

doi: 10.3788/gzxb20124111.1372

基于特征匹配的电子稳像优化技术

王敬东, 王智慧, 张春, 丁尤蓉

(南京航空航天大学 自动化学院, 南京 210016)

摘 要: 针对特征匹配的电子稳像算法存在特征点提取耗时多、准确度不高以及特征点误匹配等问题, 从特征点提取及特征点匹配两个方面采取了优化措施, 采用基于快速分割特征点提取算法更快地提取特征点, 并把特征点定位准确度提高到亚像素级; 引入随机采样一致性算法消除误匹配的特征点对, 减少场景中局部运动的影响, 提高了全局运动参量的估计准确性. 实验结果表明, 经优化技术处理后, 视频稳像准确度高、速度快, 能够在存在平移、缩放、旋转及局部有相对运动等情况时保持良好的视频稳像效果.

关键词: 电子稳像; 特征点匹配; 误匹配消除; 全局运动估计

中图分类号: TP391

文献标识码: A

文章编号: 1004-4213(2012)11-1372-5

0 引言

很多摄像系统由于其载体是晃动的, 拍摄的视频经常存在抖动现象. 稳像的目的就是消除这些抖动, 使得视频图像能够更便于观察. 电子稳像包括全局运动估计、运动参量滤波、运动补偿等, 而全局运动估计是电子稳像的重要环节, 决定了电子稳像的效果和实时性.

现有的全局运动估计算法中, 基于图像特征点匹配的全局运动估计方法^[1], 具有对旋转、缩放等较好的适应性, 可以稳定包含复杂抖动形式的视频, 所以具有较好的稳像应用前景. 但由于特征点易受干扰影响, 还存在特征点准确度不高及特征点误匹配等问题^[2-10], 因此稳像的效果还有待进一步提高. 本文研究了特征点提取及特征点匹配的一些优化方法, 通过基于快速分割特征点提取 (Features from Accelerated Segment Test, FAST) 算法更快地提取特征点, 减少干扰的影响; 通过随机采样一致性 (Random Sampling Consensus, RANSAC) 算法消除误匹配的特征点对, 提高了全局运动参量的估计准确性, 同时减少场景中局部运动的影响, 使得稳像适合的场合更加广泛.

1 特征点提取优化

全局运动估计需要准确的匹配点对, 获取一种稳定性好、抗噪能力强的特征点是基于特征匹配全

局运动估计算法的基础.

1.1 FAST 特征点提取算法

FAST 算法^[11]由 Edward Rosten、Reid Porter 和 Tom Drummond 三人在 2008~2010 年期间提出了一种新的特征点检测算法.

FAST 算法检测的特征点定义为在像素点的周围邻域内有足够多的像素点与该点处于不同的灰度区域. 应用到灰度图像中, 即有足够多的像素点的灰度值大于该点的灰度值或者小于该点的灰度值. 通常, 选择圆形区域, 图 1 给出了以该点为中心的圆形区域的模板情况. 该圆形区域为一个半径等于 3 的离散化的 Bresenham 圆.

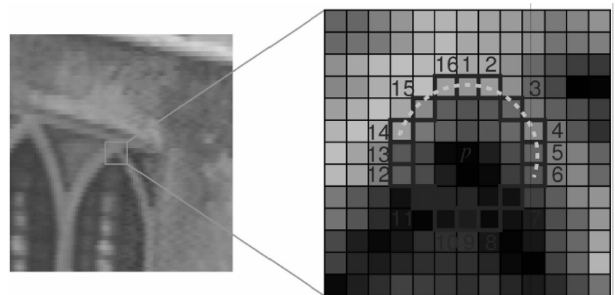


图 1 特征点检测模板示意图

Fig. 1 Template of feature points

特征点可通过定义一个特征点响应函数 (Characteristic points Response Function, CRF) 判断

$$\text{CRF} = |I(x) - I(p)| > \epsilon_a \quad (1)$$

式中 $I(x)$ 表示圆周上任意一点的图像灰度值; $I(p)$

基金项目: 国家自然科学基金 (No. 61074161) 资助

第一作者: 王敬东 (1966-), 男, 副教授, 硕士, 主要研究方向为机器视觉、计算机测控、无线传感器网络. Email: wjd1291_nuaa@163.com

收稿日期: 2012-03-26; 修回日期: 2012-07-12

表示中心像素点的图像灰度值; p 表示中心像素点即候选点; ϵ_d 为给定的一个阈值. 通过上述特征点响应函数, 可以累加出圆周上满足式(1)像素点的个数 N . 如果 N 大于给定的一个阈值, 就可以确定该候选点为特征点. 选取不同的阈值 ϵ_d 可以控制提取的特征点数目.

由于 FAST 算法对一个像素值是否为特征点的判断, 转化为了该点像素与其周围的 16 个点的像素值的大小比较, 没有复杂的数学求导和乘除计算, 因此相对其他特征点提取算法来说, FAST 算法速度快, 耗时较其他算法低一到两个数量级, 并且稳定, 抗干扰能力较强.

图 2 是特征点标准测试图, 图 2(b) 和 2(c) 分别加入高斯噪声和椒盐噪声. 结合该图采用一些典型的特征点提取算法进行了实验. 表 1~3 是实验结果的相关数据比较情况.

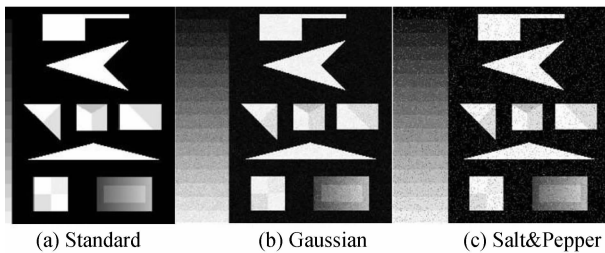


图 2 特征点标准测试图
Fig.2 Standard test chart

表 1 特征点提取算法抗噪保持率变化情况

Table 1 Retention of anti-noise

Type	KLT	Harris	SUSAN	FAST
Gaussian	55.8%	58.4%	80.8%	90.1%
Salt & Pepper	44.6%	49.6%	72.7%	82.1%

表 2 特征点提取算法噪音干扰率变化情况

Table 2 Retention of noise interference

Type	KLT	Harris	SUSAN	FAST
Gaussian	75.5%	72.7%	41.4%	30.9%
Salt & Pepper	92.1%	90.3%	50.2%	41.9%

表 3 特征点耗时情况表

Table 3 Time-consuming

Type	Time/ms	Num
KLT	16.818 00	56
Harris	38.417 50	90
Susan	30.983 90	28
FAST	3.276 70	39

本文采用了抗噪保持率 (Retention of Anti-Noise, ROAN) 和噪音干扰率 (Retention of Noise Interference, RONI) 这两个指标来反映特征点的稳定性和抗干扰能力. 假设在没有加噪音的标准检测

图上提取的特征点总数为 N_1 , 而在有噪音的图像上提取到的特征点总数为 N_2 , 两种图像上提取的特征点中相重合的个数为 N_{rep} , 有噪音的图像上因为噪音影响而提取的特征点数为 $N_2 - N_{rep}$; 则

抗噪保持率

$$K = (N_{rep} / N_1) \times 100\% \quad (2)$$

噪音干扰率

$$F = \frac{N_2 - N_{rep}}{N_2} \times 100\% \quad (3)$$

1.2 特征点优化措施

由表 1~3 可以发现 FAST 特征点具有提取速度快, 特征点稳定, 抗干扰能力强、耗时小的优点, 但 FAST 特征点容易出现局部小区域聚集现象. 提取的 FAST 特征点局部小区域聚集, 会增加后续特征匹配的耗时, 同时这些聚集的特征点代表性差, 不利于全局运动参量的计算.

本文采用了距离约束机制来减少特征点的聚集. 设置两个特征点之间的最小约束距离 d_{min} , 遍历特征点, 若两个特征点之间距离小于 d_{min} , 则剔除掉其中特征值较小的那个特征点. 最小约束距离 d_{min} 要依据处理的视频帧的尺寸来定, 如果待稳像的是高清视频 (1280×720), 则可以将 d_{min} 设置的要大些; 若待稳像的为一般大小的视频 (320×240 和 640×480), 则 d_{min} 取小一些. 一般情况下, 1280×720 , 640×480 和 320×240 大小的视频, d_{min} 分别取 90 个像素、40 个像素和 10 个像素, 这样特征点的总数将会控制在 60~100 个.

距离约束的措施使最终提取到的特征点分布相对均匀、不存在过于聚集现象等.

另外, 为了使得后续的全局运动参量计算的更加精确, 特征点的定位需要更高的准确度. 因此, 本文在 FAST 特征点检测之后, 采用了亚像素级的定位. 亚像素级特征点定位是基于对向量正交性的观测, 利用特征点附近的亮度梯度信息, 采用迭代插值的方法, 实现亚像素级特征提取. 图 3 为亚像素级特征点定位原理示意图.

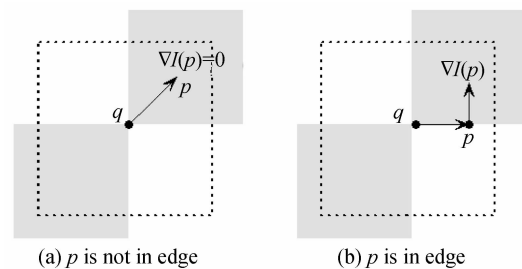


图 3 亚像素级特征点位置计算

Fig.3 Position of sub-pixel feature point

假设经亚像素级定位获得的特征点最后真实位置位于 q 点, 而 p 点是亚像素特征点 q 的邻域 W 内的任意一点, q 点的坐标为 (x, y) , p 点处的图像梯度为 ∇I , 则 p 点的梯度向量和 $q-p$ 向量点积为 0 成立, 也即

$$\nabla I \cdot (q-p) = 0 \tag{4}$$

如图 3, 如果 p 位于边缘, 则 $q-p$ 向量和 p 点的图像梯度向量方向互相垂直, 那么它们的点积为 0; 如果 p 不在位于边缘, p 点的图像梯度为 $\nabla I=0$, 因此式(4)仍成立.

式(4)反映了亚像素特征点和周围邻域中的任意一点组成的向量和该邻域中的点的图像梯度向量正交. 根据这一原理, 就可以通过迭代的方法不断逼近, 最终获得精确的亚像素级特征点精确位置.

由于使特征点定位准确度提高到亚像素级, 因此也提高了后续的全局运动参量计算的准确性, 使得稳像效果得到改善.

2 误匹配点消除技术

对提取的特征点需要进行匹配, 但难免会存在一些误匹配. 这些误匹配点一方面来源于匹配算法自身原因所导致的匹配错误, 另一方面来源于视频帧中的存在局部运动的物体. 如果不消除误匹配点, 将会导致后续全局运动参量计算错误, 以至于稳像失败. 为此本文提出了一种基于 RANSAC 算法的误匹配点消除技术, 克服了局部运动影响, 实现了更加稳定的电子稳像.

2.1 RANSAC 算法

摄像机的全局运动一般用参量模型表示, 综合考虑参量估计准确度和速度, 本文采用六参量仿射模型对摄像机运动建模, 仿射运动模型如式(5)

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} a_3 \\ a_6 \end{bmatrix} \tag{5}$$

式中 $(a_1, a_2, a_3, a_4, a_5, a_6)$ 为摄像机的全局运动参量; 其中 (a_1, a_5) 为缩放参量, (a_2, a_4) 为旋转参量, (a_3, a_6) 为平移参量; (x_i, y_i) 为当前帧的像素的坐标; (X_i, Y_i) 为对应像素经过仿射变换后的当前帧坐标.

式(5)反映了从一帧图像的特征点到另一帧图像特征点的射影关系, 这一关系也可以用射影矩阵 H 表示

$$H = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{bmatrix}$$

矩阵 H 也就是全局运动参量矩阵.

随机采样一致性算法^[12] (Random Sampling Consensus, RANSAC)是由 Fischler 和 Bolles 两人

提出的一种鲁棒性的数学估计算法. RANSAC 算法通过随机抽取尽可能少但又充分的数据确定模型参量, 然后根据模型参量对所有数据进行分类, 落在参量模型允许的误差范围之内点, 称为内点; 误差范围之外的点, 称为外点. 在 RANSAC 算法中, 假若整个待抽样数据集中“内点”的概率 e , 则一次抽样的样本数据 (k 个) 全为“内点”的概率为 e^k , 要保证在一定置信概率 p_c 下, 至少有一次抽样的数据全为“内点”, 需要的抽样的次数 N 为

$$N = \frac{\lg(1-p_c)}{\lg(1-e^k)} \tag{6}$$

2.2 误匹配点消除

本文引入 RANSAC 算法来消除误匹配点, 即通过随机选取一定数目的匹配点对, 使用奇异值分解 (Singular Value Decomposition, SVD) 求取模型参量, 计算出符合参量模型要求的“内点”(正确的匹配点对), 剔除“外点”(误匹配点对). 在 RANSAC 算法重复 N 次后, 就可以保证得到一组数据全为内点. 流程如图 4, 图中 InlierNum 是内点数, TotNum 是总的匹配点对. 本文中参量 k 取 3, p_c 一般取经验值 0.95. 每进行一次抽样, 计算内点的概率, 始终保持取最大值. 最后利用内点个数最多的那个模型所确定的内点数据, 计算出模型参量, 消除误匹配点的影响.

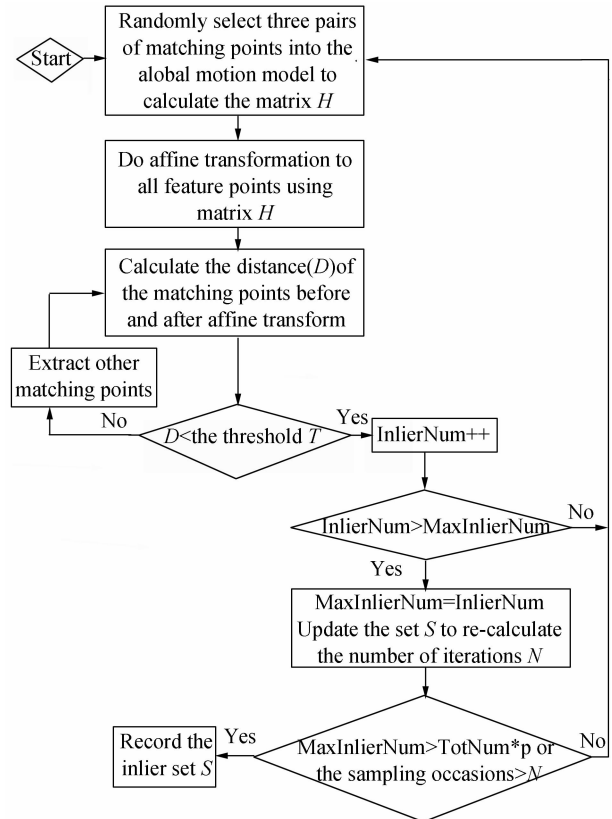


图 4 RANSAC 消除误匹配流程图

Fig. 4 Flow chart of remove the mismatch using RANSAC

另外,本文也利用 RANSAC 算法消除场景中的局部运动对全局运动参量估计的影响.全局运动是指由于摄像机位置或者参量变化引起的图像的整体变化,局部运动是指图像序列中场景中个别物体的运动.因为局部运动矢量和全局运动矢量不一致,分布在参考帧和当前帧中局部运动物体上的匹配点对,即使匹配正确,但是它们组成的运动矢量中,不仅包含全局运动分量,还包含局部运动分量.因此分布在局部运动物体上的匹配点对经过仿射变换后,其欧式距离也将会大于阈值 T ,不满足全局运动模型.因此,分布在局部运动物体上特征点也将被剔除.从而利用 RANSAC 算法消除了局部运动的影响.

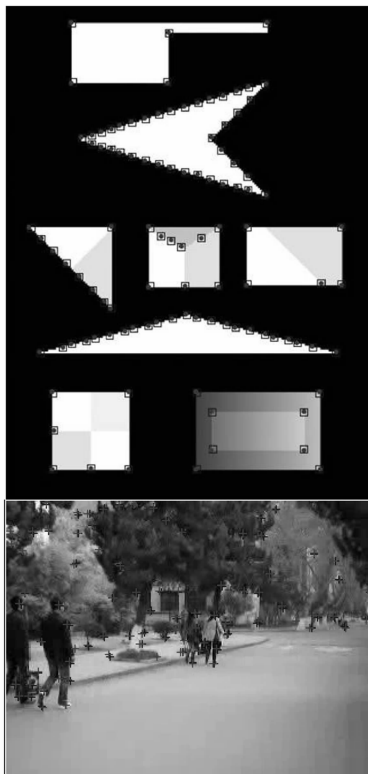
通过调整 RANSAC 算法中的阈值 T ,可以改变对特征点的限制,从而改变最终内点集 S 中匹配点对的数目.因此,可以根据实际视频大小和拍摄环境的不同,调整阈值 T ,使得最终内点集 S 中的特征点对数量适中,保持在 20~40 对.

引入 RANSAC 算法后,克服了误匹配点对及局部运动的影响,使计算出的全局运动参量更加准确.

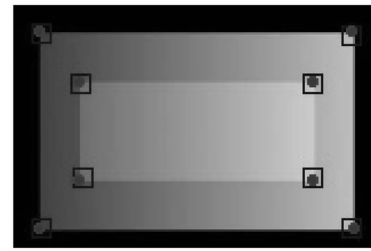
3 实验结果及分析

3.1 特征点提取实验

对标准特征点检测图和实际拍摄图进行特征点提取实验.特征点提取结果如图 5(a),其中圆点标注的为亚像素级特征点位置,方框标注的为 FAST 原特征点位置.图 5(b)是其中一个局部放大图,表 4 列出特征点检测标准图的一些特征点位置坐标.



(a) Sub-pixel feature position



(b) Partial enlargement

图 5 亚像素级特征点定位结果图及局部放大图

Fig. 5 Results map and partial enlargement of sub-pixel feature position

表 4 亚像素特征点定位表

Table 4 Characteristics of sub-pixel point positioning table

Feature points	Org-pixel	Sub-pixel
Corner1	(168,246)	(166.9, 245.2)
Corner2	(427,209)	(428.2, 209.6)
Corner3	(178,384)	(177.5, 383.4)
Corner4	(300,310)	(298.5, 311.3)
Corner5	(214,84)	(212.6, 85.7)

从图 5(b)局部放大图中可看出,圆点标注的亚像素级特征点的位置要更为精确,而方框标注的 FAST 原特征点的位置则存在一定的偏差.

3.2 消除局部运动影响实验

图 6 中左右两图是视频中的相隔数帧的图像序列帧,选取的图像序列帧中存在局部运动(行走的人).

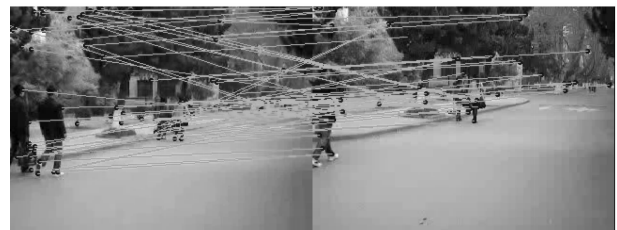


图 6 初始匹配结果图

Fig. 6 The result of first match

本文采用 KLT 算法进行初始匹配,得到的初始匹配点对的数目为 72 对.从图 6 可以发现,未加入 RANSAC 算法之前,初始匹配结果图中存在很多的交叉线,这些交叉线是误匹配产生的.

图 7 是通过 RANSAC 消除误匹配点后的结果,可以看出匹配点连线都是平行线,没有了交叉线,人体等局部运动物体上的匹配点都被剔除了.



图 7 去除误匹配后的特征点分布图

Fig. 7 Maps of remove the dis-match

最后得到的正确匹配点对数目为 29,可以满足全局运动参量求取的要求。

3.3 全局运动估计的耗时

针对 320×240 的视频,采用 FAST 算法提取特征点,耗时 2.3 ms;采用 KLT 算法做特征点初始匹配,耗时 4.3 ms;采用 RANSAC 算法消除误匹配点,耗时 0.9 ms;采用六参量仿射模型计算全局运动参量(Global Motion,GM),耗时 0.6 ms。一帧的完整稳像处理耗时 8.1 ms。详细情况如表 5 所示,可见本文提取特征点及全局运动估计等处理具有较好的实时性。

表 5 稳像算法耗时统计表

Table 5 Time-consuming tables of algorithm

Num of points	FAST	KLT	RANSAC	GM
Reference frame	72	72	29	/
Current frame	72	72	29	/
Time-consuming	2.3 ms	4.3 ms	0.9 ms	0.6 ms

4 结论

本文从特征点提取准确和特征点匹配正确两个方面开展了一些优化处理,FAST 算法提取的特征点达到了亚像素级准确度,保证了全局运动参量计算的准确性;RANSAC 算法剔除误匹配点及局部运动的影响,使得稳像更加稳定和可靠。实验结果表明,本文提出的这些优化方法具有较好的稳像效果,速度快,适合在存在平移、缩放、旋转及局部有相对运动等复杂情况下的视频稳像场合应用。

参考文献

- [1] HUANG J C. Automatic feature-based global motion estimation in video sequences [J]. *IEEE Transactions of Consumer Electronics*, 2004, **50**(3): 911-915.
- [2] LUO Shi-tu, ZHANG Qi, WANG Yan-ling. A real-time electronic image stabilizing algorithm based on feature matching [J]. *Journal of National University Technology*, 2003, **27**(3): 45-48.

- 罗诗途,张圻,王艳玲.一种基于特征匹配的实时电子稳像算法[J].国防科技大学学报,2003,**27**(3):45-48.
- [3] SHI Yang, GAO Xin-bo. A video stabilization algorithm based on feature tracking [J]. *Acta Photonica Sinica*, 2005, **34**(7): 1108-1112.
- 史阳,高新波.一种基于特征跟踪的视频序列稳像算法[J].光子学报,2005,**34**(7):1108-1112.
- [4] ZHU Juan-juan, GUO Bao-long. Electronic image stabilization method based on projection algorithm [J]. *Acta Photonica Sinica*, 2005, **34**(8): 1266-1269.
- 朱娟娟,郭宝龙.电子稳像的特征点跟踪算法[J].光子学报,2005,**34**(8):1266-1269.
- [5] PENG Xiao-jiang, ZHANG Jia-shu. Robust real-time electronic image stabilization based on feature matching and checking [J]. *Acta Photonica Sinica*, 2011, **40**(9): 1442-1446.
- 彭小江,张家树.基于特征匹配和校验的鲁棒实时电子稳像[J].光子学报,2011,**40**(9):1442-1446.
- [6] VELLA F, CASTORINA A. Digital image stabilization by adaptive block motion vector filtering [J]. *IEEE Transactions on Consumer Electronics*, 2002, **48**(3): 796-801.
- [7] YANG Shih-hsuan. An adaptive image stabilization technique [C]. 2006 IEEE International Conference on Systems, Man, and Cybernetics. 2006, Taipei, China, 2006: 1968-1973.
- [8] ZHANG Kun, XU Ting-fa, WANG Ping, et al. Real-time full-frame digital image stabilization system by SURF [J]. *Optics and Precision Engineering*, 2011, **19**(8): 1964-1972.
- 张坤,许廷发,王平,冯亮.高精度实时全帧频 SURF 电子稳像方法[J].光学精密工程,2011,**19**(8):1964-1972.
- [9] ZHANG Xu-bing, HIRAI S, ZHANG Peng-lin. SURF and spatial association correspondence applied in extraction and matching of feature points from mr images of deformed tissues [C]. Proceedings of the 2010 IEEE. International Conference on Robotics and Biomimetics, Tianjin, China, 2010: 448-453.
- [10] NEETA N, VIJAY L, BHAVITAVYA B, et al. Fast feature point detector [C]. 2008 IEEE International Conference on Signal Image Technology and Internet Based Systems. 2008: 301-306.
- [11] ROSTEN E, DRUMMOND T. Faster and better: a machine learning approach to corner detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(1): 105-119.
- [12] CHEN Tai, LIU Yun-hui. Robust structure and motion estimation by auto-scale random sample consensus [C]. Proceedings of the 2006 IEEE International Conference on Information Acquisition, Weihai, Shandong, China, 2006: 37-42.

Optimization Technique of Electronic Image Stabilization Based on Feature Matching

WANG Jing-dong, WANG Zhi-hui, ZHANG Chun, DING You-rong

(College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Electronic image stabilization still exists some problems, such as high time-consuming, low accuracy, high error matching probability and so on. The paper takes appropriate measures to solve these problems from two aspects: feature points extraction and matching. Firstly, the FAST algorithm is adopted to extract the feature points, and positional precision of the feature points is improved to the level of sub-pixel. Then the RANSAC algorithm is used to eliminate the error matched points, avoid the interference of relative movement, and the global motion parameters calculated is more accurate. Experimental results show that a more higher precision and faster speed is obtained, which can also be applied for the complex shaking video containing translation, rotation shake and relative movement.

Key words: Image stabilization; Feature points matching; Elimination of error matching; Global motion estimation