

文章编号: 1000-582X(2012)10-093-06

具有前景目标的动态场景视频快速稳像算法

何 凯, 牟聪翀, 远中文, 何海龙, 邹 刚

(天津大学 电子信息工程学院, 天津 300072)

摘 要: 为了提高具有前景目标的动态场景视频的稳像效果, 采用了一种基于块绝对差分的运动目标自动提取方法。在提取运动目标的基础上, 提出通过自适应设定相关阈值的方法, 来预判相关背景模块是否具有足够的信息, 以减少用于运动矢量估计的模块数量。仿真实验结果表明, 对于具有前景目标的动态场景视频, 在减少运动目标提取时间的基础上, 能够有效提高稳像算法的速度和精度, 对实际视频取得了比较理想的稳像效果。

关键词: 动态背景视频; 前景目标提取; 电子稳像; 运动参数估计; 块匹配

中图分类号: TP391

文献标志码: A

Fast stabilization algorithm of dynamic scene video with multiple foreground objects

HE Kai, MU Congchong, YUAN Zhongwen, HE Hailong, ZOU Gang

(School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: In order to improve the stabilization effect of dynamic scene video sequences with multiple moving foreground objects, a new method based on block absolute difference is proposed to extract the foreground objects automatically. Based on extracting the moving objects, the background blocks are selected to estimate the motion parameters with the defined threshold, so the background blocks for estimating motion parameters are greatly reduced. For moving scene with multiple moving foreground objects, due to the reduction of the time of extracting moving objects, the simulation results shown that the running time of the proposed method is reduced effectively, the stabilization effect is improved, and more ideal effect is realized compared with the traditional method.

Key words: dynamic scene video sequences; foreground objects extraction; electronic stabilization; motion parameter estimation; block matching

在实际工程中, 摄像机拍摄的内容往往不仅仅是自然背景, 还包含许多运动的前景目标, 如行人、车辆等等, 这些运动目标使得视频图像内部产生一定的局部运动。在进行稳像处理时, 全局运动矢量的估计精度会受到这些局部运动的影响而降低, 从而影响稳像的最终效果, 因此在求取运动参数时需要先对前景运动目标进行提取, 以去除干扰, 提高参

数估计的精度^[1]。

运动目标的提取技术, 是图像处理技术中一个比较热门的课题, 主要可以分为静止背景下的运动目标提取与运动背景下的目标提取 2 大类^[2]。目前常用的静止背景下的运动目标提取方法主要有背景差分法、帧间差分法和光流法^[3-5]。其中, 背景差分法是利用当前图像与背景图像的差分来检测运动区

收稿日期: 2010-10-18

基金项目: 国家自然科学基金资助项目(61002030; 61271326)。

作者简介: 何凯(1972-), 男, 天津大学, 副教授, 主要从事图像处理, 视频处理方向研究, (E-mail)hekai626@163.com

域的一种技术,主要适用于背景已知的情況^[6];相邻帧间差分法是通过相邻两帧图像的差值计算,获得运动物体位置和形状等信息的运动目标检测方法;而光流法是利用序列图像中每个像素的矢量特征来检测运动区域,在背景运动的情况下也能检测出前景运动目标,但是该方法受光照情况影响较大,运算复杂,不易于实时处理,对硬件的要求也比较高^[7-8]。

对于运动的背景,由于前景运动目标与摄像机之间存在相对运动,因此,其运动目标提取过程要比在静态背景下复杂得多^[9]。目前采取的方法通常是先对相邻各帧图像进行全局运动参数估计和补偿,再提取前景目标^[10-12]。

对于抖动的视频序列,背景必然不是静止的。一般而言,背景图像灰度分布比较均匀,在有前景目标出现的地方,背景的灰度分布会发生突变,突变处的像素点的均值和方差与周围背景相差很大^[13-15],因此,可以考虑利用目标和背景图像之间的灰度分布差异来提取前景运动目标,而不必事先进行帧间全局运动参数估计和配准。为此,提出了一种基于块绝对差分的方法来提取前景目标,并通过设定阈值来预判背景模块是否具有足够的梯度信息来估计运动矢量,从而有效减少了用于参与计算的模块数量,提高了运算速度,最后利用求解超定方程的方法获取全局运动参数,从而实现了具有前景目标的动态背景视频的稳像,取得了良好的实际效果。

1 传统前景目标提取方法

动态背景下的运动目标提取是运动目标提取领域的难点之一,目前通常采用基于监督分类和运动矢量场的方法解决。基于监督分类的方法虽然能够实现运动背景下的运动目标提取,但需要事先进行学习和训练,且只适用于对动态场景中特定目标的检测,如行人。基于运动矢量场的方法,则需要先将相邻图像进行配准,进行运动补偿后再进行差分运算,以实现前景运动目标的提取,其算法框图如图 1。

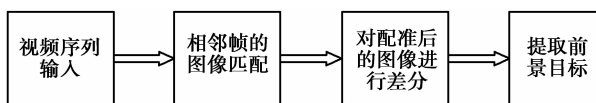


图 1 基于运动矢量场法在运动背景下提取前景目标框图

从图中可以看出,基于运动矢量场的方法虽然能够比较准确地提取出前景运动目标,但需要事先

进行整幅图像的运动参数估计,并需要基于运动场对视频图像进行分割,计算量大,这使得它的应用范围受到了很大程度的限制^[12]。

2 研究算法

2.1 前景目标提取

基于前景运动目标和背景图像之间的灰度分布差异,提出了一种基于块绝对差分的方法来提取前景目标,其具体方法如下:

1) 首先选取一帧图像,将图像粗定位出一个模板 f , f 的 4 个边距离图像的边界均大于 M 个像素点,且保证该模板能平分成 $m \times n$ 个大小为 $M \times M$ 的子块,平分后的每个子块记为 $C_{i,j}$ ($i=1, 2, \dots, m$; $j=1, 2, \dots, n$), 其中 i, j 分别代表 $C_{i,j}$ 所在的行列数。

2) 计算第 i 行(或第 j 列)的所有子块与第 $i-1$ 行(或第 $j-1$ 列)对应子块之间的像素绝对差,记为 $D_{i,j}^{\text{row}}$ (或 $D_{i,j}^{\text{col}}$)

$$D_{ij}^{\text{row}} = \sum_{m=1}^M \sum_{n=1}^M |f_k(x_{ij} + m, y_{ij} + n) - f_k(x_{ij} + m, y_{ij} - M + n)|, \quad (1)$$

$$D_{ij}^{\text{col}} = \sum_{m=1}^M \sum_{n=1}^M |f_k(x_{ij} + m, y_{ij} + n) - f_k(x_{ij} + m - M, y_{ij} + n)|, \quad (2)$$

其中 $x_{i,j}, y_{i,j}$ 为当前块 $C_{i,j}$ 左上角顶点的坐标, $f_k(x, y)$ 为当前帧 x, y 处的灰度值, $M \times M$ 为块的大小。对于当前块 $C_{i,j}$, 令其周围的块为 R_u, R_d, R_l, R_r (分别表示上、下、左、右 4 个位置), 则当前块 $C_{i,j}$ 和 R_u, R_d, R_l, R_r 之间的像素绝对差记为 $D_{i,j}^k$ ($k=1, \dots, 4$)

$$\begin{cases} D_{i,j}^1 = D_{i,j}^{\text{row}}, D_{i,j}^2 = D_{i+1,j}^{\text{row}}; \\ D_{i,j}^3 = D_{i,j}^{\text{col}}, D_{i,j}^4 = D_{i,j+1}^{\text{col}}. \end{cases} \quad (3)$$

3) 计算 $D_{i,j}^{\text{row}}, D_{i,j}^{\text{col}}$ 之和的平均值 AD

$$AD = \frac{1}{(m+1) \times (n+1)} \sum_{i=1}^{m+1} \sum_{j=1}^{n+1} (D_{i,j}^{\text{row}} + D_{i,j}^{\text{col}}), \quad (4)$$

令阈值 $T=AD$, 如果 $\min(D_{i,j}^k) > T$, 则认为当前块为前景目标, 否则认为是背景模块。

2.2 判断模块的梯度信息

若直接利用前面估计的背景模块进行全局运动参数估计, 则参与计算的模块会很多, 这会在很大程度上影响运动矢量估计速度, 采用对局部模块进行预判的方法来解决这一问题, 即通过给定一个阈值 T' ($T' < T$), 来判断当前背景模块是否具有足够的梯度信息来估计运动矢量。通过实验对比, 选取阈

值 $T' = AD/2$, 即如果满足关系 $\min(D_{i,j}^k) > T'$, 则说明当前模块具有足够的梯度信息, 可以参与计算。综合背景模块的条件, 当某个模块满足关系 $T' < \min(D_{i,j}^k) < T$ 时, 该模块可作为运动估算模块。

2.3 算法流程

算法流程可以分为以下几个步骤: 1) 将各帧图像分成若干子块, 利用块绝对差分的方法提取前景运动目标; 2) 通过设定阈值 T' , 预判提取的背景模块是否具有足够的梯度信息来估计运动矢量; 3) 利

用全搜索策略对选取的模块在帧间进行块匹配, 找出帧间特征点及对应关系; 4) 利用解超定方程的方法估计帧间全局运动参数; 5) 在帧间进行运动补偿, 得到稳像后的视频序列。

与传统方法相比, 算法在前景运动目标提取前, 不需要事先进行图像配准, 且通过设定相关阈值对获得的背景模块进行预判, 进一步去除参与计算的模块数量, 有效地减少了计算量, 提高了运算速度。算法具体框图如图 2 所示。

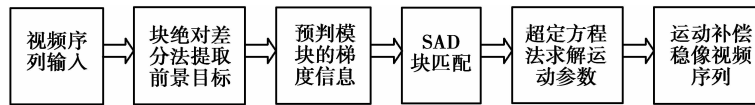


图 2 研究算法框图

3 仿真实验结果

3.1 前景运动目标提取的结果与分析

选用实际交通监控视频进行仿真实验, 其中连续 3 帧视频图像如图 3 所示, 图像分辨率大小为 320

$\times 240$ 。从图中可以看出, 视频图像出现了明显的晃动, 属动态背景, 且图中来往行进的车辆较多, 视频帧间存在较大的局部运动。



图 3 交通监控视频稳像前连续的 3 帧图像

首先将图像分成大小为 16×16 的小块, 利用研究方法对前景目标进行提取, 前景目标模块提取结

果如图 4 所示。

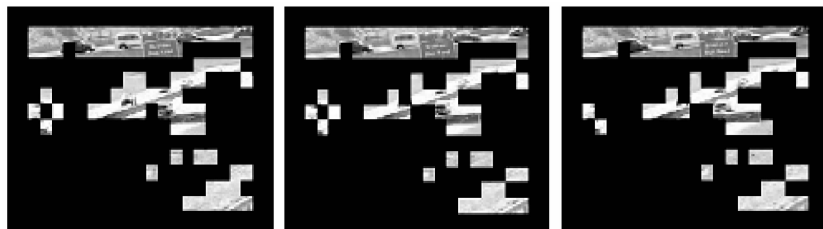


图 4 前景目标提取

3.2 特征块的选取与匹配的实验结果与分析

从图 4 中可以看出, 视频序列中的前景目标(车辆)已基本提取出来。此时, 计算得到的 AD (即阈

值 T) 在 $900 \sim 920$ 之间。如果直接选择图中黑色区域(即背景模块)用于全局运动参数估计, 则参与计算的模块会很多, 这会在很大程度上影响块匹配及

运动矢量估计的计算量。因此,采用预判局部模块的方法,在进行运动矢量估计前,通过设定预判阈值 $T' = AD/2$,对模块运动矢量估计的可靠性进行预

判,剔除可能会导致不准确运动矢量估计的模块,只选定满足条件 $T' < \min(D_{i,j}^t) < T$ 的模块用于运动参数估计,提取的模块如图 5 所示。

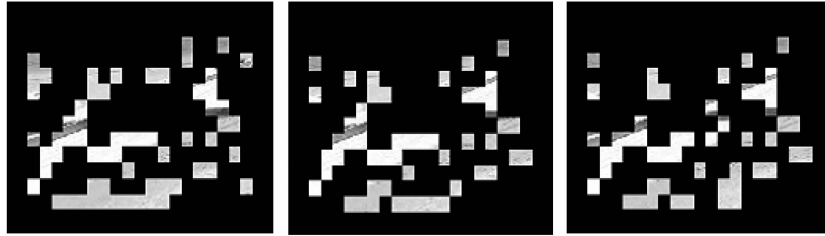


图 5 用于运动参数估算的模块选取结果

从图 4,5 中可以看出,算法成功地分离了视频序列中的前景目标,图 5 中参与运动参数估计的模块数量仅为 50 个左右,约为整幅图像模块数量的 1/6,大大减少了计算量。

为了验证选定的模块对运动矢量估计的有效性,在实际监控视频中选取了连续 30 帧图像进行稳像实验。利用模板匹配算法,获得初始阈值 $T_0 = 100$,再分别对提取前模块、图 4 中的黑色区域模块(背景模块),以及经过预判筛选后的模块进行块匹配,实验结果如表 1 所示。

表 1 选取不同运动估算模块对匹配结果的影响效果

运动估算模块	平均估计 模块数/个	平均匹配 时间/s	补偿后 PSNR /dB
原始模块	300	1.30	24.38
背景模块	220	1.05	25.73
预判筛选后模块	50	0.25	27.38

从表 1 可以看出,与原始模块和前景目标提取后的背景模块相比,经过预判筛选后的模块数量明显减少,从而大大降低了模块匹配的计算量,平均匹配时间仅约为用原始模块进行计算的 1/5,用背景模块进行计算的 1/4,且补偿后的图像帧间 PSNR 值也略高于用原始模块和背景模块进行匹配后的结果。上述结果表明,采用块绝对差分 and 预判局部模块的方法,不仅大大减小了用于运动参数估计的模块数量,减少了运算时间,而且剔除了可能会导致不准确运动矢量估计的模块,提高了模板匹配的准确度。

3.3 稳像的实验结果

利用方法对实际交通监控视频进行稳像处理,图 3 中 3 帧图像补偿后效果如图 6 所示,从图中可以看出,算法成功去除了视频画面的抖动。



图 6 交通监控视频稳像后的连续 3 帧图像

图 7 和图 8 分别是稳像前后的视频序列帧间差分结果,从图中可以看出,稳像前帧间差值结果明显,即相邻帧间具有明显的偏移量,而补偿后的相邻帧基本重合,仅有部分前景运动目标,取得了较好的

稳像效果。

从原视频序列中选取连续 30 帧视频图像,稳像前后的帧间 PSNR 如图 9 所示。



图 7 稳像前帧间差分图



图 8 稳像后帧间差分图

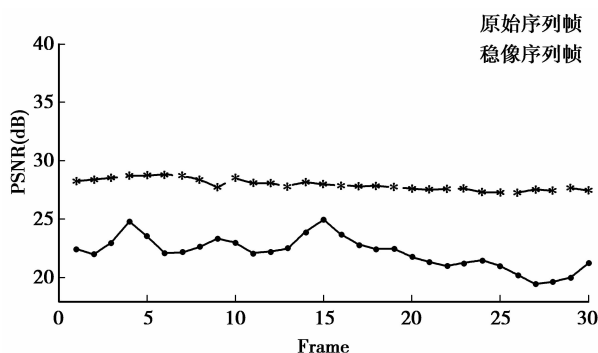


图 9 稳像前后序列帧间 PSNR 示意图

从图中可以看出,稳定后图像序列帧间的 PSNR 值明显高于原序列的 PSNR,平均高出约 7dB,说明算法对具有前景目标的动态视频序列取得了较好的稳像效果。

对上述 30 帧视频序列,分别采用经典的 harris 特征点配准方法、sift 特征点配准方法,以及算法进行稳像处理,图像配准时间和效果如表 2 所示。

从表 2 中可以看出,对于具有前景目标的动态场景视频影像,算法的图像配准时间仅约为上述 2 种算法的 1/6,大大减少了运算时间,且补偿后的帧间 PSNR 值也略高于两种算法,提高了稳像精度。

表 2 采用不同稳像算法的图像配准时间和效果

稳像算法	平均配准时间/s	补偿后 PSNR/dB
Harris 算法	1.73	25.46
Sift 算法	1.90	26.19
研究算法	0.32	27.38

4 结 论

针对动态视频序列中前景运动目标影响全局运动矢量估计精度的问题,提出了一种基于块绝对差分的方法来提取前景目标,事先不需要进行帧间的配准,同时通过设定相关阈值来确定背景模块是否具有足够的梯度信息,大大减少了参与计算的模块数量。通过对实际交通监控视频的稳像结果表明,算法在时间和稳像精度上都取得了令人满意的实际效果。

参考文献:

- [1] Chen H H, Liang C K, Peng Y C, et al. Integration of digital stabilizer with video codec for digital video cameras[J]. IEEE Transactions on Circuits Systems for Video Technology, 2007, 17(7):801-813.

- [2] Mahadevan V, Vasconcelos N. Background subtraction in highly dynamic scenes [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2008, Anchorage, AK, Piscataway: IEEE, 2008:1-6.
- [3] Yu H M, You Y S. Detecting and segmenting multiple moving objects using level-set method [J]. Journal of Zhejiang University: Engineering Science, 2007, 41(3):412-417.
- [4] 侯志强, 韩崇昭. 基于像素灰度归类的背景重构算法 [J]. 软件学报, 2005, 16(9): 1568-1576.
HOU Zhiqiang, HAN Chongzhao. A background reconstruction algorithm based on pixel intensity classification [J]. Journal of Software, 2005, 16(9): 1568-1576.
- [5] 魏志强, 纪筱鹏, 冯伟业. 基于自适应背景图像更新的运动目标检测方法 [J]. 电子学报, 2005, 33(12): 2261-2264.
WEI Zhiqiang, JI Xiaopeng, FENG Weiye. A moving object detection method based on self-adaptive updating of background [J]. ACTA Electronica Sinica, 2005, 33(12):2261-2264.
- [6] 于慧敏, 尤育赛. 基于水平集的多运动目标检测和分割 [J]. 浙江大学学报: 工学版, 2007, 41(3):411-417.
YU Huimin, YOU Yusai. Detecting and segmenting multiple moving objects using level-set method [J]. Journal of Zhejiang University: Engineering Science, 2007, 41(3): 411-417.
- [7] Kim S E, Han J K, Kim J G. An efficient scheme for motion estimation using multireference frames in H. 264/AVC [J]. IEEE Transactions on Multimedia, 2006, 8(3): 457-466.
- [8] Zan J, Ahmad M O, Swamy M N S. Comparison of wavelets for multiresolution motion estimation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3):439-446.
- [9] Patras I, Hendriks E A, Lagendijk R L. Probabilistic confidence measures for block matching motion estimation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(8):988-995.
- [10] 王长军. 基于视频的目标检测与跟踪技术研究 [D]. 浙江: 浙江大学博士学位论文, 2006.
- [11] Eng H L, Toh K A, Kam, A H, et al. An automatic drowning detection surveillance system for challenging outdoor pool environments [C]// Proceedings of 9th IEEE International Conference on Computer Vision, October 13-16, 2003, Nice, France. Piscataway: IEEE, 2003, 1:532-539.
- [12] 吴思, 张勇东, 林守勋, 等. 动态场景视频序列中的前景区域自动提取 [J]. 计算机辅助设计与图形学学报, 2005, 17(2):359-363.
WU Si, ZHANG Yongdong, LIN Shouxun. Automatic extraction of foreground areas in video sequences of a dynamic scene [J]. Journal of Computer-Aided Design & Computer Graphics, 2005, 17(2): 359-363.
- [13] Celebi A, Akbulut O, Orhan O, et al. Truncated graycoded bit-plane matching based motion estimation and its hardware architecture [J]. IEEE Transactions on Consumer Electronics, 2009, 55(3):1530-1536.
- [14] 杨昕梅, 吴钦章, 周进. 仿射不变的运动目标识别方法 [J]. 光学精密工程, 2007, 15(增刊):56-60.
YANG Xin-mei, WU Qin-zhang, ZHOU Jin. Method of motion object recognition based on affine invariant [J]. Optics and Precision Engineering, 2007, 15(增刊):56-60.
- [15] Okuda H, Hashimoto M, Sumi K, et al. Optimum motion estimation algorithm for fast and robust digital image stabilization [J]. IEEE Transactions on Consumer Electronics, 2006, 52(1):276-280.

(编辑 侯 湘)