

文章编号: 0253-2239(2009)11-3031-05

基于稳定矩阵的动态图像运动目标检测

郝志成 吴 川

(中国科学院长春光学精密机械与物理研究所图像室, 吉林 长春 130033)

摘要 为了对视频序列中的运动目标进行快速、准确地提取, 提出了一种自适应背景模型估计方法。利用背景与前景图像在时域中不同的变化特性, 构造图像的稳定矩阵函数, 通过稳定矩阵元素的变化自动区分背景点和前景点, 并对稳定矩阵设置上、下饱和值, 使算法能在短时间内自动感知背景的突变, 从而快速地建立背景图像模型并对其进行实时更新。同时还分析了动态图像序列的配准问题, 选取局部特征模板图像, 采用投影匹配原理简化计算, 快速估计出全局运动矢量。实验证明, 背景估计算法收敛速度快, 只需 10 frame 图像即可建立稳定背景, 对于 500 pixel×200 pixel 动态图像序列, 整个算法时间只需 35 ms, 完全满足工程上 25 frame/s 处理能力的要求。

关键词 图像处理; 运动目标检测; 背景估计; 稳定矩阵; 图像配准

中图分类号 TP39 **文献标识码** A **doi**: 10.3788/AOS20092911.3031

Moving Object Detection from Dynamic Image Sequence Based on Stability Matrix

Hao Zhicheng Wu Chuan

(Image Processing Laboratory, Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun, Jilin 130033, China)

Abstract To extract the moving target in video sequence fast and exactly, a self-adaptive background estimation technique was developed. Because of the different time-variability between background and foreground, a stability matrix was constructed, and according to its changes the background points could be distinguished from the foreground points. By setting two saturation values to the stability matrix, the algorithm could perceive the saltation of background automatically and update the background image quickly. Besides, in order to register the dynamic image sequences some local features were selected as templates, and projection matching technique was used to simplify the calculation and quicken the offset estimation. The experiment results show this algorithm has a high rate of convergence that only 10 frame are needed to get a stable background image, and the whole time is only 35 ms. So it meet the need of a processing ability of 25 frame/s in engineering.

Key words image processing; moving target detection; background estimation; stability matrix; image registration

1 引 言

基于动态图像序列的运动目标检测是应用视觉研究领域的一个重要课题, 捕捉和检测视频序列中的运动目标在视觉监控、交通管理、以及机器人导航等领域都有广泛的应用。由于光照强度变化以及背景本身含有运动等因素, 有效地从图像中分割出前景变得十分困难。常见的运动目标检测算法主要有基于相邻帧的帧间差分法^[1]和基于背景图像与当前

帧差的背景消减法^[2,3]。帧间差分法不受背景图像的限制和影响, 算法只对运动变化区域敏感, 检测有效且稳定, 但是当场景中运动目标比较密集、目标的运动速度较快时, 检测出的目标位置不够准确, 其外接矩形在运动方向上被拉伸。背景消减法速度快, 检测准确, 但是在有些情况下, 静止背景是不能直接获得的, 这就需要通过视频序列的帧间信息估计和恢复背景, 即背景重建。

收稿日期: 2009-02-27; 收到修改稿日期: 2009-04-17

基金项目: 中国科学院青年基金(O76031T070)资助课题。

作者简介: 郝志成(1978—), 男, 博士, 助理研究员, 主要从事图像处理、目标跟踪和模式识别等方面的研究。

E-mail: hzc972513@tom.com

1999年,Stauffer等^[4]提出的基于混合高斯模型的背景重建算法,为每个像素构建多维混合高斯分布模型,通过自适应阈值把图像中的像素点区分为背景区域和变化区域,能够较好地解决背景多模态的问题,因而成为运动检测领域的经典算法,但是该算法也存在背景模型建立时间长和检测效果不理想的缺点。于是, Lee^[5]将模型被匹配次数计数器的概念引入到模型的学习率中,提高了学习效率,但是稳定性差,并且没有涉及模型匹配问题。Zivkovic^[6]提出的一种自适应选择高斯模型的数目的方法,在一定程度上提高了运算速度,但高斯背景模型还是存在模型收敛慢,参数计算复杂等问题。

另外, Kim等^[7]根据对像素的统计,通过建立码本(Codebook)来表示压缩背景,但是该法的检测效果对参数的依赖性较强。Elgammal^[8]和 Han等^[9]用基于核密度估计的非参数方法描述任意分布的背景模型,但是该方法算法复杂,为了降低计算量,典型的核函数值需要离线计算。

有鉴于此,本文提出一种基于图像稳定矩阵的背景重建和更新方法,可快速得到稳定的背景图像,并在此基础上通过基于特征匹配的序列图像配准技术实现动态图像中运动目标的检测。

2 背景估计

运动目标检测的关键是背景图像的描述模型即背景模型,它是背景消减法分割前景目标的基础。在 t 时刻,图像 $f_t(x, y)$ 中的像素点按照运动状态可分为两类:背景点和前景运动目标点。由背景点组成的图像称为背景图像 $B_t(x, y)$,由前景运动目标点组成的图像称为前景图像 $T_t(x, y)$ 。当视场中有多个运动目标存在时,有

$$\begin{aligned} f_t(x, y) &= B_t(x, y) + T_t(x, y), \\ T_t(x, y) &= \bigcup_{i=0 \leq N_s} T_i(x, y), \end{aligned} \quad (1)$$

其中 $T_i(x, y)$ 为各运动前景图像, N_s 为目标数目。

由于运动目标和背景只有遮挡关系,其灰度特征不具有叠加性、因果性和连续性,运动目标可以看作是对背景图像的随机扰动。因此在跟踪过程中,对于背景图像而言,在时域上主要有两种变化方式:1)无目标遮挡时,其灰度特征变化缓变且连续;2)有目标遮挡时,其灰度特征将发生阶跃式跳变。根据这种特性,本文采用稳定矩阵来描述背景图像模型。

2.1 稳定矩阵

定义两幅图像 f_1 和 f_2 的距离为

$$D_{f_1, f_2} = |f_1(x, y) - f_2(x, y)|, \quad (2)$$

对于 $M \times N$ 的图像序列 $f_t(x, y)$,其稳定矩阵 $S(x, y)$ 同样为 $M \times N$,每个矩阵元素表示序列图像中对应位置像素点在时域上的稳定程度。定义 t 时刻图像 (x, y) 点对应的稳定矩阵元素为 $S_t(x, y)$,它的值随着该像素点灰度值的变化而改变

$$S_t(x, y) = \begin{cases} S_{t-1}(x, y) + 1, & D_{f_t, B_{t-1}} \geq B_{\text{TH}} \\ S_{t-1}(x, y) - 1, & D_{f_t, B_{t-1}} < B_{\text{TH}} \end{cases}, \quad (3)$$

其中 B_{TH} 为背景变化阈值。 $S_t(x, y)$ 是像素点在时域上稳定程度的一个度量, $S_t(x, y)$ 越大,说明该位置像素点越不稳定,反之则越稳定。

2.2 背景更新策略

一般来说,背景模型的更新策略是背景模型设计中最关键的技术。在实际场景中,背景图像不可避免地会发生突变,这种突变可能是由光照变化等因素引起的背景本身颜色的变化,也可能是背景区域的变化,如前景和背景的相互转化等,这时候需要背景模型自动地感知这种突变并更新背景。通过对稳定矩阵设置上、下饱和值 S_u, S_d 可实现这个功能

$$S_t(x, y) = \begin{cases} S_d, & S_t(x, y) > S_u \text{ or } S_t(x, y) < S_d \\ S_t(x, y), & \text{otherwise} \end{cases}, \quad (4)$$

可见 $S_t(x, y) > S_u$ 的区域是稳定矩阵发生“翻转”的区域,背景图像发生突变,而且其累计时间超过 S_u ,背景模型认为当前图像有别于背景但其对应像素点已达到稳定,说明以前的背景估计有误,应马上更新背景。而 $S_t(x, y) < S_d$ 的区域是背景稳定变化的区域。在此区域内,图像特征变化缓慢,并且当前图像特征和以前估计背景特征差别小于 B_{TH} ,说明当前时刻背景估计正确,且没被运动目标遮挡,可以进行背景更新(一般情况下取 $S_d = 0, S_u$ 则根据实际情况具体确定)。定义 t 时刻背景更新系数为

$$\alpha_t(x, y) = \begin{cases} 1, & S_t(x, y) = S_d \\ 0, & S_t(x, y) > S_d \end{cases}, \quad (5)$$

则背景图像可依据下式进行更新

$$\begin{aligned} B_t(x, y) &= F[\alpha_t(x, y)]f_t(x, y) + \\ &\quad \{E - F[\alpha_t(x, y)]\}B_{t-1}(x, y), \end{aligned} \quad (6)$$

其中 E 为与图像大小相同的一个全1矩阵, F 代表图像预处理操作,如形态学开运算和中值滤波等,若不进行任何操作,可令 $F(\alpha_t) = \alpha_t$ 。因此,背景的提取和更新可采用如下几个步骤来实现

1)在 $t = 0$ 时刻,令 $B_0(x, y) = f_0(x, y), S_0(x, y) = 0$;

2) 在 t 时刻, 根据(3)式, (4)式计算稳定矩阵值并对其进行更新;

3) 在 t 时刻, 根据(5)式, (6)式实时更新背景图像。

2.3 稳定矩阵的优化

在对背景模型更新时应当注意如下两个原则:

1) 背景模型对背景变化的响应速度要足够快。如果背景模型不能迅速跟上实际背景的变化, 检测结果中就会出现大范围的噪声或不合理的长时间静止前景物体。

2) 背景模型对运动目标要有较强的抗干扰能力。因为在背景模型的更新过程中, 背景模型的每个点都受到了一个灰度序列的“训练”, 背景的这种“训练”是我们所希望的, 而运动目标的“训练”则是不希望看到的。特别是当运动物体尺度较大或运动较慢时, 这种长时间的“训练”可能会引起错误的检测结果。

这两个原则是有矛盾的, 需要折中处理。本文的方法是用前一帧的跟踪结果来指导背景模型更新, 对稳定矩阵的上饱和值 $S_{\text{上}}$ 进行局部优化: 赋予背景点稳定矩阵较小的 $S_{\text{上}}$, 而赋予运动目标区域的稳定矩阵较大的 $S_{\text{上}}$, 其公式表述如下

$$S_{\text{上}}(x, y) = \begin{cases} S_{\text{上T}}, & (x, y) \in T_i(x, y) \\ S_{\text{上B}}, & (x, y) \in B_i(x, y) \end{cases}, \quad (7)$$

其中 $S_{\text{上T}} > S_{\text{上B}}$ 。经过这种改进可以提高运动目标点的耐“训练”性, 在保护背景模型不受运动目标影响的同时迅速响应背景的变化, 这实际上是用运动目标跟踪的结果来指导更新。

3 动态图像配准

当摄像机静止不动时, 相邻帧间的背景基本一致, 可以直接采用上一节的方法获取背景图像, 检测出运动目标。但如果摄像机是运动的, 就会使相邻两帧图像发生了一个整体偏移 ($\Delta X, \Delta Y$)。这时, 在两帧图像的重叠区域, 可依据(3)式, (4)式来计算稳定矩阵 $S_i(x, y)$, 而令非重叠区域的 $S_i(x, y) = 0$, 即将新进入图像部分的稳定矩阵初始化为 0; 背景模型的估计和更新方法同(6)式, 问题的关键在于偏移量 ($\Delta X, \Delta Y$) 的计算。

图像配准作为运动目标检测的一项关键技术, 必须兼顾速度与精度。块匹配法^[6]因其检测精度高、实现简单, 成为最常用的方法之一。但是传统块匹配法采用逐点像元匹配^[10], 运算量较大, 而且若模板图像选取不当, 使得检测结果易受到视场中目

标的运动以及景物变化的影响, 降低匹配精度。故而本文提出一种新的图像配准算法——局部投影匹配法, 它主要解决两个问题: 1) 选取适当的模板块图像; 2) 利用投影原理将二维像元匹配转化为一维曲线匹配, 提高算法速度。

3.1 灰度投影原理

对每帧输入的图像将二维图像信息分别在 x, y 方向上映射成独立的一维波形, 得到了反映图像灰度分布特点的各行、列的投影曲线。(8)式是以列投影计算为例说明图像映射的方法。

$$f(j) = \sum_i f(i, j), \quad n_{\text{mean}} = \left[\sum_i f(i, j) \right] / N$$

$$f'(j) = f(j) - n_{\text{mean}}, \quad (8)$$

其中 $f(j)$ 为图像 f 第 j 列的灰度值和, $f(i, j)$ 为图像 (i, j) 位置处的像素灰度值, N 是列数, $f'(j)$ 是第 j 列修正后的投影值。

3.2 模板块的选取

模板块图像的选取要考虑以下几个因素^[11]: 1) 在图像中尽可能均匀分布; 2) 灰度变化较大, 即含有丰富的细节; 3) 在其周围区域中具有一定独特性^[12]。在跟踪过程中, 目标一般都集中在视场中心, 其余背景区域表现为由摄像机运动造成的全局运动, 而目标的运动表现为全局运动与局部运动的迭加。因此图像的中心部分参与配准运算不但会增加计算时间, 而且会降低检测精度。同时由于摄像机的运动, 图像的边缘会有退出视场和新进入视场的部分, 因此应在图像四周留出适当区域作为补偿区域。

通常, 选取模板图像时要对整幅图像进行区域分割, 可以利用区域能量比和区域清晰比的概念来度量和划分区域^[13], 在各子区域内选取不同的模板图像(如图 1 所示)。

考虑到算法复杂度和计算量的因素, 文中采用了一种更为简单的划分方法, 模板块选取方法如下:

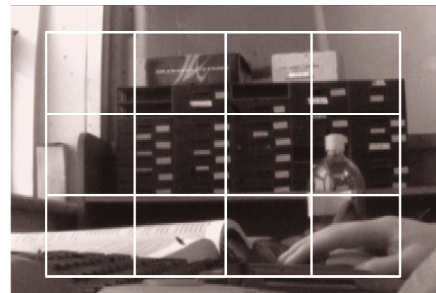


图 1 局部区域划分示意图

Fig. 1 Sketch of local area plotting

首先除掉图像每个边缘的 5%，将剩下的图像划分为 $m \times n$ 个互不重叠的区域 W ，然后将每个 W 划分为 $r \times s$ 个互不重叠的图像块 $w_{r,s}$ ，不考虑中心目标区域，计算每个图像块 $w_{r,s}$ 的灰度方差，在每个 W 中选取 a 个方差最大且超过阈值 B_{TH2} 的图像块 $w_{r,s}$ ，作为匹配的模板块。其中 B_{TH2} 可以预先设定，也可以自适应选取，以保证总能找到这样的模板块。

3.3 局部投影匹配

尽管动态图像序列帧间存在几何变化和光照变化，但是相邻帧间背景部分对应图像块的行、列投影曲线波形基本相同，利用这一原理可将二维像元匹配转化为一维曲线匹配。具体算法步骤如下：

1) 计算匹配块中行或列灰度投影的平均值，再计算它同模板块行或列投影均值的误差，如果误差值大于 B_{TH3} ，直接滤除该匹配位置；

2) 如果误差值小于 B_{TH3} ，分别计算匹配块与模板块各行、列的灰度投影误差 E_x, E_y 。误差定义为

$$E_x = \sum_i [|C'(i) - w'(i)|]$$

$$E_y = \sum_j [|C'(j) - w'(j)|], \quad (9)$$

其中 C 代表匹配块， w 代表模板。

3) 求出所有匹配位置中 $E_x + E_y$ 的最小值，并把 $E_x + E_y$ 最小的位置作为最佳匹配块。

4) 计算该模板块的局部运动矢量 $\Delta x, \Delta y$

$$\Delta x = x - x_0, \quad \Delta y = y - y_0, \quad (10)$$

其中 (x, y) 为最佳匹配块坐标， (x_0, y_0) 为模板块坐标。

3.4 计算全局运动矢量

尽管采用了局部模板匹配，但仍有个别模板块的匹配结果会受到自然因素的干扰（如树枝的摆动），与全局运动矢量有很大的偏差，必须把这些结果排除在对全局运动矢量的估计之外。考虑到干扰因素在各个区域中没有重复性，因此在得到的各局部运动矢量中选择出现频率最高的作为全局运动矢量。即作出各模板块局部运动矢量 $\Delta x, \Delta y$ 的二维直方图，找到图中峰值位置，该位置的 $\Delta x, \Delta y$ 即为全局运动矢量 $\Delta X, \Delta Y$ 。

除了上述的方法，利用特征点匹配^[14]来估计全局运动矢量也是常用方法之一。通过双向最大相关系数匹配的方法提取出初始特征点对，用随机采样符合法(Random sample consensus, RANSAC)来剔除伪特征点对，实现特征点对的精确匹配。然后，利用一定的图像空间变换关系^[15]，对帧间图像进行准确配准。

4 实验结果及讨论

算法应用在小电视跟踪器上，并进行了现场实验。实验对象为一段路面交通视频，实验平台是以 TMS320C6416DSP 为核心处理器的跟踪器，其工作频率为 1 GHz。设定算法中参数： $S_d = 0$ ， $S_{uT} = 20$ ， $S_{uB} = 10$ 。表 1 和图 2 为实验结果。

表 1 运行时间统计表

Table 1 Statistics of run time

steps	Background updating	Background subtraction	Area segmentation	Total time
Run time /ms	9.31~10.26	2.60~2.94	5.42~6.28	17.33~19.48
Average time /ms	9.77	2.75	5.85	18.37

从表 1 中的数据可以看出，利用稳定矩阵背景模型来检测运动目标平均时间不超过 20 ms，如果算上系统采集整幅图像的时间 16.4 ms，整个算法时间大约 35 ms，可以满足实时系统 25 frame/s 处理能力的要求。

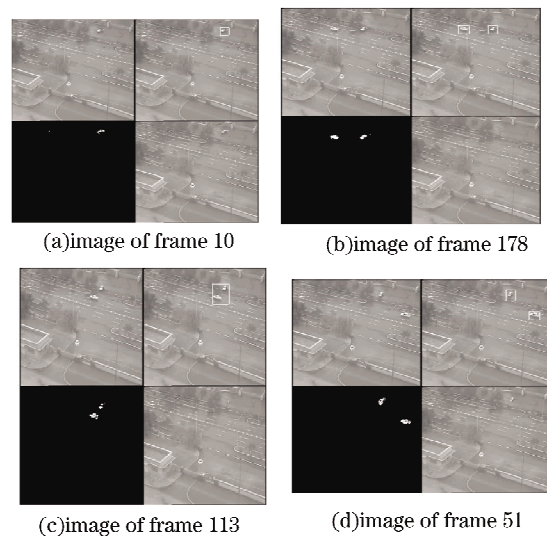


图 2 运动目标检测结果

Fig. 2 Results of moving targets detecting

图 2 是从算法检测结果视频中截取的几幅图片。(a)为原始图像序列，(b)为对应的重建背景图像，(c)为差分二值图像，(d)为运动目标检测结果图像。

场景中的树木由于风扰产生轻微地随机摆动，为了避免把这些小幅摆动的物体错认为运动目标，可以通过调节参数 S_{uT} 和 S_{uB} 来加快和减慢背景图像的更新速度，以增加背景模型对局部干扰的适应性。同时，对于减背景后的二值图像，算法中加入形态滤波进一步滤除细微运动的干扰，以保证运动目标检测的准确性。

5 结 论

提出了基于稳定矩阵的背景模型估计方法。由于采用了适当的稳定矩阵函数和背景更新策略,使得算法只对背景像素点敏感,而排斥前景的运动物体像素;并在背景发生突变时,能够在尽可能短的时间自动感知。同时为满足对动态图像序列进行实时处理的需求,提出利用局部投影匹配的图像配准方法。在合理选择局部匹配模板的基础上,采用投影相关匹配,兼顾了速度与精度的要求。实验证明,本文算法背景估计收敛速度快、实时性能好,只需 10 帧图像就能的建立起背景模型,并实时更新背景图像;最终能有效地检测出场景中的运动目标。同时,由于稳定矩阵构造简单,易实现,使算法具有每秒 25 帧 $500 \text{ pixel} \times 200 \text{ pixel}$ 图像的处理能力,完全满足工程中的要求。并且在背景较为复杂的情况下,算法也有很好的稳健性,能对运动目标稳定的识别与跟踪。

参 考 文 献

- 1 A. T. Al, E. L. Dagless. Alternative practical methods for moving object detection [C]. International Conference on Image Processing and its Applications, 1992, 77~80
- 2 I. Haritaoglu, D. Harwood, L. R. Davis. Real-time surveillance of people and their activities [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, **22** (8): 809~830
- 3 K. Toyama, J. Krumm, B. Brumitt. Principles and practice of background maintenance [C]. International Conference on Computer Vision, 1999, 255~261
- 4 C. Stauffer, W. Grimson. Adaptive background mixture models for real-time tracking [C]. IEEE Conference on Computer Vision and Pattern Recognition, 1999, 246~252
- 5 D. S. Lee. Effective gaussian mixture learning for video background subtraction [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27** (5): 827~832
- 6 Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction [C]. International Conference on Pattern Recognition, 2004, **2**: 28~31
- 7 K. Kim, T. II. Chalidabhongse, D. Harwood *et al.*. Real-time foreground-background segmentation using codebook model [J]. *Real-time Imaging*, 2005, **11**(3): 172~185
- 8 Elgammal. Background and foreground modeling using non-parametric kernel density estimation for video surveillance [C]. IEEE, 2002, **90**(7): 1151~1163
- 9 B. Han, D. Comaniciu, L. Davis. Sequential kernel density approximation through mode propagation; applications to background modeling [C]. Asian Conference on Computer Vision, 2004, 818~823
- 10 Chen Hua, Ye Dong, Chen Gang *et al.*. Digital image correlation search method based on genetic algorithm [J]. *Optics and Precision Engineering*, 2007, **15**(10): 1633~1637
陈华,叶东,陈刚等. 遗传算法的数字图像相关搜索法 [J]. *光学精密工程*, 2007, **15**(10): 1633~1637
- 11 Wang Yifeng, Zhang Hailian, Li Canwen *et al.*. A survey of registration techniques in multi-sensor data fusion [J]. *Infrared and Laser Engineering*, 1998, **27**(1): 38~41
王忆峰,张海联,李灿文等. 多传感器数据融合中的配准技术 [J]. *红外与激光工程*, 1998, **27**(1): 38~41
- 12 Jia Guimin, Wang Xiangjun, Zhang Shihai. Target tracking algorithm based on adaptive template update in complex background [J]. *Acta Optica Sinica*, 2009, **29**(3): 659~663
贾桂敏,王向军,张世海. 复杂背景下基于自适应模板更新的目标跟踪算法 [J]. *光学学报*, 2009, **29**(3): 659~663
- 13 Ye Chuanqi, Miao Qiguang, Wang Baoshu. An image fusion algorithm using region segmentation and contourlet transform [J]. *Acta Optica Sinica*, 2008, **28**(3): 447~453
叶传奇,苗启广,王宝树. 基于区域分割和 Counterlet 变换的图像融合算法 [J]. *光学学报*, 2008, **28**(3): 447~453
- 14 Liu Guixi, Liu Dongmei, Liu Fengpeng *et al.*. A robust image registration algorithm based on feature points matching [J]. *Acta Optica Sinica*, 2008, **28**(3): 454~461
刘贵喜,刘冬梅,刘凤鹏等. 一种稳健的特征点配准算法 [J]. *光学学报*, 2008, **28**(3): 454~461
- 15 Sun Tao, Wang Xiukun, Shao Gang *et al.*. Affine registration of 2D point pattern image [J]. *J. Computer-Aided Design & Computer Graphics*, 2005, **17**(7): 1497~1503
孙焘,王秀坤,邵刚等. 二维点模式图像的仿射变换配准 [J]. *计算机辅助设计与图形学学报*, 2005, **17**(7): 1497~1503