

文章编号:1004-4213(2011)09-1442-5

基于特征匹配和校验的鲁棒实时电子稳像

彭小江, 张家树

(西南交通大学 信息科学与技术学院, 成都 610031)

摘 要:提出了一种基于特征点匹配和校验的鲁棒实时电子稳像算法. 首先利用 Kanade-Lucas-Tomasi 角点检测器提取参考帧和当前帧的特征点, 并用绝对误差和准则进行特征点匹配; 在校验阶段, 提出一种能够有效剔除前景运动物体特征点和错误匹配点的空间位置不变准则; 最后, 在相似运动模型下, 利用最小二乘法求解全局运动矢量进行运动补偿. 实验证明: 该算法满足实时性要求, 对视频的平移、旋转、缩放运动都有较好的稳像效果, 并对运动物体具有鲁棒性.

关键词:电子稳像; 运动估计; 特征点匹配; 随机抽取一致性; 空间位置不变准则

中图分类号: TP391.4

文献标识码: A

doi: 10.3788/gzxb20114009.1442

0 引言

电子稳像是一种直接利用图像处理来实现视频画面稳定的技术. 它主要用于在成像设备自身抖动不可避免的领域, 如车载、机载、船载观瞄系统、无人车辆视觉导航系统、高杆监控系统、手持长焦成像设备等. 稳像算法通常包括四个过程: 运动估计、运动校正、运动补偿和图像补偿. 运动估计主要用于提取相机的全局运动矢量, 包括局部运动估计和全局运动估计; 运动校正主要是用于存在相机意向运动情况下, 从全局运动矢量中获得相机的抖动矢量; 运动补偿根据得到的抖动矢量, 对当前帧进行平移、旋转和缩放, 使视频稳定、流畅; 图像补偿则对运动补偿形成的“无定义”区域进行重构, 再现全帧视觉效果.

现有的电子稳像算法大致有: 投影法^[1]、相位相关法^[2]、块匹配法^[3]、位平面算法^[4]、对数极坐标变换(Log-Polar)法^[5]、光流法^[6]和特征点匹配法^[7-10]等. 其中, 基于特征点匹配的电子稳像算法是随着特征点提取技术发展起来的. 目前, 常用于电子稳像的特征点有: Harris 角点^[7]、KLT 角点^[8]、最小吸收同值核区点^[9] (Smallest Univalued Segment Assimilating Nucleus, SUSAN)、尺度不变点^[10] (Scale Invariant Feature Transform, SIFT)、加速的尺度不变点 (Speed Up Robust Feature, SURF). Huang^[7]和 Jiang^[9] 提取特征点后没有校验直接计算全局运动矢量; Luo 等^[8]在匹配 KLT 特征点后, 通过设定连续帧匹配的相关阈值以及判断运动矢量的标准偏差来检验特征点的有效性. 当画面旋转较大时该方法很容易将正确匹配点剔除; Yao 等人^[10]利用随机抽取一致性^[15] (Random Sample

Consensus, RANSAC) 校验提取的 SIFT 特征点; Zhu^[11]等人利用高斯-拉普拉斯算子 (Laplacian Of Gaussian, LOG) 提取特征点并匹配后, 使用了刚性物体距离不变准则 (Spatial Distance Invariant Criterion, SDIC) 进行校验, 该准则通过判断当前帧中所有特征点的组合在参考帧中的距离是否相近来判断匹配的正误, 能达到剔除误匹配和运动物体上的点, 但其校验时运算量大, 系统难以实时.

本文综合考虑稳像算法的实时性和稳像质量, 提出了一种基于 KLT 角点匹配和空间位置不变准则 (Spatial Location Invariant Criterion, SLIC) 校验的实时电子稳像算法. 该算法能够快速有效地剔除误匹配点和前景运动点, 使得最小二乘法一次计算就得到正确的运动矢量, 从而达到实时稳像效果. 另外, 本文分析了四种特征点校验准则, 从理论上说明了空间位置不变准则的优越性.

1 鲁棒性的局部运动估计

1.1 子区域 KLT 特征点检测

KLT 特征点检测^[12]主要是通过检验每个像素点的二维梯度矩阵的最小特征值来实现, 如果满足式(1)且该最小特征值为局部最大值时判为 KLT 特征点. 本文使用 Harris 矩阵^[13]作为梯度矩阵, 定义如式(2). Harris 矩阵在梯度基础上加入了高斯滤波, 有一定的抗噪能力.

$$\lambda_1 = \min\{\lambda_1, \lambda_2\} > t \quad (1)$$

$$M = \begin{bmatrix} \left(\frac{\partial I_{(x_i, y_i)}}{\partial x}\right)^2 \otimes w & \left(\frac{\partial I_{(x_i, y_i)}}{\partial x} \frac{\partial I_{(x_i, y_i)}}{\partial y}\right) \otimes w \\ \left(\frac{\partial I_{(x_i, y_i)}}{\partial x} \frac{\partial I_{(x_i, y_i)}}{\partial y}\right) \otimes w & \left(\frac{\partial I_{(x_i, y_i)}}{\partial y}\right)^2 \otimes w \end{bmatrix} \quad (2)$$

第一作者: 彭小江(1987-), 男, 博士研究生, 主要研究方向为视频图像处理、模式识别和智能车辆技术. Email: xiaojiang_p@yahoo.com
收稿日期: 2011-05-11; 修回日期: 2011-06-27

通常,为了得到均匀的特征点分布都会将图像划分为子块^[8],在子块中分别检测一定数量的特征点. 本文考虑到新视频帧的边界部分为新息,因此将图像分为如图 1 矩形框所示的四个子块. 为了实时性和正确性,特征点数目不宜过多或过少. 因此,在每个子块中预先设定提取特征点的最大个数为 20 以及特征点距离大于 5 个像素,当按前述方法提取的特征点数大于 20 时,阈值 t 减半,再重复提取过程,直到收敛到预先条件. 实际中,为了提高收敛速度,初始阈值根据经验设定. 本文提取的 KLT 特征点如图 1 中十字标记所示.

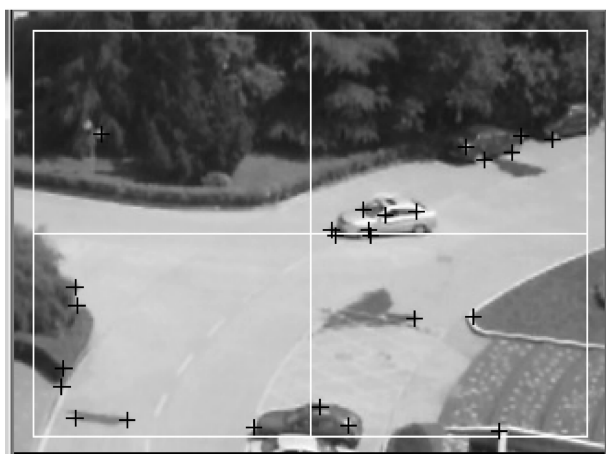


图 1 子区域 KLT 特征点检测结果

Fig. 1 Detected result of the KLT feature points in sub-regions

1.2 特征点匹配

特征点匹配通常有两种方法:一种是利用特征点自身的描述子与其它特征点描述子的距离来实现匹配,如 SIFT 特征点和 SURF 特征点的匹配;另一种是以特征点及其邻域内像素点的灰度相似性作为匹配准则,如 SAD 准则、MSE 准则等.

KLT 特征点没有提取描述子,考虑到匹配算法的复杂性以及后续存在匹配点校验,本文选用 SAD 作为匹配准则,即以当前帧的特征点为中心,定义一个搜索窗口,将当前帧的特征点及其邻域组成最小块,在参考帧的特征点中寻找最优匹配. 图 2 显示了连续两帧按 SAD 准则的特征点匹配情况.



(a) The reference frame (b) The current frame

图 2 特征点匹配结果

Fig. 2 Matching result of the feature points

1.3 匹配点校验

得到正确匹配点后,局部运动矢量即为对应点坐标之差. 通常,由于视频中的运动物体存在及错误匹配的情况,匹配后的特征点不能直接用于全局运动矢量的计算,必须先对匹配点进行校验. 比较简单的校验方法有刚性物体距离不变准则^[11],另一种在模型参量精确估计中应用广泛的方法是随机抽取一致性^[15](Random Sample Consensus, RANSAC)校验法.

1.3.1 RANSAC 校验

RANSAC 校验法利用样本一致性原则,首先通过随机抽取能求解模型的最少样本求解模型参量,然后将此模型应用到其它样本,若极少数或无样本满足此模型则选择的样本视为异常样本(局外点),否则将不满足此模型的样本视为局外点,通过迭代的方式对样本精选,获取精确的参量估计.

在特征点匹配校验中,RANSAC 校验能够精确剔除所有不满足相似模型的异常点,但是这种校验时间复杂度比较大.

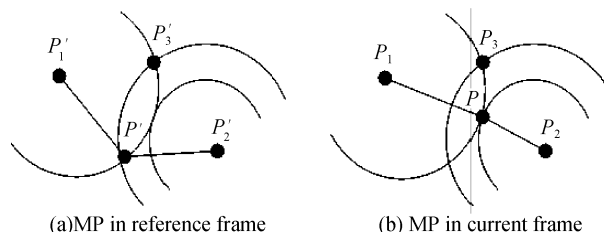
1.3.2 空间位置不变准则

本文提出的空间位置不变准则是指刚性目标点在空间中的相对位置不随仿射及缩放变换而变化,它是传统空间距离不变准则的一个推广. 传统空间距离不变准则是指刚性目标点之间的距离经仿射变换后保持不变,其数学形式定义为

$$\sqrt{(x_{1i} - x_{1j})^2 + (y_{1i} - y_{1j})^2} = \sqrt{(x_{2i} - x_{2j})^2 + (y_{2i} - y_{2j})^2} \quad (3)$$

该准则只能校验不存在缩放运动或者缩放很小的情况. 按照该准则来校验 N 对匹配点^[11],需要进行 $N(N-1)$ 次距离判断(Distance Judgment, DJ).

为了整个稳像算法的实时性,本文最初研究是利用该准则找到一个正确匹配点后,直接把那些到该点距离不满足空间距离不变准则的点剔除,并将此方法称为非穷尽空间距离不变准则. 这样,在假设正确匹配点大于总匹配点数目一半的情况下,最好的只需要 $N-1$ 次距离判断,最坏也只需要 $3N^2/8$ 次距离判断. 然而,实验发现此举改变并不能把错误



(a) MP in reference frame (b) MP in current frame

图 3 非穷尽空间距离不变准则的特殊情况

Fig. 3 An exception of non-exhausted SDIC

匹配点和运动前景上的匹配点完全剔除,原因在于有些错误匹配或运动前景点刚好出现在用于距离准则判断的正确匹配点的圆周附近,如图3中 P 和 P' 点刚好出现在 P_1 和 P'_1 点的圆周上,在存在运动物体和无规则背景(如:树叶、草地等)的视频中该现象尤为严重.

本文考虑到错误匹配点不可能出现在两个正确匹配点的圆周上,即如果匹配点相对于两个正确匹配点的空间位置不变则为正确匹配点,如图3中 P_3 和 P'_3 点.在存在缩放运动情况下,本文提出的空间位置不变准则数学形式为同时满足式(4)和(5)

$$\sqrt{(x_{1i}-x_{1j})^2+(y_{1i}-y_{1j})^2} = \lambda \sqrt{(x_{2i}-x_{2j})^2+(y_{2i}-y_{2j})^2} \quad (4)$$

$$\sqrt{(x_{1k}-x_{1j})^2+(y_{1k}-y_{1j})^2} = \lambda \sqrt{(x_{2k}-x_{2j})^2+(y_{2k}-y_{2j})^2} \quad (5)$$

(x_{1i}, y_{1i}) 和 (x_{1k}, y_{1k}) 是 (x_{2i}, y_{2i}) 和 (x_{2k}, y_{2k}) 的已经确认的正确匹配点(实际中,尽可能使它们的间距大), (x_{1j}, y_{1j}) 和 (x_{2j}, y_{2j}) 是待校验的匹配点, λ 是缩放因子.本文的校验结果如图4,粗线表示错误匹配点或运动前景点.图5为校验流程图.



(a) The reference frame (b)The current frame

图4 本文的校验结果

Fig. 4 Checked result of the paper

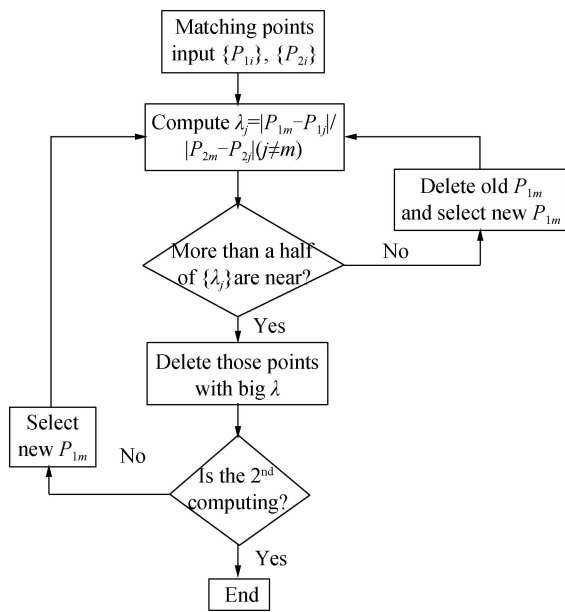


图5 空间位置不变准则校验框图

Fig. 5 Checked framework of SLIC

1.3.3 校验方法对比总结

在假设有 N 对匹配点,正确匹配点大于总数一半的情况下,上述校验方法的理论对比分析如表1.

表1 不同匹配点校验方法对比

Criterion	Cost	Performance
RANSAC	The best: 2 model computing, $N-2$ affine computing and DJS The worst: unpredictable	Good
Traditional SDIC	$N(N-1)$ DJS	Good
Non-exhausted SDIC	The best: $N-1$ DJS The worst: $3N^2/8$ DJS	Bad
SLIC	The best: less than $2(N-1)$ DJS The worst: $3N^2/8+N/2$ DJS	Good

2 全局运动估计及补偿

2.1 实际运动模型和相似运动模型

得到局部运动矢量后,根据相机的运动模型可以求出全局运动矢量.相机的实际运动模型可表示为

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \gamma \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} + \begin{bmatrix} d_x \\ d_y \end{bmatrix} \quad (6)$$

式中, (x, y) 和 (x', y') 分别为运动前后的像素点, (x_0, y_0) 为旋转中心, γ 为缩放因子, d_x 和 d_y 为水平和垂直方向的平移量, θ 为旋转角度.此模型可以简化为相似运动模型^[4],表示为

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \quad (7)$$

由式(6)和(7)可得

$$a = \gamma \cos \theta, b = \gamma \sin \theta$$

$$c = (1-a)x_0 + by_0 + d_x \quad (8)$$

$$d = (1-a)y_0 - bx_0 + d_y$$

从式(8)可知, c, d 既与平移运动相关^[4],也和旋转、缩放相关.

2.2 全局运动矢量估计

假设有 n 个正确匹配点,选用相机的相似运动模型,可以得到以下线性系统

$$\begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ x_n & -y_n & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ y_n & x_n & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \\ y'_1 \\ \vdots \\ y'_n \end{bmatrix} \quad (9)$$

简写为

$$Av = b' \quad (10)$$

式中, $\mathbf{v}=[a, b, c, d]^T$ 代表相机在相似运动模型的全局运动矢量, \mathbf{b}' 代表当前帧的特征点坐标, \mathbf{A} 包含参考帧的特征点坐标. 利用最小二乘法 (LS) 可以得到

$$\mathbf{v}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}' \quad (11)$$

2.3 运动补偿及图像补偿

本文主要针对抖动的监控视频, 不考虑相机的意向运动. 在 2.2 节求解全局运动向量时, 用参考帧 (本文为视频的第二帧) 作为变换后的帧, 因此得到的相似运动模型的全局运动矢量 \mathbf{v} 即为当前帧的变换参量. 将此变换参量按式 (7) 应用到当前帧, 即可得带“无定义区”的运动补偿图像, 然后将参考帧中相应区域对无定义区填充, 为了使得填充边界 (拼接处) 平滑, 本文加入中值滤波来完成图像补偿.

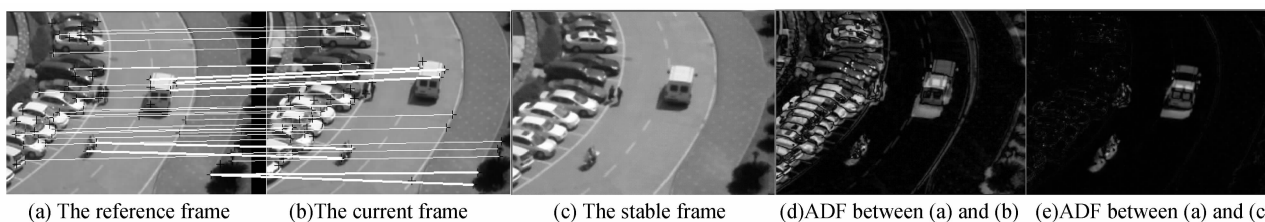


图 6 连续两帧的稳像结果

Fig. 6 Stabilization results for two real consecutive frames

3.2 视频序列

为了评估本文算法在视频中的稳像效果, 本文对一个不含相机意向运动的人为抖动视频做了实验, 利用帧间逼真度, 即峰值信噪比^[16] (Peak Signal-to-Noise Ratio, PSNR) 作为评价标准, PSNR 的定义如式 (12). 图 7 为 100 帧连续视频稳像前后的 PSNR 值对比.

$$\text{PSNR}(I_1, I_0) = 10 \log_{10} \frac{255^2}{\text{MSE}(I_1, I_0)} \quad (12)$$

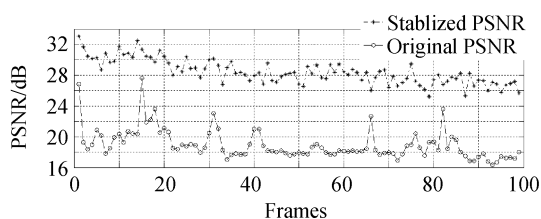


图 7 实际视频中逼真度比较

Fig. 7 PSNR comparison in the real video

4 结论

电子稳像是目前用于稳定视频输出的一个有效技术手段. 本文提出了一种基于 KLT 特征点匹配和空间位置不变校验的电子稳像方法. 该算法能够有效去除前景运动物体的影响, 对视频的平移、旋转、缩放运动都有较好的稳像效果, 在主频为 2.93 GHz 的 PC 机及 C 语言平台上, 平均单帧的稳

3 实验结果及分析

3.1 连续两帧图像序列

通过上述理论分析, 为了说明本文算法对运动前景的鲁棒性, 采集了有多个运动物体的两帧视频 (320×240) 做实验. 如图 6(a) 为参考帧, (b) 为当前帧, 粗线段表示校验后的误匹配及运动前景点, 细线为正确匹配点. 图 6(c) 为稳像帧, (d) 为当前帧和参考帧的绝对差帧 (Absolute Difference Frame, ADF), (e) 为稳像帧和参考帧的 ADF. 其中, 当前帧的运动矢量 \mathbf{v} 为 $[0.996\ 3\ 0.003\ 7\ -7.816\ 2\ 8.235\ 6]^T$, 中值滤波窗大小为 3×3 . 对比绝对差帧可以看出, 本文算法在旋转和平移运动下, 对存在前景运动物体也有很好的稳像效果.

像处理时间为 42.5 ms, 达到实时效果.

参考文献

- [1] LI Ji-tian, HE Yong-qiang, CHEN Cai-sen, *et al.* A method of vehicle-bone electronic image stabilization based on gray projection algorithm[J]. *Infrared Technology*, 2010, **32**(6): 328-332.
李计添, 何永强, 陈财森, 等. 一种基于灰度投影算法的车载电子稳像方法[J]. *红外技术*, 2010, **32**(6): 328-332.
- [2] DU Deng-chong, JIANG Xiao-yu, YAO Jun. Electronic image stabilization algorithm based on phase correlation[J]. *Laser & Infrared*, 2009, **39**(1): 63-66.
杜登崇, 蒋晓瑜, 姚军. 基于相位相关的电子稳像算法研究[J]. *激光与红外*, 2009, **39**(1): 63-66.
- [3] XU Li-dong, LIN Xing-gang. Digital image stabilization based on circular block matching [J]. *IEEE Transactions on Consumer Electronics*, 2006, **52**(2): 566-574.
- [4] GUO Bao-long, ZHU Juan-juan, SUN Wei. Hierarchical bit-plane global motion estimation for image stabilization[J]. *Acta Photonica Sinica*, 2009, **38**(11): 2993-2998.
郭宝龙, 朱娟娟, 孙伟. 电子稳像的分层位平面全局运动估计算法[J]. *光子学报*, 2009, **38**(11): 2993-2998.
- [5] ZOKAI S, WOLBERG G. Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations[J]. *IEEE Transaction on Image Processing*, 2005, **14**(10): 1422-1434.
- [6] CHANG Jyh-yeong, HU Wen-feng, CHENG Mu-huo, *et al.* Digital image translational and rotational motion stabilization using optical flow technique [J]. *IEEE Transactions on Consumer Electronics*, 2002, **48**(1): 108-115.
- [7] HUANG Keng-yen, TSAI Yi-min, TSAI Chih-chung, *et al.* Feature-based video stabilization for vehicular applications[J]. *IEEE 14th International Symposium on Consumer Electronics*, 2010: 1-2.

- [8] LUO Shi-tu, ZHANG Qi, WANG Yan-ling, *et al.* A real-time electronic image stabilizing algorithm based on feature matching[J]. *Journal of National University of Defense Technology*, 2005, **27**(3): 45-49.
罗诗途, 张纪, 王艳玲, 等. 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, 2005, **27**(3): 45-49.
- [9] JIANG Wen-tao, CHEN Wei-dong, LI Fu-liang. Electronic image stabilization algorithm based on characteristic point tracking[J]. *Journal of Applied Optics*, 2010, **31**(1): 73-77.
姜文涛, 陈卫东, 李良福. 一种基于特征点跟踪的电子稳像算法[J]. 应用光学, 2010, **31**(1): 73-77.
- [10] YAO Shen, THYAGARAJU D, BILL P B, *et al.* Video stabilization using principal component analysis and scale invariant feature transform in particle filter framework[J]. *IEEE Transactions on Consumer Electronics*, 2009, **55**(3): 1714-1720.
- [11] ZHU Juan-juan, GUO Bao-long. Features tracking algorithm for video stabilization[J]. *Acta Optica Sinica*, 2006, **26**(4): 516-521.
朱娟娟, 郭宝龙. 电子稳像的特征点跟踪算法[J]. 光学学报, 2006, **26**(4): 516-521.
- [12] SHI Jian-bo, TOMASI C. Good features to track[C]. IEEE Conference on Computer Vision and Pattern Recognition Seattle, 1994: 593-600.
- [13] HARRIS C, STEPHENS M. A combined corner and edge detector[C]. Alvey Vision Conference, 1988: 147-152.
- [14] HERBERT B, TINNE T, LUC V G. Surf: speeded up robust features[C]. Proceeding of European Conference on Computer Vision, 2006: 346-359.
- [15] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. *Communications of the ACM*, 1981, **24**(6): 381-395.
- [16] MORIMOTO C, CHELLAPPA R. Evaluation of image stabilization algorithms[C]. Proc IEEE Int Conf. Acoustics, Speech and Signal Processing, 1998, **5**: 2789-2792.

Robust Real-time Electronic Image Stabilization Based on Feature Matching and Checking

PENG Xiao-jiang, ZHANG Jia-shu

(College of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China)

Abstract: A robust real-time electronic image stabilization algorithm was proposed based on feature matching and checking. In the algorithm, feature points in reference and current frame were extracted by Kanade-Lucas-Tomasi corner detector and matched by the sum of absolute difference criterion firstly. In the process of checking, a spatial location invariant criterion was presented to delete those points which were error-matching or appeared in foreground moving objects effectively. Finally, the global motion vector was computed by least-squares algorithm in similarity motion model and applied to compensate. Experimental results show that this algorithm is able to meet the requirement in real time and has good performance to the translation, rotation and zoom, and is robust to foreground moving objects.

Key words: Electronic image stabilization; Motion estimation; Feature point matching; Random sample consensus; Spatial location invariant criterion