

# 结合三帧差分的 ViBe 运动检测算法<sup>①</sup>

严红亮, 王福龙, 刘志煌

(广东工业大学 应用数学学院, 广州 510520)

**摘要:** 提出一种以 ViBe 算法为基础, 结合三帧差分思想的运动目标检测算法. 利用 ViBe 算法对每个像素点建模, 当前帧和模型得到的差分图与前一帧得到的差分图再进行与运算, 之后运用 ViBe 的思想对模型进行实时更新; 同时在每一帧添加小波去噪处理, 去除图像高频区域. 本文算法有效地解决了光照变化对系统的影响, 消除了影子问题, 去除了闪烁背景点. 实验结果表明, 本文算法在多种环境下可以准确地提取运动目标, 达到更好的鲁棒性.

**关键词:** 目标检测; ViBe 算法; 三帧差分; 小波去噪

## Motion Object Detection Combined ViBe with Three-Frame Differential

YAN Hong-Liang, WANG Fu-Long, LIU Zhi-Huang

(School of Applied Mathematics, Guangdong University of Technology, Guangzhou 510520, China)

**Abstract:** Based on ViBe, this paper explores a new moving object detection algorithm combining with the three frame difference method. Firstly, we build model on every pixels of the background with the advantage of ViBe; then, do the logic AND operation between two differential images which have been subtracted from the current image and the previous image; lastly, update the model in real time with the idea of ViBe. Meanwhile, in order to remove the high frequency component of the image, we add wavelet denoising to every frame of the image. This algorithm effectively overcomes the effect of illumination change on the system and eliminates the ghost as well as the blinking background pixels. Experiments confirm that this algorithm can accurately extract moving objects in multiple environments and has higher robustness.

**Key words:** object detection; ViBe algorithm; three-frame differential; wavelet de-noising

## 1 引言

运动目标检测是计算机视觉领域的一个热门课题, 它在视频智能监控、人机交互、智能交通系统等许多领域有着广泛的应用. 运动目标检测实际就是从视频图像中将运动物体从背景中提取出来, 其中背景又分为静态背景和动态背景. 静态背景下的运动目标即是视频图像中的变化区域. 静态背景下的运动目标检测容易实现, 模型单一, 算法一般都具有较强的鲁棒性, 因此近年来国内外学者们纷纷把目光放在动态背景下的目标检测上. 由于天气、光照、影子及遮挡等影响, 造成了背景发生变化, 使得目标检测成为一

项相当困难的工作. 目前, 运动目标的检测算法主要分成三类: 帧间差分法<sup>[1]</sup>、背景差分法<sup>[2]</sup>和光流法<sup>[3]</sup>.

帧间差分法实现简单, 对环境变化不太敏感, 实时性较好; 但是一般难以获得运动目标的完整区域, 对于快速运动的目标, 会产生“重影”的问题. 针对“重影”问题, 有人提出三帧差分法<sup>[7,13]</sup>, 效果有了显著提高.

光流法在背景变化的情况下性能较好; 问题是对于遮挡、多光源和噪声等影响, 不能解出正确的光流场, 同时实时性差.

背景差分法思想简单, 能够提取目标较完整的特

<sup>①</sup> 基金项目: 广东省自然科学基金(S2011040004273)

收稿时间: 2014-03-17; 收到修改稿时间: 2014-04-21

征信息, 实时性较好; 可是在环境变化的时鲁棒性差. 2011 年 Olivier Barnich 发表了一篇名为 ViBe<sup>[8]</sup>的背景差分算法, 和其他背景差分法不同的是, 它不是利用多帧初始化来建立背景模型, 而是在第一帧建立一个像素模型, 提高了算法速度, 对于第一帧中没有运动目标的视频场景鲁棒性好, 但是如果第一帧中包含运动目标的话, 就会因为建立的像素模型不准确导致在当前位置留下影子, 久久不能消去. 2012 年该研究团队提出的 ViBe+<sup>[9]</sup> 对原来算法在目标区域的完整度上有所改进, 消除了部分噪声点, 但是没有解决影子问题.

以上三种方法各有优劣, 由于光流法不适合实时检测, 因此当今的研究热点是帧间差分法和背景差分法的改进算法; 也有一些算法将两者结合<sup>[4]</sup>, 将视频图像中帧间差分图像和背景差分图像直接进行与运算, 这样就加大了目标信息的权重, 得到的运动检测图像包含了更多目标的信息, 不仅包含目标轮廓而且还有目标轮廓内的目标相关点, 从而得到较完整较准确的目标区域, 但是这类算法要求背景与运动目标的对比度大于 5% 以上, 才能检测出, 同时还存在一些缺点仍然没有得到解决, 这些问题还有待进一步研究.

## 2 ViBe 算法

ViBe 背景提取算法首次提出单帧初始化建立一个像素模型, 更新机制采用随机替换像素原则, 同时运用空间传播机制使得像素模型与邻域模型互相更新, 不同于传统算法利用几十帧的视频序列初始化建立背景模型, 然后采用定期更新机制, 先入先出的准则使模型的样本在模型中的寿命保持一致. ViBe 算法有效地降低了程序的复杂度, 在加快了对于帧的处理速度的同时, 可以达到较高的检测准确度, 实时性很强.

### 2.1 像素模型及初始化

在视频图像的第一帧, 对图像的每一个像素点建立一个像素模型. 定义像素  $v(x)$  位于图像的  $x$  处, 在  $v(x)$  的八邻域中随机选取  $N$  个像素点作为  $v(x)$  对应模型的样本值, 建立像素模型  $M(x)$ .

$$M(x) = \{v_1, v_2, \dots, v_N\} \quad (1)$$

其中  $v_i (i = 1, 2, \dots, N)$  为从八邻域中随机选取的样本值, 如果  $v(x)$  是图像边界点, 则从对应的不完整的

八邻域中随机选取.

判断当前帧中的像素是否为前景目标点的准则是

$$f(x) = \begin{cases} 1 & |v_t(x) - v_i(x)| < R \\ 0 & |v_t(x) - v_i(x)| \geq R \end{cases} \quad (2)$$

$v_t(x)$  为当前帧  $x$  处的像素,  $v_i(x)$  为  $x$  处的模型样本. 当  $f(x)$  的值为 1 即为前景目标点, 为 0 则为背景点.

### 2.2 像素模型的更新

图像中背景是动态的, 因此随着时间的推移, 背景会发生一些变化, 比如相机抖动、前景遮挡、光照变化等等. 如果背景更新不能迅速跟上实际背景的变化, 检测结果中可能出现较大范围的噪声, 这要求更新算法要有较强的抗干扰能力. ViBe 提出一种新的更新机制——动态自适应更新方法. 这个机制包含了三个要点: 1) 无记忆更新, 2) 时间二次采样, 3) 空间传播<sup>[10]</sup>.

### 2.3 ViBe 优缺点

ViBe 算法是一种快速的目标检测法, 在解决诸如动态背景, 模型初始化等方面具有良好的性能; 自适应的更新机制可保证像素模型的采样值的寿命周期呈平滑指数衰减的形式, 并且能够使该方法在单模型、变速度、可接受的内存消耗的条件下, 可以处理每个像素的伴随事件.

但是该算法也存在一些缺陷, 主要是三个问题(见图 1): 1) 对光照的突然变化非常敏感; 2) 处理闪烁的背景点效果不好, 会造成大量误检; 3) 当第一帧中含有运动目标时, 会在当前位置处留下目标的影子, 直到更新许久之后才能慢慢消失.

## 3 改进算法

### 3.1 处理光照变化和影子

ViBe 建立的像素模型是基于色彩空间(R,G,B)的, 用当前帧的像素对比相应位置所建立的像素模型中的样本所代表的三个色彩通道的色彩值, 因此当背景像素点的色彩值由于光照发生改变时, 会被误检为目标. 又由于该算法用的是单帧初始化, 用第一帧建立一个初始像素模型, 所以当第一帧中含有运动目标时, 当前帧与所建立的模型相对比, 会把第一帧目标位置处的背景点误认为是前景点, 即产生目标留下的影子.

产生这两种问题的原因是 ViBe 算法模型单一且更新达不到背景变化的速度, 该算法能够随着时间的

推移,慢慢恢复误检的背景点,但是恢复的时间很长. 为了解决这两个问题,本文采用三帧差分法的思

想,用当前帧与像素模型作对比得到的差分图,和前一帧得到的差分图再做与运算. 原理图见图 2 所示.

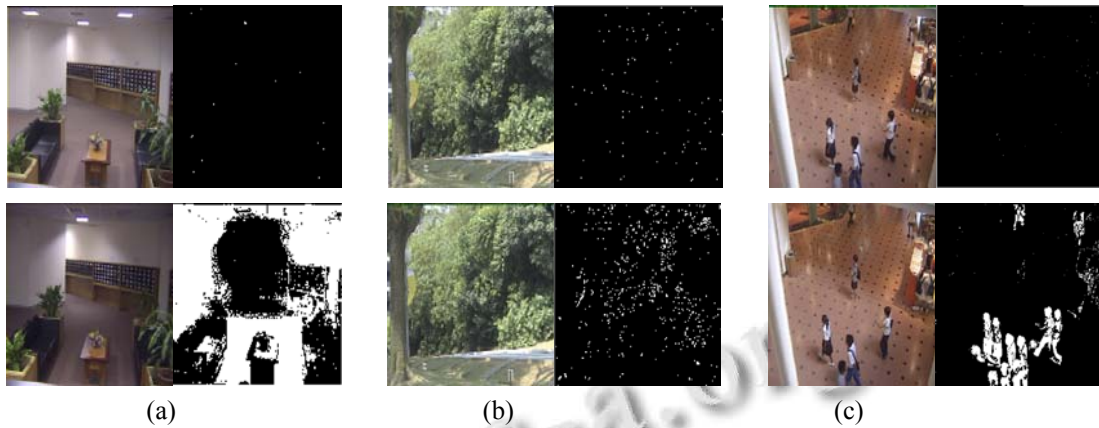


图 1 (a)为相邻两帧光照突然变暗时, ViBe 对纯背景图的检测; (b)为 ViBe 对于闪烁像素点的处理; (c)中第一帧含有运动目标,在之后帧中 ViBe 造成的影子现象.

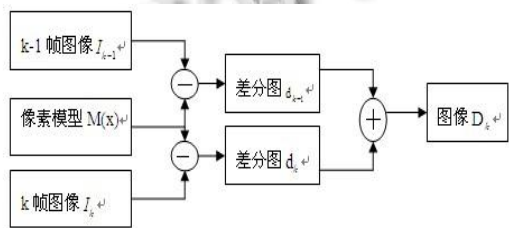


图2 三帧差分法原理图

在图2中,  $I_{k-1}$  为视频第k-1帧图像,  $I_k$  为第k帧图像,  $M(x)$ 为所建立的像素模型. 由公式(2)可以分别得到第k-1帧和第k帧的二值差分图像  $d_{k-1}$  和  $d_k$ , 再对  $d_{k-1}$  和  $d_k$  做与运算, 便得到了一个最后所求的二值图像  $D_k$ .

图像与运算的规则为:

$$\begin{cases} 0 \oplus 0 = 0 \\ 0 \oplus 1 = 1 \\ 1 \oplus 0 = 1 \\ 1 \oplus 1 = 0 \end{cases} \quad (3)$$

这里0表示背景点, 1表示前景点.

本文对加入与运算规则稍作调整, 变为:

$$\begin{cases} 0 \oplus 0 = 0 \\ 0 \oplus 1 = 1 \\ 1 \oplus 0 = 0 \\ 1 \oplus 1 = 0 \end{cases} \quad (4)$$

即在前一帧被认为是前景目标的像素点, 如果在当前帧认为它是背景点, 那么我们就把它作为背景点, 这是由实验得出的.

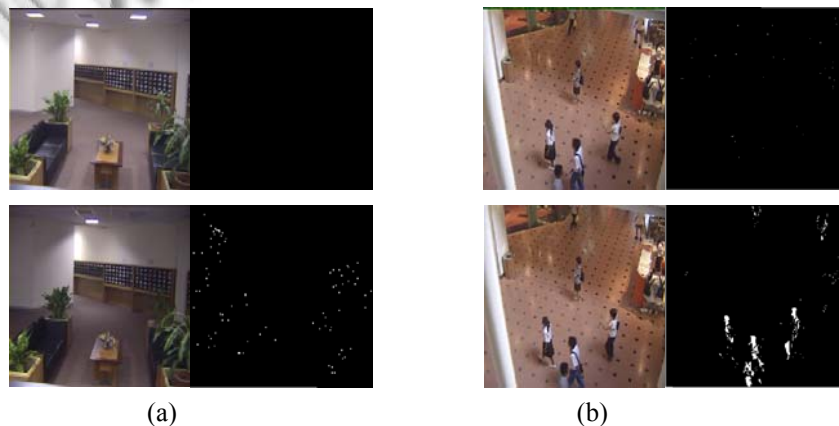


图 3 图(a)(b)分别是结合了三帧差分法对光照突然变暗的效果和目标影子的消除.

如图 3 所示, (a)(b)的原图和图 1 相对应, (a)的两幅图分别取自 Wallflower 图像库<sup>[15]</sup>“Light”中第 1449 帧和 1450 帧, 灯光突然变暗时, 结合三帧差分的处理效果, 虽然存在一小部分背景点的误检, 但是对比图 1 中 ViBe 的效果有了明显改进; (b)的两幅原图取自图像库“ShoppingMall”中第 1001 帧(即第一帧)和 1004 帧, 利用三帧差分的思想, 消除了目标在第一帧位置处留下的影子, 从(b)中可以看出有一个人未被检测出, 那是由于此人站在原处没有移动的缘故。

加入三帧差分法的思想, 有效地消除了影子问题, 对于光照改变造成的背景点的误检, 也做到及时地纠正。



图 4 小波去噪原理图

经过离散小波变换处理每一帧的高频部分, 可以有效地去除噪声点, 包括闪烁的背景像素点。

图 5 两幅原图对应图 1, 分别来自 Wallflower 图像库“Campus”中的第 1001 帧和 1006 帧, 图中没有运动目标, 只有树叶的不停摇晃。摇晃的树叶即为闪烁的背景点, 通过小波去噪之后, 对比图 1, 可以看到这些闪烁点几乎没有误检。

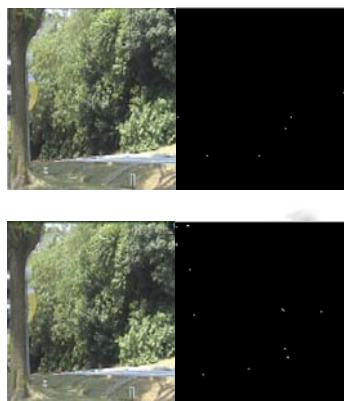


图 5 去除闪烁点

### 3.3 算法流程

- 1) 对视频每一帧进行小波去噪处理;
- 2) 用 ViBe 算法处理去噪后的视频, 得出背景减去的二值图  $d_i$ ;
- 3) 对图像  $d_i$  每三帧之间, 根据公式(4)提供的规则使用差分算法, 得到最终的背景减去图  $D_i$ 。

### 3.2 背景闪烁点的消除

如果视频中存在闪烁的背景点(如摇晃的树叶, 波动的水面), ViBe 算法用第一帧建立像素模型不具有全局通用性, 会造成这些背景点的误检。由于这些点通常是图像高频区域像素点, 因此只须对小波系数的高频部分进行处理。

小波去噪处理的实质是根据小波域上图像信号多分辨率分析的特点, 构造出最佳的规则, 尽可能的减小甚至完全剔除由噪声产生的相关系数, 同时保留真实图像信号的系数。最后将去噪后的各个尺度的小波系数进行重构而得到真实图像信号的最优估计。小波流程图如下所示:

## 4 实验结果

为了验证本文算法对多种环境的鲁棒性, 这里采用数据库中背景建模最常用的九个不同场景下的视频图像, 其中包括室内和室外的场景。测试视频每帧图像大小都为 160128 像素。在装有 VS2010(配置 opencv2.3.1), Intel(R) Celeron(R) 处理器(1.60GHz), 2G 内存的 PC 机上实现。在对视频背景减去的效果进行评价时, 我们采用主观评价。

图 6 中(1-9)分别来自 Wallflower 图像库的“Campus”“Light”“ShoppingMall”“Fountain”“Hall”“WaterSurface”“Curtain”“Bootstrap”和“Escalator”。(a)列表示原图, (b)列为真正的目标, (c)列是 ViBe 的效果, (d)列是运用的与运算的 ViBe 的实验效果, (e)列是在(d)的基础上增加了小波去噪后的效果。

从图 6 可以看出, 本文算法在(1-7)场景中的效果良好, 能够准确检测出目标位置, 去除了(1)(3)(4)(5)(7)中 ViBe 算法造成的影子, 对于(2)中光照突然变化带来的大面积误检问题, 也取得了很好的检测效果, 增加小波去噪后, 在(1)(4)(6)中可以看到消去了许多闪烁的背景点。然而本文算法在处理场景(8)(9)上不足, 只能检测出目标的位置, 但是不能提取出目标的轮廓, 特别是场景(9), 电梯的滑动造成了大量的误检, 真正的目标未能得到准确的检测。



图6 各种场景效果图

由于考量目标检测的另一个重要指标是算法的实时性, 因此本文对图 6 的场景 1“Campus”做了详细的

测试,其他场景可以类似得到相应数据。“Campus”视频共包含1438帧图像,ViBe算法能达到26.8帧/秒的检测速率,本文算法由于添加了小波去噪和三帧差分,算法上加大了复杂性,因此只能达到15.6帧/秒的速率。虽然处理速率对比ViBe算法低了许多,但是也达到了视频正常运行的速率。

## 5 总结

本文算法在对视频图像的目标检测上,对比ViBe,大大降低了对光照变化的敏感度;在第一帧出现运动目标的情况下,消除了目标的影子;加入小波去噪,减少了闪烁背景点的误检。通过实验可以看出,本文算法对比ViBe算法有了极大的改进,但也存在不足:第一,三帧差分算法容易造成目标内部的空洞现象,因此本文算法虽然能够检测出目标的位置,但是不能得到目标的完整区域;第二,算法的复杂度影响了对视频的处理速度,如何在保证检测效率的同时加快处理视频的速度,这是下一步研究的方向。

## 参考文献

- 1 Lipton AJ, Fujiyoshi H, Patil RS. Moving target classification and tracking from real-time video. Proc. of Fourth IEEE Workshop on Applications of Computer Vision, WACV'98. IEEE, 1998. 8-14.
- 2 Arseneau S, Cooperstock JR. Real-time image segmentation for action recognition. 1999 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing. IEEE, 1999. 86-89.
- 3 Barron JL, Fleet DJ, Beauchemin SS. Performance of optical flow techniques. International Journal of Computer Vision, 1994, 12(1): 43-77.
- 4 Zhang H, Wu K. A vehicle detection algorithm based on three-frame differencing and background subtraction. 2012 Fifth International Symposium on Computational Intelligence and Design (ISCID). IEEE, 2012, 1: 148-151.
- 5 Brutzer S, Hoferlin B, Heidemann G. Evaluation of background subtraction techniques for video surveillance. 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2011. 1937-1944.
- 6 Cheng FH, Chen YL. Real time multiple objects tracking and identification based on discrete wavelet transform. Pattern Recognition, 2006, 39(6): 1126-1139.
- 7 Gang L, Shangkun N, Yugan Y, et al. An improved moving objects detection algorithm. 2013 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR). IEEE, 2013. 96-102.
- 8 Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences. IEEE Trans. on Image Processing, 2011, 20(6): 1709-1724.
- 9 Van Droogenbroeck M, Paquot O. Background subtraction: Experiments and improvements for ViBe. 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2012. 32-37.
- 10 Maddalena L, Petrosino A. A self-organizing approach to background subtraction for visual surveillance applications. IEEE Trans. on Image Processing, 2008, 17(7): 1168-1177.
- 11 Tao G, Zhengguang L, Jun Z. Redundant discrete wavelet transforms based moving object recognition and tracking. Journal of Systems Engineering and Electronics, 2009, 20(5): 1115-1123.
- 12 Singh R, Purwar RK, Rajpal N. A better approach for object tracking using dual-tree complex wavelet transform. 2011 International Conference on Image Information Processing (ICIIP). IEEE, 2011. 1-5.
- 13 Su Y, Qian R, Ji Z. Surveillance video sequence segmentation based on moving object detection. Second International Workshop on Computer Science and Engineering, WCSE'09. IEEE, 2009, 1. 534-537.
- 14 Al-Berry MN, Salem MAM, Hussein AS, et al. Motion detection using wavelet-enhanced accumulative frame differencing. 2013 8th International Conference on Computer Engineering & Systems (ICCES). IEEE, 2013. 255-261.
- 15 Toyama K, Krumm J, Brumitt B, et al. Wallflower: Principles and practice of background maintenance. Proc. of the Seventh IEEE International Conference on Computer Vision, 1999. IEEE, 1999, 1: 255-261.
- 16 Wang H, Suter D. A consensus-based method for tracking: Modelling background scenario and foreground appearance. Pattern Recognition, 2007, 40(3): 1091-1105.