

doi:10.3788/gzxb20184701.0110002

基于最小生成树与改进卡尔曼滤波器的实时 电子稳像方法

谢亚晋, 徐之海, 冯华君, 李奇, 陈跃庭

(浙江大学 现代光学仪器国家重点实验室, 杭州 310027)

摘 要:为解决稳像过程中局部运动分量导致的全局运动估计不准确问题,提出了一种实时电子稳像方法.该方法提出基于最小生成树的特征点迭代筛选算法,采用相邻帧图像特征点最小生成树的相似度衡量特征点匹配精度,剔除错误匹配的特征点和局部运动前景上的特征点,避免了局部运动分量的影响.采用自适应加权法修正相邻帧之间的仿射变换矩阵,解决由于运动前景遮挡造成的背景特征点数量稀少进而导致的稳像晃动问题.针对相机跟拍与随机抖动分量混合问题,提出基于运动矢量队列的双卡尔曼滤波器,自适应地修正卡尔曼滤波器的测量噪声协方差,动态调整滤波平滑性能,有效处理同时包含相机跟拍运动与随机抖动分量的视频,保留相机跟拍分量.实验表明,该方法对于视频图像中包含局部前景运动和相机跟拍运动的情况,对比其他 3 种方法,仍可以保持良好的稳像效果;在 Intel Core i5 3.30 GHz CPU 下,对于 640×360 分辨率的彩色图像序列可达到 40FPS 的稳像帧率,并且运算过程中无需利用下一帧图像信息,具有实时稳像的优点.

关键词:视频信号处理;电子稳像;运动补偿;卡尔曼滤波;特征点筛选;运动估计;帧间保真度

中图分类号: TP391.4

文献标识码: A

文章编号: 1004-4213(2018)01-0110002-10

Real-time Video Stabilization Based on Minimal Spanning Tree and Modified Kalman Filter

XIE Ya-jin, XU Zhi-hai, FENG Hua-jun, LI Qi, CHEN Yue-ting

(State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China)

Abstract: To solve the inaccurate problem of global motion estimation due to local motion in video stabilization, a real-time stabilization method was proposed. The method proposes an approach of feature points iterative filtering algorithm based on minimal spanning tree and applies spanning tree similarity of successive frames to measure feature matching and discard wrong matched points and points in moving foreground. Then, an adaptive weighted method was applied to correct the transformation matrix to solve the wobble problem due to few feature points caused by foreground occlusion. Finally, an modified dual Kalman filters based on motion queue was proposed to correct measurement noise covariance adaptively and adjust smoothness of filter dynamically. It can process videos with intended camera motion and random jitters effectively. Experimental results show that the proposed method can perform well even in situation of local foreground motion and intended camera motion. Under circumstance of Intel Core i5 3.30 GHz CPU, the proposed method can stabilize videos of 640×360 resolution at a rate of 40FPS. And it has real-time advantage since next frame information is not needed during computation.

基金项目:装备预研教育部联合基金项目(No. 6141A02022307)和浙江省科技计划项目(No. 2017C01033)资助

第一作者:谢亚晋(1992-),男,硕士研究生,主要研究方向为光学成像工程与图像处理. Email: xieyajin@zju.edu.cn

导 师:徐之海(1964-),男,教授,博导,主要研究方向为光电成像技术及仪器,图像获取与处理. Email: xuzh@zju.edu.cn

通讯作者:冯华君(1963-),男,教授,博导,主要研究方向为光学成像工程与图像处理. Email: fenghj@zju.edu.cn

收稿日期:2017-07-25; **录用日期:**2017-09-06

<http://www.photon.ac.cn>

Key words: Video signal processing; Electronic image stabilization; Motion compensation; Kalman filters; Feature selection; Motion estimation; Interframe transformation fidelity

OCIS Codes: 100.2960; 100.4994; 110.4153; 110.4155

0 引言

拍摄过程中由于存在设备抖动,导致视频质量下降,影响后期观测和处理,稳像技术因此得到国内外研究者关注,其中电子稳像以其价格低、精度高、速度快的特点成为稳像技术的重要研究方向^[1],在航空航天、机器人、军事等领域有重要应用.电子稳像技术根据运动估计的方法可分为基于图像强度和基于特征点匹配两类^[2].

在基于图像强度的方法中,Chang 提出光流法^[3],但该方法容易受噪声和光强变化的影响,且迭代计算量大;Wang 与 Puglisi 提出的基于图像块匹配的方法^[4-5]简单、运算速度快,但全局运动矢量估计易受运动前景影响,Kwon^[6]在图像四个角落上选取图像块进行匹配,但运动物体仍会出现在所选的图像块中,影响块匹配精度.基于特征匹配法的主要思路是提取相邻帧图像的特征点并进行匹配,Battiatto 通过提取尺度不变特征(Scale-invariant feature transform, SIFT)进行匹配^[7],SIFT 具有尺度、旋转不变性,对亮度不敏感的优点^[8],但计算量大,无法实时处理;Zhu^[9]先检测出运动前景并剔除,提取背景区域的 Harris 特征进行配准来估计全局运动,但该方案无法处理背景运动较快的图像;Liu^[10]利用子空间约束平滑特征运动路径,但是单个子空间无法处理包含较大前景运动的视频,且随着相机运动大量特征点会离开视场,导致运动估计失败.Grundmann 提出 L1 相机路径优化算法^[11],即将相机运动路径分为静止、线性和抛物线三种路径并加入矩形包含、曲线贴近、显著点约束,采用 Lucas-Kanade 金字塔跟踪特征点,但是对于快速的相机移动或者特征稀少的图像会使得优化的相机路径变得不可靠.Jing 提出的实时稳像算法^[12]简化了运动模型,忽略相机运动及其带来的累计误差,但难以处理大相机运动的情况.

电子稳像技术中的一个严重问题是前景的局部运动分量会导致全局运动估计不准确,严重影响稳像效果.本文提出一种实时电子稳像方法,提出基于相邻帧最小生成树的特征点迭代筛选算法,有效剔除误匹配和运动前景上的特征点;其次,当运动前景占据图像大部分区域时,背景特征点稀少,将导致仿射变换矩阵不准确,造成图像的突然晃动,本文提出自适应仿射变换矩阵的修正方法,有效解决这个问题;最后采用双卡尔曼滤波器,提出一种基于运动矢量队列自适应调整测量噪声协方差的方法,使卡尔曼滤波随场景的运动动态地调整滤波性能,去除随机抖动噪声.实验表明,本文方法在 Intel Core i5 3.30 GHz CPU 下,对于 640×360 分辨率的彩色图像稳像帧率达到 40FPS,处理过程中无需利用下一帧信息,具有实时稳像的优点.

1 算法的理论分析

稳像的基本方法是根据相机模型计算出相机运动矢量,得到透视或仿射变换矩阵,用其稳定当前帧图像,主要分为运动估计、运动滤波与运动补偿三个步骤.运动估计基于图像序列信息计算相邻两帧图像的运动矢量和变换矩阵;运动滤波去除随机抖动噪声,保留有意运动;运动补偿利用变换矩阵补偿当前帧的偏移.

1.1 特征点筛选

拍摄过程中由于图像信息复杂、相机跟拍造成的场景切换等原因,相邻两帧的特征点在匹配过程中容易发生误匹配,同时如果出现纹理信息丰富的运动物体,特征点会落在运动物体上,如果物体运动速度快,其产生的局部运动分量在全局运动估计中将占较大比重,严重影响仿射变换矩阵精度,导致稳像失败.已有稳像方法检测出运动物体并剔除运动物体区域,但该方法适用于背景变化缓慢的情况,无法用在相机跟拍下背景变化较快的场合.

本文采用定向加速分割特征与旋转二元鲁棒特征相结合的特征检测算子(Oriented Features From Accelerated Segment Test and Rotated Binary Robust Independent Elementary Features, ORB)^[13]提取相邻帧图像特征点.ORB 具有旋转、尺度、仿射不变性,同时抗噪声能力强,运算速度比 SIFT 快两个数量级,但特征点描述的细致程度不高,为了提高匹配的精度,本文提出最小生成树约束下的特征筛选算法.

1.1.1.1 图像特征点的最小生成树

首先提取当前帧和前一帧图像的特征点,根据ORB特征之间的汉明距离(Hamming distance),即对两个特征向量异或运算,并统计结果为1的个数即为汉明距离,采用k近邻算法(k-Nearest Neighbor, kNN)^[7]进行特征匹配,得到两帧图像匹配后的特征点分布,统计匹配的特征点的位置坐标,得到特征点分布图,然后分别计算两幅特征图的最小生成树,特征点间的距离表示权重,最小生成树的基本计算步骤为:

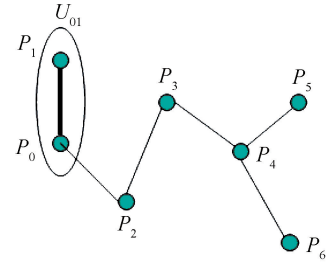


图1 最小生成树
Fig.1 Minimum spanning tree

1) 如图1所示,任选一个特征点 P_0 作为起点,考察与其直接相连且未访问过的特征点,找到其中与 P_0 距离最近的点 P_1 ,将其与 P_0 连接,同时标记 P_1 已访问过, P_0P_1 作为一个集合 U_{01} .

2) 更新其他特征点距离集合 U_{01} 的距离;

3) 继续寻找距离集合 U_{01} 最近的点 P_2 ,将其与集合 U_{01} 中相距最近的点相连,并加入到 U_{01} 得到 U_{012} ,以此类推,最后将所有特征点连接得到最小生成树.

用最小生成树表示图像特征的基本骨架,表征图像特征点分布情况,最小生成树具有旋转不变性,当图像旋转时,最小生成树随着旋转,两棵最小生成树仍保持较高的相似度,而前景运动和误匹配将会造成较大的相似度差异,当出现误匹配点如图2(a)或特征点取在运动前景上如图2(d),从图2(a)(d)可以看出两帧之间特征点匹配准确度低,存在误匹配以及将前景特征点保留的情况,图2(b)(c)与图2(e)(f)可以看出相邻两帧图像生成的最小生成树特征骨架差别较大.由最小生成树的相似度可以反应图像特征的匹配情况,当差异较大时,说明所匹配的特征点含有劣质特征点,后续运动估计的精度会受此影响,需要去除这些劣质特征.

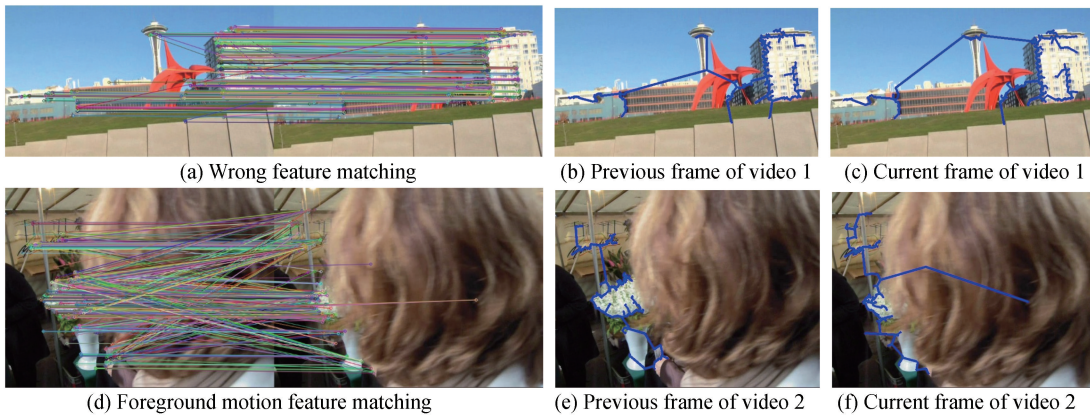


图2 相邻帧图像特征的最小权重生成树
Fig.2 Minimum spanning tree of image features of adjacent frames

1.1.1.2 迭代筛选法

为了去除劣质特征点,提出劣质特征的衡量指标,对于相机平移抖动的情况下,正确匹配并且位于图像背景上的特征点在前后两帧上的移动量稳定在某个值附近,一旦发生特征点误匹配时,特征点相对上一帧的位移较大,与正确匹配的位移量差距较大;当发生局部运动时,运动物体上的特征点随前景运动也会产生较大的位移量,基于这两个先验知识,可以将位移量大的特征点归为劣质特征点.

对于相机旋转的情况,两幅最小生成树相似,因为旋转不会造成很大的场景差异,本文不对它们进行筛选,只在最小生成树差异较大时进行特征筛选,而平移抖动和误匹配是造成最小生成树差异的主要原因.

为提高劣质特征点的筛选精度,提出一种迭代算法,算法的迭代直到最小生成树相似度大于阈值时停止.首先计算每个特征点 (x_i, y_i) 前后两帧的位移量 d_i 和所有特征点的平均位移量 \bar{d}

$$d_i = \sqrt{(x'_i - x_i)^2 + (y'_i - y_i)^2} \tag{1}$$

$$\bar{d} = \frac{1}{n} \sum_i^n d_i \tag{2}$$

统计所有位移量 d_i 大于平均位移量 \bar{d} 的特征点数量 m 占总特征点 n 的比例 E_1

$$E_1 = m/n \times 100\% \quad (3)$$

分别遍历相邻两帧图像的最小生成树,记录每条边的边长及对应的两个节点,得到点边对应表.比较两张点边对应表,计算两棵树对应边长 ξ_{ci}, ξ_{pi} 之差

$$E_2 = \sum_{i=1}^{n-1} |\xi_{ci} - \xi_{pi}| / (n-1) \quad (4)$$

根据求得的 E_1, E_2 进行置信度加权,得到两棵最小生成树的整体相似度 S

$$S = 1/(\omega_1 E_1 + \omega_2 E_2 + \epsilon) \quad (5)$$

式中 ω_1, ω_2 为置信度, ϵ 是一个很小的数,防止分母为 0.

对每个特征点作标记, $d_i > \bar{d}$ 则标记为 0,代表需要剔除的劣质特征点,反之标记为 1,代表保留的优质特征点

$$\text{flag}[i] = 0 \quad (d_i > \bar{d}), \quad \text{flag}[i] = 1 \quad (\text{else}) \quad (6)$$

下一轮迭代只对所有标记为 1 的优质特征点计算最小生成树.以此类推,直至相邻两帧图像最小生成树的相似度 S 大于阈值 S_t .

经过迭代算法后相邻帧的生成树如图 3 所示.图 3(a)对比图 2(a)可以看出,特征点匹配的准确度大大提升,对于误匹配所造成的最小生成树差异有效减小,如图 3(b)(c)所示;图 3(d)对比图 2(d)可以看出,该迭代方法去除了运动前景上的特征点,保留匹配的背景特征点,由于前景导致的最小生成树差异有效减小,如图 3(e)(f)所示.

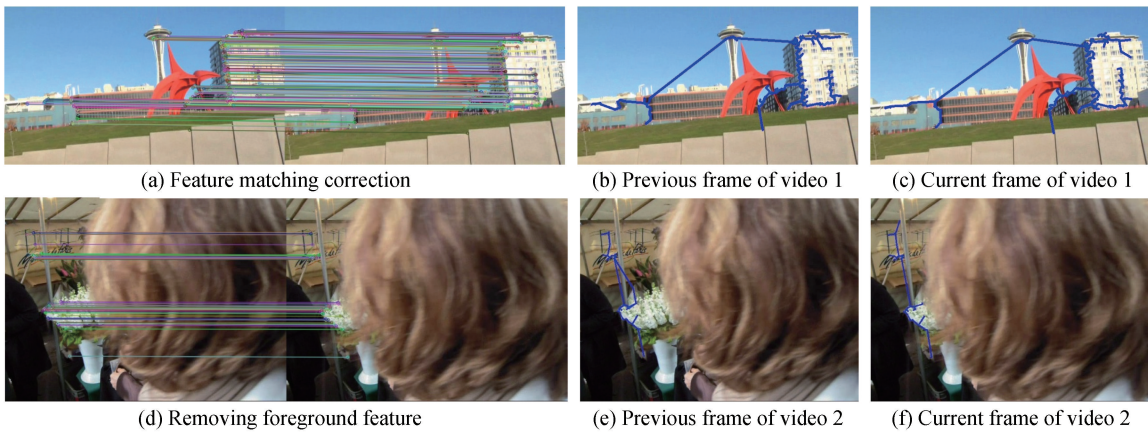


图 3 迭代筛选后相邻帧图像特征的最小生成树

Fig.3 Minimum spanning tree of adjacent image features after iterative filtering

1.2 自适应仿射变换矩阵修正

运动前景的存在不仅会导致全局运动估计不准确,在前景很大的时候,直接遮挡了大部分背景特征点,甚至导致仿射变换矩阵计算错误,如图 4(a)(b)表示修正仿射变换矩阵之前,稳像过程中的相邻两帧,图像序列中出现的大运动前景导致稳像失败.本文基于运动的时间局部性,提出一种修正仿射变换矩阵的方法.时间局部性是指前一帧的抖动信息具有一定的持续时间,可以用前一帧计算得到的运动分量来修正当前帧的运动分量,即用前一帧的仿射变换矩阵修正当前帧.



图 4 大运动前景下仿射变换矩阵修正前后的相邻两帧的对比

Fig.4 Comparison of two adjacent frames with large foreground motion before and after correcting transformation matrix

仿射变换矩阵表示为

$$\mathbf{T} = \begin{pmatrix} \cos \theta & -\sin \theta & t_x \\ \sin \theta & \cos \theta & t_y \end{pmatrix} \quad (7)$$

式中 t_x, t_y, θ 分别表示当前帧图像相对前一帧的水平、竖直移动量和旋转角度,变换矩阵的初步求解可以采用随机采样一致性算法(Random sample consensus, RANSAC)^[14].当出现大运动前景时,由于使用了基于最小生成树的特征点筛选算法,大部分前景特征点都可以有效的剔除,同时由于前景的遮挡导致背景特征点稀少,所以整幅图像特征点比无运动前景时少得多.而当特征点很少时,通过 RANSAC 求得的仿射变换矩阵不准确,本文基于运动的时间局部性,如果计算出来的 t_x, t_y 和 θ 相对上次计算变化过大,则对它们进行自适应修正,通过与前一个变换矩阵参数进行加权,得到新的变换矩阵;相反,如果变化不大,则不需要进行修正.

$$\theta_2 = \begin{cases} \lambda \theta_1 + (1 - \lambda) \theta_2, & \text{if } |\theta_1 - \theta_2| > \kappa_1 \frac{180}{\pi} \\ \theta_2, & \text{otherwise} \end{cases} \quad (8)$$

$$t_{x2} = \begin{cases} \eta t_{x1} + (1 - \eta) t_{x2}, & \text{if } |t_{x2} - t_{x1}| > \kappa_2 W \\ t_{x2}, & \text{otherwise} \end{cases} \quad (9)$$

$$t_{y2} = \begin{cases} \gamma t_{y1} + (1 - \gamma) t_{y2}, & \text{if } |t_{y2} - t_{y1}| > \kappa_3 H \\ t_{y2}, & \text{otherwise} \end{cases} \quad (10)$$

式中 t_{x2}, t_{y2}, θ_2 和 t_{x1}, t_{y1}, θ_1 分别表示当前帧和前一帧图像相对它们各自上一帧的水平偏移量、竖直偏移量和旋转角度, λ, η, γ 为当前帧运动量和前一帧运动量之间的加权因子,正比于运动量的变化率 $|\chi_2 - \chi_1|/\psi$, 3 个方向的运动量变化率分别等于水平偏移量、竖直偏移量和旋转角度相对图像长、宽和圆周率 π 的比例,其中 $\chi \in (t_x, t_y, \theta), \psi \in (W, H, 180/\pi)$, W 和 H 为图像的宽与高, $\kappa_1, \kappa_2, \kappa_3$ 为设定阈值因子,用于衡量运动参数变化量,可以根据需要进行调节.如图 4(c)(d)所示,本文的修正方法有效地防止了大前景运动导致的图像突然晃动,相比未修正时有了很大的改善.

1.3 改进的卡尔曼滤波器

1.3.1 基本原理

卡尔曼滤波是常用的运动滤波方法,它是一个线性离散控制过程,采用前一时刻对当前时刻的预测值和当前时刻实际的测量值来计算当前状态变量的估计值,卡尔曼滤波的状态转移方程和量测转移方程为

$$\begin{cases} x(k) = \Phi(k)x(k-1) + \Gamma(k)u(k) + q(k) \\ y(k) = H(k)x(k) + r(k) \end{cases} \quad (11)$$

式中 Φ 为状态转移矩阵, Γ 为系统的控制输入矩阵, H 为状态变换矩阵, x, y 分别为系统在 k 时刻的状态值、测量值, u 是系统的控制输入, q, r 分别为系统在 k 时刻的过程噪声和测量噪声,并假设为高斯噪声,协方差分别为 $Q(k), R(k)$.

卡尔曼滤波本质上是利用系统的预测值和实际的测量值得到下一状态的估计值,可分为预测阶段和校正阶段.

预测阶段:

$$x(k|k-1) = \Phi(k)x(k-1|k-1) + \Gamma(k)u(k) \quad (12)$$

$$P(k|k-1) = \Phi(k)P(k-1|k-1)\Phi(k)^T + Q(k) \quad (13)$$

校正阶段:

$$x(k|k) = x(k|k-1) + K(k)[y(k) - H(k)x(k|k-1)] \quad (14)$$

$$P(k|k) = [1 - K(k)H(k)]P(k|k-1) \quad (15)$$

$$K(k) = P(k|k-1)H(k)^T [H(k)P(k|k-1)H(k)^T + R(k)]^{-1} \quad (16)$$

$x(k|k-1)$ 是 $k-1$ 时刻对 k 时刻系统状态的预测值, $P(k|k-1)$ 是 $k-1$ 时刻对 k 时刻状态协方差矩阵的预测值, $x(k|k)$ 是 k 时刻系统状态的校正估计值, $P(k|k)$ 为 k 时刻状态协方差矩阵的校正估计值, $K(k)$ 为卡尔曼增益.根据这两组方程,利用 $x(k|k)$ 和 $P(k|k)$ 可以求得 $k+1$ 时刻 $x(k+1|k+1)$ 和 $P(k+1|k+1)$,以此类推,迭代下去.

1.3.2 参考卡尔曼滤波器

经典的卡尔曼滤波器,在过程噪声协方差 Q 值较小时无法有效的跟随相机突然的有意运动.由于卡尔曼增益 K 越大,滤波器的跟随效果越好,为了有效保留相机有意运动分量,本文引入两个卡尔曼滤波器,分别为主卡尔曼滤波器和参考卡尔曼滤波器.参考滤波器的 Q 值设成较大值,使卡尔曼增益 K 增大,有效跟踪相机的有意运动;主卡尔曼滤波器 Q 值较小,具有较小的卡尔曼增益 K ,平滑性能较高,主要去除抖动噪声.

主卡尔曼计算得到第 k 帧相对第一帧的累积位移矢量 $x_k(t_x, t_y, \theta)$ 为

$$x_k(t_x, t_y, \theta) = x(k | k) = x(k | k-1) + K(k)[y(k) - H(k)x(k | k-1)] \quad (17)$$

同理,参考卡尔曼滤波器计算的累积位移矢量为 $x'_k(t_x, t_y, \theta)$ 为

$$x'_k(t_x, t_y, \theta) = x'(k | k) = x'(k | k-1) + K'(k)[y(k) - H(k)x'(k | k-1)] \quad (18)$$

将 1.2 节中测量得到的第 k 帧相对 $k-1$ 帧的位移量测量值 (t_x, t_y, θ) 和卡尔曼滤波估计值 $x_k(t_x, t_y, \theta)$ 带入式(19),得到平滑后的相邻两帧之间的运动分量 (t'_x, t'_y, θ')

$$(t'_x, t'_y, \theta') = (t_x, t_y, \theta) + [x_k(t_x, t_y, \theta) - x'_k(t_x, t_y, \theta)] \quad (19)$$

式中 t'_x, t'_y, θ' 即平滑后的相邻两帧之间的仿射变换矩阵中的 3 个参数,代表平移和旋转,与 1.2 节中求得的 t_x, t_y, θ 区别是本节对 1.2 节得到的参数进行卡尔曼滤波,滤除了高频抖动噪声.

1.3.3 运动矢量队列

传统的卡尔曼滤波无法区别稳像过程中的相机运动的类型,为了有效分辨相机有意运动分量和随机抖动噪声,本文提出基于运动矢量队列的量测噪声协方差 R 的自适应调节方法.当连续几帧图像的特征点位移量累加量较大时,认为相机在该方向上存在有意运动,此时应该增大卡尔曼滤波的跟随性能,保留这些有意运动,本文通过对测量噪声协方差 R 的自适应修正,动态调整卡尔曼滤波器的跟随性能.

首先引入一个定长队列,长度为 L ,将 1.3.2 节每次求得的相邻帧位移 (t'_x, t'_y, θ') 存入队列;只要队列未滿,就不断将相邻两帧的位移量存入队列;一旦队列已滿,计算队列中所有位移向量总和得到近期帧的位移总量 $S(t_x, t_y, \theta)$,并根据位移总量修正测量噪声协方差 $R(t_x, t_y, \theta)$.当下一个相邻帧位移量到来时,将队列首元素删除,新元素加入至队列尾部,并重新计算 $S(t_x, t_y, \theta)$ 与 $R(t_x, t_y, \theta)$.

$$S(t_x, t_y, \theta) = [S_x, S_y, S_\theta] = \sum_{i=1}^L (t'_x, t'_y, \theta') \quad (20)$$

$$R(t_x, t_y, \theta) = [R_x, R_y, R_\theta] = \begin{bmatrix} \tau_x & \tau_y & \tau_\theta \\ S_x^2 & S_y^2 & S_\theta^2 \end{bmatrix} \quad (21)$$

式中 $\phi_x, \phi_y, \phi_\theta$ 分别为 3 个方向上的系数,当存在有意运动时,累计位移 S 较大,相应的测量噪声协方差 R 较小,滤波器的卡尔曼增益 k 增大,滤波器跟随性能增强,有效保留有意运动,相反,如果累计位移较小,则不存在相机有意运动,滤波器跟随性能将减弱,平滑性能增强,有效去除随机抖动.

2 结果与分析

2.1 评价指标

本文用帧间保真度(Interframe transformation fidelity, ITF)来评价稳像的效果,帧间保真度表示相邻帧的帧间峰值信噪比(Peak signal-to-noise ratio, PSNR)的平均值

$$\text{ITF} = \frac{1}{n-1} \sum_{i=1}^{n-1} \text{PSNR}(i) \quad (22)$$

$$\text{PSNR}(i) = 10 \log_{10} \frac{I_{\max}^2}{\text{MSE}(i)} \quad (23)$$

$$\text{MSE}(n) = \frac{1}{WH} \sum_{y=1}^H \sum_{x=1}^W [I_n(x, y) - I_{n-1}(x, y)]^2 \quad (24)$$

式中 I_{\max} 为图像中最大像素值, MSE 为帧间均方差, W, H 为图像长宽, I_n, I_{n+1} 表示相邻两帧图像.

2.2 实验结果与分析

选取的测试视频分为包含较大的前景运动、相机的跟拍运动、普通抖动视频 3 类,并与其他 3 种稳像算法进行对比,分别是 Liu 的子空间稳像算法^[10]、Grundmann 的基于 L1 范数优化的相机轨迹算法^[11]以及

Jing 的实时稳像算法^[12].实验主机的 CPU 为 Intel Core i5-4590 ,主频 3.30 GHz,8G 内存.

如图 4 为本文和另外三种方法对视频 girl 的稳像结果,分别取稳像结果的其中 3 帧对比,视频 girl 包含大前景运动,它的局部运动会影响全局运动估计.Liu 的子空间算法由于无法去除前景上的特征点,大量特征点取在运动前景上,由于单个子空间无法区别背景特征点与前景特征点的运动路径,使得全局运动估计错误,导致前景上的严重畸变,如图 5(a)所示.Grundmann 的方法将相机运动路径分为静止、线性、抛物线三种路径,用 L1 范数优化路径算法分别进行路径优化,优化过程加入 3 种约束:矩形包含约束、曲线贴近约束、显著特征点约束,并将提取的特征点离散化,分别落在不同的 50×50 像素的图像小块中,对每个小块采用随机采样一致性算法(RANSAC)进行运动估计,最后只保留符合估计模型的图像小块,但当前景较大,图像内主要存在全局和局部两种运动,由于 RANSAC 无法完全剔除前景特征点^[15],运动估计可能偏向局部运动,影响稳像效果,如图 5(b)中的背景倾斜.Jing 的实时算法为提高计算速度,简化了运动模型并忽略掉相邻几帧图像的相机运动,忽略了相机运动引入的累计误差,无法预测全局运动分量的突然变化,当视频中出现突然的移动或旋转时,将会导致明显的滞后现象,如图 4(c)所示.本文的结果相比其余 3 种算法有所改进,如图 5(d)所示,由于采用基于最小生成树的特征点迭代筛选算法,有效去除前景上的特征点,不会出现畸变与严重滞后现象.前景特征点的去除效果见图 6,其中图 6(a)为原始特征点分布,图 6(b)为经本文筛选算法之后的特征点分布,可以看出本文的基于最小生成树的迭代筛选算法有效去除了前景上的特征点,在运动估计时剔除了局部运动分量,消除局部运动对全局运动估计的影响.



图 5 包含大运动前景下 4 种稳像算法对比
Fig.5 Four different methods on a video containing foreground motion

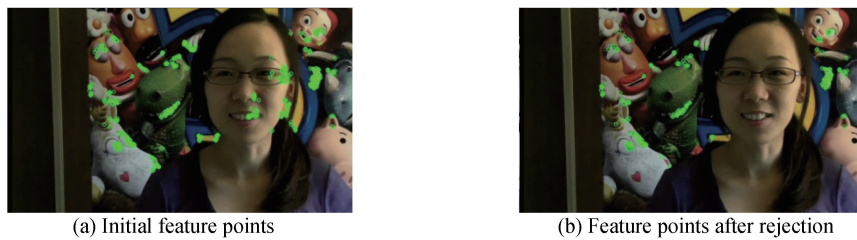


图 6 特征点迭代筛选前后对比
Fig.6 Feature points before and after iterative filtering

图 7 的测试视频 driving 包含相机的跟拍运动,本文的方法可以有效的区分相机跟拍运动与随机抖动噪声,其他 3 种方法在跟拍运动下都存在一定畸变.Liu 的方法由于相机跟拍和场景变化,导致特征点逐渐离开相机视场,没有足够长的特征跟踪轨迹,导致平滑失败.Grundmann 的方法由于场景移动时产生的



图 7 相机跟拍下的 4 种稳像方法对比
Fig.7 Comparison of 4 different stabilization methods

运动模糊, 稳像出现严重畸变. Jing 的方法无法预测下一帧的运动且无法适用于大视差场景, 对相机运动和大视差角都会产生畸变和晃动. 本文改进了卡尔曼滤波器, 用运动矢量队列动态修正滤波器的平滑性能, 在保留相机有意运动前提下, 采用双卡尔曼滤波器有效的去除随机抖动噪声, 稳像过程中不会产生严重的畸变, 有效保留了相机跟拍运动.

实验中, 4 种方法对 5 组测试视频分别进行稳像测试, 其中视频 girl 和 jump 包含前景运动, driving 和 rotation 包含相机的跟拍运动, walking 为普通抖动视频, 评价指标采用帧间保真度(ITF)衡量视频稳定性. 从表 1 可以看出本文的稳像方法相比其他方法在 4 个测试视频中具有更高的帧间保真度, 值得注意的是 Grundmann 的方法在测试视频 Driving 中虽然有最高的 ITF, 但是其畸变较为严重, 如图 7(b)所示. 表 2 为四种方法对 640×360 分辨率的测试视频的稳像帧率, 可以看出本文方法明显快于 Liu 和 Grundmann 的方法, 同时稳像速率与 Jing 的方法相差不大, 但从表 1 中可看出稳像结果优于 Jing 的方法; 本文方法在稳像过程中不需要知道下一帧信息, 仅利用当前帧和前几帧信息, 满足实时稳像要求.

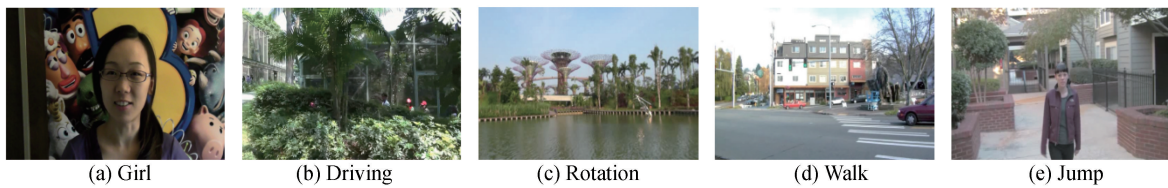


图 8 5 组测试视频

Fig.8 Five different test videos

表 1 4 种方法对 5 个测试视频稳像的帧间保真度

Table 1 ITF of 4 different methods on five test videos

ITF	Girl/dB	Driving/dB	Rotation/dB	Walk /dB	Jump/dB
Origin	28.641	16.683	23.102	20.823	21.264
Liu	34.770	17.697	22.536	22.957	22.754
Grundmann	34.378	23.671	20.998	28.446	23.356
Jing	30.498	17.872	23.430	25.199	23.372
Proposed method	34.954	17.890	25.372	28.872	24.911

表 2 4 种方法对 640×360 分辨率视频的稳像帧率Table 2 Stabilization FPS of four different methods on 640×360 resolution video

Method	FPS
Liu	16
Grundmann	20
Jing	43
Proposedmethod	40

对比表 1 和表 2 中的数据, 本文方法具有更优的稳像性能和较高的运算速度. Liu 的方法在相机跟拍时会导致跟踪的特征点离开视场, 特征跟踪路径太短影响稳像效果, 从表 1 中跟拍视频 Driving 和 Rotation 的实验结果可以看出本文比 Liu 更优, 当出现运动前景, Liu 方法单个子空间无法区分背景和前景特征, 从而影响全局运动估计, 且大数量的特征点跟踪使得计算复杂度大大增加, 从表 2 看出其帧率较低, 从前景运动视频 Girl 和 Jump 可以看出本文比 Liu 更优. Grundmann 的 L1 相机路径优化算法将相机运动路径分为静止、线性和抛物线三种路径并加入矩形包含、曲线贴合、显著点约束, 但是对于快速的相机跟拍或者特征稀少的图像会使优化的相机路径不可靠, 如图 7(b)跟拍 Driving 的严重畸变, 跟拍视频 Rotation 的结果本文更优, 对于前景运动 Girl 和 Jump, 由于 Grundmann 特征可能包含前景特征, 而本文可剔除前景特征, 本文稍优. Jing 的方法为达到算法的实时性, 简化了运动模型, 忽略相机运动及其带来的累计误差, 本文在 5 个视频的结果都优于 Jing. 本文方法在运动估计之前, 采用最小生成树对提取的特征点进行筛选, 有效提高匹配精度, 降低特征点误匹配与运动前景特征点引入的局部运动分量对全局运动估计的影响, 同时减少了劣质特征点的数量, 降低了运动估计的运算时间, 因为运动估计中的 RANSAC 算法从全部特征点中筛选出一个具有统一运动模型的最大特征点子集, 减少特征点数量可提高收敛速度, 所以本文方法比 Liu 与 Grundmann 的

方法具有更高帧率,与 Jing 的实时稳像算法速度接近.另一方面,本文采用运动矢量队列配合改进的卡尔曼滤波器对相机跟拍和随机抖动进行区分,动态调节卡尔曼增益,进而调节平滑与跟随性能,减小卡尔曼滤波对相机跟拍运动的平滑效果,保留原视频真实的跟拍运动.

2.3 不同条件下的帧间保真度

由于帧间保真度受前景运动速度、相机跟拍速度及抖动频率的影响,本节验证本文方法在这几种不同条件下,尤其在高速前景与高速跟拍下相对其他 3 种方法的优势,进一步对比本文与其他 3 种方法在不同的前景运动速度、相机跟拍速度及抖动频率下的帧间保真度.图 9 为 3 组测试视频,图 9(a)(b)为低速和高速前景运动视频,图 9(c)(d)为低速和高速相机跟拍视频,图 9(e)(f)为低频和高频抖动视频.表 3 为 4 种方法在 3 组不同视频下帧间保真度的结果.

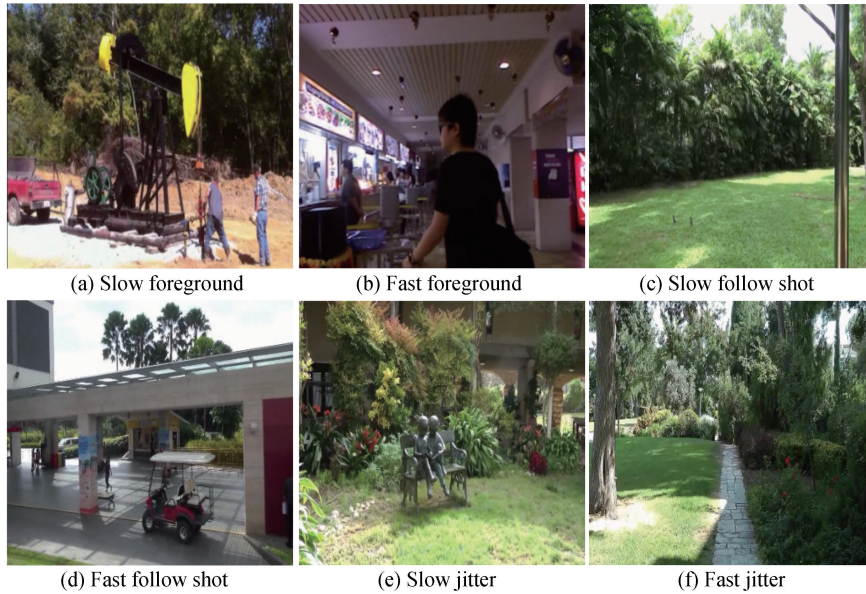


图 9 6 组测试视频

Fig.9 Six different test videos

表 3 4 种方法在不同条件下的帧间保真度

Table 3 ITF of 4 different methods under different conditions

ITF	Slow		Fast		Slow		Fast	
	foreground/dB	foreground/dB	followshot/dB	followshot/dB	jitter/dB	jitter/dB	jitter/dB	jitter/dB
Origin	22.359	20.250	21.200	23.722	24.085	14.319		
Liu	28.488	20.326	22.577	25.093	28.070	22.372		
Grundmann	30.276	25.121	22.460	25.311	28.699	23.583		
Jing	26.284	21.776	22.388	24.649	25.228	18.600		
Our method	30.927	27.718	22.684	28.677	32.073	26.319		

对比表 3 的数据,在包含高速或低速前景运动的情况下,本文方法相比 Liu 和 Jing 的方法都有较大优势,低速前景下比 Grundamn 的方法稍优,在高速前景下本文方法优势明显,因为高速前景引入的局部运动分量较大,Liu 的方法单个子空间无法区分背景和前景特征^[10],Grundmann 的方法将特征点均匀分散在不同的图像小块,但无法剔除前景特征,Jing 的方法为达到实时性了简化运动模型^[12],本文方法针对前景问题采用最小生成树的特征迭代筛选算法,有效地去除前景特征点,减小前景的局部运动对全局运动估计的影响.在低速跟拍条件下,本文的方法比其他 3 种方法稍优,在高速跟拍情况下,本文方法优势明显,Liu 方法需要跟踪特征点运动路径,而特征点会随相机跟拍离开视场导致特征跟踪路径过短,Grundmann 的方法优化的相机路径会随跟拍变得不可靠^[11],Jing 方法无法预测相机运动趋势^[12],本文方法采用相机运动矢量队列动态调整卡尔曼滤波器的平滑性能,自适应处理相机跟拍分量与随机抖动分量.在普通的抖动视频情况下,本文方法更优,因为本文基于最小生成树的特征迭代筛选算法提高了特征匹配精度,避免了特征误匹配对全局运动估计的影响.

3 结论

本文提出一种基于最小生成树和运动矢量队列的实时稳像算法,根据相邻帧特征点分布图的最小生成树相似度迭代筛选出背景特征点,有效去除前景特征点,解决了稳像过程中前景运动导致的全局运动估计不准确问题,利用自适应的仿射变换矩阵修正方法解决特征点稀少引起的稳像颠簸,利用运动矢量队列自适应调整卡尔曼滤波器的平滑与跟随能力,有效区分相机拍摄过程的有意运动与随机抖动,实验结果表明,本文方法有效地稳定包含运动前景和相机跟拍的视频,在 3.3GHzCPU 上对 640×360 分辨率图像的稳像帧率达 40FPS,同时保持较高的帧间保真度。

参考文献

- [1] CUI Shao-hui, XIE Zheng, WANG Gang, *et al.* Feature matching electronic image stabilization based on binary robust invariant scalable keypoints[J]. *Editorial Office of Optics and Precision Engineering*, 2015, **23**(9): 2715-2723.
崔少辉,谢征,王刚,等. 二进制鲁棒不变尺度特征匹配电子稳像[J]. *光学精密工程*, 2015, **23**(9): 2715-2723.
- [2] CHEN B H, KOPYLOV A, HUANG S C, *et al.* Improved global motion estimation via motion vector clustering for video stabilization[J]. *Engineering Applications of Artificial Intelligence*, 2016, **54**(C): 39-48.
- [3] CHANG H C, LAI S H, LU K R. A robust real-time video stabilization algorithm [J]. *Journal of Visual Communication and Image Representation*, 2006, **17**(3): 659-673.
- [4] WANG Peng, ZHAO Yue-jin, KONG Ling-qin *et al.* Improved fast block-matching algorithm for sensor electronic image stabilization[J]. *Acta Photonica Sinica*, 2012, **41**(9): 1065-1070.
王鹏,赵跃进,孔令琴,等. 改进块匹配宏块的快速传感器电子稳像方法[J]. *光子学报*, 2012, **41**(9): 1065-1070.
- [5] PUGLISI G, BATTIATO S. A robust image alignment algorithm for video stabilization purposes[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, **21**(10): 1390-1400.
- [6] KWON O, SHIN J, PAIK J. Edge based adaptive Kalman filtering for real-time video stabilization[C]. *IEEE Digest of Technical Papers International Conference on Consumer Electronics*, 2006: 75-76.
- [7] BATTIATO S, GALLO G, PUGLISI G, *et al.* SIFT features tracking for video stabilization[C]. *IEEE International Conference on Image Analysis and Processing*, 2007: 825-830.
- [8] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110.
- [9] ZHU Juan-juan, FAN Jing, GUO Bao-long. Adaptive electronic image stabilization algorithm resistant to foreground moving object[J]. *Acta Photonica Sinica*, 2015, **44**(6): 0610002.
朱娟娟,范静,郭宝龙. 抗前景干扰的自适应电子稳像算法[J]. *光子学报*, 2015, **44**(6): 610002.
- [10] LIU F, GLEICHER M, WANG J, *et al.* Subspace video stabilization[J]. *Acm Transactions on Graphics*, 2011, **30**(1): 4.
- [11] GRUNDMANN M, KWATRA V, ESSA I. Auto-directed video stabilization with robust L1 optimal camera paths[C]. *IEEE Computer Vision and Pattern Recognition*, 2011: 225-232.
- [12] DONG J, LIU H. Video stabilization for strict real-time applications[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, **27**(4): 716-724.
- [13] RUBLEE E, RABAUD V, KONOLIGE K, *et al.* ORB: An efficient alternative to SIFT or SURF [C]. *IEEE International Conference on Computer Vision*, 2011: 2564-2571.
- [14] FISCHIER 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.
- [15] LIU S, YUAN L, TAN P, *et al.* Steady flow: spatially smooth optical flow for video stabilization[C]. *IEEE Computer Vision and Pattern Recognition*, 2014: 4209-4216.