

文章编号: 1001-0920(2012)10-0000-00

## 电子稳像中基于运动分类的全局运动估计算法

朱娟娟, 郭宝龙, 吴宪祥

(西安电子科技大学 机电工程学院, 西安 710071)

**摘要:** 提出一种基于运动分类的全局运动估计算法. 首先, 分区提取图像中的鲁棒 Harris 特征点, 并采用特征窗匹配思路, 提高匹配速度; 其次, 对运动矢量在平移、旋转和缩放模式下的统计特性进行分析, 提出运动类型快速判定方法, 并验证特征点的有效性; 接着, 将有效点对代入运动方程, 求取全局运动参数; 最后, 结合 Kalman 滤波来补偿当前帧实现视频稳像. 实验表明, 该算法能够处理含摄像机扫描和抖动的复杂场景, 检测误差小于 0.5 像素, 且达到实时处理.

**关键词:** 电子稳像; 全局运动估计; 运动补偿; 特征点

**中图分类号:** TP391.4

**文献标识码:** A

## Electronic image stabilization based on global motion estimation with motion classification

ZHU Juan-juan, GUO Bao-long, WU Xian-xiang

(School of Electromechanical Engineering, Xidian University, Xi'an 710071, China. Correspondent: ZHU Juan-juan, E-mail: zhujoo@126.com)

**Abstract:** A global motion estimation algorithm based on motion classification is presented. Firstly, the Harris feature points are selected evenly and matched by using feature window. Hence, the statistic features of all motions are analyzed according to different motion kinds including translation, rotation and zoom. Then, the fast motion classification method is proposed to validate all points. Thirdly, the remained global feature points are brought to the affine model to compute global motion. Finally, the Kalman filter is used to compensate each current frame. Experimental results show that the algorithm can correctly detect global motion in dynamic scenes with camera scan and various dithering. The estimation error is below 1/2 pixel at real-time stabilization, which can greatly improve the stability and fidelity of videos.

**Key words:** electronic image stabilization; global motion estimation; motion compensation; feature points

### 1 引言

在移动载体摄像系统中, 载体的抖动导致视频序列出现晃动甚至模糊现象, 因而需要进行稳定处理. 电子稳像<sup>[1]</sup>就是利用数字图像处理的方法获取并去除视频帧间的干扰, 提高视频的稳定性, 为后续目标跟踪与识别<sup>[2]</sup>、分类、行为理解等提供基础, 在军事领域, 民用航测等系统中具有广泛的应用.

电子稳像的核心问题是全局运动估计和抖动参数补偿, 而前者直接决定了稳像精度. 传统的块匹配算法<sup>[3]</sup>、投影法<sup>[4]</sup>、位平面法<sup>[5]</sup>等均只能处理平移抖动, 且计算精度不高. 因此许多基于点特征的算法相继提出: 若仅在固定位置选取代表点<sup>[6]</sup>, 则会出现误匹配; 而采用尺度不变特征检测<sup>[7]</sup>, 其检测过程复杂,

难以实时处理; 另外, 若将局部点或误匹配点直接代入计算, 则会降低参数的精度<sup>[8]</sup>.

本文结合摄像机各运动的特点来分析特征点运动的统计特性, 并提出一种基于运动分类的全局运动估计算法(MCME). 摄像机在拍摄过程中, 具有方向和幅值的一致性, 对此研究在平移、旋转和缩放等复杂运动的拍摄场景中, 如何选取鲁棒的特征点并进行快速匹配, 且克服场景中局部运动和误匹配造成的不利影响, 从而提高全局运动估计的精度.

### 2 基于运动分类的全局运动估计 MCME

#### 2.1 特征点提取

采用经典的 Harris 算子<sup>[9]</sup>提取参考帧的特征点,

收稿日期: 2011-04-02; 修回日期: 2011-11-09.

基金项目: 国家自然科学基金青年基金项目(61003196); 高校基本科研业务费专项资助项目(K50510040004).

作者简介: 朱娟娟(1980—), 女, 副教授, 从事智能控制的研究; 郭宝龙(1962—), 男, 教授, 博士生导师, 从事人工智能、图像工程研究.

为了避免特征分布过于集中, 本文采用改进的分区提取 Harris 特征点的方法, 即: 将图像分成互不重叠的区域, 在各区域内选取响应最大的点作为特征点. 实验表明, 本文分区提取的  $6 \times 6$  个特征点均匀分布在整幅图像中, 绝大多数位于物体边缘处, 具有较好的独特性, 且信息量丰富.

## 2.2 特征点匹配及坐标关系

所谓特征点匹配, 就是在当前帧找到参考帧中各特征点的唯一匹配点. 本文采用特征窗的相似度测量法: 以参考帧的特征点为例, 在其邻域构造模板图像作为特征窗, 并在当前帧的搜索区域内, 采用菱形法<sup>[10]</sup>和最小绝对误差和准则<sup>[4]</sup>, 找到匹配特征窗, 其中心点即为匹配特征点. 其坐标转换为

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} m_0 & m_1 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} m_2 \\ m_5 \end{bmatrix}, \quad (1a)$$

记

$$M^T_{\text{affine}} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \end{bmatrix}. \quad (1b)$$

其中:  $m_2$  和  $m_5$  表示两图的平移量;  $m_0, m_1, m_3$  和  $m_4$  表示尺度和旋转量;  $M_{\text{affine}}$  为 6 参数仿射矩阵.

## 2.3 基于运动分类的全局运动参数计算

### 2.3.1 运动类型快速判定

假定摄像机拍摄的是静止场景, 且场景的深度变化不明显. 此时, 帧间运动类型主要分为 5 种: 静止、平移、缩放、旋转和不规则运动. 令  $[x, y]^T$  和  $[\hat{x}, \hat{y}]^T$  为匹配特征点对, 设  $MV_{i,j} = (\Delta x, \Delta y)$  表示其水平和垂直运动矢量, 即  $\Delta x = x_i - \hat{x}_i, \Delta y = y_i - \hat{y}_i$ . 在判定过程中, 取矢量的两个特征值: 方向  $\Delta x / \Delta y$  和幅值  $\sqrt{\Delta x^2 + \Delta y^2}$ .

#### 1) 静止.

计算幅值  $\sqrt{\Delta x^2 + \Delta y^2}$ , 判断是否小于某一阈值, 如果是就认为该子块为小运动块. 如果小运动矢量达到  $1/2$ , 则可判定当前帧间运动为静止, 否则继续进行其他运动类型的判定.

#### 2) 平移.

如果摄像机作横扫、竖扫、俯仰、摇头等运动, 则帧间所有子运动矢量应具有良好的幅值和方向一致性, 如图 1(a) 所示. 计算每一块的幅值和方向, 如果经统计后, 超过一半的运动矢量具有方向和幅值一致性, 则可判定是平移运动.

#### 3) 缩放.

图 1(b) 为摄像机变焦、推拉运动时的运动矢量分布, 几乎所有的运动矢量都指向扩张焦点 (FOE). 为了快速判断帧间是否存在缩放运动, 只需判断是否存在 FOE. 在排除平移运动后, 判断运动矢量场的左右

边界对称位置处是否存在  $\Delta y$  为零的运动矢量, 则沿着该对称线自左向右寻找  $\Delta x$  发生符号改变的位置为 FOE 位置.

#### 4) 旋转.

图 1(c) 为摄像机绕  $Z$  轴旋转的运动矢量分布, 几乎所有的运动矢量都绕 FOE 成圆周分布. 在排除平移运动后, 判断运动矢量场的上下边界对称位置处是否存在  $\Delta y$  为零的运动矢量, 则沿着该对称线自上向下寻找  $\Delta x$  发生符号改变的位置为 FOE 位置.

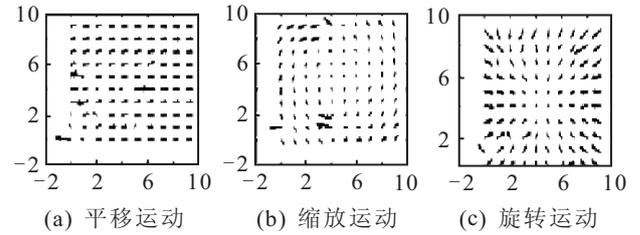


图 1 各种运动矢量分布图

根据以上分析, 帧间运动类型判定的流程为: 首先判断是否为静止帧; 其次判断平移运动; 接着, 判断是否存在 FOE 区域, 区分缩放运动还是旋转运动; 最后, 当所有情况都不满足时, 则判定为无规则运动.

### 2.3.2 基于运动类型的特征点预处理

#### 1) 静止.

在这种情况下, 直接输出全局运动估计结果, 即

$$M^T_{\text{affine}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

#### 2) 水平或垂直偏移.

本文建立了特征点有效性的概念: 如果特征点选在局部运动物体上, 特征点移出场景, 被遮挡或者纹理相似导致的误匹配等, 则将其定义为无效特征点; 而那些位于背景区域且正确匹配的则定义为有效特征点. 为了区分特征点的有效性, 提出距离一致性准则以验证特征匹配, 匹配验证步骤如下:

**Step 1:** 计算  $N$  个特征点对的纵横坐标之间的距离  $D(i) = \sqrt{\Delta x^2 + \Delta y^2}$ ;

**Step 2:** 将距离分为  $l$  ( $l \leq N$ ) 类;

**Step 3:** 逐个比较和统计每一距离类中特征点的个数  $c_i, i = 1, 2, \dots, l$ ;

**Step 4:** 判断  $c = \max\{c_i | i = 1, 2, \dots, l\}$ , 那么该距离所对应的那些特征匹配具有良好的距离一致性, 判断为有效特征点并作为候选全局特征点对.

#### 3) 缩放或旋转.

在帧间发生光轴方向比例缩放, 其运动矢量呈规则对称特性, 如图 2(a) 所示. 为了计算简便, 将  $6 \times 6$  个运动矢量分为空间对称的 4 组, 可判断出左上组 2a 中存在方向不一致的运动矢量, 则删除对应特征点, 剩

余特征点对作为候选全局特征点对。

在摄像机绕光轴旋转时, 其运动矢量呈规则对称特性, 如图2(b)所示. 将运动矢量按照圆周分为3组, 可判断出第2组中存在不一致的运动矢量, 则删除对应特征点, 剩余特征点对作为候选全局特征点对。

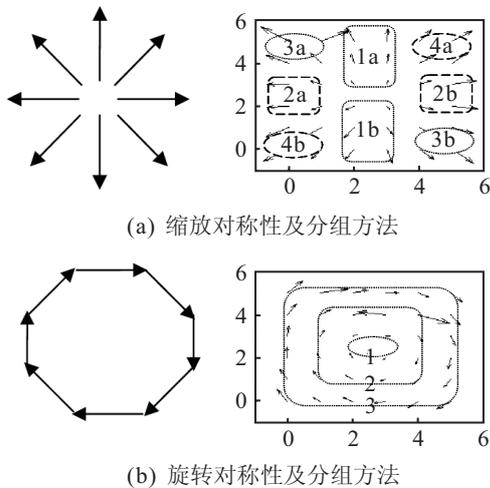


图2 运动矢量特性及分组分析方法

#### 4) 不规则运动。

当图像存在平移、旋转和缩放的复合运动, 或者FOE不在图像中间位置, 运动矢量场不存在明显的一致性, 则所有特征点均作为候选全局特征点对。

### 2.3.3 全局运动参数计算

为了提高参数估计精度, 达到各特征点的位置误差最小, 本文提出了基于求最小二乘解的迭代步骤. 首先, 计算候选全局特征点的坐标差  $\Delta x_i = x_i - \hat{x}_i$ ,  $\Delta y_i = y_i - \hat{y}_i$ , 作为初始数据集  $S$ ; 其次, 将  $S$  中对应的特征点代入运动模型 (1a), 建立仿射方程  $B = AM_{\text{affine}}$ , 即

$$\underbrace{\begin{bmatrix} \hat{x}_1 & \hat{y}_1 \\ \hat{x}_2 & \hat{y}_2 \\ \vdots & \vdots \\ \hat{x}_c & \hat{y}_c \end{bmatrix}}_B = \underbrace{\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_c & y_c & 1 \end{bmatrix}}_A \times \underbrace{\begin{bmatrix} m_0 & m_3 \\ m_1 & m_4 \\ m_2 & m_5 \end{bmatrix}}_{M_{\text{affine}}}, \quad (2)$$

以求最小二乘解  $M_{\text{affine}} = (A^T A)^{-1} A^T B$ ; 其次, 利用求出的  $M_{\text{affine}}$  计算参考帧中各特征点映射后的坐标  $(\bar{x}_i, \bar{y}_i)$ , 并计算其与参考帧中原提取特征点的坐标差

$$\Delta \bar{x}_i = x_i - \bar{x}_i, \Delta \bar{y}_i = y_i - \bar{y}_i;$$

接着, 计算位置误差

$$E_i = \sqrt{(\Delta x_i - \Delta \bar{x}_i)^2 + (\Delta y_i - \Delta \bar{y}_i)^2},$$

将  $S$  中误差大于阈值的特征点对删除并重算  $M_{\text{affine}}$ , 若误差小于阈值, 则停止迭代并输出  $M_{\text{affine}}$ 。

### 3 帧间图像补偿

在对图像进行补偿之前, 需要对全局运动矢量中的平稳分量 (代表扫描运动) 和不稳定分量 (代表抖

动) 进行区分, 即对运动矢量信号进行滤波. 传统的滤波方法, 如:  $B$  样条曲线拟和法<sup>[11]</sup>易出现过稳或欠稳现象; 惯性滤波法<sup>[12]</sup>需要设置参数, 不能自适应调整; 粒子滤波<sup>[13]</sup>过程复杂, 延迟较大。

对此, 本文选用 Kalman 滤波器来进行实时稳定运动偏移. 对于待补偿帧  $K$ , 记录其相对于第 0 帧的累积运动偏移  $Z_{\text{raw}}(k)$ , 使用 Kalman 滤波器对其绝对帧位置进行平滑滤波, 从而去除运动矢量信号中的高频抖动分量. 滤波后保留的平滑分量很好地代表了摄像系统的有意运动, 记为  $Z_{\text{Kal}}(k)$ . 因此, 修正矢量为

$$Z_{\text{cor}}(k) = Z_{\text{Kal}}(k) - Z_{\text{raw}}(k).$$

当确定各当前帧  $K$  的补偿参数后, 将其带入变换矩阵  $M_{\text{affine}}$ , 对图像进行变换, 变换时采用双线性插值来计算该点的像素值, 从而获得补偿结果图像。

## 4 实验结果及分析

实验是利用 VC++ 编程进行, 摄像机被安放在一个抖动平台上, 其水平和垂直振幅分别为 4 mm 和 20 mm, 频率为 5 Hz, 帧间比例缩放不超过 30%, 旋转角度不超过 15°。

### 4.1 特征点选取和匹配的实验结果及分析

对于抖动平台拍摄的视频序列 1, 其背景较复杂且存在局部运动物体, 图 3(a) 为参考帧中均匀提取的 6×6 个特征点, 图 3(b) 为当前帧中匹配的特征点; 对于选自航拍视频序列 2 中任两帧, 其背景复杂且存在比例变化, 但不存在运动物体, 图 3(c) 和图 3(d) 为其特征点检测和匹配结果图。

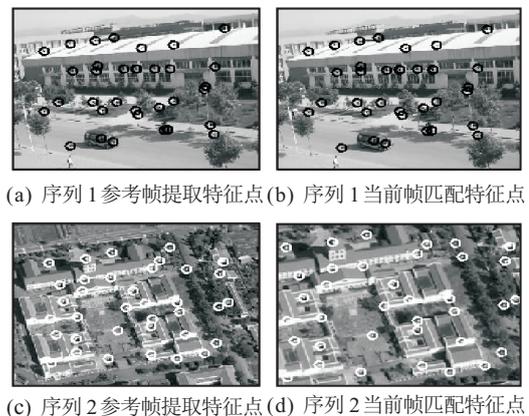


图3 序列中特征点提取和匹配结果图

### 4.2 特征点预处理的实验结果及分析

视频序列 1 中的摄像机属于平移状态, 图 4(a) 为其运动矢量分布图, 图 4(b) 为利用距离准则验证之后的结果, 其中标注“◇”的特征点为判定无效的特征点, 属于局部运动 (左下区域的运动车辆). 视频序列 2 中的摄像机属于变焦运动状态, 图 4(c) 为其运动矢量分布图, 图 4(d) 为利用方向一致性验证, 将标记“◇”的误

匹配点进行删除。

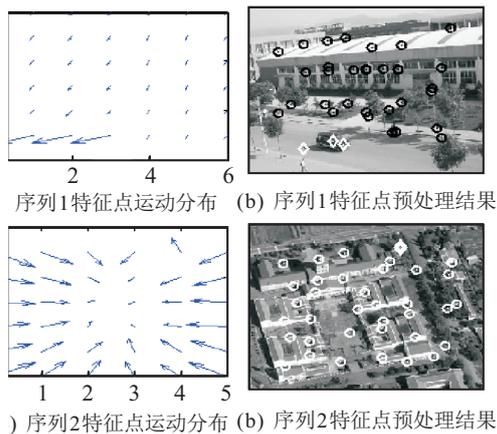


图 4 特征点预处理结果

表 1 给出了运动分类的预处理时间比较. 由表 1 可以看出, 通过分类进行预处理可有效减少矩阵计算的迭代次数, 提高运动估计的速度. 该方法对于视频场景中存在较大前景运动物体, 可快速删除无效特征点, 从而避免局部运动特征点对全局运动参数计算时间的耗费, 以及精度的不良影响. 事实上, 当对视频序列初始几帧进行运动类型判定后, 在短暂时间内 (实验设定为 100 帧) 可不用逐帧判定.

表 1 预处理时间和迭代次数比较结果

实验图片	运动类型	预处理迭代次数		类型判定+验证时间/(ms/帧)
		直接	分类	
No.1	平移	5	2	1.78+0.92
No.2	缩放	2	0	2.12+0.87

### 4.3 全局运动参数的计算结果及分析

为了验证特征点预处理对提高参数精度的特性, 求取全局运动参数时, 一是将所有特征点代入运动模型 (1) 直接计算最小二乘解<sup>[8]</sup>; 二是经特征点预处理后迭代估计参数矩阵; 同时, 利用所求参数补偿当前图像并计算帧间峰值信噪比 (PSNR) 来评价补偿前后的帧间相似度.

表 2 给出了无迭代和迭代估计得到的运动参数矩阵, 其误差超过了 1 个像素.

表 2 全局运动参数比较结果

实验图片	无迭代估计	迭代估计
No.1	$\begin{bmatrix} 1.000 & -0.054 \\ 0.0508 & 0.9822 \\ -20.036 & 11.0341 \end{bmatrix}$	$\begin{bmatrix} 1.0138 & -0.044 \\ 0.0610 & 0.9809 \\ -19.5840 & 11.6860 \end{bmatrix}$
No.2	$\begin{bmatrix} 1.0012 & -0.041 \\ 0.0011 & 1.0200 \\ -0.0223 & 1.0035 \end{bmatrix}$	$\begin{bmatrix} 0.9993 & -0.0027 \\ 0.0004 & 1.0000 \\ -1.0190 & 0.1927 \end{bmatrix}$

由表 3 可见, 迭代估计时间尽管有所增加, 但依然满足每秒处理 25 帧的要求; 且 PSNR 值的提高更具

优势, 比原始 PSNR 平均提高 46.5%, 比无迭代法获得的 PSNR 提高 4.75%, 进一步说明 MCME 算法具有较高的全局运动估计精度. 并且, 实验中采用人工设定参数合成图片后再检测的方法, 经测试, 全局运动估计的误差小于 1/2 像素.

表 3 全局运动估计时间和精度比较结果

实验图片	时间/ms		原始 PSNR/dB	补偿后 PSNR/dB
	无迭代	迭代		
No.1	18.33	28.45	19.87	27.85/29.32
No.2	19.40	21.89	22.96	31.96/33.29

### 4.4 视频稳定前后 PSNR 值的比较结果

图 5 给出了原始序列中连续 38 帧的帧间逼真度 PSNR 和稳定序列的 PSNR 的比较结果. 由图 5 可以看出, 稳定序列的帧间逼真度有很大的提高, 说明稳定后的图像序列帧间偏差变小, 达到了一定的稳定效果.

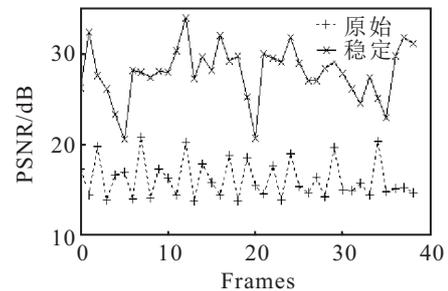


图 5 序列帧间 PSNR 比较结果

## 5 结论

本文针对电子稳像中摄像系统的复杂运动, 提出了一种基于运动分类的全局运动估计算法. 该算法通过运动统计特性对特征点进行预处理, 从而提高全局运动估计的精度. 测试结果表明, MCME 算法能达到 25 fps 的实时处理速度, 且处理后帧间相似度获得提高, 从而实现了稳像处理的效果. 今后将针对电子稳像的应用作进一步研究.

### 参考文献 (References)

- [1] Jinhee Lee, Sangkeun Lee, Joonki Paik. Digital image stabilization based on statistical selection of feasible regions[J]. IEEE Trans on Consumer Electronics, 2009, 55(4): 1748-1755.
- [2] 邢卓异, 朱齐丹, 林壮. 基于 Gabor 小波的目标跟踪方法研究[J]. 控制与决策, 2008, 23(2): 162-166. (Xing Z Y, Zhu Q D, Lin Z. Target recognizing and tracking algorithm based on Gabor wavelet[J]. Control and Decision, 2008, 23(2): 162-166.)
- [3] Tico M. Adaptive block-based approach to image stabilization[C]. 15th IEEE Int Conf on Image Processing. 2008: 521-524.

- [4] Moshe Y, Hel-Or H. Video block motion estimation based on gray-code kernels[J]. *IEEE Trans on Image Processing*, 2009, 18(10): 2243-2254.
- [5] Ko S J, Lee S H, et al. Fast digital image stabilizer based on gray-coded bit-plane matching[J]. *IEEE Trans on Consumer Electronics*, 1999, 45(3): 598-603.
- [6] Tang J, Han X W, Yuan Zh H. An approach of electronic image stabilization based on the representative point matching[C]. *Third Int Conf on Genetic and Evolutionary Computing*. 2009: 347-350.
- [7] David G Lowe. Distinctive image features from scale-invariant key-points[J]. *Int J of Computer Vision*, 2004, 60(2): 91-110.
- [8] Yao Shen, Guturu P, Damarla T, et al. Video stabilization using principal component analysis and scale invariant feature transform in particle filter framework[J]. *IEEE Trans on Consumer Electronics*, 2009, 55(3): 1714-1721.
- [9] Harris C, Stephens M. A combined corner and edge detector[C]. *Fourth Alvey Vision Conf*, 1988: 17-151.
- [10] Zhu S, Ma K K. A new diamond search algorithm for fast block-matching motion estimation[J]. *IEEE Trans on Image Processing*, 2000, 9(2): 287-290.
- [11] 孟龙, 林行刚, 等. 视频抖动矫正系统中的运动滤波[J]. *清华大学学报: 自然科学版*, 2005, 45(1): 41-43.  
(Meng L, Lin X G, et al. Motion filter for video stabilizing systems[J]. *J of Tsinghua University: Sci & Tech*, 2005, 45(1): 41-43.)
- [12] Yang S H, Jheng F M, Cheng Y C. Two dimensional adaptive image stabilization[J]. *Electronics Letters*, 2007, 43(8): 446-448.
- [13] Junlan Yang, Schonfeld D, Mohamed M. Robust video stabilization based on particle filter tracking of projected camera motion[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2009, 19(7): 945-954.