

一种内容完整的视频稳定算法

彭 艺, 叶齐祥, 黄 钧, 焦建彬

(中国科学院研究生院, 北京 100049)

摘 要: 设计了一种基于可靠特征集合匹配的内容完整的视频稳定算法。为了避免运动前景上的特征点参与运动估计, 由经典的 KLT (Kanade-Lucas-Tomasi) 算法提取特征点, 而后基于特征有效性判定规则对特征点集合进行有效性验证以提高特征点的可靠性。利用通过验证的特征点对全局运动进行估计, 得到精确的运动参数并据此对视频图像进行运动补偿。对于运动补偿造成的无定义区, 首先计算当前帧的定义区与相邻帧的光流, 以此为向导腐蚀无定义区; 利用拼接的方法, 填充仍为无定义区的像素。实验结果表明该算法对于前景物体运动具有较好的鲁棒性并能够生成内容完整的稳定视频序列。

关键词: 视频稳定; KLT 算法; 光流

中图法分类号: TP391 **文献标志码:** A **文章编号:** 1006-8961(2010)09-1384-07

Video stabilization algorithm content complete

PENG Yi, YE Qixiang, HUANG Jun, JIAO Jianbin

(Graduate University of Chinese Academy of Science, Beijing 100049)

Abstract: Although there are many feature matching and tracking methods so far, the side effect of moving foreground object, which will cause global motion estimation error, is still an open problem. In order to avoid features, located on the foreground objects, participating in motion estimation, feature effectiveness evaluation is employed to improve feature reliability for the features extracted by the traditional KLT method. Effective features are utilized to estimate global motion and obtain accurate motion parameter, based on which video frames are compensated. However, motion compensation will cause undefined area. There are some approaches to reconstruct the undefined area; nevertheless, they have not considered the effect of fast moving objects in the foreground of the video, which will decrease the video quality after stabilization and content completion. In our proposed algorithm, optical flow between defined areas of current frame and neighbor frame is first calculated, and then it is used as a guide to erode unknown areas. Finally, mosaicking on the base of reference frame is used to obtain a complete video stabilization sequence. Experiment results show that the proposed method is robust to moving foreground objects and is able to realize video frames stabilization with complete content.

Keywords: video stabilization; KLT; optical flow

0 引 言

在视频拍摄过程中, 摄像机载体的抖动会导致所拍摄视频的不稳定。这不仅会影响视频的视觉效果, 而且对于后端的图像处理也会带来不利影响。因此, 如何有效消除视频图像序列中因摄像机不稳

定而造成的图像抖动已成为视频处理领域的一个重要研究问题。常用的视频稳定算法有特征点匹配^[1-2]、灰度投影法^[3]和块匹配^[4-5]等。其中, 特征点匹配由于对不同场景中的应用有较好的适应性而得到了广泛的研究和应用。该方法不仅可以计算出图像序列帧间的平移运动, 还可估算图像的旋转和缩放运动, 实现 3 维稳像。

基金项目: 国家自然科学基金项目(60672147, 60872143)。

收稿日期: 2009-01-07; **改回日期:** 2009-05-07

第一作者简介: 彭 艺(1983—), 女。中国科学院研究生院硕士研究生。主要研究方向为数字图像处理。E-mail: pengyi06@mails.gucas.ac.cn。

在稳像算法中,运动估计是最为关键的一步。运动估计的准确与否,直接影响稳像的效果。如果只有全局运动的特征点参与运动估计的计算,无疑会得到比较高的运动估计精度。然而,当视频中存在前景运动物体时,这些物体会出现被遮挡或移出视场等现象,若前景运动物体上的特征点参与了运动估计,算法的可靠性将受到较大的影响。目前的基于特征点匹配的稳像算法^[1-2,6]大多没有考虑到运动前景物体的负面影响。可见,开发一种能够有效去除局部运动特征点的算法是非常必要的。

由于稳像只能稳定成像面内的图像,所以当摄像机晃动时,会有一部分图像信息在旋转和平移补偿后丢失,稳定后的视频上容易形成无定义区域。因此,为了达到较好的视觉效果及信息的完整性,需要一种算法实现视频完整。

考虑到以上两个问题,提出了更有效地去除异常点和运动前景特征点及相邻帧“腐蚀”和参考帧“拼接”的方法。通过大数定律和聚类算法对特征集进行有效性判定去除无效特征点,从而得到有效特征点。受 Matsushita 等人^[7]的启发,提出相邻帧“腐蚀”和参考帧“拼接”的方法实现视频内容的完整性。实验表明,该方法对于拍摄场景边缘存在局部运动物体时也具有一定的适应性。

1 稳定算法

视频稳定的关键问题是首先利用帧间图像的差别获得精确的全局运动参数。在进行全局运动估计时,首先需要建立描述摄像机运动的参数模型。综合考虑精度与实时性的要求,本文采用六参数仿射模型。该模型能很好地描述摄像机的平移、旋转和镜头摇动等运动。具体表达为

$$X' = W(X, A, D) = ZX + D = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} X + \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \quad (1)$$

式中, $X = (x, y)^T$, $X' = (x', y')^T$ 分别表示某个像素点在参考帧和当前帧的坐标, W 表示当前帧相对于参考帧的运动, $a_1 \sim a_4$ 是描述旋转、小范围缩放和形变的参数, $d_1 \sim d_2$ 是描述视频平移的参数。

确定好运动模型之后,算法流程(图1)可描述如下:

- 1) 对输入视频序列进行高斯平滑去除噪声;
- 2) 采用 Shi-Tomasi 算法提取特征点;
- 3) 利用 KLT 匹配器^[1,8]确定已选取的特征在

后续视频图像中对应的位置信息,得到特征初步匹配结果;

4) 通过特征有效性判定规则对特征集进行有效性验证以减少局部运动物体因遮挡或移出视场等产生的无效特征点,获取有效特征集合;

5) 对于有效特征集采用金字塔分层技术进行帧间运动估计,求得初始的全局运动参数;

6) 根据运动参数判断视频序列是否发生抖动,若无抖动,直接输出当前图像,并读取下一帧,否则转向运动平滑;

7) 对各个参数轨迹利用均值滤波获得平滑参数,对初始全局运动参数与平滑后参数求差,得到抖动参数;

8) 利用此参数进行运动补偿,得到具有信息缺失的视频流;

9) 最后,利用相邻帧“腐蚀”无定义区及参考帧“拼接”得到完整的稳定视频图像。

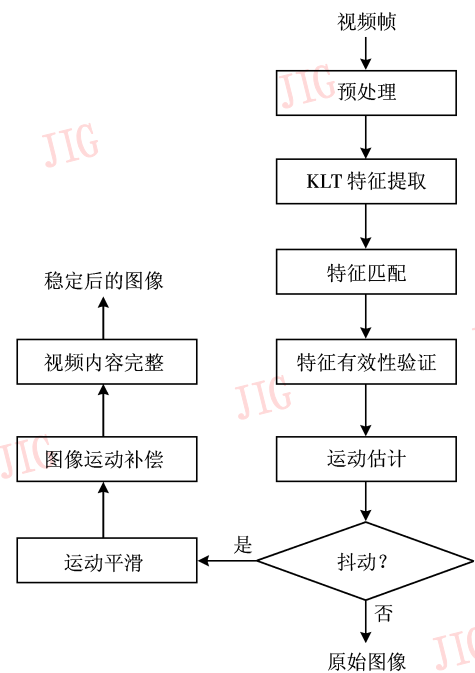


图1 算法流程图

Fig. 1 Flow chart of video stabilization algorithm

2 特征点有效性判定

2.1 基于大数定律的特征点筛选^[8]

特征被遮挡或者消失及运动模型与实际运动差异等因素会造成特征的误匹配,从而给全局运动带来偏差,进而影响稳像效果。为了避免位于运动前

景上的特征参与后续计算,使用特征有效性判定规则对提取的特征进行筛选。首先,根据大数定律对 KLT 提取的特征点进行分析,去除异常点。

假设 $\{\Omega_i\}$ ($i=0,1,\dots,N$) 表示从图像中提取的 N 个特征窗 (7×7 像素), I_k 表示第 K 帧图像坐标 X 的像素灰度值, I_T 表示参考帧图像坐标 X 的像素灰度值。 A 表示特征 Ω_i 的旋转和变形参数, D 表示平移参数。特征 Ω_i 在第 K 帧和模板帧 T 的误差平方和 ε_i 为

$$\varepsilon_i = \sum_{X \in \Omega_i} [I_k(W(X,A,D)) - I_T(X)]^2 \quad (2)$$

对于误差 ε , 根据观察和实验, 可以假定其分布服从高斯分布, μ, σ 为其期望和均方差。基于高斯分布的思想, 区间 $[\mu - 3.5\sigma, \mu + 3.5\sigma]$ 以外的数据出现的概率非常小, 是小概率事件。因此认为当特征 Ω_i 的误差 ε_i 在该区间以内时, 即

$$|\varepsilon_i - \mu| < 3.5\sigma \quad (3)$$

此特征为有效特征, 否则为无效特征。至此, 根据大数定律的思想, 根据特征匹配误差集合自身的分布规律, 提出了特征的有效性判定准则之一。

假设 V_{Ω_i} 表示某一特征 Ω_i 的运动向量, $\|V_{\Omega_i}\|$ 表示向量模, $Ang(V_{\Omega_i})$ 表示向量方向, $(\mu_{\|}, \sigma_{\|})$ 和 $(\mu_{Ang}, \sigma_{Ang})$ 分别表示它们的均值和方差, 由此可定义特征点有效性判定函数如式(4)。

根据拉依达准则^[9], 当特征数目大于 10 时, 系数 α 可以取 3.0。当 $E(\Omega_i)$ 为 1 时表示特征 Ω_i 有效, 否则特征 Ω_i 无效。即某特征点运动向量的模和方向同全局运动向量的偏差超出某一阈值, 该特征点将被剔除。

$$E(\Omega_i) = \begin{cases} 1 & \begin{cases} |\|V_{\Omega_i}\| - \mu_{\|}| < \alpha \times \sigma_{\|} \text{ 且} \\ |Ang(V_{\Omega_i}) - \mu_{Ang}| < \alpha \times \sigma_{Ang} \end{cases} \\ 0 & \text{其他} \end{cases} \quad (4)$$

2.2 基于 K-means 聚类算法的特征点筛选

根据 K-means 聚类的思想, 提出了特征点有效性判定准则之二。利用 K-means 聚类对初步验证有效的特征点 $\{\Omega_k\}$ 的运动矢量集 $\{V_{\Omega_k}\}$ 进行聚类, 运动矢量集 $\{V_{\Omega_k}\}$ 按照样本间相似性分成两类 (假定只有一个运动前景物体)。然后计算分类后特征点运动矢量集的内散度。类内散度大的, 认为其为背景; 类内散度小的, 认为其为前景, 并剔除之。

令 V_{Ω} 表示待分类的数据样本集, 设 V_{Ω} 现在在样本子集 Γ_i 中。 N_i 是第 i 聚类 Γ_i 中的样本数目,

m_i 是这些样本的均值, 即

$$m_i = \frac{1}{N_i} \sum_{V_{\Omega} \in \Gamma_i} V_{\Omega} \quad (5)$$

把 Γ_i 中的各样本 V_{Ω} 与均值 m_i 间的误差平方和对所有类 $\Gamma_1, \Gamma_2, \dots, \Gamma_C$ 相加作为聚类准则

$$J_e = \sum_{i=1}^c \sum_{V_{\Omega} \in \Gamma_i} \|V_{\Omega} - m_i\|^2 \quad (6)$$

式中, J_e 表示用 C 个聚类中心 m_1, m_2, \dots, m_c 代表 C 个样本子集 $\Gamma_1, \Gamma_2, \dots, \Gamma_C$ 时所产生的总的误差平方。

经过迭代, 寻求使 J_e 极小的最优聚类划分 $\Gamma_1, \Gamma_2, \dots, \Gamma_C$ 。

对聚类后的每一类 Γ_i , 根据其每一样本的位置 X_i^j 计算其类内散度:

$$J_i = \frac{1}{n} \sum_{j=1}^n (X_i^j - u_i)(X_i^j - u_i)^T \quad (7)$$

式中, X_i^j 为第 i 类的第 j 个样本, u_i 为第 i 类 X_i^j 的均值。

比较各个类的类内散度 J_i 的大小, 经验地认为, 值最大的一类为背景, 其余的为前景。即类内散度最大的一类为有效特征点的运动矢量集, 保存对应的特征点; 较小的类为无效特征点的运动矢量集, 剔除其对应的特征点。

实验证明, 通过以上两种算法的筛选, 异常点及前景运动物体上的特征点都能很好地被剔除。

3 稳定视频的内容完整性

对视频序列进行稳定后, 当稳定后图像的像素点不在参考图像坐标系范围内时, 该像素点灰度值为“空”, 在图像边缘形成“无定义”的待补偿区域。“无定义区”的存在, 不仅丢失了部分图像信息, 而且降低了视频的可视质量。为了获得高质量的视频, 以视频帧间具有较强的相关性为根据, 先用相邻帧腐蚀当前帧的“无定义区”, 然后再用参考帧拼接^[10]未完整的当前帧。

假定稳定视频帧有像素值的区域为“定义区”, 无像素值的区域为“无定义区”, 当前帧 I_t 无像素值而相邻帧 $I_{t'} (t-k \leq t' \leq t+k)$ 有像素值的区域为“待填充区”。实验中, 设 $k=6$ 。当前帧和相邻帧的公共“定义区”的光流用 Lucas-Kanade 光流法^[11-12]求出。

对当前帧的“待填充区”进行腐蚀, 如图 2 所

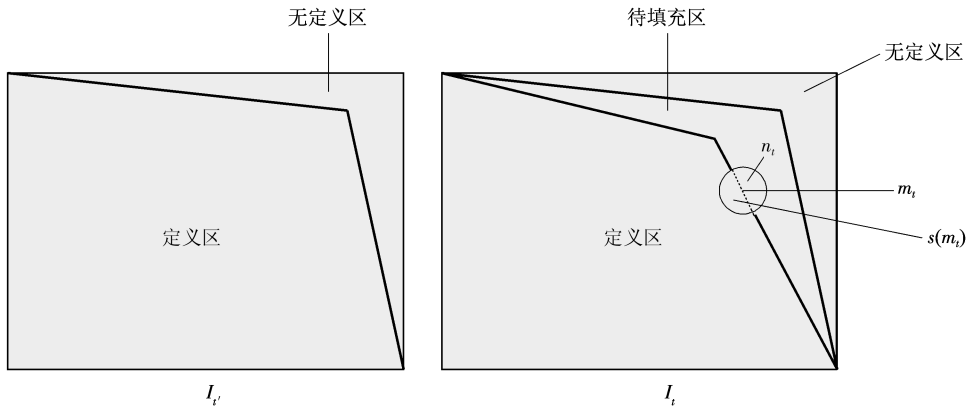


图 2 相邻帧 $I_{t'}$ 腐蚀当前帧 I_t 示意图

Fig. 2 Diagram of eroding current frame I_t with neighbor frame $I_{t'}$

示。首先,假设 m_t 为“待填充区”和“定义区”交界线上一点, $s(m_t)$ 为 m_t 的小邻域, n_t 为 $s(m_t)$ 内一点, $F_t^{t'}(n_t)$ 为第 t 帧的点 n_t 到第 t' 帧的光流。若 n_t 是在“定义区”,则其光流由金字塔中的迭代 Lucas-Kanade 求得;若 n_t 在“无定义区”,则其光流为 0。点 m_t 的光流可表示为

$$F_t^{t'}(m_t) = \frac{\sum_{n_t \in S(m_t)} w(m_t, n_t) F_t^{t'}(n_t)}{\sum_{n_t \in S(m_t)} w(m_t, n_t)} \quad (8)$$

$$w(m_t, n_t) = e^{-\frac{\|m_t - n_t\|^2}{2\sigma^2}} \quad (9)$$

式中, $F_t^{t'}(m_t)$ 为第 t 帧的点 m_t 到第 t' 帧的光流; $w(m_t, n_t)$ 是权重函数。实验中,设 σ 为 $s(m_t)$ 半径。

以“待填充区”的光流为向导,可以通过线性插值法计算当前帧的 m_t 点的灰度值:

$$I_t(m_t) = I_{t'}(W(F_t^{t'}(m_t), A, D)) \quad (10)$$

式中, W 是第 t' 帧相对于第 t 帧的运动。

同样可计算出“待填充区”和“定义区”交界线的所有像素点灰度值,并把它归为“定义区”。依此循环操作,直至“待填充区”完全被腐蚀。

用相邻帧腐蚀当前帧后,仍有缺失像素值的“无定义区”可由图像拼接技术的方法得到:

$$I_t(m_t) = I_{t'}(W(m_t, A, D)) \quad (11)$$

这样,就得到了全帧图像的的稳定视频序列。

4 实验结果与分析

选用多个分辨率为 320×240 像素的视频图像序列,在主频 3.39 GHz,内存 1 GB 的 PC 机上用 VC++ 编程进行仿真。该视频序列为移动的室内、

外环境下具有不同前景运动物体。

图 3 为运动前景特征点有效性判定的实验结果。图 3(a)是参考图像特征提取的结果,为 100 个特征点。图 3(b)是由 KLT 算法获得的初步匹配的结果,为 84 个特征点,可以看出有 5 个特征点坐落在运动前景上(运动的人)。图 3(c)为用大数定律筛选后的结果,为 82 个特征点。图 3(d)显示了本文算法提取的特征点,可以看出运动前景上的特征点大多被剔除。验证了本文算法的有效性。

图 4 为实现稳定的视频内容完整的实验过程。其中(a)为具有“无定义区”的稳定后的视频帧;(b)为用一帧相邻帧腐蚀的结果;(c)为多个相邻帧腐蚀后再被拼接的结果。

为了验证提出的有效提取特征点,消除异常点及前景局部运动点算法。采用峰值信噪比 PSNR 来定量分析视频稳定效果。其定义如下:

$$PSNR = 10 \log(255^2 / MSE) \quad (12)$$

$$MSE = \sum_{i=1}^N \sum_{X \in \Omega_i} [I_K(W(X, A, D)) - I_T(X)]^2 \quad (13)$$

式中, MSE 为当前帧和参考帧的均平方误差。 $\{\Omega_i | i=0, 1, \dots, N\}$ 表示从图像中提取的 N 个特征窗, I_K 表示第 K 帧图像坐标 X 的灰度值, I_T 表示参考帧 T 图像坐标 X 的灰度值。PSNR 越大,图像间的差别越小,图像稳定效果越好。

图 5 中,所用的视频图像是既有全局运动又有前景局部运动的复杂序列。算法 1 是提出的经过特征点有效判定后的算法,算法 2 是特征点未经有效判定的经典 KLT 稳像算法。从结果可以看出,本文算法的 PSNR 要高于经典的 KLT 算法。

图 6 为采用上述算法所得到的视频稳定结果,



(a) 参考帧的特征提取结果



(b) 特征匹配初步结果



(c) 经大数定律筛选结果



(d) 本文算法结果

图 3 特征点有效性判定

Fig.3 Feature effectiveness evaluation



(a) 具有“无定义区”的稳定视频帧



(b) 一帧相邻帧腐蚀的结果



(c) 腐蚀和拼接的最终结果

图 4 视频内容完整

Fig.4 Video content completion

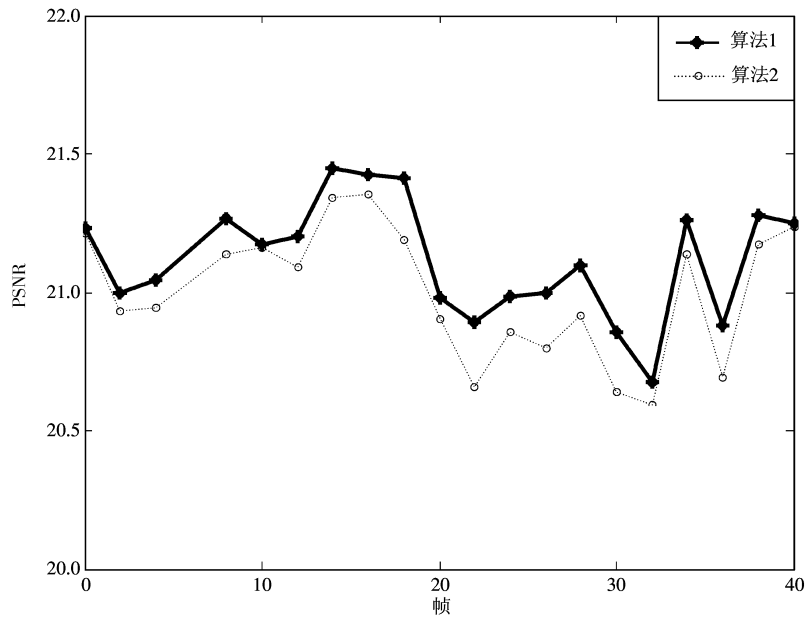
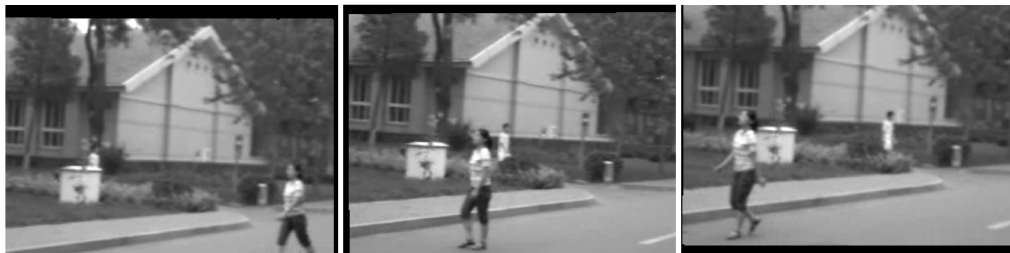


图 5 图像序列 PSNR 结果比较

Fig. 5 Comparison of image sequence PSNR result



(a) 原视频序列



(b) 稳定后的视频序列



(c) 内容完整的稳定视频序列

图 6 视频序列稳定并实现全帧的结果

Fig. 6 Result of video stabilization with content completion

即对具有抖动的视频实现了内容完整的稳定视频序列。第 1 行是具有前景运动的视频序列,第 2 行是稳定后的具有“无定义区”的稳定视频序列,第 3 行是内容补充完整的稳定视频序列。

5 结 论

提出了一种基于经典 KLT 匹配算法的视频完整的算法。在该算法中,首先用大数定律和聚类算法对提取的特征点进行有效性分析,这对于具有局部运动的抖动图像序列,更有利于进行精确的全局运动估计。然后,对稳定后的具有“无定义区”的视频帧进行完整性处理。实验结果验证了本文算法的有效性,并且算法对于在场景中出现的较快运动前景也具有较好的鲁棒性。

参考文献 (References)

- [1] Shi J, Tomasi C. Good Features to Track[DB/OL]. (1994-06-21) [2009-04-17]. http://ieeexplore.ieee.org/search/srchabstract.jsp?tp=&arnumber=323794&queryText%3DGood+Features+to+Track%26openedRefinements%3D*%26searchField%3DSearch+All.
- [2] Hu R, Shi R, Shen I, et al. Video Stabilization Using Scale-Invariant Feature. [DB/OL]. (2007-07-04) [2009-04-17]. http://ieeexplore.ieee.org/search/srchabstract.jsp?tp=&arnumber=4272080&queryText%3DVideo+Stabilization+Using+Scale-Invariant+Feature%26openedRefinements.%3D*%26searchField%3DSearch+All.
- [3] Maheux J, Cruickshank J M, Seigny L. Video-rate image Stabilization system[J]. SPIE Proceedings Series, 1998, 3414: 232-240.
- [4] Vella F, Castorina A, Mancuso M, et al. Digital image stabilization by adaptive block motion vectors filtering[J]. IEEE Transactions on Consumer Electronics, 2002, 48(3): 796-801.
- [5] Xu Lidong, Fu Fangwen, Lin Xinggong. Circular Block Matching based Video Stabilization[DB/OL]. (2005-07-12) [2009-04-17]. <http://www.engineeringvillage2.org.cn/controller/servlet/Controller?SEARCHID=98ce7e1283fcf237bM7140c3915203&CID=quickSearchDetailedFormat&DOCINDEX=1&database=49153&format=quickSearchDetailedFormat>.
- [6] Jin J S, Zhu Z G, Xu G Y. A stable vision system for moving vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2000, 1(1): 32-39.
- [7] Matsushita Y, Ofek E, Tang Xiao'ou, et al. Full-frame video stabilization with motion inpainting[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(7): 1150-1163.
- [8] Huang Yabo, Jiao Jianbin, Ye Qixiang, et al. A robust video stabilization algorithm[J]. Journal of Optoelectronics. Laser, 2008, 19(3): 394-399. [黄亚博, 焦建彬, 叶齐祥, 等. 一种鲁棒性的抖动视频稳像算法[J]. 光电子·激光, 2008, 1999(3): 394-399.]
- [9] Xiao Mingyao. Error and Treatment of Data Experiment[M]. Beijing: Science Press, 1980. [肖明耀. 试验误差估计与数据处理[M]. 北京: 科学出版社, 1980.]
- [10] Livin A, Konrad J, Karl W. Probabilistic Video Stabilization Using Kalman Filtering and Mosaic[DB/OL]. (2003-01-21) [2009-04-17]. <http://www.engineeringvillage2.org.cn/controller/servlet/Controller?SEARCHID=164debb1283f0feb9dM6ecec411510&CID=quickSearchDetailedFormat&DOCINDEX=1&database=49153&format=quickSearchDetailedFormat>.
- [11] Lucas B D, Kanade T. An Iterative Image Registration Technique with an Application to Stereo Vision[EB/OL]. (2003-01-21) [2009-04-17]. <http://dli.iit.ac.in/ijcai/IJCAI-81-VOL-2/PDF/017.pdf>.
- [12] Bouguet J Y. Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the Algorithm[EB/OL]. [2009-04-17]. http://robots.stanford.edu/cs223b04/algo_tracking.pdf.