

采用自适应缩放系数优化的块匹配运动估计*

宋传鸣^{1,2,3}, 闫小红¹, 葛明博¹, 王相海¹, 尹宝才²



¹(辽宁师范大学 计算机与信息技术学院, 辽宁 大连 116029)

²(大连理工大学 计算机科学与技术学院, 辽宁 大连 116024)

³(计算机软件新技术国家重点实验室(南京大学), 江苏 南京 210023)

通讯作者: 宋传鸣, E-mail: chmsong@lnnu.edu.cn

摘要: 尽管基于平移模型的快速块匹配运动估计算法在一定程度上解决了高计算量的问题,但却是以牺牲运动补偿质量为代价的,而高阶运动模型尚存在计算量高、收敛不稳定的不足.通过实验统计发现,视频中约有 56.21%的块包含缩放运动,进而得出缩放运动是除平移运动外最主要的视频运动形式的结论.进而借助双线性插值,在传统的块平移模型中引进一个缩放系数,将运动补偿误差表示为该缩放系数的一元二次函数,利用韦达定理推导出 1D 缩放运动下最佳缩放系数的计算方法,并将其进一步推广到 2D 等比例缩放运动的情况下.在此基础上,提出了一种采用自适应缩放系数优化的快速块匹配运动估计算法.该算法以菱形搜索计算平移矢量,再用自适应缩放系数确定待预测块的最佳匹配块.在 33 个标准测试视频上的实验结果表明,与基于平移模型的块匹配全搜索和快速菱形搜索相比,该算法的平均运动补偿峰值信噪比(peak signal-to-noise ratio,简称 PSNR)分别提高了 0.11dB 和 0.64dB,计算量比全搜索下降了 96.02%,略高于菱形搜索;与基于缩放模型的运动估计相比,该算法的平均峰值信噪比之 3D 全搜索下降了 0.62dB,但是比快速 3D 菱形搜索提高了 0.008dB,而计算量仅分别为两者的 0.11%和 3.86%,并且无需向解码端传输缩放矢量,能够实现编、解码端的自同步,不会增加边信息的码流开销.此外,该自适应缩放系数计算方法还可与菱形搜索以外的其他快速块匹配运动估计相结合,提高其运动补偿质量.

关键词: 视频编码;运动估计;块匹配;缩放模型;自适应缩放系数

中图法分类号: TP391

中文引用格式: 宋传鸣,闫小红,葛明博,王相海,尹宝才.采用自适应缩放系数优化的块匹配运动估计.软件学报,2020,31(11): 3603-3620. <http://www.jos.org.cn/1000-9825/5864.htm>

英文引用格式: Song CM, Yan XH, Ge MB, Wang XH, Yin BC. Optimized block-matching motion estimation using adaptive zoom coefficient. Ruan Jian Xue Bao/Journal of Software, 2020,31(11):3603-3620 (in Chinese). <http://www.jos.org.cn/1000-9825/5864.htm>

Optimized Block-matching Motion Estimation Using Adaptive Zoom Coefficient

SONG Chuan-Ming^{1,2,3}, YAN Xiao-Hong¹, GE Ming-Bo¹, WANG Xiang-Hai¹, YIN Bao-Cai²

¹(School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China)

²(School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China)

³(State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210023, China)

* 基金项目: 国家自然科学基金(61402214, 41671439, 61632006); 辽宁省自然科学基金(20180550570); 大连市青年科技之星支持计划(2015R069); 计算机软件新技术国家重点实验室(南京大学)开放课题(KFKT2018B07); 辽宁省高等学校创新团队支持计划(LT2017013); 辽宁省高等学校创新人才支持计划([2018]478-64)

Foundation item: National Natural Science Foundation of China (61402214, 41671439, 61632006); Natural Science Foundation of Liaoning Province (20180550570); Program for Youth Science and Technology Star of Dalian City (2015R069); Open Foundation of State Key Laboratory for Novel Software Technology (Nanjing University) (KFKT2018B07); Program for Liaoning Innovative Research Team in University (LT2017013); Program for Liaoning Excellent Talents in University ([2018]478-64)

收稿时间: 2018-10-27; 修改时间: 2019-02-25; 采用时间: 2019-05-12

Abstract: Fast block-wise motion estimation algorithm based on translational model solves the high computational complexity issue to some extent, but it sacrifices the motion compensation quality, whilst the higher-order motion model still exhibits the problems of computationally inefficiency and unstable convergence. Through a number of experiments, it is found that about 56.21% of the video blocks contain zoom motion, thus a conclusion is drawn that zoom motion is one of the most important motion forms in video except for the translational motion. Therefore, a zoom coefficient is introduced into the conventional block-wise translational model by bilinear interpolation, and model the motion-compensated error into a quadratic function with regard to the zoom coefficient. Subsequently, the approach is derived to compute the optimal zoom coefficient under the condition of 1D zoom motion through Vieta's theorem, which is further extended to the condition of 2D zoom motion with equal proportion. Based on the above, a fast block-matching motion estimation algorithm is presented and is optimized by the adaptive zoom coefficient. It first uses the diamond search (DS) to compute the translational motion vector, and then determines an optimal matching block for the block to be predicted with the adaptive zoom coefficient. Experimental results carried out on 33 standard test video sequences showed that the proposed algorithm gains separately 0.11 dB and 0.64 dB higher motion-compensated peak signal-to-noise ratio (PSNR) than those of the full search (FS) and the DS based on block-wise translational model. And its computational complexity is 96.02% lower than that of the FS, slightly higher than that of the DS. Compared with the motion estimation based on the zoom model, the average PSNR of the proposed algorithm is 0.62 dB lower than that of 3D full search, but 0.008 dB higher than that of fast 3D diamond search. And the computational complexity only amounts to 0.11% and 3.86% of the 3D full search and the 3D diamond search, respectively. Meanwhile, the proposed algorithm can realize the self-synchronization between the encoder and decoder without transmitting the zoom vectors, so it does not increase the overhead of the side information. Additionally, the proposed adaptive zoom coefficient computation can also be combined with state-of-art fast block-wise motion estimation algorithms other than the diamond search, improving their motion-compensation quality.

Key words: video coding; motion estimation; block matching; zoom model; adaptive zoom coefficient

运动估计是 AVS、H.265/HEVC 和 MPEG 等视频编码器所采用的一种时间域差分预测方法,它为“差分预测+变换”的闭环反馈编码架构带来了最主要的编码增益^[1,2]。然而文献[1,3,4]在对各个编码环节的计算量进行定量分析后发现,运动估计环节的计算开销占整个编码器所需计算资源的 40%以上。若视频编码器开启了可变块尺寸的 1/8 像素运动估计、自适应运动矢量预测等高级模式,运动估计甚至会耗费编码器全部计算资源的 80%。在这种情况下,为了达到更加合理的码率-失真-计算复杂度(rate-distortion-complexity,简称 R-D-C)性能,现有视频编码标准均采用了基于平移模型的块匹配算法来去除由物体平移运动所产生的时间域冗余,并出现了 7 类快速视频运动估计算法^[5]。

- (1) 基于候选向量下采样的运动估计:按照某种原则(如中心偏置原则),选择搜索窗口中的少数运动向量作为候选向量集合,进而确定补偿误差最小的候选向量作为运动向量,如 UMHexagonS^[6]、TZSearch^[7]、EPZS^[8]和抛物线搜索^[9]等。
- (2) 基于像素下采样的运动估计:采用某种采样矩阵(如层次采样矩阵^[10,11]、梅花形采样矩阵^[12]和自适应采样矩阵^[13])将待匹配的宏块进行下采样,进而在搜索窗口中,为尺寸缩小了的待预测宏块计算每个候选向量的补偿误差,得到最佳运动向量。
- (3) 基于像素预排序的运动估计:在计算待匹配宏块的运动补偿误差时,采用一定的预测策略确定宏块中可能产生较大帧差的像素,并优先统计其对应的运动补偿误差。若累积误差超过当前最优向量的预测误差,则可排除该候选向量成为最佳运动向量的可能,例如 PDS 算法^[14,15]等。
- (4) 基于低复杂度匹配函数的运动估计:采用异或、比较、取绝对值等运算替代均方差函数的减法、乘法、平方根操作,从而减少计算运动补偿误差所需要的 CPU 时钟周期数及硬件开销,如文献[16,17]等。
- (5) 基于低比特深度像素的运动估计:采用某种位深度映射函数,将具有较高位深(如 12bit 和 8bit)的像素转换为低位深的像素,进而将多个像素的低位深表示合并到 1 个机器字中。若与第(4)类的低复杂度匹配函数联合使用,则能达到 1 次操作即可求解多个像素的补偿误差的目的,如 1bit 运动估计^[18]和 2bit 运动估计^[19,20]等。
- (6) 基于散列表的运动估计:利用散列函数将待匹配块的像素值映射为一个散列值,再借助散列表查找最佳匹配块^[21-23],从而避免了匹配误差的重复计算,可将运动估计的时间复杂度由平方阶降低到线性

阶,尤其适用较大搜索窗口下的块匹配运动估计.

- (7) 基于块分类的运动估计:以运动补偿帧的编码率失真性能为度量指标,将当前帧的各个待匹配宏块划分为具有不同运动特性的宏块类,并为每个宏块类选取最优的运动估计/补偿算法,有助于在保证整体率失真性能的前提下,降低运动估计的软硬件开销,如文献[1,24]等.

一方面,尽管上述7类改进策略在一定程度上缓解了运动估计环节计算复杂度过高的不足,可是这些改进方法却是以牺牲运动估计/补偿质量作为代价的.具体来讲,第(1)类~第(3)类和第(7)类策略仅考察待预测块的某个像素子集的运动补偿误差,第(4)类和第(5)类策略则降低了像素表示和运算的精度,均会不可避免地使运动估计陷入局部最优;而第(6)类策略的性能依赖于散列函数的设计,且需要较大的缓存空间.另一方面,块平移模型既无法有效预测前景物体的非刚性运动、物体与摄像机相对运动产生的复合运动以及3D场景空间向2D成像平面的几何投影变换,又不能有效逼近具有复杂边缘轮廓的运动区域,以致在运动物体的边缘、轮廓周围产生大量幅值很高的补偿误差,不可避免地影响后续的变换编码和熵编码环节的效率^[25,26].于是,便有一些工作尝试将高阶运动模型引入运动估计^[27,28]来刻画复杂运动场.然而,用来编码高阶运动向量的码流开销、插值多个变形参考帧(warped reference frame)的计算开销和存储变形参考帧的缓存开销随之增加,并且高阶运动参数的求解算法尚存在计算量高、收敛不稳定的问题.在这种情况下,目前较为切实可行的一种帧间预测方案并非是直接采用高阶模型进行运动补偿,而是在平移模型的基础上,借鉴高阶模型刻画复杂运动的思想来优化传统块匹配运动估计.

因此,本文首先讨论现有高阶运动模型的优缺点,引出缩放运动模型的主要解决思路;然后,通过实验统计发现,在标准测试视频中约有56.21%的宏块包含缩放运动,得出了缩放运动是除平移运动以外、最主要运动形式的结论;进而借助双线性插值,在平移模型中引进缩放系数,将基于平移模型的块匹配运动补偿误差建模为该缩放系数的一元二次函数,并由韦达定理来预测最佳缩放系数;最后,提出一种采用自适应缩放系数优化的快速块匹配运动估计算法.实验结果验证了本文算法的有效性.

1 相关工作

为更加有效地预测视频场景和前景物体的非刚体运动,研究人员进一步引进了仿射变换和射影几何变换,利用1个或多个参考帧的几何形变完成运动补偿,提出了基于高阶运动模型的运动估计算法,从而克服平移运动模型的不足.依据几何变换基底的不同,现有的高阶运动估计算法大致可被分为4类^[5,20]:基于网格表示的运动估计^[29-35]、基于多项式函数的运动估计^[36-43]、基于弹性变换的运动估计^[28,44-47]和基于缩放变换的运动估计^[48-50].其中,网格模型和多项式模型较适用全局运动估计/补偿,弹性模型和缩放运动模型则更适用局部运动估计/补偿.而弹性运动模型的计算复杂度较高,且与块平移模型相结合时,将引入4倍数量的运动向量.由仿射变换的性质可知^[51],平移和缩放是平面仿射坐标系下的基本运动形式.虽然缩放模型不能描述错切运动和摇摄产生的刚性旋转,但后者在自然视频中所占比例很小,3D错切则更为鲜见.故此,考虑到缩放运动是自然视频中最广泛的运动形式,利用缩放模型来优化块匹配运动估计是比较理想的选择,能够有效预测更丰富的物体运动.

若按照运动形式来分类,关于缩放运动估计的研究可分为全局运动估计和局部运动估计两大类.其中:前者大多作为仿射模型中的一项(如文献[39,40,52]等),用于多参考帧的运动补偿,其目的是对摄像机的拉摄和推摄运动进行预测;后者则与块匹配运动估计相结合,用来处理摄像机与前景物体的局部相对运动.本文工作亦将精力集中于此,下面简要分析此类方法的相关研究.

若依据参数求解方法来分类,目前基于缩放模型的块匹配运动估计大致包括两个方面.

- 首先,以隐式方式计算缩放系数,无需向解码端传输同步信息,可直接与视频编码标准结合.文献[48]将参考帧进行不同比例的缩放,建立一个包含23个等级的缩放参考帧集合,再在该集合上利用典型的多参考帧运动估计展开块匹配运动补偿,从而将平均峰值信噪比(peak signal-to-noise ratio,简称PSNR)提高了1.69dB~2.34dB.但是该方法无疑将编码的时间、空间复杂度均提高了数十倍.尽管文献[48]进一步采用6-抽头有限响应滤波器和双线性插值将参考帧放大64倍以内,并将缩放比例减少到5级,再利

用亚像素精度的多参考帧运动估计中找出较优的缩放系数,可是仍未从根本上克服编码复杂度显著升高的不足,而且无法实现缩放比例的自适应计算.为了节省缩放参考帧的存储空间,文献[53]建立了全局/局部缩放运动的数学模型,进而推导出待预测块的缩放系数与相邻宏块的运动向量之间的线性回归关系,理论上仅需要 1 帧大小的空间来保存参考帧.不过,这一近似关系建立在运动连续性假设之上,对处于运动前景与静止背景交界的宏块,其估计精度有限.为此,文献[49]一方面将待预测帧划分成大小为 4×4 像素的块,以增强相邻块的运动连续性,进而得到更加精确的缩放运动向量;另一方面,利用待预测块左上部的像素组成模板,力图通过模板匹配来校正线性回归结果,取得了一定效果.然而,尺寸较小的块需占用更多的码流来编码运动向量,并且模板匹配也增加了计算开销.

- 其次,以显式方式计算缩放系数,并向解码端传输同步信息.文献[54]提出了一种多尺度运动估计方法,将待预测块缩放到不同尺度后,再利用原始参考帧完成运动估计.这样既可避免缩放参考帧所需的缓存空间,又能实现任意缩放比例下的运动补偿.但是由于要反复缩放待预测块,其双线性插值的计算量将显著增加.所以,文献[50]也把缩放比例减少到 5 级,提出了基于多尺度缩放模型的全搜索算法 ZFS (zoom full search),其时间复杂度在理论上是块匹配全搜索的 10 倍;若将插值操作考虑在内,计算量则将高出多倍(第 5.2 节的实验结果也验证了这一点).为此,文献[50]通过分析缩放运动向量的分布规律,给出了一种 3D 快速菱形搜索算法 3D-ZPDS(zoom predictive diamond search),将缩放比例限制在一个较小范围内.即便如此,其时空复杂度仍与块匹配全搜索相当,且刻画快速局部拉伸运动的能力有限.

综合上述两方面研究工作来看,为追求精确的缩放系数所导致的高计算量和缓存开销仍是制约缩放运动估计效率的瓶颈.若能在一定程度上放松对缩放系数的精度要求,实现缩放比例的自适应预测,从而减少参考帧和待预测块的双线性插值操作,则可在运动补偿质量和时空开销之间实现合理折中.

2 缩放运动模型及其比例分析

为了便于下文工作的论述,本节首先介绍缩放运动模型,然后通过实验对缩放运动在自然视频中所占的比例进行统计和分析.

2.1 缩放运动模型简介

视频运动估计的目标是将参考帧的特定范围作为搜索窗口,进而在其中为当前待匹配宏块 I 找到一个具有最小匹配误差平方和(sum of squared difference)的参考块 R ,即

$$D = \arg \min_{\mathbf{v}} \left\{ \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} [R(w(x_{m,n}, m; \mathbf{v}), w(y_{m,n}, n; \mathbf{v})) - I(x_{m,n}, y_{m,n})]^2 \right\} \quad (1)$$

其中,当前待匹配宏块的尺寸为 $B \times B$ 像素,它与 R 之间的坐标位移称为运动向量(矢量), $x_{m,n}$ 和 $y_{m,n}$ 分别表示当前块中 n 行 m 列像素的 x 坐标和 y 坐标, $0 \leq m, n \leq B-1$; $\mathbf{v} = \{v_x, v_y, z\}$ 表示运动向量; $w(\cdot)$ 表示以待预测块 I 的左上角像素 $(x_{0,0}, y_{0,0})$ 为原点的缩放运动函数,其定义为

$$w(x_{m,n}, m; \mathbf{v}) = z \cdot m + (x_{0,0} + v_x) \quad (2)$$

$$w(x_{m,n}, n; \mathbf{v}) = z \cdot n + (y_{0,0} + v_y) \quad (3)$$

其中, v_x 和 v_y 为待预测块的平移分量; z 为缩放系数,它建立了待预测块的原始尺寸 bs 及其缩放后的尺寸 zbs 之间的关系.

$$z = \frac{zbs - 1}{bs - 1} \quad (4)$$

同时,缩放运动模型由两个参数进行控制:缩放步长 z_v 和缩放尺度 z_s .前者表示块尺寸的变化量,以像素为单位;后者则表示缩放的尺度,也以像素为单位.两者共同决定了待预测块在缩放后的尺寸 zbs .

$$zbs = bs + \left(bs \times \frac{z_v}{z_s} \right), \quad -z_r \leq z_v \leq +z_r \quad (5)$$

一方面,当 $z_v < 0$ 时,表示缩小待预测块;当 $z_v > 0$ 时,表示放大待预测块;当 $z_v = 0$ 时,表示待预测块未发生缩放.

另一方面, z_v 的取值范围 z_r 决定了缩放运动估计搜索空间的大小和计算复杂度的高低,而缩放尺度 z_s 则决定了搜索最优缩放系数的精细程度.文献[50]经过实验得出,当 $z_s=64$ 且 $z_r=2$ 时,缩放运动模型能够在计算量和运动补偿质量之间达到一定折中,此时,缩放系数 z 的范围为 $[0.967,1.033]$.

直观起见,图1给出了上述缩放运动模型的示意图及其与传统块平移模型的比较.显然,前者在平移模型基础上引进了1个缩放步长 z_v 来表示物体拉伸所产生的局部运动和摄像机拉摄所产生的全局运动.

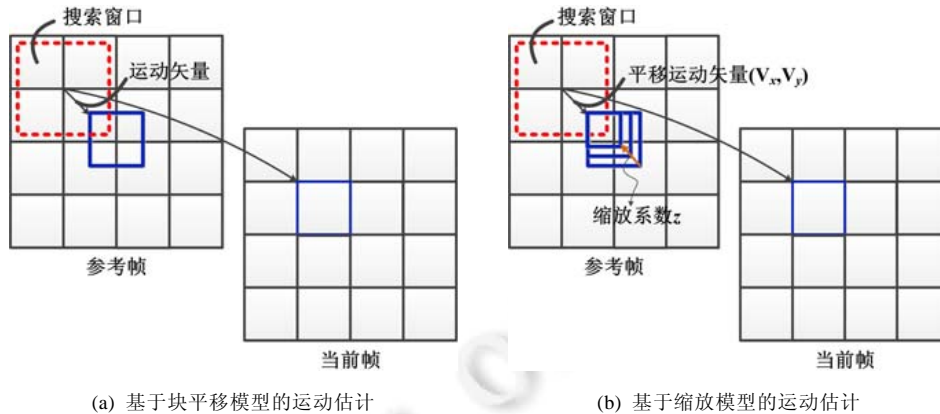


Fig.1 Comparison between zoom motion estimation and block-wise translational motion estimation

图1 缩放运动估计和块平移运动估计的比较

2.2 自然视频中缩放运动所占比例分析

由仿射变换的性质,平移、缩放和刚性旋转是平面仿射坐标系下的基本运动形式,可以通过三者的组合来表示任意的复杂运动.这意味着:缩放作为主要运动形式之一,在自然视频中应占据较高的比例.为验证这一点,本文选取了26个空间分辨率不同、运动特性迥异的标准测试视频的前90帧进行分析,采用文献[50]的3D全搜索ZFS统计各个像素块发生缩放运动的比例.实验中,令 $z_s=64, z_r=2$,搜索窗口为 64×64 像素;同时,对于每个 16×16 像素的块,若搜索到的最优缩放步长 $z_v \neq 0$,则表明其发生了局部缩放运动.统计结果见表1.

Table 1 Statistics of the zoom motion ratio in natural videos

表1 自然视频中缩放运动的比例统计

视频序列	格式	每帧的总块数	发生缩放运动的平均块数	缩放运动比例(%)
Akiyo	352×288@30fps	396	33.01	8.34
Hall	352×288@30fps	396	207.65	52.44
Tempete	352×288@30fps	396	260.11	65.68
Tennis	352×288@30fps	396	270.19	68.23
Coastguard	352×288@30fps	396	248.62	62.78
Mother & Daughter	352×288@30fps	396	178.25	45.01
Foreman	352×288@30fps	396	247.72	62.56
News	352×288@30fps	396	66.65	16.83
Paris	352×288@30fps	396	97.12	24.53
Silent	352×288@30fps	396	97.76	24.69
Students	352×288@30fps	396	64.11	16.19
Soccer	352×288@30fps	396	261.63	66.07
Vtc1nw	352×288@30fps	396	44.03	11.12
Waterfall	352×288@30fps	396	229.16	57.87
Husky	352×288@30fps	396	311.00	78.54
Mobile	352×288@30fps	396	295.82	74.70
Crew	704×576@30fps	1 584	1 330.27	83.98
City	704×576@30fps	1 584	1 317.43	83.17
Harbour	704×576@30fps	1 584	1 324.82	83.64
Flower vase	832×480@30 fps	1 560	1 021.34	65.47
ParkScene	960×540@24 fps	2 025	1 630.48	80.52

Table 1 Statistics of the zoom motion ratio in natural videos (Continued)**表 1** 自然视频中缩放运动的比例统计(续表)

视频序列	格式	每帧的总块数	发生缩放运动的平均块数	缩放运动比例(%)
BQTerrace	960×540@60 fps	2 025	1 320.24	65.20
Cactus	960×540@50 fps	2 025	1 354.58	66.89
Johnny	1280×720@60 fps	3 600	2 587.84	71.88
FourPeople	1280×720@60 fps	3 600	2 010.65	55.85
KristenAndSara	1280×720@60 fps	3 600	2 494.46	69.29
平均	-	1 135.50	742.50	56.21

从表 1 可见:

- 自然视频中,平均约有 56.21% 的像素块能用缩放模型更好地捕获其运动。
- 高清视频发生缩放运动的平均比例(72.59%)高于标清视频(45.97%),如 Harbour、ParkScene、Crew 等序列的缩放运动比例甚至已超过 80%。

综合上述统计结果,缩放运动是视频中除平移运动以外的主要运动形式,故此在平移模型中引入缩放系数是非常必要的,有利于更加有效地预测丰富的物体运动,尤其适用高清视频。

3 自适应缩放系数的计算方法

为求解第 2.1 节的缩放运动模型,文献[50]首先利用 3D 全搜索 ZFS 计算最优缩放步长 z_v ,再将其代入公式(5)得出块尺寸 zbs ,进而根据公式(4)计算最佳缩放系数 z 。该过程的时间复杂度达到了 $O(2B^2W^2zr)$ (W 表示平移运动向量的搜索窗口尺寸),为块匹配全搜索的 $2zr$ 倍。并且由于比块匹配全搜索多出大量双线性插值操作,ZFS 还需 $O(8B^2W^2zr)$ 次乘法和 $O(4B^2W^2zr)$ 次加法来计算亚像素精度的像素值,其实际计算代价将明显高于块匹配全搜索。因此,减少双线性插值次数,是在平移模型中引进缩放系数时所必须要解决的关键问题之一。

虽然亚像素精度的像素值会随着图像内容的不同而变化,但是它在本质上是由其相邻整像素精度的采样点根据像素值的某种分布假设进行插值而来^[55,56]。因此,若把整像素精度的像素值视为一个变量,则亚像素精度的像素值就是其函数。一旦整数位置的像素值确定,那么在给定插值算法的情况下,亚像素值亦是确定的。据此思路,缩放系数应可由整像素精度的像素值推导得出。

在本节中,我们以广泛采用的双线性插值为例,从讨论 1D 缩放情况下的自适应缩放系数入手,推导 2D 等比例缩放情况下自适应缩放系数的计算方法。

3.1 1D 缩放情况下自适应缩放系数的计算

本节考虑当前待预测块仅存在沿水平方向的 1D 缩放运动的情形。设待预测块的像素为 $c_{m,n}(m,n \in [0,B-1])$,且经过整像素精度的块平移全搜索后,所得到的参考帧中最佳匹配块的像素为 $r_{m,n}$;进一步地,假设当缩放系数为 z 时,与 $c_{m,n}$ 相匹配的亚像素精度的像素为 $r_{m,n}^*$,显然,它位于整像素精度的像素 $r_{m,n}$ 和 $r_{m+1,n}$ 之间,其中, $r_{m+1,n}$ 表示最佳匹配像素 $r_{m,n}$ 右侧的相邻整数位置的像素(如图 2 所示,蓝色实线框表示待预测块及其整像素精度的最佳匹配块,红色虚线框表示经过水平缩放的最佳匹配块),则由线性插值可知,

$$\begin{cases} r_{m,n}^* = pr_{m,n} + qr_{m+1,n} \\ p + q = 1 \end{cases} \quad (6)$$

其中, q 为像素 $r_{m,n}$ 和 $r_{m,n}^*$ 之间的距离, p 为像素 $r_{m+1,n}$ 和 $r_{m,n}^*$ 之间的距离,则有

$$\begin{cases} q = m \cdot z - \lfloor m \cdot z \rfloor \\ p = 1 - q \end{cases} \quad (7)$$

将公式(7)代入公式(6),就有:

$$r_{m,n}^* = [1 - (m \cdot z - \lfloor m \cdot z \rfloor)]r_{m,n} + (m \cdot z - \lfloor m \cdot z \rfloor)r_{m+1,n} \quad (8)$$

再将 $r_{m,n}^*$ 代入公式(1),则可得到当缩放系数为 z 时,当前待预测块 I 的预测误差为

$$\begin{aligned}
 D &= \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n}^*)^2 \\
 &= \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (m \cdot z - \lfloor m \cdot z \rfloor)^2 (r_{m,n} - r_{m+1,n})^2 + \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n})^2 + \\
 &\quad \left. \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (m \cdot z - \lfloor m \cdot z \rfloor) [(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2] \right\} \quad (9)
 \end{aligned}$$

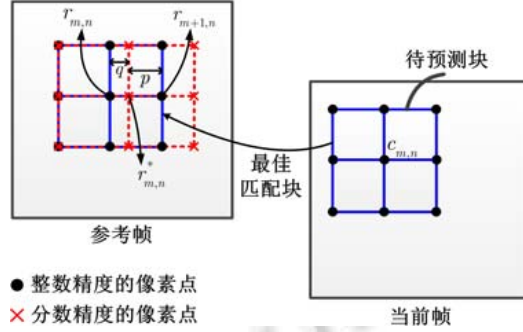


Fig.2 Diagram of 1D zoom motion along horizontal direction
图 2 沿水平方向的 1D 缩放运动示意图

首先,当 $z < 1$ 且 $(1-z) \cdot (B-1) < 1$ 时,公式(7)有:

$$\begin{cases} q = m \cdot z - \lfloor m \cdot z \rfloor = 1 - (1-z)m \\ p = 1 - q = (1-z)m \end{cases} \quad (10)$$

将公式(10)代入公式(9)并整理后,可得:

$$\begin{aligned}
 D &= z^2 \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n})^2 + \\
 &\quad z \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ (2m - 2m^2)(r_{m,n} - r_{m+1,n})^2 + m[(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2] \} + \\
 &\quad \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ (1 - 2m + m^2)(r_{m,n} - r_{m+1,n})^2 + (1 - m)[(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2] \} + \\
 &\quad \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n})^2 \quad (11)
 \end{aligned}$$

从公式(11)可见,待预测块 I 的预测误差 D 是关于缩放系数 z 的一元二次函数.由韦达定理(Vieta's theorem)可知,当

$$z = \frac{\sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ (2m - 2m^2)(r_{m,n} - r_{m+1,n})^2 + m[(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2] \}}{-2 \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n})^2} \quad (12)$$

时,预测误差 D 取得最小值.其中,分母 $\sum_n \sum_m m^2 (r_{m,n} - r_{m+1,n})^2$ 、分子第 1 项 $\sum_n \sum_m (2m - 2m^2)(r_{m,n} - r_{m+1,n})^2$ 和第 4 项 $\sum_n \sum_m m (r_{m,n} - r_{m+1,n})^2$ 可视为最佳匹配块的加权水平梯度的模长平方,而分子的第 2 项和第 3 项 $\sum_n \sum_m m [(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2]$ 则相当于待预测块的时域梯度的模长平方.

其次,当 $z > 1$ 且 $(z-1) \cdot (B-1) < 1$ 时,公式(7)化为

$$\begin{cases} q = m \cdot z - \lfloor m \cdot z \rfloor = (z-1)m \\ p = 1 - q = 1 - (z-1)m \end{cases} \quad (13)$$

将公式(13)代入公式(9)并整理,则有:

$$D = z^2 \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n})^2 + z \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{(-2m^2)(r_{m,n} - r_{m+1,n})^2 + m[(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2]\} + \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{m^2 (r_{m,n} - r_{m+1,n})^2 - m[(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2]\} + \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n})^2 \quad (14)$$

同理,此时的预测误差 D 也是关于缩放系数 z 的一元二次函数.故此,当

$$z = \frac{\sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{(-2m^2)(r_{m,n} - r_{m+1,n})^2 + m[(c_{m,n} - r_{m+1,n})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n})^2]\}}{-2 \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n})^2} \quad (15)$$

时,预测误差将取得最小值.

由公式(12)和公式(15)可见,在编/解码端已知整像素精度的块平移运动向量 (v_x, v_y) 的情况下,1D 自适应缩放系数 z 可由最佳匹配块的水平梯度和待预测块的 1 阶时域差分计算得出,既无需在编码端通过反复插值对 z 展开全搜索,又没有必要将缩放系数 z 作为边信息传输到解码端.

3.2 2D 缩放情况下自适应缩放系数的计算

第 3.1 节的结论表明,借助像素值的线性分布假设将预测误差建模为缩放系数的一元二次函数,进而求得自适应缩放系数,可有效降低引进缩放系数所需的计算复杂度和运动向量码率.在本节中,我们将该结论进一步扩展到 2D 缩放运动的情形.

设 2D 缩放运动系数为 $z=(z_x, z_y)$,则亚像素精度的像素 $r_{m,n}^*$ 可由其相邻的 4 个整像素精度的像素 $r_{m,n}, r_{m+1,n}, r_{m,n+1}$ 和 $r_{m+1,n+1}$ 通过双线性插值得到(如图 3 所示).

$$r_{m,n}^* = pp'r_{m,n} + p'qr_{m+1,n} + pq'r_{m,n+1} + qq'r_{m+1,n+1} \quad (16)$$

其中, z_x 和 z_y 分别表示水平方向和垂直方向的缩放系数; q 和 q' 分别表示 $r_{m,n}^*$ 与 $r_{m,n}$ 之间的水平距离和垂直距离,而 p 和 p' 则分别表示 $r_{m,n}^*$ 与 $r_{m+1,n+1}$ 之间的水平距离和垂直距离,且 $q=n \cdot z_x - \lfloor n \cdot z_x \rfloor, q'=n \cdot z_y - \lfloor n \cdot z_y \rfloor, p=1-q, p'=1-q'$.相应地,当前待预测块 I 的预测误差为

$$D = \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n}^*)^2 = \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - pp'r_{m,n} - p'qr_{m+1,n} - pq'r_{m,n+1} - qq'r_{m+1,n+1})^2 \quad (17)$$

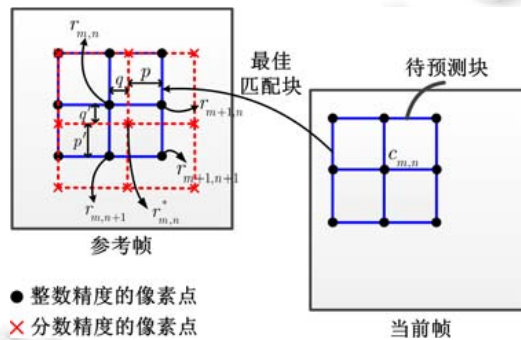


Fig.3 Diagram of 2D zoom motion along horizontal direction

图 3 沿水平方向的 2D 缩放运动示意图

显然,公式(17)是关于缩放运动系数 z_x 和 z_y 的二元 4 次多项式,其解的显式表达式难以直接给出.尽管我们可采用二分法、牛顿法等方法计算出数值解,可是由于在迭代过程中需反复计算待预测块与插值得到的参考块之间的匹配误差,势必会导致求解 z 的计算量骤增.因此,有必要进一步设计降低其时间复杂度的快速策略.

考虑到为避免物体产生形状失真,视频中的 2D 缩放运动大多应为等比例缩放(除个别特效以外),即 $z_x=z_y$.于是,有 $q=q', p=p'$.据此易推知,亚像素精度的像素 $r_{m,n}^*$ 应位于以整像素精度的像素 $r_{m,n}, r_{m+1,n+1}$ 为端点的线段上,并且与两者的欧式距离分别为 $\sqrt{2}q$ 和 $\sqrt{2}p$.通过将距离进行归一化,并由线性插值可得:

$$r_{m,n}^* = pr_{m,n} + qr_{m+1,n+1} \tag{18}$$

其中, $q=m \cdot z - \lfloor m \cdot z \rfloor, n \cdot z - \lfloor n \cdot z \rfloor, z=z_x=z_y, p=1-q$.此时,当前待预测块 I 的预测误差为

$$D = \left. \begin{aligned} & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (m \cdot z - \lfloor m \cdot z \rfloor)^2 (r_{m,n} - r_{m+1,n+1})^2 + \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n})^2 + \\ & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (m \cdot z - \lfloor m \cdot z \rfloor) [(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2] \end{aligned} \right\} \tag{19}$$

首先,当 $z < 1$ 且 $(1-z) \cdot (B-1) < 1$ 时,有:

$$\begin{cases} q = m \cdot z - \lfloor m \cdot z \rfloor = 1 - (1-z)m \\ p = (1-z)m \end{cases} \tag{20}$$

将其代入公式(19)并整理后,就得到:

$$D = z^2 \left. \begin{aligned} & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n+1})^2 + \\ & z \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ (2m - 2m^2)(r_{m,n} - r_{m+1,n+1})^2 + m[(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2] \} + \\ & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} [(1 - 2m + m^2)(r_{m,n} - r_{m+1,n+1})^2 + (c_{m,n} - r_{m,n})^2] + \\ & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (1-m)[(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2] \end{aligned} \right\} \tag{21}$$

可见,在 2D 等比例缩放运动情况下,预测误差 D 也是关于缩放系数 z 的一元二次函数.故此,当

$$z = \frac{\sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ (2m - 2m^2)(r_{m,n} - r_{m+1,n+1})^2 + m[(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2] \}}{-2 \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n+1})^2} \tag{22}$$

时,预测误差 D 将取得最小值.其中,分母以及分子的第 1 项和第 4 项是最佳匹配块沿着对角线方向的加权梯度的模长平方,分子的第 2 项和第 3 项则相当于待预测块的时域梯度的模长平方.

其次,当 $z > 1$ 且 $(z-1) \cdot (B-1) < 1$ 时,有:

$$\begin{cases} q = m \cdot z - \lfloor m \cdot z \rfloor = (z-1)m \\ p = 1 - (z-1)m \end{cases} \tag{23}$$

将公式(23)代入公式(19)并化简,则可得到待预测块 I 的预测误差为

$$D = z^2 \left. \begin{aligned} & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n+1})^2 + \\ & z \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ (-2m^2)(r_{m,n} - r_{m+1,n+1})^2 + m[(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2] \} + \\ & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{ m^2 (r_{m,n} - r_{m+1,n+1})^2 - m[(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2] \} + \\ & \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} (c_{m,n} - r_{m,n})^2 \end{aligned} \right\} \tag{24}$$

显然,由韦达定理可知,当

$$z = \frac{\sum_{n=0}^{B-1} \sum_{m=0}^{B-1} \{(-2m^2)(r_{m,n} - r_{m+1,n+1})^2 + m[(c_{m,n} - r_{m+1,n+1})^2 - (c_{m,n} - r_{m,n})^2 - (r_{m,n} - r_{m+1,n+1})^2]\}}{-2 \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n+1})^2} \quad (25)$$

时,待预测块的预测误差将取得最小值.

由公式(22)和公式(25)可见,在已知整像素精度的块平移运动向量 (v_x, v_y) 的情况下,2D 等比例缩放运动系数 z 亦可由最佳匹配块沿对角线方向的梯度和待预测块的时域差分进行预测得到,同样没必要在编码端通过反复插值和对 z 展开全搜索的方式来计算 z .

4 采用自适应缩放系数优化的块匹配运动估计步骤

第 3 节的结论表明,当平移运动向量确定后,在给定插值算法的情况下,就可由整像素精度的最佳匹配块预测得到缩放运动系数,从而避免搜索精确缩放系数所导致的高计算量和缓存开销.据此,本节提出一种采用自适应缩放系数优化的快速块匹配运动估计算法.对于每个图像组(group of picture,简称 GOP),其详细步骤如下.

Step 1. 若当前 GOP 的所有帧都已处理完毕,则算法结束;否则,选取当前 GOP 中未处理过的一帧作为当前帧,将其划分成一系列尺寸为 $B \times B$ 像素的、不重叠的块,并为当前帧确定 1 个参考帧 R .

Step 2. 若当前帧的所有块都已处理完毕,则转入 Step 1;否则,选取 1 个未处理的块作为待预测块:

$$I(x_{m,n}, y_{m,n}).$$

Step 3. 利用任何一种快速块匹配运动估计(本文选用了菱形搜索算法),在 $W \times W$ 像素的窗口中搜索待预测块 I 的平移运动分量 v_x 和 v_y ,进而得到其整像素精度的最佳匹配块 $R(x_{m,n} + v_x, y_{m,n} + v_y)$ 和预测误差 D_0 .

Step 4. 计算最佳匹配块 R 沿着对角线方向的加权梯度:

$$A = \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m^2 (r_{m,n} - r_{m+1,n+1})^2, B = \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m (r_{m,n} - r_{m+1,n+1})^2.$$

Step 5. 计算待预测块 I 与其最佳匹配块 R 的加权时域差分:

$$E = \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m (c_{m,n} - r_{m+1,n+1})^2, F = \sum_{n=0}^{B-1} \sum_{m=0}^{B-1} m (c_{m,n} - r_{m,n})^2.$$

Step 6. 根据公式(22),利用 A, B, E 和 F 预测小于 1 的缩放运动系数 $z_1: z_1 = (-2A + B + E - F) / -2A$,并将 z_1 和平移运动分量 v_x, v_y 代入公式(1)~公式(3),计算当前待预测块 I 的预测误差 D_1 ,其中,亚像素精度的像素值由双线性插值得到.

Step 7. 根据公式(25),利用 A, B, E 和 F 预测大于 1 的缩放运动系数 $z_2: z_2 = (-2A - B + E - F) / -2A$,并将 z_2 和平移运动分量 v_x, v_y 代入公式(1)~公式(3),计算当前待预测块 I 的预测误差 D_2 .

Step 8. 比较预测误差 D_0, D_1 和 D_2 ,并令 $D = \min\{D_0, D_1, D_2\}$.

Step 8.1. 若 $D = D_0$,则令待预测块 I 的自适应缩放运动系数 $z = 0$,即不发生缩放运动.

Step 8.2. 若 $D = D_1$,则令待预测块 I 的自适应缩放运动系数 $z = z_1$,即发生局部缩小运动.

Step 8.3. 否则,令待预测块 I 的自适应缩放运动系数 $z = z_2$,即发生局部放大运动.

Step 9. 输出待预测块 I 的平移运动矢量 v_x, v_y ,及其运动补偿误差 D ,转入 Step 2.

5 实验结果与分析

为测试本文算法的运动估计/补偿效率,以 MPEG 公布的 33 个国际通用测试视频流(见表 2)进行仿真实验,并将所得的客观评价结果与基于块平移模型的全搜索和菱形搜索以及基于缩放模型的全搜索 ZFS^[50]和快速搜索 3D-ZPDS^[50]的运动补偿结果进行比较.

算法参数设置如下:搜索窗口为 33×33 像素,块尺寸为 $B = 16$ 像素,运动补偿的质量采用 PSNR 进行客观评价.另外,ZFS 算法和 3D-ZPDS 算法的参数设置与文献[50]相同.

5.1 运动估计/补偿质量的比较

表2给出了5种对比算法和本文算法对每个视频序列前90帧的亮度分量进行运动估计/补偿的平均PSNR.

Table 2 Motion-compensated PSNR comparison

表2 运动补偿的PSNR比较

视频序列	块匹配全搜索	块匹配菱形搜索	块匹配全搜索+自适应缩放系数预测	基于缩放模型的全搜索 ^[50]	3D-ZPDS ^[50]	本文算法
Akiyo	42.67	42.65	43.31	43.23	43.11	43.29
Coastguard	29.71	29.19	30.22	30.45	30.32	29.69
Container	38.35	38.35	38.84	37.78	37.77	38.84
Foreman	33.52	32.47	34.04	33.94	30.59	33.10
Hall	34.48	34.37	34.93	35.02	34.92	34.85
Highway	36.08	34.63	36.62	36.83	33.56	35.15
Husky	19.70	19.17	20.26	20.64	20.53	19.78
Ice	30.99	30.11	31.54	31.70	31.51	30.89
Intros_cif	37.30	36.36	37.64	37.29	34.33	37.05
Mobile	23.96	23.58	24.93	25.59	25.29	24.59
Mother & Daughter	40.26	39.95	40.89	41.07	40.89	40.67
News	37.83	37.71	38.41	38.36	38.21	38.29
Pamphlet	43.12	41.87	43.35	43.21	43.15	42.36
Paris	30.70	30.50	31.33	31.35	31.21	31.15
Sign_irene	33.56	32.95	33.99	34.28	34.04	33.38
Silent	35.95	35.58	36.69	36.17	31.81	35.98
Soccer	28.60	27.79	30.17	30.37	30.20	28.63
Students	39.49	39.40	40.04	40.18	40.07	39.96
Tempete	27.75	27.37	28.40	28.82	28.63	28.08
Tennis	28.91	27.16	29.52	29.65	29.51	28.05
Vtc1nw	44.70	44.70	44.94	44.51	42.40	44.94
Waterfall	35.53	35.52	36.39	36.48	36.37	36.38
Crew	32.36	31.64	32.97	33.21	33.06	32.41
City	30.86	29.46	31.65	32.03	31.77	30.29
Harbour	28.10	27.79	28.84	29.04	28.86	28.58
BQmall	29.63	28.22	30.06	30.26	30.12	28.84
Flower vase	37.42	37.34	38.06	38.41	38.23	38.00
FourPeople	38.43	38.17	38.93	39.12	38.96	38.76
Johnny	38.83	38.61	39.42	39.76	39.51	39.28
KristenAndSara	39.38	39.16	40.03	40.35	40.11	39.89
BQTerrace	25.51	25.31	26.20	26.65	26.52	26.04
Cactus	29.53	28.95	30.02	30.34	30.23	29.63
ParkScene	29.98	29.57	30.51	30.93	30.81	30.06
平均	33.73	33.20	34.34	34.46	33.84	33.84

由表2可知:

- 尽管块匹配全搜索是平移模型下补偿质量最高的运动估计算法,可是由于其仅能预测物体在2D空间的刚体平移运动,其平均PSNR比基于缩放模型的全搜索ZFS和快速搜索3D-ZPDS分别降低了0.73dB和0.11dB.这说明在运动估计中引入缩放分量能够有效地提高运动补偿的质量,与表1的分析结果一致.
- 作为一种典型的快速、次最优运动估计算法,菱形搜索的平均PSNR比块匹配全搜索降低0.53dB.这源于菱形搜索只计算搜索窗口中一部分候选向量的匹配误差,难免陷入局部最优.然而,当在菱形搜索的基础上引进了自适应缩放系数后,本文算法的平均PSNR比菱形搜索提高了0.64dB,甚至优于块匹配全搜索(比后者高出0.11dB).可见,在快速块匹配运动估计中引入缩放系数来克服平移模型的不足,是一种合理且行之有效的优化方案.同时,本文算法与3D-ZPDS^[50]均是考虑了缩放因素的快速运动估计,不同之处在于,本文算法的缩放运动系数是直接预测得到的,而后者的缩放系数却是通过3D菱形搜索获得,且本文算法的运动补偿性能比后者平均高出0.008dB.这表明本文的自适应预测方法能够充分利

用像素间的线性相关关系,计算出更准确的缩放运动系数,既节省了一个维度的搜索、降低了计算量(第 5.2 节将证明这一点),又无需向解码端传输缩放运动分量,与现有编码标准兼容,同步信息的开销低于 3D-ZPDS.

- 本文算法的运动估计精度尚不及基于缩放模型的全搜索 ZFS,平均 PSNR 较其降低了 0.62dB.其根本原因是,本文算法选用菱形搜索作为初始运动估计,找到的起始搜索点不够准确.为验证这一点,以块匹配全搜索作为初始搜索,再采用本文的预测方法计算自适应缩放运动系数进行实验(见表 2 的第 4 列,简称为“块匹配全搜索+自适应缩放系数预测”),此时,平均 PSNR 比块匹配全搜索提高了 0.61dB,较 3D-ZPDS 高出 0.50dB,与 ZFS 只相差 0.12dB.并且,与以菱形搜索为初始估计的本文算法相比,块匹配全搜索+自适应缩放系数预测方法的 PSNR 亦提高了 0.50dB,尤其对快速运动的视频序列更为有效(如 Tennis 和 Soccer).这表明初始搜索的精度对本文的自适应缩放系数预测具有一定影响.
- 对于 Akiyo、Container、Foreman 和 Intros_cif 等序列,块匹配全搜索+自适应缩放系数预测方法的 PSNR 甚至高于缩放模型的全搜索 ZFS.这些视频序列的共同特点是:分辨率在标清以下,空间纹理简单或中等,其像素值分布能用双线性模型较精确地拟合,故此,本文方法计算出的自适应缩放系数获得了明显的 PSNR 增益.反之,对于 Coastguard、Highway、Flower vase 和 FourPeople 等序列,块匹配全搜索+自适应缩放系数预测方法的补偿质量仍与 ZFS 存在一定差距.究其原因在于,这些视频要么空间纹理复杂,要么分辨率较高,双线性插值无法准确逼近亚像素精度的像素值,使得本文算法得到的缩放系数仅为近似最优值.
- 对比表 1 和表 2 还可发现,较之基于缩放模型的 ZFS 和 3D-ZPDS,本文算法在缩放运动比例低于 65% 的视频序列上的 PSNR 增益较高,如 Silent、Vtc1nw 和 Foreman 序列的 PSNR 甚至比 3D-ZPDS 高出 2.5dB 以上;反之,对于缩放运动比例超过 65% 的序列,本文算法的运动估计精度则与 ZFS 存在差距.一方面,表 1 的缩放运动比例是利用 ZFS 统计获得的,该比例越高,表明相应的视频就越适合采用 ZFS 进行运动补偿;另一方面,即便表 1 给出的缩放运动比例偏低,也并不能断定相应视频较少地发生缩放运动.原因在于,文献[50]为了权衡计算量和运动补偿质量,将 ZFS 的缩放系数限定在 $\{0.967, 0.983, 1, 1.017, 1.033\}$ 之中,仅有 5 种可能的取值(分别对应 $z \in \{-2, -1, 0, 1, 2\}$),对缩放运动的刻画尚显粗糙,无法获得最优的缩放系数.而本文则将缩放系数的值域扩展为 $[0.933, 1.067]$ 区间中的任意实数,只要亚像素插值算法能够较好地逼近真实分布,本文算法就可更加准确、细致地捕获待预测块的缩放运动.

综合上述分析,通过在块匹配运动估计中引入自适应缩放系数,本文方法与块匹配全搜索和菱形搜索相比,取得了明显的性能提升.而且借助更高效的初始点搜索和插值模型,还能进一步改善自适应缩放系数的预测精度及运动补偿质量.当然,本文的目的是在低复杂度的前提下,以缩放系数来优化块匹配运动估计,而不是精确求解缩放运动模型,从这个意义上讲,本文算法在计算量、运动估计精度和边信息量之间实现了较理想的折中.

为了验证本文算法的稳定性,图 4(a)~图 4(d)进一步给出了 Akiyo、Vtc1nw、Waterfall 和 Crew 序列的 PSNR 逐帧比较情况.这 4 个序列中缩放运动的比例从 8.34% 依次增高至 83.98%(见表 1).

从图 4 可见:

- 对于纹理简单或中等、运动幅度较小,且缩放运动比例偏低的序列(如 Akiyo 和 Vtc1nw),本文的两种算法均超过块匹配全搜索和基于缩放模型的全搜索 ZFS,在所有帧上取得了最高的运动补偿质量.
- 对于纹理细节复杂、运动幅度中等且含有慢速拉摄运动的序列(如 Waterfall),本文算法在大多数帧上的 PSNR 仍高于缩放模型的 ZFS 和 3D-ZPDS,比块匹配全搜索和菱形搜索亦有稳定提高.
- 对于含有物体与摄像机快速相对运动的序列(如 Crew),ZFS 和 3D-ZPDS 获得了最高的补偿质量,块匹配全搜索+自适应缩放系数预测方法的性能与之相近,而本文算法的预测质量则高于块匹配全搜索和菱形搜索.

通过上述 4 个空间纹理特性各异、运动复杂度亦不同的视频可见,本文算法的运动补偿 PSNR 均非常稳定,与其他算法的对比关系未出现明显波动.

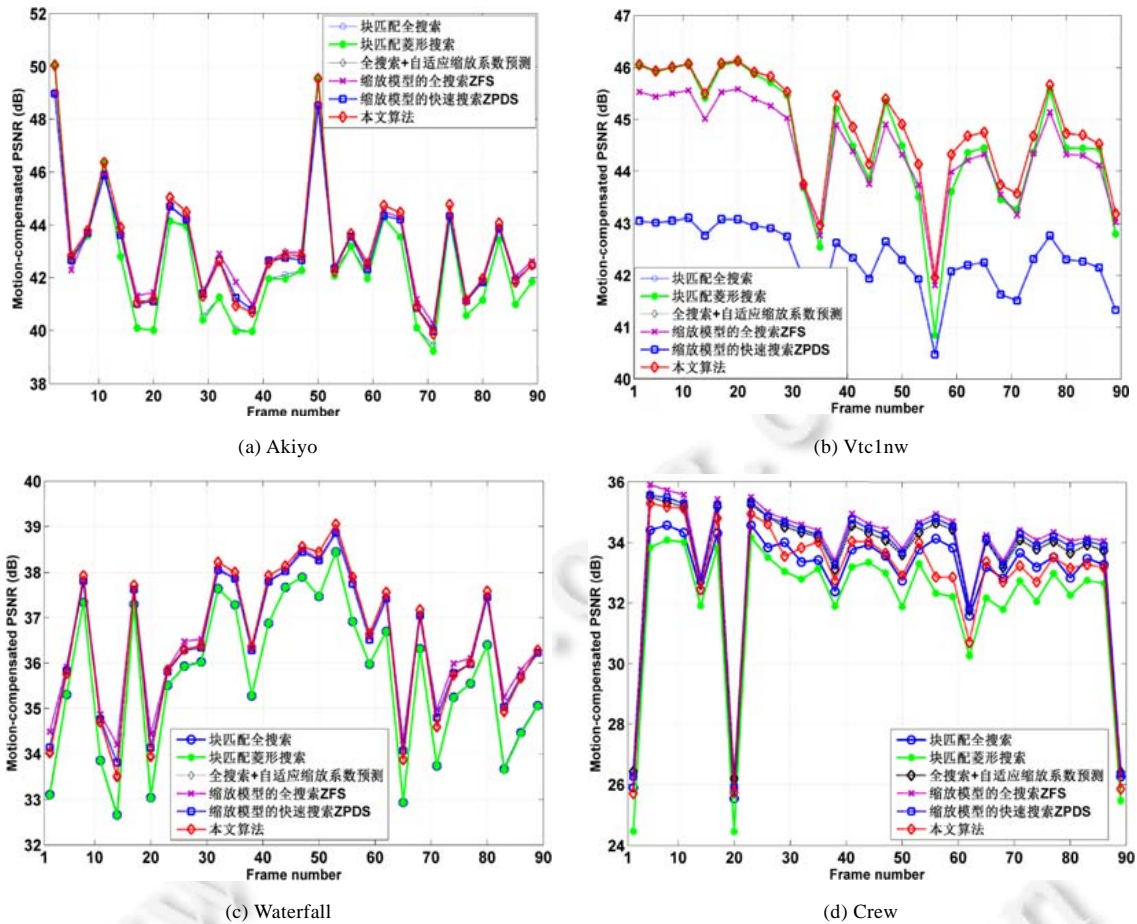


Fig.4 Frame-wise PSNR comparison of Akiyo, Vtc1nw, Waterfall, and Crew sequences

图4 Akiyo、Vtc1nw、Waterfall 和 Crew 序列的逐帧 PSNR 比较

5.2 计算复杂度分析

本节讨论上述 6 种算法的计算复杂度. 设待匹配块的尺寸为 $B \times B$ 像素, 搜索窗口的尺寸为 $W \times W$ 像素.

首先, 从处理 1 个宏块的时间复杂度来看:

- 块匹配全搜索的计算复杂度为 $O(B^2W^2)$; 块匹配全搜索+自适应缩放系数预测算法需要在全搜索的基础上, 计算 4 次空域/时域差分、2 次双线性插值和 2 次运动补偿误差, 故其复杂度为 $O(B^2W^2+8B^2)$.
- 根据文献[57,58], 块匹配菱形搜索的计算复杂度约为块匹配全搜索的 7%, 即 $O(0.07B^2W^2)$; 本文算法在菱形搜索的基础上, 还需计算 4 次空域/时域差分、2 次双线性插值和 2 次运动补偿误差, 则复杂度约为 $O(0.07B^2W^2+8B^2)$. 同时, 由于本文算法未以边信息的形式传输缩放运动系数, 解码端需在块平移运动向量的基础上计算 4 次空域/时域差分、2 次双线性插值和 2 次运动补偿误差, 从而自适应地重建该缩放系数, 其复杂度为 $O(8B^2)$.
- 由文献[50]的分析可知, 基于缩放模型的全搜索 ZFS 需要 3 重循环, 并且对于每个候选向量, 均要计算 1 轮双线性插值和 1 轮预测误差, 其时间复杂度为 $O(2B^2W^2zr)$. 而基于缩放模型的快速搜索 3D-ZPDS 的计算复杂度则约为前者的 13%^[50], 即 $O(0.26B^2W^2zr)$.

其次, 表 3 统计了 6 种算法在全部视频序列上的实际运行时间, 实验环境为: Intel Xeon(TM) i7-4790CPU @ 3.6GHz, 16GB 主存, 64 位 Windows 10 专业版, 集成开发环境为 Visual C++6.0. 从表 3 可见, 当 $B=16, W=33, zr=2$ 时,

本文算法的平均计算时间略高于块匹配菱形搜索,但是明显少于块匹配全搜索,较后者节省了 96.02%,加速 25.13 倍;并且,仅相当于缩放模型全搜索 ZFS 和快速搜索 3D-ZPDS 的 0.11%,3.86%,分别加速了 897 倍和 25.88 倍.主要原因在于,ZFS 算法为了估计出最优的缩放运动系数,需对每个像素共执行 4 356 次双线性插值,每次插值需要 8 次乘法和 4 次加法,合计 34 848 次乘法和 17 424 次加法;而本文则只需 2 次插值,另外又需 21 次乘法和 21 次加法来计算缩放运动系数 z_1 和 z_2 ,合计 37 次乘法和 29 次加法.因此,有效减少了亚像素插值操作的次数,是本文算法显著减少运算时间的关键所在.

综合理论分析和实际运行情况可见,本文算法的运动估计/补偿效率较之其他 5 种对比算法取得了有效的提升.

Table 3 Running time comparison of six algorithms on various test sequences (s/frame)
表 3 6 种算法在各个测试序列上的运行时间比较 (秒/帧)

视频序列	块匹配全搜索	块匹配菱形搜索	块匹配全搜索+自适应缩放系数	基于缩放模型的全搜索 ^[50]	3D-ZPDS ^[50]	本文算法
Akiyo	0.69	0.02	0.76	24.96	0.72	0.03
Coastguard	0.73	0.02	0.81	30.03	0.79	0.03
Container	0.80	0.01	0.84	27.09	0.85	0.03
Foreman	0.73	0.02	0.77	27.63	0.78	0.03
Hall	0.69	0.01	0.84	24.25	0.72	0.03
Highway	0.76	0.02	0.80	30.93	0.79	0.04
Husky	0.84	0.02	0.94	25.44	0.88	0.03
Ice	0.78	0.01	0.80	30.83	0.76	0.03
Intros_cif	0.72	0.02	0.76	30.67	0.68	0.03
Mobile	0.85	0.02	0.88	25.30	0.85	0.03
Mother & Daughter	0.74	0.02	0.77	24.72	0.78	0.03
News	0.76	0.01	0.82	24.30	0.77	0.03
Pamphlet	0.73	0.01	0.76	30.33	0.74	0.03
Paris	0.73	0.02	0.82	24.97	0.77	0.03
Sign_irene	0.80	0.02	0.83	26.82	0.80	0.03
Silent	0.73	0.01	0.78	25.82	0.71	0.02
Soccer	0.81	0.02	0.92	26.14	0.82	0.04
Students	0.71	0.01	0.79	25.09	0.76	0.03
Tempete	0.80	0.02	0.85	31.73	0.81	0.04
Tennis	1.02	0.02	1.05	28.34	0.85	0.03
Vtc1nw	0.73	0.01	0.74	26.42	0.67	0.03
Waterfall	0.77	0.02	0.87	25.22	0.75	0.03
Crew	2.80	0.08	3.41	118.19	2.92	0.14
City	3.52	0.10	3.73	128.80	3.51	0.16
Harbour	3.01	0.07	3.08	105.62	2.97	0.13
BQmall	3.38	0.09	3.53	123.50	3.68	0.14
Flower vase	2.51	0.06	3.01	107.76	2.66	0.11
FourPeople	7.19	0.16	8.54	271.30	8.42	0.33
Johnny	8.11	0.16	8.70	259.34	8.02	0.31
KristenAndSara	7.57	0.17	9.02	256.96	7.66	0.31
BQTerrace	4.08	0.10	5.56	152.13	4.27	0.17
Cactus	4.20	0.09	4.74	140.98	4.23	0.17
ParkScene	4.31	0.10	5.10	147.84	4.34	0.17
平均	2.01	0.05	2.28	71.76	2.07	0.08
加速比	1	40.20	0.88	0.03	0.97	25.13

6 结 论

为了克服传统平移模型无法表示复杂运动的不足,本文在整像素精度的块平移初始搜索基础上引进 1 个缩放运动系数,从而提高了块匹配运动估计的精度.为进一步减少求解缩放运动系数所需的亚像素插值次数,将运动补偿误差建模为关于该系数的一元二次函数,推导出缩放系数的预测方法,进而以菱形搜索为初始搜索,提出一种自适应缩放系数优化的快速块匹配运动估计算法.该算法可为改善现有块平移运动估计的补偿质量提供一种新思路,而且无需向解码端传输缩放运动向量.

另外,本文方法仍存在有待改进之处,如引进更高精度的亚像素插值算法和初始搜索点预测算法等,我们将

在后续研究工作中对这些思路展开深入探索。

References:

- [1] Fan R, Zhang YF, Li B. Motion classification-based fast motion estimation for high-efficiency video coding. *IEEE Trans. on Multimedia*, 2017,19(5):893–907. [doi: 10.1109/TMM.2016.2642786]
- [2] Ma SW. History and recent development of AVS video coding standards. *Journal of Computer Research and Development*, 2015, 52(1):27–37 (in Chinese with English abstract). [doi: 10.7544/issn1000-1239.2015.20140106]
- [3] Bossen F, Bross B, Sühning K, Flynn D. HEVC complexity and implementation analysis. *IEEE Trans. on Circuits and Systems for Video Technology*, 2012,22(12):1685–1696. [doi: 10.1109/TCSVT.2012.2221255]
- [4] Sinangil ME, Sze V, Zhou MH, Chandrakasan AP. Cost and coding efficient motion estimation design considerations for high efficiency video coding (HEVC) standard. *IEEE Journal of Selected Topics in Signal Processing*, 2013,7(6):1017–1028. [doi: 10.1109/JSTSP.2013.2273658]
- [5] Song CM, Min X, Yan XH, Wang XH, Yin BC. Fast elastic motion estimation using improved Levenberg-Marquardt method. *Journal of Software*, 2019,30(7):2208–2226 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5487.htm> [doi: 10.13328/j.cnki.jos.005487]
- [6] Lin WY, Panusopone K, Baylon DM, Sun MT, Chen ZZ, Li HX. A fast sub-pixel motion estimation algorithm for H.264/AVC video coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2011,21(2):237–243. [doi: 10.1109/TCSVT.2011.2106290]
- [7] Pan ZQ, Zhang Y, Kwong S, Wang X, Xu L. Early termination for TZSearch in HEVC motion estimation. In: Krishnamurthy V, Plataniotis K, eds. *Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*. Piscataway: IEEE Computer Society Press, 2013. 1389–1393. [doi: 10.1109/ICASSP.2013.6637879]
- [8] Tourapis AM, Au OC, Liou ML. Highly efficient predictive zonal algorithms for fast block-matching motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2002,12(10):934–947. [doi: 10.1109/TCSVT.2002.804894]
- [9] Fan R, Zhang YF, Li B, Wang G. Multidirectional parabolic prediction-based interpolation-free sub-pixel motion estimation. *Signal Processing: Image Communication*, 2017,53(1):123–134. [doi: 10.1016/j.image.2017.02.005]
- [10] Al-Najdawi N, Al-Najdawi MN, Tedmori S. Employing a novel cross-diamond search in a modified hierarchical search motion estimation algorithm for video compression. *Information Sciences*, 2014,268(1):425–435. [doi: 10.1016/j.ins.2013.08.009]
- [11] Song CM, Wang XH. Research progress in video motion estimation techniques in the wavelet domain. *Chinese Journal of Computers*, 2005,28(10):1716–1727 (in Chinese with English abstract).
- [12] Wang CN, Yang SW, Liu CM, Chiang T. A hierarchical N -queen decimation lattice and hardware architecture for motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2004,14(4):429–440. [doi: 10.1109/TCSVT.2004.825550]
- [13] Alzoubi H, Pan WD. Fast and accurate global motion estimation algorithm using pixel subsampling. *Information Sciences*, 2008, 178(1):3415–3425. [doi: 10.1016/j.ins.2008.05.004]
- [14] Hui KC, Siu WC, Chan YL. New adaptive partial distortion search using clustered pixel matching error characteristic. *IEEE Trans. on Image Processing*, 2005,14(5):597–607. [doi: 10.1109/TIP.2005.846020]
- [15] Xiao XP, Liu EH, Qin JJ. A fast partial distortion search algorithm for motion estimation based on the multi-traps assumption. *Signal Processing: Image Communication*, 2015,31(1):25–33. [doi: 10.1016/j.image.2014.11.007]
- [16] Yavuz S, Çelebi A, Aslam M, Urhan O. Selective gray-coded bit-plane based low-complexity motion estimation and its hardware architecture. *IEEE Trans. on Consumer Electronics*, 2016,62(1):76–84. [doi: 10.1109/TCE.2016.7448566]
- [17] AlQaralleh EA, Abu-Sharkh OMF. Low-complexity motion estimation design using modified XOR function. *Multimedia Tools and Applications*, 2016,75(24):16809–16834. [doi: 10.1007/s11042-015-2948-z]
- [18] Ertürk S. Multiplication-free one-bit transform for low-complexity block-based motion estimation. *IEEE Signal Processing Letters*, 2007,14(2):109–112. [doi: 10.1109/LSP.2006.882088]
- [19] Song CM, Guo YW, Wang XH, Liu D. Fuzzy quantization based bit transform for low bit-resolution motion estimation. *Signal Processing: Image Communication*, 2013,28(10):1435–1447. [doi: 10.1016/j.image.2013.09.007]

- [20] Song CM, Min X, Yan XH, Wang XH, Yin BC. Elastic motion estimation algorithm using two-bit-depth pixels. *Journal of Computer Research and Development*, 2019,56(11):2469–2484 (in Chinese with English abstract). [doi: 10.7544/issn1000-1239.2019.20180699]
- [21] Li B, Xu JZ, Wu F. A unified framework of hash-based matching for screen content coding. In: Farrugia R, Magli E, Zhang J, eds. *Proc. of the IEEE Visual Communications and Image Processing Conf.* Piscataway: IEEE Computer Society Press, 2014. 530–533. [doi: 10.1109/VCIP.2014.7051623]
- [22] Zhu WJ, Ding WP, Xu JZ, Shi YH, Yin BC. Hash-based block matching for screen content coding. *IEEE Trans. on Multimedia*, 2015,17(7):935–944. [doi: 10.1109/TMM.2015.2428171]
- [23] Xiao W, Shi GM, Li B, Xu JZ, Wu F. Fast Hash-based inter block matching for screen content coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2018,28(5):1169–1182. [doi: 10.1109/TCSVT.2016.2643701]
- [24] Pan ZQ, Lei JJ, Zhang Y, Sun XM, Kwong S. Fast motion estimation based on content property for low-complexity H.265/HEVC encoder. *IEEE Trans. on Broadcasting*, 2016,62(3):675–684. [doi: 10.1109/TBC.2016.2580920]
- [25] Narroschke M, Swoboda R. Extending HEVC by an affine motion model. In: Antonio O, Said A, eds. *Proc. of the Picture Coding Symp.* Piscataway: IEEE Computer Society Press, 2013. 321–324. [doi: 10.1109/PCS.2013.6737748]
- [26] Ahmmed A, Hannuksela MM, Gabbouj M. Fisheye video coding using elastic motion compensated reference frames. In: Pereira F, Sharma G, eds. *Proc. of the IEEE Int'l Conf. on Image Processing.* Piscataway: IEEE Computer Society Press, 2016. 2027–2031. [doi: 10.1109/ICIP.2016.7532714]
- [27] Liu D, Wu F. Advances and trends of video coding technologies. *Communications of the CCF*, 2016,12(6):20–23 (in Chinese with English abstract).
- [28] Muhit AA, Pickering MR, Frater MR, Arnold JF. Video coding using fast geometry-adaptive partitioning and an elastic motion model. *Journal of Visual Communication and Image Representation*, 2012,23(1):31–41. [doi: 10.1016/j.jvcir.2011.07.003]
- [29] Wang Y, Lee O. Use of two-dimensional deformable mesh structures for video coding, Part I—The synthesis problem: mesh-based function approximation and mapping. *IEEE Trans. on Circuits and Systems for Video Technology*, 1996,6(6):636–646. [doi: 10.1109/76.544735]
- [30] Wang Y, Lee O, Vetro A. Use of two-dimensional deformable mesh structures for video coding, Part II—The analysis problem and a region-based coder employing an active mesh representation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1996, 6(6):647–659. [doi: 10.1109/76.544736]
- [31] Altunbasak Y, Tekalp AM. Closed-form connectivity-preserving solutions for motion compensation using 2-D meshes. *IEEE Trans. on Image Processing*, 1997,6(9):1255–1269. [doi: 10.1109/83.623189]
- [32] Wu F, Gao P, Gao W. Motion estimation technologies based on mesh model. *Acta Electronica Sinica*, 2000,28(5):47–51 (in Chinese with English abstract).
- [33] Al-Regib G, Altunbasak Y, Mersereau RM. Hierarchical motion estimation with content-based meshes. *IEEE Trans. on Circuits and System for Video Technology*, 2003,13(10):1000–1005. [doi: 10.1109/TCSVT.2003.816520]
- [34] Cui SX, Wang YH, Fowler JE. Motion estimation and compensation in the redundant-wavelet domain using triangle meshes. *Signal Processing: Image Communication*, 2006,21(7):586–598. [doi: 10.1016/j.image.2006.03.011]
- [35] Jin SY. Research of motion object-based mesh model for motion estimation [MS. Thesis]. Xi'an: Northwestern Polytechnical University, 2006 (in Chinese with English abstract).
- [36] Li L, Li HQ, Liu D, Li Z, Yang HT, Lin SX, Chen HB, Wu F. An efficient four-parameter affine motion model for video coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2018,28(8):1934–1947. [doi: 10.1109/TCSVT.2017.2699919]
- [37] Huang H. Research on Inter/Intra prediction and optimization techniques in HEVC [Ph.D. Thesis]. Beijing: Beijing Jiaotong University, 2013 (in Chinese with English abstract).
- [38] Rath GB, Makur A. Iterative least squares and compression based estimations for a four-parameter linear global motion model and global motion compensation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1999,9(7):1075–1099. [doi: 10.1109/76.795060]
- [39] Wiegand T, Steinbach E, Girod B. Affine multipicture motion-compensated prediction. *IEEE Trans. on Circuits and Systems for Video Technology*, 2005,15(2):197–209. [doi: 10.1109/TCSVT.2004.841690]

- [40] Kordasiewicz RC, Gallant MD, Shirani S. Affine motion prediction based on translational motion vectors. *IEEE Trans. on Circuits and Systems for Video Technology*, 2007,17(10):1388–1394. [doi: 10.1109/TCSVT.2007.903777]
- [41] Huang H, Woods JW, Zhao Y, Bai HH. Control-point representation and differential coding affine-motion compensation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2013,23(10):1651–1660. [doi: 10.1109/TCSVT.2013.2254977]
- [42] Ren L. The application of deformable block matching algorithm [MS. Thesis]. Tianjin: Tianjin University, 2007 (in Chinese with English abstract).
- [43] Zheng JL, Qin TF, Ni GN. Adaptive global motion estimation method based on rate distortion optimization. *Journal of Image and Graphics*, 2011,16(8):1346–1352 (in Chinese with English abstract).
- [44] Muhit AA, Pickering MR, Frater MR, Arnold JF. Video coding using elastic motion model and larger blocks. *IEEE Trans. on Circuits and Systems for Video Technology*, 2010,20(5):661–672. [doi: 10.1109/TCSVT.2010.2045804]
- [45] Deng BY. Study on video compression based on elastic motion model [MS. Thesis]. Xiamen: Xiamen University, 2014 (in Chinese with English abstract).
- [46] Song CM, Zhao CW, Liu D, Wang XH. Elastic motion estimation of video using improved Gauss-Newton method. *Ruan Jian Xue Bao/Journal of Software*, 2016,27(11):2946–2960 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4886.htm> [doi: 10.13328/j.cnki.jos.004886]
- [47] Zhao CW. Research on motion estimation algorithm using elastic model [MS. Thesis]. Dalian: Liaoning Normal University, 2015 (in Chinese with English abstract).
- [48] Po LM, Wong KM, Cheung KW, Ng KH. Subsampled block-matching for zoom motion compensated prediction. *IEEE Trans. on Circuits and Systems for Video Technology*, 2010,20(11):1625–1637. [doi: 10.1109/TCSVT.2010.2087474]
- [49] Yuan H, Liu J, Sun JD, Liu HC, Li YJ. Affine model based motion compensation prediction for zoom. *IEEE Trans. on Multimedia*, 2012,14(4):1370–1375. [doi: 10.1109/TMM.2012.2190393]
- [50] Kim HS, Lee JH, Kim CK, Kim BG. Zoom motion estimation using block-based fast local area scaling. *IEEE Trans. on Circuits and Systems for Video Technology*, 2012,22(9):1280–1291. [doi: 10.1109/TCSVT.2012.2198137]
- [51] Qiu WS. *Analytic Geometry*. 3rd ed., Beijing: Peking University Press, 2015 (in Chinese).
- [52] Jozawa H, Kamikura K, Sagata A, Kotera H, Watanabe H. Two-stage motion compensation using adaptive global MC and local affine MC. *IEEE Trans. on Circuits and Systems for Video Technology*, 1997,7(1):75–85. [doi: 10.1109/76.554419]
- [53] Yuan H, Chang YL, Lu ZY, Ma YZ. Model based motion vector predictor for zoom motion. *IEEE Signal Processing Letters*, 2010, 17(9):787–790. [doi: 10.1109/LSP.2010.2055051]
- [54] Kim HS, Jung H, Park CS, Kim BG. Multi-scale based motion estimation for high quality video coding. In: Mozar S, ed. *Proc. of the IEEE Int'l Conf. on Consumer Electronics*. Piscataway: IEEE Computer Society Press, 2011. 703–704. [doi: 10.1109/ICCE.2011.5722819]
- [55] Zhang JY, Li B, Yao CL. A fast sub-pixel motion estimation algorithm based on best position calculation. *Chinese Journal of Computers*, 2005,28(9):1541–1548 (in Chinese with English abstract).
- [56] Lu JY, Liu YX, Hou F, Huang CH, Chao HY. High efficiency algorithm of fractional pixel motion compensation for HEVC. *Ruan Jian Xue Bao/Journal of Software*, 2017,28(8):2214–2226 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5122.htm> [doi: 10.13328/j.cnki.jos.005122]
- [57] Zhu S, Ma KK. A new diamond search algorithm for fast block-matching motion estimation. *IEEE Trans. on Image Processing*, 2000,9(2):287–290. [doi: 10.1109/83.821744]
- [58] Li RX, Zeng B, Liou ML. A new three-step search algorithm for block motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1994,4(4):438–442. [doi: 10.1109/76.313138]

附中文参考文献:

- [2] 马思伟. AVS 视频编码标准技术回顾及最新进展. *计算机研究与发展*, 2015,52(1):27–37. [doi: 10.7544/issn1000-1239.2015.20140106]
- [5] 宋传鸣, 闵新, 闫小红, 王相海, 尹宝才. 采用改进 Levenberg-Marquardt 法的快速弹性运动估计. *软件学报*, 2019,30(7):2208–2226. <http://www.jos.org.cn/1000-9825/5487.htm> [doi: 10.13328/j.cnki.jos.005487]

- [11] 宋传鸣,王相海.小波域视频运动估计研究进展.计算机学报,2005,28(10):1716-1727.
- [20] 宋传鸣,闵新,闫小红,王相海,尹宝才.采用 2b 深度像素的弹性运动估计算法.计算机研究与发展,2019,56(11):2469-2484.
- [27] 刘东,吴枫.视频编码技术与趋势.中国计算机学会通讯,2016,12(6):20-23.
- [32] 吴枫,高鹏,高文.基于网格模型的运动估计技术.电子学报,2000,28(5):47-51.
- [35] 金晟毅.基于运动对象网格模型的运动估计方法的研究[硕士学位论文].西安:西北工业大学,2006.
- [37] 黄晗.HEVC 帧间/帧内预测及优化技术研究[博士学位论文].北京:北京交通大学,2013.
- [42] 任亮.可变形块匹配算法(DBMA)的应用研究[硕士学位论文].天津:天津大学,2007.
- [43] 郑嘉利,覃团发,倪光南.结合率失真优化的自适应全局运动估计方法.中国图像图形学报,2011,16(8):1346-1352.
- [45] 邓斌攸.基于弹性运动模型的视频压缩算法研究[硕士学位论文].厦门:厦门大学,2014.
- [46] 宋传鸣,赵长伟,刘丹,王相海.采用改进高斯-牛顿法的视频弹性运动估计.软件学报,2016,27(11):2946-2960. <http://www.jos.org.cn/1000-9825/4886.htm> [doi: 10.13328/j.cnki.jos.004886]
- [47] 赵长伟.基于弹性模型的运动估计算法研究[硕士学位论文].大连:辽宁师范大学,2015.
- [51] 丘维声.解析几何.第 3 版.北京:北京大学出版社,2015.
- [55] 张蕊尹,李波,姚春莲.基于最优位置计算的快速亚像素运动估计.计算机学报,2005,28(9):1541-1548.
- [56] 陆寄远,刘宇熹,侯昉,黄承慧,朝红阳.HEVC 的高效分像素运动补偿.软件学报,2017,28(8):2214-2226. <http://www.jos.org.cn/1000-9825/5122.htm> [doi: 10.13328/j.cnki.jos.005122]



宋传鸣(1980—),男,博士,教授,CCF 专业会员,主要研究领域为图像及视频编码,多尺度几何分析,多媒体水印技术.



王相海(1964—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为图像及视频信息处理,遥感影像信息处理,计算机图形学.



闫小红(1993—),女,博士生,主要研究领域为视频编码.



尹宝才(1963—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为数字多媒体技术,多功能感知,虚拟现实与计算机图形学.



葛明博(1994—),男,硕士,主要研究领域为视频编码.