方式体现在电路(例如,加速器1444)中来提供功能,其可以代替软件或与软件一起操作,以执行本文描述的特定处理或特定处理的特定部分。在适当的情况下,对软件的提及可以包含逻辑,反之对逻辑的提及也可以包含软件。在适当的情况下,对计算机可读介质的提及可以包含存储用于执行的软件的电路(例如,集成电路(integrated circuit, 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与软件的任何合适的组合。

[0216] 尽管本公开内容已经描述了若干示例性实施方式,但是存在落入本公开内容的范围内的改变、置换和各种替换等效物。因此能够理解,本领域技术人员能够设想出多种系统和方法,这些系统和方法虽然在本文中并没有明确示出或描述,但体现公开内容的原理并且

因此落入公开内容的精神和范围内。

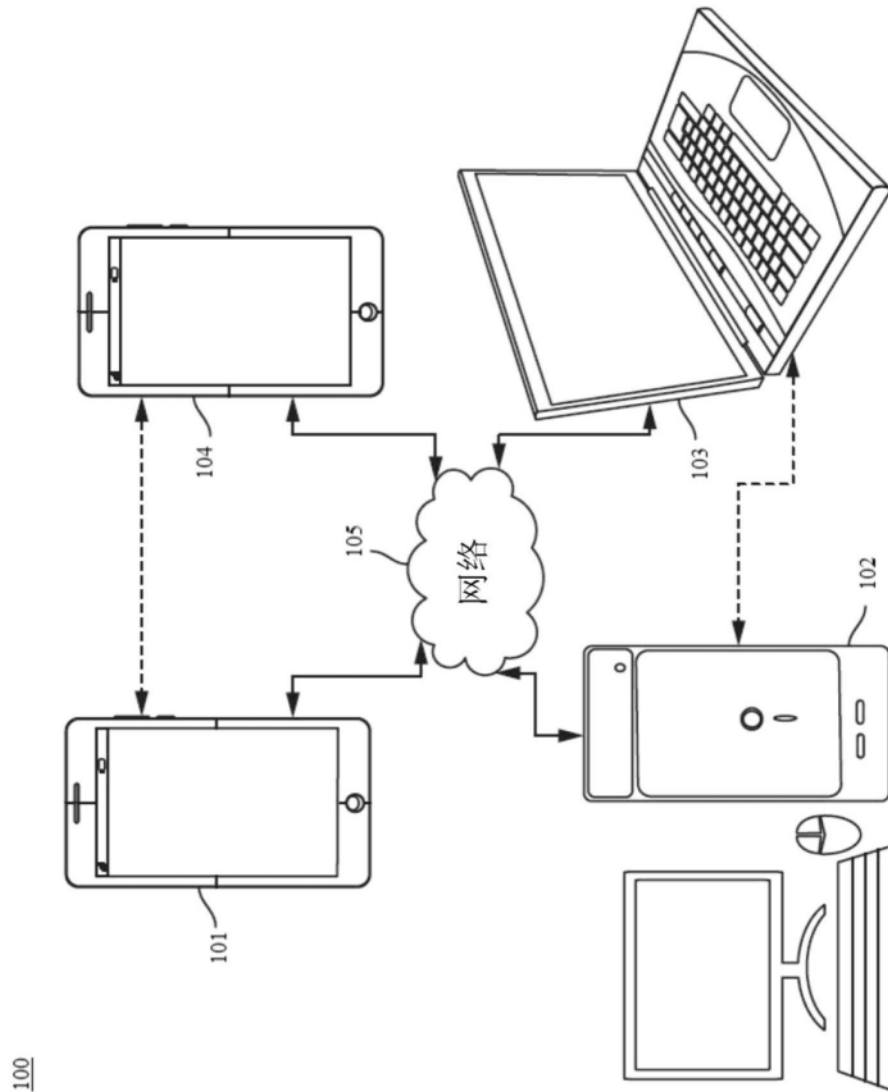


图1

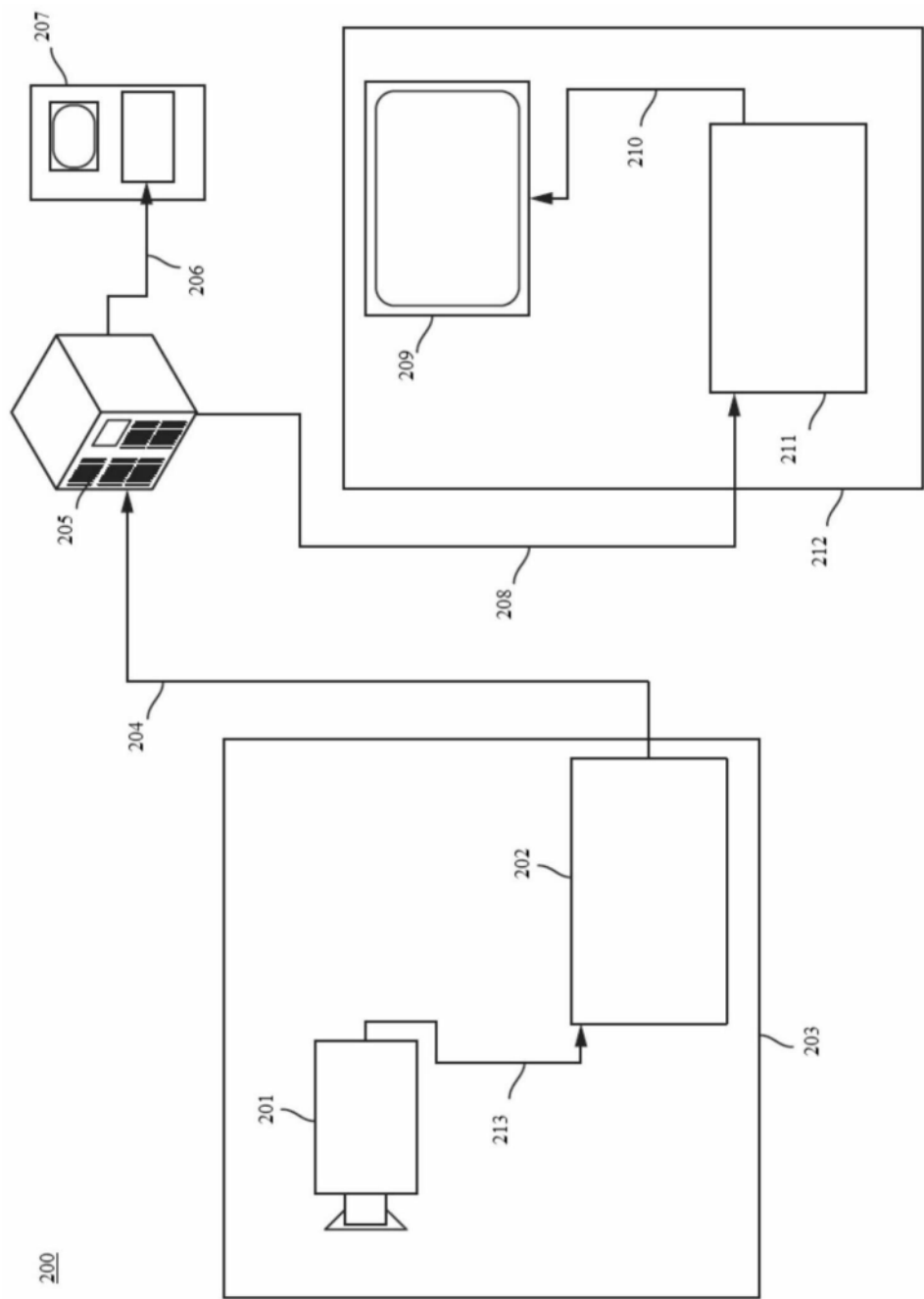


图2

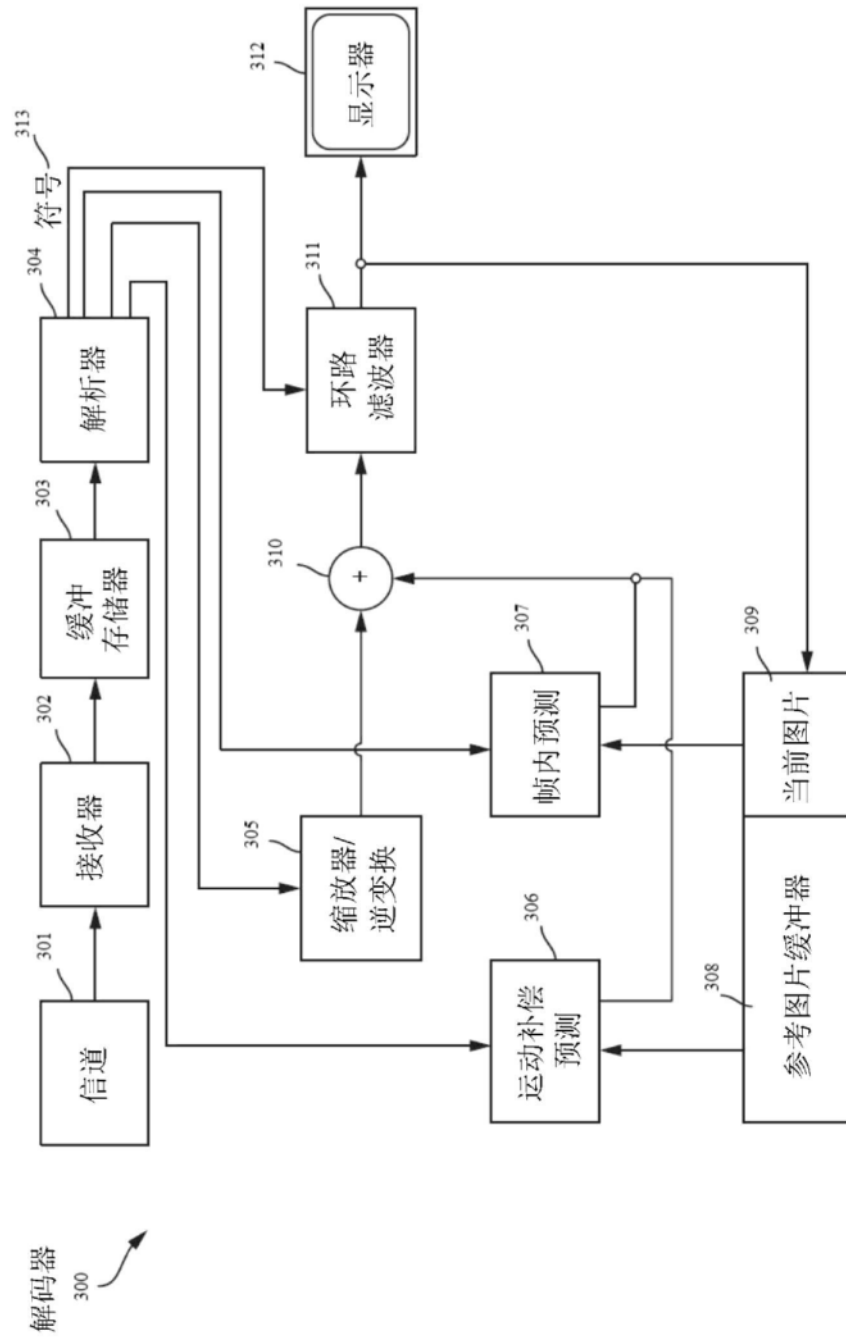


图3

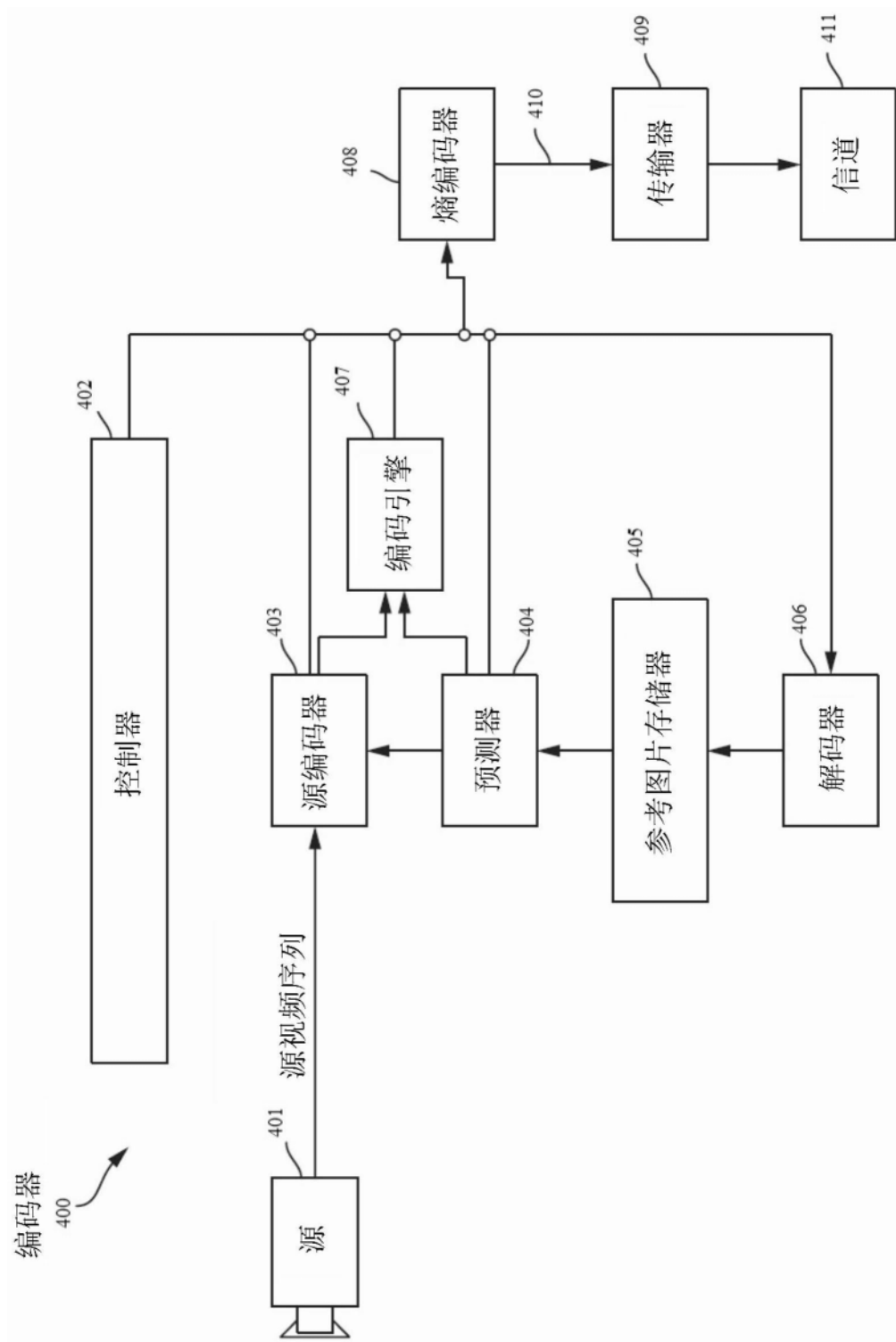


图4

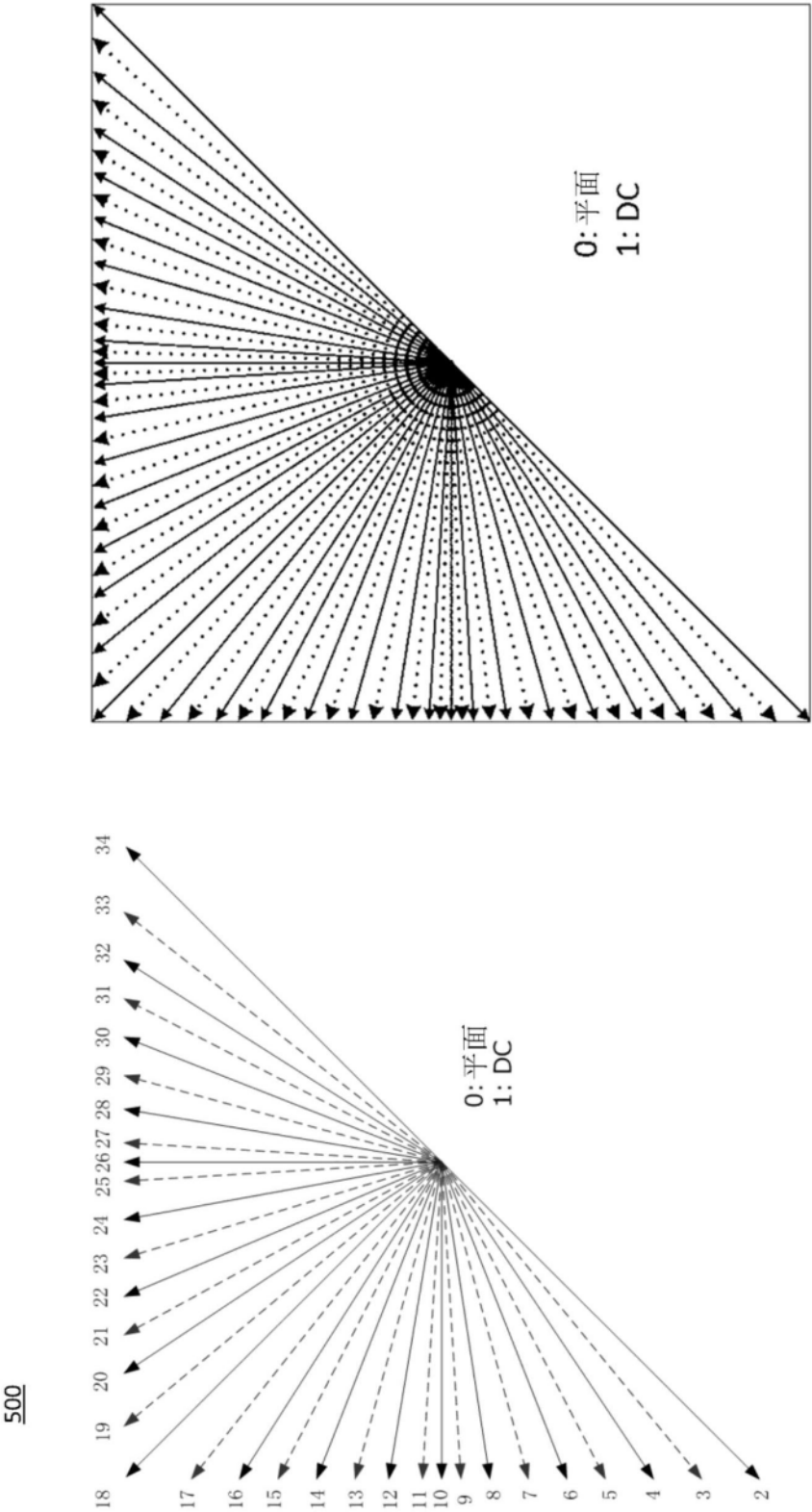


图5

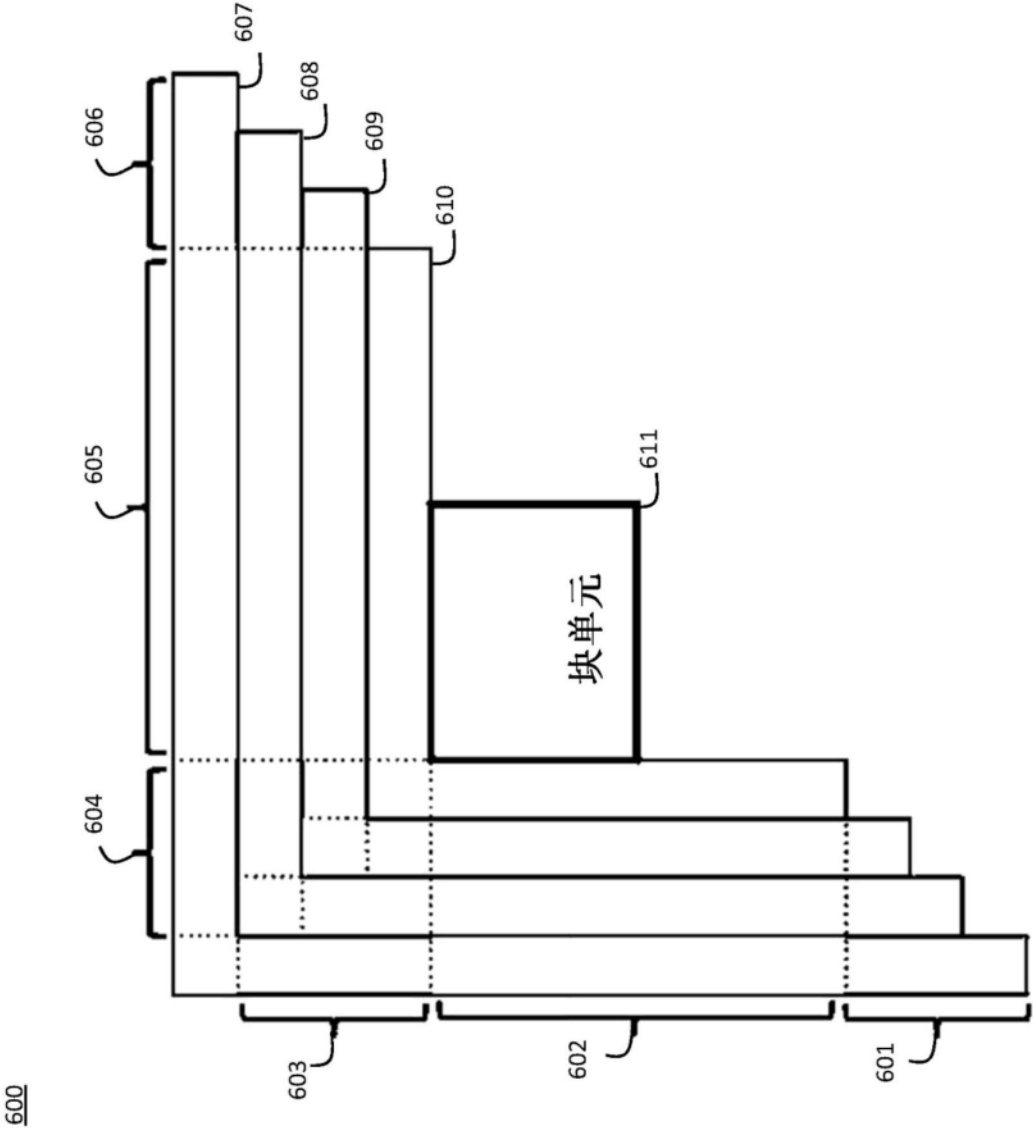


图6

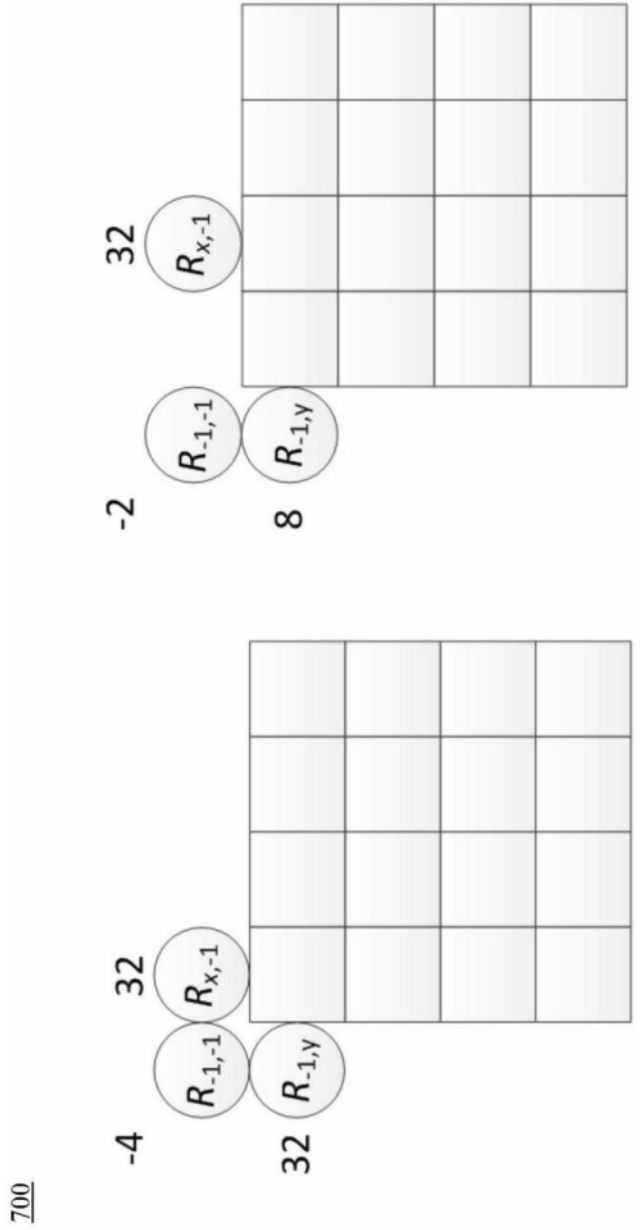


图7

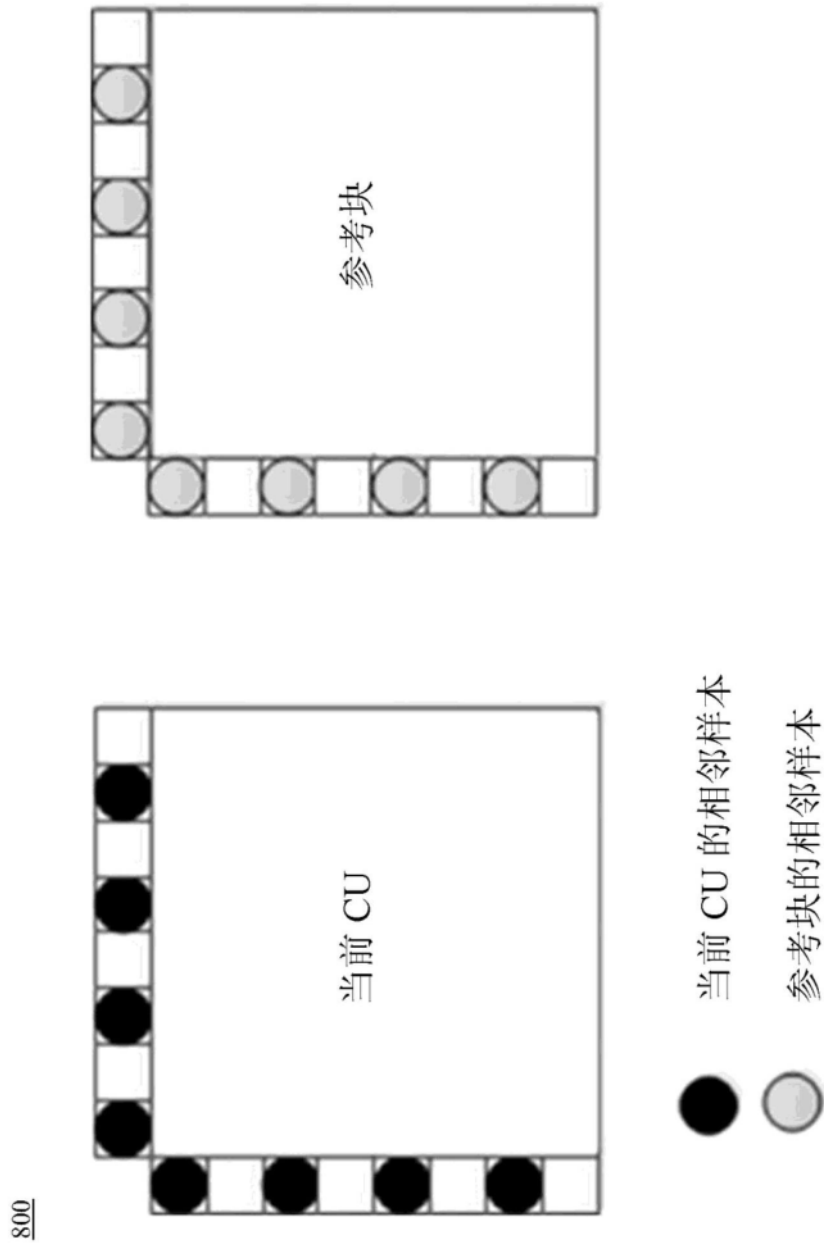


图8

900

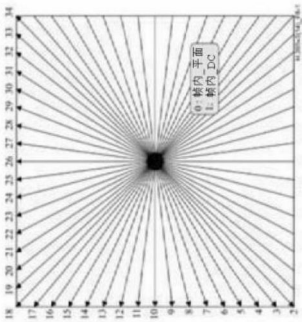


图9A

901

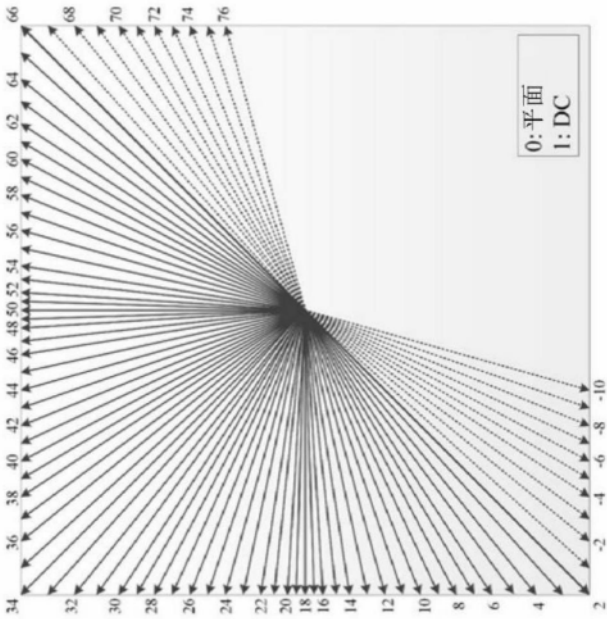


图9B

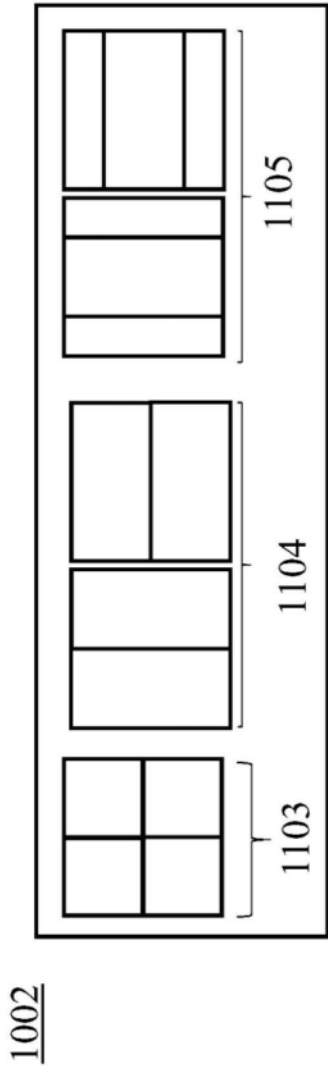


图10C

1100

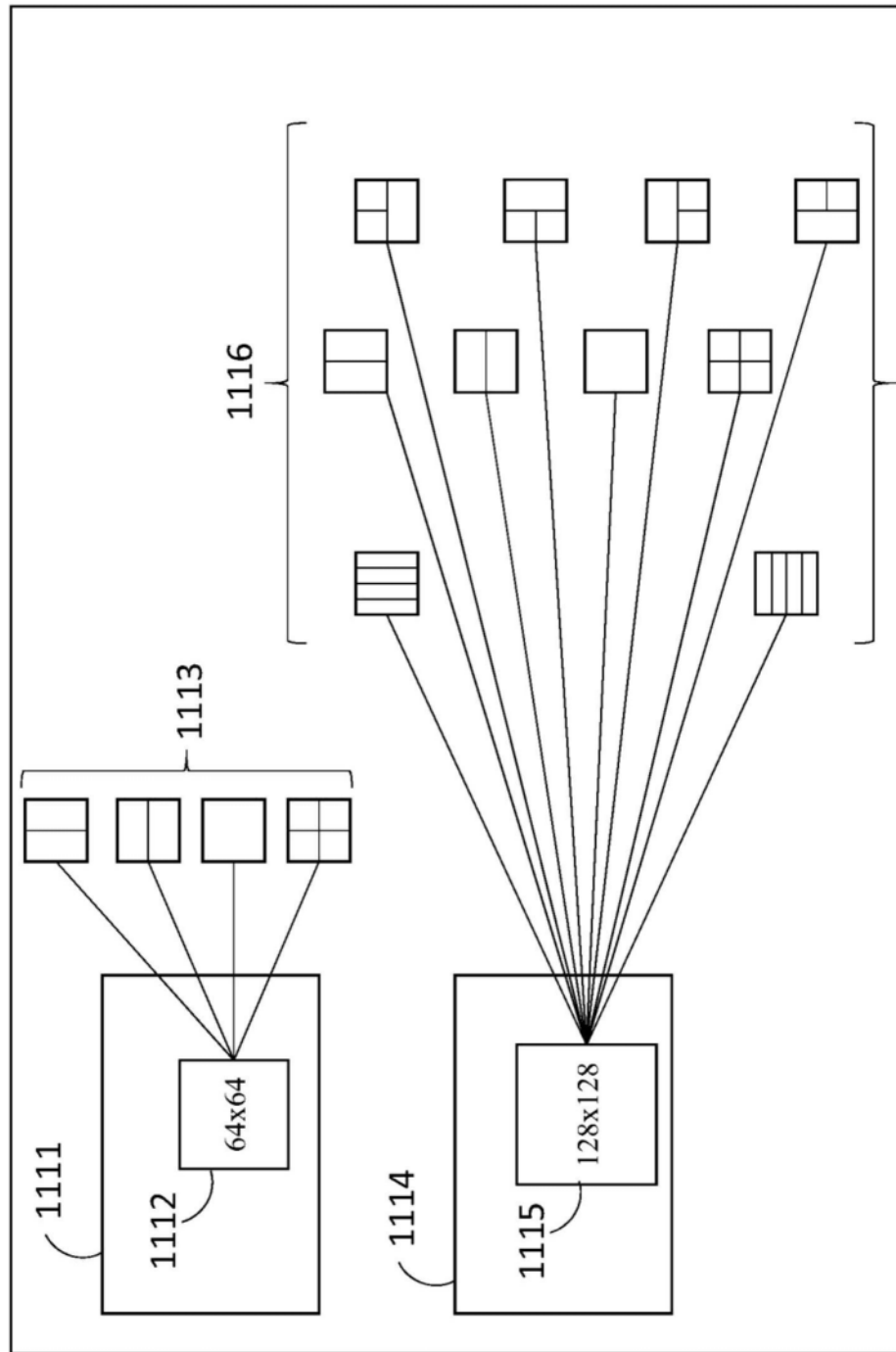


图11

1200

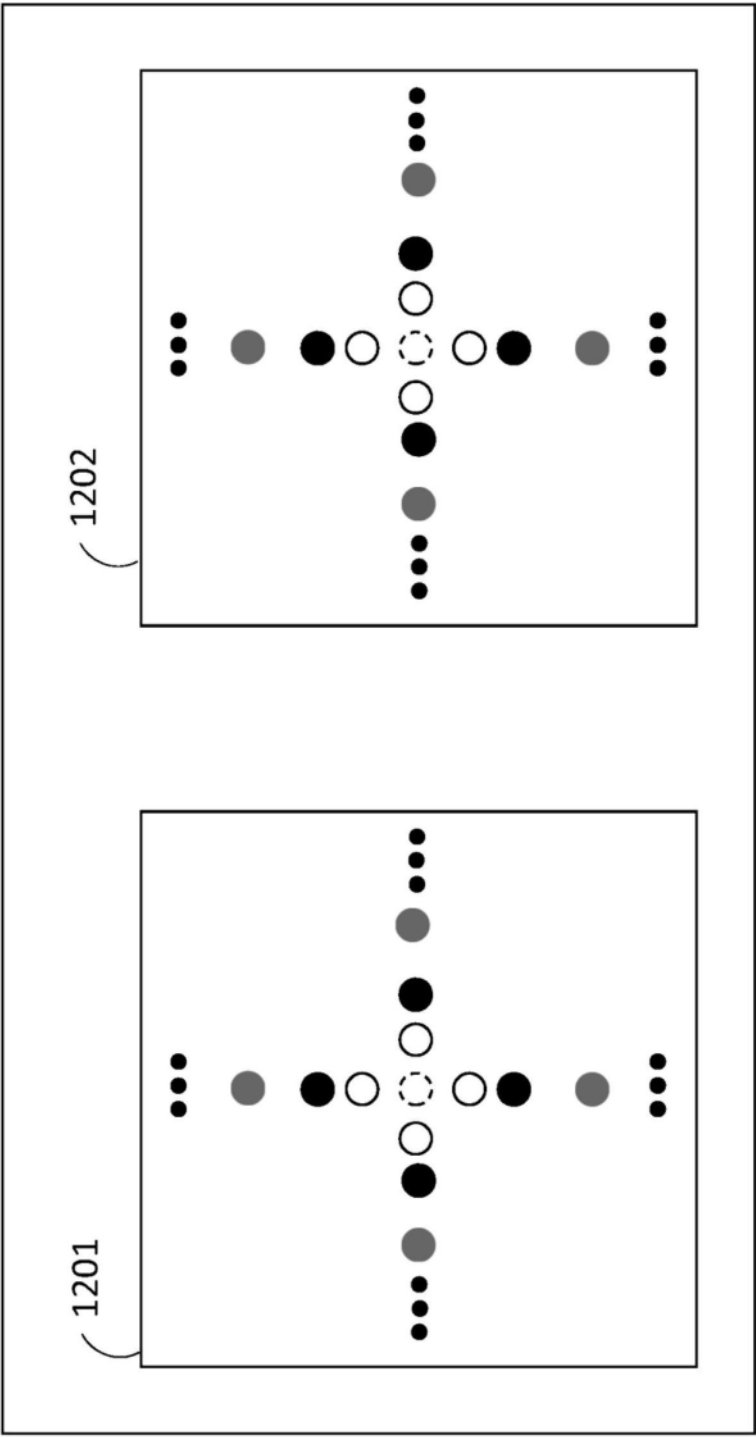


图12

1300

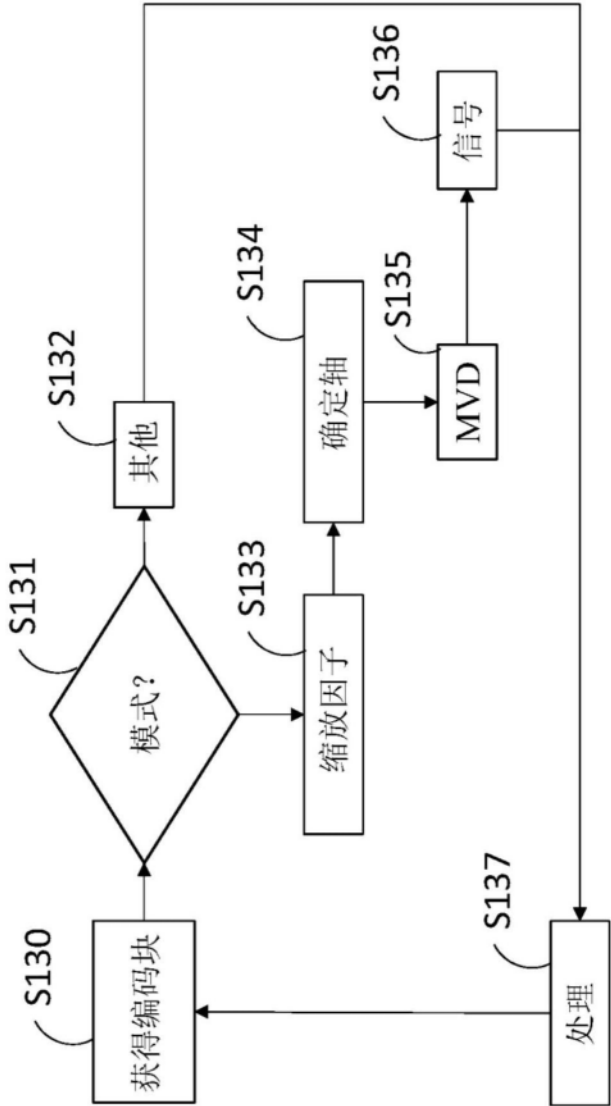


图13

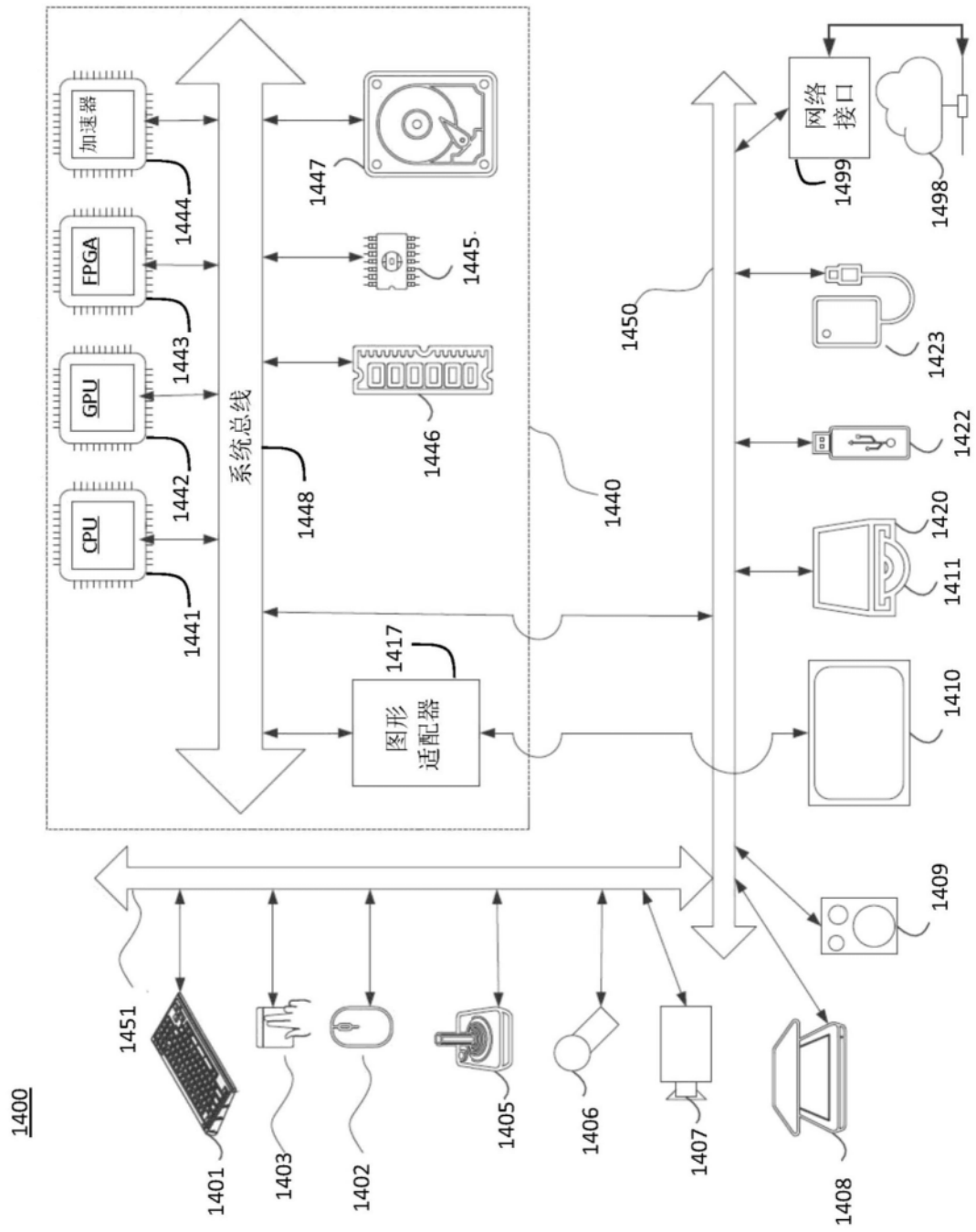


图14