

## 采用改进 Levenberg-Marquardt 法的快速弹性运动估计\*

宋传鸣<sup>1,2</sup>, 闵新<sup>1</sup>, 闫小红<sup>1</sup>, 王相海<sup>1</sup>, 尹宝才<sup>2</sup>



<sup>1</sup>(辽宁师范大学 计算机与信息技术学院, 辽宁 大连 116029)

<sup>2</sup>(大连理工大学 计算机科学与技术学院, 辽宁 大连 116024)

通讯作者: 宋传鸣, E-mail: chmsong@163.com

**摘要:** 弹性运动估计是近年来出现的一种有效的时域视频预测编码技术,但其基于高斯-牛顿法的优化求解仍存在计算量高、收敛不稳定的问题.为此提出一种基于改进 Levenberg-Marquardt(L-M)法的弹性运动估计算法.首先,根据弹性基函数和黑塞矩阵的数值对称性,给出了 L-M 黑塞矩阵的快速计算方法,将其计算量降低了 62.5%.其次,通过理论和实验分析发现,L-M 对角矩阵阻尼系数的更新因子对弹性运动估计性能有明显影响,进而采用最近 2 次迭代的搜索步长的平方商自适应地确定更新因子,并对该阻尼系数进行正、负交替更新.实验结果表明,对于具有不同空间分辨率和场景特点的视频序列,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之基于块平移模型的全搜索和基于改进高斯-牛顿法的弹性运动估计分别提高 2.54dB、1.77dB.并且,所提算法收敛速度快,一般只需 1~2 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;Levenberg-Marquardt 优化法;自适应阻尼系数

**中图分类号:** TP391

中文引用格式: 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>

英文引用格式: Song CM, Min X, Yan XH, Wang XH, Yin BC. Fast elastic motion estimation using improved Levenberg-Marquardt method. Ruan Jian Xue Bao/Journal of Software, 2019,30(7):2208-2226 (in Chinese). <http://www.jos.org.cn/1000-9825/5487.htm>

## Fast Elastic Motion Estimation Using Improved Levenberg-Marquardt Method

SONG Chuan-Ming<sup>1,2</sup>, MIN Xin<sup>1</sup>, YAN Xiao-Hong<sup>1</sup>, WANG Xiang-Hai<sup>1</sup>, YIN Bao-Cai<sup>2</sup>

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

<sup>2</sup>(School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China)

**Abstract:** Elastic motion estimation is an effective temporal predictive coding technique of video proposed in recent years. But its optimization solution based on Gauss-Newton method still exhibits the problem of high computational complexity and unstable convergence yet. Thus an elastic motion estimation algorithm is addressed based on an improved Levenberg-Marquardt (L-M) method. First, a fast implementation of the L-M Hessian matrix is designed according to the numerical symmetry of elastic basis function and the Hessian matrix, which reduces its computational complexity by 62.5%. Second, it is found that the update factor of L-M diagonal matrix's damping coefficient has obvious influence on the performance of elastic motion estimation through theoretical and experimental analyses. The squared ratio of the step size in the latest two iterations is used to adaptively determine the update factor, by which the damping

\* 基金项目: 国家自然科学基金(61402214, 41671439, 61632006); 大连市青年科技之星项目支持计划(2015R069); 辽宁省自然科学基金(20180550570); 南京大学计算机软件新技术国家重点实验室开放课题基金(KFKT2018B07)

Foundation item: National Natural Science Foundation of China (61402214, 41671439, 61632006); Dalian Foundation for Youth Science and Technology Star (2015R069); Natural Science Foundation of Liaoning Province of China (20180550570); Open Foundation of State Key Laboratory for Novel Software Technology (Nanjing University) (KFKT2018B07)

收稿时间: 2017-08-05; 修改时间: 2017-09-29; 采用时间: 2017-11-08; jos 在线出版时间: 2019-01-21

CNKI 网络优先出版: 2019-01-22 13:48:54, <http://kns.cnki.net/kcms/detail/11.2560.TP.20190122.1348.007.html>

coefficient is updated positively and negatively in turn. Experimental results show that the proposed algorithm is able to obtain stable performance for the video sequences with various spatial resolution and scene characteristics. It gains 2.54 dB and 1.77 dB higher average motion-compensated peak signal-to-noise ratio (PSNR) than those of the full search based on block-wise translational model and the elastic motion estimation based on modified Gauss-Newton method, respectively. Furthermore, the proposed algorithm converges fast. Only 1~2 iterations are needed before it achieves higher PSNR than the conventional elastic motion estimation and the block-wise translational full search.

**Key words:** video coding; motion estimation; elastic model; Levenberg-Marquardt optimization method; adaptive damping coefficient

运动估计是一项高效率的去除视频时间域冗余的预测技术,为 H.26x、MPEG 和 AVS 等系列编码标准贡献了大部分的性能提升<sup>[1,2]</sup>。然而,多项研究结果表明<sup>[1,3,4]</sup>,仅整数像素精度的运动估计所耗费的计算资源就占编码器全部资源的 40%~80%;若将分数像素精度的运动估计也考虑在内,其运算代价无疑更高。因此,为了在计算复杂度、硬件实现难度和预测精度之间进行折中,现有视频编码标准多年来始终采用了仅能刻画水平、竖直等平移运动的块匹配算法,并陆续提出了多种快速运动估计策略,大致可分为以下 7 类。

(1) 基于候选向量下采样的策略<sup>[5-13]</sup>:这类策略的基本思想是只计算搜索窗口中一部分候选向量的匹配误差,从中找出误差最小的向量作为最优运动向量,如菱形搜索、非对称十字多层次六边形格点搜索 UMHexagonS、测试域搜索 TZSearch 和抛物线搜索等。

(2) 基于像素下采样的策略<sup>[14-19]</sup>:这类策略的基本思想是只计算当前宏块中一部分像素的匹配误差,并将其作为该宏块的误差用于最优运动向量的判断,如多分辨率搜索、梅花形下采样搜索和自适应像素下采样搜索等。

(3) 基于像素预排序的策略<sup>[20-23]</sup>:这类算法的主要思路是在计算某个候选向量的匹配误差时,优先对当前宏块的较大匹配误差进行累加,当累计误差和大于已知的最小误差时就提前中止该候选向量的判断,从而达到降低计算量的目的,如部分失真搜索等。

(4) 基于低复杂度匹配函数的策略<sup>[24-28]</sup>:这类算法的主要思想是采用较低复杂度的匹配误差函数替代传统的均方误差函数,进而降低误差匹配过程的计算量,常用函数包括像素误差分类(pel difference classification, 简称 PDC)、相异像素数目(different pixel count, 简称 DPC)、改进的异或函数等。

(5) 基于低比特深度像素的策略<sup>[29-34]</sup>:这类策略的基本思路是通过某种映射方法将 8bit 深度的像素变换为低位深的像素,以此来减少运动估计的计算开销,如 1bit 搜索和 2bit 搜索等,常与第(4)类策略结合使用。

(6) 基于哈希映射的策略<sup>[35,36]</sup>:这类算法主要通过哈希函数将待匹配块映射为一个哈希值,从而借助哈希表来提高较大搜索窗口下的块匹配效率。

(7) 基于块分类的策略<sup>[1,37]</sup>:该种策略的核心思想是按照速率失真准则将待匹配块分成不同运动幅度的类,并为不同类别的待匹配块选取恰当的运动估计算法完成预测,从而在保证预测效率的情况下尽量降低计算量。

一方面,尽管众多研究人员对基于平移模型的块匹配运动估计算法进行了上述改进,但仍未从根本上解决运动估计环节计算负载过高的问题。另一方面,块平移模型既无法有效预测由物体旋转、缩放、变形和摄像机运动产生的非刚性复合运动,又不能准确表示具有复杂形状的运动区域,导致在运动物体边缘产生大幅值的预测残差,影响后续的熵编码效率。为此,H.264/AVC(advanced video coding)和 H.265/HEVC(high efficiency video coding)等新一代编码标准将早期的正方形宏块改进为对称或非对称的精细矩形块,并进一步采用复杂的分层可变尺寸块结构、分数像素运动向量、多参考帧和广义 B 帧等多种优化手段来逼近复杂运动场和运动物体。然而,文献[4]通过实验统计发现,随着块尺寸的减小和运动向量精度的提高,用于编码运动向量、块划分方式的码流开销和各种软/硬件计算开销也逐渐增加,尤其是当矩形块尺寸减小至 4×4 像素、向量精度达到 1/16 像素时,软/硬件开销的增加幅度甚至超过了率失真性能的提升幅度。这个结论说明,仅仅依靠块平移模型来实现运动估计愈来愈无法很好地满足高清/超高清视频、面向视频通信的桌面视频和面向虚拟现实应用的全景视频等视频编码的需求<sup>[38,39]</sup>。因此,越来越多的学者认为,研究能有效表示复杂运动场的高阶运动模型及其参数求解策略是帧间运动估计的未来发展方向之一<sup>[40-42]</sup>,对于下一代视频编码效率的大幅提升有重要意义。

本文首先对比分析典型的高阶运动模型的优势和不足,进而介绍弹性运动模型的基本思想.然后,针对弹性运动估计仍然存在的计算量高、收敛不稳定的问题,引进 Levenberg-Marquardt 方法进行优化求解,并从黑塞矩阵(Hessian matrix)的快速计算和对角矩阵阻尼系数的自适应更新两方面做出改进,提出一种基于改进 Levenberg-Marquardt 法的视频弹性运动估计算法.实验结果验证了本文算法的有效性.

## 1 相关工作

鉴于平移模型的不足,研究人员将高阶运动模型引入到运动估计中,利用高阶函数产生 1 个或多个扭曲的参考帧,实现更高质量的运动补偿.依据模型显式表现形式的不同,本文将这些运动估计/补偿算法划分为 4 类.

(1) 基于网格模型的运动估计<sup>[43-49]</sup>:该类算法主要利用三角形网格或四边形网格来刻画视频图像中的运动物体或内容,能够实现较为可靠的运动/静止区域划分,并可缓解运动补偿帧的块效应,取得更高的主观质量.但是,文献[50]发现,基于网格模型的运动估计对于相邻网格共有节点的控制和对编码率失真的优化较为困难,尚缺少有效的优化方法,而且需要额外传输网格划分方式的同步信息.

(2) 基于多项式模型的运动估计<sup>[38,50-56]</sup>:这类算法的基本思路是采用 4-参数的一元一次缩放变换模型、6-参数的二元一次仿射变换模型、8-参数的二元二次投影变换模型、12-参数的二元二次变换模型及其混合模型产生全局扭曲的参考帧,能够比平移模型更加有效地捕获物体的平移、旋转、拉伸、仿射、透视和景深变化等丰富的运动形式.但是,随着模型参数的增多,其搜索复杂度明显提高,并且对局部运动的刻画能力不足.

(3) 基于缩放模型的运动估计<sup>[57-59]</sup>:考虑到多参数模型的运算复杂度,该类算法简化了基于多项式模型的运动估计,通过在平移模型基础上引进缩放因子来表示物体的拉伸运动和景深变化等全局运动,但却不能描述 3D 错切和由于摄像机的俯仰而产生的物体旋转以及物体的局部变形运动等.

(4) 基于弹性模型的运动估计<sup>[42,60-63]</sup>:该类算法采用离散余弦函数、小波基函数或样条函数刻画物体的平移、错切和扭曲运动,既能表示全局和局部运动,又可通过调整模型参数个数来控制运动估计的计算复杂度.文献[60]的实验结果表明,基于弹性模型的运动估计在相同码率下获得了比块平移运动估计高 0.7dB 的运动补偿峰值信噪比(peak signal-to-noise ratio,简称 PSNR);文献[61]的结论显示,弹性运动模型可将 H.265 的输出码率降低 3%~12%,而运动补偿失真仅损失 1%;文献[62,63]则通过优化弹性模型的求解方法取得了更加理想的预测效率,其平均 PSNR 比传统弹性运动估计提高了 1.42dB,并且比基于块平移模型的全搜索算法高出 1.73dB.

因此,综合各个模型之间的对比和现有研究结论可知,弹性运动模型是一种表示复杂运动场的高效率模型.目前,关于视频弹性运动估计的研究主要集中在以下两个方面.

(1) 弹性运动模型与视频编码标准的结合方式.文献[42]将弹性运动估计引进到 H.264 中,依据图像的几何特征,采用不同斜率的线段划分待预测块,获得其三角形或四边形网格表示,从而使弹性模型能够更准确地描述复杂形状的运动区域,更好地适应多样的局部运动.文献[39]将弹性运动估计作为 HEVC 的一种可选模式,通过已解码帧计算弹性运动场,再利用该运动场重建弹性变形后的参考帧,进而根据率失真准则在平移模型和弹性模型之间进行自适应的选择.但是,文献[39,42]简单地使用传统的高斯-牛顿法求解弹性运动向量,既未能避免弹性运动估计的高计算量,又无法避免搜索陷入局部最优,这会从根本上影响弹性运动估计的有效性和实用性.

(2) 弹性运动向量的优化求解.文献[64]提出了基于 1bit 深度像素的高斯-牛顿迭代法,文献[65,66]又进一步将其推广到了 2bit 深度像素的情况下.尽管这 3 种算法通过避免黑塞矩阵及其逆矩阵的计算并固定迭代步长的方式实现了较快的运动估计速度,但是由于只采用了两个梯度下降方向,并且低位深像素的梯度往往不同于 8 bit 深度像素,其预测质量与基于 8bit 深度像素的弹性运动估计尚存在较大差距.文献[62,63]则通过大量实验发现,弹性运动模型的高斯-牛顿解法对初始迭代点和迭代步长较为敏感,即固定的初始迭代点和迭代步长无法求解出全局最优解,进而采用 2bit 深度像素和均匀搜索模板将初始迭代点置于全局最优解的单调区间内,再利用离散余弦变换的低频能量比率和黄金分割法调整迭代步长,使之适应目标函数的线性程度,明显提升了弹性运动估计的计算效率和补偿质量.然而,作为一类牛顿型优化求解方法,文献[62,63]的快速算法在本质上仍不能避免牛顿型方法存在的不足,也就是说,目标函数偏离线性的程度越大,初始迭代点距离全局最优点越远,高斯-牛

顿法的收敛速度就越慢,甚至出现远离最优点或不收敛的现象<sup>[67]</sup>.事实上,视频数据以及运动补偿误差的复杂性,匹配误差曲面往往不会呈现给我们期望的理想线性性,文献[62]也已验证了这一点.

因此,对于上述两方面的研究工作来讲,将现有弹性运动模型的高斯-牛顿求解方法做出改进乃至改变是非常必要的,进而在降低其计算量的同时,使之收敛到更优解.这一问题的解决将有助于弹性运动估计走向实用.而从我们所掌握的文献来看,目前尚鲜见相关研究.

## 2 弹性运动模型以及传统弹性运动估计算法

为了便于下文工作的论述,本节首先介绍弹性运动模型,然后介绍运动向量的典型求解方法.

### 2.1 弹性运动估计模型简介

视频运动估计的目标是在参考帧的某个搜索窗口内,为当前待预测块  $\mathbf{I}$ (尺寸为  $B \times B$  像素)搜索到一个运动向量(矢量),使得  $\mathbf{I}$  与其最佳匹配块  $\mathbf{R}$  之间的误差平方和(sum of squared difference)最小,即:

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

其中,  $x_{ij}$  和  $y_{ij}$  分别表示当前块中  $i$  行  $j$  列像素的  $x$  坐标和  $y$  坐标;  $\mathbf{m}$  表示弹性运动向量;  $w(\cdot)$  表示弹性运动函数,其定义为

$$x'_{ij} = w(x_{ij}, \mathbf{m}) = x_{ij} + \sum_{k=1}^{p/2} \mathbf{m}_k \varphi_k(i, j) \quad (2)$$

$$y'_{ij} = w(y_{ij}, \mathbf{m}) = y_{ij} + \sum_{k=p/2+1}^p \mathbf{m}_k \varphi_k(i, j) \quad (3)$$

其中,  $p$  表示运动矢量的分量数目;  $\varphi_k$  表示弹性运动模型的基函数.虽然样条函数、仿射函数、小波函数都可作为基函数,但是文献[60]经过对比发现,选取离散余弦函数作为基函数对连续运动场具有较高的表示效率,即:

$$\varphi_k(i, j) = \varphi_{k+p/2}(i, j) = \cos \left[ \frac{(2i+1)\pi u}{2B} \right] \cos \left[ \frac{(2j+1)\pi v}{2B} \right] \quad (4)$$

其中,  $k = su + v + 1; u, v = 0, 1, \dots, s-1; s = \sqrt{p/2}; 0 \leq i, j < B$ . 文献[60]中取  $p=8$ .

### 2.2 弹性运动估计的高斯-牛顿求解方法

为了获得弹性运动向量  $\mathbf{m}$ ,文献[39,42,60-63]均采用了高斯-牛顿法进行求解,其主要思想是对匹配误差函数(公式(1))进行线性逼近,再通过反复迭代求出极小值.具体地,假设当前迭代点是  $\mathbf{m}$ ,并令像素  $(x_{ij}, y_{ij})$  处的预测误差为  $\mathbf{e}_{ij}(\mathbf{m}) = \mathbf{R}(w(x_{ij}, \mathbf{m}), w(y_{ij}, \mathbf{m})) - \mathbf{I}(x_{ij}, y_{ij})$ ,则函数  $\mathbf{e}_{ij}(\mathbf{m})$  可用其在  $\mathbf{m}$  处的 1 阶泰勒展开式近似表示.

$$\mathbf{e}_{ij}(\mathbf{m} + \Delta \mathbf{m}) \approx \mathbf{e}_{ij}(\mathbf{m}) + \mathbf{e}'_{ij}(\mathbf{m}) \Delta \mathbf{m} \quad (5)$$

其中,  $\mathbf{e}'_{ij}(\mathbf{m})$  表示  $\mathbf{e}_{ij}(\mathbf{m})$  在  $\mathbf{m}$  处的 1 阶偏导数  $\partial \mathbf{e}_{ij} / \partial \mathbf{m}$ . 将公式(5)代入公式(1),则有:

$$D = \arg \min_{\mathbf{m}} \left\{ \sum_{i=1}^B \sum_{j=1}^B [\mathbf{R}(w(x_{ij}, \mathbf{m}), w(y_{ij}, \mathbf{m})) - \mathbf{I}(x_{ij}, y_{ij})]^2 \right\} \approx \arg \min_{\Delta \mathbf{m}} \sum_{i=1}^B \sum_{j=1}^B [\mathbf{e}_{ij}(\mathbf{m}) + \mathbf{e}'_{ij}(\mathbf{m}) \Delta \mathbf{m}]^2 \quad (6)$$

为取得上式的最小值,将其对  $\Delta \mathbf{m}$  取导并令导数为 0,整理后,则有:

$$\sum_{i=1}^B \sum_{j=1}^B [\mathbf{e}'_{ij}(\mathbf{m})]^T \mathbf{e}'_{ij}(\mathbf{m}) \Delta \mathbf{m} = - \sum_{i=1}^B \sum_{j=1}^B [\mathbf{e}'_{ij}(\mathbf{m})]^T \mathbf{e}_{ij}(\mathbf{m}) \quad (7)$$

其中,上标  $T$  表示向量转置.令

$$\mathbf{b} = - \sum_{i=1}^B \sum_{j=1}^B [\mathbf{e}'_{ij}(\mathbf{m})]^T \mathbf{e}_{ij}(\mathbf{m}) \quad (8)$$

$$\mathbf{H} = \sum_{i=1}^B \sum_{j=1}^B [\mathbf{e}'_{ij}(\mathbf{m})]^T \mathbf{e}'_{ij}(\mathbf{m}) \quad (9)$$

显然,  $\mathbf{H}$  是一个  $p \times p$  阶的黑塞矩阵,并且根据求导的链式法则和  $\mathbf{e}_{ij}(\mathbf{m})$ 、公式(2)、公式(3)的定义,有:

$$\mathbf{H} = \sum_{i=1}^B \sum_{j=1}^B \left[ \frac{\partial \mathbf{R}}{\partial x'_{ij}} \frac{\partial w}{\partial \mathbf{m}} + \frac{\partial \mathbf{R}}{\partial y'_{ij}} \frac{\partial w}{\partial \mathbf{m}} \right]^T \left[ \frac{\partial \mathbf{R}}{\partial x'_{ij}} \frac{\partial w}{\partial \mathbf{m}} + \frac{\partial \mathbf{R}}{\partial y'_{ij}} \frac{\partial w}{\partial \mathbf{m}} \right] \quad (10)$$

其中,  $\partial \mathbf{R}/\partial x'_{ij}$  和  $\partial \mathbf{R}/\partial y'_{ij}$  分别表示匹配块沿着水平和垂直方向的梯度分量;  $\partial w/\partial \mathbf{m}$  是一个雅克比矩阵, 并且  $\partial w/\partial \mathbf{m}_k = \partial w/\partial \mathbf{m}_{p/2+k} = \varphi_k(i, j), k \in [1, p/2]$ . 于是有  $\Delta \mathbf{m} = \mathbf{H}^{-1} \mathbf{b}$ , 进而得到更新后的弹性运动向量  $\mathbf{m} \leftarrow \mathbf{m} + \Delta \mathbf{m}$ .

### 3 基于 Levenberg-Marquardt 法的弹性运动估计

对于弹性运动估计这类非线性无约束最小二乘问题, 主要有两类典型的优化解法: 牛顿型方法(如牛顿法和高斯-牛顿法)和负梯度方法(如最速下降法和对角线近似法)<sup>[67]</sup>. 理论表明, 当初迭代点距离局部极小点较近时, 牛顿法和高斯-牛顿法的收敛速度更快; 反之, 牛顿法和高斯-牛顿法则收敛较慢, 甚至会由于黑塞矩阵奇异或病态导致无法收敛的情形发生<sup>[68]</sup>, 而此时最速下降法和对角线近似法的收敛效率更高. 故而, 弹性运动估计的高斯-牛顿解法对初始搜索点存在一定的敏感性, 文献[62,63]的研究也验证了这一结论. 在这种情况下, 文献[69,70]将对角线近似法与高斯-牛顿法相结合, 修正了高斯-牛顿黑塞矩阵, 提出了一种 Levenberg-Marquardt(下文简称 L-M)黑塞矩阵:

$$\mathbf{H}_{\text{LM}} = \sum_{i=1}^N \left[ \frac{\partial \mathbf{R}}{\partial \mathbf{m}} \right]^T \left[ \frac{\partial \mathbf{R}}{\partial \mathbf{m}} \right] + \delta \sum_{i=1}^N \begin{bmatrix} \left( \frac{\partial \mathbf{R}}{\partial \mathbf{m}_1} \right)^2 & 0 & \dots & 0 \\ 0 & \left( \frac{\partial \mathbf{R}}{\partial \mathbf{m}_2} \right)^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \left( \frac{\partial \mathbf{R}}{\partial \mathbf{m}_p} \right)^2 \end{bmatrix} \quad (11)$$

从其定义可见, 当  $\delta \ll 1$  时,  $\mathbf{H}_{\text{LM}}$  近似为高斯-牛顿黑塞矩阵  $\mathbf{H}$ ; 而当  $\delta \gg 1$  时,  $\mathbf{H}_{\text{LM}}$  则近似为对角矩阵(仅差 1 个步长因子  $1/\delta$ ).

在  $\mathbf{H}_{\text{LM}}$  的基础上, 本文提出了基于 L-M 算法的弹性运动估计, 其计算步骤如下所示.

- Step 1. 输入当前待预测块  $\mathbf{I}$  和迭代次数  $T$ , 并将弹性运动向量  $\mathbf{m}$  初始化为 0, 令  $\delta \leftarrow 1$ , 迭代次数  $t \leftarrow 1$ .
- Step 2. 根据弹性运动向量  $\mathbf{m}$  和弹性运动函数(公式(2)、公式(3))计算  $\mathbf{I}$  的匹配块  $\mathbf{R}(w(x_{ij}, \mathbf{m}), w(y_{ij}, \mathbf{m}))$ .
- Step 3. 计算初始运动补偿误差  $\mathbf{e}_{ij}(\mathbf{m})$  及其平方和  $D_0$ .
- Step 4. 计算匹配块的梯度  $\nabla \mathbf{R}$ .
- Step 5. 计算雅克比矩阵  $\partial w/\partial \mathbf{m}$ .
- Step 6. 计算  $\frac{\partial \mathbf{R}}{\partial \mathbf{m}_k} = \frac{\partial \mathbf{R}}{\partial x'} \frac{\partial w}{\partial \mathbf{m}_k} + \frac{\partial \mathbf{R}}{\partial y'} \frac{\partial w}{\partial \mathbf{m}_k}, k \in [1, p]$ .
- Step 7. 计算高斯-牛顿黑塞矩阵  $\mathbf{H}$ , 其中,  $\mathbf{H}_{a,b} = \sum_{i=1}^B \sum_{j=1}^B [(\partial \mathbf{R}/\partial \mathbf{m}_a)^T (\partial \mathbf{R}/\partial \mathbf{m}_b)], a, b \in [1, p]$ .
- Step 8. 计算向量  $\mathbf{b}$ , 其中,  $\mathbf{b}_k = -\sum_{i=1}^B \sum_{j=1}^B [(\partial \mathbf{R}/\partial \mathbf{m}_k) \mathbf{e}_{ij}], k \in [1, p]$ .
- Step 9. 根据公式(11)计算 L-M 黑塞矩阵  $\mathbf{H}_{\text{LM}}$ .
- Step 10. 计算运动向量增量  $\Delta \mathbf{m} \leftarrow \mathbf{H}_{\text{LM}}^{-1} \mathbf{b}$ , 并更新弹性运动向量  $\mathbf{m} \leftarrow \mathbf{m} + \Delta \mathbf{m}$ .
- Step 11. 更新当前块  $\mathbf{I}$  的匹配块  $\mathbf{R}(w(x_{ij}, \mathbf{m}), w(y_{ij}, \mathbf{m}))$ , 并计算它与当前块  $\mathbf{I}$  的匹配误差平方和  $D_t$ .
- Step 12. 若  $D_t > D_{t-1}$ , 则令  $\delta \leftarrow \delta \times \lambda$  ( $\lambda$  为对角矩阵阻尼系数的更新因子, 传统 L-M 算法一般将其设置为 10), 转入 Step 9; 否则, 转入 Step 13.
- Step 13. 令  $\delta \leftarrow \delta/\lambda, t \leftarrow t+1$ . 若  $t > T$  或增量  $\|\Delta \mathbf{m}\| < T_m$ , 则输出运动向量  $\mathbf{m}$  和块匹配误差, 算法结束. 其中,  $\|\cdot\|$  表示向量模长,  $T_m$  表示预设的增量阈值; 否则, 转入 Step 4.

与高斯-牛顿优化算法相比, L-M 算法对于光滑目标函数体现出收敛速度快、稳定性好等优点. 但是, 一

方面它在每次迭代时均需计算黑塞矩阵及逆矩阵,其渐近时间复杂度达到了  $\max\{O(p^3), O(p^2B^2)\}$ ; 另一方面,现有文献往往将对角矩阵阻尼系数的更新因子  $\lambda$  设置成正数,这对于阶数较低的、形式简单的显式目标函数较为有效.但是,视频运动补偿的误差曲面非常复杂,甚至无法用任何一个显式函数表示出来.实验结果表明,  $\lambda$  为正数在某些情况下反而会增大运动补偿误差,而将其设置为负数却更有效.基于上述考虑,下文从黑塞矩阵的快速计算和对角矩阵系数的自适应更新两方面改进 L-M 算法,进一步提高弹性运动估计/补偿的效率.

#### 4 黑塞矩阵的快速计算方法

由第 2 节和第 3 节可知,黑塞矩阵的计算与参考块梯度、弹性基函数有关,并且黑塞矩阵本身也具有明显的对称性.为此,本文从两个角度加快黑塞矩阵的计算:(a) 根据 2D 离散余弦函数矩阵的特点减少基函数的计算量;(b) 利用黑塞矩阵的对称性避免重复计算.

##### 4.1 弹性运动基函数的快速计算

根据公式(4),弹性运动模型共有 4 种基函数.以  $B=8$  为例,图 1 给出了 4 种基函数对应的  $8 \times 8$  矩阵的数值分布.其中,符号  $A \sim H$  分别代表  $\cos(\pi/16)$ 、 $\cos(3\pi/16)$ 、 $\cos(5\pi/16)$ 、 $\cos(7\pi/16)$ 、 $\cos(9\pi/16)$ 、 $\cos(11\pi/16)$ 、 $\cos(13\pi/16)$  和  $\cos(15\pi/16)$ .



Fig.1 Numerical distribution of four basis function matrixes of elastic motion model

图 1 弹性运动模型 4 种基函数矩阵的数值分布

从图 1 可见:① 第 1 个基函数矩阵只包含元素 1;② 第 2 个基函数矩阵的元素只与纵坐标有关;③ 第 3 个基函数矩阵的元素只与横坐标有关;④ 第 4 个基函数矩阵的元素为第 2 个和第 3 个基函数矩阵的对应元素之积.故此,基函数矩阵可通过少量计算和大量赋值得到,即:首先,  $\varphi_1$  对应的矩阵只需全部赋值为 1;其次,  $\varphi_2$  对应的矩阵只需计算一行,其余各行通过赋值得到,转置后再赋值给  $\varphi_3$  的各列;最后,将  $\varphi_2$  和  $\varphi_3$  矩阵的上三角元素对应相乘得到  $\varphi_4$  矩阵的上三角元素,而下三角元素可通过转置赋值获得.这样,需要计算的元素个数仅占全部基函数

矩阵元素数量的 17%.

### 4.2 高斯-牛顿黑塞矩阵的快速计算

由黑塞矩阵的定义可知,其上三角元素和下三角元素关于主对角线对称,即  $\mathbf{H}_{a,b}=\mathbf{H}_{b,a}(1\leq a\leq b\leq B)$ .

此外,根据黑塞矩阵的计算过程和第 4.1 节的基函数矩阵数值分布特点还可进一步发现高斯-牛顿黑塞矩阵中各元素之间的相等关系,从而降低黑塞矩阵的计算量,下面以  $\mathbf{H}_{1,4}$  和  $\mathbf{H}_{2,3}$  为例来分析.由两个元素的定义,有:

$$\mathbf{H}_{1,4} = \sum_{i=1}^B \sum_{j=1}^B \left[ \left( \frac{\partial \mathbf{R}}{\partial \mathbf{m}_1} \right)^T \left( \frac{\partial \mathbf{R}}{\partial \mathbf{m}_4} \right) \right] = \sum_{i=1}^B \sum_{j=1}^B \left[ \left( \frac{\partial \mathbf{R}}{\partial x_{ij}}, \frac{\partial \mathbf{R}}{\partial y_{ij}} \right) \cdot \begin{pmatrix} \frac{\partial x_{ij}}{\partial \mathbf{m}_1} \\ 0 \end{pmatrix} \cdot \left( \frac{\partial \mathbf{R}}{\partial x_{ij}}, \frac{\partial \mathbf{R}}{\partial y_{ij}} \right) \cdot \begin{pmatrix} \frac{\partial x_{ij}}{\partial \mathbf{m}_4} \\ 0 \end{pmatrix} \right] = \sum_{i=1}^B \sum_{j=1}^B \left[ \left( \frac{\partial \mathbf{R}}{\partial x_{ij}} \right)^2 \cdot \varphi_1(i, j) \cdot \varphi_4(i, j) \right] \quad (12)$$

同理可得,

$$\mathbf{H}_{2,3} = \sum_{i=1}^B \sum_{j=1}^B [(\partial \mathbf{R} / \partial x_{ij})^2 \cdot \varphi_2(i, j) \cdot \varphi_3(i, j)] \quad (13)$$

而由图 1 所示的基函数矩阵数值分布可知,  $\varphi_1$  和  $\varphi_4$  对应元素的乘积与  $\varphi_2$  和  $\varphi_3$  对应元素的乘积相同,即对于  $\forall i, j \in [1, 8]$ , 有  $\varphi_1(i, j)\varphi_4(i, j) = \varphi_2(i, j)\varphi_3(i, j)$ . 因此有  $\mathbf{H}_{1,4} = \mathbf{H}_{2,3}$ .

采用与上述相同的推导过程以及公式(14)的等价关系,就可得到图 2 所示的高斯-牛顿黑塞矩阵的数值分布.可见,我们只需计算 64 个元素中的 27 个  $A \sim H, a \sim s$ , 其余的 37 个元素只需简单赋值即可得到.同时, L-M 黑塞矩阵(见公式(11))第 2 项的对角元素与高斯-牛顿黑塞矩阵的对角元素  $A \sim H$  相同,无需计算.

$$\left. \begin{aligned} \varphi_1(i, j)\varphi_4(i, j) &= \varphi_1(i, j)\varphi_8(i, j) = \varphi_5(i, j)\varphi_4(i, j) = \varphi_5(i, j)\varphi_8(i, j) \\ &= \varphi_2(i, j)\varphi_3(i, j) = \varphi_2(i, j)\varphi_7(i, j) = \varphi_6(i, j)\varphi_3(i, j) = \varphi_6(i, j)\varphi_7(i, j), \quad \forall i, j \in [1, 8] \end{aligned} \right\} \quad (14)$$

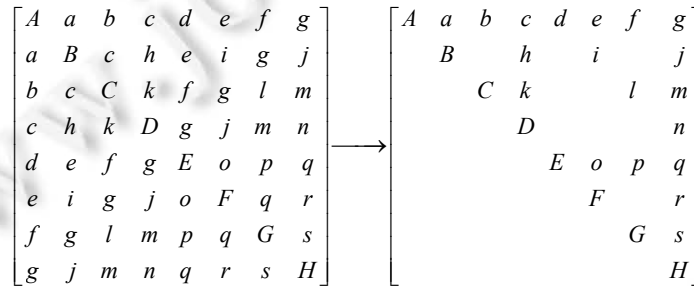


Fig.2 Numerical distribution of the Gauss-Newton Hessian matrix

图 2 高斯-牛顿黑塞矩阵的数值分布

综上, L-M 黑塞矩阵  $\mathbf{H}_{LM}$  中需要计算的元素个数仅占全部元素数量的 37.5%.

## 5 对角矩阵阻尼系数的自适应更新方法

### 5.1 $\lambda$ 的取值对弹性运动估计效率的影响

由  $\mathbf{H}_{LM}$  的定义(见公式(11))可知, L-M 算法的搜索方向和更新步长取决于高斯-牛顿黑塞矩阵和对角矩阵的向量和.然而,传统 L-M 算法把对角矩阵阻尼系数  $\delta$  的更新因子  $\lambda$  设置成正常数(一般设置为 10),这使得  $\delta$  在迭代过程中以 10 的幂级迅速改变.一方面,考虑到视频帧的相邻像素值之间存在较高的相关性,其梯度幅值往往较小,过大的  $\lambda$  会导致弹性运动向量  $\mathbf{m}$  的搜索方向快速偏离或靠近负梯度方向,更新步长也发生大幅变化.对于多峰分布的运动补偿误差曲面,这种粗糙的参数更新方式难免会错过最佳的搜索方向.直观起见,图 3 以二维向量和  $\lambda=5$  为例给出了示意图.其中,  $\mathbf{H}_{GN}$  表示高斯-牛顿黑塞矩阵对应的下降方向;  $\mathbf{H}_{D1}$ 、  $\mathbf{H}_{D2}$  和  $\mathbf{H}_{D3}$  分别表示对角矩阵在连续 3 次迭代中对应的下降方向;  $\mathbf{H}_{LM1}$ 、  $\mathbf{H}_{LM2}$  和  $\mathbf{H}_{LM3}$  则表示由 2 个矩阵确定的 3 个迭代搜索方向,而  $\mathbf{H}_{LM}^*$  却是最优下降方向.可见,  $\lambda$  变化过快会导致搜索空间划分得不够精细,使运动向量的搜索偏离最佳方向.另一方

面,将 $\lambda$ 设置成正常数还会将运动向量的搜索限制在整个向量空间的一个子空间内.如图 3 所示,下降方向仅局限在第 1 象限内,而最优搜索方向  $\mathbf{H}_{LM}^*$  却可能位于第 2 象限.显然,上述两个方面都会影响运动估计的收敛.

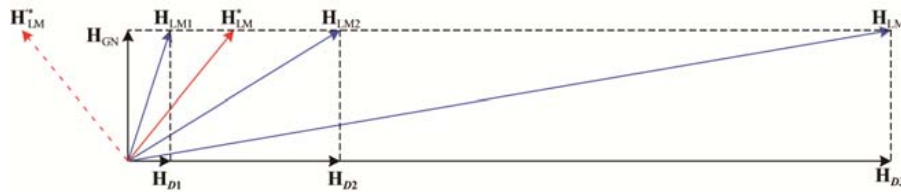
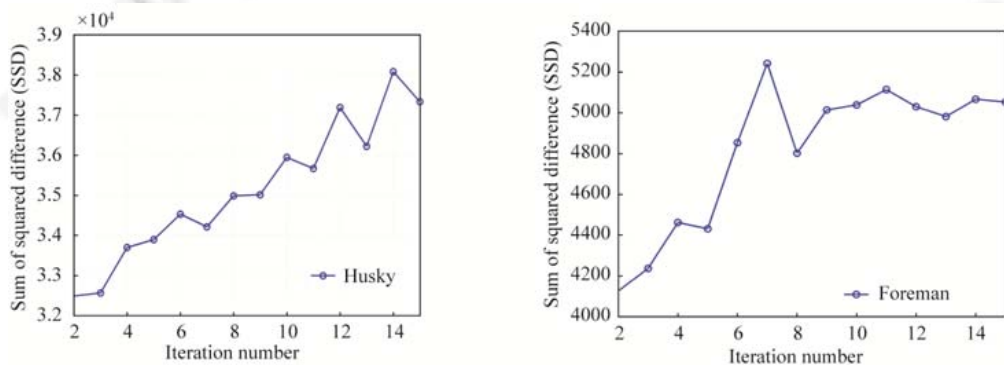


Fig.3 Iteration direction diagram of Levenberg-Marquardt algorithm when  $\lambda$  is set to be a large positive constant  
图 3 将 $\lambda$ 设置为较大的正常数时,Levenberg-Marquardt 算法的迭代方向示意图

为了验证以上分析,以 Husky 序列和 Foreman 序列的第 2 帧为例进行实验:首先利用菱形搜索确定最优的平移向量作为初始迭代点,使之尽量靠近全局最优解;然后采用第 3 节的 L-M 算法( $\lambda=10$ )和文献[62]中的方法分别计算  $\mathbf{m}$  的搜索方向和阻尼步长,并统计前 15 次迭代的运动补偿误差.图 4 给出了 Husky 序列以(16,0)像素为左上角的宏块和 Foreman 序列以(112,16)像素为左上角的宏块在第 2~15 次迭代的误差曲线(第 1 次迭代的误差较大,为了清晰地展现曲线变化的细节,这里未予绘制).可见,尽管基本排除了初始迭代点与全局最优解的距离对优化过程的影响,但是无论对于纹理、运动复杂的 Husky 序列,抑或对纹理、运动中等的 Foreman 序列,其运动补偿误差随着迭代的进行非但没有降低,反而均呈现出波动上升的趋势.事实上,本文在几乎所有的测试序列上都发现了类似现象.究其原因,存在 3 种可能情况:① 黑塞矩阵  $\mathbf{H}_{LM}$  奇异或病态使得迭代无法正常收敛;② 迭代步长  $\|\Delta\mathbf{m}\|$  不准确以致更新后的运动向量  $\mathbf{m}$  未达到或超过了极值点;③ 搜索过程偏离了最佳下降方向.由于 L-M 算法已在理论上避免了  $\mathbf{H}_{LM}$  的奇异和病态,并且一维线搜索保证了  $\|\Delta\mathbf{m}\|$  的有效性,我们可以基本排除①、②两种情况.综合上述理论分析和实验数据可以推知,将对角矩阵阻尼系数  $\delta$  的更新因子  $\lambda$  固定为正常数对于弹性运动估计不够合理.



(a) Husky 第 2 帧(16,0)宏块的运动补偿误差变化曲线  
(b) Foreman 第 2 帧(112,16)宏块的运动补偿误差变化曲线  
Fig.4 Motion-compensated difference curves of the Levenberg-Marquardt method in its 2~15 iteration when  $\lambda=10$   
图 4 当 $\lambda=10$ 时,Levenberg-Marquardt 法第 2~15 次迭代的运动补偿误差曲线

为此,文献[71,72]分别将每次迭代后的目标函数值及其 0.01 倍作为阻尼系数  $\delta$ .但当初始迭代点距离全局最优解较远时,文献[71,72]的方法将由于目标函数值过大导致迭代步长很小而无法快速收敛;反之,该方法又会产生过小的  $\delta$ ,使得对角矩阵丧失梯度下降作用.尽管文献[73]进一步采用目标函数值的  $s$ -范数( $s \in [1,2]$ )来更新  $\delta$ ,能够在一定程度上克服上述不足,但是如何根据迭代点与全局最优解的距离选择合适的  $s$  也并不容易,并且不能有效避免目标函数值在迭代后期出现反复振荡.于是,文献[74]提出一种基于对数线性函数的对角矩阵系数更新方法,增强了传统 L-M 算法的收敛稳定性,而其初始迭代速度却较慢.针对这一不足,文献[75]提出通过迭代向



量  $\mathbf{m}$  之间的内积来计算更新因子  $\lambda$ , 加快了初始迭代效率, 同时引进对角优势矩阵(diagonally dominant matrix), 减少了为使矩阵  $\mathbf{H}_{LM}$  正定而反复更新  $\delta$  所需的尝试次数; 文献[76]根据目标函数值的实际下降量与预测下降量之比确定更新因子  $\lambda$ ; 文献[77]则利用测地线距离来修正迭代步长. 然而, 这 3 种方法需要的计算量均较大, 分别需多次判断  $\mathbf{H}_{LM}$  的正定性、根据泰勒展开式计算匹配误差的预测下降量、计算目标函数的测地线距离, 不能满足弹性运动对低运算量的要求. 故此, 下文提出一种用来自适应更新  $\delta$  的快速策略.

## 5.2 基于步长平方商的对角矩阵阻尼系数交替更新

因为文献[62]的一维线搜索方法能够根据视频纹理特征估计出较准确的阻尼搜索步长, 所以如果某次迭代的搜索步长  $\|\Delta\mathbf{m}\|$  较大, 就说明运动补偿误差的实际下降幅度与预测下降幅度较为一致; 反之, 表明二者之间存在一定偏差, 因而采用 2 次曲面不能有效逼近运动补偿误差曲面. 进一步地, 若相邻 2 次迭代的搜索步长差异较大, 则可推知运动补偿误差曲面存在明显波动, 其阶数高于 2 阶, 且搜索将从一个峰值的单调区间进入另一个峰值的单调区间. 根据文献[75]的研究结论, 此时迭代点距离全局最优点较远, 无论运动补偿误差是否减小都应增大对角矩阵阻尼系数的更新因子  $\lambda$ , 从而提高梯度下降法的权重; 相反地, 若相邻 2 次迭代的搜索步长相近, 则表明运动补偿误差曲面沿着相应搜索方向的数值分布较为一致, 当前搜索方向就是合理的下降方向, 只需令对角矩阵阻尼系数不变或者缓慢变化来保持高斯-牛顿法和梯度下降法的相对权重、朝着同一方向继续搜索即可.

根据上述分析, 本文采用相邻 2 次迭代的搜索步长的平方商作为自适应计算  $\lambda$  的依据. 具体地, 假设当前为第  $t(1 \leq t \leq T)$  次迭代, 则  $\delta$  在该次迭代过程中的自适应因子  $\lambda_t^*$  为

$$\lambda_t^* = \begin{cases} \lambda_{\min}, & \text{当 } t=1, 2 \\ \min \left\{ \frac{\|\Delta\mathbf{m}\|_{t-1}^2 / \|\Delta\mathbf{m}\|_{t-2}^2 + \lambda_{\min}}{2}, \lambda_{\max} \right\}, & \text{当 } t > 2 \text{ 且 } \|\Delta\mathbf{m}\|_{t-1} \geq \|\Delta\mathbf{m}\|_{t-2} \\ \min \left\{ \frac{\|\Delta\mathbf{m}\|_{t-2}^2 / \|\Delta\mathbf{m}\|_{t-1}^2 + \lambda_{\min}}{2}, \lambda_{\max} \right\}, & \text{当 } t > 2 \text{ 且 } \|\Delta\mathbf{m}\|_{t-1} < \|\Delta\mathbf{m}\|_{t-2} \end{cases} \quad (15)$$

其中,  $\|\Delta\mathbf{m}\|_{t-1}^2$  和  $\|\Delta\mathbf{m}\|_{t-2}^2$  分别表示第  $(t-1)$  次和第  $(t-2)$  次迭代时  $\Delta\mathbf{m}$  的模长的平方;  $\lambda_{\min}$  和  $\lambda_{\max}$  分别为更新因子的下限和上限. 考虑到  $\lambda$  过小会导致对角矩阵阻尼系数更新较慢, 而过大的  $\lambda$  又会错过可能的最佳搜索方向, 二者都将影响 L-M 优化的运算效率, 本文依据文献[78], 令  $\lambda_{\min}=2$  和  $\lambda_{\max}=10$ .

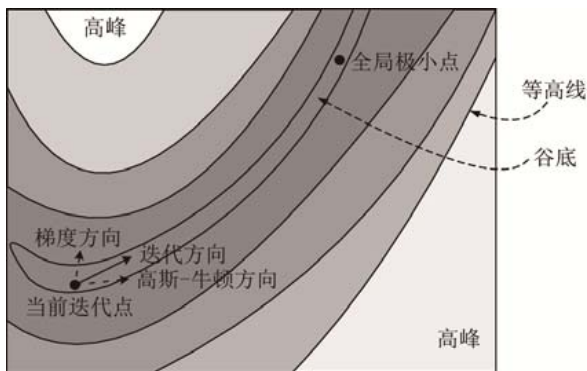


Fig.5 Diagram of controlling iteration using both gradient direction and Gaussian-Newton direction

图 5 利用梯度方向和高斯-牛顿方向控制迭代的示意图  
 度慢甚至无法收敛的情况, 而此时利用梯度方向与高斯-牛顿方向的矢量和控制迭代却有望将搜索引导向全局最优点. 具体地,  $\delta$  的正、负交替更新方法是

但是, 公式(15)给出的自适应因子依然是正数, 其目的是通过保证  $\mathbf{H}_{LM}$  正定性使迭代求解沿着运动补偿误差曲面的下降方向进行. 尽管这是一种稳妥的做法, 但文献[77]的研究结论却表明, 当迭代点陷入由多个运动补偿误差曲面峰所围的狭长谷底时, 保持自适应因子为正数会形成一条锯齿状搜索路径, 若适当允许朝着函数值增长的方向迭代则可更快地收敛到更小点. 基于这一思路, 本文在搜索中借助梯度方向提出一种  $\delta$  的正、负交替更新策略. 直观起见, 图 5 给出了这一过程的示意图, 其中, 灰度越浅的区域表示补偿误差越大. 可见, 当前迭代点处于两峰之间的谷底, 其负梯度指向远离全局极小点的方向, 且由于距离全局最优解较远, 沿高斯-牛顿方向搜索会出现收敛速度

$$\delta = \begin{cases} -\delta \times \lambda_t^*, & \text{当 } D_t > D_{t-1} \text{ 时} \\ \delta / \lambda_t^*, & \text{否则} \end{cases} \quad (16)$$

其中,  $D_t$  和  $D_{t-1}$  分别表示第  $t$  次迭代和第  $(t-1)$  次迭代的匹配误差平方和, 从而以高斯-牛顿方向  $\mathbf{H}_{GN}$  为中心, 沿着当前搜索方向的对称方向进行搜索(如图 3 所示的  $\mathbf{H}_{LM}^*$  和  $\mathbf{H}_{LM}^*$ ), 使搜索方向可达整个运动向量空间, 提高了算法收敛到全局最优解的概率。

### 6 采用改进 L-M 法的弹性运动估计步骤

在第 3 节算法的基础上, 结合第 4 节和第 5 节的改进思路, 下面给出基于改进 L-M 法的弹性运动估计步骤。对于每个待预测块  $\mathbf{I}$ , 其运动估计/补偿的详细步骤如下。

- Step 1. 输入  $\mathbf{I}$ 、迭代次数  $T$ 、增量阈值  $T_m$ 、 $\lambda_{\min}$  和  $\lambda_{\max}$ , 并令  $t \leftarrow 1, \delta \leftarrow 1, \lambda_1^* \leftarrow \lambda_{\min}$ 。
- Step 2. 利用整数像素精度的菱形搜索估计平移分量  $\mathbf{m}_1$  和  $\mathbf{m}_{p/2+1}$ , 将其余分量置 0, 并计算  $\mathbf{I}$  的初始运动补偿误差  $\mathbf{e}_{ij}(\mathbf{m})$  及其平方和  $D_0$ 。
- Step 3. 根据第 4.1 节计算弹性运动基函数  $\varphi_1 \sim \varphi_p$ 。
- Step 4. 计算匹配块的梯度  $\nabla \mathbf{R}$  以及雅克比矩阵  $\partial w / \partial \mathbf{m}$ 。
- Step 5. 根据第 4.2 节计算高斯-牛顿黑塞矩阵  $\mathbf{H}$ 。
- Step 6. 将  $\mathbf{H}$  及其主对角线元素、 $\mathbf{e}_{ij}(\mathbf{m})$ 、 $\delta$  代入公式(8)、公式(11), 计算向量  $\mathbf{b}$  和 L-M 黑塞矩阵  $\mathbf{H}_{LM}$ 。
- Step 7. 计算弹性运动向量增量  $\Delta \mathbf{m} \leftarrow \mathbf{H}_{LM}^{-1} \mathbf{b}$ , 进而更新运动向量  $\mathbf{m} \leftarrow \mathbf{m} + \Delta \mathbf{m}$ 。
- Step 8. 将运动向量  $\mathbf{m}$  代入公式(2)和公式(3), 建立  $\mathbf{I}$  的每个像素与其匹配像素的坐标映射, 并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  $\mathbf{e}_{ij}(\mathbf{m})$  及其平方和  $D_t$ 。
- Step 9. 若  $D_t > D_{t-1}$ , 则令  $\delta \leftarrow -\delta \times \lambda_t^*$ , 转入 Step 6; 否则, 转入 Step 10。
- Step 10. 令  $\delta \leftarrow \delta / \lambda_t^*, t \leftarrow t + 1$ 。若  $t > T$  或增量  $\|\Delta \mathbf{m}\| < T_m$ , 则输出运动向量  $\mathbf{m}$  和块匹配误差, 待预测块  $\mathbf{I}$  的运动估计结束; 否则, 根据公式(15)更新  $\lambda_t^*$ , 转入 Step 4。

### 7 实验结果与分析

为了验证本文算法的有效性, 以 CIF(common intermediate format)、4CIF 和高清格式的 37 个标准视频序列(见表 1)的 1~90 帧为例进行大量实验, 并将基于块平移模型的全搜索(full search, 简称 FS)、基于改进高斯-牛顿法的弹性运动估计<sup>[62]</sup>、基于菱形搜索(diamond search, 简称 DS)+传统 L-M 法的弹性运动估计、基于菱形搜索+对角优势矩阵的 L-M 法的弹性运动估计<sup>[75]</sup>和基于改进 L-M 法的弹性运动估计的预测结果进行比较。

Table 1 Test video sequences' names and their formats

表 1 测试视频序列名称及其格式

空间分辨率	视频序列名称
352×288	Akiyo, Bowling, Coastguard, Container, Flower, Foreman, Hall, Highway, Husky, Ice, Intros_cif, Mobile, Mother_daug, News, Pamphlet, Paris, Sign_irene, Silent, Soccer, Students, Tempete, Tennis, Vtc1nw, Waterfall
704×576	Crew, City, Harbour
832×480	BQmall, Flowervase
1280×720	sc_Robot, FourPeople, Johnny, KristenAndSara
1920×1080	BQTerrace, Cactus, ParkScene, VenueVu

实验参数设置如下: 搜索窗口为 33×33 像素, 块尺寸为 16×16 像素,  $\lambda_{\min}=2, \lambda_{\max}=10, T_m=0.0001, T=15$ (这与文献[60,62]的设置相同); 文献[75]的参数与原文相同。运动补偿帧的质量采用 PSNR 进行评价。

#### 7.1 运动估计/补偿质量的比较

表 2 给出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 2** Motion-compensated PSNR comparison  
表 2 运动补偿的 PSNR 比较

视频序列名称	块匹配全搜索 (16×16 块)	块匹配全搜索 (8×8 块)	基于改进高斯-牛顿法的弹性运动估计 <sup>[62]</sup>	基于 DS+传统 L-M 法的弹性运动估计	基于 DS+对角优势矩阵的 L-M 法的弹性运动估计 <sup>[75]</sup>	基于 DS+步长平方商 L-M 法的弹性运动估计	基于 DS+交替更新 L-M 法的弹性运动估计	基于改进 L-M 法的弹性运动估计
Akiyo	42.67	-	40.81	42.31	42.23	42.28	46.01	46.04
Bowing	42.39	-	40.54	41.32	42.21	41.40	42.54	43.48
Coastguard	29.71	-	32.02	32.34	32.35	32.35	32.35	32.38
Container	38.35	-	38.91	39.72	39.73	39.80	41.01	41.05
Flower	26.03	-	26.32	26.77	26.77	26.78	26.81	27.25
Foreman	33.52	-	30.65	31.44	31.56	31.42	35.52	36.03
Hall	34.48	-	35.22	37.14	37.19	37.21	37.23	37.29
Highway	36.08	-	31.18	33.70	33.84	34.16	36.27	36.79
Husky	19.10	-	21.20	21.82	21.86	21.85	21.89	21.97
Ice	30.99	-	34.83	36.13	36.15	36.07	36.13	36.15
Intros_cif	37.30	-	33.86	34.71	35.22	35.22	38.15	38.51
Mobile	23.96	-	27.74	28.56	28.60	28.56	28.65	28.70
Mother_daug	40.26	-	41.86	42.89	42.98	42.88	42.96	42.98
News	37.83	-	37.82	38.78	38.79	38.79	41.28	41.38
Pamphlet	43.12	-	43.41	43.29	43.50	43.27	43.31	43.37
Paris	30.70	-	33.56	34.00	34.03	34.03	34.19	34.51
Sign_irene	33.56	-	36.06	36.21	36.40	36.22	36.40	36.41
Silent	36.27	-	31.44	32.58	32.67	33.91	39.08	39.11
Soccer	29.61	-	31.74	32.18	32.14	32.12	32.12	32.13
Students	39.49	-	41.23	42.17	42.18	42.26	42.78	42.94
Tempete	27.75	-	30.56	30.61	30.80	30.72	30.71	30.87
Tennis	28.91	-	30.48	31.11	31.12	31.06	31.08	31.12
Vtc1nw	44.70	-	42.06	42.75	42.75	42.76	45.43	45.44
Waterfall	35.53	-	39.62	39.80	39.80	39.82	39.82	39.86
Crew	32.36	-	33.50	34.42	34.43	34.42	34.40	34.43
City	30.86	-	32.13	32.98	32.98	33.06	32.97	33.00
Harbour	28.10	-	30.99	31.41	31.42	31.41	31.42	31.43
BQmall	29.63	32.10	30.20	32.18	31.72	32.19	32.18	32.20
Flowervase	37.42	37.99	36.75	39.88	39.36	39.86	39.90	39.91
sc_Robot	32.20	33.88	34.46	34.92	35.42	35.87	35.86	35.88
FourPeople	38.43	39.46	39.84	40.01	39.80	40.39	40.26	40.46
Johnny	38.83	39.79	39.69	40.12	40.12	40.73	40.70	40.75
KristenAndSara	39.38	40.20	40.37	41.16	41.41	41.45	41.42	41.46
BQTerrace	25.51	26.20	26.64	28.10	27.22	28.12	28.15	28.16
Cactus	29.53	30.82	31.17	31.50	31.33	31.76	31.77	31.91
ParkScene	29.98	30.62	32.09	32.11	32.26	32.30	32.32	32.34
VenueVu	27.12	28.51	29.17	28.11	28.11	28.11	28.11	28.12
平均	33.56	33.96	34.33	35.11	35.15	35.26	35.98	36.10

由表 2 可知:

• 尽管 FS 是目前编码标准中预测精度最高的运动估计/补偿算法,但是由于运动模型本身的局限,其平均 PSNR 比本文算法低 2.54dB.

• 基于改进高斯-牛顿法的弹性运动估计利用初始迭代点预测和一维线搜索取得了优于 FS 的性能(比后者高出 0.77dB),但其 PSNR 却比本文提出的基于菱形搜索+传统 L-M 法的弹性运动估计低 0.79dB,这表明,L-M 法比高斯-牛顿法更适用于弹性运动模型的求解.

• 基于菱形搜索+对角优势矩阵的 L-M 法的弹性运动估计利用迭代向量内积更新对角矩阵阻尼系数,其 PSNR 比基于菱形搜索+传统 L-M 法的弹性运动估计提高了 0.04dB,表明自适应选取阻尼系数有利于获得更高的运动补偿质量.然而,由于文献[75]修改 L-M 黑塞矩阵的方式无法保证其是严格对角占优的,且被修改的对角元素难免不会使迭代过程偏离真实的梯度下降方向,其 PSNR 较之基于改进 L-M 法的弹性运动估计低 0.95dB,这说明本文的系数自适应交替更新策略可更有效地提高弹性运动估计预测精度.

• 表 2 第 7 列和第 8 列进一步对比了基于步长平方商和基于交替策略( $\lambda=10$ )的对角矩阵阻尼系数更新方法的性能增益,两者的平均 PSNR 比基于菱形搜索+传统 L-M 法的弹性运动估计分别提高了 0.15dB 和 0.87dB.可见,正、负交替更新策略为本文算法贡献了大部分的性能提升.

另外,对于高清视频的运动估计,H.264 和 HEVC 通过缩小块尺寸的方式提高了预测准确率;而且,在相同的块尺寸条件下,块平移模型的运动向量数量仅为弹性模型的 1/4.为此,本文统计了当块尺寸为  $8 \times 8$  像素时,块匹配 FS 对 10 个高清视频序列的运动补偿 PSNR(结果见表 2 第 3 列),虽然它在这 10 个序列上的平均 PSNR 比  $16 \times 16$  像素的块匹配 FS 提高了 1.15dB,但是仍较基于改进 L-M 法的弹性运动估计( $16 \times 16$  块)低 1.16dB.可见,即使采用了更精细的块结构,在同等的运动向量开销下,基于块平移模型的全搜索性能仍与本文算法存在一定差距.

需要指出,本文算法在高清 VenueVu 序列上的 PSNR 与  $8 \times 8$  像素的块匹配 FS、基于改进高斯-牛顿法的弹性运动估计存在一定差距.其原因在于:该视频是由计算机生成的动画序列,纹理细腻且存在快速的全局形变,而本文算法在初始搜索时选用的菱形运动估计精度低于后两种算法的全搜索.如果采用与基于改进高斯-牛顿法的弹性运动估计相同的初始搜索,本文算法对 VenueVu 序列的运动补偿 PSNR 将达到 29.51dB.这表明,本文的对角矩阵自适应更新方法在个别情况下也不能完全克服传统 L-M 算法对于初始迭代点的敏感性.

图 6(a)~图 6(d)进一步给出了 Akiyo、Foreman、Mobile 和 Husky 序列的 PSNR 逐帧比较情况.

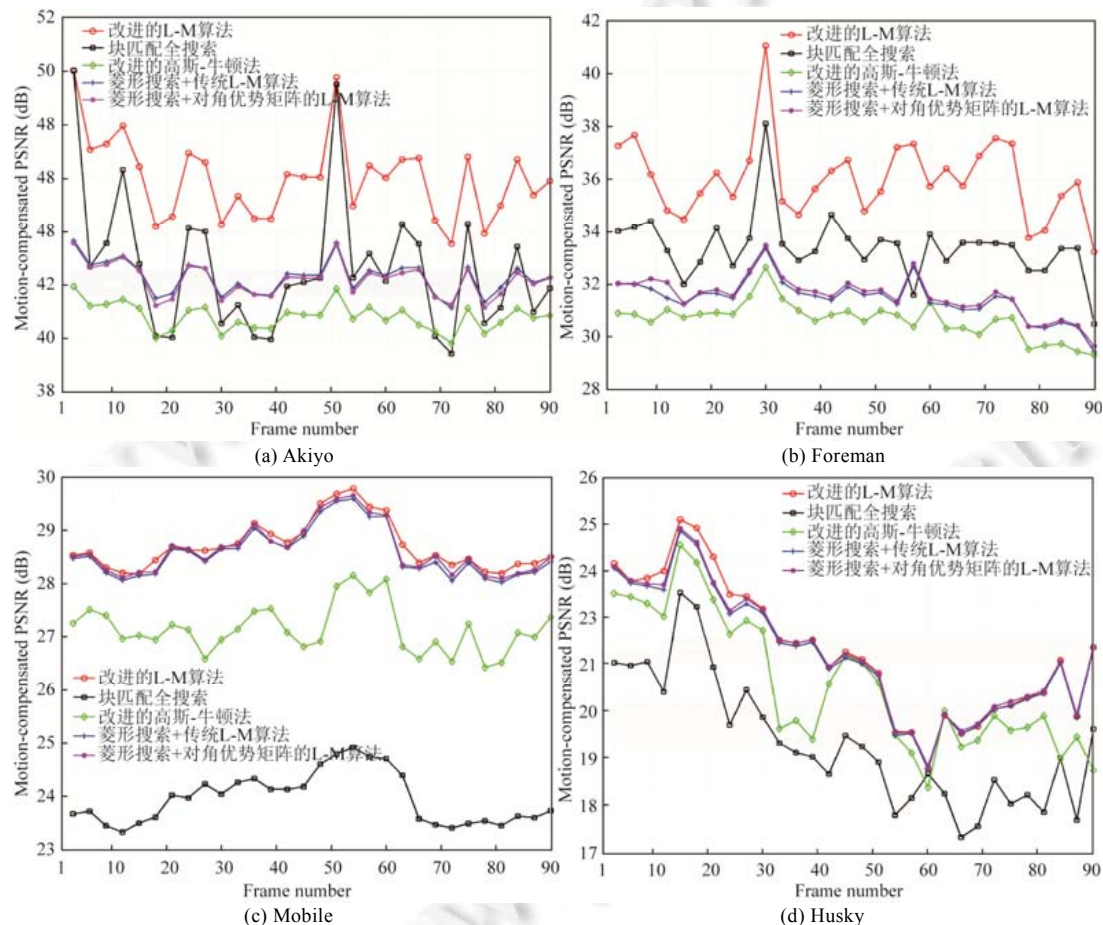


Fig.6 Frame-wise PSNR comparison of Akiyo, Foreman, Mobile, and Husky sequences

图 6 Akiyo、Foreman、Mobile 和 Husky 序列的逐帧 PSNR 比较

如图 6 所示,这 4 个序列的纹理细节和运动复杂度依次增高,且包含摄像机的拉摄、摇摄运动和物体的局部形变.从图中可见,对于具有不同纹理、运动特点的视频序列,基于改进 L-M 法的弹性运动估计均取得了最高的运动补偿质量.并且,对于快速摇摄、纹理细节复杂的序列(如 Mobile、Husky),本文提出的两种算法均明显优

于块匹配 FS 和基于改进高斯-牛顿法的弹性运动估计;而对于慢速运动或存在局部形变的序列(如 Akiyo、Foreman),基于改进 L-M 法的弹性运动估计较之块匹配 FS 亦有稳定提高。

## 7.2 收敛性和收敛速度比较

收敛性和收敛速度是衡量弹性运动估计效率的重要指标之一.图 7 所示分别为 Husky 和 Akiyo 序列采用基于改进高斯-牛顿法的弹性运动估计、基于改进 L-M 法的弹性运动估计的前 14 次迭代结果.可见,前者在纹理和运动较为复杂的序列(如 Husky)上收敛效率较高,反之,却不够理想(如 Akiyo).其次,本文算法最多经过 4 次迭代就可达到与最优解相差不到 2% 的 PSNR,而经过 1~2 次迭代即可取得明显高出块匹配 FS 的补偿质量.这表明,本文算法的阻尼步长和更新因子符合匹配误差函数的分布特点,能够高效地逼近其最优解。

图 8 给出了 Husky 序列第 2 帧以(48,304)像素为左上角的宏块和 Foreman 序列第 2 帧以(0,336)像素为左上角的宏块在第 2~15 次迭代的误差曲线(第 1 次迭代的误差较大,为了清晰地展现曲线变化的细节,这里未予绘制),以及在此过程中本文算法的  $\delta$  值与  $\lambda$  值的变化情况.不难发现,尽管传统 L-M 算法和基于对角优势矩阵的 L-M 算法<sup>[75]</sup>的运动补偿误差均出现了振荡,但是本文算法却能保持稳定的下降趋势.并且,随着算法的收敛, $\delta$  与  $\lambda$  也逐渐趋于定值,从而保证了本文的改进 L-M 算法在靠近全局最优解时表现出高斯-牛顿法的快速收敛特性。

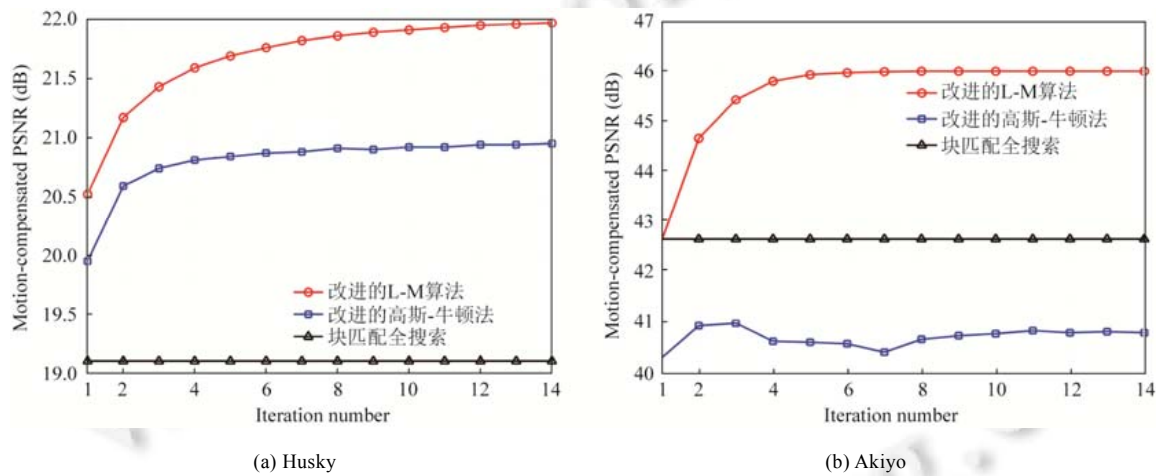


Fig.7 Comparison between our method and modified Gauss-Newton method in terms of convergence speed

图 7 本文算法与改进的高斯-牛顿算法的收敛速度比较

## 7.3 计算复杂度分析

下面对比分析基于块平移模型的全搜索、基于改进高斯-牛顿法的弹性运动估计和基于改进 L-M 法的弹性运动估计处理 1 个宏块的计算量.设块尺寸为  $B \times B$  像素,搜索窗口为  $W \times W$  像素,弹性基函数为  $p$  个。

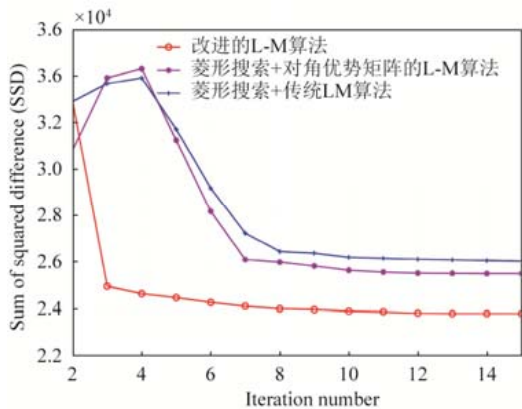
首先,基于块平移模型的全搜索的计算复杂度为  $O(B^2W^2)$ 。

其次,根据文献[62],基于改进高斯-牛顿法的弹性运动估计的计算复杂度为  $O(Tp^2B^2+73B^2)$ 。

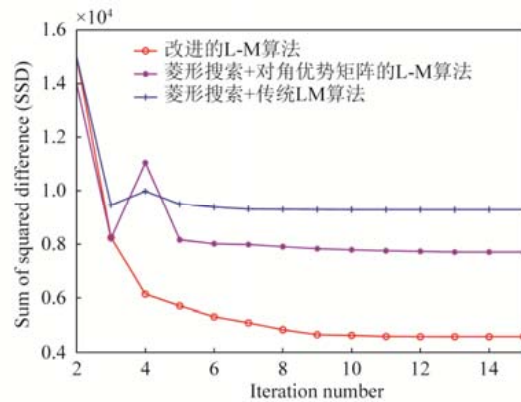
最后,基于改进 L-M 法的弹性运动估计需进行 1 次菱形搜索,其复杂度约为块匹配 FS 的 6%<sup>[6]</sup>,即  $O(0.06B^2W^2)$ ;每次迭代需计算对角矩阵阻尼系数、黑塞矩阵、逆矩阵、矩阵乘法、双线性插值和运动补偿误差,结合第 4 节的快速实现,其计算复杂度分别为  $O(1)$ 、 $O(0.375p^2B^2)$ 、 $O(p^3)$ 、 $O(pB^2)$ 、 $O(8B^2)$ 和  $O(B^2)$ , $T$  次迭代的渐近时间复杂度为  $O(0.375Tp^2B^2)$ 。

具体地,当  $B=16$ ,  $W=33$ ,  $T=15$ ,  $p=8$  时,本文算法的总计算量约为基于块平移模型的全搜索的 65%,为基于改进高斯-牛顿法的弹性运动估计的 51%。此外,由第 7.2 节可知,本文算法一般仅需 1~2 次迭代就能取得高于两者的 PSNR,从这个意义上讲,基于改进 L-M 法的弹性运动估计的计算量约是块匹配 FS 的 9.9%~13.9%,是基于改进

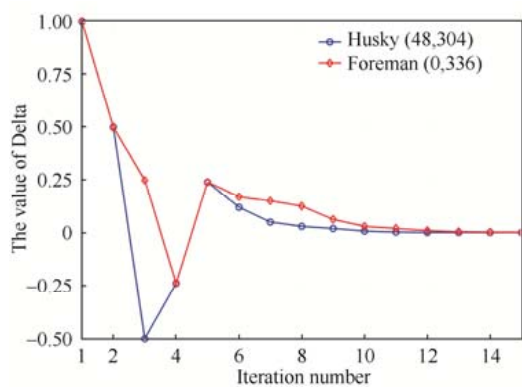
高斯-牛顿法的弹性运动估计的 7.8%~10.8%.



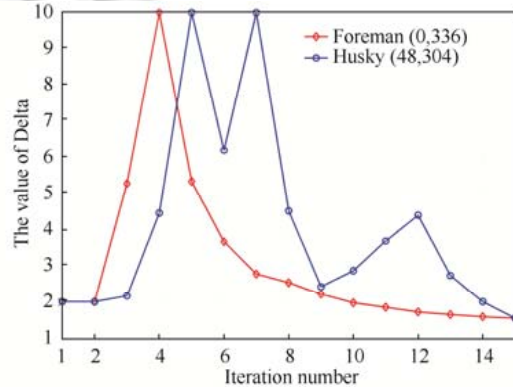
(a) Husky 第 2 帧(408,304)宏块的运动补偿误差变化曲线



(b) Foreman 第 2 帧(0,336)宏块的运动补偿误差变化曲线



(c) 1~15 次迭代过程中  $\delta$  的变化曲线



(d) 1~15 次迭代过程中  $\lambda$  的变化曲线

Fig.8 Curves of motion-compensated difference,  $\delta$ , and  $\lambda$  of the proposed method in first 15 iterations

图 8 本文算法在前 15 次迭代过程中的运动补偿误差、 $\delta$ 和 $\lambda$ 的曲线

### 8 结 论

弹性运动模型是一种表示非刚性复合运动的有效方法,现有工作均采用高斯-牛顿法进行求解,但仍存在迭代计算量高、收敛不稳定的不足.为避免高斯-牛顿黑塞矩阵出现奇异或病态,本文引进 L-M 优化算法求解弹性运动模型,并进行了两方面的改进:第一,根据黑塞矩阵的数值分布,给出一种快速计算黑塞矩阵的方法,降低了多次迭代所需的计算量;第二,通过理论和实验探讨了 L-M 对角矩阵阻尼系数的更新因子对弹性运动估计的影响,设计了一种自适应交替更新策略.在此基础上,提出一种基于改进 L-M 法的视频弹性运动估计算法.实验结果验证了算法的有效性.同时,本文算法对于视频处理中相关的无约束最优化问题求解具有一定的参考价值.

另外,本文方法仍存在若干问题可臻完善,如减少逆矩阵的反复计算、快速预测初始迭代点等,我们将在今后的工作中进一步深入研究这些问题的解决思路.

### 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, Sühring 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] Kuo CM, Kuan YH, Hsieh CH, Lee YH. A novel prediction-based directional asymmetric search algorithm for fast block matching motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2009,19(6):893–898. [doi: 10.1109/TCSVT.2009.2017420]
- [8] 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]
- [9] 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]
- [10] 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]
- [11] Nguyen T, Nguyen P, Dinh C. A novel search pattern for motion estimation in high efficiency video coding. In: Arumugam S, ed. *Proc. of the Int'l Conf. on Computer Communication and Informatics*. Piscataway: IEEE Computer Society Press, 2016. 1–6. [doi: 10.1109/CCCI.2016.7479971]
- [12] Medhat A, Shalaby A, Sayed MS, Elsabrouty M, Mehdipour F. Adaptive low-complexity motion estimation algorithm for high efficiency video coding encoder. *IET Image Processing*, 2016,10(6):438–447. [doi: 10.1049/iet-ipr.2015.0666]
- [13] 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]
- [14] Lengwehasatit K, Ortega A. Probabilistic partial-distance fast matching algorithms for motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 2001,11(2):139–152. [doi: 10.1109/76.905981]
- [15] 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]
- [16] 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).
- [17] 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]
- [18] 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]
- [19] Saha A, Mukherjee J, Sural S. New pixel-decimation patterns for block matching in motion estimation. *Signal Processing: Image Communication*, 2008,23(10):725–738. [doi: 10.1016/j.image.2008.08.004]
- [20] 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]
- [21] 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]

- [22] 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]
- [23] Yang LB, Yu KM, Li J, Li SP. An effective variable block-size early termination algorithm for H.264 video coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 2005,15(6):784–788. [doi: 10.1109/TCSVT.2005.848306]
- [24] 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]
- [25] 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]
- [26] 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]
- [27] 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.
- [28] 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]
- [29] 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]
- [30] 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]
- [31] 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]
- [32] 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]
- [33] 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]
- [34] 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]
- [35] 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]
- [36] 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]
- [37] Park CS. Level-set-based motion estimation algorithm for multiple reference frame motion estimation. *Journal of Visual Communication and Image Representation*, 2013,24(8):1269–1275. [doi: 10.1016/j.jvcir.2013.08.008]
- [38] 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]
- [39] Ahmmed A, Hannuksela MM, Gabbouj M. Fisheye video coding using elastic motion compensated reference frames. In: Pereira F, Sharma G, eds. *Proc. of the Int'l Conf. on Image Processing.* Piscataway: IEEE Computer Society Press, 2016. 2027–2031. [doi: 10.1109/ICIP.2016.7532714]
- [40] 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).
- [41] Ahmmed A, Xu R, Naman AT, Alam MJ, Pickering M, Taubman D. Motion segmentation initialization strategies for bi-directional inter-frame prediction. In: Florencio D, Pérez-González F, eds. *Proc. of the Int'l Workshop on Multimedia Signal Processing.* Piscataway: IEEE Computer Society Press, 2013. 58–63. [doi: 10.1109/MMSP.2013.6659264]
- [42] 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]



- [43] 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]
- [44] 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]
- [45] 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]
- [46] 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).
- [47] 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]
- [48] 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]
- [49] 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).
- [50] 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).
- [51] 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]
- [52] 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]
- [53] 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]
- [54] 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: 0.1109/TCSVT.2013.2254977]
- [55] Ren L. The application of deformable block matching algorithm [MS. Thesis]. Tianjin: Tianjin University, 2007 (in Chinese with English abstract).
- [56] 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).
- [57] 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]
- [58] 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]
- [59] 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]
- [60] 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]
- [61] Deng BY. Study on video compression based on elastic motion model [MS. Thesis]. Xiamen: Xiamen University, 2014 (in Chinese with English abstract).
- [62] 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]
- [63] Zhao CW. Research on motion estimation algorithm using elastic model [MS. Thesis]. Dalian: Liaoning Normal University, 2015 (in Chinese with English abstract).

- [64] Haque MN, Biswas M, Pickering MR, Frater MR. A low complexity algorithm for global motion parameter estimation targeting hardware implementation. In: Shi H, Zhang YC, Bottema MJ, Lovell BC, Maeder AJ, eds. Proc. of the Digital Image Computing: Techniques and Applications. Piscataway: IEEE Computer Society Press, 2009. 1–5. [doi: 10.1109/DICTA.2009.11]
- [65] 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]
- [66] 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]
- [67] Gao L. Numerical Optimization Method. Beijing: Peking University Press, 2014 (in Chinese).
- [68] 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.f0]
- [69] Levenberg K. A method for the solution of certain nonlinear problems in least squares. Quarterly of Applied Mathematics, 1944,2(2):164–168. [doi: 10.1090/qam/10666]
- [70] Marquardt DW. An algorithm for least squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, 1963,11(2):431–441.
- [71] Yamashita N, Fukushima M. On the rate of convergence of the Levenberg-Marquardt method. In: Alefeld G, Chen XJ, eds. Topics in Numerical Analysis, with Special Emphasis on Nonlinear Problems. Vienna: Springer-Verlag, 2001. 239–249. [doi: 10.1007/978-3-7091-6217-0\_18]
- [72] Suratgar AA, Tavakoli MB, Hoseinabadi A. Modified Levenberg-Marquardt method for neural networks training. Int'l Journal of Computer and Information Engineering, 2007,1(6):1745–1747.
- [73] Fan JY, Yuan YX. On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption. Computing, 2005,74(1):23–39. [doi: 10.1007/s00607-004-0083-1]
- [74] Chen TC, Han DJ, Au FTK, Tham LG. Acceleration of Levenberg-Marquardt training of neural networks with variable decay rate. In: Wunsch II DC, Hasselmo ME, Venayagamoorthy GK, Wang D, eds. Proc. of the Int'l Joint Conf. on Neural Networks. Amsterdam: Elsevier Science, 2003. 1873–1878. [doi: 10.1109/IJCNN.2003.1223693]
- [75] Kwak YT, Hwang JW, Yoo CJ. A new damping strategy of Levenberg-Marquardt algorithm for multilayer perceptrons. Neural Network World, 2011,21(4):327–340. [doi: 10.14311/NNW.2011.21.020]
- [76] Qi LY, Xiao XT, Zhang LW. A parameter-self-adjusting Levenberg-Marquardt method for solving nonsmooth equations. Journal of Computational Mathematics, 2016,34(3):317–338. [doi: 10.4208/jcm.1512-m2015-0333]
- [77] Transtrum MK, Sethna JP. Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization. 2012. <http://arxiv.org/abs/1201.5885>
- [78] Li MH, Bai M, Lü YJ. Application of LM algorithm in parameter estimation for second-order over-damped system. Process Automation Instrumentation, 2016,36(7):90–93 (in Chinese with English abstract).

#### 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]
- [16] 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.
- [33] 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]
- [40] 刘东, 吴枫. 视频编码技术发展及趋势. 中国计算机学会通信, 2016,12(6):20–23.
- [46] 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.
- [49] 金晟毅. 基于运动对象网格模型的运动估计方法研究[硕士学位论文]. 西安: 西北工业大学, 2006.
- [50] 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.
- [55] 任亮. 可变形状匹配算法(DBMA)的应用研究[硕士学位论文]. 天津: 天津大学, 2007.
- [56] 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.
- [61] 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.

- [62] 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]
- [63] 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.
- [67] 高立.数值最优化方法.北京:北京大学出版社,2014.
- [78] 李敏花,柏猛,吕英俊.LM 算法在二阶过阻尼系统参数估计中的应用.自动化仪表,2016,36(7):90-93.



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.



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.



闵新(1992—),男,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,计算机图形学.



闫小红(1993—),女,硕士生,主要研究领域为视频编码.