

文章编号: 1003-501X(2011)03-0138-07

多尺度空间基于集中度判定的二维稳像算法

王 洪^{1,2}, 戴 明¹, 柏旭光¹

(1. 中国科学院长春光学精密机械与物理研究所, 长春 130033;
2. 中国科学院研究生院, 北京 100039)

摘要: 二轴运动平台由于震动、非一致性摩擦等因素造成视频序列出现抖动, 同时外部噪声, 光照变化等对图像特征提取、匹配等存在严重影响导致运动向量估计错误, 本文提出一种在多尺度空间基于集中度判定的二维稳像算法。通过多尺度空间下提取视频序列的不变特征, 采用深度优先最邻近搜索算法, 寻找匹配点对, 然后计算匹配点对的集中度, 通过最优集中化原则, 建立不同层次的计算模型估算二维运动空间的补偿参数, 实现抖动视频的亚像素精度补偿。实验采用 320×240 pixels 灰度图像序列测试, 本文算法稳像精度高, 运动估计时间仅为随机抽样一致性算法的 13%, 有效提升了整个稳像算法的性能。

关键词: 电子稳像; 集中度; 尺度空间; 不变特征; 运动估计

中图分类号: TP391

文献标志码: A

doi: 10.3969/j.issn.1003-501X.2011.03.025

Multi-scale Space Two Dimension High Precision Video Stabilization Based on Concentration Degree

WANG Hong^{1,2}, DAI Ming¹, BAI Xu-guang¹

(1. Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun 130033, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100039, China)

Abstract: The vibration and friction factors of two-axis motion platform cause video jitter and because of image feature extraction and matching influenced by illumination variation, noise and so on, the wrong motion vector will be estimated. The multi-scale space two dimension high precision video stabilization algorithm is proposed based on concentration degree. Invariant features are extracted, and then the depth first nearest neighbor search algorithm is used to get the matching double points. At last, concentration degree is computed, and different computing models are built to estimate compensation parameters by optimal concentration degree principle. 320×240 pixels gray video sequence is used in experiment, and the algorithm proposed in this paper has high precision. The estimation time is just 13% of Random Sample Consensus (RANSAC), which improves the performance of video stabilization.

Key words: electric image stabilization; concentration degree; scale space; invariant feature; motion estimation

0 引 言

对于机载、车载、船载等成像设备, 往往会因为载体受外界风力、大气压差、交通状况以及内部发动机等因素的影响, 使采集到的视频序列出现抖动, 无法友好地用于观察被侦查目标的实时细节, 影响指挥系统的决策。现阶段大部分成像设备均以二维机械平台为基础, 图像抖动主要以二维方向为主, 即平移抖

收稿日期: 2010-11-12; 收到修改稿日期: 2011-01-06

基金项目: 国家 863 高技术研究发展计划资助项目(2007AA12Z113); 973 国家自然科学基金项目(2009CB72400105)

作者简介: 王洪(1982-), 男(汉族), 四川射洪人。博士研究生, 主要研究工作是图像与信号处理, 嵌入式设计。E-mail: wedon121@163.com。

导师简介: 戴明(1965-), 男(汉族), 硕士, 博士生导师, 主要从事光电平台稳定技术和图像稳定技术及航空光电成像技术方面的研究。

E-mail: daim@vip.sina.com。

动。目前电子稳像算法主要采用基于特征^[1]、灰度信息^[2]以及滤波算法等进行运动估计, 获取运动补偿参数, 实现运动补偿。但是, 这些算法, 均在单一尺度空间, 获取的特征信息无法承受外部噪声, 光照等因素的影响, 使补偿精度无法保证。

LOWE^[3]等人利用 Lindeberg 等人对图像进行多尺度理论分析的基础上, 提出了旋转、缩放、平移等尺度不变特征(Scale Invariant Feature Transform, SIFT)算法, 具有丰富的多尺度信息特征, 能够抵抗外部光照变化、随机噪声的影响, 同时具有亚像素特征定位精度; 而后 Herbert Bay^[4]对 SIFT 算法进行了成功的改进, 提出了快速鲁棒特征(Speed-Up Robust Features, SURF)算法, 对 SIFT 算法进行了优化, 使得算法性能很大程度上提高, 能更好地压制噪声, 并为实时处理连续视频序列提供了可能。

本文利用 SURF 算法提取图像序列亚像素精度尺度不变特征, 并提出基于集中度原则的最优全局运动估计算法。通过对运动向量建立直方图, 并利用不同层次的数学模型进行参数估计, 最终实现运动参数快速、准确估计。实验利用 320×240 pixels 图像序列进行测试, 与 RANSAC(Random Sample Consensus)全局运动参数估计算法对比, 本文在保证高精度稳定的同时, 时间性能优势明显。

1 多尺度不变特征

1.1 不变特征提取

多尺度不变特征具有丰富梯度与曲率等信息, 特征周围点亮度往往变化剧烈, 同时存在一个最大的梯度方向, 并对于外界光照变化以及噪声不敏感。Lindeberg 等人^[5]建立了多尺度图像分析理论, 采用均值为 0, 方差为 σ^2 的高斯核函数 $G(x, y, \sigma)$ 作为最佳的尺度函数, 定义如下:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x^2 + y^2)}{\sigma^2}\right) \quad (1)$$

在 SIFT 算法中, 采用 $G(x, y, \sigma)$ 与图像进行迭代卷积运算, 得到高斯差函数(Difference of Gaussian, DOG)以建立不同尺度下的尺度空间, 在该 DOG 尺度空间中, 寻求极值点。而 Herbert Bay 等人提出了 SURF 算法, 在不改变图像尺度的前提下, 通过改变离散截断后的高斯函数与图像做迭代卷积运算, 建立尺度空间, 寻求在多尺度空间下的极值点, 相比于 SIFT 算法, 计算复杂度得以简化, 本文采用 SURF 算法进行特征提取。

特征定位利用具有仿射不变性与投影不变性的自相关 Hessian 矩阵特征值估算特征强度, 并通过阈值法选取出稳定的图像特征点, 对于图像中给定点 $X = (x, y)^T$, Hessian 矩阵定义如下:

$$\mathbf{H}(X, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (2)$$

$L(X, \sigma)$ 为图像 I 与高斯函数的卷积, $L_{xx}(X, \sigma)$ 为其二阶导数 $\frac{\partial^2 L(X, \sigma)}{\partial x^2}$, 同理 $L_{xy}(X, \sigma)$, $L_{yy}(X, \sigma)$ 。

特征点位置定位完成以后, 为使每一个特征点具有平移, 旋转, 缩放, 仿射以及投影不变性, 需对特征点进行描述, 建立特征描述向量。通常采用紧支撑一阶导 Harr 小波作用于积分图像 x - y 两个方向上的响应来提取, 计算 64 维空间描述向量。具体算法: 以兴趣点为中心, 半径为 $6s$ (s 为被提取特征点的尺度因子) 的区域, 利用 Harr 小波对积分图像进行快速滤波。在完成 Harr 小波变换以及兴趣点高斯加权处理以后, 可以获得水平响应强度与垂直响应强度, 通过计算窗口大小为 $\pi/3$ 的滑动窗内所有响应点之和可以获得特征点的主方向。接着以该兴趣点为中心, 以该点的主方向为主轴, 取窗口大小为 $20s$ 的区域, 并将该区域划分为 4×4 的子区域, 计算水平方向 dx , 垂直方向 dy 的 Harr 小波响应。在计算 $\sum dx$ 与 $\sum dy$ 的同时, 计算 $\sum |dx|$ 与 $\sum |dy|$, 这样构成了 4 维描述子向量 $\mathbf{v} = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$ 。计算 4×4 子区域内的所有向量, 最后将获得 64 维描述子。

1.2 特征匹配

特征匹配通过对待匹配点特征建立搜索树, 常见的如 K-D 树, 球形树, 词汇树等, 利用深度优先或宽度优先等原则, 寻找与待匹配点最邻近的特征点, 使得匹配点向量与待匹配点向量距离 d 最小。

$$d = |P_i - P_j|_L \quad (3)$$

其中: d 为 Lerburg 测度, P_i, P_j 分别为待匹配点集。

2 集中度判定算法

运动向量的集中度即指局部运动估计参数的密集程度, 是全局运动参数集中程度的测量指标, 可以用来衡量局部运动参数对全局运动参数的差异, 是估计全局运动参数的重要参数集。集中度的判定以概率统计为基础, 定义匹配点对为采样样本, 对于匹配点对 $\{p_1, p_2, \dots, p_n\}$, 其中 $p_i (i=1, \dots, n)$ 为 $\{p_{ij}(x_{\text{pos}}, y_{\text{pos}}, V_{\text{descriptor}}), (j=1, 2)\}$ ($x_{\text{pos}}, y_{\text{pos}}$ 表示特征点的位置信息, $V_{\text{descriptor}}$ 表示特征点的向量描述子), 利用匹配点对, 可以寻找抖动视频序列的二维运动向量 V :

$$V = p_{i_1} - p_{i_2} \quad (4)$$

将二维运动向量以统计直方图表示, 横轴代表特征点的运动量, 纵轴代表对应运动量的数目。假定二维运动向量的任意一维运动向量均满足正态分布:

$$f(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right) \quad (5)$$

其中 μ 为常数, 其意义为直方图分布的峰值位置 (即运动估计的最可能位置), σ 值一般较小, 因此, 正态分布函数的峰值较大, 而宽度较窄。如第三节中图 2 所示。对于样本集 V 中的成员变量视为观测值, 分布函数 $F(v)$ 可表示为

$$F(v) = \begin{cases} 0 & v < v_{(1)} \\ \sum_{j=1}^i f_j & v_{(i)} \leq v < v_{(i+1)} (i=1, 2, \dots, k-1) \\ 1 & v \geq v_{(k)} \end{cases} \quad (6)$$

其中: f_j 为 $v_{(j)}$ 出现的频率 $f_j = \frac{n_j}{n}$ (n 为样本总数, n_j 为某一个样本出现的频数)。

各个特征点的运动向量的分布状况可以用集中度进行简化, 对于密集程度不高, 过于发散, 认为当前帧没有全局运动参数。本文设计如下模型来判定集中度: 如果局部特征点运动参数变化范围 (LV_{range}) 大于动态阈值 (LV_{Thr}), 认为当前局部特征点运动参数集中度高, 否则发散, 不存在全局运动参数。

$$LV_{\text{DynThr}}[n] = \max \left\{ \begin{array}{l} 0.9 \times LV_{\text{Thr}}[n-1] \\ 1.25 \times \max \{ LV_{\text{range}}[n-m] \} \end{array} \right\} \quad (7)$$

其中: $LV_{\text{range}}[n] = LV_{\text{max}}[n] - LV_{\text{min}}[n]$, m 取过去的 1, ..., 5 帧, n 为当前帧。

$$LV_{\text{Thr}}[n] = \text{sat}(LV_{\text{DynThr}}[n], LV_{\text{Thr}}^{\text{max}}, LV_{\text{Thr}}^{\text{min}}) \quad (8)$$

其中: $LV_{\text{Thr}}^{\text{max}} = \text{round}(0.6 \times 2 \times LV_{\text{max}})$, $LV_{\text{Thr}}^{\text{min}} = \text{round}(0.35 \times 2 \times LV_{\text{max}})$, $LV_{\text{max}}[n]$ 为当前帧运动向量中的最大值, 为 LV_{max} ; $LV_{\text{min}}[n]$ 为当前帧运动向量中的最小值。利用 $LV_{\text{DynThr}}[n]$ 来表示运动向量的动态阈值。

在满足集中度判定的前提下, 全局运动参数估计按如下步骤进行:

1) 建立直方图: 当直方图中最大直方条的值大于等于其他所有直方条之和或大于阈值 $N/3$ (N 为匹配的特征点对数), 该直方条被确定为当前帧的全局运动参数, 若最大值小于等于 $N/4$ 时, 认为估计失效。

2) 加窗: 若大于 $N/4$ 时, 在建立好统计直方图后, 利用自适应窗对其进行累加运算, 其中窗口大小通过对先前估计的运动向量进行中值滤波预测得到, 将窗口内的所有非中心位置运动向量累加到窗口中心位置的运动向量上, 这样可以有效避免窗口内相邻直方条等高现象, 在完成对原始直方图的滑动窗运算后, 便产生了被加窗处理后的直方图, 若直方图中的最大值大于 $2N/3$, 认为该最大值为全局运动参数。

3) 累加: 如果第二步无法产生全局运动参数, 则选择高度最高的直方条为中心, 建立一个窗口区域, 窗口大小与先前一致, 将窗口内的所有直方条累加到窗口的中心直方条上, 一旦累加完成, 用于窗口内累加的直方条就不再参与运算, 若直方图中的最大值大于 $4N/3$, 则认为该最大值为全局运动参数。

4) 滤波: 若前 3 步均无法产生全局运动参数, 对直方图进行滤波:

$$LV_{\text{valid}}^{\text{filter}}[n][id] = \sum_{k=0}^4 lk \times LV_{\text{valid}}[n-k][id] \tag{9}$$

其中: n 为帧数, id 为特征点的索引, lk 为当前帧与前 5 帧运动向量的相关系数。

5) 平滑处理: 对于全局运动参数估计失效, 可以对最后有效估计的全局运动参数进行逐步衰减直至为 0。衰减系数分别取 1、0.75、0.5、0.25 和 0。如果衰减过程中出现有效全局估计, 或过去 5 帧均无效, 或当前帧与前一帧的差值 GV_{diff} 大于阈值 GV_{DynThr} , 衰减过程仍然采用先前最近有效帧, 其中

$$GV_{\text{DynThr}}[n] = \max \left\{ \begin{array}{l} 0.9 \times GV_{\text{DynThr}}[n-1] \\ 1.5 \times \max \{GV_{\text{diff}}[n-m]\} \end{array} \right\} \tag{10}$$

其中: $GV_{\text{diff}}[n] = GV[n] - GV[n-1]$, m 取过去的 1, 2, ..., 5 帧, n 为当前帧;

$$GV_{\text{Thr}}[n] = \text{sat}(GV_{\text{DynThr}}[n], GV_{\text{Thr}}^{\text{max}}, GV_{\text{Thr}}^{\text{min}}) \tag{11}$$

其中: $GV_{\text{Thr}}^{\text{max}} = \text{round}(0.8 \times 2 \times GV_{\text{max}})$, $GV_{\text{Thr}}^{\text{min}} = \text{round}(0.4 \times 2 \times GV_{\text{max}})$, $GV_{\text{max}}[n]$ 为过去 5 帧全局运动参数中的最大值, 为 GV_{max} ; $GV_{\text{min}}[n]$ 为过去 5 帧全局运动参数中的最小值。利用 $GV_{\text{DynThr}}[n]$ 来表示全局运动参数的动态阈值。

3 集中度参数估计

通过提取具有多尺度特性的特征点, 利用深度优先最邻近查找匹配点对, 常见的利用 RANSAC 等算法进行最优逼近, 估算全局运动参数(将全局运动参数符号取反可以计算运动补偿参数), 这是一种迭代式的矩阵奇异值分解算法, 可以寻找到与运动模型最接近的估计结果。本文提出的基于集中度运动参数估计, 是基于考虑特征提取在整个稳像中占用 70%左右的时间, 保证足够的稳像精度的前提下提出的一种最优算法。参数估计基本流程如图 1。

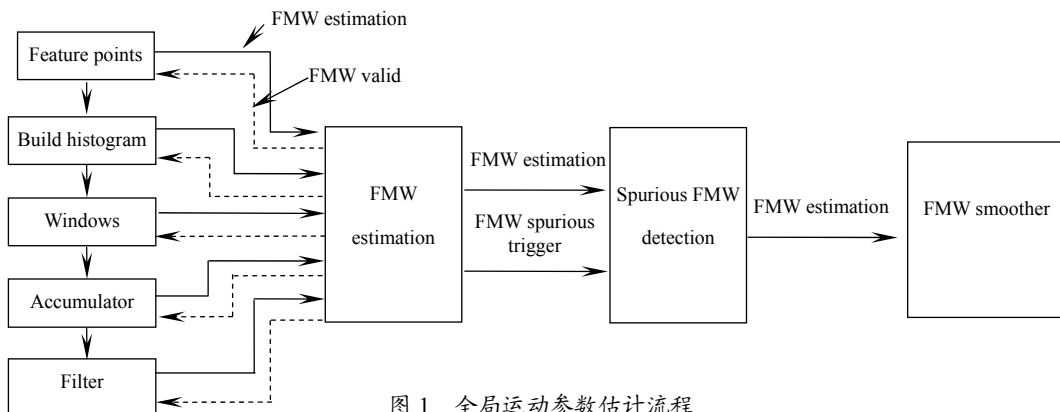


图 1 全局运动参数估计流程

Fig.1 Global motion estimation process

以 132 帧灰度连续视频 shaky_car.avi 文件, 任取一帧图像作为起始帧作参考, 这里取第 3 帧, 以寻找第 4 帧与第 3 帧之间的运动向量关系来说明基于集中度判定的全局参数估计流程。

首先利用 SURF 算法提取图像特征, 并通过特征匹配, 获得匹配点对。根据匹配点对, 以横轴为运动向量, 纵轴为对应运动向量的数目, 建立直方图, 如图 2。根据第二节第一步判定, 如果最大直方图的值大于 $N/3$ (即 $32/3 \approx 11$) 由图 2 可知, 最大值 $H_{\text{max}} = 15$, 满足条件。若不满足条件, 可以通过对运动直方图加窗、累加、滤波等处理, 判定是否满足条件, 得到全局运动参数。当全局运动估计参数结果无法满足式(8)给出的条件, 便触发伪运动向量估计, 通过参数平滑等方式防止图像大幅抖动。运动补偿参数由于通过具有亚像素定位精度的匹配点对获得, 经过集中度估计算法处理后, 补偿参数精度结果不受

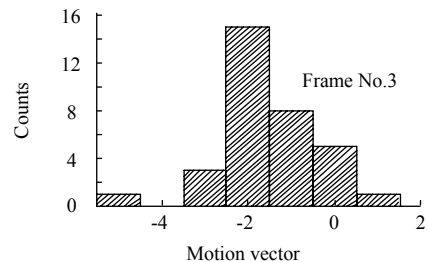


图 2 运动向量直方图

Fig.2 Motion vector histogram

影响，仍具有亚像素精度。

4 稳像实验

本文将采用 132 帧灰度连续视频 shaky_car.avi 文件，在 Windows 平台利用 VS2008 编程实现，CPU P4 3.0 GHz, 1G 内存。按照本文提出的稳像流程，首先以第一帧作为参考帧，并利用多尺度快速鲁棒算法提取具有亚像素定位精度的不变特征，通过基于最小测度距离的深度优先匹配算法，获得匹配点对，然后采用本文提出的基于集中度最优原则的参数估计算法，获得补偿参数，最后实现对图像序列的每帧图像进行运动补偿，达到稳像目的。

4.1 特征提取匹配

实验采取第一节所描述的算法，对图像序列进行不变特征抽取，如图 3。所提取特征具有平移，旋转，缩放等不变特性，特征定位具有亚像素精度通过对图像序列的特征抽取，利用匹配算法可以得到对应的特征点匹配结果，给出图像序列的第 3 帧到第 7 帧对应的匹配结果，误匹配率≤10%。如图 4 所示。



图 3 图像不变特征抽取

Fig.3 Image invariant feature extraction

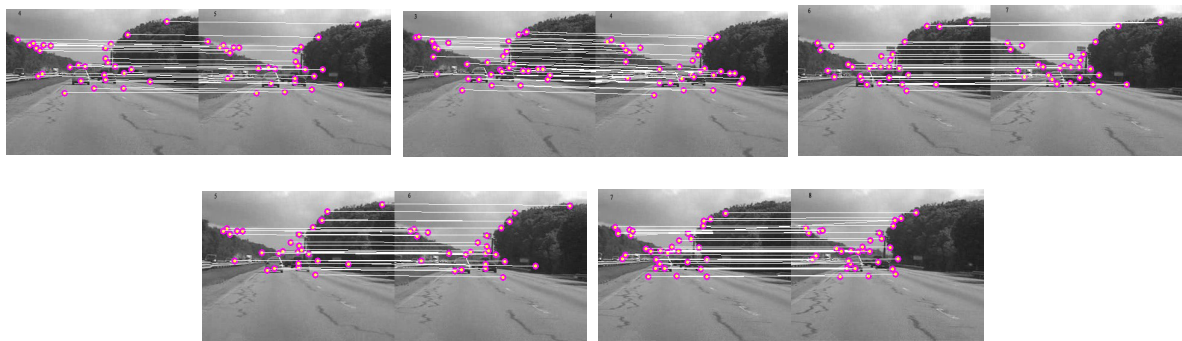


图 4 不变特征匹配结果

Fig.4 Invariant feature matching result

4.2 基于集中度最优化原则的运动参数估计

采用第 2 节所提出的基于集中度最优化原则，估计抖动图像序列帧间平移运动向量，以运动曲线方式描述如图 5、6 所示。通过对运动曲线与 4.3 节给出的稳像结果对比，本文算法估计结果正确，使得输出图像平滑稳定度高，验证了算法的实用性。

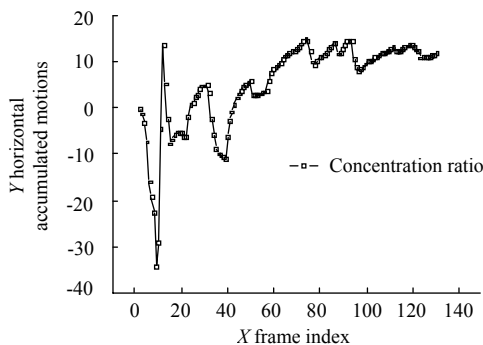


图 5 水平方向累加运动估计

Fig.5 Horizontal accumulated motion

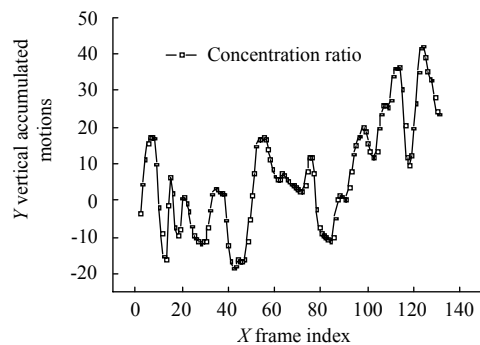


图 6 垂直方向累加运动估计

Fig.6 Vertical accumulated motion

4.3 稳像结果

给出第 3 帧到第 12 帧连续 10 帧具有明显抖动的图像序列的稳像效果, 如图 7。



图 7 稳像实验结果

Fig.7 Stabilization result

4.4 集中度最优原则算法与 RANSAC 比较

本文所提出的运动参数估计算法计算复杂度为 n 次加法, 至多 $4n$ 次乘法, 而采用文献[6]的 RANSAC 算法, 在进行奇异值分解的同时, 还需判定符合估计模型的特征点数量, 整个计算复杂度为 $O(T_{iter}(C_{estimate}(k) + NC_{fitting}))$, 其中 T_{iter} 为迭代次数, $C_{estimate}(k)$ 为利用最小采样集计算模型参数的计算代价, $NC_{fitting}$ 为所有特征点符合模型的计算代价。随着特征点与迭代次数的增加, RANSAC 算法运算复杂度将会明显的增加, 相比之下, 本文所提出的运动参数估计算法具有明显的计算复杂度优势。以测试不同点对匹配数据为例, 其中 RANSAC 算法中迭代次数 $T_{iter} = 100$, 给出本文实验结果, 如表 1。

由图表 1 可知, 基于集中度的运动估计参数算法计算时间平均为 RANSAC 算法的 13%。

表 1 算法性能比较

Table 1 Comparison of algorithm performance

Matching points	RANSAC	Concentration
Matching points No.1	500	500
Time of estimation/ms	16	2
Matching points No.2	1 000	1 000
Time of estimation /ms	40	4
Matching points No.3	4 000	4 000
Time of estimation/ms	137	18

5 结 论

由于电子稳像过程复杂, 往往需要利用特征进行运动参数估计, 然后才能达到稳定抖动视频序列的目的, 整个过程计算复杂度较高, 对每个环节在不影响稳像性能的前提下, 均需要更优化的算法。因此本文提出了在多尺度空间基于集中度的稳像算法, 对于二维抖动图像序列, 利用运动向量建立直方图, 并对直方图通过不同层次模型的建立、计算, 最终能够快速、准确的估计出全局运动补偿参数。实验结果表明, 本文算法稳像精度高, 具有较强的计算复杂度优势, 增强了算法实时性, 同时对误匹配, 参数错误估计等

引入的干扰均具有很好的鲁棒性能。

参考文献:

- [1] BATTIATO S, GALLO G, PUGLISI G, *et al.* SIFT Features Tracking for Video Stabilization [C]// **International Conference on Image Analysis and Processing**, Modena, Sept 10-14, 2007: 825-830.
- [2] 赵红颖, 晏磊, 熊经武. 舰船图像序列电子稳定算法的研究 [J]. **光学精密工程**, 2003, **11**(6): 602-606.
ZHAO Hong-ying, YAN Lei, XIONG Jing-wu. Image composition based on EIS model [J]. **Optics and Precision Engineering**, 2009, **17**(1): 202-206.
- [3] Lowe D G. Distinctive image features from scale invariant keypoints [J]. **International Journal of Computer Vision**(S0302-9743), 2004, **60**(2): 91-110.
- [4] HERBERT BAY, ANDREAS ESS, TINNE TUYTELAAR, *et al.* Speeded Up Robust Features(SURF) [J]. **Computer Vision and Image Understanding (CVIU)** (S0302-9743), 2008, **110**(3): 346-359.
- [5] LINDEBERG T. Feature detection with automatic scale selection [J]. **International Journal of Computer Vision** (S0302-9743), 1998, **30**(2): 79-116.
- [6] Zhang Z. Parameter estimation techniques: a tutorial with application to conic fitting [J]. **Image and Vision Computing Journal**(S0302-9743), 1997, **1**(25): 59-76.
- [7] Matas J, Chum O. Randomized RANSAC with $T(d, d)$ test [J]. **Image and Vision Computing** (S0302-9743), 2004, **22**(10): 837-842.

~~~~~  
(上接第 137 页)

- [6] MALLAT S, HWANG W L. Singularity detection and processing with wavelets [J]. **IEEE Transaction on Information Theory**(S0018-9448), 1992, **38**(2): 617-643.
- [7] SADLER B M, SWAMI A. Analysis of multiscale products for step detection and estimation [J]. **IEEE Transactions on Information Theory**(S0018-9448), 1999, **45**(30): 1043-1051.
- [8] 屈小波, 闫敬文, 杨贵德. 改进拉普拉斯能量和的尖锐频率局部化 Contourlet 域多聚焦图像融合方法 [J]. **光学精密工程**, 2009, **17**(5): 1203-1212.  
QU Xiao-bo, YAN Jing-wen, YANG Gui-de. Multifocus image fusion method of sharp frequency localized Contourlet transform domain based on sum-modified-Laplacian [J]. **Optics and Precision Engineering**, 2009, **17**(5): 1203-1212.
- [9] QU G H, ZHANG D L, YAN P F. Information measure for performance of image fusion [J]. **Electronics Letters**(S0013-5194), 2002, **38**(7): 313-315.
- [10] PETROVIC V, XYDEAS C. On the effects of sensor noise in pixel level image fusion performance [C]// **Proceedings of the Third International Conference on Image Fusion**, Paris, France: IEEE, 2000, **2**: 14-19.
- [11] LOZA A, BULL D, CANAGARAJAH N, *et al.* Non-gaussian model-based fusion of noisy images in the wavelet domain [J]. **Computer Vision and Image Understanding**(S1077-3142), 2010, **114**(1): 54-65.