

# 基于视觉背景提取的自适应运动目标提取算法

吕嘉卿\*, 刘立程, 郝禄国, 张文忠

(广东工业大学 信息工程学院, 广州 510006)

(\* 通信作者电子邮箱 katenan@126.com)

**摘要:**在复杂场景下的视频运动目标提取是视频分析技术的首要工作。为了解决前景运动目标提取的精确度不高的问题,提出一种基于视觉背景提取(ViBE)的改进视频运动目标提取算法(ViBE\*)。首先,在背景模型初始化阶段采用像素的菱形邻域来简化样本信息;其次,在前景运动目标提取阶段引入自适应分割阈值来适应场景的动态变化;最后,在更新阶段提出背景重建和调整更新因子方法来处理光照变化的情形。实验结果表明,对于复杂视频场景 LightSwitch 的运动目标提取结果在相似度指标上,改进后的算法与混合高斯模型(GMM)算法、码本模型算法以及原始 ViBE 算法相比,分别提高了 1.3 倍、1.9 倍以及 3.8 倍。所提算法能够在有效时间内对复杂场景具有较好的自适应性,且性能明显优于对比算法。

**关键词:**前景提取;视觉背景提取;背景建模;自适应阈值;更新因子

**中图分类号:** TP391.413 **文献标志码:** A

## Adaptive moving object extraction algorithm based on visual background extractor

LYU Jiaqing\*, LIU Licheng, HAO Lugu, ZHANG Wenzhong

(College of Information Engineering, Guangdong University of Technology, Guangzhou Guangdong 510006, China)

**Abstract:** The prior work of video analysis technology is video foreground detection in complex scenes. In order to solve the problem of low accuracy in foreground moving target detection, an improved moving object extraction algorithm for video based on Visual Background Extractor (ViBE), called ViBE\*, was proposed. Firstly, in the model initialization stage, each background pixel was modeled by a collection of its diamond neighborhood to simply the sample information. Secondly, in the moving object extraction stage, the segmentation threshold was adaptively obtained to extract moving object in dynamic scenes. Finally, for the sudden illumination change, a method of background rebuilding and update-parameter adjusting was proposed during the process of background update. The experimental results show that, compared with the Gaussian Mixture Model (GMM) algorithm, Codebook algorithm and original ViBE algorithm, the improved algorithm's similarity metric on moving object extracting results increases by 1.3 times, 1.9 times and 3.8 times respectively in complex video scene LightSwitch. The proposed algorithm has a better adaptability to complex scenes and performance compared to other algorithms.

**Key words:** foreground extraction; Visual Background Extractor (ViBE); background modeling; self-adaptive threshold; update-parameter

## 0 引言

随着模式识别、机器学习以及人工智能等学科的飞速发展,利用计算机视觉与图像处理技术来获取对客观世界的感知成为当下计算机视觉领域中最热门的课题之一。其中,提取视频图像中感兴趣的部分,如移动的人和车辆,也称为运动目标检测,一直是视频图像分析技术的研究热点。

传统的运动目标检测算法主要有帧间差法<sup>[1]</sup>、光流法<sup>[2]</sup>和背景差法<sup>[3]</sup>。其中,帧间差法的原理及实现技术相对简单,但目标的轮廓信息不能完整提取,容易产生空洞现象;光流法能较好检测出前景运动目标,但需要一定的硬件支持才能达到实时处理的效果;背景差法需要建立良好的背景模型以适应动态背景的变化,采用该算法提取的前景目标较为完整且噪点较少。综合考虑,背景差法是一个较为折中的算法,

成为当前视频处理领域中运动目标检测的热门研究方法。

背景差法的核心思想是建立一个鲁棒性的背景模型,将当前帧图像和背景图像进行差分,得到感兴趣的前景运动目标。从参数角度看,背景模型具体分为有参数背景建模和非参数背景建模。作为参数模型的经典代表,混合高斯模型(Gaussian Mixture Model, GMM)<sup>[4-5]</sup>对每个像素建立多个独立的高斯分布,在复杂场景下能较好地分割出运动目标。但是,该算法需要一段时间来训练样本,而且由于参数是固定的,当光照发生突变时混合高斯模型很难建立有效的背景模型,因此,一些学者提出了基于非参数的背景模型,比较有代表性的是以下两种:1) 码本(Codebook)模型<sup>[6]</sup>。该模型为每个像素建立一个码本结构,实时性较强,但消耗大量内存,同时易受背景中存在的细微扰动影响。2) 视觉背景提取(Visual Background Extractor, ViBE)。该算法采用随机样本

收稿日期:2015-02-02;修回日期:2015-03-26。

**作者简介:**吕嘉卿(1991-),女,广东顺德人,硕士研究生,主要研究方向:视频图像处理;刘立程(1972-),男,福建上杭人,副教授,博士,主要研究方向:无线通信、数字信号处理;郝禄国(1968-),男,河北霸县人,讲师,博士,主要研究方向:多媒体通信、数字信号处理;张文忠(1991-),男,广东汕头人,硕士研究生,主要研究方向:信号与信息系统。

模型<sup>[7]</sup>,能较快地分割出完整的运动目标,对噪声也具备一定的抗扰能力。

然而 ViBE 算法本身也有局限性,表现在以下 3 个方面:1)背景模型的样本值是重复选取的;2)目标检测过程中,固定的目标分割阈值无法适应复杂视频场景中背景的动态变化;3)固定的更新因子无法有效消除光照变化所引起的噪点。针对 ViBE 算法存在的问题,文献[8]提出扩大邻域来避免样本的重复选取,文献[9]和文献[10]提出结合梯度值和样本值来抑制鬼影传播,但同时增加了算法的复杂度和运算时间。文献[11]采用基于空间邻域的阈值选取方法来适应各类视频场景,但该方法增加了计算量。

针对 ViBE 存在的局限性,本文提出一种基于 ViBE 的视频前景提取算法 ViBE<sup>+</sup>。首先,在初始化阶段,背景模型的样本信息通过以当前像素为中心的菱形邻域来选取,以此减少建立背景模型所需的样本点和运算时间;其次,采用自适应阈值代替单一固定的阈值来提取前景目标;最后,根据光照变化的剧烈程度,进行相应的背景重建和更新因子调整。

## 1 ViBE

ViBE 算法由文献[12]首次提出,该算法采用随机选择策略进行背景模型的建立,同时利用空间传播特性来更新邻域像素点的模型样本。ViBE 算法<sup>[13]</sup>的工作原理是为每个像素点建立一个样本集,样本集中的采样值是从与当前像素点空间相邻的像素值中随机选取的,然后将图像帧的每一个像素点及其对应的样本集中所有的像素点进行比较来区分前景点和背景点。下面简单地描述一下 ViBE 算法的处理步骤。

### 1) 建立背景模型。

ViBE 算法仅用单帧视频图像序列来初始化背景模型,根据相近像素点具有较强的空间相关特性,采用随机策略采样相邻像素点来填充当前像素的样本集,即对于每一个像素点  $x_i$ ,从它的  $m$  邻域  $N_G(x_i)$  随机选择  $N$  个像素值来作为它的背景模型样本值。图像中像素点  $x_i$  在  $t$  时刻的背景模型定义如式(1)所示:

$$B(x_{i,t}) = \{v_{1,t}^C, v_{2,t}^C, \dots, v_{N,t}^C \mid v_{j,t}^C \in N_G^m(x_i), j = 1, 2, \dots, N\} \quad (1)$$

其中:  $v_{j,t}^C$  表示在  $t$  时刻从像素  $x_i$  的邻域中随机选取的样本值,且每个样本值有  $C$  个通道(灰度图像:  $C = 1$ ; RGB 颜色空间:  $C = (r, g, b)$ )。  $m$  和  $N$  为经验值,原始算法中取  $m = 8, N = 20$ 。

### 2) 前景运动目标检测。

统计像素点  $x_i$  与其背景模型样本值距离较近的样本数目,如果满足式(2),则把像素点  $x_i$  判断为背景点;否则判断为前景点:

$$\#\{ |x_{i,t+1} - B(x_{i,t})| \leq R \} \geq \#min \quad (2)$$

其中:  $\#\{\cdot\}$  表示满足条件表达式的样本值的个数;  $R$  为目标分割阈值,一般取 20;  $\#min$  表示数量阈值,一般取 2。

### 3) 背景模型更新。

ViBE 算法采用保守的背景模型更新策略,利用随机选取法把判断为背景点的像素点以  $1/\varphi$  的概率( $\varphi$  一般取 16)去更新自己的模型样本值,同时为了保持像素邻域空间的一致性,也以  $1/\varphi$  的概率去更新它的邻居点的背景模型样本值,即经

过  $dt$  时间后,某一样本值  $v_{j,t}^C$  仍保留在背景模型中的概率为:

$$P(v_{j,t}^C, v_{j,t+dt}^C) = e^{-\ln(\frac{N}{N-1})dt} \quad (3)$$

式(3)表明了背景模型中每一个样本值的生命周期呈指数衰减,随机策略是合适的。但保守机制容易引起死锁,所以当像素点连续  $num$  次检测为前景点,则将其更新为背景点。

## 2 改进的视频前景提取算法

ViBE 算法原理简单,提取视频前景的效果也很好,且实时性好,但仍有不足之处。针对引言中指出的 ViBE 算法存在的问题,本文将逐一进行分析,同时提出相应的改进策略,作为本文所提出的 ViBE<sup>+</sup> 算法。

### 2.1 简化样本信息

原始 ViBE 算法是从像素点  $x_i$  的 8 邻域中随机选取 20 个样本值来构成背景模型的,在简单的场景下,这种方法的目标检测效果较好。但在复杂多变的场景下,由于样本值的重复选取,容易增加背景像素点误判为前景像素点的概率。针对这一问题,本文采取以当前像素点为中心的固定菱形邻域来建立背景模型,如图 1 所示的像素采样布局,即对于每个像素点  $x_i$ ,其背景模型只有当前像素点  $x_i$  及其四菱形邻域  $N_G^4(x_i)$  共 5 个点的样本信息,如式(4)所示。根据人的视觉特性研究得知,人眼往往容易忽略视频图像边缘处的细节变化,所以对于边缘处的像素点,本文仅仅选取当前像素点来构成其对应的背景模型。

$$B(x_{i,t}) = \{v_{1,t}^C, v_{2,t}^C, x_{i,t}^C, v_{4,t}^C, v_{5,t}^C \mid v_{j,t}^C \in N_G^4(x_i)\} \quad (4)$$

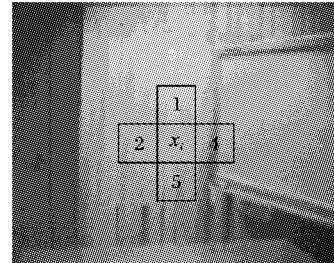


图1 像素采样布局

### 2.2 自适应阈值

在前景运动目标提取阶段,目标分割阈值  $R$  为单一固定值,优点是考虑了整体的视频场景情形,但是容易忽略图像的细微信息,从而出现局部不清晰的检测效果。由于实际的视频场景都是多模态的,如摇曳的树叶、晃动的水波等,而固定的阈值无法适应这种动态变化的视频场景。所以,针对这一问题,本文引入了自适应的目标分割阈值,其实现步骤如下。

1) 统计像素点  $x_i$  在第  $t$  时刻的模型样本值,求得样本模型中最大值  $max_{i,t}$ 。

2) 遍历  $t$  时刻图像帧的每一个像素点,求得当前像素点的矩形邻域的 9 个像素点的平均值  $\mu_{i,t}$  和标准差  $\sigma_{i,t}$ 。

3) 求判断条件,如式(5)所示:

$$T = \mu_{i,t} + \lambda \sigma_{i,t} \quad (5)$$

4) 获得自适应的目标分割阈值,如式(6)所示:

$$R(x_i) = \begin{cases} R(x_{i-1}) \cdot (1 + \alpha), & max_{i,t} < T \\ R(x_{i-1}) \cdot (1 - \alpha), & max_{i,t} \geq T \\ R_{init}, & R(x_i) > R_{max} \text{ 或 } R(x_i) < R_{min} \end{cases} \quad (6)$$

其中:  $\alpha \in [0, 1]$ ;  $\lambda$  为常数, 一般取  $[2, 6]$ ;  $R_{\max}$  和  $R_{\min}$  分别表示目标分割阈值的最大值和最小值, 使得阈值始终保持在最佳的适用范围内, 通过多次实验验证取  $R_{\max} = 35$ ,  $R_{\min} = 5$ ;  $R_{\text{init}}$  表示目标分割阈值的初始值, 本文取 18。

### 2.3 光照变化

在前景运动目标检测过程中, 视频场景光照的突然变化和缓慢变化往往会引起不佳的运动目标检测效果。而原始 ViBE 算法没有考虑场景光照变化的情形, 且随机采样更新方法相对缓慢, 从而使得前景提取效果不理想, 甚至可能导致整个区域都被检测为前景点。针对这一问题, 本文对视频场景的光照变化进行实时判断, 即将图像进行分块处理, 然后统计每个区域的前景点个数, 如果满足式 (5), 则认为场景光照发生变化, 采取以当前图像重建背景模型并根据式 (8) 来调整更新因子; 否则认为没发生变化, 不进行处理。

$$\frac{\# \{x_i(\text{foreground})\}}{\text{block}} > Th \quad (7)$$

其中:  $\text{block}$  表示分块后的区域,  $x_i(\text{foreground})$  表示某区域内的前景点,  $Th$  为判断场景光照变化的经验临界值, 通过多次实验验证取  $Th = 0.65$ 。

$$\varphi = \begin{cases} \varphi, & Th < 0.65 \\ \varphi/2, & 0.65 \leq Th < 0.8 \\ \varphi/4, & 0.8 \leq Th \leq 1 \end{cases} \quad (8)$$

其中  $\varphi$  表示更新因子, 本文取 16。

### 3 实验结果及分析

为了验证本文所提出的改进算法 ViBE<sup>+</sup> 的有效性, 本次实验在 Intel Core i3 2.53 GHz CPU 及 2.00 GB 内存的配置环境下, 采用 VS2010 及 OpenCV 平台对 5 个帧率为 30 帧/s 的复杂场景视频<sup>[14-15]</sup>进行仿真。

为了验证本文算法的性能, 本文与 GMM 算法、Codebook 算法、ViBE 算法等进行比较, 如图 2 所示。检测效果都是对原视频序列进行相应的算法处理, 然后经过相同的形态学后处理。其中: Curtain 为背景窗帘飘动的视频序列; Hall 为商城中人来人往的视频序列; Campus 为复杂背景下的视频序列 (阳光下树叶及其阴影剧烈晃动, 且有车辆和行人运动); Lobby 为前景运动目标受障碍物阻挡的视频序列; LightSwitch 为光照突变情形的视频序列, 用于测试算法对光照的敏感性。

窗帘的飘动 (序列 Curtain)、人来人往的商场 (序列 Hall) 以及在更复杂的视频背景下 (序列 Campus), GMM 算法和 Codebook 都出现了把背景误检为前景目标的现象, 而 ViBE 算法能消除这些干扰。较之原始 ViBE 算法, 本文所提出的采用菱形邻域来构建背景模型以及引入自适应阈值的 ViBE<sup>+</sup> 算法更胜一筹。受障碍物阻挡的影响 (序列 Lobby), ViBE 算法得到的运动目标出现了空洞, GMM 算法和 Codebook 算法基本能获取完整的运动目标, 相比之下, ViBE<sup>+</sup> 算法的前景检测效果更好。对于场景光线发生剧烈变化的 LightSwitch 序列 (视频开始时房间是漆黑的, 然后灯光突然开启, 同时目标出现), 其他算法几乎无法检测出运动目标, 误判率较高, 而本文算法考虑了光照变化情形, 检测效果远远优于其他算法。

同时, 本文根据检测运行时间和以下 4 个评价指标<sup>[16-17]</sup>对实验过程计算得出的相关数据进行处理并分析。

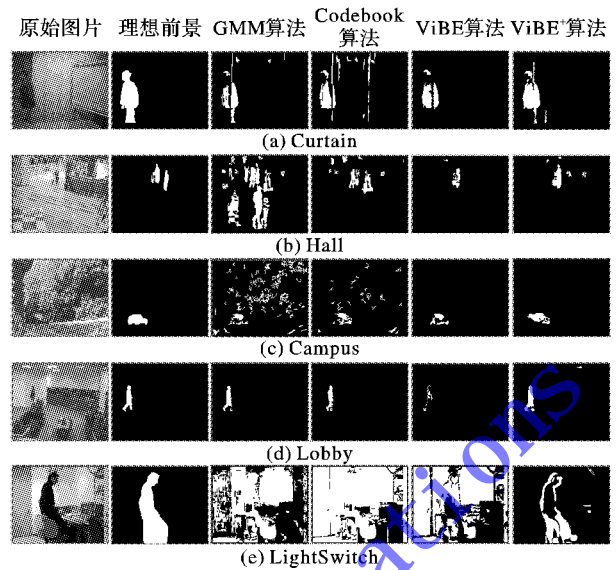


图2 4种算法的检测效果

1) 查全率 (Recall), 计算公式如下:

$$\text{Recall} = tp / (tp + fn) \quad (9)$$

2) 查准率 (Precision), 计算公式如下:

$$\text{Precision} = tp / (tp + fp) \quad (10)$$

3)  $F_1$ -准则, 计算公式如下:

$$F_1 = 2 \times \text{Recall} \times \text{Precision} / (\text{Recall} + \text{Precision}) \quad (11)$$

4) 相似度 (Similarity), 计算公式如下:

$$\text{Similarity} = tp / (tp + fp + fn) \quad (12)$$

其中:  $tp$  表示判断正确的前景点,  $fn$  表示判断错误的背景点,  $fp$  表示判断错误的前景点。

从表 1 和图 3 可以看出, 在 5 类视频场景中, 每一个算法的性能指标都与仿真实验的检测效果相匹配。其中, 在 Lobby 视频序列中, 由于 GMM 算法对视频进行背景模型训练处理, 因而能够提取出较好的前景目标, 但算法计算量大, 运行时间较长, 实时性相对而言较差。Codebook 算法的总体运行时间较短, 但需要存储大量码本结构, 消耗内存。ViBE<sup>+</sup> 算法比 ViBE 算法略复杂, 处理速度较慢, 但基本也满足实时性要求。而且在复杂视频序列 LightSwitch 中, 与其他 3 种算法相比, ViBE<sup>+</sup> 算法的运动目标提取结果在 Similarity 指标上分别提高了 1.3 倍、1.9 倍和 3.8 倍。总体来说, 本文所提出的 ViBE<sup>+</sup> 算法表现更稳定, 更占优势。

表1 4种算法的运行时间比较

视频序列	GMM	Codebook	ViBE	ViBE <sup>+</sup>
Curtain	36.717	6.476	9.583	17.011
Hall	76.268	10.261	22.576	37.294
Campus	24.191	3.463	5.222	9.672
Lobby	29.988	4.339	7.137	14.983
LightSwitch	41.808	8.839	10.777	21.241

### 4 结语

本文对 ViBE 算法的优点与缺点进行详细的分析, 并针对其存在的问题逐一给出改进方法, 最终实现 ViBE<sup>+</sup> 算法。为了简化样本信息, 本文所提出的 ViBE<sup>+</sup> 算法采用更具相关性的菱形邻域布局来填充背景样本集; 在前景提取阶段, 引入



自适应阈值来处理各种动态变化的场景,提高了背景与目标分割的准确性;并且通过背景重建和调整更新因子的策略,使得算法在一定程度上对光照变化所引起的干扰具有较强的抵抗能力。与其他3种经典算法比较的实验结果表明,ViBE<sup>+</sup>算法具有较好的检测效果,且满足实时性要求。在后续研究中,将针对其他突变的场景进行考虑。

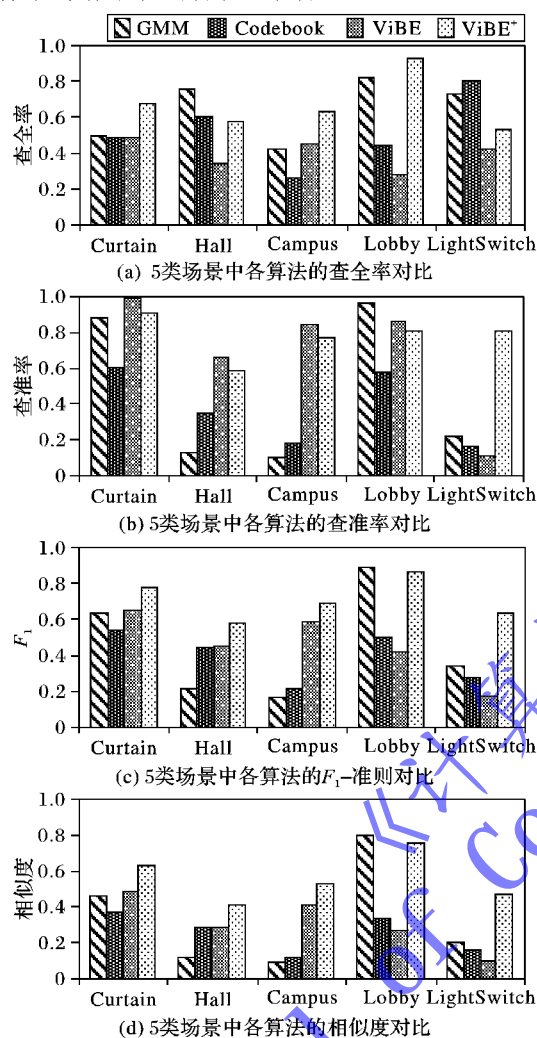


图3 4种算法的性能指标

#### 参考文献:

- [1] SHEN J, YANG W, LU Z, *et al.* Information integration for accurate foreground segmentation in complex scenes [J]. *IET Image Processing*, 2012, 6(5): 596–605.
- [2] KUMAR A, TUNG F, WONG A, *et al.* A decoupled approach to illumination-robust optical flow estimation [J]. *IEEE Transactions on Image Processing*, 2013, 22(10): 4136–4147.
- [3] LI B, YANG G. Adaptive foreground detection approach of Gaussian mixture mode [J]. *Journal of Image and Graphics*, 2013, 18(12): 1620–1627. (李百惠, 杨庚. 混合高斯模型的自适应前景提取[J]. *中国图象图形学报*, 2013, 18(12): 1620–1627.)
- [4] STAUFFER C, GRIMSON W E L. Adaptive background mixture models for real-time tracking [C]// *CVPR 1999: Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 1999: 246–252.
- [5] ZIVKOVIC Z. Improved adaptive Gaussian mixture model for background subtraction [C]// *ICPR 2004: Proceedings of the 17th International Conference on Pattern Recognition*. Piscataway: IEEE, 2004: 28–31.
- [6] KYUNG NAM K, CHALIDABHONGSE T H, HARWOOD D, *et al.* Background modeling and subtraction by codebook construction [C]// *Proceedings of the 2004 International Conference on Image Processing*. Piscataway: IEEE, 2004: 3061–3064.
- [7] LIU H, HOU X, HAO X, *et al.* Moving object detection algorithm based random sample model [J]. *Journal of Electronic Measurement and Instrument*, 2013, 27(12): 1155–1161. (刘红海, 侯向华, 郝秀兰, 等. 采用随机样本进行运动目标检测的算法研究[J]. *电子测量与仪器学报*, 2013, 27(12): 1155–1161.)
- [8] YU Y, CAO M, YUE F. EViBE: an improved ViBE algorithm for detecting moving objects [J]. *Chinese Journal of Scientific Instrument*, 2014, 35(4): 924–931. (余烨, 曹明伟, 岳峰. EViBE: 一种改进的 ViBE 运动目标检测算法[J]. *仪器仪表学报*, 2014, 35(4): 924–931.)
- [9] van DROOGENBROECK M, PAQUOT O. Background subtraction: experiments and improvements for ViBE [C]// *CVPRW 2012: Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Piscataway: IEEE, 2012: 32–37.
- [10] GE W, DONG Y, GUO Z, *et al.* Background Subtraction with dynamic noise sampling and complementary learning [C]// *ICPR 2014: Proceedings of the 22nd International Conference on Pattern Recognition*. Piscataway: IEEE, 2014: 2341–2346.
- [11] SU Y, LI A, JIANG K, *et al.* Improved visual background extractor model for moving objects detecting algorithm [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2014, 26(2): 232–240. (苏延召, 李艾华, 姜柯, 等. 改进视觉背景提取模型的运动目标检测算法[J]. *计算机辅助设计与图形学学报*, 2014, 26(2): 232–240.)
- [12] BARNICH O, van DROOGENBROECK M. ViBE: a powerful random technique to estimate the background in video sequences [C]// *ICASSP 2009: Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. Piscataway: IEEE, 2009: 945–948.
- [13] BARNICH O, van DROOGENBROECK M. ViBE: a universal background subtraction algorithm for video sequences [J]. *IEEE Transactions on Image Processing*, 2011, 20(6): 1709–1724.
- [14] FISHER R B. Change detection in color images [EB/OL]. [2014-12-18]. <http://homepages.inf.ed.ac.uk/rbf/PAPERS/iccv99.pdf>.
- [15] TOYAMA K, KRUMM J, BRUMITT B, *et al.* Wallflower: principles and practice of background maintenance [C]// *Proceedings of the 7th IEEE International Conference on Computer Vision*. Piscataway: IEEE, 1999: 255–261.
- [16] GOYETTE N, JODOIN P, PORIKLI F, *et al.* Changedetection.net: a new change detection benchmark dataset [C]// *CVPRW 2012: Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Piscataway: IEEE, 2012: 1–8.
- [17] AGARWAL S, AWAN A, ROTH D. Learning to detect objects in images via a sparse, part-based representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(11): 1475–1490.