



(12) 发明专利

(10) 授权公告号 CN 111479110 B

(45) 授权公告日 2022.12.13

(21) 申请号 202010293694.8
 (22) 申请日 2020.04.15
 (65) 同一申请的已公布的文献号
 申请公布号 CN 111479110 A
 (43) 申请公布日 2020.07.31
 (73) 专利权人 郑州轻大产业技术研究院有限公司
 地址 450000 河南省郑州市高新技术产业
 开发区长椿路6号西美大厦A座20-23
 层
 专利权人 郑州轻工业大学
 (72) 发明人 张秋闻 黄立勋 蒋斌 王祎菡
 赵进超 吴庆岗 常化文 王晓
 张伟伟 赵永博 崔腾耀 郭睿
 孟颖辉 李祖贺 黄伟 甘勇
 (74) 专利代理机构 郑州优质知识产权代理有限
 公司 41125
 专利代理师 张真真

(51) Int.Cl.
 H04N 19/109 (2014.01)
 H04N 19/119 (2014.01)
 H04N 19/51 (2014.01)
 H04N 19/543 (2014.01)
 H04N 19/567 (2014.01)
 (56) 对比文件
 CN 104320658 A, 2015.01.28
 CN 1934871 A, 2007.03.21
 CN 110213584 A, 2019.09.06
 WO 2018124332 A1, 2018.07.05
 吴庆岗等. 基于随机森林和多特征融合的青
 苹果图像分割. 《信阳师范学院学报(自然科学
 版)》. 2018, 681-686.
 审查员 臧萍

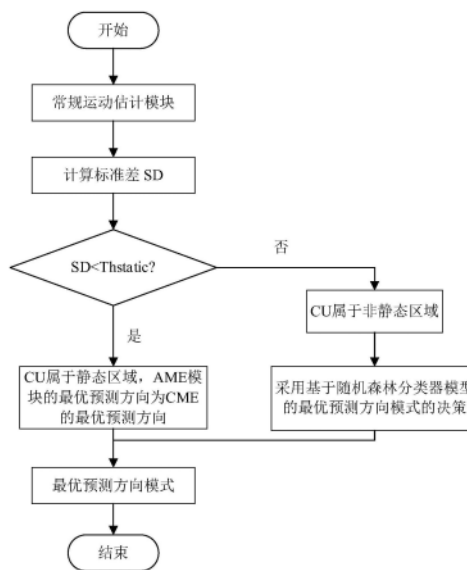
权利要求书2页 说明书11页 附图4页

(54) 发明名称

针对H.266/VVC的快速仿射运动估计方法

(57) 摘要

本发明提出了一种针对H.266/VVC的快速仿射运动估计方法,其步骤为:利用标准差计算当前CU的纹理复杂度,并根据纹理复杂度将当前CU分为静态区域或非静态区域;对于静态区域的CU,跳过仿射运动估计,直接利用运动估计对当前CU进行预测,并通过率失真优化的方法选择最佳的预测方向模式;对于非静态区域的CU,利用训练好的随机森林分类器RFC模型对当前CU进行分类,输出最佳的预测方向模式。对于静态区域的CU,本发明跳过仿射运动估计,降低了计算复杂度;对于非静态区域的CU,本发明通过提前训练好的模型直接进行预测方向模式的预测,避免了仿射运动估计的计算,从而降低仿射运动估计模块的复杂度。



CN 111479110 B

1. 一种针对H.266/VVC的快速仿射运动估计方法,其特征在于,其步骤如下:

S1、利用标准差计算当前CU的纹理复杂度SD,并根据纹理复杂度SD将当前CU分为静态区域或非静态区域;

S2、对于静态区域的CU,跳过仿射运动估计AME,直接利用运动估计CME对当前CU进行预测,并通过率失真优化的方法选择最佳的预测方向模式;

S3、对于非静态区域的CU,利用训练好的随机森林分类器RFC模型对当前CU进行分类,输出最佳的预测方向模式。

2. 根据权利要求1所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述利用标准差计算当前CU的纹理复杂度SD的方法为:

$$SD = \sqrt{\frac{1}{W \times H} \sum_{a=1}^W \sum_{b=1}^H P(a, b)^2 - \left(\frac{1}{W \times H} \sum_{a=1}^W \sum_{b=1}^H P(a, b) \right)^2},$$

其中,W代表CU的宽度,H代表CU的高度,P(a,b)表示在CU中位置为(a,b)的像素值。

3. 根据权利要求1所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述利用运动估计CME对当前CU进行预测,并通过率失真优化的方法选择最佳的预测方向模式的方法为:

S21、当前CU首先经过单向预测Uni-L0,然后经过单向预测Uni-L1,最后经过双向预测Bi;

S22、利用率失真优化分别计算步骤S21中的当前CU分别经过单向预测Uni-L0、单向预测Uni-L1和双向预测Bi的率失真代价;

S23、将率失真代价最小的预测模式作为最佳的预测方向模式。

4. 根据权利要求3所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述单向预测Uni-L0、单向预测Uni-L1的率失真代价的计算方法均为:

$$\arg \min_{j \in \Gamma(\phi) | \phi = L0, L1} \{J(\phi(j))\},$$

所述双向预测Bi的率失真代价的计算方法为:

$$\arg \min_{j \in \Gamma(\phi) | \phi = L0, L1} \left\{ \frac{J(\phi(j) | \phi = L0)}{2} + \frac{J(\phi(j) | \phi = L1)}{2} \right\},$$

其中, $\Gamma(\phi)$ 表示所有可用参考列表集合, ϕ 表示参考列表集,L0和L1表示两个参考帧列表, $\phi(j)$ 表示参考列表中的参考帧, $J(\cdot)$ 为率失真代价函数,且 $J(\phi(j)) = D(\phi(j)) + \lambda \cdot R(\phi(j))$, $D(\cdot)$ 表示CU编码的失真程度, λ 表示拉格朗日乘子, $R(\cdot)$ 表示CU编码消耗的比特数。

5. 根据权利要求1所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述步骤S3中的随机森林分类器RFC模型的训练方法为:

S31、从通用测试序列中选用不同分辨率下的Traffic、Kimono、BQSquare、RaceHorseC、和FourPeople视频序列,在VTM上分别编码前M帧,同时记录VTM中CU的形状、CU的纹理复杂度及CU的三种预测方向模式作为数据集,数据集包括样本集S和测试集T,其中,三种预测方向模式包括单向预测Uni-L0、单向预测Uni-L1和双向预测Bi;

S32、利用Bootstrap法重采样样本集S,生成K个训练样本集 $\{S_1^*, S_2^*, \dots, S_K^*\}$,将生成的

每个训练集 S_i^* 作为根节点,生成对应的决策树 $\{T_1, T_2, \dots, T_K\}$, 其中, $i=1, 2, \dots, K$ 表示第 i 个训练样本, K 表示训练样本集的大小;

S33、从根节点开始训练,在决策树的每个中间节点上随机选择 m 个特征属性,计算每个特征属性的Gini指标系数,从中选择Gini指标系数最小的特征属性作为当前节点的最优分裂属性,以最小Gini指标系数为分裂阈值,将 m 个特征属性划分为左子树、右子树;

S34、重复步骤S33,训练 K' 次,直到 K' 棵决策树训练完成,每棵决策树都完整生长而不进行剪枝;

S35、生成的多棵决策树即为随机森林分类器RFC模型,并利用随机森林分类器RFC模型对测试集 T 进行判别分类,分类结果采用投票方式,将 K' 棵决策树输出最多的类别作为测试集 T 的所属类别,得到当前CU的最佳的预测方向模式。

6. 根据权利要求5所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述步骤S31中获得数据集的方法为:

S31.1、利用运动估计CME对视频序列进行预测;

S31.2、利用4参数仿射运动模型对步骤S31.1中预测后的视频序列进行仿射预测,其中,仿射预测包括单向预测Uni-L0、单向预测Uni-L1和双向预测Bi;

S31.3、利用6参数仿射运动模型对步骤S31.2中仿射预测后的视频序列进行仿射预测;

S31.4、分别计算步骤S31.2和S31.3进行仿射预测后的率失真代价,将最小的率失真代价对应的预测模式为视频序列的预测方向模式。

7. 根据权利要求5所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述特征属性包括二维哈尔小波变换水平系数、二维哈尔小波变换垂直系数、二维哈尔小波变换角度系数、角二阶矩、对比度、熵、逆差矩、最小差值和和梯度。

8. 根据权利要求6所述的针对H.266/VVC的快速仿射运动估计方法,其特征在于,所述4参数仿射运动模型,CU中样本位置 (x, y) 的运动矢量为:

$$\begin{cases} mv_x = \frac{mv_{1x} - mv_{0x}}{W} x + \frac{mv_{1y} - mv_{0y}}{W} y + mv_{0x} \\ mv_y = \frac{mv_{1y} - mv_{0y}}{W} x + \frac{mv_{1x} - mv_{0x}}{W} y + mv_{0y} \end{cases},$$

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, W 表示CU的宽;

所述6参数仿射运动模型,CU中样本位置 (x, y) 的运动矢量为:

$$\begin{cases} mv_x = \frac{mv_{1x} - mv_{0x}}{W} x + \frac{mv_{2x} - mv_{0x}}{H} y + mv_{0x} \\ mv_y = \frac{mv_{1y} - mv_{0y}}{W} x + \frac{mv_{2y} - mv_{0y}}{H} y + mv_{0y} \end{cases},$$

其中, (mv_{2x}, mv_{2y}) 是左下角的运动矢量控制点, H 表示CU的高。

针对H.266/VVC的快速仿射运动估计方法

技术领域

[0001] 本发明涉及图像处理技术领域,特别是指一种针对H.266/VVC的快速仿射运动估计方法。

背景技术

[0002] 在如今的信息化时代,三维影像、超高清视频和虚拟现实等视频服务需求日益增长,高清晰度视频的编码和传输日益成为研究的热点问题。随着H.266/VVC标准的发展与完善,视频处理效率的提高也带动了视频产业的发展,为新一代视频编码技术的发展奠定了基础。高度密集的数据给带宽和存储带来巨大挑战,当前主流的视频编码标准开始无法满足目前新兴的应用,因此,新一代视频编码标准H.266/VVC应运而生,满足人们对视频的清晰度、流畅度以及实时度的要求。国际标准化组织ISO/IEC MPEG和ITU-T VCEG成立了联合视频小组(Joint Video Exploration Team, JVET),负责进行下一代视频编码标准H.266/通用视频编码(Versatile Video Coding, VVC)的开发。H.266/VVC的制定是针对4K及以上的高清视频,位深以10比特为主,这与H.265/HEVC的定位不同,这导致目前编码器的最大块尺寸变为128,编码中间处理的像素都为10比特,即使输入的8比特的序列,都会转为10比特处理。

[0003] H.266/VVC使用混合编码技术框架,图像划分从单一、固定划分不断朝着多样、灵活的划分结构发展,能够更加高效的适配高分辨率图像的编解码处理。此外,H.266/VVC针对新一代视频数据对原有H.265/HEVC编码器的帧间-帧内预测,预测信号滤波,变换,量化/缩放,熵编码等新元素进行了扩展,并考虑新一代视频编码标准的特性,添加了新的模型预测模式。具体地,H.266/VVC沿用了高效视频编码H.265/HEVC帧间编码的运动估计、运动补偿和运动矢量预测技术,并在此基础上引入了一些新的技术。如对Merge模式进行扩展添加基于历史的预测运动矢量,添加了新的预测方法如仿射变换技术,自适应运动矢量精度方法,1/16采样精度的运动预测补偿等。众多先进编码工具的引入,大大提高新一代视频编码标准H.266/VVC的编码效率。但也因率失真成本计算而显著提高H.266/VVC帧间编码的运算复杂性,从而显著降低了新一代视频的编码速度。

[0004] 帧间预测的主要原理是为当前图像的每个像素块在之前已经编码的图像中找一个最佳匹配块,该过程称为运动估计ME,其中,用于预测的图像称为参考图像,参考块为参考图像中最佳匹配块,即参考像素块,参考块到当前像素块的位移称为运动矢量MV,当前像素块与参考块的差值称为预测残差。其中,运动估计ME算法是H.266/VVC视频编码过程中最为关键的算法,它占用了整个视频编码一半以上的计算量和绝大部分的运算时间,是决定视频压缩效率的主导因素。运动估计ME通过有效地去除连续图像间的时间冗余而成为视频压缩技术中的一个研究热点。为了提高压缩效率,最近的视频编解码器尝试估计不同形状和大小的运动。此外,通过添加多类型树,可以对非常薄的块执行运动估计ME(例如,宽度是高度的八分之一)。因此,在多类型树(multi-type trees, MTT)的各个模块中,运动估计ME是VVC中编码复杂度最高的工具。由于在MTT的精细划分块中递归地执行更高级的帧间预测

方案,运动估计ME的计算复杂度甚至比HEVC中增加更多,因为未来视频编码(Future video coding,FVC)中ME还尝试了仿射运动估计等新技术。仿射运动估计AME以旋转和缩放等非平移运动为特征,以牺牲较高的编码复杂度为代价,在率失真(Rate distortion,RD)性能上是有效的。在整个运动估计ME处理时间中,仿射运动估计AME的计算复杂度占很大一部分,因此降低其复杂度是非常重要的。因此,要降低VTM编码器的复杂度,需要加快AME模块。

[0005] 事实上,针对H.265/HEVC帧间预测复杂度高的问题,许多文献已经进行了大量的研究。J.Xiong等人根据椎体运动散度提出了一种快速CU选择算法,该算法可以提前跳过H.265/HEVC中个别帧间CU。L.Shen等人提出了一种用于H.265/HEVC的自适应模式间决策算法,该算法联合利用了层间和时空相关性,基于统计分析提出了早期跳过模式决策,基于预测大小相关的模式决策和基于速率失真(RD)成本相关的模式决策三种方法。H.Lee等人提出了在计算 $2N \times 2N$ 的Merge模式的RD代价后,利用其失真特性提出了一种早期的跳过模式决策方法。Q.Zhang等人针对纹理视频和深度图内容的高相关性,提出了提前判决编码单元深度级别和自适应模式判决方法,用以降低视频编码的计算复杂度。Q.Hu等人提出了一种基于Neyman-Pearson规则的快速帧间模式决策算法,该算法包括早期的SKIP模式决策和快速的CU大小决策来降低H.265/HEVC复杂度。Z.Pan等人提出了一种基于内容相似度的快速参考帧选择算法,以减少基于多个参考帧的帧间预测的计算复杂性。Z.Pan等人基于不同尺寸预测模式之间最佳的运动矢量选择相关性,提出一种快速运动估计ME方法,以降低H.265/HEVC编码器的编码复杂度。J.Zhang等人提出了一种基于贝叶斯方法和条件随机场的两阶段快速帧间CU决策方法,以降低HEVC编码器的编码复杂度。T.S.Kim等人提出了一种基于HEVC的快速运动估计算法,该算法支持高度灵活的块分区结构。通过搜索多个精确运动矢量预测周围的狭窄区域,该算法大大降低了其计算复杂性。C.Ma等人提出了一种基于神经网络的算术编码方法,对HEVC中的帧间预测信息进行编码。L.Shen等人提出了一种快速模式决策算法来减少编码器的计算复杂度。所提出方法利用了三个调整参数后的优化编码器,即SKIP/Merge模式的条件概率,运动特性和模式复杂度。D.Wang等人提出了快速的深度级别和帧间模式预测算法。该算法使用层间相关性,空间相关性及其相关程度来加快HEVC帧间编码。以上的快速帧间方法都能够保证编码性能的同时,有效降低H.265/HEVC的计算复杂度。但是,这些方法并非是为H.266/VVC编码器设计的,而H.266/VVC编码器采用了新帧间预测技术,如采用更先进的仿射运动补偿预测、扩展Merge模式、自适应运动矢量精度、三角划分模式等技术。基于此,H.266/VVC与基于H.265/HEVC帧间预测的空间和层间相关性必然存在着较大的差别,因而需要重新研究基于H.266/VVC的低复杂度帧间编码方法。

[0006] 针对H.266/VVC帧间编码复杂度高的问题,极少一部分文献对此进行了探索。S.Park等人提出了一种有效限制正常运动估计以及仿射运动估计的参考帧搜索范围的方法,它主要利用了H.266/VVC预测结构内的依赖性。该方法基于父节点的预测信息来最小化CU的参考帧搜索范围的最大值来降低编码复杂度。Z.Wang等人提出了一种基于置信区间的四叉树加二叉树(quadtree plus binary tree,QTBT)划分结构的提前终止方案,建立了基于运动发散场的率失真(rate distortion,RD)模型,来估计每个分区模式的率失真RD成本;并基于该模型早期终止了H.266/VVC的块划分,以消除不必要的分区迭代,使H.266/VVC编码性能和编码复杂度之间取得良好的平衡。Z.Wang等人提出了一种面向卷积神经网络(convolutional neural networks,CNN)的快速QTBT分区决策算法,用于H.266/VVC帧间编

码,该算法以统计方式来分析QTBT,从而设计卷积神经网络CNN的体系结构,并利用时间相关性来控制错误预测风险,以提高卷积神经网络CNN方案的鲁棒性。D.García-Lucas等人提出了一种用于提取帧运动信息的预分析算法,该算法在运动估计模块中用于加速H.266/VVC编码器。S.Park等人提出了一种快速H.266/VVC帧间编码方法,以有效降低使用多类型树MTT时VTM中仿射运动估计的编码复杂度。该方法包括两个过程:提前终止方案和减少仿射运动估计的参考帧的数量。H.Gao等人提出了一种低复杂度的解码器侧运动矢量细化方案,通过在先前解码的参考图片中搜索匹配成本最小的块,从Merge模式中优化初始运动矢量MV,并被添加到的基于双边匹配的解码器侧运动矢量细化方法中。N.Tang等人针对H.266/VVC帧间编码提出了一种快速块划分算法,使用三帧差来判断当前块是否为静态对象;当前块为静止时,无需进一步拆分,从而提前终止分区以提高帧间编码速度。然而,在VVC中减轻仿射运动估计AME复杂度的工作很少。对于VTM,有很大的空间进一步降低多类型树MTT结构中的运动估计ME复杂度,特别是在仿射运动估计AME中。

发明内容

[0007] 针对上述背景技术中存在的不足,本发明提出了一种针对H.266/VVC的快速仿射运动估计方法,解决了在VTM中的仿射运动估计AME编码复杂度高的技术问题。

[0008] 本发明的技术方案是这样实现的:

[0009] 一种针对H.266/VVC的快速仿射运动估计方法,其步骤如下:

[0010] S1、利用标准差计算当前CU的纹理复杂度SD,并根据纹理复杂度SD将当前CU分为静态区域或非静态区域;

[0011] S2、对于静态区域的CU,跳过仿射运动估计AME,直接利用运动估计CME对当前CU进行预测,并通过率失真优化的方法选择最佳的预测方向模式;

[0012] S3、对于非静态区域的CU,利用训练好的随机森林分类器RFC模型对当前CU进行分类,输出最佳的预测方向模式。

[0013] 所述利用标准差计算当前CU的纹理复杂度SD的方法为:

$$[0014] \quad SD = \sqrt{\frac{1}{W \times H} \sum_{a=1}^W \sum_{b=1}^H P(a, b)^2 - \left(\frac{1}{W \times H} \sum_{a=1}^W \sum_{b=1}^H P(a, b) \right)^2},$$

[0015] 其中,W代表CU的宽度,H代表CU的高度,P(a,b)表示在CU中位置为(a,b)的像素值。

[0016] 所述利用运动估计CME对当前CU进行预测,并通过率失真优化的方法选择最佳的预测方向模式的方法为:

[0017] S21、当前CU首先经过单向预测Uni-L0,然后经过单向预测Uni-L1,最后经过双向预测Bi;

[0018] S22、利用率失真优化分别计算步骤S21中的当前CU分别经过单向预测Uni-L0、单向预测Uni-L1和双向预测Bi的率失真代价;

[0019] S23、将率失真代价最小的预测模式作为最佳的预测方向模式。

[0020] 所述单向预测Uni-L0、单向预测Uni-L1和双向预测Bi的率失真代价的计算方法均为:

$$[0021] \quad \arg \min_{j \in \Gamma(\phi) | \phi = L0, L1} \{J(\phi(j))\},$$

$$[0022] \quad \arg \min_{j \in \Gamma(\phi) | \phi=L0, L1} \left\{ \frac{J(\phi(j) | \phi=L0)}{2} + \frac{J(\phi(j) | \phi=L1)}{2} \right\},$$

[0023] 其中, $\Gamma(\phi)$ 表示所有可用参考列表集合, ϕ 表示参考列表集, L0和L1表示两个参考帧列表, $\phi(j)$ 表示参考列表中的参考帧, $J(\cdot)$ 为率失真代价函数, 且 $J(\phi(j)) = D(\phi(j)) + \lambda \cdot R(\phi(j))$, $D(\cdot)$ 表示CU编码的失真程度, λ 表示拉格朗日乘子, $R(\cdot)$ 表示CU编码消耗的比特数。

[0024] 所述步骤S3中的随机森林分类器RFC模型的训练方法为:

[0025] S31、从通用测试序列中选用不同分辨率下的Traffic、Kimono、BQSquare、RaceHorseC、和FourPeople视频序列, 在VTM上分别编码前M帧, 同时记录VTM中CU的形状、CU的纹理复杂度及CU的三种预测方向模式作为数据集, 数据集包括样本集S和测试集T, 其中, 三种预测方向模式包括单向预测Uni-L0、单向预测Uni-L1和双向预测Bi;

[0026] S32、利用Bootstrap法重采样样本集S, 生成K个训练样本集 $\{S_1^*, S_2^*, \dots, S_K^*\}$, 将生成的每个训练集 S_i^* 作为根节点, 生成对应的决策树 $\{T_1, T_2, \dots, T_K\}$, 其中, $i=1, 2, \dots, K$ 表示第i个训练样本, K表示训练样本集的大小;

[0027] S33、从根节点开始训练, 在决策树的每个中间节点上随机选择m个特征属性, 计算每个特征属性的Gini指标系数, 从中选择Gini指标系数最小的特征属性作为当前节点的最优分裂属性, 以最小Gini指标系数为分裂阈值, 将m个特征属性划分为左子树、右子树;

[0028] S34、重复步骤S33, 训练K' 次, 直到K' 棵决策树训练完成, 每棵决策树都完整生长而不进行剪枝;

[0029] S35、生成的多棵决策树即为随机森林分类器RFC模型, 并利用随机森林分类器RFC模型对测试集T进行判别分类, 分类结果采用投票方式, 将K' 棵决策树输出最多的类别作为测试集T的所属类别, 得到当前CU的最佳的预测方向模式。

[0030] 所述步骤S31中获得数据集的方法为:

[0031] S31.1、利用运动估计CME对视频序列进行预测;

[0032] S31.2、利用4参数仿射运动模型对步骤S31.1中预测后的视频序列进行仿射预测, 其中, 仿射预测包括单向预测Uni-L0、单向预测Uni-L1和双向预测Bi;

[0033] S31.3、利用6参数仿射运动模型对步骤S31.2中仿射预测后的视频序列进行放射预测;

[0034] S31.4、分别计算步骤S31.2和S31.3进行仿射预测后的率失真代价, 将最小的率失真代价对应的预测模式为视频序列的预测方向模式。

[0035] 所述特征属性包括二维哈尔小波变换水平系数、二维哈尔小波变换垂直系数、二维哈尔小波变换角度系数、角二阶矩、对比度、熵、逆差矩、最小差值和和梯度。

[0036] 所述4参数仿射运动模型, CU中样本位置 (x, y) 的运动矢量为:

$$[0037] \quad \begin{cases} mV_x \frac{mV_{1x} - mV_{0x}}{W} X + \frac{mV_{1y} - mV_{0y}}{W} Y + mV_{0x} \\ mV_y \frac{mV_{1y} - mV_{0y}}{W} X + \frac{mV_{1x} - mV_{0x}}{W} Y + mV_{0y} \end{cases},$$

[0038] 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, W 表示CU的宽;

[0039] 所述6参数仿射运动模型, CU中样本位置 (x, y) 的运动矢量为:

$$[0040] \begin{cases} mv_x \frac{mv_{1x} - mv_{0x}}{W} x + \frac{mv_{2x} - mv_{0x}}{H} y + mv_{0x} \\ mv_y \frac{mv_{1y} - mv_{0y}}{W} x + \frac{mv_{2y} - mv_{0y}}{H} y + mv_{0y} \end{cases},$$

[0041] 其中, (mv_{2x}, mv_{2y}) 是左下角的运动矢量控制点, H 表示CU的高。

[0042] 本技术方案能产生的有益效果: 本发明首先利用标准差SD将CU分为静态区域和非静态区域, 如果CU属于静态区域, 选择SKIP模式进行帧间预测的概率较高, 并且倾向于选择SKIP模式进行帧间预测的静态区域不需要进行仿射预测, 因此, 在静态区域可以提前终止仿射运动估计AME模块, 并且当前CU的最佳方向模式为运动估计CME的最佳方向模式; 如果CU属于非静态区域, 则根据随机森林分类模型判断CU的帧间预测模式, 最终提前得到最优的预测方向模式; 因此, 本发明降低了计算复杂度并节省了编码时间, 从而实现H.266/VVC的快速编码。

附图说明

[0043] 为了更清楚地说明本发明实施例或现有技术中的技术方案, 下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍, 显而易见地, 下面描述中的附图仅仅是本发明的一些实施例, 对于本领域普通技术人员来讲, 在不付出创造性劳动的前提下, 还可以根据这些附图获得其他的附图。

[0044] 图1为本发明的流程图;

[0045] 图2为本发明的预测方向模式复杂度分布图;

[0046] 图3为本发明的4-参数仿射模型;

[0047] 图4为本发明的6-参数仿射模型;

[0048] 图5为本发明的运动估计ME的整体过程图;

[0049] 图6为本发明方法与FAME方法的整体运行时间对比结果图。

具体实施方式

[0050] 下面将结合本发明实施例中的附图, 对本发明实施例中的技术方案进行清楚、完整地描述, 显然, 所描述的实施例仅仅是本发明一部分实施例, 而不是全部的实施例。基于本发明中的实施例, 本领域普通技术人员在没有付出创造性劳动前提下所获得的所有其他实施例, 都属于本发明保护的范围。

[0051] 如图1所示, 本发明实施例提供了一种针对H.266/VVC的快速仿射运动估计方法, 具体步骤如下:

[0052] S1、在图像编码过程中, 单一区域的图像内容往往采用较大的CU进行编码。相反, 具有丰富细节的区域通常使用较小的CU进行编码。由此可知, 利用编码块的纹理复杂度程度决定CU是否使用SKIP模式进行帧间预测。在图像编码过程中, 图像内容单一的区域更趋向于使用SKIP模式进行帧间预测进行编码, 而细节丰富的区域很小几率使用SKIP模式进行

帧间预测。一个CU的方差代表当前块的两个像素间能量的分散程度,因此一个块的纹理复杂度可以粗略的用它标准差SD来衡量,因此,利用标准差计算当前CU的纹理复杂度SD,并根据纹理复杂度SD将当前CU分为静态区域或非静态区域;标准差的公式为:

$$[0053] \quad SD = \sqrt{\frac{1}{W \times H} \sum_{a=1}^W \sum_{b=1}^H P(a, b)^2 - \left(\frac{1}{W \times H} \sum_{a=1}^W \sum_{b=1}^H P(a, b) \right)^2},$$

[0054] 其中,W代表CU的宽度,H代表CU的高度,P(a,b)表示在CU中位置为(a,b)的像素值。由于相邻块的纹理复杂度与CU具有相关性,通过相邻块的纹理复杂度来导出分类的阈值。根据大量的实验数据,将CU的相邻块的标准差SD中的最小值作为 Th_{static} 是合理的。通过阈值可以将CU进行分类。如果当前标准差SD小于阈值 Th_{static} ,则表明当前CU是静态区域。相反,如果标准差SD的值大于 Th_{static} ,则当前的CU属于非静态区域。

[0055] S2、现有的视频编码标准(如H.265/HEVC)对于运动估计CME,使用覆盖平移运动的运动矢量MV,然而,仿射运动估计AME不仅可以预测平移运动,还可以预测线性变换运动,如缩放和旋转。如果相机缩放或旋转以捕获视频,仿射运动估计AME比运动估计CME更准确地预测运动。在H.266/VVC中,仿射运动估计AME和运动估计CME一样,从单向预测Uni-prediction L0开始,然后是单向预测Uni-prediction L1,最后是双向预测Bi-prediction。在计算了三种预测方向模式后,利用率失真优化(rate distortion optimization,RDO)的方法选择最佳的预测方向模式。图2示出了仿射运动估计AME帧间预测模式的复杂度的分布,并且单向预测Uni-prediction L0比单向预测Uni-prediction L1需要更多的预测时间。如果单向预测Uni-prediction L0和单向预测Uni-prediction L1的参考帧不同,单向预测的编码复杂度是一样。否则,如果单向预测Uni-prediction L0和单向预测Uni-prediction L1之间的参考帧相同,则从单向预测Uni-prediction L0的运动矢量复制单向预测Uni-prediction L1的运动矢量MV,以避免冗余仿射运动估计AME处理。因此,单向预测Uni-prediction L0预测比单向预测Uni-prediction L1消耗更多的预测时间。虽然在单向预测中需要最大的编码复杂度,但双向预测模式作为最佳帧间预测模式的概率较高。在仿射运动估计AME模块中,计算率失真RD代价是导致复杂度大的重要原因。

[0056] 为了获得最佳运动矢量MV和最佳参考帧,编码器搜索多个可用的参考帧,使用拉格朗日乘子方法计算率失真RD代价 $J(\cdot)$ 并比较预测结果的代价,拉格朗日乘子方法计算率失真RD代价函数 $J(\cdot)$ 表示为: $J(\varphi(j)) = D(\varphi(j)) + \lambda \cdot R(\varphi(j))$,其中, $D(\cdot)$ 表示CU编码的失真程度, λ 表示拉格朗日乘子, $R(\cdot)$ 表示CU编码消耗的比特数。因为用单向预测Uni-L0和单向预测Uni-L1表示的两个参考帧列表用于运动预测,所以应该用两个列表测试用于单向预测的运动估计ME过程,从而在两个列表中生成所有可用的帧。

[0057] 对于静态区域的CU,跳过仿射运动估计AME,直接利用运动估计CME对当前CU进行预测,并通过率失真优化的方法选择最佳的预测方向模式;具体方法为:

[0058] S21、当前CU首先经过单向预测Uni-L0,然后经过单向预测Uni-L1,最后经过双向预测Bi。

[0059] S22、利用率失真优化分别计算经过单向预测Uni-L0、单向预测Uni-L1和双向预测Bi的率失真代价;

[0060] 所述单向预测Uni-L0、单向预测Uni-L1和双向预测Bi的率失真代价分别为:

[0061]
$$\arg \min_{j \in \Gamma(\phi) | \phi=L0, L1} \{J(\phi(j))\},$$

[0062]
$$\arg \min_{j \in \Gamma(\phi) | \phi=L0, L1} \left\{ \frac{J(\phi(j) | \phi = L0)}{2} + \frac{J(\phi(j) | \phi = L1)}{2} \right\},$$

[0063] 其中, $\Gamma(\phi)$ 表示所有可用参考列表集合, ϕ 表示参考列表集, L0和L1表示两个参考帧列表, $\phi(j)$ 表示参考列表中的参考帧, $J(\cdot)$ 为率失真代价函数。

[0064] S23、将率失真代价最小的预测模式作为最佳的预测方向模式。

[0065] S3、对于非静态区域的CU, 则不满足跳过仿射运动估计AME进程的条件, 利用训练好的随机森林分类器RFC模型对当前CU进行分类, 输出最佳的预测方向模式, 来进一步降低计算复杂性。随机森林算法基于Bootstrap重采样生成K个自助样本集, 每个样本集的数据生长为一棵决策树; 在每棵树的节点处, 基于随机子空间方法RSM, 从M' 个特征向量中随机抽取m ($m < M'$) 个特征。按照一定的节点分裂算法, 从m个特征属性中选择最优属性进行分支生长; 最终将K' 棵决策树组合起来进行众数投票。随机森林分类器生成后, 对随机森林分类器模型进行测试, 森林中的每棵树都会独立判定分类结果, 最终决策取相同判定最多的分类类别, 用公式表示如下,

[0066]
$$H(t) = \arg \max_Y \sum_{i=1}^K I(h_i(t) = Y),$$

[0067] 其中, $H(t)$ 表示组合分类模型, $h_i(t)$ 是单个分类树模型, t 表示决策树的特征属性, Y 表示输出变量, $I(\cdot)$ 表示集合性示性函数 (即当集合内有出现某个分类结果时, 函数值为1, 否则为0)。

[0068] 在遍历CU时, 记录CU的特征及CU的预测方向模式, 不干扰正常编码过程。通过Bagging集成方法重采样产生多个训练集, 从原始训练样本集中随机等量抽取样本, 重复有放回抽取生成K个新的训练样本集, 最终得到K个新的训练样本集。样本提取后, 进入随机森林分类器模型的训练模块。表1显示了随机森林分类器RFC模型建立的相关训练参数设置。

[0069] 表1训练参数配置

参数名称	参数设置
参数集	Traffic, Kimono, BQSquare, RaceHorseC and FourPeople.
特征属性的数量	10
决策树的数量	25
决策树的最大深度	30
分类的数量	3

[0071] 根据表1的参数, 随机森林分类器RFC模型的训练方法为:

[0072] S31、训练分类器的关键之一是样本集的选取, 从通用测试序列中选能够涵盖丰富的纹理复杂度的用不同分辨率下的Traffic、Kimono、BQSquare、RaceHorseC和FourPeople视频序列, 在VTM上分别编码前M=50帧, 同时记录VTM中CU的形状、CU的纹理复杂度及CU的

三种预测方向模式作为数据集,数据集包括样本集 $S=20$ 和测试集 $T=30$,其中,三种预测方向模式包括单向预测Uni-L0、单向预测Uni-L1和双向预测Bi;

[0073] 在VTM中,仿射运动的块也通过三种方式进行预测:单向预测Uni-L0、单向预测Uni-L1和双向预测Bi。同时,仿射预测还包括4-参数和6-参数的仿射模型。仿射运动估计AME模块的单向预测或双向预测都需要相关的参考帧,从而增加了VTM的编码复杂度。当仅计算每个仿射运动估计AME模块所需的参考帧数时,仿射运动估计AME进程需要两倍的运动估计CME进程的参考帧数。整个运动估计ME过程如图5所示。由图5可知,步骤S31中获得数据集的方法为:

[0074] S31.1、利用运动估计CME对视频序列进行预测,预测方法同步骤S21;

[0075] S31.2、利用4参数仿射运动模型对步骤S31.1中预测后的视频序列进行仿射预测,其中,仿射预测包括单向预测Uni-L0、单向预测Uni-L1和双向预测Bi;

[0076] 如图3所示,4参数仿射运动模型的CU中样本位置 (x,y) 的运动矢量为:

$$[0077] \begin{cases} mV_x \frac{mV_{1x} - mV_{0x}}{W} x + \frac{mV_{1y} - mV_{0y}}{W} y + mV_{0x} \\ mV_y \frac{mV_{1y} - mV_{0y}}{W} x + \frac{mV_{1x} - mV_{0x}}{W} y + mV_{0y} \end{cases},$$

[0078] 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, W 表示CU的宽;

[0079] S31.3、利用6参数仿射运动模型对步骤S31.2中仿射预测后的视频序列进行放射预测;

[0080] 如图4所示,6参数仿射运动模型的块中样本位置 (x,y) 的运动矢量为:

$$[0081] \begin{cases} mV_x \frac{mV_{1x} - mV_{0x}}{W} x + \frac{mV_{2x} - mV_{0x}}{H} y + mV_{0x} \\ mV_y \frac{mV_{1y} - mV_{0y}}{W} x + \frac{mV_{2y} - mV_{0y}}{H} y + mV_{0y} \end{cases},$$

[0082] 其中, (mV_{2x}, mV_{2y}) 是左下角的运动矢量控制点, H 表示CU的高。

[0083] S31.4、分别计算步骤S31.2和S31.3进行仿射预测后的率失真代价,将最小的率失真代价对应的预测模式为视频序列的预测方向模式。

[0084] S32、利用Bootstrap法重采样样本集 S ,生成 K 个训练样本集 $\{S_1^*, S_2^*, \dots, S_K^*\}$,将生成的每个训练集 S_i^* 作为根节点,生成对应的决策树 $\{T_1, T_2, \dots, T_K\}$,其中, $i=1, 2, \dots, K$ 表示第 i 个训练样本, K 表示训练样本集的大小;

[0085] S33、从根节点开始训练,在决策树的每个中间节点上随机选择 m 个特征属性,计算每个特征属性的Gini指标系数,从中选择Gini指标系数最小的特征属性作为当前节点的最优分裂属性,以最小Gini指标系数为分裂阈值,将 m 个特征属性划分为左子树、右子树;

[0086] 机器学习的有效性与训练数据集的多样性和相关性高度相关。尽管随机森林分类器RFC可以处理超高维特征数据,但选取出真正相关的特征向量可以更好地推广分类模型。由于CU的预测方向模式和图像的纹理、纹理方向以及运动状态有关,因此将这些作为分类依据,即作为随机森林分类器模型的特征向量。本发明选取的特征属性包括二维哈尔小波

变换水平系数(2D Haar wavelet transform horizontal coefficient,HL)、二维哈尔小波变换垂直系数(2D Haar wavelet transform vertical coefficient,LH)、二维哈尔小波变换角度系数(2D Haar wavelet transform angle coefficient,HH)、角二阶矩(angular second moment,ASM)、对比度(contrast,CON)、熵(entropy,ENT)、逆差矩(inverse difference moment,IDM)、最小差值和(Sum of Absolute Difference,SAD)和梯度(gradient)作为随机森林分类器模型的特征属性,特征属性的计算如下:

[0087] 图像的二维哈尔小波变换水平系数HL表示图像水平方向的纹理,值越大说明水平方向的纹理越丰富,值越小表示水平方向的纹理越平坦;图像的二维哈尔小波变换垂直系数LH表示图像垂直方向的纹理,值越大说明垂直方向的纹理越丰富,值越小表示垂直方向的纹理越平坦;图像的二维哈尔小波变换角度系数HH表示图像垂直方向的纹理,值越大说明45°方向的纹理越丰富,值越小表示45°方向的纹理越平坦,二维哈尔小波变换水平系数HL、二维哈尔小波变换垂直系数LH和二维哈尔小波变换角度系数HH分别表示为:

$$[0088] \quad HL = \frac{4}{W \times H} \times \sum_{a=1}^{W/2} \sum_{b=1}^{H/2} \left| P(2 \times a, 2 \times b) - P(2 \times a, 2 \times b - 1) + \right. \\ \left. P(2 \times a + 1, 2 \times b) - P(2 \times a + 1, 2 \times b + 1) \right|,$$

$$[0089] \quad LH = \frac{4}{W \times H} \times \sum_{a=1}^{W/2} \sum_{b=1}^{H/2} \left| P(2 \times a, 2 \times b) + P(2 \times a, 2 \times b + 1) - \right. \\ \left. P(2 \times a + 1, 2 \times b) - P(2 \times a + 1, 2 \times b + 1) \right|,$$

$$[0090] \quad HH = \frac{4}{W \times H} \times \sum_{a=1}^{W/2} \sum_{b=1}^{H/2} \left| P(2 \times a, 2 \times b) - P(2 \times a, 2 \times b + 1) - \right. \\ \left. P(2 \times a + 1, 2 \times b) + P(2 \times a + 1, 2 \times b + 1) \right|,$$

[0091] 其中,W代表CU的宽,H代表CU的高,P(a,b)代表在位置为(a,b)的像素值。

[0092] 角二阶矩ASM反应灰度分布均匀程度和纹理粗细度,值越大说明图像纹理分布越均匀;对比度CON反应图像的纹理深度,值越大说明纹理深度越大;熵ENT表示图像的信息量,值越大说明图像的信息量越大;逆差矩IDM反应图像局部纹理变化的大小,图像的纹理的不同区域间较均匀,变化缓慢,角二阶矩ASM、对比度CON、熵ENT和逆差矩IDE分别表示为:

$$[0093] \quad ASM = \sum_{a=1}^W \sum_{b=1}^H P(a, b)^2,$$

$$[0094] \quad CON = \sum_{a=1}^W \sum_{b=1}^H (a - b)^2 P(a, b),$$

$$[0095] \quad ENT = \sum_{a=1}^W \sum_{b=1}^H P(a, b) \log_2 P(a, b),$$

$$[0096] \quad IDM = \sum_{a=1}^W \sum_{b=1}^H \frac{P(a, b)}{1 + (a - b)^2}$$

[0097] 在基于块匹配的运动估计算法中,最佳匹配块的判断准则有很多,我们使用最小差值和SAD,SAD越小,表明参考块越接近当前预测块,最小差值和SAD表示为:

$$[0098] \quad SAD(i', j') = \sum_{m=1}^W \sum_{n=1}^H |P_k(m, n) - P_{k-1}(m+i', n+j')|,$$

[0099] 其中, $P_k(a, b)$ 代表当前像素的值,(a,b)表示当前像素的坐标, $P_{k-1}(a+i', b+j')$ 是

参考像素值, $(a+i', b+j')$ 表示参考像素的坐标。

[0100] 梯度表示CU的纹理方向, 使用亮度样本的水平和垂直方向的梯度作为特征属性。水平和垂直方向的梯度表示为:

$$[0101] \quad G_x(a, b) = P(a+1, b) - P(a, b) + P(a+1, b+1) - P(a, b+1),$$

$$[0102] \quad G_y(a, b) = P(a, b) - P(a, b+1) + P(a+1, b) - P(a+1, b+1),$$

[0103] 其中 $G_x(a, b)$ 和 $G_y(a, b)$ 分别表示当前像素在水平和垂直方向上的梯度分量。 (a, b) 代表像素的坐标, $P(a, b)$ 代表像素值。

[0104] S34、重复步骤S33, 训练 $K' = 25$ 次, 直到 K' 棵决策树训练完成, 每棵决策树都完整生长而不进行剪枝;

[0105] S35、生成的多棵决策树即为随机森林分类器RFC模型, 并利用随机森林分类器RFC模型对测试集T进行判别分类, 分类结果采用投票方式, 将 K' 棵决策树输出最多的类别作为测试集T的所属类别, 得到当前CU的最佳的预测方向模式, 降低仿射运动估计AME模块的计算复杂度。

[0106] 为了评估本发明的方法, 在最新的H.266/VVC编码器(VTM 7.0)上进行了仿真测试。测试视频序列在“Random Access”配置中使用默认参数进行编码。BDBR反映了本发明的压缩性能, 时间的下降体现了复杂性的降低。表2给出了本发明的编码特性, 本发明的总编码时间平均减少到87%, 仿射运动估计AME时间平均减少到56%。因此, 本发明可以有效地节省编码时间, 并且RD性能的损失可以忽略不计。

[0107] 表2本发明的编码特性

视频序列	提出的方法 (VTM7.0)				
	BDBR-Y (%)	BDDR-U (%)	BDBR-V (%)	AME (%)	Total time (%)
Tango2	0.04	-0.17	0.09	56	88
FoodMarket4	0.07	0.08	0.03	57	90
CatRobot1	0.05	0.52	0.15	55	83
DaylightRoad2	0.11	0.04	0.13	59	86
ParkRunning3	0.08	0.08	0.06	60	84
MarketPlace	0.07	0.14	0.09	54	92
RitualDance	0.07	0.09	0.16	61	87
Cactus	0.09	-0.03	0.03	55	85
[0108] BasketballDrive	0.08	-0.01	0.11	54	89
BQTerrace	0.06	0.03	-0.22	52	85
BasketballDrill	0.04	-0.05	0.02	58	89
BQMall	0.08	-0.21	-0.19	56	90
PartyScene	0.21	0.11	0.03	57	88
RaceHorses	0.06	0.08	0.14	53	89
BasketballPass	0.08	0.06	0.13	57	85
BQSquare	0.43	-0.09	0.03	55	85
BlowingBubbles	0.12	-0.15	0.11	54	88
RaceHorses	0.08	-0.13	-0.04	55	84
平均值	0.10	0.02	0.05	56	87

[0109] 从表2可以看出本发明与VTM相比RD性能和节省的编码运行时间。对于不同的测试视频, 可能实验结果可能会有所波动, 但是本发明提出的方法是有效的。与VTM相比, 本发明可以有效地降低仿射运动估计AME模块的复杂度, 并且具有良好的RD性能。

[0110] 仿射运动估计AME模块时间是根据不同的量化参数(Quantization parameter, QP)测量的。当量化参数QP为22时, 从图6可以看出, 所有视频序列的仿射运动估计AME模块

时间总计约为36小时。但是,在本发明的方法中,仿射运动估计AME模块的时间减少了大约9个小时。可以看出,在其他量化参数QPs下,这种趋势是相似的。因此,从图6更直观地观察到,所提出的方法减少了仿射运动估计AME模块的编码时间,从而降低了计算复杂度。

[0111] 以上结合附图详细说明了本发明的技术方案,本发明的技术方案提出了一种针对H.266/VVC的快速仿射运动估计方法,有效地降低了在VTM中的仿射运动估计AME编码复杂度。首先利用标准差SD将CU分为静态区域和非静态区域,如果CU属于静态区域,选择SKIP模式进行帧间预测的概率较高,并且倾向于选择SKIP模式进行帧间预测的静态区域不需要进行仿射预测,因此,在静态区域可以提前终止仿射运动估计AME模块,并且当前CU的最佳方向模式为运动估计CME的最佳方向模式。如果CU属于非静态区域,则根据随机森林分类模型判断CU的帧间预测模式,最终提前得到最优的预测方向模式。因此,本发明降低了计算复杂度并节省了编码时间,从而实现H.266/VVC的快速编码。

[0112] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

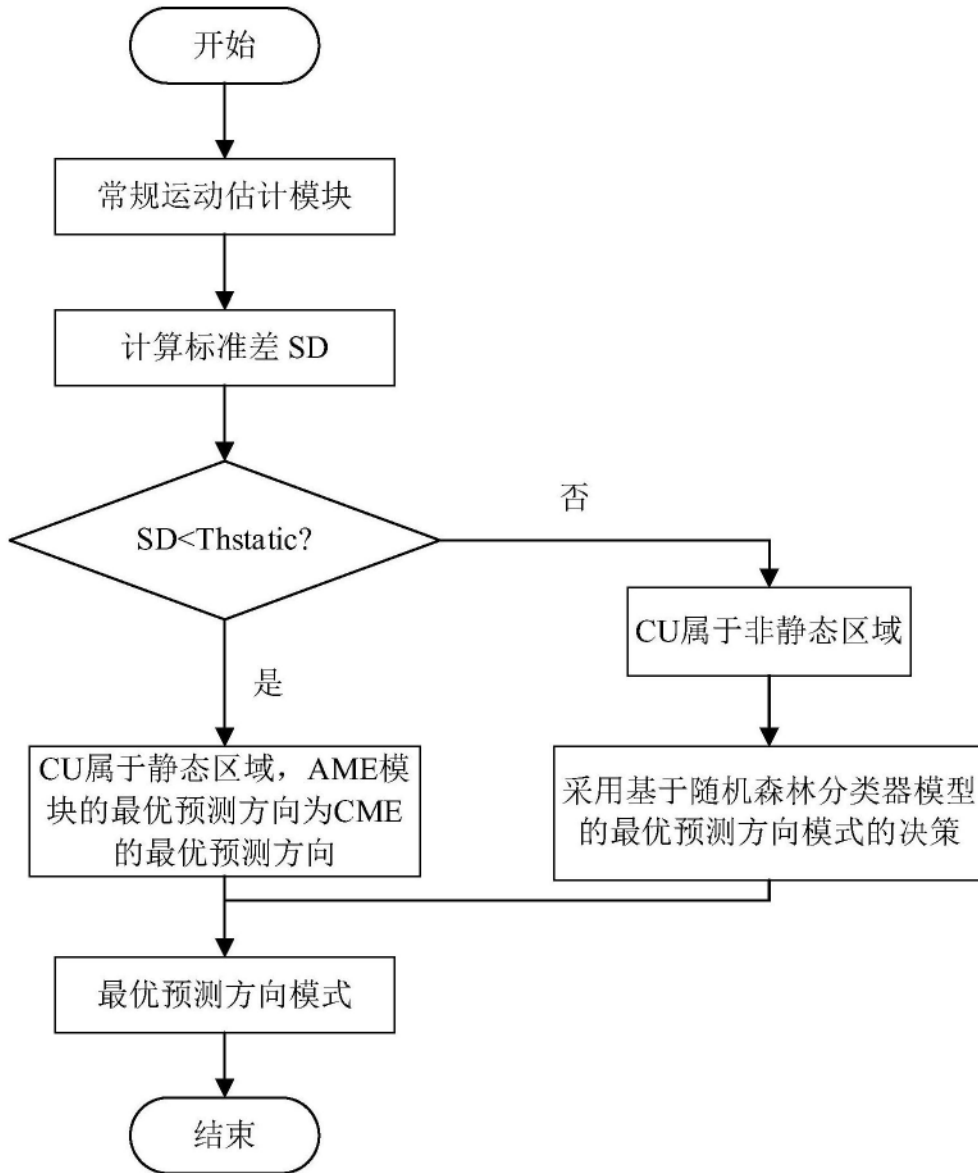


图1

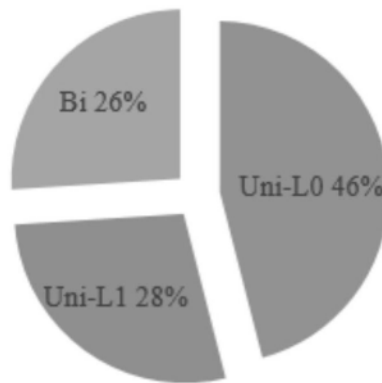


图2

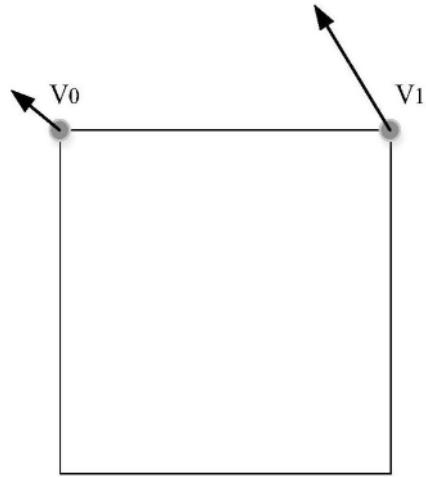


图3

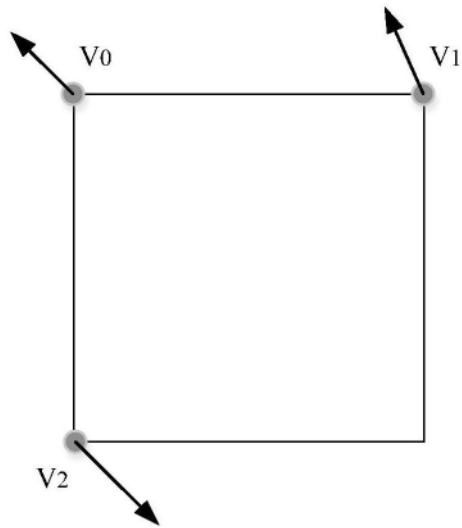


图4

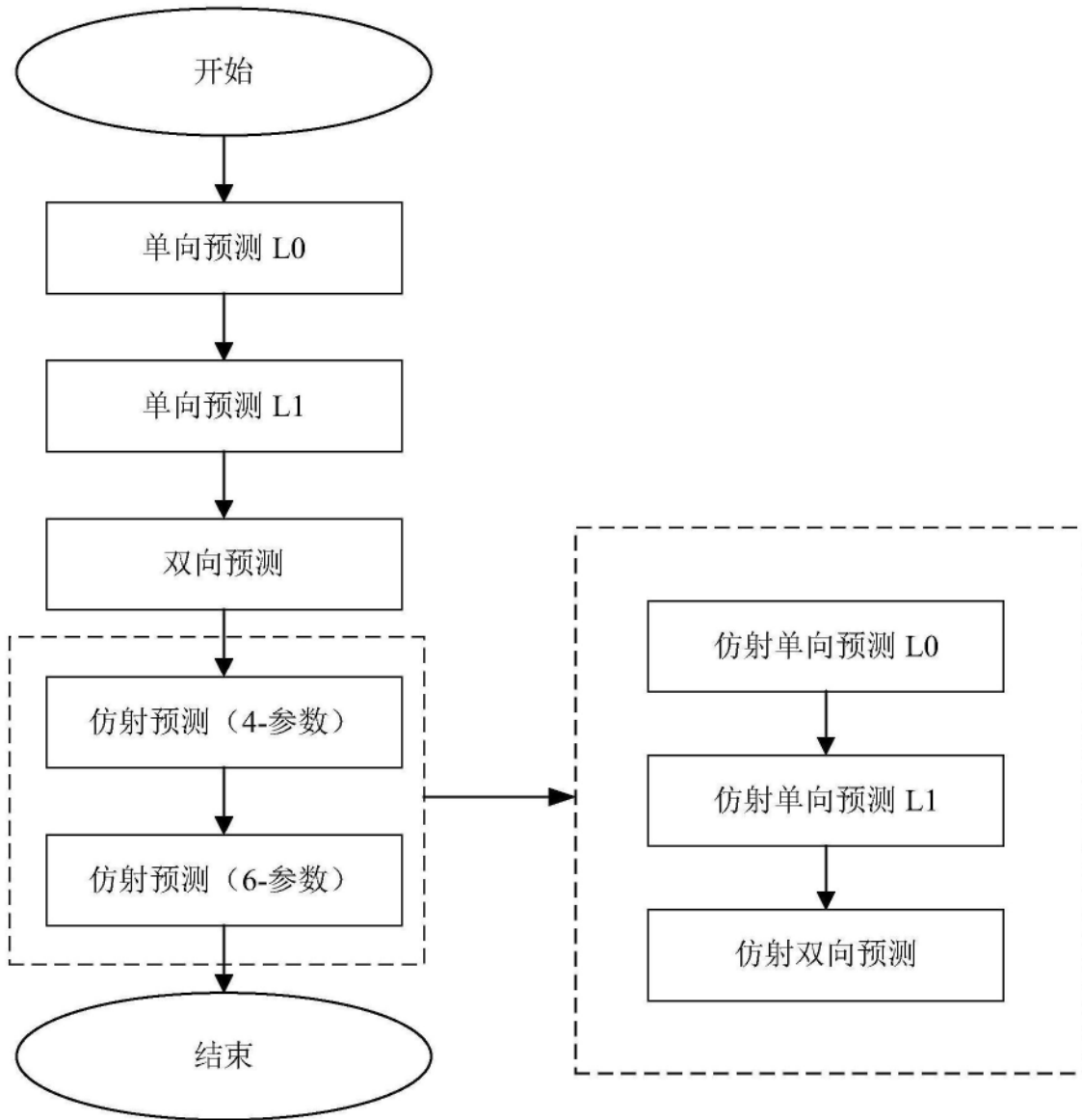


图5

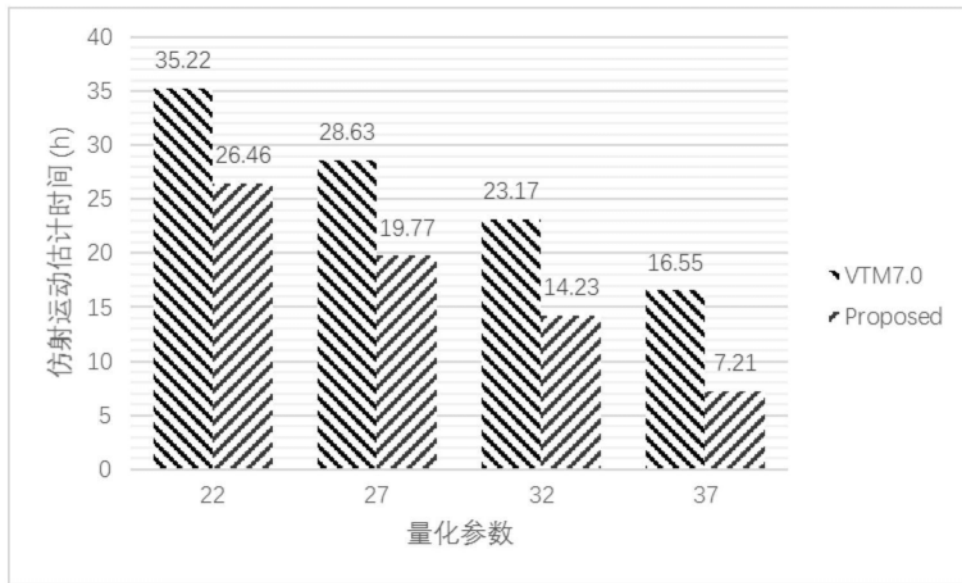


图6