

采用改进高斯-牛顿法的视频弹性运动估计*

宋传鸣¹, 赵长伟¹, 刘丹^{1,2}, 王相海¹



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 信息与通信工程学院, 辽宁 大连 116024)

通讯作者: 刘丹, E-mail: liudan_dl@163.com, http://mtl.lnpu.edu.cn/member/songchuanming/homepage.html

摘要: 运动估计是去除视频时间维冗余的编码技术,而目前通用的平移运动模型无法有效地表示物体的局部非刚性复杂运动.为此,提出一种基于改进高斯-牛顿法的弹性运动估计方法.首先,通过分析初始迭代点对高斯-牛顿迭代结果的影响,采用基于 2bit 深度像素的均匀搜索预测初始迭代点;其次,通过理论和实验分析发现,不同的迭代步长对弹性运动估计/补偿性能有明显的影响,采用离散余弦变换的低频能量比率估计步长的上限,再利用黄金分割法对步长进行求精.实验结果表明,对于具有不同场景特点的视频序列,该算法始终能够保持较高的估计精度,运动补偿的平均峰值信噪比,比基于块平移模型的全搜索算法和传统弹性运动估计算法分别提高 1.73dB 和 1.42dB.并且,该算法具有更快的收敛速度,一般仅需 1~3 次迭代就能取得高于传统弹性运动估计和块平移全搜索的峰值信噪比.

关键词: 视频编码;运动估计;弹性模型;高斯-牛顿法

中图法分类号: TP391

中文引用格式: 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

英文引用格式: 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). http://www.jos.org.cn/1000-9825/4886.htm

Elastic Motion Estimation of Video Using Improved Gauss-Newton Method

SONG Chuan-Ming¹, ZHAO Chang-Wei¹, LIU Dan^{1,2}, WANG Xiang-Hai¹

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

²(School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: Motion estimation is a coding technique to eliminate the temporal redundancy of video. However, state-of-the-art translational motion model is not able to efficiently represent objects' local non-rigid complex motion. To address the issue, an elastic motion estimation algorithm is developed in this paper based on modified Gauss-Newton method. The effect of initial iteration point is first analyzed on the result of the Gauss-Newton method, and a two bit-depth pixel based uniform search is used to predict the initial iteration point. Subsequently, it is found that different step size has obvious influence on the performance of the elastic motion estimation by both theoretical and experimental analyses. The ratio of low-frequency energy of discrete cosine transform is employed to estimate the upper bound of the step size which is then refined by the golden ratio method. Experimental results show that the proposed algorithm is able to obtain stable performance for video sequences with various scene characteristics. It gains 1.73dB and 1.42dB higher average motion-compensated peak signal-to-noise ratio (PSNR) than those of the full search algorithm based on block-wise translational model and conventional elastic motion estimation, respectively. Furthermore, the proposed algorithm has faster convergence speed. Only 1~3

* 基金项目: 国家自然科学基金(61402214, 41271422); 高等学校博士学科点专项科研基金(20132136110002); 辽宁省教育厅科学研究一般项目(L2013406); 大连市科学技术基金(2013J21DW027)

Foundation item: National Natural Science Foundation of China (61402214, 41271422); Specialized Research Fund for the Doctoral Program of Higher Education (20132136110002); Foundation of Science and Technology Plan for Higher Education of Liaoning Province (L2013406); Dalian Foundation for Science and Technology (2013J21DW027)

收稿时间: 2015-06-06; 修改时间: 2015-08-11; 采用时间: 2015-08-17

iterations are needed before the proposed algorithm achieves higher PSNR than conventional elastic motion estimation and block-wise translational full search method.

Key words: video coding; motion estimation; elastic model; Gauss-Newton algorithm

运动估计是一种有效的时域预测技术,而且多年来,视频编码效率的大部分提升均是由于其采用了更为优秀的运动估计算法^[1,2]。然而,运动估计环节所消耗的计算资源往往占编码器全部资源的 50%以上,甚至可达 80%^[3,4]。故此,为了在复杂度和预测精度之间达到较好的折中,现有视频编码标准均采用基于平移模型的块匹配运动估计算法,并陆续提出了多种快速块匹配策略。

- 基于候选向量下采样的策略^[5-11]:菱形搜索和非对称十字多层次六边形格点搜索 UMHexagonS 等绝大多数快速运动估计均属此类。
- 基于像素下采样的策略^[12-17]:如部分失真搜索、多分辨率搜索和自适应像素下采样搜索等。
- 基于低复杂度匹配函数的策略^[18-21]:常用函数包括误差绝对值之和(sum of absolute difference,简称 SAD)、像素误差分类(pel difference classification,简称 PDC)和相异像素数目(different pixel count,简称 DPC)等。
- 基于低比特深度像素的策略^[22-26]:如 1bit 搜索、2bit 搜索等。

但是,平移模型无法有效预测由物体的旋转、缩放、变形和摄像机运动产生的非刚性复合运动。于是,H.264/AVC 和 H.265/HEVC 等编码标准进一步采用基于四叉树的分层次可变尺寸块结构来逼近复杂运动场,最小可以支持 4×4 像素的块运动估计。而随着块尺寸的减小,用于编码运动矢量的码流开销和各种软硬件计算开销也逐渐增加。根据文献[4]的研究结论:当块尺寸减小至 4×4 像素时,软、硬件开销的增加幅度已经超过其率失真性能的提升幅度。这说明,仅仅依靠传统块平移模型来实现时域预测的视频编码方法越来越无法很好地满足视频快速发展的需求。因此,研究一种能够有效地表示复杂运动场的运动模型及其参数求解策略,对于新一代视频编码的研究具有重要意义。

为了弥补平移运动模型的不足,一些研究人员多年来尝试将高阶运动模型引入到运动估计/补偿中。依据所采用模型的不同,现有的基于非平移模型的运动估计方法可以分为 4 类:基于网格模型的运动估计^[27-33]、基于多项式模型的运动估计^[34-41](如四参数模型^[34]、六参数仿射模型^[35-38]、八参数透视或双线性模型^[39,40]、十二参数变换模型^[41]及其混合模型^[41])、基于缩放模型的运动估计^[42,43]和基于弹性模型的运动估计^[44-47]。基于网格模型的运动估计方法避免了块效应,且视觉效果更佳,但是对相邻网格的共有控制点进行运动估计和率失真优化的难度较大,尤其是当网格尺寸不一致时,目前还缺乏行之有效的优化方法^[38]。基于多项式模型的运动估计方法可描述丰富的运动形式,运动补偿精度高于块匹配方法。但是随着模型参数的增多,其搜索复杂度明显提高,并且该模型大多用于视频全局运动估计,对局部复杂运动的刻画能力不足。基于缩放模型的运动估计方法可有效捕获视频中的物体缩放,却不能表示旋转、错切等运动。基于弹性模型的运动估计可有效弥补上述模型的不足,它不仅便于表示物体的全局和局部运动,而且可依据块尺寸和视频的具体特性灵活调整模型参数的个数,并根据需要选择离散余弦函数、小波函数或样条函数等作为基函数,进而控制运动估计/补偿的效率^[45]。文献[45]研究表明,在相同码率下,基于弹性运动估计的解码视频的峰值信噪比(peak signal-to-noise ratio,简称 PSNR)比块匹配运动估计高 0.7dB。文献[47]则进一步将弹性运动估计引进到 H.265 中,其实验结果显示,在图像失真度损失约 1%的情况下,弹性运动模型可将 H.265 的码率降低 3%~12%。故此,理论和实验数据均表明,弹性运动模型是一种表示复杂运动场的高效率模型。

目前,有关基于弹性模型进行运动估计/补偿的研究主要集中在两个方面:首先,弹性运动模型与现有视频编码标准的结合方式,如文献[46]提出应依据图像的几何特征确定帧划分方式,使得弹性模型能够更准确地捕获局部运动;其次,弹性运动模型的加速求解策略,如文献[48]采用 1bit 深度像素讨论了弹性运动模型的高斯-牛顿迭代法,文献[49,50]则进一步将其推广到了 2bit 深度像素的情况。然而,上述工作均忽视了视频数据的复杂性对高斯-牛顿法收敛性的影响:第一,理论证明,高斯-牛顿法收敛到全局最优解的前提是初始迭代点接近全局最优解(即位于全局最优解的单调区间内)^[51]或目标函数在解空间内具有单调性;第二,文献[46-50]均采用了文献

[52]提出的高斯-牛顿法完成弹性运动估计,该方法在每次迭代中的步长都为 1,而文献[53]指出,目标函数偏离线性的距离越远,高斯-牛顿法的收敛速度就越慢,此时需选用更小的步长.事实上,视频数据以及运动补偿误差的复杂性既不能保证匹配误差曲面满足第 1 条单调性,也不会呈现第 2 条期望的理想线性(第 2.1 节的实验将证明这一点).在这种情况下,现有的高斯-牛顿法并不能给出弹性运动估计的最优解.

本文在回顾弹性运动估计及高斯-牛顿法的基础上,将通过实验论证初始迭代点和迭代步长的选取对高斯-牛顿收敛性的影响,并进一步探讨基于 2bit 深度像素的初始迭代点预测方法和基于低频能量比率、黄金分割法的步长计算策略,从而给出一种基于改进高斯-牛顿法的视频弹性运动估计方法.

1 弹性运动估计及其高斯-牛顿解法

为了便于下文工作的论述,本节首先介绍弹性运动模型,然后讨论模型的高斯-牛顿迭代解法.

1.1 弹性运动模型

视频运动估计的目标是最小化如下的公式:

$$e(\mathbf{v}) = \sum_{i=1}^N f[I(x_i, y_i) - R(g(x_i, \mathbf{v}_x), g(y_i, \mathbf{v}_y))] \quad (1)$$

其中, I 表示视频的当前帧, R 表示参考帧, N 表示视频帧的像素总数, f 表示某种匹配误差的度量准则,如 SAD, SSD(sum of squared difference)等; g 表示某种运动模型; \mathbf{v} 则表示在某特定运动模型下的运动向量(或运动参数), \mathbf{v}_x 和 \mathbf{v}_y 为其沿着 x 轴和 y 轴的分量.例如,传统块平移运动模型对应的 g 函数为 $g(x_i, \mathbf{v}) = x_i + \mathbf{v}_x$ 和 $g(y_i, \mathbf{v}) = y_i + \mathbf{v}_y$. 而弹性运动模型对应的 g 函数为

$$g(x_i, \mathbf{v}) = x_i + \sum_{k=1}^{p/2} \mathbf{v}_k \varphi_k(x_i, y_i) \quad (2)$$

$$g(y_i, \mathbf{v}) = y_i + \sum_{k=p/2+1}^p \mathbf{v}_k \varphi_k(x_i, y_i) \quad (3)$$

其中, φ_k 为基函数, p 为运动参数的个数.考虑到视频中的运动矢量场大多是平滑的,离散余弦函数基能为其提供紧致表示,而且离散余弦变换又是现有视频编码的必备模块之一,文献[45]采用离散余弦函数基作为弹性运动模型的基函数,即

$$\varphi_k(x_i, y_i) = \varphi_{k+p/2}(x_i, y_i) = \cos\left(\frac{(2x_i+1)\pi u}{2M}\right) \cos\left(\frac{(2y_i+1)\pi v}{2N}\right) \quad (4)$$

其中, $k = su + v + 1, u, v = 0, 1, \dots, s-1, s = \sqrt{p/2}$. 当然,也可以根据应用需要选择样条函数、小波函数和多项式函数等作为基函数.另外,文献[45]又选取 SSD 作为匹配误差的度量准则.这样,公式(1)就变成了一个无约束的非线性最小二乘问题.

$$\min_{\mathbf{v}} e(\mathbf{v}) = \frac{1}{2} \sum_{i=1}^N [I(x_i, y_i) - R(g(x_i, \mathbf{v}_x), g(y_i, \mathbf{v}_y))]^2 \quad (5)$$

注意,公式(5)中的 1/2 是为了后文证明方便.

1.2 高斯-牛顿迭代求解算法

高斯-牛顿迭代法是求解上述无约束最小二乘问题的一种有效方法^[51,53],其基本思想是:采用 1 阶泰勒(Taylor)展开式对匹配误差函数进行线性逼近,从而通过反复迭代求得其极小值.假设当前的迭代点是 \mathbf{v}_k , 并令 $r_i(\mathbf{v}) = I(x_i, y_i) - R(g(x_i, \mathbf{v}_x), g(y_i, \mathbf{v}_y))$, 那么可利用函数 $r_i(\mathbf{v})$ 在 \mathbf{v}_k 处的 1 阶泰勒展开式来近似 $r_i(\mathbf{v})$, 即

$$r_i(\mathbf{v}_k + \mathbf{d}) \triangleq r_i(\mathbf{v}_k) + r_i'(\mathbf{v}_k) \mathbf{d} \quad (6)$$

其中, $\mathbf{d} = \mathbf{v}_{k+1} - \mathbf{v}_k, \mathbf{v}_{k+1}$ 表示下一个迭代点; $r_i'(\mathbf{v}_k)$ 表示 $r_i(\mathbf{v})$ 在 \mathbf{v}_k 处的 1 阶导数.将上式代入公式(5),就有:

$$\min_{\mathbf{d}} e_k(\mathbf{d}) = \frac{1}{2} \sum_{i=1}^N r_i^2 \triangleq \frac{1}{2} \sum_{i=1}^N [r_i(\mathbf{v}_k) + r_i'(\mathbf{v}_k) \mathbf{d}]^2 \quad (7)$$

于是有:

$$e_k(\mathbf{d}) = \frac{1}{2} \sum_{i=1}^N [d^T (r_i)^T r_i' d + 2d^T (r_i)^T r_i + r_i^T r_i] \quad (8)$$

其中,上标 T 表示转置.令 $e_k'(\mathbf{d}) = 0$,则有:

$$(r_i')^T r_i' d = -(r_i')^T r_i \quad (9)$$

再令 $\mathbf{H} = (r_i')^T r_i', \mathbf{b} = -(r_i')^T r_i$,进而解得:

$$\mathbf{d} = \mathbf{H}^{-1} \mathbf{b} \quad (10)$$

以及下一个迭代点 \mathbf{v}_{k+1} :

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \mathbf{d} = \mathbf{v}_k + \mathbf{H}^{-1} \mathbf{b} \quad (11)$$

此时,若 $e(\mathbf{v}_{k+1})$ 满足迭代终止条件,则停止迭代;否则,利用公式(10)和公式(11)计算产生下一步迭代点,以此循环下去.

2 基于 2bit 深度像素的初始点预测

文献[51]从理论上证明:高斯-牛顿法能够收敛到全局最优解的基础是,初始迭代点位于全局最优解的单调区间内或目标函数在解空间内具有单调性.而匹配误差的曲面分布往往是多峰的,存在多个局部最优解^[54],这就不可避免地影响高斯-牛顿法的计算结果.本节首先通过实验分析初始迭代点对收敛性的影响,然后讨论一种初始迭代点的快速预测算法.

2.1 初始迭代点对高斯-牛顿迭代结果的影响

为了讨论初始迭代点的重要性,选取 Akiyo 和 Bus 视频序列的第 2 帧进行实验,以 $[-16,+16] \times [-16,+16]$ 窗口中的每个整数点为初始迭代点,依据文献[45]的高斯-牛顿法步骤迭代 15 次,再分别统计其运动补偿的 PSNR.图 1(a)所示是 Akiyo 第 2 帧以(160,144)像素为左上角的宏块从每个迭代点出发所得 PSNR 的分布情况,图 1(b)所示为 Bus 第 2 帧以(64,64)像素为左上角的宏块从每个迭代点出发所得 PSNR 的分布情况.

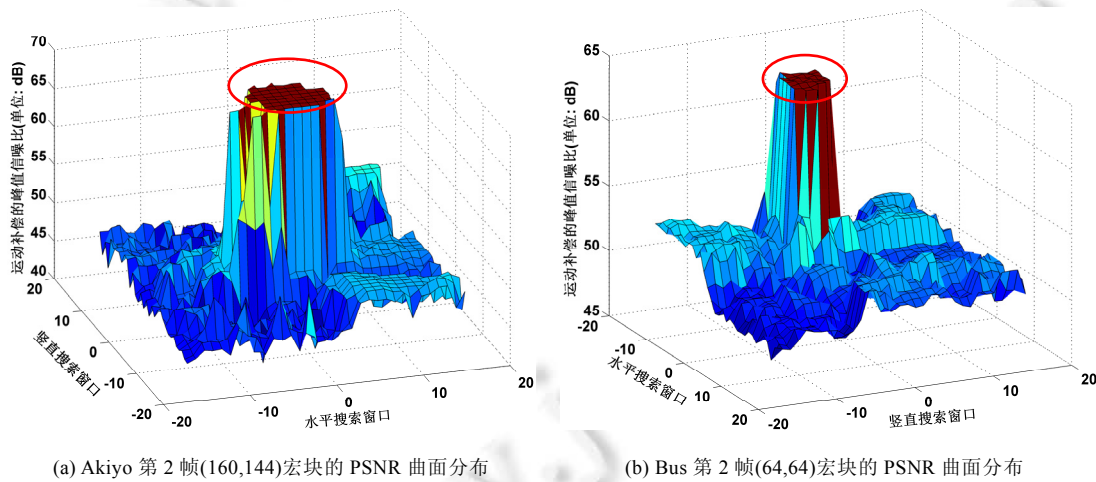


Fig.1 Surface distribution of PSNR from different initial vectors

图 1 不同初始迭代点所得 PSNR 的曲面分布

由图 1 可以看出,首先,从不同初始迭代点出发,所得的运动补偿结果迥异,且存在多处局部极小点,这表明合理选取初始迭代点对于最终的运动补偿质量是至关重要的;其次,无论是运动复杂度和空间复杂度都低的视频序列(如 Akiyo),还是运动量大、空间纹理细节复杂的视频序列(如 Bus),其运动补偿 PSNR 的全局极大点附近都形成了一个平滑的峰顶(如图 1(a)和图 1(b)中用椭圆圈出的部分所示).进一步统计显示:物体的运动复杂度越

低,其平滑峰顶的面积就越大,且绝大多数平滑峰顶的面积大于 5×5 像素.这说明,只要所选初始迭代点位于全局极大点周围、 5×5 像素范围内,就可以获得最优的运动补偿结果.故此,在弹性运动估计之前,有必要对其初始迭代点进行预测.

2.2 初始迭代点的快速预测算法

由公式(2)~公式(4),弹性运动模型的本质是对运动矢量场进行离散余弦分解,其中, \mathbf{v}_1 和 $\mathbf{v}_{p/2+1}$ 为直流分量,其余 $(p-2)$ 个运动参数为交流分量.根据离散余弦变换的基本理论,平滑运动场经过离散余弦分解后绝大部分的能量将集中在直流分量上,即 \mathbf{v}_1 和 $\mathbf{v}_{p/2+1}$,而这两个分量恰好对应平移运动.于是,预测初始迭代点主要是确定宏块的水平和竖直位移量.为了以尽可能低的计算复杂度预测平移分量,本文采用类似于文献[55]的基于 2bit 深度像素的起点预测算法,基本思想是:将当前帧和参考帧转换成 2bit 深度表示,再根据均匀搜索起点分布模板搜索初始的平移向量.具体步骤如下:

- Step 1. 计算当前帧和参考帧的均值 $\mu_c = E[I_c], \mu_r = E[I_r]$ 和方差 $\sigma_c^2 = E[I_c^2] - E^2[I_c], \sigma_r^2 = E[I_r^2] - E^2[I_r]$.
- Step 2. 对当前帧和参考帧分别进行直方图均衡变换得到 g' .
- Step 3. 对 g' 施以 4 级均匀量化,得到初始量化阈值 $\{T_j\} (j \in \{1, 2, 3\})$.
- Step 4. 通过逆变换 $T_j = e^{-1}(T'_j)$ 调整 $\{T_j\}$, 得到最终的量化阈值:

$$\{T_j\} : e^{-1}(T'_j) = \min\{k\} \quad \text{s.t.} \quad \{e(k) \geq T'_j\} \tag{12}$$

其中, $k \in \{0, 1, 2, \dots, 255\}$.

- Step 5. 利用 $\{T_j\}$ 对当前帧和参考帧分别进行 2bit 转换:

$$t(g) = \begin{cases} 00, & -1 < g \leq T_1 \\ 01, & T_1 < g \leq T_2 \\ 10, & T_2 < g \leq T_3 \\ 11, & T_3 < g \leq 255 \end{cases} \tag{13}$$

Step 6. 对于当前帧的每个宏块,按照图 2 的搜索模板计算各候选初始迭代点的匹配误差(图中每个图案表示一个待匹配的候选初始迭代点).考虑到平滑峰顶的面积大于 5×5 像素(半径为 2 个像素),模板将候选初始迭代点的搜索步长设置为 4,以保证全局最优解的单调区间内至少分布着 1 个候选初始迭代点.同时,由于自然视频序列的运动向量具有中心偏置特性^[56],本文依据到原点的距离进行分层次搜索,相同图案的点表示其位于同一搜索层次.在连续的 3 轮搜索中,若第 i 层 ($1 \leq i \leq 7$) 上的最小匹配误差 $MinE_i$ 连续小于第 $(i+1)$ 层和第 $(i+2)$ 层上的最小匹配误差 $MinE_{i+1}, MinE_{i+2}$, 则将 $MinE_i$ 对应的向量选定为初始迭代点,转入 Step 7; 否则,继续第 $(i+3)$ 层的搜索.若最外层搜索结束时仍未有满足上述条件的向量,则将全局的最小匹配误差点作为初始迭代点,转入 Step 7.

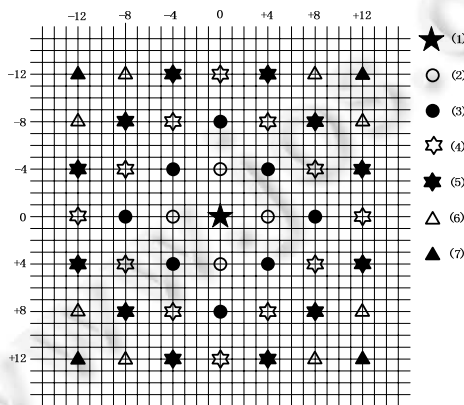


Fig.2 Uniform starting search point pattern

图 2 均匀起始搜索点模式

此外,在搜索过程中,匹配误差衡量准则为基于位截断的 SAD 准则:

$$BTSAD(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |I_c^{q:q+1}(x, y) - I_r^{q:q+1}(x + u, y + v)| \quad (14)$$

其中,下标 c 和 r 仍表示当前宏块和参考宏块,上标 $q:q+1$ 表示减法和绝对值的操作仅在第 q 至第 $q+1$ 位上进行. 这样,可将多个像素的 2bit 表示存储在同一个机器字中,以并行计算多个像素的匹配误差.

Step 7. 输出初始迭代点,算法结束.

3 低频能量比率和黄金分割法的步长预测

3.1 步长对高斯-牛顿迭代结果的影响

文献[52]提出的高斯-牛顿法指出了参数的更新方向,但是并未给出其准确的迭代步长.实际上,文献[46-50]在求解弹性运动模型的每次迭代中均将步长设置为 1.而文献[53]却指出:目标函数偏离线性的距离越远(也就是泰勒展开式的余项越大),高斯-牛顿法的收敛速度就越慢,此时,宜选用更小的迭代步长.如图 3 所示:沿着 $f'(x_0)$ 方向逼近 $f_1(x)$ 时,可选取较大步长;而逼近 $f_2(x)$ 时,则需用尽可能小的步长.

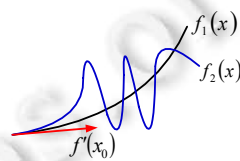
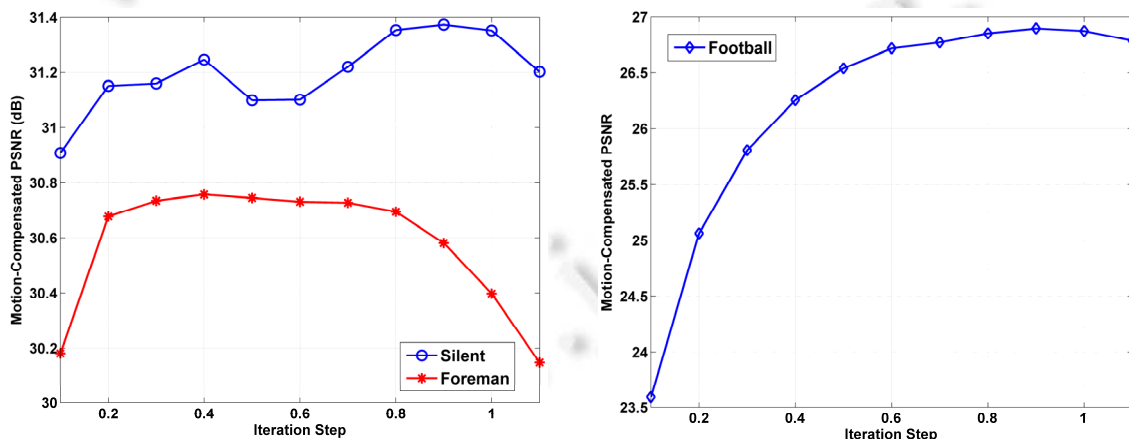


Fig.3 Diagram of 1-degree linear approximation

图 3 一阶线性逼近示意图

为了证明这一点,利用传统弹性运动估计算法对 Football,Silent 和 Foreman 序列的前 90 帧进行运动估计,并将迭代步长分别设置成 0.1~1.1,其运动补偿的平均 PSNR 如图 4 所示.从中可以清楚地看到:不同的迭代步长对运动估计/补偿性能有明显的影响,而且不同序列的最佳迭代步长不尽相同.Football,Silent 和 Foreman 的最佳迭代步长分别为 0.9,0.9 和 0.4,均不等于默认的迭代步长 1.所以,文献[46-50]将所有视频序列的步长简单设置成 1 不尽合理.



(a) Silent 和 Foreman 序列的运动估计/补偿结果

(b) Football 序列的运动估计/补偿结果

Fig.4 Influence of iteration step on elastic motion estimation

图 4 迭代步长对弹性运动估计性能的影响

3.2 基于DCT低频能量比率与黄金分割的步长预测

为了确定最佳的迭代步长,本文进一步引入带阻尼的高斯-牛顿法,采用步长因子 $\alpha(0<\alpha\leq 1)$ 来调节参数更新速度:

$$v_{k+1}=v_k+\alpha d=v_k+\alpha H^{-1}b \tag{15}$$

此时,弹性模型的求解效率主要依赖于 α 的选取:视频场景和运动越复杂,匹配误差函数偏离其线性预测值越远, α 就应越小;反之, α 则应越大.

图 5 给出了弹性运动模型的前 4 个基函数^[45],用来控制一个大小为 4×4 像素的块沿着水平方向的运动,其中,每个箭头表示一个像素的运动模式.

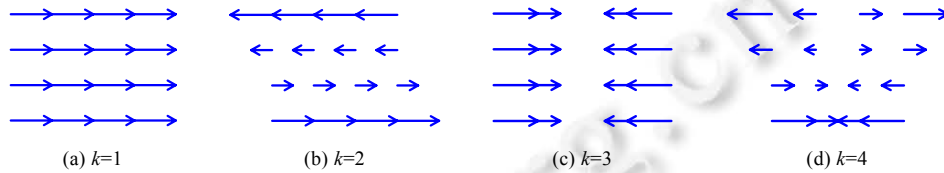


Fig.5 First four basis functions of elastic motion model

图 5 弹性运动模型的前 4 个基函数

可见,弹性运动估计/补偿的过程实际上是对参考帧进行离散余弦扭曲.若参考帧包含与上述余弦基函数相同频率的信号主分量,即包含能够被基函数高效表示的几何正则性,则参考帧经过离散余弦扭曲后,其像素值的变化量往往不会发生剧烈变化.直观地讲,当图 6 中的宏块沿水平方向运动时,其像素值不会发生变化,原因在于其运动与宏块本身的几何正则性一致;反之,若参考帧不包含与上述余弦基函数频率近似的主分量,扭曲运动后的像素值则可能发生明显的变化,进而偏离像素的线性预测值(如图 3 中的 $f_2(x)$),导致参数更新过快,收敛到局部最优解.故此,步长因子 α 的取值应考虑视频内容的几何正则性与弹性运动模型基函数的相关性.而对比弹性运动模型基函数与二维离散余弦变换基函数易知,前者就是后者的前 4 个基函数,于是,通过视频帧的余弦变换频谱,即可判断其几何特征与弹性运动模型的吻合程度.

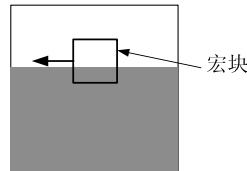


Fig.6 Diagram of motion direction consistent with principal component direction of signal

图 6 运动方向与信号主分量方向相同示意图

依据上述分析,本文采用二维离散余弦频谱的前 4 个低频系数占全部系数的能量比率作为步长因子 α 的上限 $\bar{\alpha}$,即

$$\bar{\alpha} = \frac{\sum_{i=0}^1 \sum_{j=0}^1 F^2(i, j)}{\sum_{i=0}^B \sum_{j=0}^B F^2(i, j)} \tag{16}$$

其中, F 表示某宏块的二维离散余弦变换系数集合, B 表示宏块的高度(或宽度). $\bar{\alpha}$ 越大,表明与弹性模型基函数同频率的信号分量越主要,也就是视频中包含的能够被基函数高效表示的几何正则性越强,此时,选取较大的迭代步长不会导致弹性模型收敛到局部最优解;反之, $\bar{\alpha}$ 越小,表明视频内容的几何正则性较为复杂,匹配误差曲面可能存在多峰性,不适宜用大的迭代步长.

由于小步长不会导致局部收敛,最多可使得收敛速度过慢.于是,本文将步长下限 $\underline{\alpha}$ 设置成 0.05,这样就得到了步长因子的取值区间 $\alpha \in [\underline{\alpha}, \bar{\alpha}]$.然而,仅有该区间还不足以得到准确的步长.考虑到第 2 节的迭代点预测方法已经将初始向量放入最佳向量的单调区间(如图 7 所示),本文进一步采用黄金分割法来完成对最佳步长 α^* 的搜索,其基本思想是根据 $[\underline{\alpha}, \bar{\alpha}]$ 区间中匹配误差的单峰性,反复按照黄金分割的比例将 $[\underline{\alpha}, \bar{\alpha}]$ 的长度缩小,进而从区间两端不断逼近 α^* .具体过程如下.

Step 1. 令 $i=0, \underline{\alpha}_0 = \underline{\alpha}, \bar{\alpha}_0 = \bar{\alpha}, \lambda = 0.618$.

Step 2. 计算:

$$\begin{aligned} \underline{\alpha}' &= \underline{\alpha}_i + (1-\lambda)(\bar{\alpha}_i - \underline{\alpha}_i), \\ \bar{\alpha}' &= \underline{\alpha}_i + \lambda(\bar{\alpha}_i - \underline{\alpha}_i). \end{aligned}$$

Step 3. 若匹配误差 $e(v; \underline{\alpha}') < e(v; \bar{\alpha}')$, 则令 $\underline{\alpha}_{i+1} = \underline{\alpha}_i, \bar{\alpha}_{i+1} = \bar{\alpha}'$; 否则,令 $\underline{\alpha}_{i+1} = \underline{\alpha}', \bar{\alpha}_{i+1} = \bar{\alpha}_i$.

Step 4. $i=i+1$. 若 $i < It$, 则转入 Step 2; 否则,令 $\alpha^* = (\underline{\alpha}_i + \bar{\alpha}_i)/2$, 算法结束.

为直观起见,图 7 给出了第 1 轮迭代的示意图.此外,Step 4 中 It 表示预设的迭代次数.由于迭代次数与计算量直接相关,为了在计算量和预测性能之间取得平衡,本文将其设置为 3,此时 $[\underline{\alpha}, \bar{\alpha}]$ 区间的长度已缩小 90%,能够满足模型求解的需要.

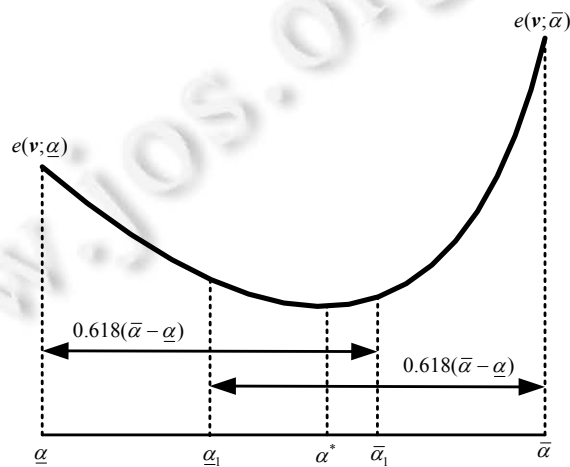


Fig.7 Diagram of golden ratio search

图 7 黄金分割搜索的示意图

4 改进高斯-牛顿法的弹性运动估计步骤

在第 2 节和第 3 节算法的基础上,本节提出一种基于改进高斯-牛顿法的运动估计方法.对于每个图像组 (group of picture, 简称 GOP), 其运动估计/补偿的详细步骤如下:

Step 1. 根据公式(16),计算当前 GOP 第 1 帧中所有宏块的步长因子上限,并求出其平均值 $\bar{\alpha}$. 考虑到同一 GOP 的场景内容相差不大,该 GOP 的每个宏块都选用 $\bar{\alpha}$ 作为其步长上限,进而保证较低的计算量.

Step 2. 若 GOP 的所有宏块都已处理完毕,则算法结束;否则,选取当前 GOP 中未处理过的一个宏块,利用第 2.2 节的初始迭代点预测算法计算其平移分量 \mathbf{v}_i 和 $\mathbf{v}_{p/2+1}$,并将迭代计数器 i 初始化为 1.

Step 3. 利用第 3.2 节的黄金分割搜索算法求解最优迭代步长因子 α^* .

Step 4. 根据公式(10)、公式(15)和 α^* 更新弹性运动参数 \mathbf{v}_i ,并将其代入公式(2)和公式(3),建立当前宏块中每个像素与其匹配像素的坐标映射.

Step 5. 根据像素坐标映射关系,利用双线性插值计算匹配像素的值,从而得到运动补偿误差.

Step 6. 令 $i=i+1$.如果 $i > Th$ (Th 表示预设的迭代次数阈值),那么当前宏块处理完毕,转入 Step 2; 否则,转入

Step 4.

5 实验结果与分析

为了验证本文所提出算法的性能,分别以 CIF(common intermediate format)@30 帧/秒格式的 Football, Coastguard, Bus, Soccer, Tennis, Flower, Tempete, Paris, Mobile, Stefan, News, Sign Irene, Husky, Pamphlet 和 4CIF@30 帧/秒格式的 Crew, City, Harbour 等序列的 1~90 帧为例进行了大量实验,并将结果与传统的弹性运动估计算法^[45]、基于块平移模型的全搜索进行比较.同时,为了考察本文算法的每个预测阶段的作用,还将对比分析仅结合了初始点预测的弹性运动估计性能.

实验参数设置如下:搜索窗口为 33×33 像素,宏块尺寸为 16×16 像素, $l_t=3$, $l_h=15$ (这与文献[45]的设置相同).重构帧的质量采用 PSNR 进行评价.

5.1 运动估计/补偿质量的比较

表 1 罗列了各个测试序列的亮度分量采用不同运动估计算法所得到的平均 PSNR.

Table 1 Motion-Compensated PSNR comparison
表 1 运动补偿的 PSNR 比较

视频序列名称	格式	块匹配全搜索算法	传统弹性运动估计	初始点预测+弹性运动估计	本文算法
Football	352×288@30fps	27.46	26.88	29.27	29.64
Coastguard	352×288@30fps	29.71	31.48	31.87	32.02
Bus	352×288@30fps	24.87	21.85	25.79	25.92
Soccer	352×288@30fps	29.61	29.91	31.44	31.74
Tennis	352×288@30fps	28.91	29.30	29.96	30.48
Flower	352×288@30fps	26.03	26.13	25.97	26.42
Tempete	352×288@30fps	27.75	29.43	29.62	30.56
Paris	352×288@30fps	30.70	32.61	32.90	33.56
Mobile	352×288@30fps	23.96	26.91	26.52	27.74
Stefan	352×288@30fps	25.90	26.15	29.06	29.16
News	352×288@30fps	37.83	37.65	37.54	37.82
Sign irene	352×288@30fps	33.56	35.22	35.74	36.06
Husky	352×288@30fps	19.10	20.26	20.59	21.20
Pamphlet	352×288@30fps	43.12	36.14	36.99	43.41
Crew	704×576@30fps	32.36	33.17	33.44	33.50
City	704×576@30fps	30.86	30.24	31.84	32.13
Harbour	704×576@30fps	28.10	30.78	30.94	30.99

由表 1 可知:

- 基于块平移模型的全搜索算法性能最为逊色,平均 PSNR 比本文算法低 1.73dB.
- 其次是传统弹性运动估计,平均 PSNR 比本文算法低 1.42dB.
- 传统弹性运动估计与初始迭代点预测结合后,其平均 PSNR 提高了 0.97dB,表明初始迭代点预测算法能够有效地将搜索起点放入最优向量的单调区间,进而改善传统弹性模型的收敛性.
- 在此基础上,本文提出的改进弹性运动估计算法的平均 PSNR 又比初始点预测+传统弹性运动估计算法提高了 0.46dB,这说明本文优化控制迭代步长的策略是有效的.

图 8(a)~图 8(d)给出了 Bus, Football, Stefan 和 News 这 4 个序列的 PSNR 逐帧比较情况.其中, Bus, Football 和 Stefan 序列都包含前景物体的快速运动,又分别有摄像机的拉摄运动、快速摇摄运动和慢速摇摄运动; News 序列则同时包含了播音员的局部慢速运动和背景物体的局部中、快速旋转运动.

从图中可见,尽管 4 种序列的运动复杂,可是本文算法的性能均较为稳定,取得了最高的运动补偿 PSNR(少数帧除外).

需要说明的是,对于 News 序列,本文算法的 PSNR 比块匹配全搜索算法低 0.01dB;然而后者的运动估计性能非常不稳定(如图 8(d)所示),容易造成码流的频繁波动,不利于视频主客观质量的控制.这表明,弹性运动模型能够比平移模型更加有效地表示物体和摄像机的运动.

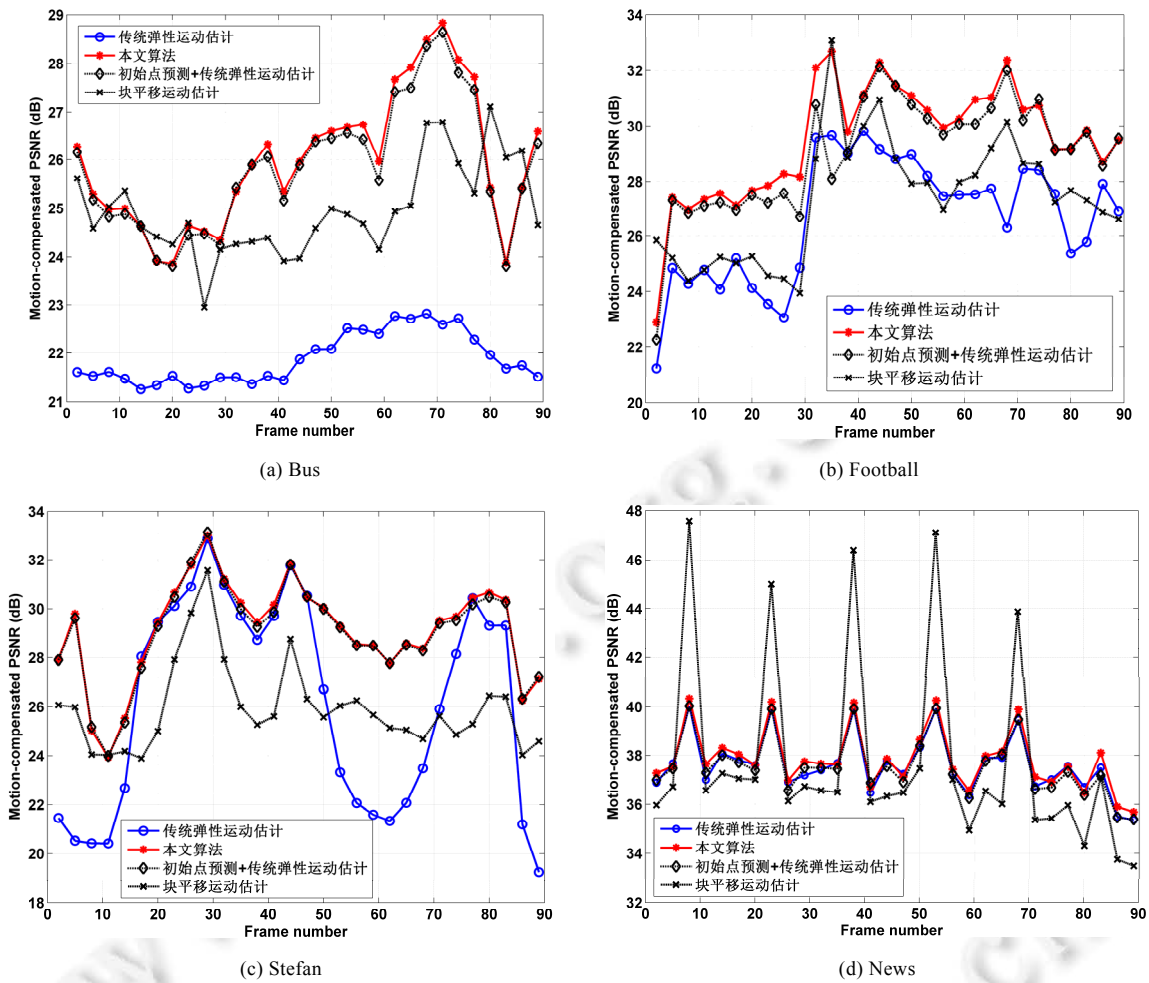


Fig.8 Frame-Wise PSNR comparison of Bus, Football, Stefan, and News sequences

图 8 Bus, Football, Stefan 和 News 序列的逐帧 PSNR 比较

5.2 收敛性和收敛速度比较

收敛性是衡量弹性运动估计效率的重要指标之一.图 9 所示为 Football, Mobile & Calendar 和 Tempete 序列分别采用传统高斯-牛顿法与本文算法求解弹性运动模型的前 14 次迭代结果.

从图 9 中可以看出:

- 一方面,由于匹配误差曲面的复杂性,传统高斯-牛顿法的收敛很不稳定,在有些情况下能够趋于全局最优解(如 Football),但是在有些情况下,会由于迭代步长过大而出现运动估计/补偿 PSNR 的下降(如 Mobile 和 Tempete),无法收敛到全局最优解.而本文算法的补偿质量则随着迭代次数的增加而稳定提高.
- 另一方面,本文算法的收敛速度优于传统高斯-牛顿法.对于 Football 序列,本文算法经过 8 次迭代后获得的补偿 PSNR 与最优解仅相差不到 1%,传统高斯-牛顿法达到同样的性能误差则需 12 次迭代.而对于 Mobile 和 Tempete 序列,本文算法只需 4 次迭代即可.分析其原因在于:本文初始迭代点和步长因子的选取符合各个测试视频序列的匹配误差曲面分布特性,能够高效地逼近其最优运动参数.

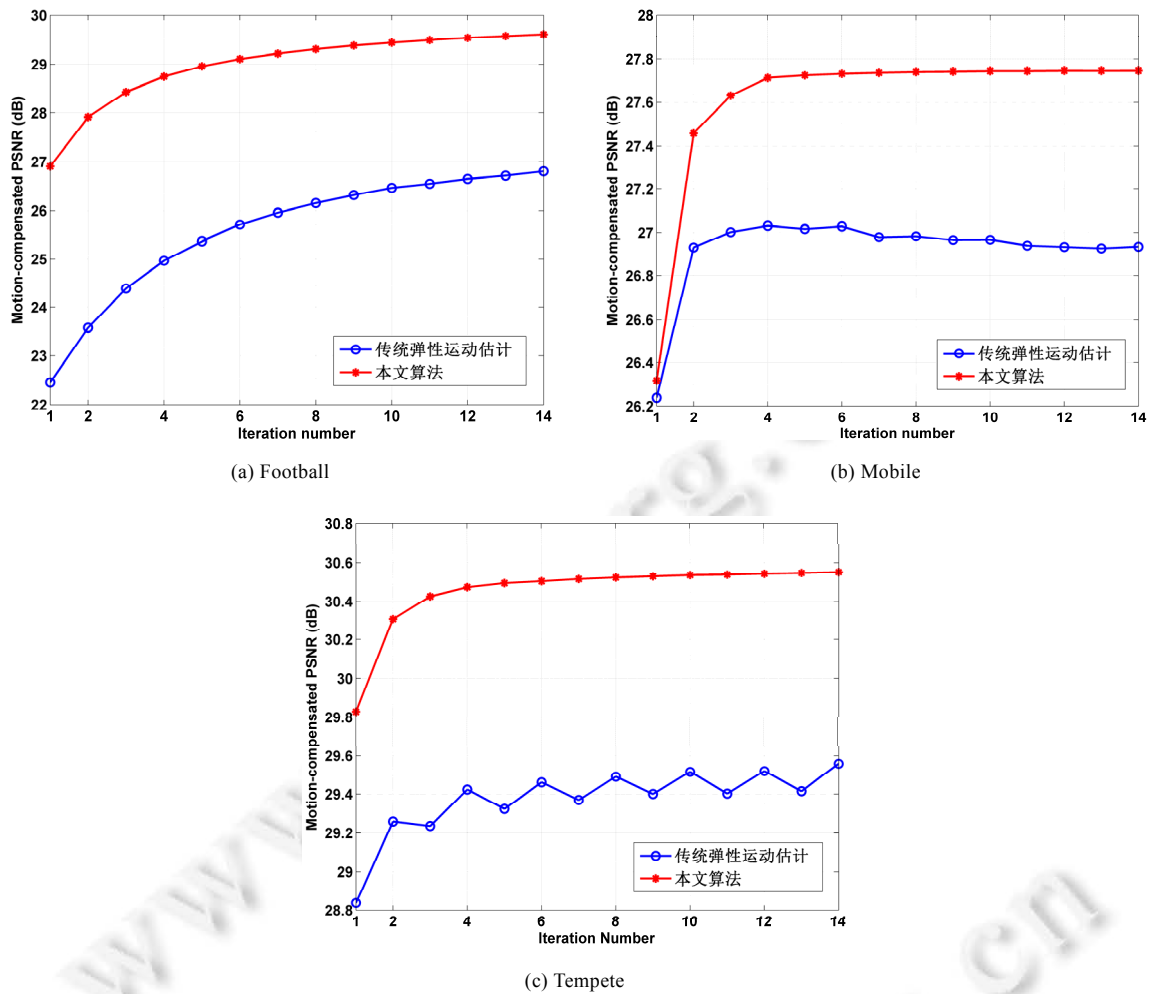


Fig.9 Comparison between traditional Gauss-Newton method and our method in terms of convergence speed

图9 传统高斯-牛顿算法与本文算法的收敛速度比较

5.3 计算复杂度分析

计算复杂度是衡量运动估计效率的重要指标,下面讨论4种运动估计算法处理一个宏块的计算量.

- 首先,设宏块尺寸为 $B \times B$ 像素、搜索窗口为 $W \times W$ 像素,则基于块平移模型的全搜索的计算复杂度为 $O(B^2W^2)$.
- 其次,用传统高斯-牛顿法求解弹性运动模型的每次迭代包括计算偏导数、黑塞矩阵、逆矩阵、矩阵乘法、双线性插值和运动补偿误差,其计算复杂度分别为 $O(pB^2)$, $O(p^2B^2)$, $O(p^3)$, $O(pB^2)$, $O(8B^2)$ 和 $O(B^2)$,共 Th 次迭代,所以总复杂度为 $O(Th \cdot p^2B^2)$.
- 最后,本文算法比传统弹性模型的求解增加了初始点预测、低频能量比率和黄金分割搜索的计算.其中,初始点预测阶段的计算量至多为 $O(49B^2)$;由于离散余弦变换是视频编码的必备环节,变换不需引入额外的计算量,而且每个 GOP 仅需计算1帧的低频能量比率,故此该阶段的计算复杂度为 $O(B^2/16)$;黄金分割搜索阶段需计算3次双线性插值和运动补偿误差,其复杂度为 $O(24B^2)$.

综上,在本文的参数设置下,即 $W=33, Th=15, p=8$,传统弹性运动模型解法的计算复杂度比全搜索提高约22%,本文算法的计算复杂度至多比传统弹性运动模型解法提高约5%.而若在初始点预测阶段采用并行实现,

本文算法的复杂度至多比传统弹性运动模型解法提高约 3%。

需要特别指出的是,通过第 5.2 节的分析可知:本文算法的收敛速度明显优于传统高斯-牛顿法,一般仅需 1~3 次迭代就能取得高于传统弹性运动估计和块平移全搜索的 PSNR。故此,从这个意义上讲,本文算法的计算量是传统弹性运动估计的 8.33%~21.00%,是块平移全搜索的 8.54%~25.62%。

6 结 论

弹性运动估计比块平移运动估计能够更有效地表示物体的局部非刚性复合运动,但是计算效率尚待改善。本文通过理论和实验分析,讨论了初始迭代点和迭代步长的选取对高斯-牛顿法求解结果有明显影响,从而提出利用像素的 2bit 深度表示和均匀搜索模板实现初始迭代点的预测,利用离散余弦变换的低频系数能量比率估计迭代步长的上限,再采用黄金分割法计算合理的步长。与典型的高斯-牛顿法相比,本文算法为弹性运动估计的求解以及视频处理中类似的无约束优化问题的求解提供了新思路。

另外,该方法仍存在若干问题有待解决,如减少黑塞矩阵和逆矩阵的反复计算、迭代步长的快速准确求解等,我们将在今后的工作中进一步深入研究这两个问题的解决方案。

References:

- [1] Sullivan GJ, Ohm JR, Ortega A, Delp E, Vetro A, Barni M. Future of video coding and transmission. *IEEE Signal Processing Magazine*, 2006,23(6):76–82. [doi: 10.1109/SP-M.2006.248722]
- [2] Gao W, Zhao DB, Ma SW. *Principles of Digital Video Coding Technology*. Beijing: Science Press, 2010 (in Chinese).
- [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] 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]
- [6] Cheung CH, Po LM. Novel cross-diamond-hexagonal search algorithms for fast block motion estimation. *IEEE Trans. on Multimedia*, 2005,7(1):16–22. [doi: 10.1109/TMM.2004.840609]
- [7] Chen ZB, Xu JF, He Y, Zheng JL. Fast integer-pel and fractional-pel motion estimation for H.264/AVC. *Journal of Visual Communication and Image Representation*, 2006,17(2):264–290. [doi: 10.1016/j.jvcir.2004.12.002]
- [8] Lou CC, Lee SW, Jay Kuo CC. Adaptive motion search range prediction for video encoding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2010,20(12):1903–1908. [doi: 10.1109/TCSVT.2010.2087551]
- [9] 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]
- [10] 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]
- [11] Ismail Y, McNeely JB, Shaaban M, Mahmoud H, Bayoumi MA. Fast motion estimation system using dynamic models for H264/AVC video coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2012,22(1):28–42. [doi: 10.1109/TCSVT.2011.2148450]
- [12] Cheung CK, Po LM. Normalized partial distortion search algorithm for block motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2000,10(3):417–422. [doi: 10.1109/76.836286]
- [13] 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). [doi: 10.3321/j.issn:0254-4164.2005.10.019]

- [14] Zhang R, Comer ML. Rate distortion performance of pyramid and subband motion compensation based on quantization theory. *IEEE Trans. on Circuits and Systems for Video Technology*, 2010,20(12):1876–1881. [doi: 10.1109/TCSVT.2010.2077478]
- [15] Bao XN, Zhou DJ, Liu PL, Goto S. An advanced hierarchical motion estimation scheme with lossless frame recompression and early-level termination for beyond high-definition video coding. *IEEE Trans. on Multimedia*, 2012,14(2):237–249. [doi: 10.1109/TMM.2011.2171677]
- [16] 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]
- [17] Montrucchio B, Quaglia D. New sorting-based lossless motion estimation algorithms and a partial distortion elimination performance analysis. *IEEE Trans. on Circuits and Systems for Video Technology*, 2005,15(2):210–220. [doi: 10.1109/TCSVT.2004.841689]
- [18] Gharavi H, Mills M. Blockmatching motion estimation algorithms-new results. *IEEE Trans. on Circuits and Systems*, 1990,37(5):649–651. [doi: 10.1109/31.55010]
- [19] Chen MJ, Chen LG, Chiueh TD, Lee YP. A new block-matching criterion for motion estimation and its implementation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1995,5(3):231–236. [doi: 10.1109/76.401100]
- [20] Xiong B, Zhu C. A new multiplication-free block matching criterion. *IEEE Trans. on Circuits and Systems for Video Technology*, 2008,18(10):1441–1446. [doi: 10.1109/TCSVT.2008.2002824]
- [21] Tu CJ, Tran TD, Prince JL, Topiwala P. Projection-Based block matching motion estimation. In: Tescher AG, ed. *Proc. of the SPIE 4115, Applications of Digital Image Processing XXIII*. Bellingham: SPIE Press, 2000. 1–10. [doi: 10.1117/12.411605]
- [22] Ertürk A, Ertürk S. Two-Bit transform for binary block motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2005,15(7):938–946. [doi: 10.1109/TCSVT.2005.848340]
- [23] 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]
- [24] Lee H, Jeong J. Early termination scheme for binary motion estimation. *IEEE Trans. on Consumer Electronics*, 2007,53(4):1682–1686. [doi: 10.1109/TCE.2007.4429270]
- [25] Çelebi A, Akbulut O, Urhan O, Hamzaoglu İ, Ertürk S. An all binary sub-pixel motion estimation approach and its hardware architecture. *IEEE Trans. on Consumer Electronics*, 2008,54(4):1928–1937. [doi: 10.1109/TCE.2008.4711255]
- [26] Song CM, Guo YW, Wang XH, Liu D. Motion estimation algorithm using 2 bit-depth pixel and fuzzy quantization. *Journal on Communications*, 2013,34(7):59–70 (in Chinese with English abstract). [doi: 10.3969/j.issn.1000-436x.2013.07.007]
- [27] 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]
- [28] 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]
- [29] Nakaya Y, Harashima H. Motion compensation based on spatial transformations. *IEEE Trans. on Circuits and System for Video Technology*, 1994,4(3):339–356. [doi: 10.1109/76.305878]
- [30] 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]
- [31] 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). [doi: 10.3321/j.issn:0372-2112.2000.05.013]
- [32] 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]
- [33] 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]

- [34] 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]
- [35] 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]
- [36] 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]
- [37] 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]
- [38] 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).
- [39] Narroschke M, Swoboda R. Extending HEVC by an affine motion model. In: Ortega A, Said A, eds. *Proc. of the Picture Coding Symp.* Piscataway: IEEE Computer Society Press, 2013. 321–324. [doi: 10.1109/PCS.2013.6737748]
- [40] Ren L. The application of deformable block matching algorithm [MS. Thesis]. Tianjin: Tianjin University, 2007 (in Chinese with English abstract).
- [41] Zheng JL, Qin TF, Ni GG. 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). [doi: 10.11834/jig.20110816]
- [42] 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]
- [43] 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]
- [44] Kybic J, Unser M. Fast parametric elastic image registration. *IEEE Trans. on Image Processing*, 2003,12(11):1427–1442. [doi: 10.1109/TIP.2003.813139]
- [45] 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]
- [46] 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]
- [47] Deng BY. Study on video compression based on elastic motion model [MS. Thesis]. Xiamen: Xiamen University, 2014 (in Chinese with English abstract).
- [48] Haque MN, Biswas M, Pickering MR, Frater MR. A low complexity algorithm for global motion parameter estimation targeting hardware implementation. In: Zhang YC, Bottema MJ, eds. *Proc. of the Digital Image Computing: Techniques and Applications.* Piscataway: IEEE Computer Society Press, 2009. 1–5. [doi: 10.1109/DICTA.2009.11]
- [49] Haque MN, Biswas M, Pickering MR, Frater MR. An adaptive low-complexity global motion estimation algorithm. In: Saito T, Asai K, Aizawa K, Yoshida T, eds. *Proc. of the Picture Coding Symp.* Piscataway: IEEE Computer Society Press, 2010. 598–601. [doi: 10.1109/PCS.2010.5702574]
- [50] Haque MN, Biswas M, Pickering MR, Frater MR. A low-complexity image registration algorithm for global motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2012,22(3):426–433. [doi: 10.1109/TCSVT.2011.2163983]
- [51] He JY. *Optimization Method.* Beijing: Tsinghua University Press, 2007 (in Chinese).
- [52] Baker S, Matthews I. Lucas-Kanade 20 years on: A unifying framework. *Int'l Journal of Computer Vision*, 2004,56(3):221–255. [doi: 10.1023/B:VISI.0000011205.11775.fd]
- [53] Gao L. *Numerical Optimization Method.* Beijing: Peking University Press, 2014 (in Chinese).
- [54] Song CM, Wang XH, Zhang FY. Binary alpha-plane assisted fast motion estimation scheme of video object. *Ruan Jian Xue Bao/ Journal of Software*, 2008,19(4):829–841 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/19/829.htm> [doi: 10.3724/SP.J.1001.2008.00829]
- [55] Song CM, Tan R, Wang XH, Liu D. Wavelet-Domain motion estimation with initial search point prediction using low bit-depth pixels. *Journal of Computer Research and Development*, 2013,50(4):776–786 (in Chinese with English abstract).

- [56] 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:

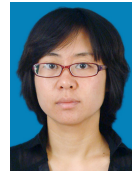
- [2] 高文,赵德斌,马思伟.数字视频编码技术原理.北京:科学出版社,2010.
- [13] 宋传鸣,王相海.小波域视频运动估计研究进展.计算机学报,2005,28(10):1716–1727. [doi: 10.3321/j.issn:0254-4164.2005.10.019]
- [26] 宋传鸣,郭延文,王相海,刘丹.基于模糊量化和 2bit 深度像素的运动估计算法.通信学报,2013,34(7):59–70. [doi: 10.3969/j.issn.1000-436x.2013.07.007]
- [31] 吴枫,高鹏,高文.基于网格模型的运动估计技术.电子学报,2000,28(5):47–51. [doi: 10.3321/j.issn:0372-2112.2000.05.013]
- [38] 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.
- [40] 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.
- [41] 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. [doi: 10.11834/jig.20110816]
- [47] 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.
- [51] 何坚勇.最优化方法.北京:清华大学出版社,2007.
- [53] 高立.数值最优化方法.北京:北京大学出版社,2014.
- [54] 宋传鸣,王相海,张福炎.二值 α 平面辅助的视频对象快速运动估计算法.软件学报,2008,19(4):829–841. <http://www.jos.org.cn/1000-9825/19/829.htm> [doi: 10.3724/SP.J.1001.2008.00829]
- [55] 宋传鸣,谭茹,王相海,刘丹.结合低位深像素预测起点的小波域运动估计.计算机研究与发展,2013,50(4):776–786.



宋传鸣(1980—),男,辽宁沈阳人,博士,副教授,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.



赵长伟(1989—),男,硕士,主要研究领域为视频运动估计.



刘丹(1981—),女,博士生,讲师,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.



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.