

文章编号:1004-4213(2010)s1-0023-6

一种基于检测块动态选择的稳像算法

邱家涛¹, 李玉山¹, 王彩玲², 刘学斌²

(1 西安电子科技大学 电路 CAD 研究所, 西安 710071)

(2 中国科学院西安光学精密机械研究所, 西安 710119)

摘 要:针对包含运动物体视频序列的稳像问题,提出了一种基于运动目标检测的数字稳像算法。该算法运用估计的全局运动参数补偿帧间全局运动,得到运动补偿前帧和当前帧的差值图像,并将差值图像分割成一系列的前景块和背景块,仅跟踪背景块来估计下一轮的全局运动参数。仿真结果表明,提出的稳像算法不仅计算复杂度低,而且估计的全局变换参数具有相当高的帧间变换保真度,可以用于占据半个像面的动目标的视频序列。

关键词:数字稳像;图像分割;全局运动估计;运动目标

中图分类号:TP391.4

文献标识码:A

doi:10.3788/gzxb201039s1.0023

0 引言

在视频图像序列的捕获过程中,人为抖动或载体震动等原因,会造成图像序列的不稳定性。解决图像序列不稳定常采用两种方案:采用惯性运动传感器并相应控制棱镜或者 CCD/CMOS 传感器进行物理补偿^[1]和采用数字图像处理领域的数字稳像^[2](Digital Image Stabilization, DIS)技术。

DIS 主要分三步进行^[3]:全局运动估计(Global Motion Estimation, GME)、运动参量滤波和图像变换。在全局运动估计中由于误匹配产生的错误运动参量估计,将会传递到运动补偿矩阵里,导致原图像非准确补偿,从而产生不自然的修补边缘图像。因此,全局运动估计作为第一步是至关重要的。现有的全局运动估计算法大致可分为两类:基于图像特征的方法^[1]和直接利用像素灰度信息的图像配准方法^[4-5]。近来也有把两种方法结合到一起的研究成果报道^[6]。基于图像特征的方法由于只用少量的匹配特征对,拥有比灰度图像配准方法更好的实时性。但图像特征的方法更易于受视频序列中运动物体的影响,针对这个问题一些学者利用运动物体和背景运动的不同,提出了一些基于运动矢量分类的方法,然而运用光流分析对运动矢量进行分类通常实时性不好。DIS 中另一种经常使用的块匹配算法(Block Matching Algorithm, BMA)为灰度图像配准方法。虽然 BMA 算法在许多应用中能准确地估计出全局

运动参量,但当视频图像中包含大运动物体时,对应于运动物体的局部块矢量往往会被叠加到最终的全局运动估计参量中,导致不正确的全局运动估计。

针对包含运动物体视频序列的稳像问题,本文提出了一种基于直方图运动目标分割的全局运动估计方法,形成一种新颖的数字稳像算法。提出的全局运动估计方法为:

1) 利用估计的帧间全局运动参量补偿摄像机的运动,得到运动补偿前帧与当前帧的绝对差分图像。

2) 把绝对差分图像的中央区域分割成一系列极坐标块,并计算每个极坐标块上的平均绝对差值(Mean Absolute Difference, MAD);

3) 用这些 MAD 构建平均灰度直方图,并用一个最优阈值对其分割,从而把差分图像的中央区域分割为前景块和背景块;

4) 跟踪从这些背景块中选取的检测块以估计下一轮的全局运动参量。

提出的全局运动估计方法,只用少量的极坐标块的平均灰度差来构建直方图,计算量小且有相当高的准确度,能够满足数字图像稳定的要求。

1 基于检测块动态选择的稳像算法

1.1 全局运动估计

通常设定一个坐标变换模型来配准相邻两帧图像,以估计摄像机的帧间运动参量。本文采用的简化仿射变换模型为

基金项目:国家重点基础研究发展计划(No. 2009CB724005)资助

第一作者:邱家涛(1964-),男,高级工程师,主要研究方向为图像处理。Email:jtqiu@mail.xidian.edu.cn

责任作者:李玉山(1945-),男,教授,主要研究方向为图像处理和信号完整性。Email:yshli@mail.xidian.edu.cn

收稿日期:2010-11-01; **修回日期:**2010-12-16

$$\mathbf{v} = \mathbf{F}_i^j(\mathbf{x}; \mathbf{p}) = \mathbf{M}(\mathbf{W}_i^j) \mathbf{u}, \mathbf{x} = (x, y)^T, \mathbf{u} = (x, y, 1)^T \quad (1)$$

式中 \mathbf{W}_i^j 为坐标变换矩阵, 定义为

$$\mathbf{W}_i^j = \begin{bmatrix} 1+p_1 & -p_2 & p_3 \\ p_2 & 1+p_1 & p_4 \\ 0 & 0 & 1 \end{bmatrix}$$

式中 \mathbf{x} 表示图像 I_i 中像素的坐标矢量, \mathbf{v} 表示图像 I_j 中的亚像素位置, $\mathbf{p} = (p_1 \ p_2 \ p_3 \ p_4)^T$ 为变换参量矢量, $\mathbf{M}(\cdot)$ 为取 3×3 矩阵前两行构成 2×3 矩阵的算子. 式(1)只把图像 I_i 的坐标变换到 I_j 的坐标, 像素 \mathbf{x} 上的灰度值 $I_i(\mathbf{x})$ 被映射到 I_j 中与其对应的 $I_j(\mathbf{v})$. 为简化分析, 用 $\mathbf{W}_i^j \mathbf{x}$ 表示 $\mathbf{M}(\mathbf{W}_i^j) \mathbf{u}$. \mathbf{F}_i^j 的逆变换表示为 $\mathbf{F}_i^j = (\mathbf{F}_i^j(\mathbf{x}; \mathbf{p}))^{-1}$, 即 $\mathbf{x} = (\mathbf{W}_i^j)^{-1} \mathbf{v}$.

1.1.1 检测块的提取

由于图像背景的运动(全局运动)才能反映摄像机的运动, 因此只能从背景中选取估计摄像机运动的检测块, 而不能从叠加有自身运动分量的前景(运动物体)中选取. 在提出算法中, 用全局运动估计把前帧 I_{t-1} 映射到当前帧 I_t , 以补偿背景的运动, 用

\mathbf{W}_{t-1}^t 表示使得运动补偿前帧 I_{t-1}^w 与当前帧 I_t 最佳匹配的全局运动估计, 即: $I_{t-1}^w(\mathbf{x}) \leftarrow I_{t-1}((\mathbf{W}_{t-1}^t)^{-1} \mathbf{x})$, \mathbf{x} 表示当前帧中像素的坐标矢量, $(\mathbf{W}_{t-1}^t)^{-1} \mathbf{x}$ 表示 I_{t-1} 中像素的坐标矢量. 然后从补偿前帧 I_{t-1}^w 与 I_t 的差分图像中分割出运动物体, 并从剩余区域中选取背景检测块用于下一轮的全局运动估计. 绝对差分图像可表示为

$$I_t^d(\mathbf{x}) = |I_t(\mathbf{x}) - I_{t-1}^w(\mathbf{x})| \quad (2)$$

用绝对差分图像分割运动目标的实例如图 1. 前帧经过摄像机运动补偿后, 差分图像中与背景运动不同的运动目标区域显示出强烈的灰度变化, 即运动目标所在的高亮区域比背景区域具有高得多的绝对差值. 因此, 利用这种强烈的灰度反差提出了一种基于直方图的快速图像分割方法来检测运动目标, 然后从剩下的背景中选取检测块. 在本文中, 为防止摄像机的大幅度的运动导致在像面中搜索不到匹配块, 在图像的中央区域选取检测块.

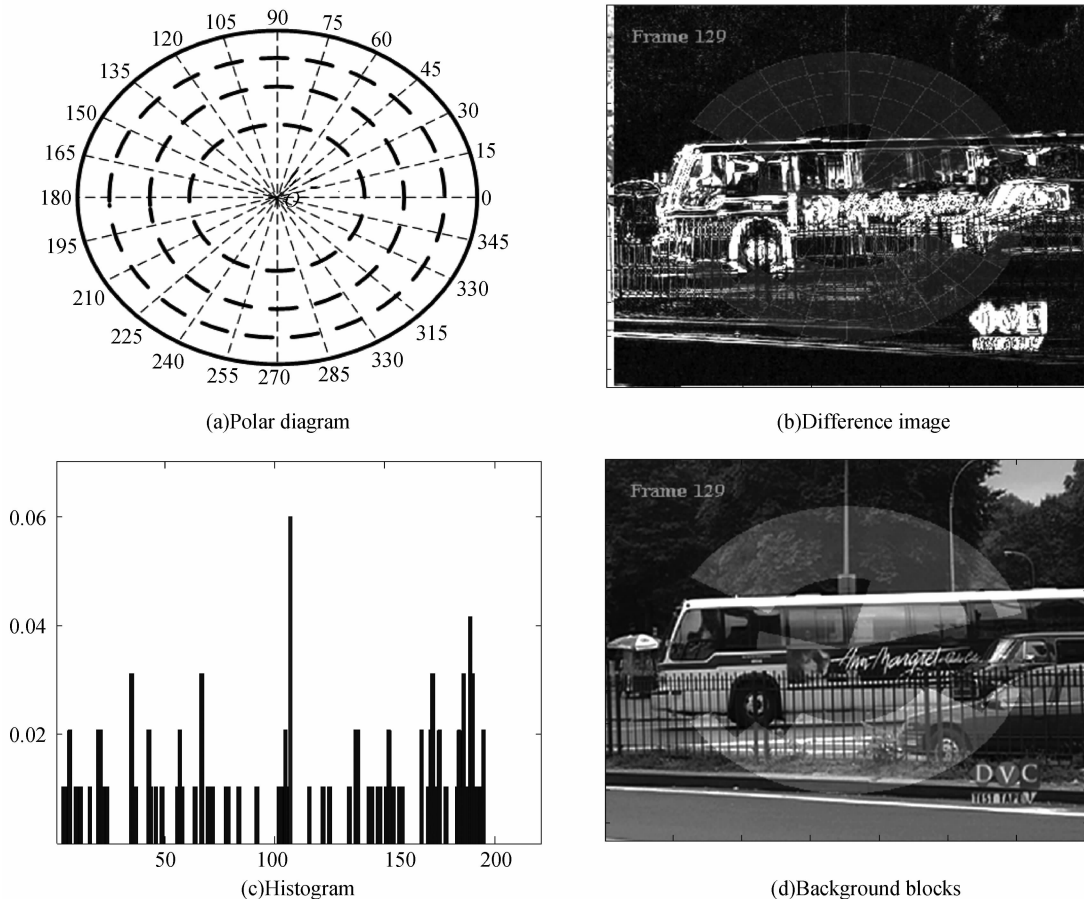


图 1 检测块提取示意

Fig. 1 Illustration of the detection blocks extraction

本文运用直方图图像分割方法分割绝对差分图像的中央区域(如图 1). 直方图的构建分为三步. 首先, 把这个中央区域分割成一系列互不重叠的子块, 采用极坐标系 (r, θ) 进行分割. 图像中央区域被分割

成一系列极坐标块 $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_i$, 设定 $n=96$. 每个极坐标块 \mathbf{R}_i 的 MAD 为

$$m_i = \left(\sum_{\mathbf{x} \in \mathbf{R}_i} |I_t^d(\mathbf{x})| \right) / \# \mathbf{R}_i \quad (3)$$

式中 $I_t^d(\mathbf{x})$ 表示由式(2)定义的绝对差分图像, $\# \mathbf{R}_i$

表示极坐标块 R_i 中像素的个数, m_i 表示 R_i 的平均差分. 并用这些平均差分构建一个矢量

$$\mathbf{v} = \mathbf{e}^{-\frac{[m_1, \dots, m_i, \dots, m_n]}{2\sigma_d^2}} \quad (4)$$

式中 σ_d^2 表示由试验确定的高斯密度函数的方差. 最后用矢量中的元素构建统计直方图, 即

$$\mathbf{H} = \{h_j\}_{j=0, \dots, L-1} \quad (5)$$

分割灰度直方图的 Otsu^[7] 方法可以直接用于求取 \mathbf{H} 的最优阈值 t^* . 用阈值 t 将 \mathbf{H} 分成两类 $C_1 = \{0, 1, \dots, t\}$ 和 $C_2 = \{t+1, t+2, \dots, L-1\}$, C_1 和 C_2 分别对应于运动目标和背景. 最优阈值 t^* 为

$$t^* = \arg \max_{0 \leq t < L} [\omega_1(t)\mu_1(t)^2 + \omega_2(t)\mu_2(t)^2] \quad (6)$$

式中 $\mu_1(t)$ 和 $\mu_2(t)$ 分别表示两类的数学期望, $\omega_1(t)$ 和 $\omega_2(t)$ 分别表示两类的概率.

图 1 给出了视频剪辑 bus_cif 中运动补偿第 128 帧与第 129 帧的差分图像, 按上述方法构建的直方图如图 1(c), 横轴第 103 个 bin 上的 bar 指示其相应的最优阈值 t^* . 图 2 给出了一个检测块提取的实例. 格子区域背景上的极坐标块对应于高于阈值 t^* 的背景块, 理论上, 所有的这些背景块都可被当作检测块, 用于下一轮的全局运动估计, 为提高算法的实时性, 检测块只在极坐标区域的最外三圈上选取, 而且每隔一个角度选一次.

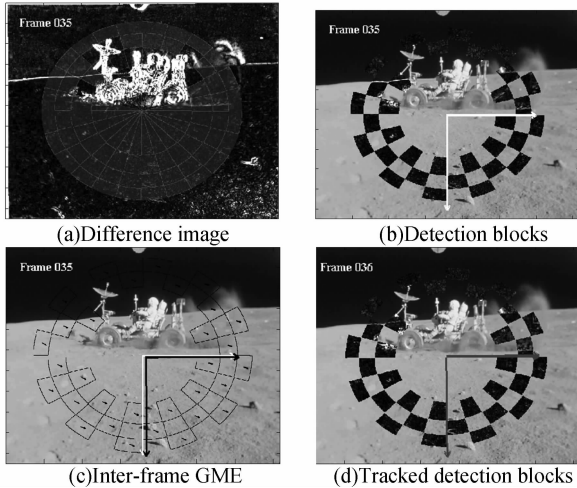


图 2 帧间全局运动参量估计示意

Fig. 2 Illustration of the inter-frame GME

1.1.2 帧间运动参量估计

为提高算法的实时性, 本文通过跟踪图像的一个背景子区域(所有检测块中的像素集合, 如图 2)来估计全局运动参量. 提出的算法采用可逆复合算法(Inverse Compositional Algorithm, ICA)的框架^[8]. 对于彩色图像, 3 个 RGB 颜色通道的像素值用下式转换成灰度像素值: $0.30R + 0.59G + 0.11B$ ^[3].

用可逆复合算法迭代求解全局运动参量, 目标函数为

$$\sum_{x \in \Omega} [I_t(\mathbf{F}(x; \Delta p)) - I_{t+1}(\mathbf{F}(x; p))]^2 \quad (7)$$

为使目标函数最小化, 每一步迭代更新变换矩阵, 即

$$\mathbf{F}(x; p) \leftarrow \mathbf{F}((\mathbf{F}(x; \Delta p))^{-1}; p) \quad (8)$$

式中 $\mathbf{F}(\cdot)$ 表示定义的简化仿射变换, 为表达清晰, 上下标被忽略, Ω 表示所有检测块中的像素集合, $(\mathbf{F}(x; \Delta p))^{-1}$ 表示 $\mathbf{F}(x; \Delta p)$ 的逆变换. 迭代结束后的变换 $\mathbf{F}(\cdot)$ 最优, 使得当前帧 $I_t(x)$ 和 $I_{t+1}(\mathbf{F}(x; p))$ 达到最优匹配. 图 2 给出了一个帧间变换参量估计的实例. Apollo 剪辑第 35 帧上的检测块的中点的运动矢量如图 2(c), 通过跟踪这些检测块到第 36 帧, 可以准确地估计出两帧间的全局运动参量. 估计出所有帧的帧间全局运动参量, 就能构成一个全局运动参量链.

1.2 运动补偿

在全局运动参量链中包含两个全局运动分量, 一个是对应于摄像机扫描运动的有意运动分量^[5], 另一个是摄像机随机抖动引起的高频无意运动分量. 如图 3, 用运动补偿矩阵 S_t 补偿高频无意运动分

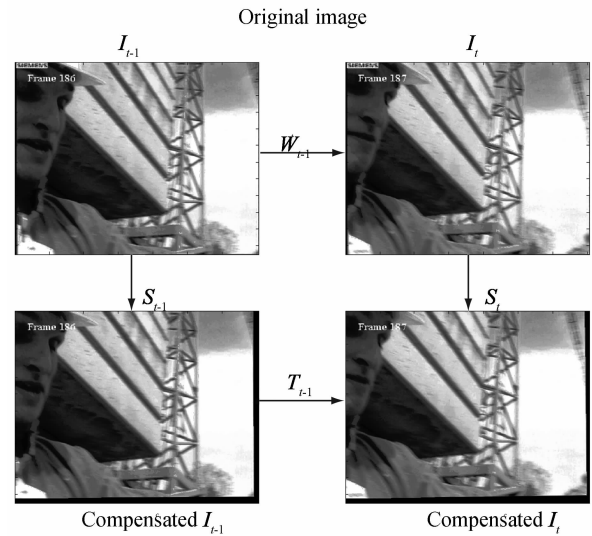


图 3 变换矩阵示意

Fig. 3 Illustration of the transformation matrices

量, S_t 可表示为

$$S_t = \frac{[\sum_{k \in N_1} g(k-t)(W_k^t)^{-1} + \sum_{k \in N_2} g(k-t)W_k^t]}{\sum_{j \in N_1 \cup N_2} g(j-t)} \quad (9)$$

式中: $N_1 = \{k: t-\lambda \leq k < t\}$; $N_2 = \{k: t < k \leq t+\lambda\}$.

式(9)中 W_k^t 表示从图像 I_t 到 I_k 的全局运动估计, 即坐标变换矩阵, $g(\Delta)$ 表示高斯函数, 2λ 是设定的当前帧 I_t 的邻帧的帧数. 把 I_t 到其邻帧的坐标变换矩阵对称相加则使得对称的稳定分量相互抵消, 而随机的抖动分量得以保留, 从而得到运动补偿矩阵 S_t . 这样就可以用一系列运动补偿矩阵 $\{S_0, \dots, S_t\}$, 通过图像变换把原始图像序列变换到补偿图像序列, 即

$$I_i'(\mathbf{x}) \leftarrow I_i((\mathbf{S}_i)^{-1}\mathbf{x}) \quad (10)$$

式中 \mathbf{x} 表示补偿帧 $I_i'(\mathbf{x})$ 中像素的坐标, $(\mathbf{S}_i)^{-1}\mathbf{x}$ 表示原始帧 I_i 中像素的坐标. 补偿后平滑运动参量链 $\{\mathbf{T}_0^i, \dots, \mathbf{T}_{i-1}^i\}$ 中的 \mathbf{T}_{i-1}^i 表示为

$$\mathbf{T}_{i-1}^i = \mathbf{S}_i \mathbf{W}_{i-1}^i (\mathbf{S}_{i-1})^{-1} \quad (11)$$

2 实验结果与分析

用多个视频剪辑对提出的稳像算法进行测试. 本文只列出四个视频剪辑, 它们分别是 ship、foreman_cif、bus_cif 和 Apollo 视频剪辑. 视频剪辑的分辨率被放大到 720×576 , 并对原始图像帧不规则的边缘进行了裁剪, 以生成连续的修补边缘图像.

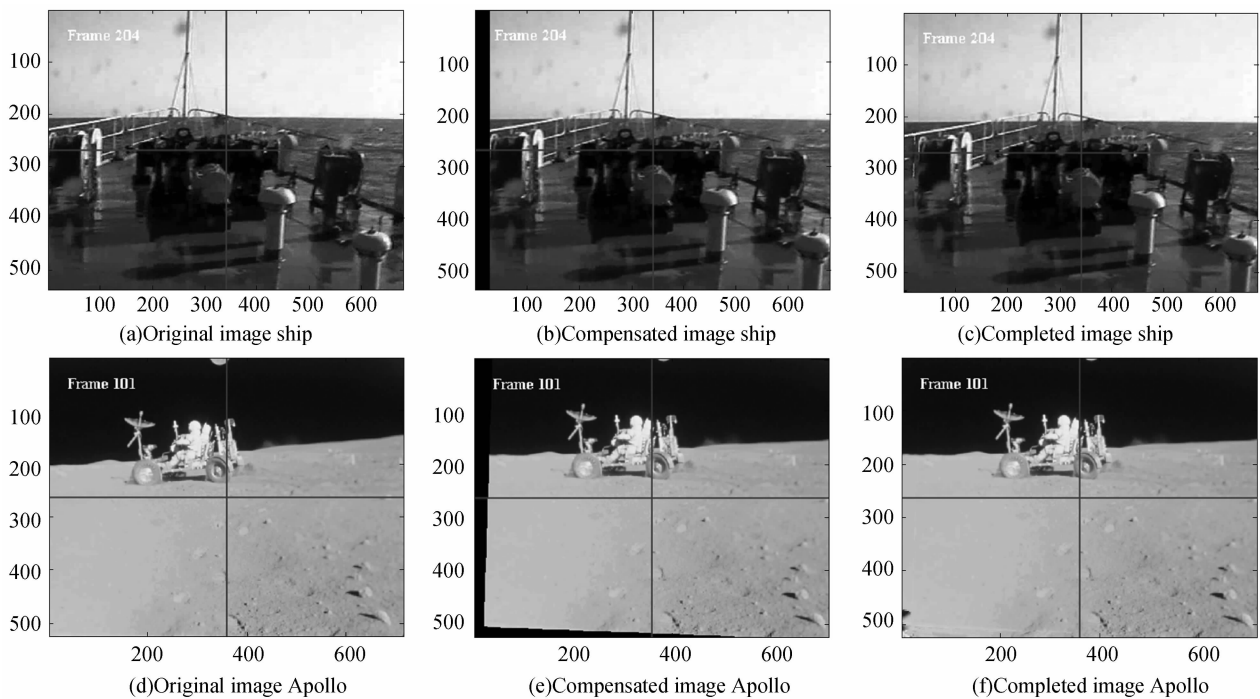


图 4 提出稳像算法的稳像实例

Fig. 4 Stabilizing examples by the proposed method

两帧图像的重合度通常用峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 测量. 例如, 用背景峰值信噪比测量运动补偿前帧与当前帧的重合度^[9]. 本文也采用背景峰值信噪比, 然而仅仅计入网格背景子区域

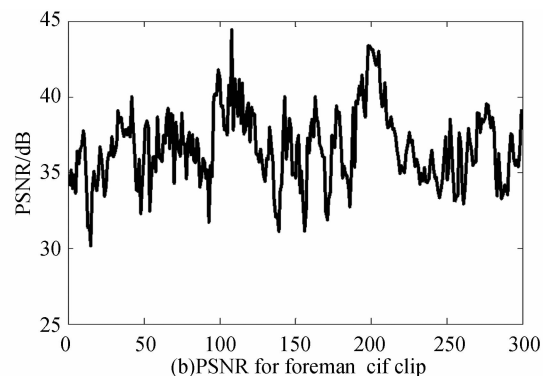
图 4 给出了稳定 ship 和 Apollo 视频剪辑的两个稳像实例. 对于 Apollo 视频剪辑, 提出的稳像算法能稳定剧烈摇晃的太空图像. 在 ship 视频剪辑中, 一条大船和起伏的波浪占据了大半个像面, 然而天空是静止的, 甲板也是相对静止的, 对其上提取的检测块仍然能准确地跟踪.

如图 4(b) 和 4(e), 由于运动补偿, 最后得到的补偿图像的边缘会存在未定义区域. 用补偿图像帧的邻帧图像的变换图像来重新定义缺失的图像像素, 采用静态图像拼接方法和光流传播动态图像修补方法^[3], 这样生成的完整的图像如图 4(c) 和 (f).

上的像素而排除前景区域上的像素, 如图 2(a) 和 5(a). 网格背景子区域是用 1.1.1 节的检测块提取方法得到的.



(a) Difference image



(b) PSNR for foreman_cif clip

图 5 提出全局运动估计算法性能的评估

Fig. 5 Evaluation performance of the proposed GME

补偿前帧 I_{t-1}^c 和当前帧 I_t 之间的背景峰值信噪比可定义为

$$\text{PSNR}(t) = 10 \log_{10} \frac{I_{\max}^2}{\text{MSE}(t)} \quad (12)$$

式中 $\text{MSE}(t) = \frac{1}{l} \sum_{x \in \chi} (I_t^d(x))^2$, $I_t^d(x)$ 表示式(2)定义的差分图像, χ 表示网格背景子区域上的像素集合,如图 2(a) 和 5(a). l 表示 χ 中包含的像素的个数. $\text{MSE}(t)$ 表示 χ 中的像素的均方误差 (Mean-Square-Error). I_{\max} 表示最大像素灰度值 (255). 图 5(b) 给出了 foreman_cif 剪辑的背景峰值信噪比.

视频序列的平均背景峰值信噪比可表示为

$$\text{PSNR}_{\text{ave}} = \frac{1}{N_f - 1} \sum_{t=2}^{N_f} \text{PSNR}(t) \quad (13)$$

式中 N_f 表示视频序列包含的帧数.

对提出稳像算法和 Sift-VS 稳像算法^[1] 的平均背景峰值信噪比进行了比较. 在提出算法中计算 PSNR_{ave} 仅仅计入网格背景子区域上的像素, 在 Sift-VS 算法中, 帧间变换保真度 (Interframe Transformation Fidelity, ITF) 的计算采用背景差分规则^[9], 尽管 Sift-VS 算法也采用部分与本文相同的视频剪辑, 但没有考虑动目标问题, 因此只能仍然沿用其采用的背景差分规则. 其定义的原始 ITF 与式(13)定义的 RSNR_{ave} 相对应. 表 1 列出了两个算法的平均背景峰值信噪比, 其中包含四个视频序列的仿真结果.

表 1 Sift-VS 稳像算法与本文算法 PSNR_{ave} 的对比

Table 1 Comparison of PSNR_{ave} between Sift-VS and the proposed algorithms

Video Clip	Proposed	Sift-VS
	$\text{PSNR}_{\text{ave}}/\text{dB}$	
Foreman_cif	36.65	25.75
Apollo	41.58	28.35
Bus_cif	29.38	15.11
Ship	38.22	24.31

显然, 提出算法的 PSNR_{ave} 相当高, 充分说明相继补偿前帧和当前帧重合背景上的差值近乎为零, 即估计的帧间全局变换参量具有很高的 ITF. 因此, 与 Sift-VS 稳像算法相比, 本文算法对运动物体有较强的鲁棒性. 检测块的提取会带来一些额外的计算资源开销, 然而其计算复杂度较低, 本文用 2.0 GHz CPU 仿真, 其仿真时间仅为 6 ms 左右, 这并不妨碍把这种方法应用到对实时性要求较高的稳

像系统. 仿真结果表明, 对上述四个包含运动物体的视频剪辑均能成功的实现稳像.

3 结论

本文针对包含大运动物体视频序列的稳像问题, 提出了一种对于运动物体鲁棒的稳像算法. 用一种直方图图像分割方法分割出对应于局部运动的运动物体区域, 只跟踪从剩余背景区域选择的检测块, 以估计帧间全局运动参量. 通过仿真对提出稳像算法对于运动物体视频序列的稳像性能进行评估, 仿真结果表明, 检测块的提取效率高, 估计的全局变换参量也具有较高的 ITF.

然而提出稳像算法尚需在进一步的研究中完善. 例如, 图像块选取为固定位置是有局限性的, 当特大的运动物体完全遮挡住这个区域时会找不到背景检测块, 也需要克服图像差分方法有时只能够分割出运动物体部分边缘的局限性.

参考文献

- [1] BATTIATO S, GALLO G, PUGLISI G, *et al.* Sift features tracking for video stabilization [C]. *Proceedings of 14th ICIAP*, 2007: 825-830.
- [2] CHO W H, HONG K S. A fast CIS still image stabilization method without parallax and moving object problems [J]. *IEEE Transactions on Consumer Electronics*, 2008, **54** (2): 197-205.
- [3] MATSUSHITA Y, OFEK E, GE W, *et al.* Full-frame video stabilization with motion inpainting [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(7): 1150-1163.
- [4] QI B, GHAZAL M, AMER A. Robust global motion estimation oriented to video object segmentation [J]. *IEEE Transactions on Image Processing*, 2008, **17**(6): 958-967.
- [5] LITVIN A, KONRAD J, KARL W C. Probabilistic video stabilization using kalman filtering and mosaicking [C]. *Proceedings of IS&T/SPIE Symp on Electronic Imaging, Image and Video Communications*, 2003, **1**: 663-674.
- [6] SHI Lei, SU Xiu-qin, XIANG Jing-bo. Electronic image stabilizing algorithm based on feature block matching [J]. *Acta Photonica Sinica*, 2008, **37**(1): 202-205.
石磊, 苏秀琴, 向静波. 一种基于特征块匹配的电子稳像方法 [J]. *光子学报*, 2008, **37**(1): 202-205.
- [7] NG H F. Automatic thresholding for defect detection [J]. *Pattern Recognition Letters*, 2006, **27**(14): 1644-1649.
- [8] BAKER S, MATTHEWS I. Lucas-kanade 20 years on: a unifying framework [J]. *International Journal of Computer Vision*, 2004, **56**(3): 221-255.
- [9] MERCENARO L, VERNAZZA G, REGAZZONI C. Image stabilization algorithms for video-surveillance application [C]. *Proceedings of 8th ICIP*, 2001, **1**: 349-352.

A Dynamic Detection Blocks Selection-based Image Stabilization Algorithm

QIU Jia-tao¹, LI Yu-shan¹, WANG Cai-ling², LIU Xue-bin²

(1 *Institute of Electronic CAD, Xidian University, Xi'an 710071, China*)

(2 *Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China*)

Abstract: A moving object detection-based image stabilization algorithm was proposed for addressing the moving object problem of image stabilization. In the algorithm, the estimated global motion parameters were used to compensate the inter-frame global motion, and the difference image between the motion-compensated previous and current frames was obtained. Then, the difference image was divided into a series of foreground and background blocks. And the global motion parameters in the next run were estimated only by tracking the background blocks. Simulation results show that the proposed algorithm has low computational complexity, and achieves quite high interframe transformation fidelity, which is suitable for stabilizing video clips with large moving objects.

Key words: Digital image stabilization; Image segmentation; Global motion estimation; Moving object