

## 动态场景中基于绝对差值和的前景检测算法

刘 敏, 赵丹丹, 武明虎, 王 娟

(湖北工业大学 太阳能高效利用湖北省协同创新中心, 武汉 430068)

**摘 要:** 针对动态场景中前景检测误报率较高的问题, 结合视频序列中背景图像存在时域高冗余的特点, 提出一种改进的前景检测算法。基于滑动窗口机制组建微视频单元(MVBE), 采用时空绝对差值和计算 MVBE 的冗余性, 对冗余性较低的 MVBE 采用动态纹理(DT)进行建模。利用批量主成分分析方法估计 DT 参数, 并引入观测性测量方法计算 MVBE 中每个像素点的显著度以实现前景检测。实验结果表明, 该算法相比 KDE, GMM 等检测算法对场景的动态变化具有更好的抑制效果, 并且取得了较鲁棒的前景检测结果。

**关键词:** 绝对差值和; 微视频元; 动态纹理; 前景; 动态冗余性

**中文引用格式:** 刘 敏, 赵丹丹, 武明虎, 等. 动态场景中基于绝对差值和的前景检测算法[J]. 计算机工程, 2017, 43(7): 250-252, 260.

**英文引用格式:** Liu Min, Zhao Dandan, Wu Minghu, et al. Foreground Detection Algorithm Based on Sum of Absolute Difference in Dynamic Scenes[J]. Computer Engineering, 2017, 43(7): 250-252, 260.

## Foreground Detection Algorithm Based on Sum of Absolute Difference in Dynamic Scenes

LIU Min, ZHAO Dandan, WU Minghu, WANG Juan

(Hubei Collaborative Innovation Center for High-efficiency Utilization of Solar Energy,  
Hubei University of Technology, Wuhan 430068, China)

**[Abstract]** In order to reduce the high false alarm rate on foreground detection in dynamic scenes, this paper proposes an improved foreground detection algorithm, considering the high temporal redundancy of background images in the video sequence. Firstly, the Micro Video Block Element(MVBE) groups are formed by the sliding window technique and the redundancy values of each MVBE are computed by Spatio-Temporal Sum of Absolute Difference(STSAD) method. Secondly, The MVBE groups of low redundancy are modeled by Dynamic Texture(DT), and the DT parameters are estimated by Batch-Principal Component Analysis(Batch-PCA). Thirdly, the saliency value of every pixel in the MVBE is calculated by the observability measure method to realize foreground detection. Experimental results show, compared with the KDE, GMM algorithm, the algorithm has better inhibitory effect in dynamic scenes and a robust foreground detection result is obtained.

**[Key words]** Sum of Absolute Difference(SAD); Micro Video Element(MVE); Dynamic Texture(DT); foreground; dynamic redundancy

**DOI:** 10.3969/j.issn.1000-3428.2017.07.042

### 0 概述

前景检测在智能视频监控<sup>[1-2]</sup>、视频压缩<sup>[3-4]</sup>、行为分析和理解<sup>[5]</sup>等领域具有广泛的应用。对于固定照相机拍摄的相对静态的场景, 传统的背景建模方法具有较高的前景检测率, 如单高斯(Single Gaussian,

SG)<sup>[6]</sup>、混合高斯(Mixture of Gaussian, MOG)<sup>[7]</sup>、非参数核密度估计(Non Parametric Kernel Density Estimation, KDE)<sup>[8]</sup>等。然而, 在一些动态场景中存在背景元素运动(如下雪, 下雨, 狂风下持续摇摆的树枝、喷泉、波浪等), 照相机运动或自运动(ego-motion), MOG 和 KDE 检测性能较差。

**基金项目:** 国家自然科学基金面上项目(61471162); 湖北省科技支撑计划项目(2015BAA115)。

**作者简介:** 刘 敏(1979—), 女, 讲师、博士, 主研方向为计算机视觉、智能视频分析; 赵丹丹, 硕士研究生; 武明虎, 教授、博士; 王 娟, 讲师、博士。

**收稿日期:** 2016-11-29 **修回日期:** 2016-12-29 **E-mail:** lmsnowlove@aliyun.com

基于动态纹理<sup>[9]</sup>的方法由于考虑到场景的动态性变化具有较强的鲁棒性,因此近年来在前景检测方面应用较广泛。文献[10]基于生物视觉 Center-Surround 机制,采用 DT 建模方法,通过计算 KL 距离来实现显著性像素点的检测,具有较高的准确率,但计算复杂度较高,并且实时性有待改善。SO (Sustain Observability) 算法<sup>[11]</sup>通过引入控制理论中可观测性测量方法来检测动态场景中的显著性运动,但对于存在 ego-motion 的动态场景,误报率较高。上述方法均采用批量主成分分析 (Batch-Principal Component Analysis, Batch-PCA) 进行 DT 参数的估计, Batch-PCA 的复杂度主要取决于数据矩阵  $D \in \mathbb{R}^{m \times n}$  的奇异值分解,其计算复杂度为  $O(nm^2 + n^3)$ 。Batch-PCA 基于分块的思想,采用 DT 局部建模的方法,数据矩阵  $D$  较小,但模型个数较多,平均每帧的计算开销较大。而采用 DT 全局建模的方法,数据矩阵  $D$  为高维矩阵,奇异值分解时计算复杂度较高。

在此基础上,本文根据视频序列中一些背景块在时域上的高冗余特点,假设像素信息的帧间变化是线性的。基于滑动窗口机制与分块的思想,设滑动窗口长度为  $\tau$  帧,将输入的连续  $\tau$  帧中相同像素位置的块组建为一个微视频基元 (Micro Video Block Element, MVBE),采用时空绝对差值和 (Spatio-Temporal Sum of Absolute Difference, STSAD) 对一个滑动窗口内的所有 MVBE 进行过滤处理,仅对那些具有较小动态冗余性<sup>[12]</sup>的 MVBE 进行 DT 建模,并采用观测测量的方法来判定像素点的背景或前景属性。

## 1 动态冗余性计算

将输入的视频序列通过长度为  $\tau$  帧的滑动窗口,在每一时刻滑动窗口内相同像素位置的视频块作为一个微基元,每个 MVBE 由  $\tau$  帧组成,其中  $\tau = 2k + 1$ ,包括  $k$  个历史块、当前帧中的视频块、 $k$  个未来块。若视频块大小选取为  $N \times N$  个像素点,左上角像素坐标位置为  $(u, v)$ ,则每个 MVBE 中第  $i$  帧中的视频块可表示为  $(u, v, i)$ ,其中  $i \in [1, \tau]$ 。

本文结合视频序列中一些背景块在时域上存在高冗余,而运动具有持久性的特点,对文献[13]提出的绝对差值和 (Sum of Absolute Difference, SAD) 方法在时域上进行扩展,采用  $k$  个历史块和  $k$  个未来块的像素信息来评估第  $i$  帧中块  $(u, v, i)$  的动态冗余性大小。若  $f(x, y, i)$  为 MVBE 中第  $i$  帧图像块中的像素信息,  $M(u, v, i)$  为 MVBE 中的动态冗余性大小,采用 STSAD 方法计算  $M(u, v, i)$ ,具体公式如下:

$$M(u, v, i) = \frac{1}{\tau} \frac{1}{N^2} \sum_{n=1}^{\tau-1} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} |f(u+j, v+k, n) - f(u+j, v+k, n+1)| \quad (1)$$

$$S(u, v, i) = \begin{cases} 1, & M(u, v) > \sigma_s \\ 0, & M(u, v) \leq \sigma_s \end{cases} \quad (2)$$

对于同一个滑动窗口内的 MVBE,动态冗余性最小的记作  $M(u, v, i)_{\min}$ ,最大的记作  $M(u, v, i)_{\max}$ ,均值记作  $M(u, v, i)_{\text{median}}$ 。对于背景动态性较小的场景,  $\sigma_s = (M(u, v, i)_{\min} + M(u, v, i)_{\text{median}})/2$ 。当场景中存在照相机运动或 ego-motion 时,  $\sigma_s = (M(u, v, i)_{\max} + M(u, v, i)_{\text{median}})/2$ 。 $\sigma_s$  用来判定 MVBE 的动态冗余性大小。若  $M(u, v, i) < \sigma_s$ ,  $S(u, v, i) = 0$ ,则左上角像素点为  $(u, v)$  的 MVBE 动态冗余度大, MVBE 对应的空间区域隶属于背景。若  $M(u, v) > \sigma_s$ ,  $S(u, v, i) = 1$ ,则 MVBE 中可能包含运动物体。本文方法仅对  $S(u, v, i) = 1$  的 MVBE 进行 DT 建模,这样可减少 DT 模型个数,降低每帧图像处理的计算开销。

## 2 动态纹理

动态纹理<sup>[9]</sup>作为一种自回归的状态空间模型,假设视频序列在任一时刻  $t$  的输出代表  $t$  时刻的帧图像,  $y(t) \in \mathbb{R}^{m \times 1}$  是  $t$  时刻的图像向量,  $x(t) \in \mathbb{R}^{n \times 1}$  是  $t$  时刻线性系统的内部状态向量,DT 模型通过构建线性系统将观测到的图像空间与内部状态空间关联起来,表示为:

$$\begin{aligned} x(t) &= Ax(t-1) + v_t \\ y(t) &= Cx(t) + w_t \end{aligned} \quad (3)$$

其中,  $m$  是一帧图像或视频块中的像素个数 (本文中  $m = N^2$ );  $n$  是隐状态空间维数;  $v_t$  是均值为 0、分布为  $N(0, Q)$  的高斯状态;  $w_t$  是可观测的、分布为  $N(0, R)$  的高斯白噪声。

$\Theta = \{A, C, Q, R\}$  是 DT 参数,其中,  $A \in \mathbb{R}^{n \times n}$  是转移矩阵;  $C \in \mathbb{R}^{m \times n}$  是观测矩阵。本文采用文献[9]提出的 Batch-PCA 方法来估计参数  $\Theta = \{A, C, Q, R\}$ 。设一个 MVBE 的数据矩阵为  $Y_1^T = [y(1), y(2), \dots, y(\tau)] \in \mathbb{R}^{m \times \tau}$ , 它的零均值矩阵为  $\bar{Y}_1^T$ , 则  $\bar{Y}_1^T = USV^T$ ,  $U = [u_1, u_2, \dots, u_m] \in \mathbb{R}^{m \times m}$ ,  $C = [u_1, u_2, \dots, u_n]$  是转移矩阵。取矩阵  $U$  的前  $n$  列向量, 得到  $X_1^T = C^T \bar{Y}_1^T$ ,  $A = X_2^T (X_1^T)^+$ ,  $(M)^+$  是矩阵  $M$  的广义逆矩阵。

## 3 观测性测量

由文献[11]可知,前景一般是指规律性的、容易被观测的或者可预测的显著运动刚性物体。而背景相对前景,其动态变化宏观上是无规律、不容易被观测的,且具有非显著性。SO 算法中通过计算系统的可观测性来实现前景运动的显著性检测。观测性大表明该像素点属于前景运动,观测性小表明该像素点属于背景运动。

本文采用文献[14]中改进的观测性测量方法来实现前景检测。若  $m_i$  是观测矩阵  $C$  和状态矩阵  $A$  的乘积矩阵的第  $i$  个列向量,则总观测向量为  $m_o$ :

$$m_o = C \times A = \sum_{i=1}^n |m_i| \quad (4)$$

在  $m_o$  正规化处理后,根据每一个像素点对应的观测值测量大小,可以将该点判定为前景或背景。

#### 4 实验与结果分析

实验在 3.4 GHz 和 4 GB RAM 的 PC 机, Windows 操作系统的 Matlab7.0 环境下进行仿真。测试视频选自数据集<sup>[15]</sup>中的一些动态场景视频。仿真时参数设置为: $\tau = 15, n = 5, N = 80$ 。

图 1 给出了本文算法与 GMM<sup>[7]</sup>, KDE<sup>[8]</sup> 和 SO<sup>[11]</sup> 算法对“骑车者”进行检测的前景图像。“骑车者”场景中存在 ego-motion,由此可知,传统的减背景算法 GMM, KDE 对于存在 ego-motion 的场景前景检测性能较差,SO 算法对由于 ego-motion 造成的背景变化抑制性较好,而本文算法由于对高动态冗余性的 MVBE 进行了过滤筛选处理,具有较高的前景准确率和较低的误报率,综合鲁棒性能相对较好。

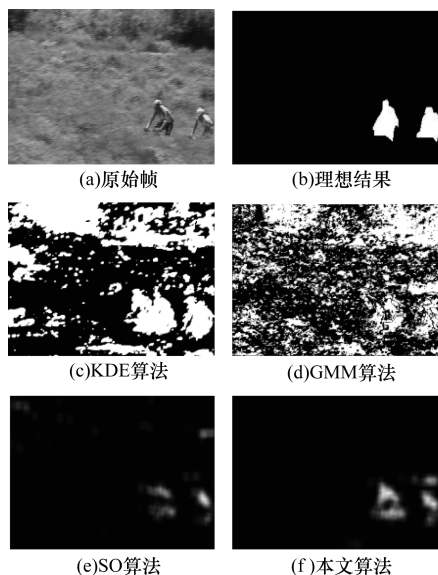


图 1 前景检测结果

为客观评估本文算法的性能,采用分类正确率 (Percentage of Correct Classification, PCC) 指标进行评价: $PCC = (TP + TN) / (TP + FP + FN) \times 100\%$ , 其中  $TP, TN, FP$  和  $FN$  定义如下:

- 1)  $TP$  表示检测正确的前景像素数;
- 2)  $FP$  表示检测为前景的背景像素数;
- 3)  $TN$  表示检测正确的背景像素数;
- 4)  $FN$  表示错误检测为前景的背景像素数。

表 1 给出了 KDE, GMM, SO 和本文算法对“瓶

子”场景中第 12 帧图像进行检测的 PCC 指标参数,“瓶子”场景中存在背景元素起伏波动的水面,其动态变化性较大。由表 1 中数据可知,本文算法、SO 算法、KDE 算法和 GMM 算法的 PCC 数值分别为 95.3%, 95.2%, 91.7% 和 87.9%。本文算法对背景元素的动态变化抑制性较强,检测准确率较 KDE 和 GMM 算法有明显提高,比 SO 算法略高。

表 1 各算法的检测性能对比

算法	TP	TN	FP	FN	PCC/%
KDE 算法	1 722	60 741	5 276	357	91.7
GMM 算法	896	58 959	7 058	1 183	87.9
SO 算法	2 054	62 786	3 231	25	95.2
本文算法	2 031	62 974	3 043	48	95.3

本文算法根据一些背景块在时域上存在高冗余的特点,采用 SAD 算法进行动态筛选,减少了 DT 局部建模的模型数量,降低了单帧图像处理的时间开销,算法运行效率明显提高。在本文的实验环境下,对于 320 像素  $\times$  240 像素的视频图像,平均单帧处理时间约 50 ms,而 SO 算法处理同样大小的视频图像,平均每帧处理时间约 370 ms。

#### 5 结束语

本文根据背景图像存在时域高冗余的特点,采用滑动窗口机制组建 MVBE,基于 STSAD 方法对 MVBE 进行动态筛选,减少 DT 局部建模个数,从而降低单帧图像处理的时间开销。另外,通过 DT 建模构建动态线性系统,将前景运动与线性系统的可观测性相关联,并采用线性系统中观测性测量方法实现前景像素点的检测。实验结果表明,本文算法对由于背景元素变化、照相机移动或 ego-motion 造成的场景动态性具有较强的抑制效果,前景检测准确率较高,实时性较好。但由于该算法需要对视频帧进行分块,块的大小选择对前景检测性能有一定影响,因此如何自适应地选择合适的视频帧块是下一步研究的内容。

#### 参考文献

- [1] Shah M, Javed O, Shafique K. Automated Surveillance in Realistic Scenarios [J]. IEEE Multi-media, 2007, 14(1):30-39.
- [2] 刘云鹏,张三元,王仁芳,等. 视觉注意模型的道路监控视频关键帧提取[J]. 中国图象图形学报, 2013, 18(8):933-943.
- [3] Guo C, Zhang L. A Novel Multi Resolution Spatiotemporal Saliency Detection Model and Its Applications in Image and Video Compression [J]. IEEE Transactions on Image Processing, 2010, 19(1):185-198.

(下转第 260 页)

## 参考文献

- [1] Li Zhaohui, Wang Xiaorui, Zhang Jianqi. Novel Spatiotemporal Filter for Dim Point Targets Detection in Infrared Image Sequences[J]. Journal of Electrical and Computer Engineering, 2015, 12(2): 1201-1209.
- [2] 凌强, 黄树彩, 吴潇. 基于核各向异性扩散的红外小目标检测[J]. 强激光与粒子束, 2015, 27(1): 93-98.
- [3] Ahmadi K, Salari E. Small Dim Object Tracking Using Frequency and Spatial Domain Information[J]. Pattern Recognition, 2016, 58(10): 227-234.
- [4] 吴涛, 何文忠, 陈晓露. 基于局部特征的单帧红外小目标检测算法[J]. 激光与红外, 2016, 8(3): 45-49.
- [5] 方义强, 程正东, 樊祥. 一种基于方差标记的形态学红外小目标检测算法[J]. 电子学报, 2015, 43(2): 338-343.
- [6] Wang Luping, Zhang Luping, Zhao Ming. An Algorithm for Moving Target Detection in IR Image Based on Grayscale Distribution and Kernel Function[J]. Journal of Central South University, 2014, 21(11): 4270-4278.
- [7] Bai Xiangzhi, Zhou Fugen, Xue Bindang. Multiple Linear Feature Detection Based on Multiple-structuring-element Center-surround Top-Hat Transforms[J]. Applied Optics, 2012, 51(21): 5201-5211.
- [8] 白相志, 周付根, 解永春. 新型 Top-Hat 变换及其在红外小目标检测中的应用[J]. 数据采集与处理, 2009, 24(5): 643-645.
- [9] 郑绍华, 陈健, 潘林, 等. 基于形态学的可变权值匹配自适应图像增强算法[J]. 电子与信息学报, 2014, 36(11): 2586-2592.
- [10] 田维. 复杂背景下红外小目标检测算法研究[D]. 沈阳: 东北大学, 2012.
- [11] 王鑫. 复杂背景下红外目标检测与跟踪算法研究[D]. 南京: 南京理工大学, 2011: 26-37.
- [12] 胡亦, 王琳娜, 朱恭生. 锯齿空间填充曲线耦合压缩感知的彩图灰度化实时加密算法[J]. 激光杂志, 2015, 6(2): 12-18.
- [13] Hai Jing, Zhang Gang, Cheng Jiang. Estimation of Mass Transfer Coefficient in Ozone Absorption by Linear Least Square Fitting and Simplex Search Methods[J]. Journal of Central South University, 2012, 19(12): 3396-3399.
- [14] 何乃甩. 基于红外图像的内河船舶目标检测算法研究[D]. 重庆: 重庆大学, 2008.
- [15] Cui Zheng, Yang Jingli, Jiang Shouda. An Infrared Small Target Detection Algorithm Based on High-speed Local Contrast Method[J]. Infrared Physics and Technology, 2015, 76(5): 474-481.
- [16] 赵爱罡, 王宏力, 杨小冈. 基于极端学习机背景预测的红外小目标检测算法[J]. 中国惯性技术学报, 2016, 24(1): 36-44.
- [17] Bianchi T, Piva A. Image Forgery Localization via Block-grained Analysis of JPEG Artifacts[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(3): 1003-1017.
- [18] Gao Chenqiang, Meng Deyu. Infrared Patch-image Model for Small Target Detection in a Single Image[J]. IEEE Transactions on Image Processing, 2013, 22(12): 4996-5008.

编辑 陆燕菲

(上接第 252 页)

- [4] Tong Lingling, Dai Feng, Zhang Dongming, et al. Encoder Combined Video Moving Object Detection[J]. Neuro-computing, 2014, 139(2): 150-162.
- [5] 徐光祐, 陶霖密, 邸慧军. 跨学科前沿与应用的交汇点: 日常活动理解[J]. 中国图象图形学报, 2013, 18(2): 132-140.
- [6] Wren C, Azarbayejani A, Darrell T, et al. Pfinder: Real-time Tracking of the Human Body[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.
- [7] Stauffer C, Grimson W E L. Adaptive Background Mixture Models for Real-time Tracking[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 1999: 246-252.
- [8] Elgammal A, Harwood D, Davis L. Non-parametric Model for Background Subtraction[C]//Proceedings of ECCV'00. Berlin, Germany: Springer, 2000: 751-767.
- [9] Doretto G, Chiuso A, Wu Y N, et al. Dynamic Texture[J]. International Journal of Computer Vision, 2003, 51(2): 91-109.
- [10] Mahadevan V, Vasconcelos N. Spatiotemporal Saliency in Dynamic Scenes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(1): 171-177.
- [11] Gopalakrishnan V, Rajan D, Hu Y. A Linear Dynamic System Framework for Salient Motion Detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(5): 683-692.
- [12] 刘敏. 视频序列中运动物体检测算法研究[D]. 武汉: 华中科技大学, 2013.
- [13] Farin D, de With P H N, Effelsberg W. Robust Estimation for Complex Video Sequences[C]//Proceedings of International Conference on Image Processing. Washington D. C., USA: IEEE Press, 2003.
- [14] Liu M, Zhang D. A New Approach for Salient Motion in Dynamic Scenes[C]//Proceedings of the 5th International Conference on Machine Vision. Washington D. C., USA: IEEE Press, 2012: 406-407.
- [15] Mahadevan V, Vasconcelos N. Background Subtraction in Dynamic Scenes[EB/OL]. [2016-08-10]. [http://www.svcl.ucsd.edu/projects/background\\_subtraction](http://www.svcl.ucsd.edu/projects/background_subtraction).

编辑 陆燕菲