

# 基于多层背景模型的运动目标检测

曹明伟, 余 焯

(合肥工业大学计算机与信息学院, 安徽合肥 230009)

**摘要:** 复杂背景下的运动目标检测一直是计算机视觉领域中一个具有挑战性的问题, 本文提出一种基于多层背景模型的运动目标检测算法. 该算法首先从视频序列的第一帧中提取每个像素的邻域样本, 用于初始化背景模型, 只需一帧图像即可完成背景模型的初始化; 其次, 为实现背景模型的自适应更新, 引入随机采样技术, 随机选取一个不匹配的码字, 采用新的背景像素取而代之, 避免错误分类的码字长时间驻留在背景模型中; 为处理动态场景中多种干扰因素的影响, 提出了多层背景模型策略, 每个像素经过多层背景模型的逐层验证, 保证了背景模型的精确性. 实验结果表明, 该算法能够有效克服复杂背景下的多种干扰因素影响, 且检测率和识别率均高于现有经典算法.

**关键词:** 动态背景; 目标检测; 随机采样; 视频监控; 像素分类器

**中图分类号:** TP391      **文献标识码:** A      **文章编号:** 0372-2112 (2016)09-2126-08

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2016.09.016

## Moving Object Detection Based on Multi-layer Background Model

CAO Ming-wei, YU Ye

(School of Computer & Information, Hefei University of Technology, Hefei, Anhui 230009, China)

**Abstract:** Moving object detection under complex-background is always a challenging issue, and in order to defend these challenges, this paper proposed an algorithm named MMBM (Moving object detection based on Multi-layer Background Model). First, samples are selected from neighbors of each pixel of the first frame to initialize background model. Only one frame image is needed for initialization. Second, in order to update the background model adaptively, random sampling technique is introduced, i. e., selecting one code word randomly from the background model and then updating it with new background pixel, which overcomes the deficiency of the wrong classified code word overstaying in the background model. Multi-layer background model is proposed in order to overcome the influence of multi-disturbance in dynamic background, in which one pixel is tested through multi-layers, so as to guarantee and improve the accuracy of background pixels. Finally, Experimental results show that this algorithm can overcome the influence of multi-disturbance existing in dynamic outside scenes effectively, and at the same time, achieve a higher detection rate and recognition rate over the existing classical algorithms.

**Key words:** dynamic background; object detection; random sampling; video surveillance; pixel classifier

## 1 引言

智能视频监控为当今社会的治安管理和维护社会稳定起到了重要作用, 被广泛应用于民用和军工业中. 运动目标检测作为智能视频监控系统的核心, 受到了国内外学者的广泛关注, 并取得一些研究成果<sup>[1-3]</sup>. 如 Stauffer 等<sup>[4]</sup>提出的混合高斯建模算法 (Mixture of Gaussian, MOG), Olivier 等<sup>[5]</sup>提出的可视背景提取子算法 (Visual Background Extractor, VIBE)、Kim 等<sup>[6]</sup>提出的码书模型算

法 (CodeBook, CB) 和基于像素的自适应分割 (Pixel-Based Adaptive Segmenter, PBAS) 算法<sup>[7]</sup>, 局部二进制相似性分割 (Local Binary Similarity segmenter, LOBSTER) 算法<sup>[8]</sup>和基于像素的自适应单词一致性分割 (Pixel-based Adaptive Word Consensus Segmenter, PAWCS) 算法<sup>[9]</sup>均为目前常用的动态背景下的运动目标检测算法<sup>[10-14]</sup>.

MOG 算法采用多个高斯分布构造背景模型, 通过设置学习率实现背景模型的自适应更新. VIBE 采用真实的像素集合来表示背景模型, 采用 2-D 空间内的欧氏

距离对像素进行分类,算法效率较高.为了提高 VIBE 算法像素分类准确性, Van 提出了带有反馈信息的 PBAS 算法. CodeBook 采用少量码字组成的码书训练视频序列的背景模型.然而,此算法采用学习率实现背景模型更新,必然会面对“权衡问题”. Geng 等<sup>[15]</sup>采用像素的亮度值范围代替原始的分类标准,提高了算法对阴影的鲁棒性. Wu 等<sup>[16]</sup>将时空上下文(Spatio-Temporal Context)信息加入到 CodeBook 算法中,提高了算法的精确性. Charles 等<sup>[17]</sup>采用时空特征描述子构造每个像素背景模型,提出一种具有较强鲁棒性的 SuBSENSE 算法,但是此算法实时性较低. LOBSTER 和 PAWCS 算法采用局部二进制相似模式(Local Binary Similarity Patterns, LBSP)构造背景模型,然而 LBSP 的计算量较大使得算法的实时性降低.以上这些背景建模算法被广泛应用于复杂运动目标检测系统<sup>[4,5,6,12]</sup>.

尽管取得了上述研究成果,在复杂背景下运动目标检测算法仍存在如下不足:(1)背景模型初始化时间过长;(2)背景模型的更新系数设定不具有通用性;(3)单层背景模型难以处理复杂背景干扰问题.为弥补上述不足,本文在 CodeBook 模型基础之上提出一种“基于多层背景模型的运动目标检测(Moving Object Detection Based on Multi-layer Background Model, MMBM)”算法.首先,从第一帧图像的邻域空间内选取 20 个样本用于初始化背景模型,避免背景模型初始化时间过长的问題;其次,为了实现背景模型自适应更新,提出等概率替代的自适应背景更新方法,有效解决传统方法无法避免的“权衡问题”;再次,通过多层背景模型对多种干扰因素逐层验证,有效的克服了多干扰因素同时存在的影响.实验结果表明该算法在室外多干扰因素存在的复杂背景下,能够实现运动目标的实时检测.

## 2 CodeBook 模型介绍

为了便于分析 CodeBook 算法实现运动目标检测的原理,给出以下两个术语的示例说明:

(1)前景像素:属于运动目标区域的像素,如图 1 中的行人区域.

(2)背景像素:运动目标以外的像素,如图 1 中的非行人区域.



图1 前景与背景像素示意图

### 2.1 CodeBook 模型描述

记  $X$  表示一个像素的训练序列,此序列由  $N$  个  $V_{RGB}$  向量组成,  $X = \{x_1, x_2, \dots, x_N\}$ .  $CB = \{c_1, c_2, \dots, c_L\}$ , 由  $L$  个码字(CodeWord)组成,表示一个像素的码书(CodeBook)模型.每一个码字  $c_i$  由一个 3 维的向量  $V_{RGB}(v_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i))$  和一个六元组( $\text{aux}_i = \langle \hat{I}_i, \hat{I}_i, f_i, \lambda_i, p_i, q_i \rangle$ )的辅助信息组成.  $c_i$  中每个参数的详细信息如下所示:

(1)  $v_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$ , 表示像素的 RGB 颜色分量组成的 3 维向量.

(2)  $\hat{I}_i, \hat{I}_i$ , 分别表示匹配成功码字的最小亮度值和最大亮度值.

(3)  $f_i$ , 表示码字出现的频率.

(4)  $\lambda_i$ , 表示最大反面运行长度(Maximum Negative Run-Length, MNRL),即码字没有匹配的最大时间间隔.

(5)  $p_i, q_i$ , 分别表示码字第一次和最后一次匹配成功的时间.

### 2.2 CodeBook 模型的像素分类标准

CodeBook 算法采用像素分类思想区分前景像素与背景像素,其像素分类标准是运动目标检测算法的核心.如图 2 所示,如果像素  $X_i$  在判别边界(decision boundary)内部,表明  $X_i$  是匹配码字;否则不匹配.具体分类过程采用像素分类器 1(Classifier<sub>1</sub>)完成:

$$\text{Clf}_1 = \begin{cases} 1, & \text{if } \begin{cases} \text{colordist}(x_i, v_i) > T \\ \text{bttness}(I, \langle \hat{I}, \hat{I} \rangle) = \text{true} \end{cases} \\ 0, & \text{else} \end{cases} \quad (1)$$

其中各变量计算方法如下:

$$\text{colordist}(x_i, v_i) = \delta = \sqrt{\|x_i\|^2 - p^2} \quad (2)$$

$$\|x_i\|^2 = R^2 + G^2 + B^2 \quad (3)$$

$$p^2 = \|x_i\|^2 \cos^2 \theta \quad (4)$$

$$\text{bttness}(I, \langle \hat{I}, \hat{I} \rangle) = \begin{cases} \text{true}, & \text{if } I_{\text{low}} \leq \|x_i\| \leq I_{\text{hi}} \\ \text{false} & \text{otherwise} \end{cases} \quad (5)$$

$$I_{\text{low}} = \alpha \hat{I}, I_{\text{hi}} = \min \left\{ \beta \hat{I}, \frac{\hat{I}}{\alpha} \right\} \quad (6)$$

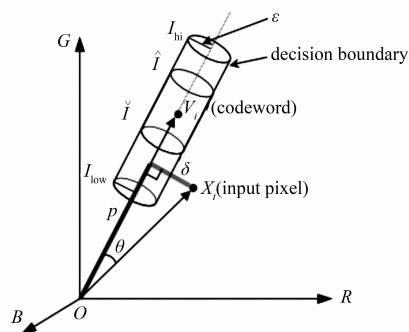


图2 像素分类模型

### 3 MMBM 算法

MMBM 算法的流程如图 3 所示,其核心思想为:对于给定的图像序列,从第一帧中提取训练样本,采用像素分类器 1 提取背景像素,构造背景模型. 创建多层背景模型,在背景模型的更新上采用等概率替代的自适应背景更新方法. 基于多层背景模型,利用像素分类器 2 (将在 3.3 节中进行定义) 对像素进行分类,判断哪

些属于背景像素,构造多层背景模型. 同时,基于像素分类器 2 对多层背景模型中的主模型进行更新,将缓存模型中满足条件的像素提升至主模型. 通过多层背景模型的逐层验证,准确地检测出运动目标. 算法的三个主要部分:背景模型初始化、自适应背景更新方法、多层背景模型的创建与更新,将在下面的小节中进行详细阐述.

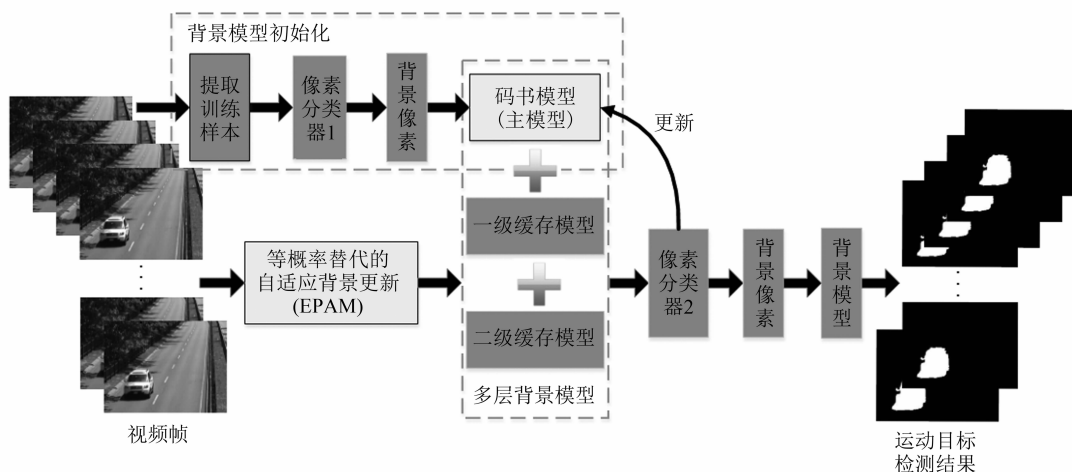


图3 MMBM算法流程图

#### 3.1 基于邻域的背景模型初始化

模型初始化的本质是“像素表达问题”,由于邻域内的像素具有空间一致性<sup>[12]</sup>,这些像素能够很好表达背景模型. 因此,基于空间一致性原理,本文提出了基于邻域的背景模型初始化方法,即:读取视频序列的第一帧,利用邻域中的像素来初始化背景模型,算法只需

一帧图像即可完成背景模型的初始化,提高了算法的时间效率. 如图 4 所示,首先从像素的 24 邻域内随机选取不重复的 20 个样本,然后采用这些样本构造对应的码字,对码字进行验证,将满足条件的码字组成像素的码书模型,遍历整幅图像即可构造完整的背景模型.

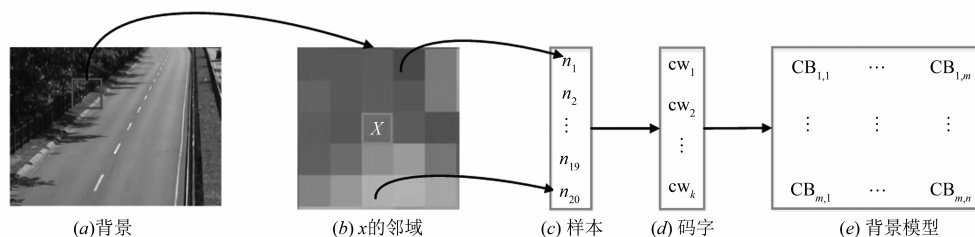


图4 模型初始化过程示意图

#### 3.2 等概率替代的自适应背景更新

为避免室外环境中动态背景的干扰,需要背景模型能够自适应更新,因此,提出“等概率替代的自适应背景更新方法”(Equal Probability Alternative Method, EPAM). 假设在背景模型  $M(x)$  中找到与像素  $x_i$  匹配的码字  $c_m$ , EPAM 算法不是对码字  $c_m$  进行更新,而是采用像素  $x_i$  创建新的码字,取代  $[1, m-1]$  之间的某一个码字. 由于像素  $x_i$  与  $[1, m-1]$  之间的码字不匹配,说明  $x_i$  是一个更优的背景像

素,因此,基于  $x_i$  创建一个新的码字插入到背景模型中.

背景更新过程如图 5 所示. EPAM 方法的详细步骤如下:

**Step1** 在背景模型  $M(x)$  中,通过像素分类器 1 寻找与像素  $x_i$  匹配的码字  $c_m$ ;

**Step2** 采用均匀分布的随机数生成算法,产生一个范围在  $[1, m-1]$  之间的随机数  $rnd$ ,以确定被替换的码字为  $c_{rnd}$ ;

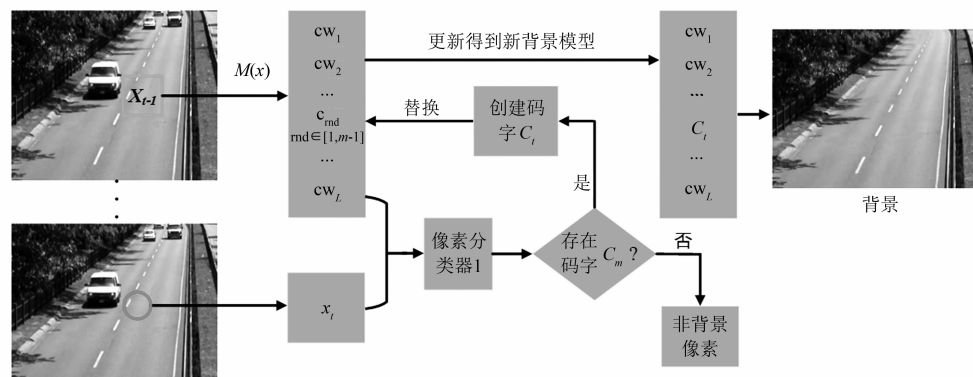


图5 背景更新过程

**Step3** 使用像素  $x_t$  创建一个新的码字  $c_t$ , 如下所示:

$$v_t = x_t, \text{aux}_t = \langle I, I, 1, t-1, t, t \rangle$$

**Step4** 采用新的码字  $c_t$  替代码字  $c_{\text{rnd}}$ , 即可实现背景模型的自适应更新.

### 3.3 多层背景模型的创建与更新

为了克服室外场景中多种干扰因素影响, 本文提出了多层背景模型的策略. 按照干扰因素对背景影响的强度进行分类, 定义了 3 层背景模型, 分别是:  $M(x)$  (主模型)、 $H(x)$  (一级缓存模型) 和  $U(x)$  (二级缓存模型),  $H(x)$  和  $U(x)$  的数据结构定义与  $M(x)$  相同, 模型的创建分为三个阶段, 在训练阶段创建主模型  $M(x)$ , 在检测阶段根据第一帧输入序列与主模型的匹配情况来创建一级缓存模型  $H(x)$ , 然后根据第二帧输入序列与一级缓存模型的匹配情况来创建二级缓存模型  $U(x)$ . 当二级缓存模型  $U(x)$  中的码字满足一定时间, 就将  $U(x)$  中码字提升到  $H(x)$  中; 当  $H(x)$  中的码字满足一定的时间, 则将  $H(x)$  中的码字提升到  $M(x)$  中, 从而得到更加精确的主模型  $M(x)$ . 多层背景模型的创建过程可以采用“像素分类器 2” ( $\text{Clf}_2$ ) 实现, 如式 (7) 所示.

$$\text{Clf}_2 = \begin{cases} \text{De}_1, \lambda_1 > \text{all}(T_U, T_{\text{add}1}, T_{\text{delete}1}) \\ \text{De}_2, \lambda_2 > \text{all}(T_H, T_{\text{add}2}, T_{\text{delete}2}) \end{cases} \quad (7)$$

其中,  $\text{De}_1$  和  $\text{De}_2$  分别表示删除二级缓存模型和一级缓存模型的码字.  $T_U$ 、 $T_{\text{add}1}$ 、 $T_{\text{delete}1}$ 、 $T_H$ 、 $T_{\text{add}2}$  和  $T_{\text{delete}2}$  为实验值.

采用递归思想描述多层背景模型的创建和更新过程, 如下所示:

**Step1** 基于邻域初始化方法创建主模型  $M(x)$ , 同时, 分别创建空的一级缓存模型  $H(x)$  和二级缓存模型  $U(x)$ .

**Step2** 对于  $t$  时刻给定的输入像素  $I_t$ , 在  $M(x)$  中寻找与  $I_t$  匹配的码字  $\text{CW}_M$ . 如果  $\text{CW}_M$  在  $M(x)$  中, 则认为  $I_t$  属于背景像素, 采用 EPAM 方法对主模型  $M(x)$  进

行更新, 否则跳转到 Step3.

**Step3** 如果没有找到匹配的码字, 按以下步骤进行处理:

(1) 将像素  $I_t$  标记为前景像素.

(2) 在一级缓存模型  $H(x)$  中查找匹配码字  $\text{CW}_H$ , 如果找到匹配的码字  $\text{CW}_H$ , 则使用 EPAM 方法对  $H(x)$  进行更新. 否则采用  $I_t$  创建一个新的码字  $\text{CW}_t$  插入到一级缓存模型  $H(x)$  中.

(3) 在二级缓存模型  $U(x)$  中寻找与  $I_t$  匹配的码字  $\text{CW}_U$ , 如果找到匹配的码字  $\text{CW}_U$ , 则使用 EPAM 算法对模型  $U(x)$  进行更新, 否则使用  $I_t$  创建一个新的码字  $\text{CW}_t$  插入到二级缓存模型  $U(x)$  中.

**Step4** 采用像素分类器 2, 将漏检的真实背景码字提升到主模型  $M(x)$  中, 过程如下所示:

(1) 移除二级缓存模型  $U(x)$  中码字的最长时间间隔  $\lambda_1$  大于  $T_U$  的所有码字; 将  $U(x)$  中码字的最长时间间隔  $\lambda_1$  大于  $T_{\text{add}1}$  的所有码字提升到  $H(x)$  中, 同时删除  $H(x)$  中码字的最大时间间隔大于  $T_{\text{delete}1}$  的所有码字.

(2) 移除一级缓存模型  $H(x)$  中码字的最长时间间隔  $\lambda_2$  大于  $T_H$  的所有码字; 将  $H(x)$  中码字的最长时间间隔  $\lambda_2$  大于  $T_{\text{add}2}$  的所有码字提升到  $M(x)$  中, 同时删除  $M(x)$  中码字的最大时间间隔  $\lambda_2$  大于  $T_{\text{delete}2}$  的所有码字.

**Step5** 对于新的像素  $I_{t+1}$  跳转到 Step2 进行处理.

通过 Step1 ~ Step5 即可实现多层背景模型的创建与更新, 为提高算法的精确性奠定基础. 采用 Camarim 序列对多层背景模型的有效性进行了验证, 图 6(a) 为 Camarim 序列的第 132 帧, 图 6(b) 为基于单层背景模型的检测结果, 图 6(c) 为多层背景模型检测结果, 可以看

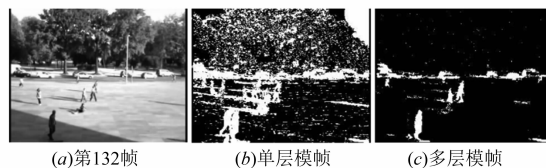


图6 多层背景模型结果

出经过多层背景模型的逐层检验,从而避免摇晃树叶和光照的影响,同时提高了算法对阴影的鲁棒性.

#### 4 实验结果与分析

MMBM 算法的开发环境为 Visual C++ 2010、OpenCV SDK 2.4.6、Nvidia CUDA SDK 5.5,所用 PC 配置为 Intel i5 CPU、8.0GB RAM. 实验使用的数据集为 ChangeDetection<sup>[18,19]</sup>和 Cross (本文创建). 为了进行算法精度的对比分析,采用与本文算法相同的软硬件开发环

境实现了 MOG 和 VIBE 算法;PBAS、SuBSense、LOBSTER 和 PWACS 算法采用作者提供的源代码;CodeBook 算法采用 OpenCV SDK 2.4.6 提供的可调用类. 为了公正、有效的对算法进行比较,MOG、VIBE、PBAS 等算法的参数取值与原文相同,且实验结果没有经过后处理.

##### 4.1 定性分析

在 ChangeDetection 数据集上,对 MMBM 和 MOG 等算法进行测试,结果如图 7 所示.

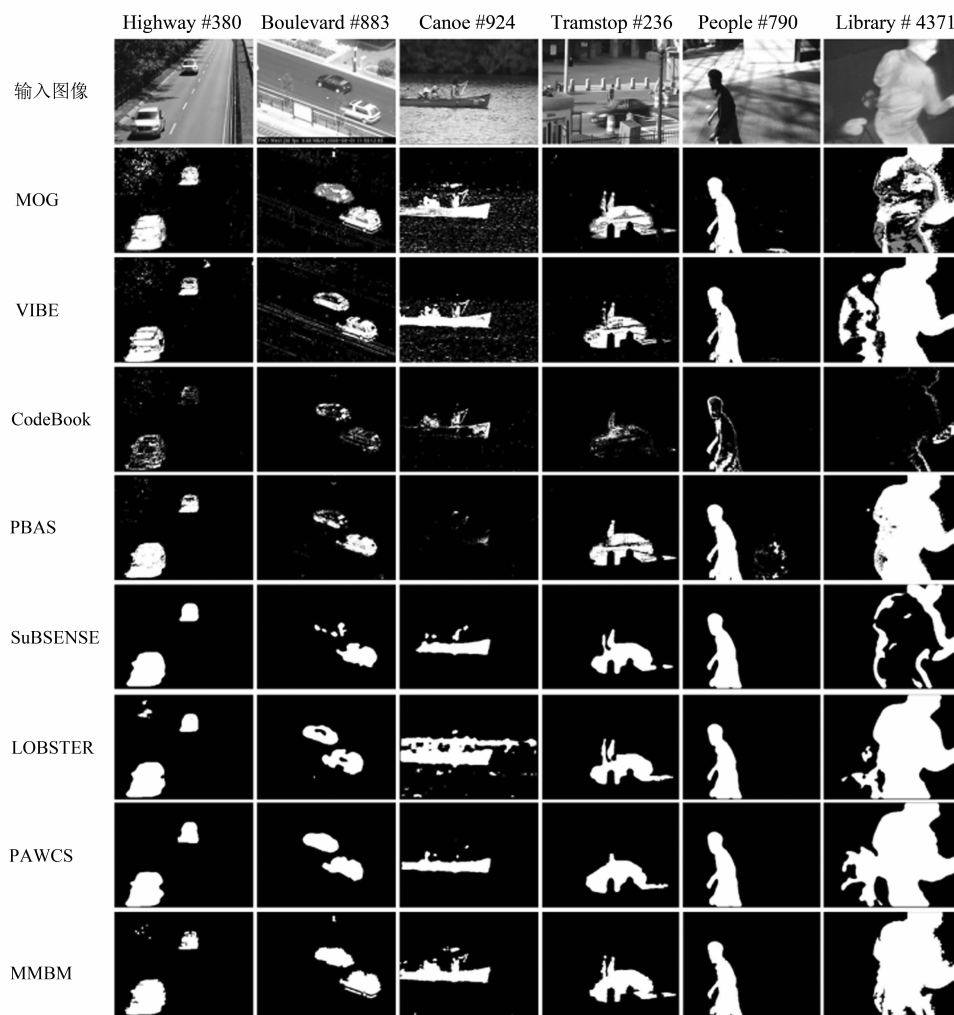


图 7 ChangeDetection数据集测试结果

从 Highway 序列的测试结果可以看出,MOG、VIBE 和 LOBSTER 算法都受到了背景中摇晃树叶的干扰,导致检测结果中出现一些小的白色亮点;CodeBook 算法结果很模糊;SuBSense、PWACS 和本文算法的检测结果清晰且不存在残缺.

从 Boulevard 序列检测结果可以看出,MOG 和 VIBE 算法受到序列中光照变化的影响,在地面上形成了白色条纹;CodeBook 算法的检测结果在路面位置没

有出现白色条纹,但是检测出的车辆较模糊;PBAS、SuBSense 和 LOBSTER 算法结果中目标存在残缺;本文方法的检测结果没有受到光照的影响,结果清晰可见.

从 Canoe 序列检测结果可以看出,MOG、VIBE 和 LOBSTER 算法都能清晰的检测出运动的小船,然而受到湖面水波运动的影响,检测结果中出现了大量白点;CodeBook 和 PWACS 算法检测出的小船出现了大

面积残缺;PBAS 算法的结果中小船的尾部漏检;SuB-SENSE 算法没有检测出小船上的行人. 本文方法既没有受到水波影响,检测结果也不存在残缺,清晰可见.

从 Tramstop 序列检测结果可以看出,MOG、VIBE、PBAS 以及本文算法都能很好的处理遮挡问题,且本文方法比 SuBSENSE 算法的结果更加精确;由于受光照和遮挡的影响,导致 CodeBook 算法的检测结果较模糊;LOBSTER、PAWCS 和本文方法虽然受到行人遗留轨迹的影响,相对而言本文算法结果更加精确.

从 People 序列检测结果可以看出,MOG、VIBE、SuBSENSE、LOBSTER、PAWCS 以及本文算法的检测结果都很清晰;在 CodeBook 算法结果中行人的上身出现了大面积残缺;PBAS 算法受到了阴影的影响,在行人背后出现大面积虚假信息.

从 Library 序列检测结果可以看出,MOG、PBAS 和本文算法均可用于检测红外序列中的运动目标,且本文方法的检测结果相对精确一些;VIBE、LOBSTER 和 PAWCS 算法检测出了行人运动的遗留轨迹;CodeBook 和 SuBSENSE 算法只检测出行人的部分边缘信息.

从上述分析可知,本文算法对“多干扰复杂动态背景”具有较强的鲁棒性.

## 4.2 定量分析

### 4.2.1 算法精度分析

为了公正的比较算法性能,采用文献[19]中的评价标准对算法的性能进行定量分析,评价标准如下:

- (1) P(Precision) 表示算法的准确率;
- (2) R(Recall) 表示识别率;
- (3) Sp(Specificity) 表示背景像素的识别率;
- (4) FPR(False Positive Rate) 表示背景像素被错误

分类为前景像素的比率;

(5) FNR(False Negative Rate) 表示前景被错误分类为背景像素的比率;

(6) PWC(Percentage of Wrong Classification) 表示错误分类的比率;

(7) FM(F-Measure) 表示算法的测度值,其值越高越好.

基于上述评价标准,在 ChangeDetection 和 Cross 数据集上对各检测算法进行评估,并采用 ChangeDetection 数据集提供的算法性能统计代码,分别对每个算法的精度进行统计分析,结果如表 1 所示.可以看出,本文算法无论是在准确率(P)还是在识别率(R)上均高于 MOG、VIBE、CodeBook、PBAS 和 SuBSENSE 等算法.

### 4.2.2 时间效率分析

假设图像的宽度为  $W$ ,高度为  $H$ ,本文算法的时间复杂度为  $O(W \times H \times (N_s \times N_s + S_s))$ ,其中背景模型初始化的时间复杂度为  $O(N_s \times N_s)$ , $N_s$  表示背景模型初始化时所用样本的邻域空间,实际测试时取值为 5. 运动目标检测时对每个像素进行分类的时间复杂度为  $O(S_s)$ ,由于本文采用 20 个样本表达一个像素的背景模型,因此在算法中  $S_s$  的取值为 20. CodeBook 算法的时间复杂度为  $O(W \times H \times (W \times H \times N))$ ,其中  $N$  表示训练背景模型所用样本数.由于 MOG、VIBE、PBAS 和 SuBSENSE、LOBSTER 和 PAWCS 算法的作者没有明确给出时间复杂度分析,因此,为了公正比较,我们在 ChangeDetection 和 Cross 数据集上对每个算法进行测试,计算出每个算法在不同序列上的处理时间,结果(每个算法计算 3 次取平均值)如表 2 所示,可以看出本文算法在所有测试序列上所消耗时间均最少.

表 1 算法精度测试结果

Method	P	R	SP	FPR	FNR	PWC	F-Measure
MOG	0.5726	0.6564	0.9701	0.0299	0.3436	4.0265	0.6116
Vibe	0.7153	0.6706	0.9765	0.0235	0.3294	3.2064	0.6922
CodeBook	0.4736	0.5837	0.8832	0.1168	0.4163	8.2459	0.5229
PBAS	0.7967	0.7842	0.9901	0.0099	0.2158	1.7654	0.7904
SuBSENSE	0.7632	0.8084	0.9887	0.0113	0.1916	1.8204	0.7852
LOBSTER	0.8023	0.7847	0.9856	0.0144	0.2153	2.0157	0.7934
PWACS	0.8087	0.8106	0.9917	0.0083	0.1894	1.8579	0.8096
MMBM	0.8106	0.8123	0.9936	0.0064	0.1274	1.7625	0.8405

表 2 算法时间效率

测试序列	检测时间/秒							
	MOG	VIBE	CodeBook	PBAS	SuBSENSE	LOBSTER	PWACS	MMBM
Highway	27257.6	557.3	128.3	570.4	889.5	202.9	123.9	98.2
Boulevard	45196.2	678.8	368.6	645.5	182.1	210.4	287.6	240.4
Canoe	19095.1	1100.6	39.5	464.5	156.0	76.9	79.0	22.3
Tramstop	51884.2	869.5	771.7	605.6	568.3	1234.4	597.4	530.2
People	19480.7	401.9	1190.0	599.6	535.4	345.1	653.0	504.3
Library	78404.3	1233.3	867.3	155.7	1252.0	225.7	513.9	458.7
Leaves	326195.0	1048.0	275.3	570.3	843.8	280.6	5910.2	152.3
Crowd	371491.0	834.2	56.8	753.9	1146.7	320.4	7239.4	32.5
Traffic	119776.0	1155.0	77.3	70.6	1061.7	853.8	3553.2	42.6
FarAway	118711.0	765.9	964.2	214.7	1061.8	856.7	1601.7	623.7
Fog	132773.0	817.1	1047.5	782.6	1214.4	686.8	841.8	756.7
Slow	255036.0	596.7	303.0	537.5	302.5	988.0	10930.1	200.7

## 5 总结与展望

针对复杂背景下的运动目标检测进行研究,提出了一种适用于复杂背景下的运动目标检测算法—MMBM.采用从第一帧图像每个像素的邻域中选取样本,对背景模型进行初始化的方案,算法的初始化过程得到了简化和加速;通过采用 EPAM 背景模型更新策略,避免了被错误分类的像素长期保留在背景模型中;由于多层背景模型方法的使用,即在初始化阶段创建一个主模型,在检测阶段创建两个缓存模型,通过缓存模型的逐层验证,将漏检的背景像素提升到主模型,增强了主模型的精确性,提高了算法在复杂背景下的鲁棒性.

在 ChangeDetection 和 Cross 数据集上开展的实验结果表明,与经典的 MOG、VIBE、CodeBook 等算法相比,MMBM 算法的检测率和识别率均有较大提高且在多干扰复杂动态背景环境下具有很好的鲁棒性.

### 参考文献

- [1] 丁莹,李文辉,范静涛,等.基于 Choquet 模糊积分的运动目标检测算法[J].电子学报,2010,38(2):263-268.  
Ding Ying, Li Wenhui, Fan Jingtao, et al. A moving object detection algorithm based on choquet integrate[J]. Acta Electronica Sinica, 2010, 38(2): 263-268. (in Chinese)
- [2] 李宏友,汪同庆,叶俊勇,等.基于 3DOG 的视频运动目标检测算法[J].电子学报,2008,36(3):605-610.  
Li Hongyou, Wang Tongqing, Ye Junyong, et al. An approach based on 3D orthogonal Gaussian-Hermite moments for detecting moving objects[J]. Acta Electronica Sinica, 2008, 36(3): 605-610. (in Chinese)
- [3] 魏志强,纪筱鹏,冯业伟.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 Yewei. A moving object detection method based on self-adaptive updating of background[J]. Acta Electronica Sinica, 2005, 33(12): 2261-2264. (in Chinese)
- [4] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[A]. Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on[C]. Fort Collins, CO: IEEE, 1999. 2246-2253.
- [5] Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences[J]. Image Processing, IEEE Transactions on, 2011, 20(6): 1709-1724
- [6] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time foreground-background segmentation using codebook model[J]. Real-time Imaging, 2005, 11(3): 172-185.
- [7] Hofmann M, Tiefenbacher P, Rigoll G. Background segmentation with feedback; The pixel-based adaptive segmenter[A]. Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on[C]. Providence, RI: IEEE, 2012. 38-43.
- [8] St-Charles P L, Bilodeau G A. Improving background subtraction using local binary similarity patterns[A]. Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on[C]. Steamboat: IEEE, 2014. 509-515.
- [9] St-Charles P L, Bilodeau G A, Bergevin R. A self-adjusting approach to change detection based on background word consensus[A]. Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on[C]. Waikoloa Beach, HI:

- IEEE,2015.990-997.
- [10] Sobral A, Vacavant A. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos[J]. *Computer Vision and Image Understanding*,2014,122(5):4-21.
- [11] Brutzer S, Hoferlin B, Heidemann G. Evaluation of background subtraction techniques for video surveillance[A]. *Computer Vision and Pattern Recognition (CVPR)*,2011 IEEE Conference on[C]. Colorado Springs, CO: IEEE, 2011. 1937-1944.
- [12] Bouwmans T. Traditional and recent approaches in background modeling for foreground detection: An overview [J]. *Computer Science Review*,2014,58(3):1-37.
- [13] Xue K, Liu Y, Ogunmakin G, et al. Panoramic Gaussian mixture model and large-scale range background subtraction method for PTZ camera-based surveillance systems [J]. *Machine Vision and Applications*,2013,24(3):477-492.
- [14] Valera M, Velastin S A. Intelligent distributed surveillance systems: a review[J]. *IEE Proceedings-Vision, Image and Signal Processing*,2005,152(2):192-204.
- [15] Geng L, Xiao Z T. Real time foreground-background segmentation using two-layer codebook model[A]. *Control, Automation and Systems Engineering (CASE)*,2011 International Conference on[C]. Singapore: IEEE, 2011. 1-5.
- [16] Wu M, Peng X. Spatio-temporal context for codebook-based dynamic background subtraction[J]. *AEU-International Journal of Electronics and Communications*,2010,64(8):739-747.
- [17] St-Charles P L, Bilodeau G A, Bergevin R. Flexible background subtraction with self-balanced local sensitivity [A]. *Computer Vision and Pattern Recognition Workshops (CVPRW)*,2014 IEEE Computer Society Conference on[C]. Columbus, OH: IEEE, 2014. 414-419.
- [18] Goyette N, Jodoin P, Porikli F, et al. Changedetection.net: A new change detection benchmark dataset [A]. *Computer Vision and Pattern Recognition Workshops (CVPRW)*,2012 IEEE Computer Society Conference on[C]. Providence, RI: IEEE, 2012. 1-8.
- [19] Wang Y, Jodoin P M, Porikli F, et al. CDnet 2014: An expanded change detection benchmark dataset [A]. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*[C]. Columbus, OH: IEEE, 2014. 387-394.

#### 作者简介



曹明伟 男,1986年4月出生于安徽凤阳,现为合肥工业大学计算机与信息学院博士研究生,主要研究方向为目标检测与跟踪,多视图三维重建。  
E-mail:cmwqq2008@163.com



余 焯 (通讯作者) 女,1982年3月出生,安徽安庆人,副教授,硕士生导师,2010年获合肥工业大学工学博士学位。主要研究方向为:场景重建、计算机视觉。  
E-mail:yuyeyue3841@sina.com