(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119629340 A
(43)申请公布日 2025.03.14

(21)申请号 202411748761.5

(22)申请日 2020.07.08

(30)优先权数据
10-2019-0081926 2019.07.08 KR
10-2019-0170099 2019.12.18 KR

(62)分案原申请数据
202080049464.5 2020.07.08

(71)申请人 现代自动车株式会社
地址 韩国首尔
申请人 起亚株式会社
梨花女子大学校产学协力团

(72)发明人 姜制远 朴相孝 朴胜煜 林和平

(74)专利代理机构 北京戈程知识产权代理有限公司 11314
专利代理师 程伟 甄雁翔

(51)Int.Cl.
H04N 19/105(2014.01)
H04N 19/122(2014.01)
H04N 19/132(2014.01)
H04N 19/139(2014.01)
H04N 19/154(2014.01)
H04N 19/176(2014.01)
H04N 19/44(2014.01)
H04N 19/46(2014.01)
H04N 19/513(2014.01)
H04N 19/523(2014.01)
H04N 19/53(2014.01)
H04N 19/55(2014.01)
H04N 19/573(2014.01)
H04N 19/577(2014.01)
H04N 19/70(2014.01)

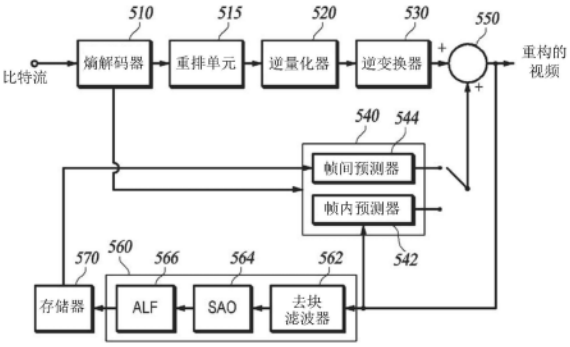
权利要求书3页 说明书25页 附图10页

(54)发明名称

一种视频解码方法、视频编码方法和提供与视频数据相关的比特流的方法

(57)摘要

本发明涉及一种视频解码方法、视频编码方法和提供与视频数据相关的比特流的方法。本实施方案基于更高级语法在编码器/解码器侧上使用编码工具以调整运动矢量的校正,其中,所述编码工具改善了帧间预测的压缩性能。另外,为了简单地执行以下步骤,提供了这样一种方法:整数样本搜索步骤,其用于搜索整数偏移,特别是与编码工具中的运动矢量细化相关的整数偏移;以及分数样本细化步骤,其用于搜索子像素偏移。



CN 119629340 A

1.一种视频解码装置,其用于对编码的图像序列逐块解码,所述视频解码装置包括至少一个处理器,所述至少一个处理器配置为:

从比特流解码至少一个第一高级语法元素,所述第一高级语法元素指示是否允许至少一个编码工具的每个,所述至少一个编码工具包括运动矢量细化;

根据至少一个第一高级语法元素,从比特流解码第二高级语法元素,并且在图像级别处设置是否允许至少一个编码工具的每个,所述第二高级语法元素与至少一个编码工具的每个相对应;

当基于图像级别处的设置识别出对于包括目标块的当前图像允许运动矢量细化时,通过细化从比特流获得初始运动矢量,生成目标块的细化的运动矢量;

利用双向参考图像和细化的运动矢量来生成目标块的帧间预测样本;

获得用于目标块的残差信号;

基于帧间预测样本和残差信号来重构目标块,

其中,生成细化的运动矢量包括:

整数样本搜索操作,其包括:

计算与由初始运动矢量指示的成对的初始位置以及从整数样本单元中的初始位置移位的成对的两个相邻位置相对应的成对的两个候选块的度量,其中,所述整数样本单元中的相邻位置是在每个初始位置周围具有正方形分布的24个样本位置,

在所述度量之中搜索与具有最小度量的位置相对应的整数偏移;

分数样本细化操作,其仅当整数偏移不位于正方形分布的最外侧时执行,并且所述分数样本细化操作包括:

利用具有最小度量的位置以及具有最小度量的位置的顶部、底部、左侧和右侧相邻位置,基于参数误差曲面方程来推导子像素偏移,

通过将子像素偏移与整数偏移相加来生成偏移,

通过将所述偏移应用于初始运动矢量来生成细化的运动矢量。

2.根据权利要求1所述的视频解码装置,其中,从序列头解码第一高级语法元素。

3.根据权利要求2所述的视频解码装置,其中,解码第二高级语法元素包括:

当第一高级语法元素指示出允许运动矢量细化时,从属于序列的每个图像的图像头解码指示是否允许运动矢量细化的图像级别标志。

4.根据权利要求3所述的视频解码装置,其中,当第一高级语法元素指示出不允许运动矢量细化时,图像级别标志设置为指示出不允许运动矢量细化的值,而无需从图像头解码所述图像级别标志。

5.根据权利要求1所述的视频解码装置,其中,当第一高级语法元素指示出不允许运动矢量细化时,对于序列中的任何图像,不允许用于运动矢量细化的编码工具。

6.根据权利要求1所述的视频解码装置,其中,当与初始运动矢量相对应的两个候选块之间的度量小于预设的第一阈值时,跳过整数样本搜索操作和分数样本细化操作两者。

7.根据权利要求1所述的视频解码装置,其中,当整数偏移为零时,跳过分数样本细化操作。

8.根据权利要求1所述的视频解码装置,其中,当最小度量与在整数样本搜索操作中的两个或更多个位置的整数坐标处获取的度量之间的每个差小于预设的第二阈值时,跳过分

数样本细化操作。

9.根据权利要求1所述的视频解码装置,其中,当目标块的初始运动矢量通过合并模式生成时,根据被认为是合并候选的相邻块的运动矢量的大小或应用于相邻块的编码模式来确定是执行整数样本搜索操作还是分数样本细化操作。

10.一种视频编码装置,其用于对编码的图像序列逐块编码,所述视频编码装置包括至少一个处理器,所述至少一个处理器配置为:

在序列头中编码至少一个第一高级语法元素,所述第一高级语法元素指示是否允许至少一个编码工具的每个,所述至少一个编码工具包括运动矢量细化;

根据至少一个第一高级语法元素编码图像级别的第二高级语法元素,所述图像级别的第二高级语法元素与是否允许至少一个编码工具的每个有关;

当基于图像级别的设置识别出对于包括目标块的当前图像允许运动矢量细化时,通过细化对于目标块所获得的初始运动矢量,生成目标块的细化的运动矢量;

利用双向参考图像和细化的运动矢量来生成目标块的帧间预测样本;

基于帧间预测样本来生成用于目标块的残差信号,并且对残差信号的信息进行编码,

其中,生成细化的运动矢量包括:

整数样本搜索操作,其包括:

计算与由初始运动矢量指示的成对的初始位置以及从整数样本单元中的初始位置移位的成对的两个相邻位置相对应的成对的两个候选块的度量,其中,所述整数样本单元中的相邻位置是在每个初始位置周围具有正方形分布的24个样本位置,

在所述度量之中搜索与具有最小度量的位置相对应的整数偏移;

分数样本细化操作,其仅当整数偏移不位于正方形分布的最外侧时执行,并且所述分数样本细化操作包括:

利用具有最小度量的位置以及具有最小度量的位置的顶部、底部、左侧和右侧相邻位置,基于参数误差曲面方程来推导子像素偏移,

通过将子像素偏移与整数偏移相加来生成偏移,

通过将所述偏移应用于初始运动矢量来生成细化的运动矢量。

11.一种提供与视频数据相关的比特流的装置,所述装置包括至少一个处理器,所述至少一个处理器配置为:

将视频数据编码至比特流中;

向视频解码装置发送比特流,

其中,所述处理器配置为通过执行以下处理对视频数据进行编码,所述处理包括:

在序列头中编码至少一个第一高级语法元素,所述第一高级语法元素指示是否允许至少一个编码工具的每个,所述至少一个编码工具包括运动矢量细化;

根据至少一个第一高级语法元素编码图像级别的第二高级语法元素,所述图像级别的第二高级语法元素与是否允许至少一个编码工具的每个有关;

当基于图像级别的设置识别出对于包括目标块的当前图像允许运动矢量细化时,通过细化对于目标块所获得的初始运动矢量,生成目标块的细化的运动矢量;

利用双向参考图像和细化的运动矢量来生成目标块的帧间预测样本;

基于帧间预测样本来生成用于目标块的残差信号,并且对残差信号的信息进行编码,

其中,生成细化的运动矢量包括:

整数样本搜索操作,其包括:

计算与由初始运动矢量指示的成对的初始位置以及从整数样本单元中的初始位置移位的成对的两个相邻位置相对应的成对的两个候选块的度量,其中,所述整数样本单元中的相邻位置是在每个初始位置周围具有正方形分布的24个样本位置,

在所述度量之中搜索与具有最小度量的位置相对应的整数偏移;

分数样本细化操作,其仅当整数偏移不位于正方形分布的最外侧时执行,并且所述分数样本细化操作包括:

利用具有最小度量的位置以及具有最小度量的位置的顶部、底部、左侧和右侧相邻位置,基于参数误差曲面方程来推导子像素偏移,

通过将子像素偏移与整数偏移相加来生成偏移,

通过将所述偏移应用于初始运动矢量来生成细化的运动矢量。

一种视频解码方法、视频编码方法和提供与视频数据相关的比特流的方法

[0001] 本申请是中国专利申请号为202080049464.5,发明名称为“一种视频解码方法、视频编码方法和提供与视频数据相关的比特流的方法”,申请日为2020年7月8日的进入中国的PCT专利申请的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求2019年7月8日提交的韩国专利申请No.10-2019-0081926和2019年12月18日提交的韩国专利申请No.10-2019-0170099的优先权,其全部内容通过引用合并于本文中。

技术领域

[0004] 本发明涉及视频的编码和解码,更具体地,本发明涉及配置为改善帧间预测的压缩性能的编码工具以及基于高级语法在编码器/解码器侧利用编码工具来调整运动矢量的校正的方法。

背景技术

[0005] 本部分中的陈述仅提供与本发明有关的背景技术信息,并不构成现有技术。由于视频数据量大于语音数据量或静止图像数据量,因此在不进行压缩处理的情况下存储或传输视频数据需要大量的硬件资源(包括存储器)。

[0006] 相应地,在存储或传输视频数据时,通常利用编码器来压缩视频数据以进行存储或传输。然后,解码器接收压缩的视频数据,解压并再现视频数据。用于这种视频的压缩技术包括H.264/AVC和高效率视频编码(High Efficiency Video Coding,HEVC),所述高效率视频编码(HEVC)比H.264/AVC的编码效率提高了大约40%。然而,图像大小、分辨率和帧速率逐渐增加,并且相应地,要编码的数据量也在增多。因此,需要一种与现有压缩技术相比具有更好的编码效率和更高的影像质量的新压缩技术。

[0007] 在视频编码中,预测编码主要用于改善压缩性能。存在用于基于当前图像中的预先重构的样本来预测要编码的目标块的帧内预测以及用于利用预先重构的参考图像来预测当前块的帧间预测。特别地,帧间预测由于相比帧内预测展现出更好的压缩性能而广泛用于视频编码。

[0008] 本发明提出了用于改善现有帧间预测的压缩性能的编码工具。

发明内容

[0009] 技术问题

[0010] 在本发明中,关于用于改善帧间预测的压缩性能的编码工具,基于高级语法在编码器/解码器侧利用编码工具来调整运动矢量的校正。本发明的主要目的是提供一种用于简单地执行与运动矢量细化的其他编码工具相关的整数样本搜索操作和分数样本细化操作的方法。

[0011] 根据本发明的一个方面,提供了一种视频解码的方法,用于通过视频解码装置对编码的图像序列逐块解码。所述方法包括:解码至少一个第一高级语法元素,所述第一高级语法元素指示是否允许来自比特流的至少一个编码工具的每个,所述至少一个编码工具包括运动矢量细化;根据至少一个第一高级语法元素,从比特流解码第二高级语法元素,并且在图像级别处设置是否允许至少一个编码工具的每个,所述第二高级语法元素与至少一个编码工具的每个相对应;当基于图像级别处的设置识别出对于包括目标块的当前图像允许运动矢量细化时,通过计算与目标块的双向参考图像中的初始运动矢量相距预设偏移范围内的两个候选块之间的度量,并且搜索具有最小度量的运动矢量候选,生成目标块的细化的运动矢量;以及利用双向参考图像和细化的运动矢量来生成目标块的帧间预测样本。

[0012] 根据本发明的另一个方面,提供了一种视频解码装置,用于对编码的图像序列逐块解码。所述装置包括解码器和预测器,所述解码器配置为从比特流解码至少一个第一高级语法元素,所述第一高级语法元素指示是否允许至少一个编码工具的每个,以及根据至少一个第一高级语法元素,从比特流解码第二高级语法元素,并且在图像级别处设置是否允许至少一个编码工具的每个,所述第二高级语法元素与至少一个编码工具的每个相对应,所述至少一个编码工具包括运动矢量细化;所述预测器配置为基于图像级别处的设置,响应于识别出对于包括目标块的当前图像允许运动矢量细化,通过计算与目标块的双向参考图像中的初始运动矢量相距预设偏移范围内的两个候选块之间的度量,并且搜索具有最小度量的运动矢量候选,生成目标块的细化的运动矢量,其中所述预测器利用双向参考图像和细化的运动矢量来生成目标块的帧间预测样本。

[0013] 根据本发明的另一个方面,提供了一种视频编码方法,用于通过视频编码装置对编码的图像序列逐块编码。所述方法包括:在序列头中编码至少一个第一高级语法元素,所述第一高级语法元素指示是否允许至少一个编码工具的每个,所述至少一个编码工具包括运动矢量细化;根据至少一个第一高级语法元素,编码图像级别的第二高级语法元素,所述图像级别的第二高级语法元素与是否允许至少一个编码工具的每个有关;当基于图像级别的设置识别出对于包括目标块的当前图像允许运动矢量细化时,通过计算与目标块的双向参考图像中的初始运动矢量相距预设偏移范围内的两个候选块之间的度量,并且搜索具有最小度量的运动矢量候选,生成目标块的细化的运动矢量;以及利用双向参考图像和细化的运动矢量来生成目标块的帧间预测样本。

[0014] 有益效果

[0015] 如从以上描述显然的是,根据本实施方案,提供了一种方法,用于在编码工具中简单地执行与运动矢量细化有关的整数样本搜索操作和分数样本细化操作,从而改善帧间预测的压缩性能。由此,可以提高编码器/解码器的编码效率,并且相应地可以降低复杂性。

附图说明

[0016] 图1是能够实现本发明技术的视频编码装置的示例性框图。

[0017] 图2示例性地示出了利用QTBTTT结构的块分区结构。

[0018] 图3a和图3b示例性地示出了包括宽角度帧内预测模式的多个帧内预测模式。

[0019] 图4示例性地示出了当前块周围的相邻块。

[0020] 图5是能够实现本发明的技术的视频解码装置的示例性框图。

[0021] 图6是示出根据本发明实施方案的BDOF的概念的示例图。

[0022] 图7是示出用于推导双向光流中的块边界样本的梯度的方法的示例图。

[0023] 图8a和图8b是示出根据本发明实施方案的仿射运动预测的示例图。

[0024] 图9是示出根据本发明实施方案的逐子块仿射运动预测的示例图。

[0025] 图10是示出用于从相邻块的平移运动矢量来推导用于仿射运动预测的合并候选的方法的示例图。

[0026] 图11是示出根据本发明实施方案的运动矢量细化编码工具的概念的示例图。

[0027] 图12是示出根据本发明实施方案的用于降低运动矢量细化编码工具的搜索复杂性的方法的示例图。

[0028] 图13是示出根据本发明的另一个实施方案的用于降低运动矢量细化编码工具的搜索复杂性的方法的示例图。

[0029] 图14是示出根据本发明的另一个实施方案的用于降低运动矢量细化编码工具的搜索复杂性的方法的示例图。

[0030] 附图标记

[0031] 124:帧间预测器

[0032] 155:熵编码器

[0033] 510:熵解码器

[0034] 544:帧间预测器。

具体实施方式

[0035] 在下文中,将参考所附附图对本发明的一些实施方案进行详细描述。应当注意,在将附图标记添加到各个附图中的组成元件时,尽管元件在不同的附图中示出,但是相同的附图标记也表示相同的元件。此外,在本发明实施方案的以下描述中,将省略并入本文中的已知功能和配置的详细描述,以避免模糊实施方案的主题。

[0036] 图1是能够实现本发明技术的视频编码装置的示例性框图。在下文中,将参考图1来描述视频编码装置以及该装置的元件。

[0037] 视频编码装置包括:图像分割器110、预测器120、减法器130、变换器140、量化器145、重排单元150、熵编码器155、逆量化器160、逆变换器165、加法器170、环路滤波单元180和存储器190。视频编码装置的每个元件可以以硬件或软件、或者硬件和软件的组合来实现。各个元件的功能可以实现为软件,并且微处理器可以实现为执行对应于各个元件的软件功能。

[0038] 一个视频由包括多个图像的一个或更多个序列组成。每个图像分割为多个区域,并且对每个区域执行编码。例如,一个图像分割为一个或更多个瓦片(tile)或/和切片(slice)。特别地,一个或更多个瓦片可以被定义为瓦片组。每个瓦片或切片分割为一个或更多个编码树单元(coding tree unit,CTU)。每个CTU通过树结构分割为一个或更多个编码单元(coding unit,CU)。应用于每个CU的信息被编码为CU的语法,并且共同应用于包括在一个CTU中的CU的信息被编码为CTU的语法。另外,共同应用于一个切片中的所有块的信息被编码为切片头的语法,而应用于构成一个或更多个图像的所有块的信息被编码在图像参数集(Picture Parameter Set,PPS)或图像头中。

[0039] 另外,由一个或更多个SPS共同参考的信息被编码在视频参数集(Video Parameter Set,VPS)中。此外,由多个图像共同参考的信息被编码在序列参数集(Sequence Parameter Set,SPS)中。此外,共同应用于一个瓦片或瓦片组的信息可以被编码为瓦片头或瓦片组头的语法。包括在SPS、PPS、切片头以及瓦片头或瓦片组头中的语法可以被称为高级语法。

[0040] 图像分割器110配置为确定编码树单元(CTU)的大小。关于CTU的大小(CTU尺寸)的信息被编码为SPS或PPS的语法,并且被传输至视频解码装置。图像分割器110将构成视频的每个图像分割为具有预定大小的多个CTU,然后利用树结构来递归地分割CTU。在树结构中,叶节点用作编码单元(CU),所述编码单元(CU)是编码的基本单元。

[0041] 树结构可以是四叉树(QuadTree,QT)、二叉树(BinaryTree,BT)、三叉树(TernaryTree,TT)、或由两个或更多个QT结构、BT结构和TT结构的组合形成的结构,所述四叉树(QT)即节点(或父节点)被分割为相同大小的四个从节点(或子节点),所述二叉树(BT)即节点被分割为两个从节点,所述三叉树(TT)即节点以1:2:1的比率被分割为三个从节点。例如,可以使用四叉树加二叉树(QuadTree plus BinaryTree,QTBT)结构,或者可以使用四叉树加二叉树三叉树(QuadTree plus BinaryTree TernaryTree,QTBTTT)结构。这里,BTTT可以统称为多类型树(multiple-type tree,MTT)。

[0042] 图2是示出用于利用QTBTTT结构来分割块的方法的示意图。如图2所示,CTU可以首先分割为QT结构。可以重复QT分割,直到分割块的大小达到QT中允许的叶节点的最小块大小MinQTSize。由熵编码器155对指示QT结构的每个节点是否被分割为下层的四个节点的第一标志(QT_split_flag)进行编码,并将其用信号通知视频解码装置。当QT的叶节点等于或小于BT中允许的根节点的最大块大小(MaxBTSize)时,可以进一步将其分割为一个或更多个BT结构或TT结构。BT结构和/或TT结构可以具有多个分割方向。例如,可以存在两个方向,即,水平地分割节点的块的方向和竖直地分割块的方向。如图2所示,当MTT分割开始时,通过熵编码器155对指示节点是否被分割的第二标志(mtt_split_flag)、指示分割情况下的分割方向(竖直或水平)的标志、和/或指示分割类型(二叉或三叉)的标志进行编码,并将其用信号通知视频解码装置。

[0043] 替选地,在对指示每个节点是否被分割为下层的4个节点的第一标志(QT_split_flag)进行编码之前,可以对指示节点是否被分割的CU分割标志(split_cu_flag)进行编码。当CU分割标志(split_cu_flag)的值指示出没有执行分割时,节点的块成为分割树结构中的叶节点,并用作编码单元(CU),编码单元(CU)是编码的基本单元。当CU分割标志(split_cu_flag)的值指示出执行分割时,视频编码装置开始以上述方式从第一标志起对标志编码。

[0044] 当利用QTBT作为树结构的另一个示例时,可以存在两种分割类型,即将块水平地分割为相同大小的两个块的类型(即,对称水平分割)和将块竖直地分割为相同大小的两个块的类型(即,对称竖直分割)。由熵编码器155对指示BT结构的每个节点是否被分割为下层的块的分割标志(split_flag)和指示分割类型的分割类型信息进行编码,并将其传输至视频解码装置。可以存在将节点的块分割为两个非对称块的额外类型。非对称分割类型可以包括以1:3的大小比率将块分割为两个矩形块的类型,或者对角线地分割节点的块的类型。

[0045] 另外,CU可以根据CTU的QTBT或QTBTTT分割而具有各种大小。在下文中,与要编码

或解码的CU(即,QTBTTT的叶节点)相对应的块被称为“当前块”。在采用QTBTTT分割时,当前块的形状可以是正方形或矩形。预测器120配置为对当前块进行预测以生成预测块。预测器120包括帧内预测器122和帧间预测器124。

[0046] 通常,图像中的每个当前块可以分别被预测地编码。另外,可以利用帧内预测技术(其利用来自包括当前块的图像的数据)或帧间预测技术(其利用在包括当前块的图像之前被编码的图像的数据)来执行当前块的预测。帧间预测包括单向预测和双向预测。

[0047] 帧内预测器122配置为利用在包括当前块的当前图像中位于当前块周围的像素(参考像素)来预测当前块中的像素。根据预测方向,存在多个帧内预测模式。例如,如图3a所示,多个帧内预测模式可以包括含有平面(planar)模式和直流(DC)模式的两种非方向模式以及65种方向模式。针对每种预测模式不同地定义要使用的相邻像素和等式。

[0048] 为了对矩形形状的当前块进行有效的方向预测,可以额外使用由图3b中的虚线箭头指示的方向模式(帧内预测模式67至80和-1至-14)。这些模式可以称为“宽角度帧内预测模式(wide angle intra-prediction modes)”。在图3b中,箭头指示用于预测的相应参考样本,而非指示预测方向。预测方向与由箭头指示的方向相反。宽角度帧内预测模式是在当前块具有矩形形状时,在与特定方向模式相反的方向上执行预测而无需额外的比特传输的模式。

[0049] 具体地,在宽角度帧内预测模式中,可以基于矩形当前块的宽度与高度的比率来确定可用于当前块的一些宽角度帧内预测模式。例如,在当前块的矩形形状的高度小于其宽度时,可以利用角度小于45度的宽角度帧内预测模式(帧内预测模式67至80)。在当前块的矩形形状的宽度大于其高度时,可以利用角度大于-135度的宽角度帧内预测模式(帧内预测模式-1至-14)。

[0050] 帧内预测器122可以配置为确定对当前块进行编码时要使用的帧内预测模式。在一些示例中,帧内预测器122可以配置为利用若干帧内预测模式来对当前块进行编码,并且从测试模式中选择要使用的适当的帧内预测模式。例如,帧内预测器122可以配置为利用若干测试的帧内预测模式的率失真(rate distortion)分析来计算率失真值,并且在测试的模式中选择具有最佳率失真特性的帧内预测模式。

[0051] 帧内预测器122配置为从多个帧内预测模式中选择一个帧内预测模式,并且利用根据选择的帧内预测模式确定的相邻像素(参考像素)和等式来预测当前块。由熵编码器155对关于选择的帧内预测模式的信息进行编码,并将其传输至视频解码装置。

[0052] 帧间预测器124配置为通过运动补偿来生成当前块的预测块。帧间预测器124配置为在比当前图像更早已被编码和解码的参考图像中搜索与当前块最相似的块,并且利用搜索到的块来生成当前块的预测块。然后,帧间预测器配置为生成与当前图像中的当前块和参考图像中的预测块之间的位移(displacement)相对应的运动矢量(motion vector,MV)。通常,对亮度(luma)分量执行运动估计,并且基于亮度分量计算的运动矢量用于亮度分量和色度分量两者。由熵编码器155对包括关于参考图像的信息和有关用于预测当前块的运动矢量的信息的运动信息进行编码,并将其传输至视频解码装置。

[0053] 帧间预测器124可以配置为在参考图像或参考块上执行插值,以增加预测准确性。换句话说,通过向包括两个整数样本的多个连续整数样本应用滤波器系数来在两个连续整数样本之间插值子样本。当对插值的参考图像执行搜索与当前块最相似的块的操作时,运

动矢量可以以分数样本单元的精度级表达,而不是以整数样本单元的精度级表达。

[0054] 对要编码的每个目标区域,例如,诸如切片、瓦片、CTU或CU的每个单元,可以不同地设置运动矢量的精度或分辨率。当应用这样的自适应运动矢量分辨率(adaptive motion vector resolution,AMVR)时,应当针对每个目标区域用信号通知关于要应用于每个目标区域的运动矢量分辨率的信息。例如,当目标区域是CU时,用信号通知关于应用于每个CU的运动矢量分辨率的信息。关于运动矢量分辨率的信息可以是指示运动矢量差的精度的信息,这将随后描述。

[0055] 帧间预测器124可以配置为利用双向预测来执行帧间预测。在双向预测中,帧间预测器124利用表示各个参考图像中与当前块最相似的块位置的两个运动矢量和两个参考图像。帧间预测器124配置为分别从参考图像列表0(RefPicList0)和参考图像列表1(RefPicList1)选择第一参考图像和第二参考图像,在各个参考图像中搜索与当前块相似的块,以及生成第一参考块和第二参考块。然后,帧间预测器124配置为通过对第一参考块和第二参考块进行平均或加权来生成用于当前块的预测块。然后,帧间预测器124配置为向编码器155传输包括关于用于预测当前块的两个参考图像和两个运动矢量的信息的运动信息。具体地,RefPicList0可以由在重构的图像中按显示顺序在当前图像之前的图像组成,并且RefPicList1可以由在重构的图像中按显示顺序在当前图像之后的图像组成。然而,实施方案不限于此。按显示顺序在当前图像之后的预先重构的图像可以进一步包括在RefPicList0中,相反地,在当前图像之前的预先重构的图像可以进一步包括在RefPicList1中。

[0056] 可以利用各种方法来使对运动信息进行编码所需的比特数最少。例如,在当前块的参考图像和运动矢量与相邻块的参考图像和运动矢量相同时,关于当前块的运动信息可以通过用于识别相邻块的编码信息发送到视频解码装置。这种方法称为“合并模式(merge mode)”。

[0057] 在合并模式中,帧间预测器124配置为从当前块的相邻块中选择预定数量的合并候选块(在下文中称为“合并候选”)。如图4所示,与当前图像中的当前块相邻的左侧块L、上侧块A、右上侧块AR、左下侧块BL和左上侧块AL的全部或部分可以用作用于推导合并候选的相邻块。另外,除了当前块所在的当前图像之外,位于参考图像(其可以与用于预测当前块的参考图像相同或不同)内的块可以用作合并候选。例如,在参考图像中与当前块处于相同位置处的同位块(co-located block)或与该同位块相邻的块可以额外地用作合并候选。

[0058] 帧间预测器124利用这样的相邻块来配置包括预定数量的合并候选的合并列表。帧间预测器124配置为从合并列表中包括的合并候选之中选择要用作关于当前块的运动信息的合并候选,并且生成用于识别选择的候选的合并索引信息。生成的合并索引信息由编码器155编码并传输至视频解码装置。

[0059] 对运动信息进行编码的另一种方法是高级运动矢量预测(advanced motion vector prediction,AMVP)模式。在AMVP模式中,帧间预测器124配置为通过利用当前块的相邻块来推导用于当前块的运动矢量的预测的运动矢量候选。在图2中的当前图像中,与当前块相邻的左侧块L、上侧块A、右上侧块AR、左下侧块BL和左上侧块AL的全部或部分可以用作用于推导预测的运动矢量候选的相邻块。

[0060] 此外,除了包括当前块的当前图像之外,位于参考图像(其可以与用于预测当前块

的参考图像相同或不同)内的块可以用作用于推导预测的运动矢量候选的相邻块。例如,可以利用在参考图像中与当前块处于相同位置处的同位块或与该同位块相邻的块。帧间预测器124配置为利用相邻块的运动矢量来推导预测的运动矢量候选,并且利用预测的运动矢量候选来确定用于当前块的运动矢量的预测的运动矢量。然后,通过将当前块的运动矢量减去预测的运动矢量来计算运动矢量差。

[0061] 可以通过将预定义函数(例如,用于计算中值、平均值等的函数)应用于预测的运动矢量候选来获得预测的运动矢量。特别地,视频解码装置还知道预定义的功能。由于用于推导预测的运动矢量候选的相邻块已经被编码和解码,所以视频解码装置也已经知道相邻块的运动矢量。相应地,视频编码装置不需要对用于识别预测的运动矢量候选的信息进行编码。因此,在这种情况下,对关于运动矢量差的信息和关于用于预测当前块的参考图像的信息进行编码。

[0062] 可以通过选择预测的运动矢量候选的任何一个来确定预测的运动矢量。具体地,用于识别选择的预测的运动矢量候选的信息与要用于预测当前块的关于运动矢量差的信息以及关于参考图像的信息一起被进一步编码。减法器130配置为将当前块减去由帧内预测器122或帧间预测器124生成的预测块以生成残差块。

[0063] 变换器140配置为将空域中具有像素值的残差块中的残差信号变换为频域中的变换系数。变换器140可以配置为利用残差块的整个大小作为变换单元来变换残差块中的残差信号。替选地,变换器140可以配置为将残差块分割为多个子块,并且利用子块作为变换单元来执行变换。替选地,可以通过将块分为两个子块,即变换区域和非变换区域,并且仅利用变换区域子块作为变换单元来变换残差信号。

[0064] 特别地,变换区域子块可以是基于水平轴线(或竖直轴线)的具有大小比率为1:1的两个矩形块的一个。在这种情况下,由熵编码器155对指示仅变换了子块的标志(cu_sbt_flag)、方向(竖直/水平)信息(cu_sbt_horizontal_flag)和/或位置信息(cu_sbt_pos_flag)进行编码,并且将其用信号通知视频解码装置。另外,变换区域子块的大小可以具有基于水平轴线(或竖直轴线)1:3的大小比率。特别地,由熵编码器155对区分相应分割的标志(cu_sbt_quad_flag)进行额外地编码,并且将其用信号通知视频解码装置。

[0065] 变换器140可以配置为在水平方向和竖直方向上单独地变换残差块。对于变换,可以利用各种类型的变换函数或变换矩阵。例如,用于水平变换和竖直变换的成对变换函数可以被定义为多变换集(multiple transform set,MTS)。变换器140可以配置为选择MTS中具有最佳变换效率的一对变换函数,并且分别在水平方向和竖直方向上变换残差块。由熵编码器155对关于在MTS中选择的变换函数对的信息(mts_idx)进行编码,并且将其用信号通知视频解码装置。

[0066] 量化器145配置为利用量化参数对从变换器140输出的变换系数进行量化,并且将量化的变换系数输出到熵编码器155。对于一些块或帧,量化器145可以配置为不经变换直接对相关残差块进行量化。量化器145可以配置为根据变换块中的变换系数的位置来应用不同的量化系数(缩放值)。可以对应用于二维布置的量化的变换系数的量化的系数的矩阵进行编码,并且将其用信号通知视频解码装置。

[0067] 重排单元150可以配置为对用于量化的残差值的系数值进行重新分类。重排单元150可以配置为通过系数扫描(coefficient scanning)将2维系数阵列改变为1维系数序

列。例如,重排单元150可以配置为利用锯齿形扫描(zig-zag scan)或对角线扫描(diagonal scan)从直流(DC)系数向高频区域中的系数对系数进行扫描,以输出1维系数序列。根据变换单元的大小和帧内预测模式,可以利用竖直扫描或水平扫描代替锯齿形扫描,所述竖直扫描即在列方向上对二维系数阵列进行扫描,所述水平扫描即在行方向上对二维块形状的系数进行扫描。换句话说,可以根据变换单元的大小和帧内预测模式在锯齿形扫描、对角线扫描、竖直扫描和水平扫描中确定要使用的扫描模式。

[0068] 熵编码器155配置为利用诸如基于上下文的自适应二进制算术编码(Context-based Adaptive Binary Arithmetic Code,CABAC)和指数哥伦布(exponential Golomb)的各种编码技术对从重排单元150输出的一维量化的变换系数进行编码,以生成比特流。熵编码器155配置为对与块分割相关的信息(例如,CTU大小、CU分割标志、QT分割标志、MTT分割类型和MTT分割方向)进行编码,使得视频解码装置可以以与视频编码装置相同的方式来分割块。

[0069] 另外,熵编码器155配置为对关于指示当前块是由帧内预测编码进行编码还是由帧间预测进行编码的预测类型的信息进行编码,并且根据预测类型对帧内预测信息(即,关于帧内预测模式的信息)或帧间预测信息(用于合并模式的合并索引,关于用于AMVP模式的参考图像索引和运动矢量差的信息)进行编码。熵编码器155还配置为对与量化有关的信息(即,关于量化参数的信息和关于量化矩阵的信息)进行编码。

[0070] 逆量化器160配置为对从量化器145输出的量化的变换系数进行逆量化,以生成变换系数。逆变换器165配置为将从逆量化器160输出的变换系数从频域变换到空域,并且重构残差块。加法器170配置为将重构的残差块和由预测器120生成的预测块相加,以重构当前块。重构的当前块中的像素在执行后续块的帧内预测时用作参考像素。

[0071] 环路滤波单元180配置为对重构的像素进行滤波,以减少由于基于块的预测和变换/量化而产生的块伪影(blocking artifacts)、振铃伪影(ringing artifacts)和模糊伪影(blurring artifacts)。环路滤波单元180可以包括去块滤波器182、采样自适应偏移(sample adaptive offset,SAO)滤波器184和自适应环路滤波器(adaptive loop filter,ALF)186的至少一个。去块滤波器182配置为对重构的块之间的边界进行滤波,以去除由逐块编码/解码引起的块伪影,并且SAO滤波器184配置为对去块滤波的视频执行额外的滤波。

[0072] 具体地,SAO滤波器184是用于对由有损编码(lossy coding)引起的重构的样本与原始的样本之间的差进行补偿的滤波器,并且以将相对应的偏移和每个重构的样本相加的方式来执行滤波。ALF 186配置为通过对目标样本和目标样本的相邻样本应用滤波器系数来对要滤波的目标样本执行滤波。ALF 186可以配置为将包括在图像中的样本划分为预定组,然后确定要应用于相应组的一个滤波器,以对每个组差分地执行滤波。可以对关于要用于ALF的滤波器系数的信息进行编码,并且将其用信号通知视频解码装置。通过环路滤波单元180滤波的重构的块存储在存储器190中。一旦一个图像中的所有块被重构,重构的图像就可以用作后续要被编码的图像中的块的帧间预测的参考图像。

[0073] 图5是能够实现本发明的技术的视频解码装置的示例性功能框图。在下文中,将参考图5来描述视频解码装置和该装置的元件。

[0074] 视频解码装置可以包括:熵解码器510、重排单元515、逆量化器520、逆变换器530、预测器540、加法器550、环路滤波单元560和存储器570。类似于图1的视频编码装置,视频解

码装置的每个元件可以用硬件、软件、或硬件和软件的组合来实现。此外,每个元件的功能可以用软件来实现,并且微处理器可以实现为执行对应于每个元件的软件功能。

[0075] 熵解码器510配置为通过对由视频编码装置生成的比特流解码并提取与块分割相关的信息来确定要解码的当前块,并且提取重构当前块所需的预测信息和关于残差信号的信息等。熵解码器510配置为从序列参数集(SPS)或图像参数集(PPS)中提取关于CTU大小的信息,确定CTU的大小,并且将图像分割为确定大小的CTU。然后,解码器配置为将CTU确定为树结构的最高层(即,根节点),并且提取关于CTU的分割信息,以利用树结构来分割CTU。

[0076] 例如,当利用QTBTTT结构来分割CTU时,提取与QT的分割相关的第一标志(QT_split_flag),以将每个节点分割为子层的四个节点。对于与QT的叶节点相对应的节点,提取与MTT的分割有关的第二标志(MTT_split_flag)以及关于分割方向(竖直/水平)和/或分割类型(二叉/三叉)的信息,从而以MTT结构分割对应的叶节点。由此,以BT或TT结构来递归地分割QT的叶节点下方的每个节点。

[0077] 作为另一个示例,当利用QTBTTT结构来分割CTU时,可以提取指示是否分割CU的CU分割标志(split_cu_flag)。当分割对应块时,可以提取第一标志(QT_split_flag)。在分割操作中,在零个或更多个递归QT分割之后,每个节点都可能发生零个或更多个递归MTT分割。例如,CTU可以直接经历MTT分割而不经历QT分割,或者仅经历QT分割多次。作为另一个示例,当利用QTBT结构来分割CTU时,提取与QT分割相关的第一标志(QT_split_flag),并且将每个节点分割为下层的四个节点。然后,提取指示是否以BT进一步分割与QT的叶节点相对应的节点的分割标志(split_flag)以及分割方向信息。

[0078] 一旦通过树结构进行分割来确定要解码的当前块,熵解码器510就配置为提取关于指示当前块是被帧内预测还是被帧间预测的预测类型的信息。当预测类型信息指示帧内预测时,熵解码器510配置为提取用于当前块的帧内预测信息(帧内预测模式)的语法元素。当预测类型信息指示帧间预测时,熵解码器510配置为提取用于帧间预测信息的语法元素,即指示运动矢量和由运动矢量参考的参考图像的信息。

[0079] 熵解码器510还配置为提取关于量化的当前块的变换系数的信息作为与量化有关的信息和关于残差信号的信息。重排单元515可以配置为以由视频编码装置执行的系数扫描的相反顺序,将由熵解码器510熵解码的一维量化的变换系数的序列改变为2维系数阵列(即,块)。

[0080] 逆量化器520配置为对量化的变换系数进行逆量化,并且利用量化参数对量化的变换系数进行逆量化。逆量化器520可以配置为向布置为二维的量化的变换系数应用不同的量化系数(缩放值)。逆量化器520可以配置为通过将来自视频编码装置的量化系数(缩放值)矩阵应用于二维量化的变换系数阵列来执行逆量化。逆变换器530配置为将逆量化的变换系数从频域逆变换到空域,以重构残差信号,从而生成当前块的残差块。

[0081] 另外,当逆变换器530仅对变换块的部分区域(子块)进行逆变换时,逆变换器530提取指示出仅变换块的子块已进行变换的标志(cu_sbt_flag)、关于子块的方向性(竖直/水平)信息(cu_sbt_horizontal_flag)、和/或关于子块的位置信息(cu_sbt_pos_flag),并且将子块的变换系数从频域逆变换到空域。然后,逆变换器530重构残差信号,并且用“0”值填充未逆变换的区域作为残差块,从而生成当前块的最终残差块。

[0082] 另外,当应用MTS时,逆变换器530配置为利用从视频编码装置用信号通知的MTS信

息(mts_idx)来确定要分别在水平方向和竖直方向上应用的变换函数或变换矩阵,并且利用确定的变换函数在水平方向和竖直方向上对变换块中的变换系数进行逆变换。预测器540可以包括帧内预测器542和帧间预测器544。在当前块的预测类型是帧内预测时,激活帧内预测器542,而在当前块的预测类型是帧间预测时,激活帧间预测器544。

[0083] 帧内预测器542配置为基于从熵解码器510提取的帧内预测模式的语法元素,在多个帧内预测模式中确定当前块的帧内预测模式,并且根据帧内预测模式,利用当前块周围的参考像素来预测当前块。帧间预测器544配置为利用从熵解码器510提取的帧内预测模式的语法元素来确定当前块的运动矢量和由该运动矢量参考的参考图像,并且基于运动矢量和参考图像来预测当前块。

[0084] 加法器550配置为通过将从逆变换器输出的残差块与从帧间预测器或帧内预测器输出的预测块相加来重构当前块。在对后续要被解码的块进行帧内预测时,重构的当前块中的像素用作参考像素。环路滤波单元560可以包括去块滤波器562和SAO滤波器564。去块滤波器562对重构的块之间的边界进行去块滤波,以去除由逐块解码引起的块伪影。

[0085] SAO滤波器564配置为在去块滤波之后对重构的块执行额外滤波,以补偿由有损编码引起的重构的像素与原始的像素之间的差。ALF 566配置为通过对目标样本和目标样本的相邻样本应用滤波器系数来对要滤波的目标样本执行滤波。ALF566可以配置为将图像中的样本划分为预定组,然后确定要应用于相应组的一个滤波器,以对每个组差分地执行滤波。基于关于从比特流解码的滤波器系数的信息来确定ALF的滤波器系数。通过环路滤波单元560滤波的重构的块存储在存储器570中。当一个图像中的所有块被重构时,重构的图像用作对后续要被编码的图像中的块进行帧间预测的参考图像。

[0086] 本实施方案涉及如上所述的视频的编码和解码。更具体地,关于改善帧间预测的压缩性能的编码工具,编码器/解码器基于高级语法利用编码工具来调整运动矢量的校正。此外,本实施方案提供了一种方法,用于简单地执行编码工具中与运动矢量细化有关的搜索整数偏移的整数样本搜索操作以及搜索子像素偏移的分数样本细化操作。

[0087] 以下实施方案涉及用于改善帧间预测的压缩性能的编码工具,其可以由视频编码装置的帧间预测器124和视频解码装置的帧间预测器544来操作。如本文中所使用的,术语“目标块”可以具有与以上所述的当前块或编码单元(CU)相同的含义,或者可以表示CU的局部区域。

[0088] I.双向光流(Bi-directional Optical Flow,BDOF)

[0089] 双向光流是在假定构成视频的样本或物体以恒定的速度移动并且样本值几乎没有变化的情况下,对利用双向运动预测而预测的样本的运动进行额外补偿的技术。图6是示出根据本发明实施方案的BDOF的概念的示例性图。

[0090] 假设通过用于目标块的(正常的)双向运动预测,在参考图像Ref_0和Ref_1中已经确定了指向与当前图像中的要编码的目标块最相似的相应区域(即,参考块)的双向运动矢量MV_0和MV_1。这两个运动矢量具有表示整个目标块的运动的值。

[0091] 在图6的示例中,P_0是由运动矢量MV_0指示的并与目标块中的样本P相对应的参考图像Ref_0中的样本,而P_1是由运动矢量MV_1指示的并与目标块中的样本P相对应的参考图像Ref_1中的样本。此外,假设图6中样本P的运动与目标块的整体运动略有不同。例如,当位于图6的Ref_0中的样本A处的物体经由当前图像的目标块中的样本P向Ref_1中的样本B移动时,

样本A和样本B可以具有彼此非常相似的值。相应地,在这种情况下,Ref_0中与目标块中的样本P最相似的点不是由运动矢量MV_0指示的P_0,而是从P_0已经移位了预定位移矢量$(v_x t_0, v_y t_1)$的样本A。Ref_1中与目标块中的样本P最相似的点不是由运动矢量MV_1指示的P_1,而是从P_1已经移位了预定位移矢量$(-v_x t_0, -v_y t_1)$的样本B。具体地,t_0和t_1分别表示Ref_0和Ref_1相对于当前图像的时间轴距离,并且是基于图像顺序计数(picture order count,POC)来计算的。在下文中,(v_x, v_y)被称为“光流”。

[0092] 在预测当前图像中的当前块的样本P的值时,两个参考样本A和B能够实现比由双向运动矢量MV_0和MV_1指示的参考样本P_0和P_1更准确的预测。$I^{(0)}(i,j)$表示如与目标块中的样本(i,j)相对应的由运动矢量MV_0指示的参考图像Ref_0中的样本的值,而$I^{(1)}(i,j)$表示如与目标块中的样本(i,j)相对应的由运动矢量MV_1指示的参考图像Ref_1中的样本的值。

[0093] 如与目标块中的样本相对应的由BDOF运动矢量(v_x, v_y)指示的参考图像Ref_0中的样本A的值可以定义为$I^{(0)}(i+v_x t_0, j+v_y t_0)$,并且参考图像$Ref_1$中的样本B的值可以定义为$I^{(1)}(i-v_x t_1, j-v_y t_1)$。特别地,当仅利用泰勒级数的一阶项执行线性逼近时,A和B可以表达为等式1。

[0094] 等式1

[0095]
$$\begin{aligned} A &= I^{(0)}(i+v_x\tau_0, j+v_y\tau_0) \\ &\approx I^{(0)}(i,j) + v_x\tau_0 I_x^{(0)}(i,j) + v_y\tau_0 I_y^{(0)}(i,j) \\ B &= I^{(1)}(i-v_x\tau_1, j-v_y\tau_1) \\ &\approx I^{(1)}(i,j) - v_x\tau_1 I_x^{(1)}(i,j) - v_y\tau_1 I_y^{(1)}(i,j) \end{aligned}$$

[0096] 其中,$I_x^{(k)}$和$I_y^{(k)}$(k=0,1)是在Ref_0和Ref_1的位置(i,j)处在水平方向和竖直方向上的梯度值。如上所述,t_0和t_1分别表示相对于当前图像的Ref_0和Ref_1的时间轴距离,并且基于POC表示为t_0=POC(当前)-POC(Ref_0),和t_1=POC(Ref_1)-POC(当前)。

[0097] 块中每个样本的双向光流(v_x, v_y)确定为使Δ最小化的解,所述Δ定义为样本A与样本B之间的差。可以利用从等式1推导的A和B的线性近似由等式2来定义Δ。

[0098] 等式2

[0099]
$$\begin{aligned} \Delta &= A - B \\ &= \left(I^{(0)} - I^{(1)}\right) + v_x\left(\tau_0 I_x^{(0)} + \tau_1 I_x^{(1)}\right) + v_y\left(\tau_0 I_y^{(0)} + \tau_1 I_y^{(1)}\right) \end{aligned}$$

[0100] 为了简单起见,从等式2的每个项中省略样本位置(i,j)。

[0101] 为了实现更稳健的光流估计,假设运动与相邻样本局部一致。考虑到在以样本(i,j)为中心的一定大小的掩膜Ω中存在的所有样本(i',j')的等式2中的差Δ,计算当前要预测的样本(i,j)的BDOF运动矢量。如等式3所示,当前样本(i,j)的光流可以确定为使掩膜Ω中为各个样本获得的差Δ[i',j']的平方和最小化的矢量,如等式3所示。

[0102] 等式3

[0103]
$$\Phi(v_x, v_y) = \sum_{[i',j']\in\Omega} \Delta^2[i',j']$$

[0104] 本实施方案的双向光流可以应用于这样的情况,用于双向预测的两个参考图像的一个按显示顺序在当前图像之前,另一个在当前图像之后,并且两个参考图像与当前图像

的距离彼此相等,也就是说,每个参考图像与当前图像之间的图像顺序计数(POC)的差彼此相等。因此,t_0和t_1可以排除。

[0105] 此外,本实施方案的双向光流可以仅应用于亮度组件。对于应用了双向预测的目标块,基于子块而非基于像素来执行本实施方案的双向光流。在下面的描述中,为了简单起见,可以具有2×2、4×4和8×8的各种大小的子块假设具有4×4的大小。

[0106] 在执行光流之前,视频编码装置的帧间预测器124配置为利用上述双向预测来生成用于目标块的两个参考块。两个参考块的第一参考块表示由利用目标块的第一运动矢量MV_0从参考图像Ref_0生成的预测的样本组成的块,并且第二参考块表示由利用运动矢量MV_1从参考图像Ref_1生成的预测的样本组成的块。

[0107] 帧间预测器124配置为利用第一参考块和第二参考块中的预测的样本的水平梯度值和竖直梯度值,为构成目标块的4×4子块的每个来计算光流(v_x,v_y)。光流(v_x,v_y)确定为使得来自参考图像Ref_0的预测的样本与来自参考图像Ref_1的预测的样本之间的差最小化。帧间预测器124配置为通过利用针对4×4子块计算的(v_x,v_y)和4×4子块中的预测的样本的梯度来推导用于修改4×4子块的双向预测的样本的样本偏移。

[0108] 如等式4所示,帧间预测器124配置为计算位置(i,j)处的样本值的水平梯度和竖直梯度。

[0109] 等式4

$$\frac{\partial I^{(k)}}{\partial x}(i,j) = \left(I^{(k)}(i+1,j) - I^{(k)}(i-1,j)\right) \gg shift1$$

[0110]

$$\frac{\partial I^{(k)}}{\partial y}(i,j) = \left(I^{(k)}(i,j+1) - I^{(k)}(i,j-1)\right) \gg shift1$$

[0111] 其中,k为0或1,$I^{(0)}(i,j)$和$I^{(1)}(i,j)$分别表示第一参考块和第二参考块中的位置(i,j)处的样本值。另外,shift1是从亮度分量的比特深度推导的值,例如Shift1=max(6,比特深度-6)。

[0112] 为了推导位于每个参考块的边界处的样本的梯度,需要在第一参考块和第二参考块的边界之外的样本。相应地,如图6所示,每个参考块向左方和向右方扩展一列,并且向上侧和向下侧扩展一行。为了减少计算量,扩展部分中的每个样本可以用参考块中最相邻的位置处的样本或整数样本来填充。另外,在每个参考块的边界之外的样本位置处的梯度可以用与最相邻的位置处的样本相对应的梯度来填充。

[0113] 帧间预测器124配置为如图7所示利用覆盖4×4子块的6×6窗口中的水平梯度和竖直梯度,如等式5所示计算与梯度的自相关和互相关相对应的S1、S2、S3、S5和S6。

[0114] 等式5

[0115] $S_1=\sum_{(i,j)\in\Omega}\psi_x(i,j)\cdot\psi_x(i,j)$,$S_3=\sum_{(i,j)\in\Omega}\theta(i,j)\psi_x(i,j)$

[0116] $$S_2 = \sum_{(i,j)\in\Omega}\psi_x(i,j)\cdot\psi_y(i,j)$$

[0117] $S_5=\sum_{(i,j)\in\Omega}\psi_y(i,j)\cdot\psi_y(i,j)$ $S_6=\sum_{(i,j)\in\Omega}\theta(i,j)\psi_y(i,j)$

[0118] 如上所述,Ω表示一定大小的掩码。另外,如下面的等式6所示,$\Psi_x(i,j)$表示第一参考块和第二参考块中位置(i,j)处的水平梯度值的和,$\Psi_y(i,j)$表示第一参考块和第二参考块中位置(i,j)处的竖直梯度值的和,而$\theta(i,j)$表示第二参考块中位置(i,j)处的样本

值与第一参考块中位置(i,j)处的样本值之间的差。

[0119] 等式6

$$\psi_x(i,j)=\left(\frac{\partial I^{(1)}}{\partial x}(i,j)+\frac{\partial I^{(0)}}{\partial x}(i,j)\right)\gg n_a$$

[0120]

$$\psi_y(i,j)=\left(\frac{\partial I^{(1)}}{\partial y}(i,j)+\frac{\partial I^{(0)}}{\partial y}(i,j)\right)\gg n_a$$

[0121] $\theta(i,j)=(I^{(1)}(i,j)>n_b)-(I^{(0)}(i,j)>n_b)$

[0122] 其中,n_a和n_b是从比特深度推导的值,并且具有min(1,比特深度-11)和min(4,比特深度-8)的值。

[0123] 帧间预测器124配置为基于S1、S2、S3、S5和S6来计算4×4子块的光流(v_x,v_y),如等式7所示。

[0124] 等式7

$$v_x=S_1>0?\,clip3\left(-th'_{BIO},th'_{BIO},-\left((S_3\cdot 2^{n_b-n_a})\gg\lfloor\log_2 S_1\rfloor\right)\right):0$$

[0125]

$$v_y=S_5>0?\,clip3\left(-th'_{BIO},th'_{BIO},-\left(\left(S_6\cdot 2^{n_b-n_a}-\left((v_xS_{2,m})\ll n_{S_2}+v_xS_{2,s}\right)/2\right)\gg\lfloor\log_2 S_5\rfloor\right)\right):0$$

[0126] 其中,$S_{2,m}=S_2\gg n_{S_2}$, $S_{2,s}=S_2\&(2^{n_{S_2}}-1)$和$th'_{BIO}=2^{\max(5,BD-7)}$。$\lfloor\cdot\rfloor$是向下取整函数(floor function),并且$n_{S_2}=12$。

[0127] 可以利用针对4×4子块计算的光流和样本位置(x,y)处的梯度值来计算用于修改目标块中的4×4子块中的每个样本位置(x,y)处的预测的样本的样本偏移,如等式8所示。

[0128] 等式8

[0129]

$$b(x,y)=rnd\left(\left(v_x\left(\frac{\partial I^{(1)}(x,y)}{\partial x}-\frac{\partial I^{(0)}(x,y)}{\partial x}\right)/2\right)\right)+rnd\left(\left(v_y\left(\frac{\partial I^{(1)}(x,y)}{\partial x}-\frac{\partial I^{(0)}(x,y)}{\partial x}\right)/2\right)\right)$$

[0130] 其中,rnd()表示舍入运算。

[0131] 帧间预测器124配置为利用位置(x,y)处的样本偏移b(x,y)和第一参考块中的预测样本$I^{(0)}(x,y)$以及第二参考块中的预测的样本$I^{(1)}(x,y)$来生成最终预测的样本pred(x,y),如等式9所示。

[0132] 等式9

[0133] $pred(x,y)=(I^{(0)}(x,y)+I^{(1)}(x,y)+b(x,y)+o_{offset})>>shift$

[0134] 其中,shift是Max(3,15-比特深度),而o_{offset}是用于舍入运算的值,并且是shift的一半。

[0135] 如上所述,双向光流技术使用利用用于双向预测的运动信息(两个运动矢量和两个参考图像)来预测的样本的值。因此,视频解码装置的帧间预测器544也可以利用从视频编码装置接收的用于双向预测的运动信息(运动矢量、参考图像)以与视频编码装置相同的方式执行双向光流。相应地,不需要视频编码装置将用于双向光流处理的附加信息用信号

通知视频解码装置。

[0136] II.仿射运动预测

[0137] 上述帧间预测是反映平移运动模型的运动预测。换句话说,这是用于在水平方向(x轴方向)和竖直方向(y轴方向)上预测运动的技术。然而,实际上,可能存在各种类型的运动,例如,除平移运动之外,还要旋转、放大或缩小。本实施方案的一个方面提供能够反映这样的各种类型的运动的仿射运动预测。

[0138] 图8a和图8b是示出根据本发明实施方案的仿射运动预测的示例图。

[0139] 可能存在用于仿射运动预测的两种类型的模型。一种模型是利用当前要编码的目标块的左上角和右上角的两个控制点的运动矢量的模型,也就是说,四个参数,如图8a所示。另一种模型是利用目标块的左上角、右上角和左下角的三个控制点的运动矢量的模型,也就是说,六个参数,如图8b所示。

[0140] 四参数仿射模型表达为如等式10所示。在目标块中的样本位置(x,y)处的运动可以根据等式10计算。假设目标块的左上方样本的位置为(0,0)。

[0141] 等式10

[0142]
$$\begin{cases} mv_x = \dfrac{mv_{1x} - mv_{0x}}{W} x - \dfrac{mv_{1y} - mv_{0y}}{W} y + mv_{0x} \\ mv_y = \dfrac{mv_{1y} - mv_{0y}}{W} x + \dfrac{mv_{1x} - mv_{0x}}{W} y + mv_{0y} \end{cases}$$

[0143] 六参数仿射模型由等式11表达。在目标块中的样本位置(x,y)处的运动可以根据等式11计算。

[0144] 等式11

[0145]
$$\begin{cases} mv_x = \dfrac{mv_{1x} - mv_{0x}}{W} x + \dfrac{mv_{2x} - mv_{0x}}{W} y + mv_{0x} \\ mv_y = \dfrac{mv_{1y} - mv_{0y}}{W} x + \dfrac{mv_{2y} - mv_{0y}}{W} y + mv_{0y} \end{cases}$$

[0146] 其中,(mv_0x, mv_0y)是左上角控制点的运动矢量,(mv_1x, mv_1y)是右上角控制点的运动矢量,并且(mv_2x, mv_2y)是左下角控制点的运动矢量。此外,W是目标块的水平长度,并且H是目标块的竖直长度。

[0147] 可以利用根据等式10或等式11计算的运动矢量对目标块中的每个样本执行仿射运动预测。替选地,为了降低计算复杂性,可以对从目标块分割的每个子块执行预测,如图9所示。例如,子块的大小可以是4×4、2×2或8×8。在下面的示例性实施方案中,基于4×4子块来执行用于目标块的仿射运动预测。该示例仅是为了便于说明,并且本发明不限于此。

[0148] 在基于子块的仿射运动预测中,通过将每个子块的中心位置代入等式10或等式11的(x,y)中来计算每个子块的运动矢量(仿射运动矢量)。具体地,中心位置可以是子块的实际中心点或中心点的右下方样本位置。例如,在4×4子块中左下方样本的坐标为(0,0)的情况下,子块的中心位置可以是(1.5,1.5)或(2,2)。利用子块的仿射运动矢量(mv_x, mv_y)来生成用于每个子块的预测的块。运动矢量(mv_x, mv_y)可以设置为具有1/16样本精度。具体地,根据等式1或等式2计算的运动矢量(mv_x, mv_y)可以以1/16样本单位四舍五入。

[0149] 不仅可以对亮度分量而且还可以对色度分量执行仿射运动预测。在4:2:0视频格式的情况下,当基于4×4子块执行用于亮度分量的仿射运动预测时,可以基于2×2子块来

执行用于色度分量的仿射运动预测。可以从相应的亮度分量的运动矢量推导色度分量的每个子块的运动矢量(mv_x,mv_y)。替选地,用于色度分量的仿射运动预测的子块的大小可以与用于亮度分量的子块的大小相同。

[0150] 当基于4×4子块执行用于亮度分量的仿射运动预测时,还可以基于4×4子块来执行用于色度分量的仿射运动预测。特别地,由于用于色度分量的4×4子块对应于用于亮度分量的四个4×4子块,可以通过计算亮度分量的四个相应的子块的运动矢量的平均值来计算用于色度分量的子块的运动矢量(mv_x,mv_y)。

[0151] 视频编码装置配置为执行帧内预测、帧间预测(平移运动预测)、仿射运动预测等,并且计算率失真(rate-distortion,RD)代价以选择最佳预测方法。为了执行仿射运动预测,视频编码装置的帧间预测器124确定使用两种类型的模型的哪种类型的模型,并且根据确定的类型来确定两个或三个控制点。帧间预测器124配置为利用控制点的运动矢量来计算目标块中的4×4子块中每个子块的运动矢量(mv_x,mv_y)。然后,帧间预测器124配置为利用每个子块的运动矢量(mv_x,mv_y)以逐子块为基础在参考图像中执行运动补偿,以生成用于目标块中的每个子块的预测的块。

[0152] 视频编码装置的熵编码器155配置为对包括指示是否将仿射运动预测应用于目标块的标志、指示仿射模型的类型的类型信息以及指示每个控制点的运动矢量的运动信息的仿射相关的语法元素进行编码,并且将其发送到视频解码装置。当执行仿射运动预测时,可以用信号通知关于控制点的类型信息和运动信息,并且可以用信号通知与根据类型信息确定的数量一样多的控制点的运动矢量。

[0153] 视频解码装置配置为利用用信号通知的语法元素来确定仿射模型的类型和控制点运动矢量,并且利用等式10或等式11来计算目标块中每个4×4子块的运动矢量(mv_x,mv_y)。当用信号通知关于目标块的仿射运动矢量的运动矢量分辨率信息时,通过诸如舍入的操作将运动矢量(mv_x,mv_y)校正到由运动矢量分辨率信息识别的精度。

[0154] 视频解码装置配置为通过利用用于每个子块的运动矢量(mv_x,mv_y)在参考图像内执行运动补偿来生成用于每个子块的预测的块。为了减少编码控制点的运动矢量所需的比特量,可以应用如在上述正常帧间预测(平移运动预测)中使用的方法。

[0155] 作为示例,在合并模式中,视频编码装置的帧间预测器124配置为从目标块的相邻块推导每个控制点的运动矢量。例如,帧间预测器124配置为通过从图4所示的目标块的相邻样本L、BL、A、AR和AL推导预定义数量的合并候选来生成合并候选列表。包括在列表中的每个合并候选与两个或三个控制点的成对运动矢量相对应。

[0156] 首先,帧间预测器124配置为从相邻块中以仿射模式预测的相邻块的控制点运动矢量来推导合并候选。在一些实施方案中,可以限制从以仿射模式预测的相邻块来推导的合并候选的数量。例如,帧间预测器124可以配置为从以仿射模式预测的相邻块来推导两个合并候选:L和BL的一个以及A、AR和AL的一个。可以按照L和BL的顺序以及A、AR和AL的顺序来分配优先级。当合并候选的总数大于或等于3时,帧间预测器124可以配置为从相邻块的平移运动矢量来推导必要数量的合并候选。

[0157] 图10是示出从相邻块的平移运动矢量来推导用于仿射运动预测的合并候选的方法的示例图。帧间预测器124配置为分别从相邻块组{B2,B3,A2}、相邻块组{B1,B0}和相邻块组{A1,A0}推导控制点运动矢量CPMV1,CPMV2,CPMV3。作为示例,可以按照B2、B3、A2的顺

序、B1和B0的顺序以及A1和A0的顺序来分配每个相邻块组中的优先级。另外,从参考图像中的同位块T推导另一控制点运动矢量CPMV4。帧间预测器124配置为通过四个控制点运动矢量中的两个或三个控制点运动矢量的组合来生成与需要的数量一样多的合并候选。组合的优先级分配如下所示。按照控制点运动矢量的左上角、右上角和左下角的顺序列出每个组中的元素。

[0158] {CPMV1,CPMV2,CPMV3}、{CPMV1,CPMV2,CPMV4}、{CPMV1,CPMV3,CPMV4}、{CPMV2,CPMV3,CPMV4}、{CPMV1,CPMV2}、{CPMV1,CPMV3}

[0159] 帧间预测器124配置为选择合并候选列表中的合并候选,并且对目标块执行仿射运动预测。当选择的候选由两个控制点运动矢量组成时,利用四参数模型执行仿射运动预测。另一方面,当选择的候选由三个控制点运动矢量构成时,利用六参数模型来执行仿射运动预测。视频编码装置的熵编码器155编码索引信息并将其用信号通知给视频解码装置,所述索引信息指示在合并候选列表中的合并候选中选择的合并候选。

[0160] 视频解码装置的熵解码器510对从视频编码装置用信号通知的索引信息进行解码。视频解码装置的帧间预测器544以与视频编码装置相同的方式构建合并候选列表,并且配置为利用与由索引信息指示的合并候选相对应的控制点运动矢量来执行仿射运动预测。

[0161] 作为另一示例,在AMVP模式中,视频编码装置的帧间预测器124配置为确定用于目标块的仿射模型的类型和控制点运动矢量。然后,帧间预测器124配置为确定运动矢量差,该运动矢量差是目标块的实际控制点运动矢量与各个控制点的预测的运动矢量之间的差,并且发送分别对应于控制点的运动矢量差。相应地,视频编码装置的帧间预测器124配置预定义数量的仿射AMVP的列表。当目标块是4参数类型时,包括在列表中的候选各自由成对的两个控制点运动矢量组成。另一方面,当目标块是6参数类型时,包括在列表中的候选各自由成对的三个控制点运动矢量组成。可以以与上述构建合并候选列表的方法相似的方式,利用相邻块的控制点运动矢量或平移运动矢量来推导仿射AMVP列表。

[0162] 然而,为了推导要包括在仿射AMVP列表中的候选,可能存在这样的限制:仅考虑相邻块,所述相邻块参考与图4的相邻块中的目标块相同的参考图像。

[0163] 另外,在AMVP模式中,应当考虑目标块的仿射模型类型。当目标块的仿射模型类型是4参数类型时,视频编码装置利用相邻块的仿射模型来推导两个控制点运动矢量(目标块的左上角和右上角控制点运动矢量)。当目标块的仿射模型类型是6参数类型时,所述装置利用相邻块的仿射模型来推导三个控制点运动矢量(目标块的左上角、右上角和左下角控制点运动矢量)。

[0164] 当相邻块为4参数类型时,利用相邻块的两个控制点运动矢量,根据目标块的仿射模型类型来预测两个或三个控制点运动矢量。例如,可以使用由等式10表达的相邻块的仿射模型。在等式10中,(mv_{0x},mv_{0y})和(mv_{1x},mv_{1y})分别由相邻块的左上角和右上角控制点运动矢量代替。此外,W由相邻块的水平长度代替。通过将目标块的相应控制点的位置与相邻块的左上角的位置之间的差置入(x,y)中来推导用于目标块的每个控制点的预测的运动矢量。

[0165] 当相邻块为6参数类型时,利用相邻块的三个控制点运动矢量,根据目标块的仿射模型类型来预测两个或三个控制点运动矢量。例如,可以使用由等式11表达的相邻块的仿射模型。在等式11中,(mv_{0x},mv_{0y})、(mv_{1x},mv_{1y})和(mv_{2x},mv_{2y})分别由相邻块的左上角、右上角

和左下角的控制点运动矢量代替。通过将目标块的相应控制点的位置与相邻块的左上角的位置之间的差置入(x,y)中来推导用于目标块的每个控制点的预测的运动矢量。

[0166] 视频编码装置的帧间预测器124配置为在仿射AMVP列表中选择一个候选,并且生成每个实际控制点的运动矢量与选择的候选的相应控制点的预测的运动矢量之间的运动矢量差。视频编码装置的熵编码器155配置为对指示目标块的仿射模型类型的类型信息、指示仿射AMVP列表中的候选中的选择的候选的索引信息以及与每个控制点相对应的运动矢量差进行编码,并将其发送到视频解码装置。

[0167] 视频解码装置的帧间预测器544配置为利用从视频编码装置用信号通知的信息来确定仿射模型类型,并且生成每个控制点的运动矢量差。然后,帧间预测器配置为以与视频编码装置相同的方式生成仿射AMVP列表,并且选择由仿射AMVP列表中用信号通知的索引信息所指示的候选。视频解码装置的帧间预测器544配置为通过将选择的候选的每个控制点的预测的运动矢量与相应的运动矢量差相加来计算每个控制点的运动矢量。

[0168] III.仿射运动预测样本的逐样本调整

[0169] 上面已经描述了针对目标块的逐子块仿射运动预测。本发明的另一方面涉及基于逐样本调整由逐子块仿射运动预测生成的预测的样本的样本值。在形成仿射运动预测的基础的每个子块中额外补偿根据每个样本的位置的运动。

[0170] 当作为针对目标块的逐子块仿射运动预测的结果而生成的任何一个子块中的样本值为I(x,y)时,视频编码装置计算每个样本位置处的水平梯度gx(i,j)和竖直梯度gy(i,j)。等式12可以用于计算梯度。

[0171] 等式12

[0172] $g_x(i,j)=I(i+1,j)-I(i-1,j)$

[0173] $g_y(i,j)=I(i,j+1)-I(i,j-1)$

[0174] 用于调整预测样本的样本偏移ΔI(i,j)由以下等式计算。

[0175] 等式13

[0176] $\Delta I(i,j)=g_x(i,j)*\Delta mv_x(i,j)+g_y(i,j)*\Delta mv_y(i,j)$

[0177] 其中,Δmv(i,j)表示运动偏移,也就是说,样本(i,j)中的仿射运动矢量与子块的中心位置处的仿射运动矢量之间的差,并且可以根据目标块的仿射模型类型通过应用等式10或等式11来计算。换句话说,可以根据等式10或等式11通过将(i,j)置入(x,y)时给出的运动矢量减去将子块中心位置置入(x,y)时给出的运动矢量来计算Δmv(i,j)。另外,可以根据通过将从子块中心位置到样本位置(i,j)的水平偏移和竖直偏移代入等式11和等式12中的(x,y)除去最后项“$+mv_{0x}$”和“$+mv_{0y}$”来计算Δmv(i,j)。中心位置可以是子块的实际中心点,或者可以是中心点的右下方样本位置。

[0178] 用于计算Δmv(i,j)的目标块的每个控制点的运动矢量以及子块的样本位置(i,j)与中心位置之间的差对于所有子块是相同的。相应地,Δmv(i,j)的值可以仅用于计算一个子块(例如,第一子块),并且可以重复用于其他子块。

[0179] 本发明的技术基于物体以恒定速度移动并且样本值的变化均匀的假设。因此,通过将Δmv(i,j)的x分量(Δmv_x)和y分量(Δmv_y)分别乘以水平样本梯度值和竖直样本梯度值来获得水平方向上的样本变化量和竖直方向上的样本变化量。通过将两个样本变化量相加来计算样本偏移ΔI(i,j)。预测的样本的最终值计算如下。

[0180] 等式14

[0181] $I'(i,j)=I(i,j)+\Delta II(i,j)$

[0182] 当应用用于仿射运动预测样本的逐样本调整时,视频编码装置的帧间预测器124和视频解码装置的帧间预测器544执行上述处理以修改通过仿射运动预测生成的预测的样本的样本值。从通过仿射运动预测生成的预测的样本推导梯度值,并且从目标块的控制点运动矢量推导$\Delta mv(i,j)$。因此,不需要视频编码装置将用于本技术的附加信息用信号通知视频解码装置。

[0183] IV.运动矢量细化

[0184] 如上所述,在利用双向预测的帧间预测的情况下,确定从参考图像列表L_0选择的参考图像Ref_0、关于参考图像Ref_0的运动矢量MV_0、从参考图像列表L_1选择的参考图像Ref_1、以及关于参考图像Ref_1的运动矢量MV_1。本实施方案中描述的运动矢量细化是通过精细地调整用于双向预测的两个运动矢量MV_0和MV_1来提高帧间预测的精度的编码工具。运动矢量细化可以应用于合并模式,在所述合并模式中相邻块的运动矢量用作目标块的运动矢量。替选地,运动矢量细化可以应用于AMVP模式(正常帧间预测(normal inter-prediction))或仿射模式。

[0185] 图11是示出根据本发明实施方案的运动矢量细化编码工具的概念的示例图。在双向预测中,视频编码装置的帧间预测器124配置为在参考图像Ref_0和参考图像Ref_1的初始运动矢量周围搜索细化的运动矢量。特别地,初始运动矢量表示用于双向预测的两个运动矢量MV_0和MV_1。利用块匹配来计算参考图像Ref_0和Ref_1中的两个候选块之间的失真。如图11所示,在参考图像Ref_0和Ref_1中,计算在由各自的初始运动矢量确定的位置处的两个候选块之间以及在所述位置周围的两个候选块之间的绝对差的和(sum of absolute difference,SAD)。

[0186] 换句话说,计算两个候选块中彼此对应的样本值之间的差值的绝对值的和。帧间预测器124配置为选择具有最小SAD的运动矢量候选作为细化的运动矢量,并且根据帧间预测使用细化的运动矢量以生成双向预测信号。尽管SAD用于表示两个候选块之间的差,但实施方案不限于此。可以使用能够表达两个候选块之间的差的任何度量,例如误差平方和(sum of squared error,SSE)。

[0187] 当一个参考图像按显示顺序在当前图像之前,另一个参考图像在当前图像之后,并且两个参考图像到当前图像的距离相同,也就是说,参考图像与当前图像的图像顺序计数的差(POC)相同时,可以应用运动矢量细化。如图11所示,对于两个参考图像Ref_0和Ref_1中的两个初始运动矢量MV_0和MV_1,相应参考图像列表中的细化的双向运动矢量候选MV_0'和MV_1'可以表达为等式15。

[0188] 等式15

[0189] $MV_0'=MV_0+MV_offset$

[0190] $MV_1'=MV_1-MV_offset$

[0191] 其中,MV_offset是根据运动矢量细化的初始运动矢量的偏移,并且表示运动矢量候选与相应的初始运动矢量之间的差。该偏移可以是以整数样本为单位的整数偏移和以子像素或子像素样本为单位的子像素偏移的和。如等式15所示,两个运动矢量的候选符合MV差分的镜像规则。搜索用于细化的运动矢量的最佳偏移的处理包括两个操作:用于搜索整

数偏移的整数样本搜索,以及用于搜索子像素偏移的分数样本细化。

[0192] 首先,在整数样本搜索操作中,帧间预测器124配置为搜索所有的25个相邻点,包括如图12所示的细化的中心。特别地,细化的中心与由初始运动矢量指示的位置相对应。当根据初始运动矢量确定的两个位置处的成对的两个候选块之间的SAD小于特定阈值时,可以跳过整数样本搜索之后的操作。通过计算与由初始运动矢量指示的位置的相邻点相对应的两个候选块之间的SAD,帧间预测器124配置为搜索具有最小SAD的点,也就是说,minSAD。根据整数样本搜索的结果,具有minSAD的运动矢量候选的整数偏移表达为(intOffX,intOffY)。

[0193] 在整数样本搜索操作完成之后,当整数偏移的intOffX不是-2或2且intOffY不是-2或2{(intOffX!=(2||-2))&&(intOffY!=(2||-2))}时,帧间预测器124将作为应用分数样本细化操作的指示标志的subPelFlag设置为“真(true)”值。当将subPelFlag设置为“真”时,帧间预测器124执行分数样本细化操作。帧间预测器124配置为通过求解二维空间中的参数误差曲面方程来生成最佳子像素偏移,如等式16所示。

[0194] 等式16

[0195] $E(x,y)=A(x-x_{min})^2+B(y-y_{min})^2+C$

[0196] 其中,E(x,y)是位置(x,y)处的代价,并且可以定义为与位置(x,y)相对应的两个候选块之间的SAD。另外,(x_{min},y_{min})是具有最小代价C的样本位置,并且表示最佳的子像素偏移。

[0197] 假设根据整数样本搜索,与整数偏移相对应的样本位置为(0,0),则可以用与所有的四个位置(顶部、底部、左侧和右侧位置)(-1,0)、(1,0)、(0,1)和(0,1)相对应的块的代价来计算最佳子像素偏移,如等式17所示。

[0198] 等式17

[0199] $x_{min}=(E(-1,0)-E(1,0))/(2(E(-1,0)+E(1,0)-2E(0,0)))$

[0200] $y_{min}=(E(0,-1)-E(0,1))/(2(E(0,-1)+E(0,1)-2E(0,0)))$

[0201] 对代价E(x,y)的约束可以由等式18表达。

[0202] 等式18

[0203] E(x,y)≥0&&E(0,0)最小

[0204] 基于等式18,“$-0.5<x_{min},y_{min}<0.5$”得到满足。

[0205] 帧间预测器124可以配置为通过将子像素偏移和整数偏移相加来生成以子像素为单元的最佳偏移,然后利用等式15将最佳偏移与初始运动矢量相加以获得细化的双向运动矢量。此外,帧间预测器124可以配置为利用细化的双向运动矢量和双向参考图像来生成目标块的帧间预测的样本。

[0206] 如上所述,在运动矢量细化处理中,利用用于双向预测的运动信息(两个运动矢量和两个参考图像)来细化运动矢量。相应地,视频解码装置的帧间预测器544也可以配置为利用用于双向预测并从视频编码装置接收的运动信息(运动矢量和参考图像)以与视频编码装置相同的方式执行运动矢量细化。因此,不需要从视频编码装置向视频解码装置用信号通知用于运动矢量细化处理的附加信息。

[0207] 当应用运动矢量细化处理时,由于对于每个目标块必须计算候选块的SAD值,因此特别是在解码器侧可能会引起各种复杂性问题。当基于与整数偏移相对应的样本对运动矢

量执行分数样本细化操作时,需要求解如等式16所示的复杂等式。这个过程可能难以用硬件进行优化。当合并模式适用于帧间预测时,在运动矢量细化中对分数样本细化操作的需要可能相对较低。相应地,对于在合并模式中编码的每个块的运动矢量细化处理增加了视频编码/解码装置的复杂性,并且使得难以期望相应的压缩效率。

[0208] 鉴于上述,本发明提供了一种通过以下操作降低视频编码/解码装置的复杂度的方法,所述操作通过利用在整数样本搜索操作中获取的整数偏移信息来确定以子像素为基础的运动矢量调整的需要,并且根据确定的结果来跳过分数样本细化操作。在上述运动矢量细化处理中,当初始运动矢量对的SAD大于或等于特定阈值时,执行整数样本搜索和分数样本细化操作。根据目标块的大小来确定阈值。例如,可以将阈值设置为sbHeight×sbWidth的值,也就是说,目标块的面积。

[0209] 在根据本实施方案的另一个示例中,可以基于量化参数以及目标块的面积来确定阈值。在另一个示例中,可以基于在合并模式中搜索的相邻块的索引值来确定是否执行整数样本搜索和分数样本细化操作。在运动矢量细化处理的整数样本搜索操作中,在水平地和竖直地从-2至+2的总共25个整数坐标(包括与初始运动矢量相对应的细化的中心)上执行搜索操作,如图12所示。相应地,具有minSAD的块的位置的整数偏移表达为(intOffX,intOffY)。

[0210] 为了跳过根据如等式16和等式17所示的基于子像素的运动矢量细化操作的复杂等式求解处理,在整数样本搜索操作中执行到特定条件的整数样本的搜索之后,可以提前终止分数样本细化操作。当intOffX和intOffY两者均为0时,视频编码装置的帧间预测器124可以配置为跳过搜索子像素偏移的分数样本细化操作,并且将如等式17所示的子像素偏移设置为零。替选地,当(intOffX,intOffY)不为(0,0)时,帧间预测器124可以将subPelFlag设置为“真”值,以应用分数样本细化操作。由此,当(intOffX,intOffY)为(0,0)时,可以提前终止分数样本细化操作。

[0211] 作为本发明实施方案的另一方面,当在整数样本搜索操作中获取的SAD小于特定阈值时,帧间预测器124配置为跳过搜索矢量的子像素偏移的分数样本细化操作。例如,除了minSAD之外,将第二小和第三小的SAD分别表示为SAD_1和SAD_2。当表达为等式19的条件得到满足时,可以跳过分数样本细化操作。

[0212] 等式19

[0213] Abs(minSAD-SAD_1)<threshold&&Abs(minSAD-SAD_2)<threshold

[0214] 其中,Abs(•)是将输入改变为绝对值的函数。

[0215] 作为根据本发明的实施方案的另一方面,帧间预测器124配置为基于在合并模式中搜索的相邻块的运动矢量的大小来确定是否跳过搜索运动矢量的子像素偏移的分数样本细化操作。例如,当在合并模式中搜索的两个或更多个相邻块的两个或更多个运动矢量相同或相似(例如,运动矢量之间的差小于一定阈值)时,可以跳过分数样本细化操作。

[0216] 在合并模式中,基于在一个或更多个特定索引处搜索的运动矢量的大小,帧间预测器124配置为确定是否跳过搜索运动矢量的子像素偏移的分数样本细化操作。例如,当在左上侧块、左下侧块和右上侧块中搜索的运动矢量的值相等或彼此相差很小时,在合并模式中可以跳过分数样本细化操作。作为另一个示例,当在左上侧块、左下侧块和右上侧块中搜索的运动矢量的大小等于0或小于特定阈值时,可以跳过分数样本细化操作。

[0217] 在合并模式中,帧间预测器124配置为基于已经推导出在一个或更多个特定索引处搜索的运动矢量的处理来确定是否跳过搜索运动矢量的子像素偏移的分数样本细化操作。例如,当在没有运动矢量细化处理的情况下已经推导出左上侧块、左下侧块和右上侧块的运动矢量的值时,在当前块中不附加执行运动矢量细化处理。作为另一个示例,当在推导左上侧块、左下侧块和右上侧块的运动矢量的处理中没有使用运动矢量细化处理的分数样本细化操作时,可以在当前块的运动矢量细化处理中跳过分数样本细化操作。

[0218] 在当前块的一个或更多个相邻块(例如,左上侧块、左下侧块和右上侧块)使用合并模式并且以跳过模式(即,对于块,cu_skip_flag==真(TRUE))进行编码时,影像编码/解码装置配置为确定出对当前块的运动矢量的精度要求较低,并且因此跳过运动矢量细化处理。具体地,在跳过运动矢量细化处理时,可以只跳过分数样本细化操作,或者可以跳过包括分数样本细化操作的整数样本搜索操作。

[0219] 如上所述,在根据本实施方案的运动矢量细化处理中,可以通过提前终止分数样本细化操作来降低帧间预测器124的搜索复杂性。

[0220] 帧间预测器124可以配置为基于用于预测的运动信息(两个运动矢量和两个参考图像)、在合并模式中搜索的相邻块的运动矢量等来执行运动矢量细化处理的提前终止。相应地,视频解码装置的帧间预测器544还可以配置为基于用于从视频编码装置接收的双向预测的运动信息(运动矢量和参考图像)、在合并模式中搜索的相邻块的运动矢量等,以与视频编码装置相同的方式执行运动矢量细化处理的提前终止。相应地,不需要从视频编码装置向视频解码装置用信号通知用于运动矢量细化处理的提前终止的附加信息。

[0221] 在下文中,作为根据本发明的实施方案的另一方面,将描述用于在运动矢量细化处理的搜索范围的边界处执行附加的分数样本细化操作的方法。

[0222] 如上所述,当intOffX为-2或2或者intOffY为-2或2{(intOffX==(2||-2))||(intOffY=(2||-2))}时,跳过运动矢量细化处理的分数样本细化操作。由于根据等式18所示的条件,运动矢量细化的结果限制在-1.5至1.5,因此视频编码/解码装置的压缩性能可能会受到限制。相应地,根据本发明的运动矢量细化可以在将整数样本搜索的范围限制为-2至2的同时,允许从-2至-1.5和从1.5至2的以子像素为单位的运动矢量的分数样本细化操作,从而改善压缩性能。

[0223] 即使当intOffX为-2或2或者intOffY为-2或2时,视频编码装置的帧间预测器124也可以配置为将subPelFlag设置为真值,以便应用分数样本细化操作。当将subPelFlag设置为真时,帧间预测器124可以配置为执行搜索以至少1/4像素为单位的子像素偏移的分数样本细化操作,以对小于或等于-1.75或者大于或等于1.75的坐标执行运动矢量细化。作为另一个示例,当将subPelFlag设置为真值时,帧间预测器124可以配置为执行搜索以至少1/8像素为单位的子像素偏移的分数样本细化操作,以执行坐标为-1.875或1.875的运动矢量细化。

[0224] 替选地,通过删除subPelFlag,帧间预测器124可以配置为即使当intOffX为-2或2或者intOffY为-2或2时,也执行搜索子像素偏移的分数样本细化操作。如上所述,当intOffX为-2或2或者intOffY为-2或2时,帧间预测器124可以配置为执行搜索以至少1/4像素为单位的子像素偏移的分数样本细化操作,以细化坐标为-1.75或1.75的运动矢量。此外,当intOffX为-2或2或者intOffY为-2或2时,不执行坐标为-1.25或1.25的运动矢量细

化。

[0225] 为了应用等式17,(intOffX,intOffY)需要顶部、底部、左侧和右侧的4个位置。然而,当intOffX为-2或2或者intOffY为-2或2时,可能不存在一个或更多个所需的位置。相应地,可以通过在水平方向上应用Max(-2,intoffx-1)、intOffX和Min(intOffX+1,2),并且在竖直方向上应用Max(-2,intOffY-1)、intOffY和Min(intOffY+1,2)来设置(intOffX,intOffY)的顶部、底部、左侧和右侧的4个位置。换句话说,当(intOffX,intOffY)位于25个点的最外侧时,最外侧的点(intOffX,intOffY)可以用作顶部、底部、左侧或右侧位置。

[0226] 上述运动矢量细化编码工具需要将通过由初始运动矢量确定的25个样本位置的SAD计算来执行的整数样本搜索,以及将根据参数误差曲面方程来执行的分数样本细化,因此可能会增加复杂性。在下文中,将描述简化的运动矢量细化处理。当附加地执行分数样本细化操作,甚至包括intOffX为-2或2或者intOffY为-2或2的情况时,简化的运动矢量细化处理可以降低视频编码/解码装置的复杂性。

[0227] 在本发明的实施方案中,如图12所示,帧间预测器124配置为在整数样本搜索操作中计算位置A1、A2、A3和A4的SAD,以找到具有minSAD的位置。当具有minSAD的位置为A4时,计算A4周围的位置A41、A42、A43和A44的SAD。基于计算出的SAD,帧间预测器124配置为利用等式16和等式17所示的参数误差曲面方程来执行分数样本细化操作,从而生成以1/8像素为单位的(x_{min},y_{min})。由于上述处理不满足在等式18所示的约束中中心点的代价应当最小的约束,从而可以以1/8像素为单位对搜索范围内的所有位置进行搜索。该搜索方法将整数样本搜索操作中的搜索次数从25次减少到8次,从而可以显著地降低搜索复杂性。

[0228] 在本发明的另一个实施方案中,如图13所示,帧间预测器124配置为计算整数样本搜索操作中的位置A1、A2、A3、A4和A5的SAD,并且基于计算出的SAD来执行分数样本细化操作,从而生成以1/4像素为单位的(x_{min},y_{min})。在第一至第四象限中检查(x_{min},y_{min})的大致位置。然后,当(x_{min},y_{min})位于例如第二象限中时,帧间预测器124配置为额外地计算位置B1、B2、B3、B4和B5的SAD。基于计算出的SAD,帧间预测器124配置为附加地执行分数样本细化操作,从而生成以1/8像素为单位的(x_{min},y_{min})。上述处理不满足在等式18所示的约束中中心点的代价应当最小的约束,相应地可以以1/8像素为单位对搜索范围内的所有位置进行搜索。利用该搜索方法,用10个样本位置的搜索和附加的分数样本细化来代替整数样本搜索操作中所需的25个样本位置的搜索。由此,可以降低复杂性。

[0229] 在本发明的另一个实施方案中,如图13所示,帧间预测器124配置为计算整数样本搜索操作中的位置A1、A2、A3、A4和A5的SAD,并且基于计算出的SAD来执行分数样本细化操作,以生成以1/4像素为单位的(x_{min},y_{min})。由于上述处理不满足在等式18所示的约束中中心点的代价应当最小的约束,从而可以以1/8像素为单位对搜索范围内的所有位置进行搜索。利用这种搜索方法,整数样本搜索操作中的搜索次数可以从25次减少到5次。由此,可以显著地降低搜索复杂性。

[0230] 在本发明的另一个实施方案中,如图14所示,帧间预测器124配置为在整数样本搜索操作中计算位置A1至A9的SAD,以找到具有minSAD的位置。例如,当A5是具有minSAD的位置时,帧间预测器124配置为基于为位置A2、A4、A6和A8预先计算出的SAD执行分数样本细化操作,以生成以1/8像素为单位的(x_{min},y_{min})。由于上述处理不满足在等式18所示的约束中中心点的代价应当最小的约束,从而可以以1/8像素为单位对搜索范围内的所有位置进行

搜索。

[0231] 当根据执行整数样本搜索的结果,A7是具有minSAD的位置时,额外地计算位置B1和位置B2的SAD。基于位置A4、A7、A8、B1和B2的SAD,帧间预测器124通过执行分数样本细化操作来生成以1/8像素为单位的(x_{min}, y_{min})。由于上述处理不满足在等式18所示的约束中中心点的代价应当最小的约束,从而可以以1/8像素为单位对搜索范围内的所有位置进行搜索。利用该搜索方法,整数样本搜索操作中的搜索次数可以从25次减少到最小8次最大11次。由此,可以降低搜索复杂性。

[0232] 如上所述,在根据本实施方案的运动矢量细化处理中,可以通过扩展子像素偏移的搜索范围来提高压缩性能,并且可以通过简化整数样本搜索来降低搜索复杂性。

[0233] 视频编码装置可以利用用于预测的运动信息(两个运动矢量和两个参考图像)来扩展运动矢量细化处理的搜索范围并简化搜索操作。因此,视频解码装置的帧间预测器544也可以利用用于从视频编码装置接收的双向预测的运动信息(运动矢量和参考图像),以与视频编码装置相同的方式扩展运动矢量细化处理的搜索范围并简化搜索操作。因此,不需要从视频编码装置向视频解码装置用信号通知附加信息,以简化运动矢量细化处理的搜索操作。

[0234] V.用信号通知高级语法

[0235] 本发明的另一个方面涉及用于将控制上述各种工具的开/关的高级语法用信号通知的技术。上述光流、用于仿射运动预测样本的逐样本调整和运动矢量细化是用于提高视频编码效率的编码工具。然而,对于诸如以下的特定内容,例如屏幕内容,前述各种编码工具可能无助于改善压缩性能。相应地,应用基于用信号通知是否应用每个编码工具的编码单元或者基于确定是否应用工具的编码单元可能反而降低编码效率或增加计算复杂性。本发明提供了用于有效地控制上述编码工具的发信号技术。

[0236] 为了编码一系列图像,视频编码装置配置为对指示是否禁用上述多个编码工具的每个的序列级别的第一高级语法元素进行编码。第一高级语法元素可以被编码在序列头中,也就是说,SPS中。特别地,多个编码工具包括光流、逐样本调整和运动矢量细化的至少一个。

[0237] 当第一高级语法元素指示出不启用多个编码工具的每个时,将相应的编码工具应用于序列的编码。换句话说,上述编码工具的每个不应用于序列中的任何图像。当第一高级语法元素指示出启用多个编码工具的每个时,可以使用多个编码工具的每个,以编码相应的序列。例如,光流适用于序列中的所有图像。另外,运动矢量细化适用于序列中的所有图像。

[0238] 视频编码装置配置为根据第一高级语法元素的值对图像级别标志进行编码,所述图像级别标志指示在属于序列的每个图像的图像头(即,PPS)中是否允许相应的编码工具。例如,当第一高级语法元素指示出在序列级别上允许相应的编码工具时,对用于指示在图像级上是否允许相应编码工具的图像级别标志进行编码。当第一高级语法元素指示出在序列级别上不允许相应的编码工具时,不单独对图像级别标志进行编码。具体地,视频解码装置配置为自动地将图像级别标志设置为指示出不允许相应编码工具的值。

[0239] 视频解码装置可以配置为在图像级上根据从视频编码装置用信号通知的图像级别标志或自动设置值的图像级别标志来设置是否允许相应的编码工具。当用于根据本实施

方案的编码工具的图像级别标志不包括在比特流中时,视频解码装置将每个图像级别标志设置为指示出在图像级上不允许相应编码工具的值。

[0240] 根据上述语法结构,视频编码装置可以配置为自适应地确定在序列级别和图像级别上是否允许编码工具。此外,视频解码装置可以配置为通过对从视频编码装置用信号通知的语法进行解码来自适应地确定在序列级别和图像级别上是否允许编码工具。

[0241] 在本发明的一些其他的实施方案中,上述实施方案的图像级别标志可以用切片级别标志代替。换句话说,当第一高级语法元素指示出在序列级别上允许相应的编码工具时,视频编码装置在切片头中对指示在切片级别上是否允许相应编码工具的切片级别标志进行编码。当第一高级语法元素指示出不允许相应的编码工具时,不对切片级别标志进行编码。视频编码装置可以配置为自适应地确定在序列级别和切片级别上是否允许编码工具。视频解码装置可以配置为根据第一高级语法元素或切片级别标志来确定在序列级别和切片级别上是否允许每个工具。

[0242] 在本发明的另一个实施方案中,为了编码一系列图像,视频编码装置配置为对指示是否启用上述所有编码工具的序列级别的第一高级语法元素进行编码。第一高级语法元素可以被编码在序列头中,也就是说,SPS中。当序列级别的第一高级语法元素指示出不启用多个编码工具的任何一个时,相应的编码工具不应用于序列的编码。换句话说,上述编码工具均不应用于序列中的任何图像。当序列级别的第一高级语法元素指示出对所有的多个编码工具的使用没有约束时,视频编码装置配置为对指示是否启用所有的多个编码工具的切片级别的第二高级语法元素进行编码。第二高级语法元素可以在切片头中编码。

[0243] 当切片级别的第二高级语法元素指示出不启用多个编码工具的任何一个时,相应的编码工具不应用于相应切片的编码。换句话说,上述编码工具均不应用于切片。当切片级别的第二高级语法元素为0时,对所有的多个编码工具的使用没有约束。此外,当序列级别的第一高级语法元素指示出禁用所有的多个编码工具时,将切片级别的第二高级语法元素设置为不启用多个编码工具的任何一个,并且相应的编码工具不应用于相应切片的编码。

[0244] 另外,在根据本发明的一些其他实施方案中,如上所述的切片级别的第二高级语法元素可以用图像级的第二高级语法元素代替。第二高级语法元素可以编码在图像头(即,PPS)中。当图像级的第二高级语法元素指示出禁用所有的多个编码工具时,相应的编码工具不应用于图像的编码。换句话说,上述编码工具的任何一个不应用于图像。当图像级的第二高级语法元素为0时,对所有的多个编码工具的使用没有约束。此外,当序列级别的第一高级语法元素指示出禁用所有的多个编码工具时,将图像级的第二高级语法元素设置为不启用多个编码工具的任何一个,并且相应的编码工具不应用于图像的编码。

[0245] 如上所述,根据本实施方案,通过提供一种用于提高帧间预测的压缩性能的简单地执行与编码工具中的运动矢量细化有关的整数样本搜索操作和分数样本细化操作的方法,可以提高编码器/解码器的编码效率,因此可以降低复杂性。

[0246] 本发明中描述的各种功能或方法可以用存储在非易失性记录介质中的指令来实现,所述指令可以由一个或更多个处理器读取和执行。非易失性记录介质包括,例如,所有类型的记录器件,其中数据以计算机系统可读的形式存储。例如,非易失性记录介质包括存储介质,例如可擦除可编程只读存储器(EPROM)、闪存驱动器,光盘驱动器、磁性硬盘驱动器和固态驱动器(SSD)。

[0247] 尽管已经出于说明的目的描述了示例性实施方案,但是本领域的技术人员将理解,在不脱离实施方案的思想和范围的情况下,各种修改和变化是可能的。为了简洁和清楚起见,已经描述了示例性实施方案。相应地,普通技术人员应当理解,实施方案的范围不受以上明确描述的实施方案限制,而是包括在权利要求及其等同形式内。

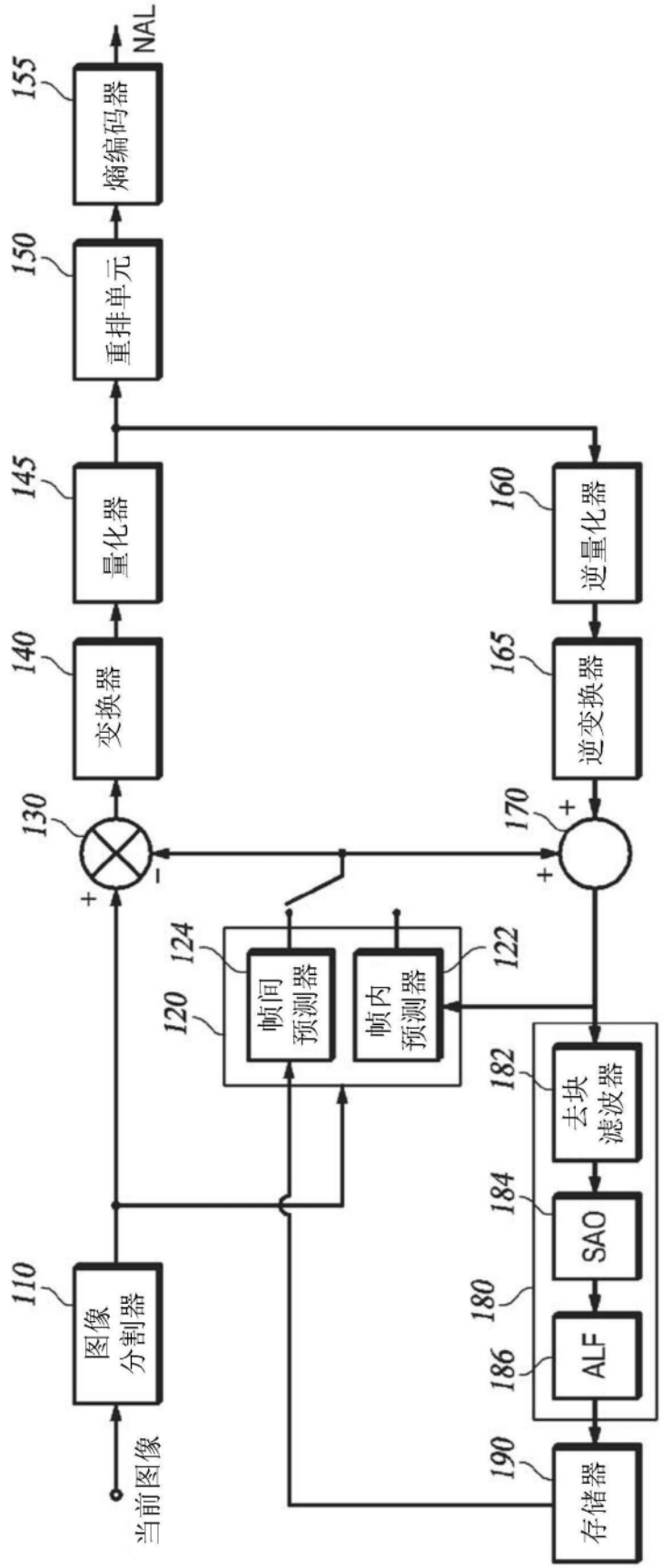

NAL
155
熵编码器
150
重排单元
145
量化器
140
变换器
130
110
图像分割器
当前图像
160
逆量化器
165
逆变换器
170
124
帧间预测器
122
帧内预测器
120
182
去块滤波器
184
SAO
180
186
ALF
190
存储器

图1

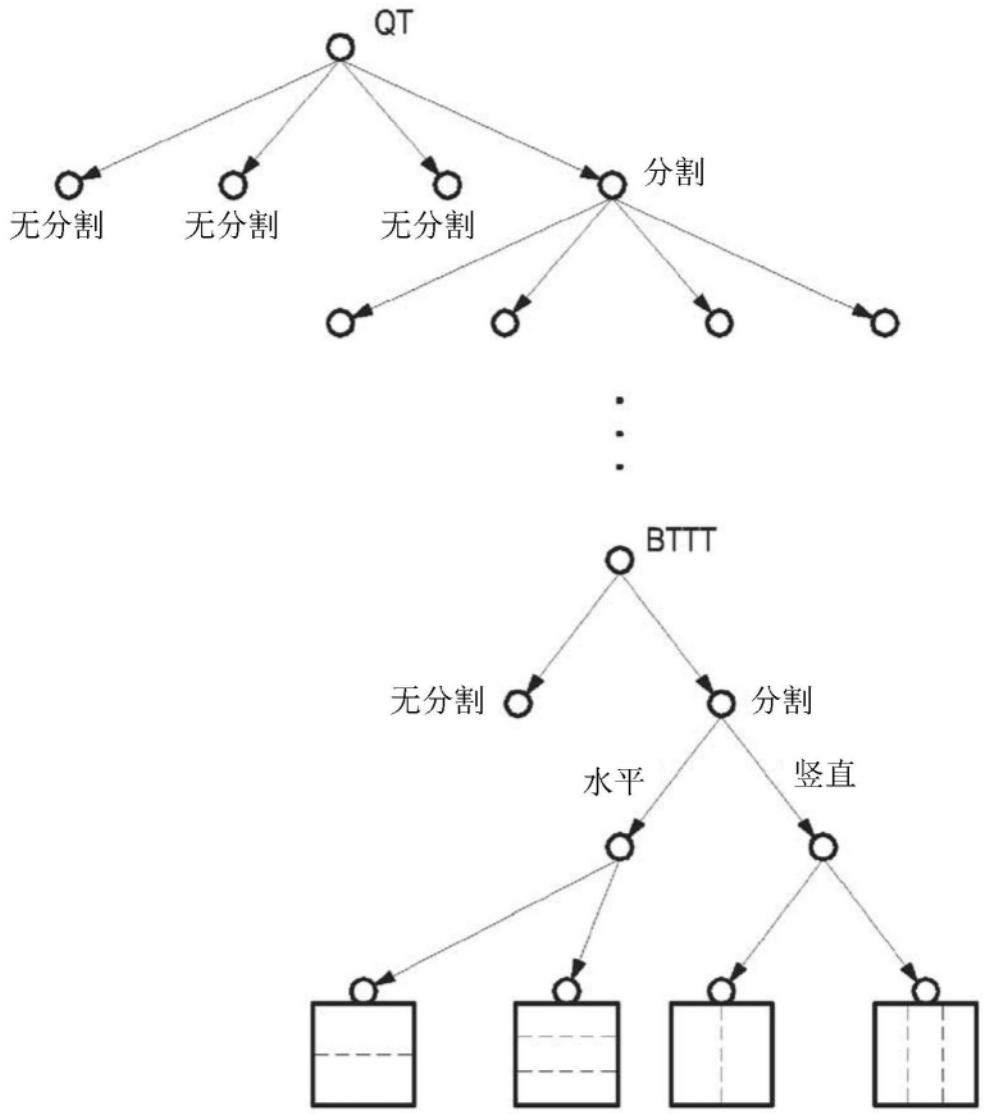


图2

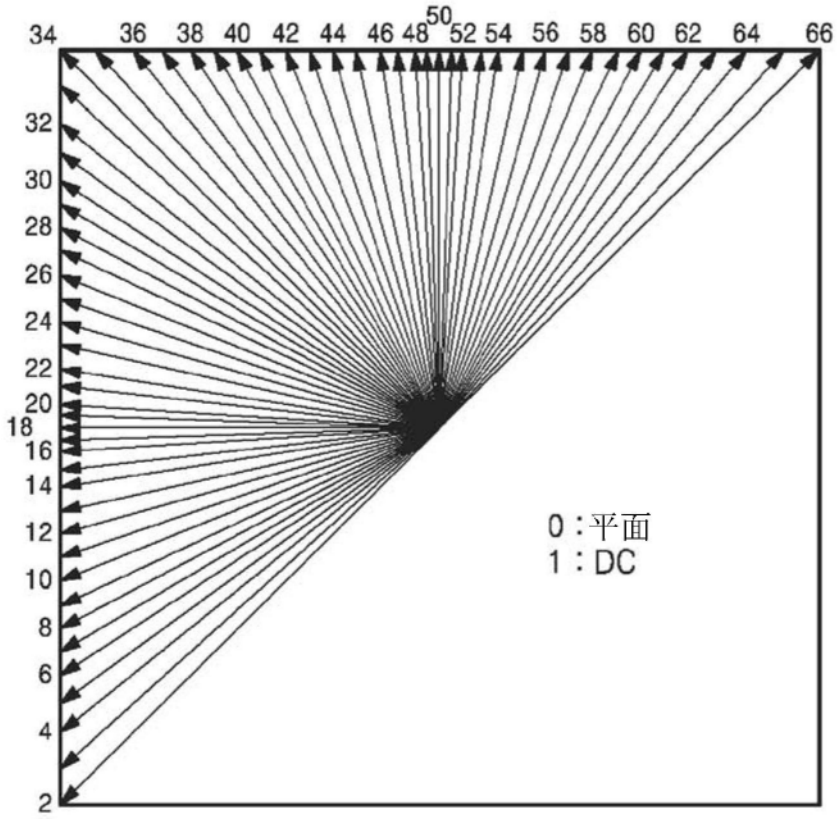

34
36
38
40
42
44
46
48
50
52
54
56
58
60
62
64
66
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0：平面
1：DC

图3a

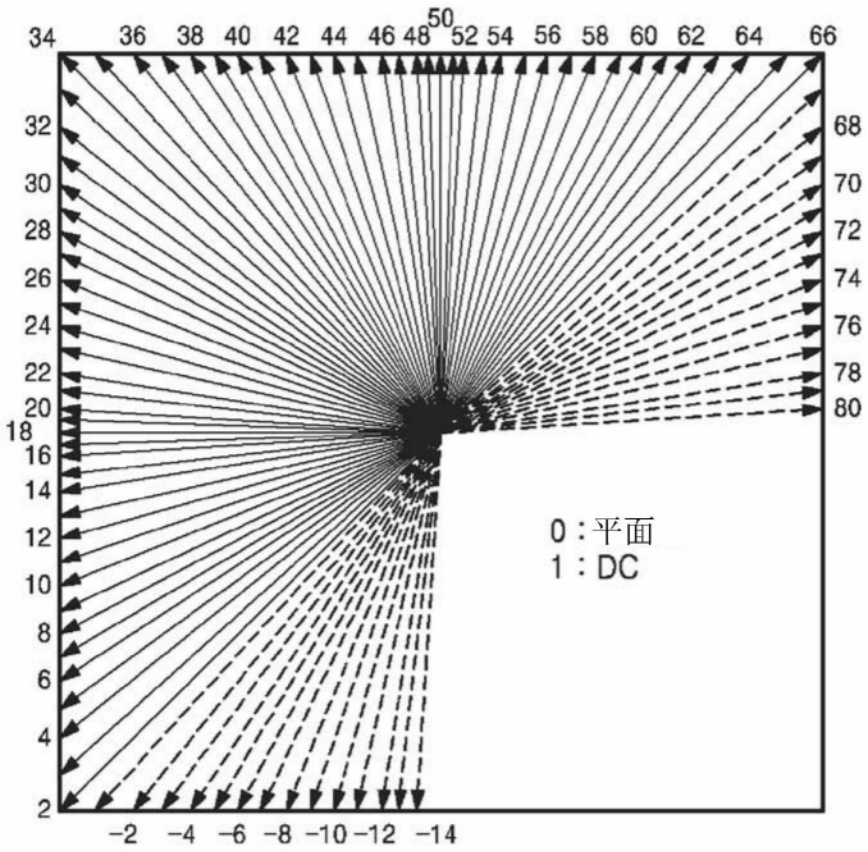

34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 64 66
32 30 28 26 24 22 20 18 16 14 12 10 8 6 4 2
68 70 72 74 76 78 80
-2 -4 -6 -8 -10 -12 -14
0：平面
1：DC

图3b

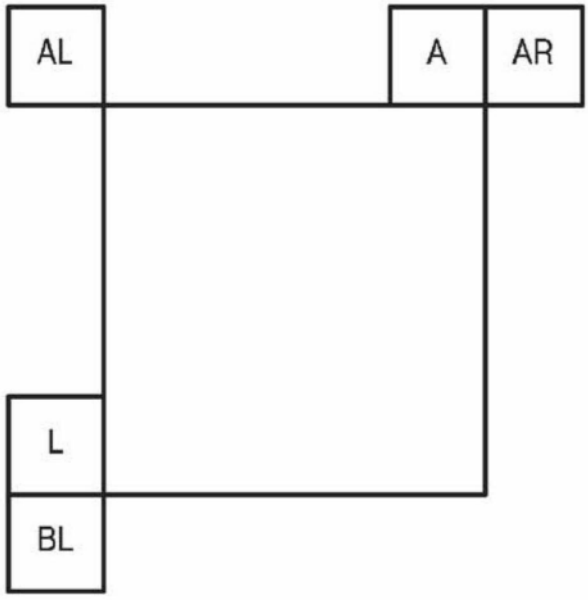

AL
A
AR
L
BL

图4

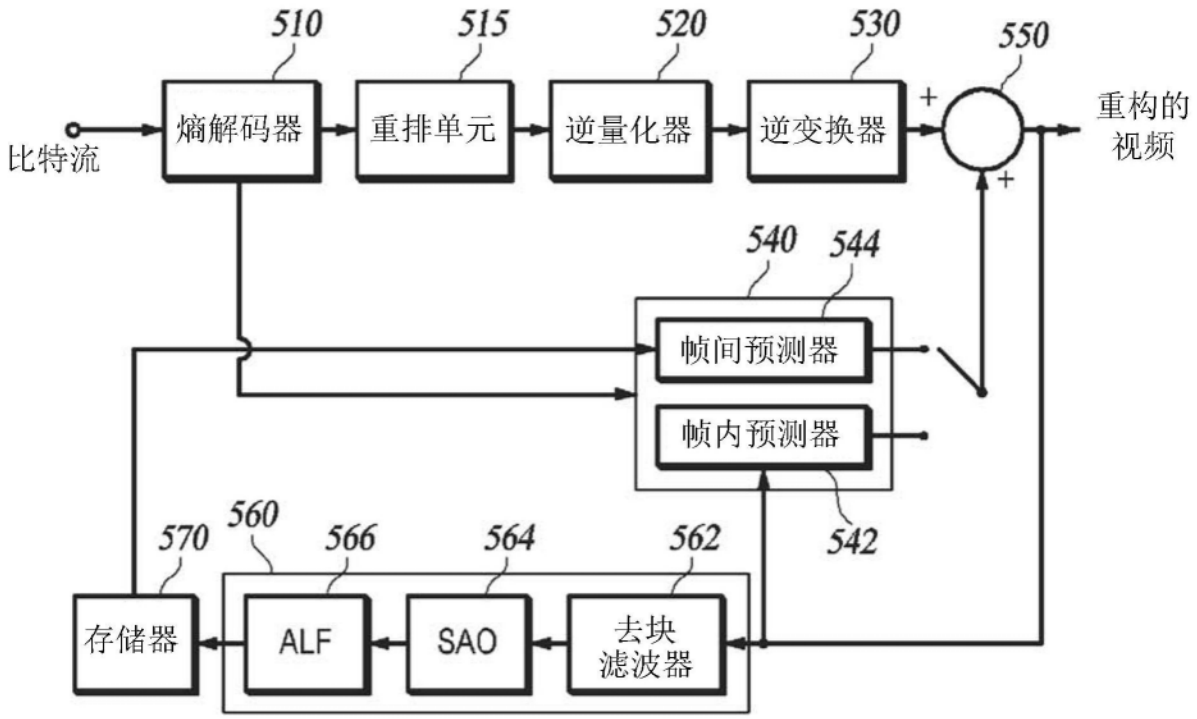

510
515
520
530
550
比特流
熵解码器
重排单元
逆量化器
逆变换器
+
+
重构的
视频
540
544
帧间预测器
帧内预测器
542
570
560
566
564
562
存储器
ALF
SAO
去块
滤波器

图5

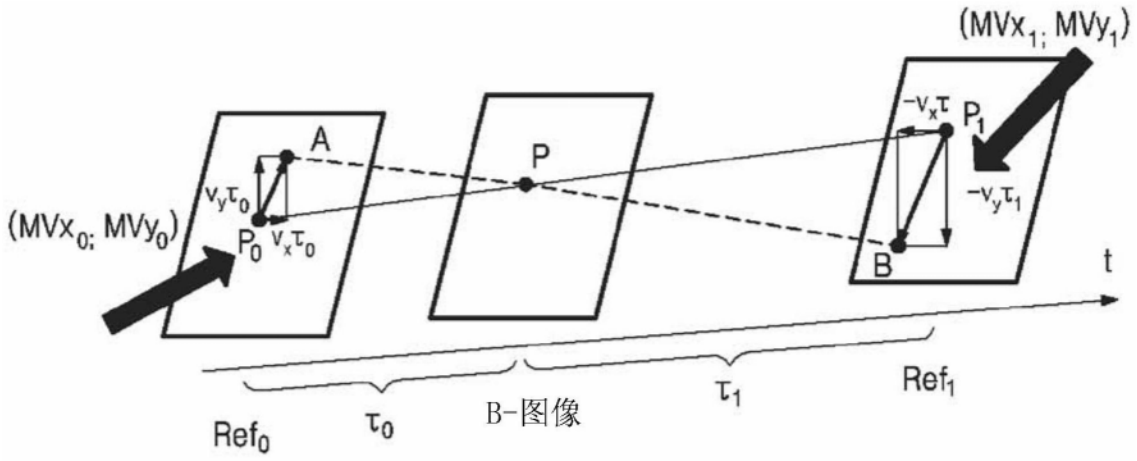

(MVx1; MVy1)
A
P
P1
B
(MVx0; MVy0)
P0
t
Ref0
τ0
B-图像
τ1
Ref1

图6

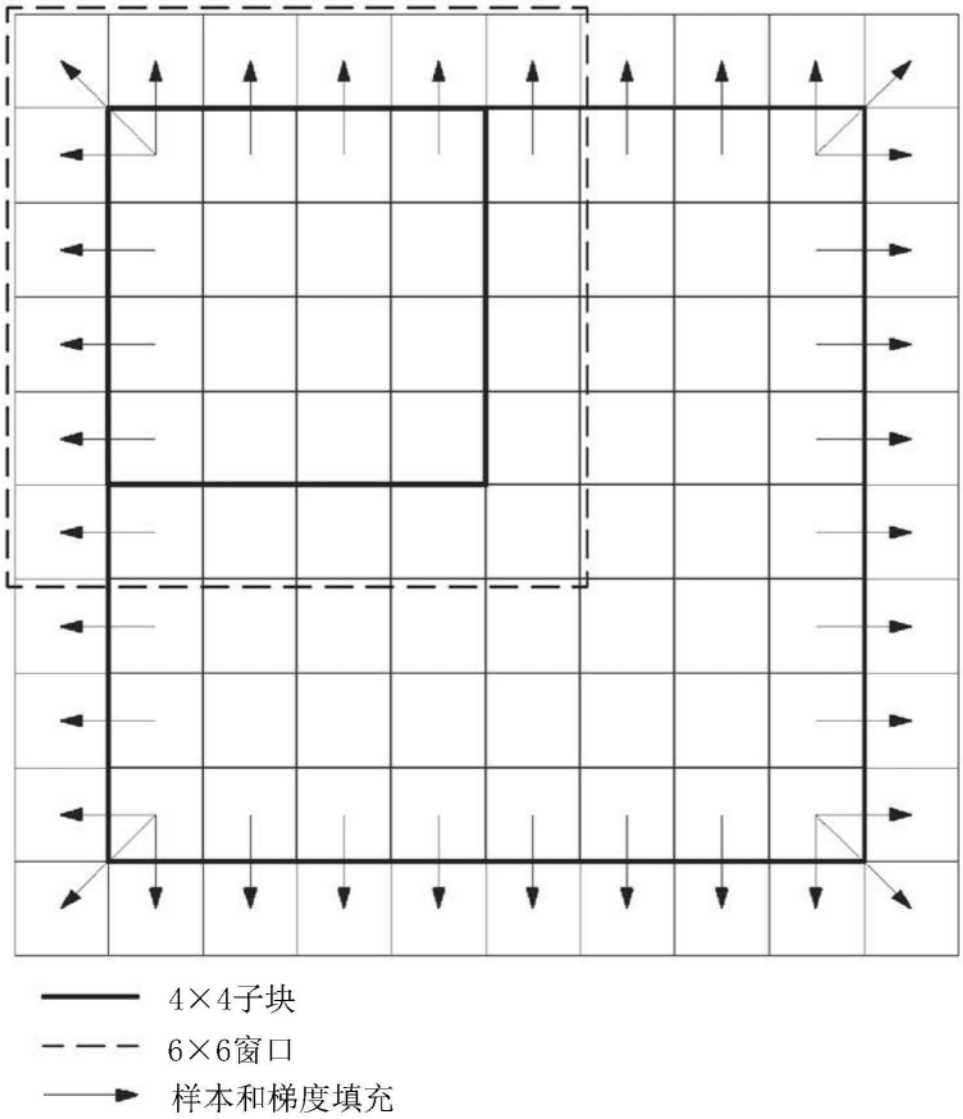


图7

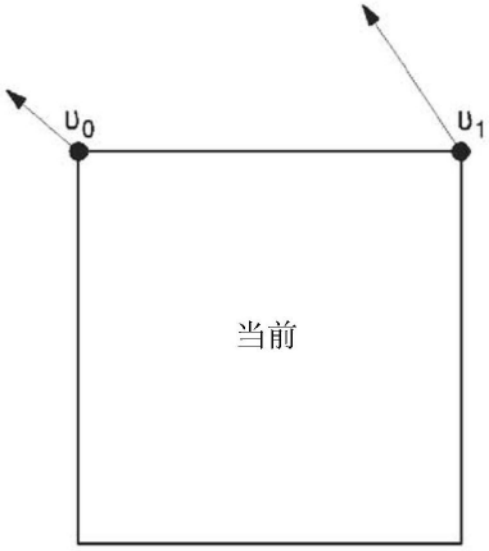


图8a

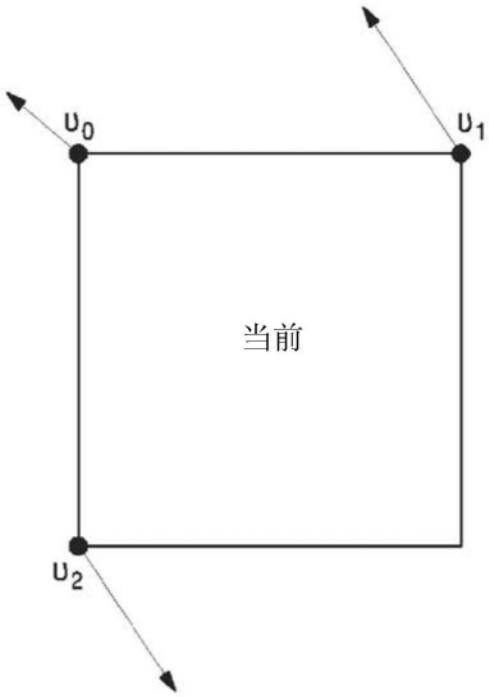


图8b

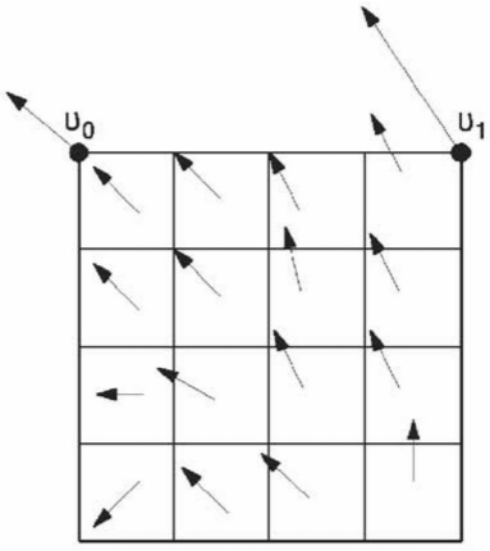


图9

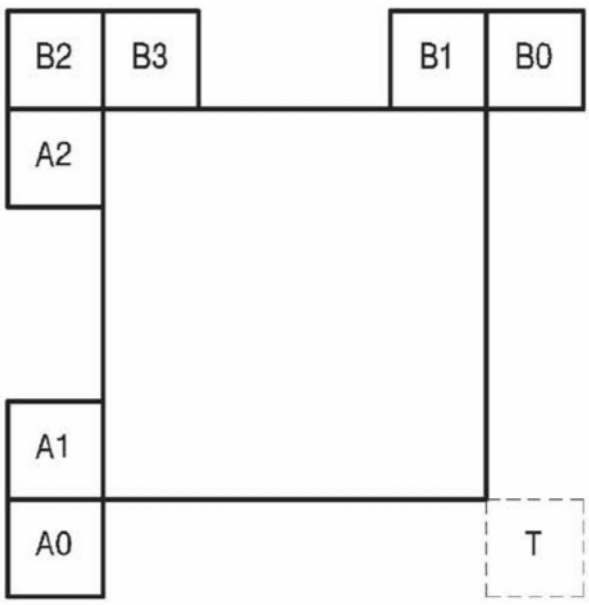


图10

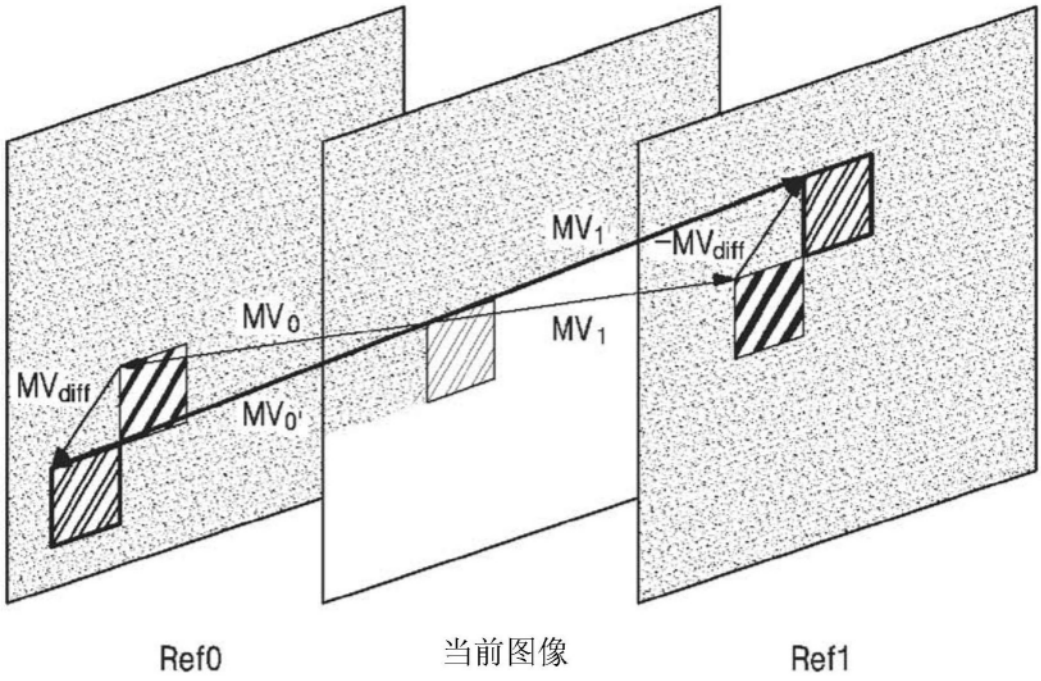

MV1
-MVdiff
MV0
MV1
MVdiff
MV0'
Ref0
当前图像
Ref1

图11

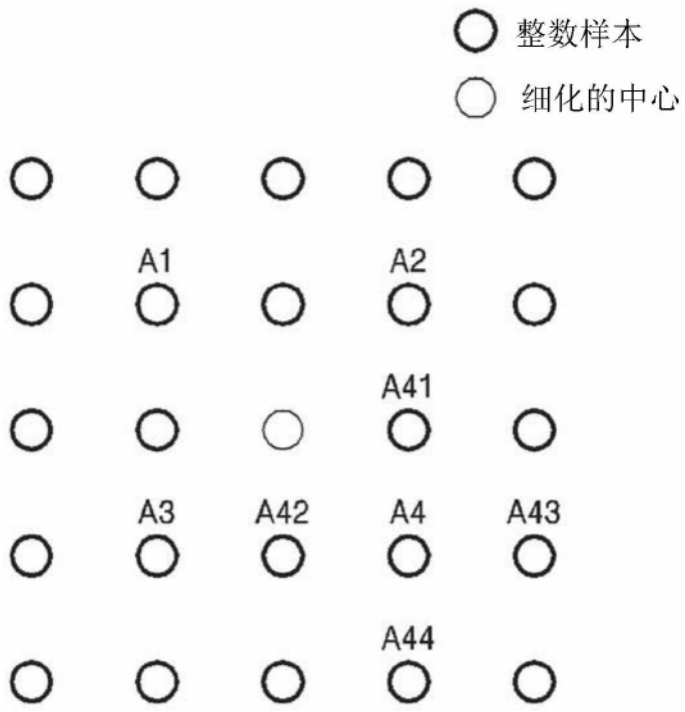

整数样本
细化的中心
A1
A2
A41
A3
A42
A4
A43
A44

图12

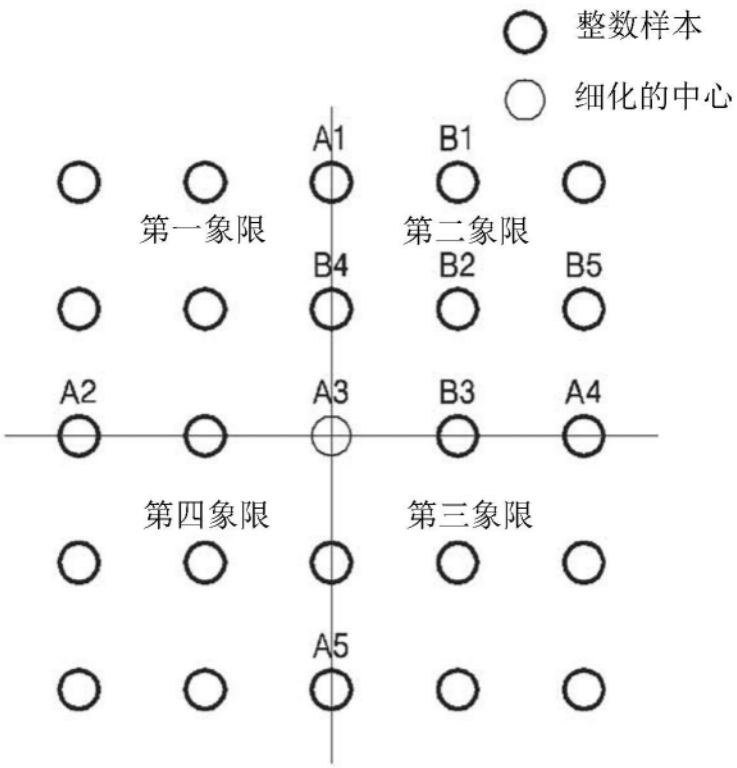

整数样本
细化的中心
A1
B1
第一象限
第二象限
B4
B2
B5
A2
A3
B3
A4
第四象限
第三象限
A5

图13

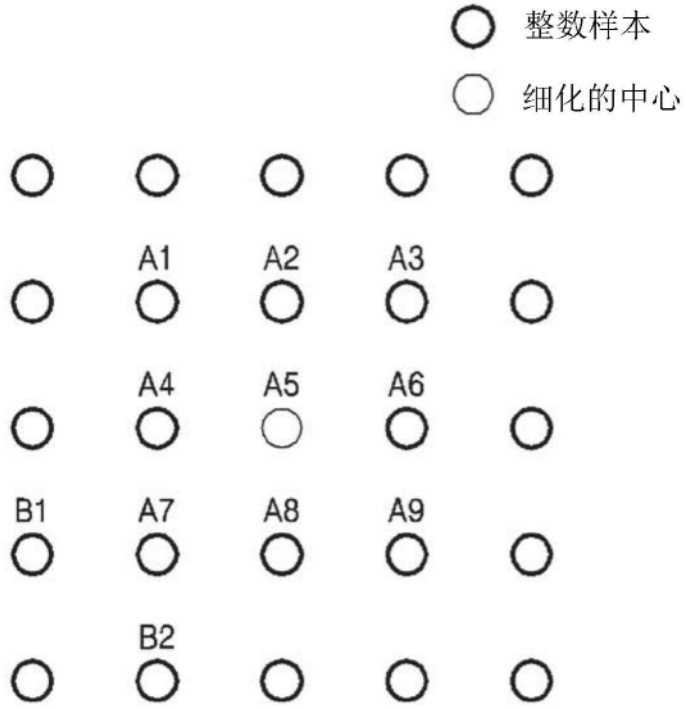

整数样本
细化的中心
A1
A2
A3
A4
A5
A6
B1
A7
A8
A9
B2

图14