

一种基于特征跟踪的视频序列稳像算法*

史 阳 高新波

(西安电子科技大学电子工程学院, 西安 710071)

摘 要 提出一种基于特征跟踪的视频序列稳像算法. 该算法从视频序列的参考帧中提取出一组角点特征, 然后在后续帧中基于模糊 Kalman 滤波进行特征窗跟踪, 通过比较各帧图像中特征窗间的对应关系计算出补偿摄像机运动所必需的参数, 使用这些参数将后续帧向参考帧对准, 从而得到稳定的视频序列. 实验结果表明该算法稳像效果好, 运算复杂度低, 且具有较强的鲁棒性.

关键词 稳像算法; 角点检测; 模糊 Kalman 滤波; 运动估计; 运动补偿

中图分类号 TP751 **文献标识码** A

0 引言

很多时候, 摄像机的意外晃动会产生抖动的视频序列, 导致观看者产生头晕等不适. 数字稳像就是用来补偿抖动图像获得序列稳定显示的技术, 现已广泛应用于军事侦察, 机器人导航, 全景图拼接, 目标跟踪等方面, 是计算机视觉领域的研究热点.

关于稳像问题, 研究人员提出了许多算法^[1~9], 包括代表点匹配法^[1], 投影算法^[2], 位平面匹配法^[3], 光流场法^[7], 频域估计法^[8]等等. 其中特征跟踪法^[4,5]受到广泛关注. 特征跟踪法的基本思想是通过跟踪特征点求取摄像机运动参数, 进行运动补偿, 从而得到稳定的显示. 该方法的难点在于特征点的选取和跟踪算法的精度, 不理想的特征检测和跟踪算法都无法得到好的稳像效果. 为此, 本文提出一种基于角点检测和特征窗跟踪的稳像算法. 该方法以参考帧中检测到的角点为特征, 在后续帧中基于模糊 Kalman 滤波实现特征窗的跟踪, 同时利用 X84 规则进行特征窗的优选, 最后综合对应特征窗的运动信息估计摄像机的运动参数, 从而利用运动补偿技术将后续帧向参考帧对准, 消除视频序列图像间的抖动.

1 特征检测

本文提出的稳像算法包括特征窗选取、特征跟踪和运动补偿三个主要模块, 算法流程图如图 1. 其中, 特征选取的好坏对于后续的特征跟踪和运动补偿等步骤具有重要意义. 基于二维特征的方法由于度量运动时不受孔径问题影响^[10], 因而被广泛采用, 最具代表性的二维特征就是角点.

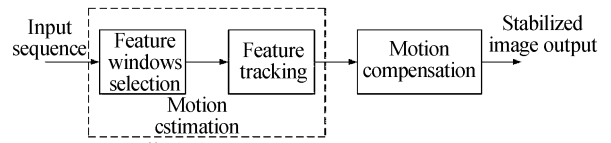


图 1 算法流程图

Fig. 1 Flow chart of the proposed algorithm

这里, 采用 SUSAN (Smallest Univalued Segment Assimilating Nucleus) 准则^[11]完成角点检测. 它直接对图像灰度值进行操作, 方法简单, 定位准确, 对多个区域的结点也能精确检测, 并且对局部噪声不敏感, 抗噪能力强.

SUSAN 准则的原理如图 2, 用一个圆形模板遍历图像, 若模板内其他任意像素的灰度值与模板中心像素(核)的灰度值的差小于一定阈值, 就认为该点与核具有相同(或相近)的灰度值, 满足这样条件的像素组成的区域称为 USAN(Univalued Segment Assimilating Nucleus)区域. 把图像中的每个像素与具有相近灰度值的局部区域相联系是 SUSAN 准则的基础.

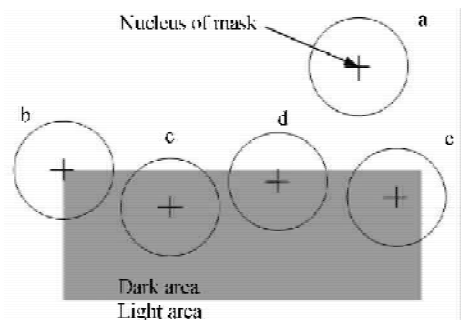


图 2 SUSAN 原理

Fig. 2 The principle of SUSAN algorithm

式(1)定义了模板内属于 USAN 区域的像素的判别函数

$$c(\mathbf{r}, \mathbf{r}_0) = \exp \left(- \left(\frac{I(\mathbf{r}) - I(\mathbf{r}_0)}{t} \right)^6 \right) \quad (1)$$

式中, $I(\mathbf{r}_0)$ 是核的灰度值, $I(\mathbf{r})$ 为模板内其他任意像素的灰度值, t 是灰度差门限, 决定了检测出的特征的最小对比度和可忽略的噪声的最大量.

*国家自然科学基金(No. 60202004)资助项目
Tel: 029-88201838 Email: xbgao@ieee.org
收稿日期: 2004-05-20

图像中某一点 r_0 对应的 USAN 区域的大小由式(2)确定

$$n(r_0) = \sum_r c(r, r_0) \quad (2)$$

USAN 区的大小反映了图像局部特征的强度, 当模板完全处于背景或目标中时, USAN 区域最大(如图 2 中 a), 当模板移向目标边缘时, USAN 区域逐渐变小(如图 2 中 c, d, e), 当模板中心处于角点位置时, USAN 区域很小(如图 2 中 b). 得到每个像素对应的 USAN 区域大小后, 利用下式产生初始角点响应

$$R(r_0) = \begin{cases} g - n(r_0) & \text{if } n(r_0) < g \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

式中, g 为几何门限, g 越小, 检测到的角点越尖锐.

去除一些误判后, 执行非最大抑制操作即可得到最终的角点响应. 以每个检测到的角点位置为中心(称为特征点), 开辟一个 $N \times N$ 的窗 w 作为特征匹配和跟踪的对象, 就是特征窗.

2 特征跟踪

完成特征提取后, 就需要在后续帧中跟踪这些特征窗以预测和估计特征点的运动轨迹, 从而得到当前帧和参考帧之间的对应关系, 实现运动补偿. 本文采用模糊 Kalman 滤波器^[12,13]进行特征窗跟踪.

假设视频序列为 $f_0, f_1, f_2 \dots f_k \dots$, Kalman 滤波的状态矢量定义为

$$X_k = [x_k, y_k, u_k, v_k, \alpha_k, \beta_k] \quad (4)$$

这里 $(x_k, y_k), (u_k, v_k), (\alpha_k, \beta_k)$ 分别表示第 k 帧跟踪到的任一特征点的位置, 速度和加速度. 则 Kalman 滤波的状态转移矩阵和观测矩阵分别为

$$\Phi = \begin{bmatrix} 1 & 0 & 1 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (6)$$

给出经典 Kalman 滤波递推公式如下(Q_k 和 R_k 是噪声协方差矩阵)

$$\bar{X}_k = \Phi \bar{X}_{k-1} + K_k (Z_k - H \Phi \bar{X}_{k-1}) \quad (7)$$

$$K_k = P'_k H^T (H P'_k H^T + R_k)^{-1} \quad (8)$$

$$P'_k = \Phi P_{k-1} \Phi^T + Q_{k-1} \quad (9)$$

$$P_k = (I - K_k H) P'_k \quad (10)$$

如果在经典 Kalman 滤波中, 能够利用一些额外信息, 形成某种形式的隶属度函数和模糊推理规则, 且加以有效利用, 那么就相当于充分融合了信

息, 使 Kalman 滤波执行的精度提高. 因此, 文献[12]提出了一种基于模糊控制的 Kalman 滤波方法. 其本质就是基于滤波处理的数据残差构造模糊算法, 以自适应的控制 Kalman 滤波的相应加权系数, 用来剔除异常的观测数据和对估计结果进行加权调整.

图 3 描述了所构造的模糊规则的隶属度函数.

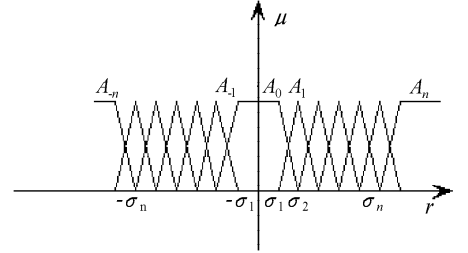


图 3 隶属度函数

Fig. 3 The membership functions

其中 $A_i; i=0, \pm 1, \pm 2, \dots, \pm n$ 是论域的模糊子集, $0 \leq \mu \leq 1$ 表示 r_k 落在 A_i 内的可能性.

数据残差 r_k 定义为

$$r_k = \| Z_k - H \Phi \bar{X}_{k-1} \| \quad (11)$$

模糊规则: 如果 r_k 落在 A_i 内, 则相应的观测数据 Z_k 无效的可能性为 F_i , 那么 $\alpha = g(F_i)$, 其中, α 是将在迭代算法中用于增益矩阵 K_k 的加权系数.

算法的具体步骤如下

1) 执行常规的 Kalman 滤波算法; 2) 根据式(11)计算数据残差 r_k ; 3) 构造模糊算法, 计算 σ_i 和 α ; 4) 更新一步递推公式(7)为

$$\bar{X}_k = \Phi \bar{X}_{k-1} + \alpha K_k (Z_k - H \Phi \bar{X}_{k-1}) \quad (12)$$

5) 重复上述步骤作下一时刻的估计.

状态协方差矩阵 P_k 描述了当前状态的不确定性. 中心为估计状态 \bar{X} , 包含真实状态的可能性为 c^2 的区域是如下表达式描述的一个椭圆

$$(X - \bar{X}) P_k^{-1} (X - \bar{X})^T \leq c^2 \quad (13)$$

在此椭圆范围内依据最小误差平方和 SSD (sum of squared differences, 见式(15)), 准则搜索匹配窗, 匹配窗中心位置就是相应的观测矢量, 经过滤波得到的状态矢量给出了所跟踪的特征点在当前帧中的相应位置.

3 运动补偿

假设摄像机面对一个近似平坦的场景, 或者绕着其光轴旋转拍摄, 则不同帧对应的图像点间为线性投影变换. 为了补偿摄像机的相对运动, 我们需要在当前帧和参考帧之间建立反映其像素对应关系的变换矩阵. 我们假设两帧 F_k 和 F_l 中相应的 n 对特征点分别为 $P_k^1, P_k^2, \dots, P_k^n$ 和 $P_l^1, P_l^2, \dots, P_l^n$, 变换矩阵为 M , 则 $P_k = M P_l$, 即

$$\begin{bmatrix} x_k^1, x_k^2, \dots, x_k^n \\ y_k^1, y_k^2, \dots, y_k^n \\ \omega_k^1, \omega_k^2, \dots, \omega_k^n \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} x_l^1, x_l^2, \dots, x_l^n \\ y_l^1, y_l^2, \dots, y_l^n \\ \omega_l^1, \omega_l^2, \dots, \omega_l^n \end{bmatrix} \quad (14)$$

这里的点用齐次坐标表示,相应的笛卡儿坐标为 $(x/\omega, y/\omega)$. 求得描述运动的 8 个参数用于运动补偿. 我们假设序列中每一帧都和参考帧有部分重叠,则补偿后,重叠部分的静态景物点在理想情况下是不动的.

为了减少稳像处理过程对于可能出现的跟踪错误(由于特征窗被遮挡或者消失),我们采用一个简单有效的方法,叫做 X84 规则^[14],来去除虚假跟踪的特征.

对应特征窗间的 SSD 定义如下

$$\epsilon = \sum_w [I_k - I_0]^2 \quad (15)$$

I_k 和 I_0 分别代表当前帧和参考帧内对应特征窗的像素亮度值.

根据分析, ϵ 符合自由度为 N^2 的 χ^2 分布,即 $\epsilon \equiv \chi^2(N^2)$, 而自由度大于 30 的 χ^2 分布近似为高斯分布. 特征窗内的像素点个数远远大于 30, 因此, 我们假设跟踪良好的特征窗的 ϵ 值都符合高斯分布: $\epsilon \equiv \eta(N^2, 2N^2)$.

此规则抛弃当前帧中那些 ϵ 值偏离中值大于 $k * MAD$ (Median Absolute Deviations) 的特征窗, MAD 定义为

$$MAD = med_i \{ | \epsilon_i - med_j \epsilon_j | \} \quad (16)$$

式中, ϵ_i 是当前帧的第 i 个特征窗和参考帧对应特征窗之间的 SSD 值. 当 $k = 5.2$ 时, 对应高斯分布大约 3.5 倍的标准差, 而 $[\mu - 3.5\sigma, \mu + 3.5\sigma]$ 区间包含多于 99.9% 的数据. 抛弃虚假跟踪的特征点后, 使用剩余的特征点来计算变换矩阵 M .

4 实验结果及分析

为了验证本文提出的稳像算法的有效性, 我们选用三段代表不同运动情况的视频序列作为测试数据(来自意大利 Udine 大学机器视觉实验室, <http://mvl.dimi.uniud.it/>).

图 4 给出镜头平移的 PARK 序列的特征窗跟踪效果, 白色框代表特征窗. 图 4(a) 标识出从参考帧中提取的特征窗; 图 4(b) 标识出在第 40 帧跟踪到的特征窗. 两幅图像中对应特征窗包含相同的内容, 反映了跟踪算法的效果. 图 5 以镜头旋转的 TOWN 序列为例, 给出了稳定后的图像帧和全景图拼接的效果, 白色框代表参考帧的位置. 图 5(a) 显示第 50 帧经过稳像处理后的效果; 图 5(b) 给出序列的拼接图效果. 经过稳像处理后, 所有图像落在

白色框内的内容是一致的, 达到了稳像目的. 图 6 以镜头内有运动目标的 BOOK 序列为例(图 6(b) 中黑色物体 BOOK 进入画面), 给出序列的特征窗跟踪效果, 白色框代表特征窗. 图 6(a) 标识出从参考帧中提取的特征窗, 图 6(b) 第 15 帧中运动物体遮挡了背景上的部分特征窗, 无法正确跟踪, 因而根据 X84 规则抛弃这些特征窗.

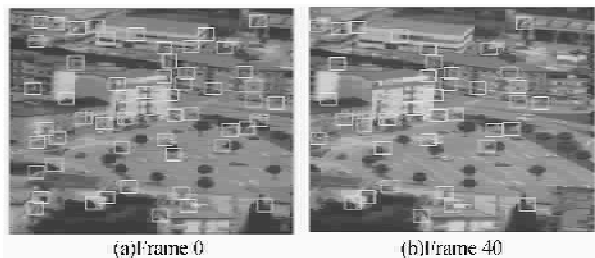


图 4 PARK 序列的特征跟踪

Fig. 4 Feature tracking of PARK sequence

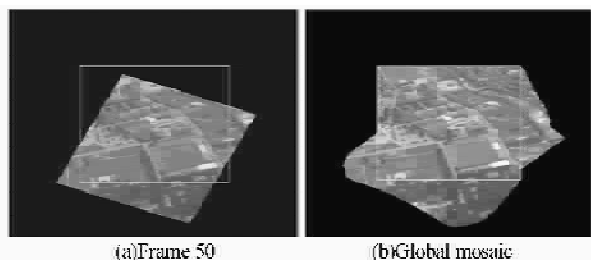


图 5 TOWN 序列稳定图像和拼接图像

Fig. 5 The stabilized image and global mosaic of TOWN sequence

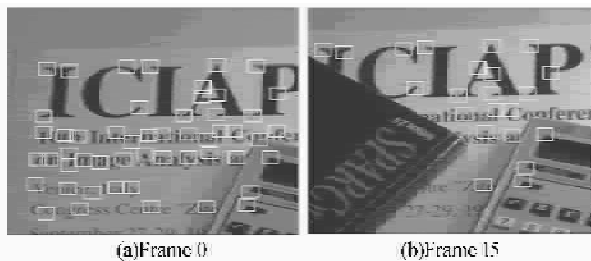


图 6 BOOK 序列的特征跟踪

Fig. 6 Feature tracking of BOOK sequence

根据 Morimoto 和 Chellappa^[15] 提出的评价准则, 使用峰值信噪比 (PSNR) 来评价稳像算法的性能. 我们知道, 当帧间运动被完全补偿时, 稳定后的图像的重叠部分在理想情况下对应的静态像素的差值都为零. 而实际情况往往并非如此, 这可能是由于噪声, 估计误差, 运动模型限制引起的失真等因素造成的. 考虑 PSNR 作为一个测度, PSNR 的定义式(假设对于 256 级灰度图像来说)如下

$$PSNR = 10 \log \frac{255^2}{MSE(I_1, I_0)} \quad (17)$$

式中

$$MSE(I_1, I_0) = \frac{1}{n} \sum_{i=1}^n (I_{1i} - I_{0i})^2 \quad (18)$$

为两幅图像之间的均方差.

如果 I_1 和 I_0 代表连续的两帧, 则称为帧间变

换保真度 ITF (Inter-frame Transformation Fidelity); 如果 I_1 代表当前帧, I_0 代表参考帧, 则称为全局变换保真度 GTF (Global Transformation Fidelity). 图 7 分别给出三段视频序列的 ITF 和

GTF 随帧数的变化情况. 其中 BOOK 序列中由于有运动物体, 中间若干帧图像中存在遮挡物, 而参考帧中没有, 对应像素值变化较大, 所以相应 GTF 值会有明显下降.

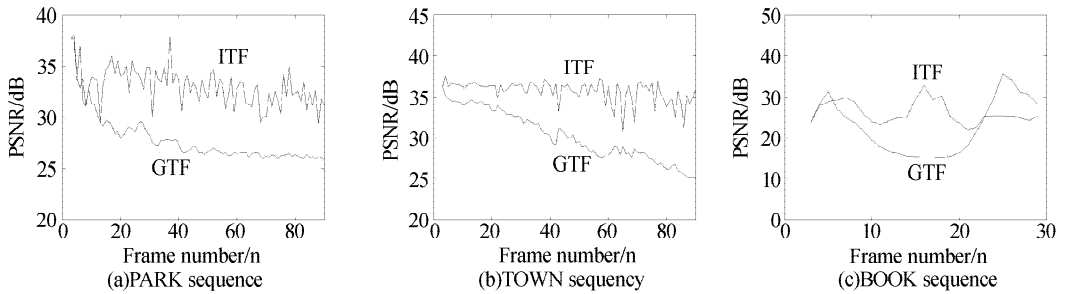


图 7 各序列的 ITF 和 GTF

Fig. 7 ITF and GTF of sequences

5 结论

本文提出了一种基于特征跟踪的视频序列稳像算法. 首先, 应用 SUSAN 准则在参考帧中提取特征, 然后, 基于模糊 Kalman 滤波在后续帧中跟踪特征窗, 根据对应的特征点坐标得到变换矩阵进行运动补偿, 从而达到稳定视频序列的目的. 实验证明, 该方法稳像效果好, 运算复杂度低, 且具有较强的鲁棒性.

参考文献

- Uomori K, Morimura A, Ishii H, *et al.* Automatic image stabilizing system by full-digital signal processing. *IEEE Trans on Consumer Electronics*, 1990, **36**(3): 510~519
- Sauer K, Schwartz B. Efficient block motion estimation using integral projections. *IEEE Trans on Circuits and Systems for Video Technology*, 1996, **6**(5): 513~518
- Ko S J, Lee S H, Lee K H. Fast digital image stabilizer based on gray-coded bit-plane matching. *IEEE Trans on Consumer Electronics*, 1999, **45**(3): 598~603
- Shi J, Tomasi C. Good features to track. *In CVPR'94*, 1994. 593~600
- Censi A, Fusiello A. Image stabilization by features tracking. 10th Inter Conf. on image analysis and processing, Venice Italy, 1999, 27~29, 665~667
- 刘雅轩, 苏秀琴, 王萍. 一种基于局部投影熵的图像匹配新算法. *光子学报*, 2004, **33**(1): 105~108
Liu Y X, Su X Q, Wang P. *Acta Photonica Sinica*, 2004, **33**(1): 105~108
- Chang J Y, Hu W F, Cheng M H, *et al.* Digital image translational and rotational motion stabilization using optical flow technique. *IEEE Trans on Consumer Electronics*, 2002, **48**(1): 108~115
- Ertürk S. Digital image stabilization with sub-image phase correlation based global motion estimation. *IEEE Trans on Consumer Electronics*, 2003, **49**(4): 1320~1325
- Vella F, Castorina A, Mancuso M, *et al.* Digital image stabilization by adaptive block motion vectors filtering. *IEEE Trans on Consumer Electronics*, 2002, **48**(3): 796~801
- Trucco E, Verri A. Introductory techniques for 3-D computer vision. Upper Saddle River, NJ, USA: Prentice-Hall, 1998
- Smith S M, Brady J M. SUSAN—a new approach to low level image processing. *Int Journal Computer Vision*, 1997, **23**(1): 45-78
- Chao-Yin H, Chi-Chih L. Analysis and design of fuzzy filter algorithms. *Int IEEE / IAS Conf*, 1995, 22~27, 413~420
- Kosko B. Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence. Upper Saddle River, NJ, USA: Prentice-Hall, 1992
- Fusiello A, Trucco E, Tommasini T, *et al.* Improving feature tracking with robust statistics. *Pattern Analysis & Applications*, 1999, **2**(4): 312~320
- Morimoto C, Chellappa R. Evaluation of image stabilization algorithms. *ICASSP'98*, 5: 2789~2792

A Video Stabilization Algorithm Based on Feature Tracking

Shi Yang, Gao Xinbo

School of Electronic Engineering, Xidian University, Xi'an 710071

Received date: 2004-05-20

Abstract A video stabilization algorithm is presented based on the feature window tracking. First, the proposed algorithm extracts the corner-points as feature from the specified reference frame. Then, based on the fuzzy Kalman filtering, the extracted features are tracked with window-matching method from the succeed frames. By registering the corresponding feature windows, the motion parameters of the camera can be calculated between the reference frame and succeed frames. Finally, vibration of images due to camera shake is stabilized with motion compensation. The experimental results illustrate that the proposed algorithm is effective and robust with low computational complexity.

Keywords Video stabilization; Corner detection; Fuzzy Kalman filtering; Motion estimation; Motion compensation



Shi Yang was born in 1981, in Shaanxi Province, China. He received Bachelor degree from Electronic Engineering, Xidian University in 2002. Currently he is working toward the Master degree in signal and information processing. His research interests include video processing and computer vision.