

doi: 10.3788/gzxb20154406.0610002

# 抗前景干扰的自适应电子稳像算法

朱娟娟, 范静, 郭宝龙

(西安电子科技大学 空间科学与技术学院, 西安 710071)

**摘 要:**针对场景中存在前景目标运动的抖动视频,提出一种抗前景干扰的自适应电子稳像算法.算法以视觉对运动的感知为指导,采用基于块的三帧间差分,利用时空一致性快速剔除运动前景区域;改进传统 Harris 算子,用网格筛选和显著度排序,对背景区域进行全局显著特征点的提取和配准,保证全局配准准确度;用统计分布的距离准则去除误匹配点,无需特征点迭代运算,提高了全局运动估计的速度和准确度.在 Sage-Husa 自适应运动滤波方法的基础上,改进了修正过程噪音和观测噪音的统计特性,模拟摄像机低频匀速运动的视觉平滑效果,有效解决摄像机抖动中存在的扫描运动.在 Intel 酷睿 2 四核 2.33 GHz 的微机上用 VC++ 进行实验,结果表明,该算法对  $320 \times 240$  像素的视频序列能够达到 22 fps 的处理能力,可以实时稳定含较大或多运动前景目标的复杂抖动视频,输出视觉完整流畅的真实扫描场景.

**关键词:**视频信息处理;电子稳像;全局运动估计;自适应运动滤波;运动补偿;Harris 算子;峰值信噪比  
**中图分类号:**TP391.4      **文献标识码:**A      **文章编号:**1004-4213(2015)06-0610002-8

## Adaptive Electronic Image Stabilization Algorithm Resistant to Foreground Moving Object

ZHU Juan-juan, FAN Jing, GUO Bao-long

(School of Aerospace Science and Technology, Xidian University, Xi'an 710071, China)

**Abstract:** An adaptive electronic image stabilization algorithm was proposed to remove jitter from videos with foreground moving object. Visual attention on motion was used to make fast foreground segmentation. It made difference between three adjacent frames based on blocks, and the temporal-spatial consistency was validated to label foreground moving object. The grid-sampling and significant degree sorting were proposed to detect global distinct features in the background area of reference and match in current frame. The distance criterion based on statistical distribution was used to delete mismatched points, which improved the speed and accuracy of global motion estimation without iteration process. Then Sage-Husa adaptive filtering method was described and improved, which corrected the process and observation noise by estimating their statistical property in real time. The low-frequency visual smooth as uniform motion was simulated. Thus, it can remove camera jitter and retain camera scanning simultaneously. With VC++ implementation and a computer (Intel Core (TM) 2, 2.33 GHz), processing image of size  $320 \times 240$  reached 22 fps speed. The experimental results illustrate that the proposed algorithm can stabilize the inter-frame jitter with large or multi moving targets and track the real scene with smooth visibility.

**Key words:** Video signal processing; Electronic image stabilization; Global motion estimation; Adaptive motion filter; Motion compensation; Harris operator; Peak signal to noise ratio

**OCIS Codes:** 100.2960; 100.3008; 100.4999; 110.4155; 110.4153

**基金项目:**国家自然科学基金(No. 61201290)和高校基本科研业务费专项资金(No. JB141301)资助  
**第一作者:**朱娟娟(1980-),女,副教授,博士,主要研究方向为视觉信息处理. Email: zhujuo@126.com  
**收稿日期:**2014-12-15; **录用日期:**2015-03-17

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

## 0 引言

人类视觉系统具有视觉暂留特性,当摄像设备出现频率较高抖动时,视频呈现模糊等不稳定现象,降低主观视觉质量.由此,出现了电子稳像技术,它采用图像处理来估计视频帧间偏移量并进行补偿,可作为光学稳像的二级稳定.目前电子稳像技术已广泛应用于目标检测与跟踪、机器人、视频拼接、视觉校正稳像<sup>[1]</sup>等领域.

对电子稳像系统的研究主要集中在运动估计和运动补偿.运动估计的任务是确定帧间的全局运动偏移量,当前景中存在不一致局部运动物体时,基于块匹配的改进算法<sup>[2-3]</sup>利用统计特性决策最优模式,其算法实现简单但是准确度不高.特征匹配法<sup>[4]</sup>可以处理平移、旋转和变焦情况,传统特征点易集中于局部物体的边界上,仍需利用 RANSAC 算法<sup>[5]</sup>进行迭代,计算量较大从而会降低算法的处理速度.运动补偿对运动参量滤波提取抖动分量,其重点在于去除抖动的同时保留扫描运动.常用的补偿方法有运动衰减法<sup>[6]</sup>、均值滤波法<sup>[7]</sup>、Kalman 滤波法<sup>[8]</sup>等.其中,运动衰减法中的衰减系数是实验中凭经验设定的,无法适用于所有的视频序列;均值滤波采用求平均运算,会引入多余的低频噪声;Kalman 滤波则要求过程噪声和观测噪声先验已知,并服从零均值的高斯分布,这在实际系统中难以满足.

考虑稳像系统实际应用中不可避免出现前景运动的干扰<sup>[9]</sup>,必须提高全局运动估计的准确度;同时,模拟视觉平滑特性,只需摄像机保持低频匀速,即对全局运动矢量序列中的高频分量进行滤除,获得低频平滑运动模式.因此,本文提出了一种融合全局特征匹配和自适应滤波的电子稳像算法(Global Features Matching and Adaptive Filter, GFM-AF).在特征提取前增加预处理环节,结合连续三帧间差分图像,采用基于图像块的时域和空域相似性来快速标记运动子块;将传统 Harris 整幅提取改进为网格筛选和显著度排序后,在背景区域上检测全局显著特征点;然后结合特征点对距离分布特性,提出距离准则来进一步剔除误匹配点,从而获得准确的全局运动矢量.提出自适应滤波方法进行运动补偿,改进 Sage-Husa 的预测方程,通过新息序列来判断并修正预测误差的方差阵,有效防止滤波发散,得到兼顾跟踪的平滑效果.为保证系统的实时性,根据图像线性存储结构,采用快速线性坐标计算的补偿方法来实现实时稳像.

## 1 全局运动估计

视觉系统最易捕捉图像中的显著部分,视频整个画面的抖动与局部目标相比也更易引起人眼注意,从

而导致视觉模糊,难以观察目标.视频序列中的运动可以分为两类:目标运动和摄像机运动,前者为局部运动,后者为全局运动,视频序列的不稳定正是由摄像机高频抖动引起的,需要对全局运动进行平滑稳定处理,因此必须保证全局运动估计的准确度.由于常用的特征点检测<sup>[5]</sup>是针对整幅图像的,因而无法避免特征点选在了局部目标上,尤其当前景运动目标较大时,局部特征点数量较多,将严重干扰全局配准;同时,纹理重复区域的特征点容易出现错误匹配,这些均会导致全局运动估计的准确度下降.本文采用前景运动区域标记和排除,保证了提取的特征点位于背景区域,同时改进 Harris 算子<sup>[10]</sup>,使得特征点为视觉观察到的具有独特信息的显著特征点;结合特征匹配的距离验证,进一步提高全局运动估计的准确度.全局运动估计的具体流程如图 1.

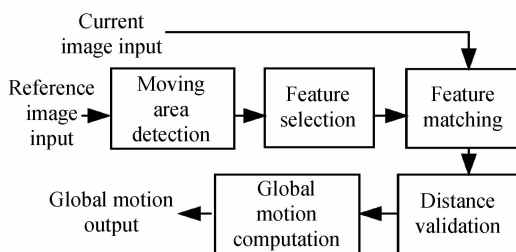


图 1 全局运动估计流程

Fig. 1 Flow chart of global motion estimation

### 1.1 前景运动区域检测

在动态场景中检测前景运动目标可采用帧差法<sup>[11]</sup>,将相邻两帧逐点进行灰度求差,二值化后提取前景目标.但是单个像素点作为差分对象,容易存在很多细小的残留噪声,运动目标区域会出现不同程度的碎化.另外,由于光照等影响,背景区域也会出现单个噪声点.

根据视觉特性,人眼无法观察到单个像素点的差异,而是以整个区域的持续时间变化即空域和时域特性来判断运动,故可利用基于块的差异来判断时空一致性.空域一致性是指若某子块为运动区域,则其周围邻域块必然存在运动块;时域一致性是指若当前帧内该块为运动块,则在下一帧该块位置及其邻域内也应包含运动,否则可能是孤立子块,应予以删除以避免误检.空域一致性解决了准确定位目标边缘的问题,时域一致性确定运动目标的位置,从而时空联合特性能够准确获得目标的相对位置和边界轮廓.

因此本文采用三帧间差分,以块作为差分单位,统计块内的平均差分,目标区域差分的值相对较大,从而可以迅速判断该块是否属于运动目标.分析单帧内运动区域内部的差分应具有空域一致性,即运动子块的 8 邻域块同时存在运动子块.同时分析连续两帧内运动区域的时域一致性,即运动子块在下一帧的类似

位置应出现运动子块,若不满足时空一致性,则判断该块应为背景中残差较大的孤立块,予以删除.其具体检测步骤为:

Step 1. 对相邻三帧  $f_{t-1}$ 、 $f_t$  和  $f_{t+1}$ , 分别计算它们之间的差分图像

$$D_t(x, y) = \text{abs}[f_t(x, y) - f_{t-1}(x, y)] \quad (1)$$

$$D_{t+1}(x, y) = \text{abs}[f_{t+1}(x, y) - f_t(x, y)] \quad (2)$$

Step 2. 实验设定差分图像分块尺寸为  $16 \times 16$ , 计算各块的平均绝对残差, 标记为  $B_t(m, n)$  和  $B_{t+1}(m, n)$

$$B_t(m, n) = \frac{1}{16 \times 16} \sum D_t(x, y) \quad (3)$$

$$B_{t+1}(m, n) = \frac{1}{16 \times 16} \sum D_{t+1}(x, y) \quad (4)$$

Step 3. 初步判断各块是否存在运动 (Moving Object, MO), 残差大于阈值的块为候选运动块, 其中  $Th_1$  和  $Th_2$  为采用最大类间方差法设定残差  $B$  的阈值, 判断公式为

$$\begin{cases} \mathbf{MO}_t(m, n) = \begin{cases} 1, & \text{if } B_t(m, n) > Th_1 \\ 0, & \text{else} \end{cases} \\ \mathbf{MO}_{t+1}(m, n) = \begin{cases} 1, & \text{if } B_{t+1}(m, n) > Th_2 \\ 0, & \text{else} \end{cases} \end{cases} \quad (5)$$

Step 4. 对候选运动块, 进行时空一致性检测, 即统计连续两帧各子块的 8 相邻子块, 如标记的运动子块个数大于 3, 说明该块周围同属于前景, 且时间上连续, 是真实的运动前景, 即

$$\mathbf{MO}_t(m, n) = \begin{cases} 1, & \text{if } \sum_{i=0}^1 \mathbf{MO}_{t-1}(m \pm i, n \pm i) \geq 3 \\ & \& \sum_{i=0}^1 \mathbf{MO}_{t+1}(m \pm i, n \pm i) \geq 3 \\ 0, & \text{else} \end{cases} \quad (6)$$

对两段标准视频序列 Coastguard 进行测试, 其前景检测结果如图 2, 图中标记方块为前景运动子块, 可以快速有效地标记局部运动区域.

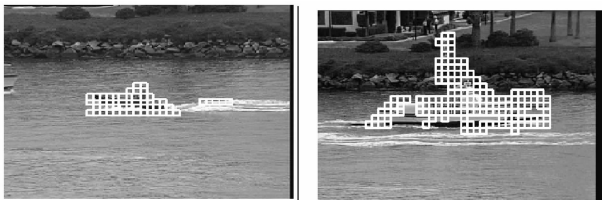


图 2 Coastguard 中前景运动子块标记

Fig. 2 Moving blocks in coastguard sequence

## 1.2 参考帧提取全局显著特征点

特征点检测采用 Harris 角点算子, 它具有良好的平移、旋转不变性, 同时对噪声及光照条件等有很好的鲁棒性. Harris 角点检测原理为

$$\mathbf{M} = \mathbf{w} \otimes \begin{bmatrix} \mathbf{I}_x^2 & \mathbf{I}_x \mathbf{I}_y \\ \mathbf{I}_x \mathbf{I}_y & \mathbf{I}_y^2 \end{bmatrix} \quad (7)$$

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (8)$$

式中  $\mathbf{w}$  为离散二维零均值高斯函数,  $\mathbf{I}_x$ 、 $\mathbf{I}_y$  分别表示图像在水平和垂直方向上的梯度,  $\lambda_1$ 、 $\lambda_2$  为自相关矩阵  $\mathbf{M}$  的特征值, 参量  $k=0.05$ ,  $R$  为响应值,  $R$  值较大的点为图像的角点.

传统 Harris 算子提取的特征点数量过大且密集排列, 其迭代运算复杂耗时. 减少特征点数量最直接的方法是提高 Harris 角点响应的阈值  $T$ , 但其保留的特征点依然密集, 且易落于前景物体上, 如图 3. 其中, 图 3(a) 为实际拍摄的 Ground 序列中一幅图片, 图 3(b) 为传统 Harris 检测结果, 其响应阈值  $T=5000$ , 提高阈值后检测结果如图 3(c). 显然, 很多特征点集中在前景运动的人身体上, 这些局部运动点必然会干扰全局参量的计算. 另一种减少特征点数量的方法是网格筛选, 即将图像划分为网格后, 每个网格内只选择一个特征点. 如图 3(d), 网格筛选后的特征点均匀分布在整幅图像中, 且网格筛选可以控制特征点的数量, 以保证运算速度. 对网格筛选后的结果进一步分析, 部分特征点依然选在了运动目标上; 另有大量特征点位于纹理重复的天空和地面区域, 其较低的显著度信息会造成误匹配的干扰.

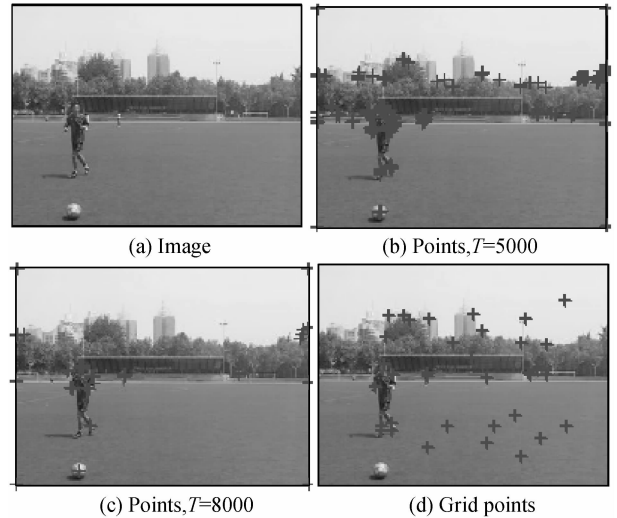


图 3 Harris 特征点检测结果比较

Fig. 3 Comparison of Harris feature points selection

因此, 将传统 Harris 提取的特征点作为初选特征点, 综合考虑其数量、显著度和分布因素, 对初选特征点进行筛选处理, 本文提出简单有效的改进方法. 其基本原理是: 1) 将参考帧分成互不重叠的网格, 找到各网格内最大 Harris 响应值, 并对其进行排序, 选择较高  $R$  值的点为显著特征点. 这样可以减少特征点个数, 保证特征点均匀分布且具有独特信息, 减少密集分布或纹理重复导致的误匹配. 2) 结合前景子块标记结果, 进一步判断该特征点位置, 若位于运动区域, 则予以删除, 从而保证特征点的位置尽可能不在前景运动物体中, 去除局部运动的干扰. 满足这些特性的特征点称之为全局显著特征点, 其具体筛选步骤为

Step 1. 参考帧  $f_i$  分块, 找到各子块内  $R$  的最大值, 作为该子块的特征响应值  $R_{\text{HMAX}}(m, n)$ ;

Step 2. 将  $R_{\text{HMAX}}(m, n)$  进行由高到低的排序, 取出前 20% 较大的响应值, 其对应的位置可以标记为图像中具有较多信息的显著特征点  $(x_i, y_i)$ ;

Step 3. 对  $(x_i, y_i)$  进行判定, 该特征点对应位置  $\text{MO}_i(m, n)$  若为 1, 则表明该特征点为前景局部特征点, 予以删除, 保留的即为全局显著特征点.

### 1.3 当前帧特征点匹配

视觉判断特征点对是否正确匹配, 一般是将特征点周围像素看成一个整体块, 判断其整体信息是否一致. 故采用模板匹配的方法, 在参考帧中以特征点为中心, 构建模板块, 在当前帧中利用最小误差和准则和全搜索策略进行块匹配, 则配准块的中心点为匹配特征点. 匹配过程中, 采用序贯相似性检测, 以随机不重复的顺序选取像素对, 计算部分误差和, 只要超过阈值, 则停止本次计算, 并进入下一个位置的测试, 直至找到最小误差位置点. 由于匹配过程中可能出现误匹配, 因此须对匹配特征点对的正确性进行验证. 本文采用距离准则对特征点对进行验证.

定义欧式距离<sup>[12]</sup>  $d_i = \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$ , 其中  $\Delta x_i, \Delta y_i$  分别为相邻两帧中的第  $i$  个特征点对在水平方向和垂直方向上的平移量. 结合实验统计,  $d$  近似服从正态分布, 即

$$\begin{cases} P(d_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(d_i - \mu)^2}{2\sigma^2}\right) \\ \mu = \frac{1}{n} \sum_{i=1}^n d_i \\ \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i - \mu)^2} \end{cases} \quad (9)$$

式中  $\mu, \sigma$  分别为  $d$  的均值和标准差,  $n$  为特征点数.

特征点对距离近似服从正态分布, 这组距离数据中含有误差, 对其进行计算处理得到均值  $\mu$  和标准差  $\sigma$ , 按一定概率确定一个区间, 认为凡超过这个区间的误差, 就不属于随机误差而是粗大误差, 对应的数据应予以剔除. 这种正态分布剔除粗大误差的原理是基于样本数据充分大或者正确数据较大量为前提的, 文中预处理检测的全局显著特征点的正确匹配度高满足这一前提. 根据拉依达准则 ( $3\sigma$  准则), 在  $[\mu - 3\sigma, \mu + 3\sigma]$  区间上的数据占总数据的 99.7%, 如图 4 所示. 故剔除误匹配的距离准则为: 当特征点对的距离误差  $|d_i - \mu| > 3\sigma$  时, 即数据落于  $[\mu - 3\sigma, \mu + 3\sigma]$  之外, 此特征为无效数据, 对应的误匹配点将其剔除, 由此得到正确匹配的  $N$  对特征点对. 预处理能有效保证特征点对匹配的正确度, 从而可以直接代入运动模型进行全局运动参量计算, 避免传统的特征点对迭代, 大大提高了运算速度.

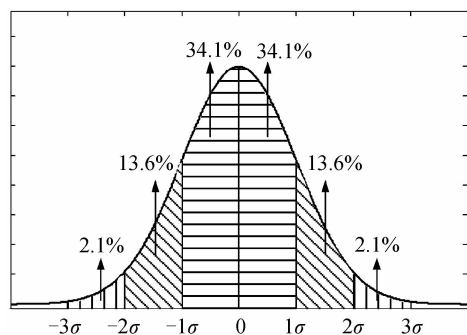


图 4 正态分布“ $3\sigma$  准则”  
Fig. 4  $3\sigma$  criterion

### 1.4 运动参量的确定

假设摄像机对着一个平坦的面进行拍摄, 并且摄像机的光轴与该平面垂直, 沿  $x$  和  $y$  轴的平移、俯仰和摇头均映射为像平面的平移, 沿  $z$  轴的平移为比例缩放, 绕  $z$  轴的转动导致图像的整体旋转, 建立二维变换模型

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u \\ v \end{bmatrix} \quad (10)$$

式中,  $(\hat{x}, \hat{y})$  和  $(x, y)$  分别为匹配点对的坐标,  $(u, v)$  为水平和垂直方向上的平移量,  $m_1, m_2, m_3, m_4$  表示旋转和缩放. 该参量仿射模型能够描述视频中绝大部分运动类型, 具有适中的计算复杂度和准确度<sup>[13]</sup>.

将特征点距离验证后得到的  $N$  对特征点代入式 (10) 中, 得到含有 6 个未知参量的  $2N$  个线性方程. 对方程进行展开、移项, 整理得到  $\mathbf{B} = \mathbf{A}\mathbf{M}$ , 即

$$\begin{bmatrix} \hat{x}_1 & \hat{y}_1 \\ \hat{x}_2 & \hat{y}_2 \\ \vdots & \vdots \\ \hat{x}_N & \hat{y}_N \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & 1 \end{bmatrix} \times \begin{bmatrix} m_1 & m_2 & m_3 \\ m_4 & u & v \end{bmatrix} \quad (11)$$

根据解超定线性方程组的方法求解  $\mathbf{M}$ ,  $\mathbf{M}$  的最小二乘解由  $\mathbf{M} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B}$  得到.

## 2 帧间运动矢量补偿

帧间摄像机的运动通常由正常扫描运动和随机抖动构成, 对视频序列的运动滤波是基于假设: 摄像机的扫描运动变化平稳、低速且方向性明确, 而抖动的变化随机性大且方向具有任意性, 即扫描分量是低频分量而抖动分量是高频分量. 因此, 通过运动矢量滤波来模拟摄像机低频匀速运动的视觉平滑效果, 而待补偿的抖动分量即可看成是原始运动矢量与平滑分量的差值.

### 2.1 Sage-Husa 滤波算法

采用经典 Kalman 滤波技术时, 由于无法准确获知过程噪声和观测噪声, 常常导致滤波准确度降低甚至发散. 因此, 利用 Sage-Husa 滤波算法<sup>[14]</sup>, 通过观测数据不断修正预测值, 同时将过程噪声和观测噪声的统

计特性代入标准卡尔曼滤波器中,进行实时估计和修正,从而达到降低模型误差、提高滤波准确度的目的。

考虑摄像机的实际扫描状态,帧间的旋转运动对扫描运动的影响较小,令状态矢量  $\mathbf{S}(k)=[\mathbf{u}(k), \mathbf{v}(k), d\mathbf{u}(k), d\mathbf{v}(k)]^T$ ,分别表示水平和垂直平移以及对应的瞬时速度; $\mathbf{Z}(k)$ 为观测矢量,  $\mathbf{Z}(k)=[\mathbf{u}(k), \mathbf{v}(k)]^T$ . 构造线性离散系统模型为

$$\begin{cases} \mathbf{S}(k) = \mathbf{F} \cdot \mathbf{S}(k-1) + \mathbf{w} \\ \mathbf{Z}(k) = \mathbf{H} \cdot \mathbf{S}(k) + \mathbf{v} \end{cases} \quad (12)$$

式中,  $\mathbf{F}$  是状态转移矩阵,  $\mathbf{H}$  是观测矩阵

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Sage-Husa 状态预测与更新过程为:

Step 1. 建立系统状态预测方程

$$\mathbf{S}(k|k-1) = \mathbf{F} \cdot \mathbf{S}(k-1|k-1) \quad (13)$$

Step 2. 对  $\mathbf{S}(k|k-1)$  的协方差矩阵  $\mathbf{P}(k|k-1)$  进行预测

$$\mathbf{P}(k|k-1) = \mathbf{F} \cdot \mathbf{P}(k-1|k-1) \cdot \mathbf{F}^T + \mathbf{Q} \quad (14)$$

Step 3. 建立状态更新方程,得到最优估算值,即滤波结果

$$\mathbf{S}(k|k) = \mathbf{S}(k|k-1) + \mathbf{K}_g(k) \cdot \boldsymbol{\varepsilon}(k) \quad (15)$$

式中,  $\mathbf{K}_g(k)$  为 Kalman 增益,  $\boldsymbol{\varepsilon}(k)$  为新息序列

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

$$\boldsymbol{\varepsilon}(k) = \mathbf{Z}(k) - \mathbf{H} \cdot \mathbf{S}(k|k-1) \quad (17)$$

Step 4. 更新  $k$  状态下  $\mathbf{S}(k|k)$  的协方差矩阵

$$\mathbf{P}(k|k) = (\mathbf{I} - \mathbf{K}_g(k) \cdot \mathbf{H}) \cdot \mathbf{P}(k|k-1) \quad (18)$$

当噪音  $\mathbf{Q}$  和  $\mathbf{R}$  均未知时, Sage-Husa 法不能把它们都估计出来;而且,随着  $\mathbf{Q}$  和  $\mathbf{R}$  失去半正定性和正定性,会出现滤波发散现象,因此,无法保证稳定性和收敛性。

## 2.2 改进的自适应平滑滤波方法

为了防止滤波发散现象的产生,提出了改进的自适应平滑滤波方法 (Adaptive Filter, AF): 通过对预测误差协方差阵  $\mathbf{P}(k|k-1)$  进行修正,增强对突变状态的跟踪能力. 主要方法是利用新息序列的性质来判断滤波器是否发散,即

$$\boldsymbol{\varepsilon}(k)^T \boldsymbol{\varepsilon}(k) \leq \gamma \cdot \text{Trace}[\mathbf{H} \mathbf{P}(k|k-1) \mathbf{H}^T + \mathbf{R}] \quad (19)$$

式中,  $\gamma$  为可调系数且  $\gamma > 1$ .

当式(19)成立时,滤波器处于正常工作状态,直接利用式(14)并获得最优估计值;当式(19)不成立时,表明实际误差将超过理论估计值的  $\gamma$  倍,滤波将发散,此时要通过加权系数  $\mathbf{C}(k)$  对协方差矩阵  $\mathbf{P}(k|k-1)$  进行修正,修正公式为

$$\mathbf{P}(k|k-1) = \mathbf{C}(k) \mathbf{F} \mathbf{P}(k-1) \mathbf{F}^T + \mathbf{Q} \quad (20)$$

$$\mathbf{C}(k) = \frac{\boldsymbol{\varepsilon}(k)^T \boldsymbol{\varepsilon}(k) - \text{Trace}[\mathbf{H} \mathbf{Q} \mathbf{H}^T + \mathbf{R}]}{\text{Trace}[\mathbf{H} \cdot \mathbf{F} \cdot \mathbf{P}(k) \cdot \mathbf{F}^T \cdot \mathbf{H}^T]} \quad (21)$$

通过  $\mathbf{P}(k|k-1)$  的修正,可以获得自适应的最佳平滑滤波结果。

## 2.3 快速运动补偿方法

运动补偿是将当前帧的每一像素点,利用补偿参量进行坐标运算,补偿参量为滤波前后平移之差  $(u, v)_{\text{jitter}} = (u, v) - (u, v)_{\text{filter}}$  作为抖动参量,结合已计算的全局运动仿射矩阵  $\mathbf{M}$  中的其余旋转和缩放参量,待补偿参量为  $\bar{\mathbf{M}} = [m_1, m_3; m_2, m_4; u_{\text{jitter}}, v_{\text{jitter}}]$ .

传统补偿时逐点运算过程费时,效率低,根据像素点间的相对位置具有旋转不变性的特点,在实际坐标计算时只对当前帧首行首列像素进行旋转计算;对其余行列的像素,进行坐标线性加减运算,获得新坐标. 这样避免了对整幅图像的每个像素点作矩阵乘法运算,有效地节省了运算时间. 实验时,采用线性坐标计算方法,平均每帧补偿耗时仅为 4.025 ms,相对于传统的逐点计算耗时 17.633 ms,节约了 76.2% 的时间。

## 3 实验结果及分析

实验时摄像机被安放在一个抖动平台上,该平台存在水平和垂直方向的随机抖动,旋转运动,同时在水平方向存在慢速扫描运动. 对实际拍摄的多组抖动视频进行实验,图像分辨率为  $320 \times 240$ . 在 Intel 酷睿 2 四核 Q8200, CPU 主频: 2.33 GHz 的微机利用 VC++ 编程,统计整个算法的平均执行时间为 45.2 ms,能够达到 22 fps 的实时处理能力。

### 3.1 前景运动区域检测结果分析

实际拍摄的 Ground 序列中存在较大运动前景目标,其检测结果如图 5. 图 5(a)和图 5(b)为参考帧和当前帧,计算三帧间差分,以块内的平均差分 MAD 进

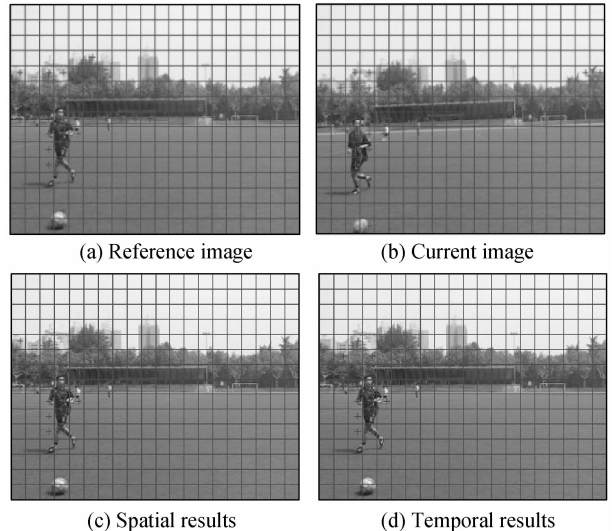


图 5 前景运动区域检测结果

Fig. 5 Moving area detection in Ground video

行阈值比较,迅速初步判断该块是否属于运动目标,图中标记为“+”的那些子块为候选运动块。

先对参考帧内子块的 8 邻域块进行空域相似性分析,若周围块存在不少于 3 个运动子块,说明该块和邻域块都属于前景运动区域,候选前景子块如图 5(c)。再对运动子块进行时域相似性分析,若在连续的当前帧内,类似的 8 邻域位置仍然存在运动子块,则判断为真实的运动;否则,应视为偶尔出现的误检,需要删除,最终保留的子块为检测的运动前景区域,如图 5(d),运动子块标记的平均耗时为 10.7 mspf。

### 3.2 全局显著特征点检测和配准结果

#### 3.2.1 较大单目标简单场景测试

对存在较大单目标的 Ground 视频进行全局运动估计,该视频存在快速运动的前景,背景中有地面和天空等纹理区域。图 6 是全局显著特征点提取、配准的结果,其中图 6(a)为参考帧网格筛选的所有特征点。可以看出,特征点均匀分布在整幅图像中,但是有一部分特征点选在了前景运动目标上,另有一部分特征点位于纹理重复的天空和地面区域,这两类特征点会造成误匹配的干扰。

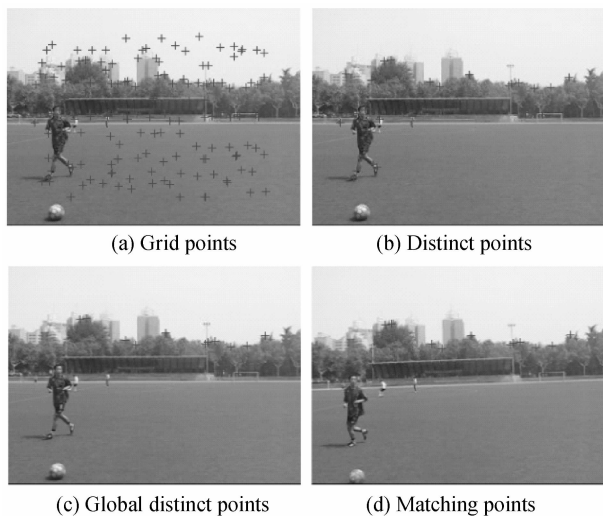


图 6 全局显著特征点提取和匹配结果图

Fig. 6 Result of global distinctpoints matching

图 6(b)为显著特征点提取结果,保留特征响应较大的前 20% 个特征点,提高特征点信息的独特性。结合图 5(d)的实验结果,剔除前景区域内的干扰特征点,得到如图 6(c)所示的全局显著特征点。图 6(d)为当前帧中配准的特征点结果图,特征点成功配准。对测试序列连续 50 帧处理,平均每帧的特征点运动估计的处理时间为 15.68 ms。

#### 3.2.2 较小单目标和多目标场景测试

图 7 为对两组实际高空拍摄抖动视频 Car 和 Cars 序列进行实验的结果,图中标记“◇”的特征点为删除的局部运动点,可见对于视频序列中存在单目标和多目标,均可以通过前景运动子块标记的预处理来去除

局部特征点,从而保证特征匹配的全局性。

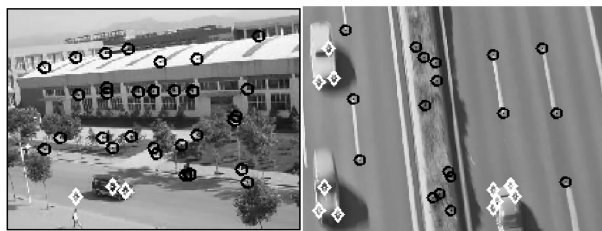


图 7 全局显著特征点配准结果

Fig. 7 Result of global distinct feature matching

表 1 为剔除前景区域的预处理和直接运算的时间比较,预处理可保证特征点对的全局性和正确性,无需迭代即可获得全局运动参量,和直接处理需要进行多次迭代相比,其时间平均缩减 30%。

表 1 预处理时间和迭代次数比较结果

Table 1 Comparison of preprocessing time

Images	Motion type	Iterations(direct/preprocessing)	Time(direct/preprocessing/mspf)
Ground	Rotation	5/0	39.57/26.38
Car	Translation	2/0	21.65/15.89
Cars	Scan	4/0	39.83/28.45

### 3.3 运动矢量自适应滤波结果比较

以对 Ground 序列的水平 and 垂直偏移分量滤波为例,图 8 给出了自适应滤波法 (Adaptive Filter, AF) 和经典 Kalman 滤波法<sup>[5]</sup>的滤波结果,由于水平方向上存在近匀速的摄像机扫描运动,因此图 8(a)实验曲线呈

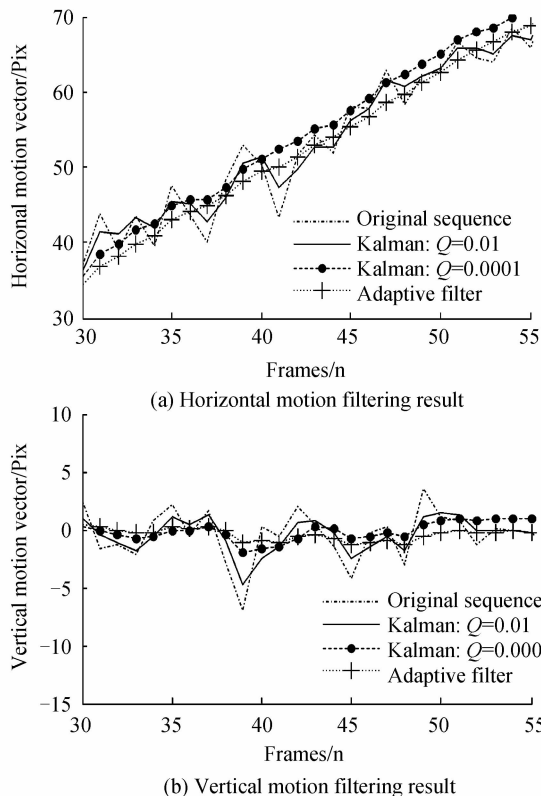


图 8 运动矢量滤波结果比较

Fig. 8 Comparison of motion filtering

稳定上升状态;摄像机在垂直方向仅存在随机抖动,因此图 8(b) 实验曲线在 0 位置波动. 由该图可见, Kalman 滤波中选择不同的过程噪声  $Q$  对补偿结果影响较大: 当  $Q$  值较大时, 无明显的滤波效果; 当  $Q$  值较小时, 由图 8(b) 观察 52 帧后原始抖动已经结束, 而其滤波结果却偏离 0 矢量, 从而造成滤波发散. 而自适应滤波法有效避免这种滤波发散的现象, 既能很好地平滑抖动分量, 同时有效跟踪摄像机的真实扫描运动, 其处理时间为 14.8 mspf.

### 3.4 补偿后帧间 PSNR 和差分图比较

#### 3.4.1 PSNR 值比较

采用帧间峰值信噪比 (Peak Signal to Noise Ratio, PSNR)<sup>[15]</sup> 来进一步检验稳像前后的帧间全局逼真度, 并与 Shakoor<sup>[9]</sup> 快速稳像方法进行对比. 由图 9 可以看出稳定后的视频序列帧间逼真度有了明显提高, 由于存在扫描运动, 故即使完全补偿图像的抖动分量, 图像仍不能完全重合<sup>[12]</sup>, 因而 GFM-AF 算法的 psnr 曲线在均值为 24 dB 处波动; 而 Shakoor<sup>[9]</sup> 快速稳像方法假设背景在图像的四个角落, 当前景目标落于这四个角落时会干扰运动估计, 从而影响补偿准确度.

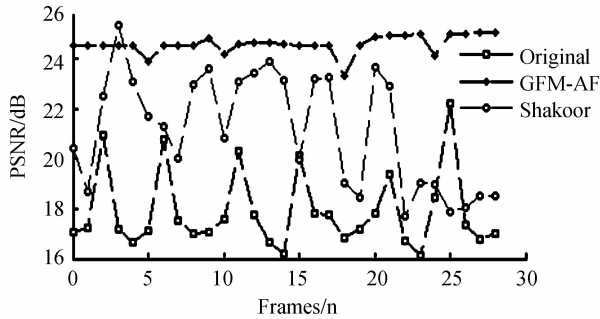


图 9 视频序列帧间 PSNR 曲线比较

Fig. 9 Comparison of PSNR

#### 3.4.2 运算时间和 PSNR 值比较

表 2 为 GFM-AF 算法与 Kumar<sup>[6]</sup> 和 Angelos<sup>[7]</sup> 处理各类抖动视频的结果比较, 可知 GFM-AF 算法在保证一定补偿准确度的条件下, 可以达到实时处理的能力.

表 2 各算法时间和补偿 PSNR 比较

Table 2 Comparison of time and PSNR

Video	Index	GFM-AF	Angelos	Kumar
Ground	PSNR	28.7	28.5	27.5
	Time/ms	43.7	127	91
Car	PSNR	28.8	27.9	28.2
	Time/ms	38.4	108	78
Cars	PSNR	25.3	25	25.2
	Time/ms	45.9	143	115
Coastguard	PSNR	28.8	27.6	28.3
	Time/ms	29.8	93	32

#### 3.4.3 大运动目标场景稳像差分图比较

图 10 为相邻两帧图像稳像后的差分图比较, 该视频中摄像机存在镜头垂直方向的抖动, 还含有较大的

旋转运动; 同时, 视场中存在较大尺寸前景目标. 任选连续的两帧作为参考帧和当前帧, 如图 10(a) 和 10(b); 图 10(c) 为对所有特征点直接求取运动参量补偿后的差分图, 可见, 由于该场景中前景目标尺寸较大, 局部特征点数量太多, 无法准确计算全局运动参量导致补偿失败; 利用 GFM-AF 算法对前景区域进行剔除后, 很好地消除了前景干扰, 得到精确的全局运动参量进行补偿, 图 10(d) 为补偿后的差分图. 可以看出, GFM-AF 算法能够很好地处理含较大运动前景的视频抖动, 具有优异的抗局部运动干扰的性能.

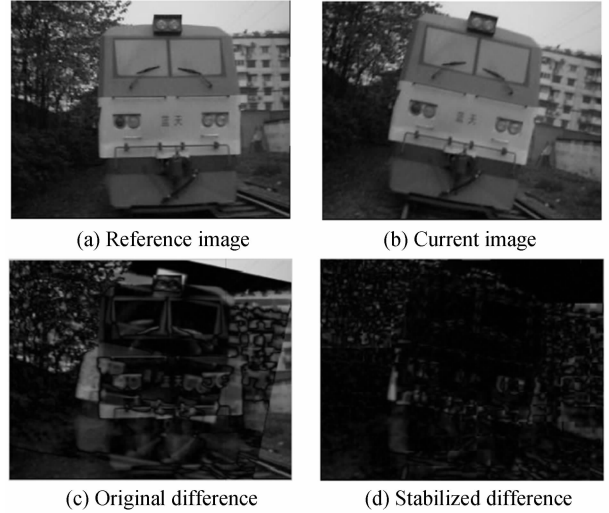


图 10 稳像后差分结果比较

Fig. 10 Comparison of image difference

## 4 结论

本文提出了抗前景干扰的自适应快速稳像算法, 通过快速剔除前景、网格筛选背景的全局显著特征点, 来提高特征匹配和全局运动估计的准确度. 自适应滤波不依赖预先设定的噪音方差阵, 提高滤波平滑性的同时防止滤波发散, 改善视觉观察的稳定性和平滑性, 且快速线性补偿方法保证了实时处理. 实验表明, 本算法在 Intel 酷睿 2 四核 2.33 GHz 的微机机上利用 VC++ 编程, 对拍摄视频 (320 × 240 像素) 达到实时处理能力, 且对于平移、大角度的旋转以及含较大或多个前景运动目标的复杂场景, 都有很好的稳像处理效果.

对于存在更为复杂的扫描运动, 可以结合粒子滤波的方法做进一步的研究与探讨; 并将进一步研究视觉平滑特性, 分析视觉对抖动的分辨率, 研究稳定程度的评价准则.

#### 参考文献

- [1] LI Peng-fei, XU Jin-kai, HAN Wen-bo, *et al.* The design and realization of dynamically stabilizing retinal image in human eye aberration wave front correction setup [J]. *Acta Photonica Sinica*, 2014, **43**(6): 0611002.  
李鹏飞, 许金凯, 韩文波, 等. 基人眼像差校正仪视网膜动态稳像研究 [J]. *光子学报*, 2014, **43**(6): 0611002.

- [2] DI Xiao-guang, JIN Wan-xin, YU Ying. Digital video stabilization for large-scale moving foreground object and rotation jitter[J]. *Optics and Precision Engineering*, 2014, **22**(1): 177-185.  
迪晓光, 靳万鑫, 余颖. 大运动前景和旋转抖动视频的快速数字稳定[J]. *光学精密工程*, 2014, **22**(1): 177-185.
- [3] 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.
- [4] KIM S K, KANG S J. Feature point classification based global motion estimation for video stabilization [ J ]. *IEEE Transactions on Consumer Electronics*, 2013, **59**(1): 267-272.
- [5] LIU K X, QIAN J, YANG R K. Block matching algorithm based on RANSAC algorithm [ C ]. 2010 International Conference on IASP: 223-227.
- [6] SANJEEV K, HALEH A, MAINAK B, *et al.* Real-time affine global motion estimation using phase correlation and its application for digital image stabilization [ J ]. *IEEE Transactions on Image Processing*, 2011, **20**(2): 3406-3418.
- [7] ANGELOS A. Amanatiadis, ioannis andreadis. digital image stabilization by independent component analysis [J]. *IEEE Transactions on Instrumentation and Measurement*, 2010, **59**(7): 1755-1763.
- [8] WANG C T, KIM J H. Robust digital image stabilization using the Kalman filters[J]. *IEEE Transactions on Consumer Electronics*, 2009, **55**(1): 6-14.
- [9] SHAKOOR M H, MOATTARI M. A new fast method for digital image stabilization [C]. International Conference on ACTE, 2010, 2:350-354.
- [10] MORITZ G. On the peripheral point spectrum and the asymptotic behavior of irreducible semigroups of Harris operators[J]. *Journal Positivity*, 2013, **17**(3): 875-898.
- [11] QU Jing-jing, XIN Yun-hong. Combined continuous frame difference with background difference method for moving object detection[J]. *Acta Photonica Sinica*, 2014, **43**(7): 071002.  
屈晶晶, 辛云宏. 连续帧间差分与背景差分相融合的运动目标检测方法[J]. *光子学报*, 2014, **43**(7): 071002.
- [12] ZHU Juan-juan, GUO Bao-long, WU Xian-xiang. Electronic image stabilization based on global motion estimation with motion classification [J]. *Control and Decision*, 2012, **27**(10): 1575-1578.  
朱娟娟, 郭宝龙, 吴宪祥. 电子稳像中基于运动分类的全局运动估计算法[J]. *控制与决策*, 2012, **27**(10): 1575-1578.
- [13] XIONG X, QIN K H. Linearly estimating all parameters of affine motion using radon transform[J]. *IEEE Transactions on Image Process*, 2014, **23**(10): 4311-4321.
- [14] FORBES J R. Extended Kalman filter and sigma point filter approaches to adaptive filtering [ C ]. AIAA Guidance, Navigation and Control Conference, 2010: 1-15.
- [15] ALEXANDER T. Visual-PSNR measure of image quality [ J ]. *Journal of Visual Communication and Image Representation*, 2014, **25**(5): 874-878.