

改进的 ViBe 算法及其在运动目标检测中的应用

张钦礼, 邱 杰, 杨秀兰

(玉林师范学院计算机科学与工程学院, 广西, 玉林 537000)

[摘要] 针对 ViBe 算法用于运动目标检测时, 前景提取结果容易出现鬼影、噪音以及检测结果不完整等问题, 本文提出了一种改进的 ViBe 算法. 首先, 利用多帧图像的中位数代替单帧图像对背景模型进行初始化; 第二, 引入分散度系数构造自适应半径阈值代替原来固定的距离判定阈值; 第三, 引入空间一致系数构造动态自适应更新因子代替原来固定更新因子. 实验结果表明: 与传统及改进的 ViBe 算法相比, 本文提出的算法能够适应动态背景及光照变化, 有效抑制了鬼影, 降低了误检率且使检测目标更加完整.

[关键词] ViBe 算法, 中位数, 分散度系数, 空间一致系数

[中图分类号] TN911.73 **[文献标志码]** A **[文章编号]** 1001-4616(2020)04-0104-09

Improved ViBe Algorithm and Its Application in Moving Objects Detection

Zhang Qinli, Qiu Jie, Yang Xiulan

(College of Computer Science & Engineering, Yulin Normal University, Yulin 537000, China)

Abstract: When ViBe algorithm is used to detect moving objects, the results of foreground extraction are prone to ghost, noise and incomplete detection results. To solve these problems, this paper proposes an improved ViBe algorithm. Firstly, the background model is initialized by using the median of multiframe images instead of single frame image; secondly, the dispersion coefficient is introduced to replace the original fixed distance determination threshold with the adaptive radius threshold; thirdly, the spatial consistency coefficient is introduced to construct the dynamic adaptive updating factor instead of the original fixed updating factor. The experimental results show that compared with the traditional ViBe algorithm and some improved ViBe algorithms, our improved ViBe algorithm can adapt to the dynamic background and illumination change, effectively suppress ghost, reduce the false detection rate and make the detection target more complete.

Key words: ViBe algorithm, median, dispersion coefficient, spatial consistency coefficient

随着互联网及人工智能技术的飞速发展, 智能监控设备已经分布我们生活的各个角落, 而智能监控设备的核心技术是运动目标检测. 目前, 主流的运动目标检测算法大致分为帧差法、光流法和背景减除法 3 种^[1-3]. 帧差分法计算量小, 实时性好, 但对缓慢运动的物体效果不佳, 检测到的运动物体往往存在空洞现象^[4]. 光流法能精确计算运动目标的速度和处理目标的旋转, 但对目标运动速度及光线敏感, 计算量大不能直接用于实时场景处理^[5]. 背景减除法是最常用的目标检测方法之一, 其基本原理是为图像中的每个像素建立一个背景模型, 并利用当前点与其背景模型的比较来确定该点是背景点还是前景点. 它算法简单, 计算量小且易于实现, 可用于实时监测, 但其对视频外界环境敏感^[6]. 常用的背景减法主要有高斯混合模型和 ViBe 算法^[7-8]. 高斯混合模型能够适应相对复杂的场景, 对外界干扰具有较强的鲁棒性. 但它需要计算很多高斯模型, 计算量大, 估计参数时间长, 很难满足实时性的要求.

ViBe 算法通常利用视频序列的第一帧完成背景建模, 思想简单, 易于实现, 因而计算量小. 但算法自身也存在着一些局限性, 如在第一帧有前景目标的情况下会出现鬼影现象; 在前景分割阶段, 由于 ViBe 算法采用固定的分割阈值, 因而在静态背景下还能有效地检测出运动目标, 但在动态背景下(如水波、摇摆树枝、喷泉等)会产生大量的误判, 造成检测目标不完整; ViBe 算法以固定的更新率更新背景模型, 当受到外界干扰时, 会出现鬼影、死锁及缓慢移动目标被整合到背景中的情况.

收稿日期: 2020-07-08.

基金项目: 玉林师范学院高层次人才科研启动基金项目(G2018015).

通讯作者: 张钦礼, 博士, 教授, 研究方向: 计算机视觉. E-mail: zqlynu@163.com

针对 ViBe 算法存在的不足,许多学者提出了改进算法.瞿中等提出了一种融合时域信息的自适应 ViBe 算法^[9].该算法虽然实现了半径阈值和更新因子的自适应计算,但在计算中引入了两个经验参数,而这两个参数的值并不好确定.杨丹等提出了一种结合背景差分、帧差法和 OTSU 算法的自适应计算半径阈值的改进 ViBe 算法^[10],但该算法背景初始化采用平均值法并不能完全消除鬼影.张磊等提出了一种基于最小误差阈值的自适应 ViBe 改进算法^[11].该算法通过计算每一帧图像的最小误差分割阈值来对鬼影像素点进行判别并消除,有效抑制了鬼影,但对动态背景效果差.

在分析前人研究工作的基础上,本文提出了一种基于中位数、分散度系数和空间一致系数的改进 ViBe 算法.实验结果表明,改进后的 ViBe 算法有效抑制了鬼影,提高了对动态背景的适应性,使得检测目标更加完整准确.

1 ViBe 算法

ViBe 算法是由 Olivier Barnich 和 Marc Van Droogenbroeck 在 2011 年提出的一种背景建模方法^[8].该算法为运动目标检测领域提供了一种新思路,是一种新颖、快速及有效的运动目标检测算法.该算法采用邻域像素来创建背景模型,通过比对背景模型和当前输入像素值来检测前景,可以细分为 3 个步骤:

(1) 背景模型初始化

ViBe 算法利用单帧视频图像初始化背景模型,对于一个像素点,结合相邻像素点拥有相近像素值的空间分布特性,随机选择邻域点的像素值作为模型样本值.实际上,初始化就是填充像素样本集的过程.由于在一帧图像中不可能包含像素点的时空分布信息,所以只能利用像素点的空间分布信息.

设 $p(x, y)$ 表示 (x, y) 处的像素值,则 $p(x, y)$ 的背景模型 $B(x, y)$ 可定义如下:

$$B(x, y) = \{p_1(x, y), p_2(x, y), \dots, p_N(x, y)\}, \quad (1)$$

式中, $p_i(x, y)$ ($i = 1, \dots, N$) 为随机选择的 $p(x, y)$ 的邻居点.

(2) 前景检测过程

背景模型为每个点存储了一个含有 N 个样本的样本集,然后每个新的像素值和样本集比较判断是否属于背景.具体方法是计算新像素值和样本集中每个样本值的距离,若距离小于阈值 R ,则近似样本点数目增加.如果近似样本点数目大于阈值 T ,则认为新的像素点为背景.所以,检测过程主要由这 3 个参数决定.一般,这 3 个参数分别取值为经验值: $N = 20$ 、 $T = 2$ 、 $R = 20$.

(3) 背景模型更新

背景模型更新策略通常有两种:保守更新策略和盲目更新策略.

保守更新策略:像素样本只有被分类为背景样本时,才能被包含在背景模型中,而前景点永远不能被用于填充背景模型.该方法保证了对移动对象的精确检测,但前提是它们与背景颜色不相似.缺点是会导致死锁情况和鬼影.

盲目更新策略:不管样本是否被归类为背景都被添加到背景模型中.优点是对死锁不敏感;缺点是对运动慢的目标检测能力差,会使慢速运动目标逐渐融入到背景模型中.

ViBe 算法采用的是保守更新策略加前景点计数加随机子采样.所谓前景点计数就是对像素点进行统计,如果某个像素点连续多次被检测为前景,则将其更新为背景点.所谓随机的子采样就是不在每一个新的视频帧中都去更新背景模型,而是当一个像素点被分类为背景点时,它有 $1/\varphi$ 的概率去更新背景模型. φ 为模型固定更新因子,通常取经验值 16.

具体的更新方法:每一个背景点有 $1/\varphi$ 的概率去更新自己的模型样本值,同时也有 $1/\varphi$ 的概率去更新它的邻居点的模型样本值.同时当前景点计数达到临界值时将其变为背景,并有 $1/\varphi$ 的概率去更新自己的模型样本值.在选择要替换的样本集中的样本值时候,也是随机选取一个样本值进行更新.这样做可以充分利用像素值的空间传播特性,使得背景模型逐渐向外扩散.

2 改进的 ViBe 算法

针对 ViBe 算法的缺点,本文提出了一种基于中位数、分散度系数和空间一致性因子的改进 ViBe 算法.首先,采用多帧图像像素值的中位数初始化背景.由于前景目标不断移动,因此每个位置背景占多数,

这能帮助消除鬼影;其次,利用分散度系数自适应计算半径阈值,这能很好地适应动态环境.最后,引入空间一致性因子自适应计算模型更新因子,降低算法的误检率和增强算法的鲁棒性.

2.1 背景初始化

中位数又称中值,是按顺序排列的一组数据中居于中间位置的数,可将数值集合划分为相等的上下两部分.对于有限的数集,可以通过把所有观察值高低排序后找出正中间的一个作为中位数.如果观察值有偶数个,通常取最中间的两个数值的平均数作为中位数.在含有动态目标的视频中,背景占主导地位,选择多帧图像的中位数来初始化背景有助于消除鬼影现象.

首先,选择视频的前若干帧图像求其中位数得到初始背景 IB ;然后,利用该初始背景随机选取邻居像素点完成背景模型的初始化,即点 $p(x, y)$ 的初始化背景模型为:

$$B(x, y) = \{p_1(x, y), p_2(x, y), \dots, p_N(x, y)\}, \quad (2)$$

式中, $p_i(x, y) \in IB$ 且为 $p(x, y)$ 的邻居像素点.

2.2 自适应分割阈值

ViBe 算法使用全局固定分割阈值对像素进行分类.对于静态场景,固定的分割阈值可以很好地对像素进行分类.然而,在动态场景中,全局固定分割阈值会导致检测结果出现较大误差.为了提高 ViBe 算法在动态背景下的鲁棒性,引入自适应机制,提出了一种基于分散度系数的自适应分割阈值.其基本思想是利用分散度系数动态调整分割阈值,即在高动态背景区域适当提高分割阈值,消除高频成分的干扰;在低动态背景区域降低分割阈值,以检测细微变化.

(1) 分散度系数

分散度系数是指一组数据的标准差与其平均值的比值,是衡量数据离散程度的相对指标^[12],其定义如下:

$$d = \sigma / \mu (\mu > 0), \quad (3)$$

式中, σ 表示数据标准差, μ 表示均值.

(2) 自适应分割阈值

在 ViBe 算法中,半径阈值 R 是固定的,经验值通常为 20,灰度图像像素的范围为 $[0, 255]$.依据式(3),本文的自适应分割阈值定义如下:

$$R(x, y) = \begin{cases} \frac{\sigma(x, y)}{\mu(x, y)} R & \mu(x, y) > 0 \\ R & \mu(x, y) = 0 \end{cases}, \quad (4)$$

式中,

$$\mu(x, y) = \frac{v_1(x, y) + v_2(x, y) + \dots + v_N(x, y)}{N}, \quad (5)$$

$$\sigma(x, y) = \sqrt{\frac{(v_1(x, y) - \mu(x, y))^2 + (v_2(x, y) - \mu(x, y))^2 + \dots + (v_N(x, y) - \mu(x, y))^2}{N}}. \quad (6)$$

自适应分割阈值与固定分割阈值的对比如图 1 所示.图 1(a) 为低动态背景固定阈值 R ,图 1(b) 为高动态背景固定阈值 R ,图 1(c) 为高动态背景自适应阈值 $R(x, y)$.其中, v_p 为新帧像素值, R 为固定阈值,

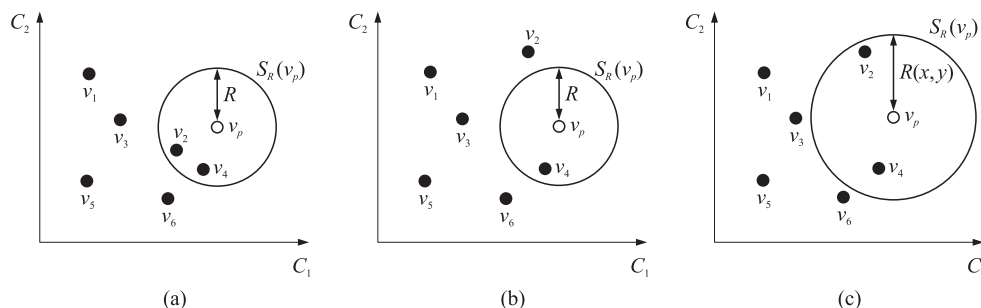


图 1 自适应阈值和固定阈值对不同动态背景的比较图

Fig. 1 Comparison of adaptive threshold and fixed threshold for different dynamic background

v_1, v_2, \dots 为样本集中的像素值, $R(x, y)$ 为自适应阈值, $S_R(v_p)$ 是以 v_p 为圆心, 以 R 或者 $R(x, y)$ 为半径的集合. 当样本集和 $S_R(v_p)$ 的交集大于给的阈值 (通常取经验值 2) 时, 可以认为此像素点为背景像素点, 交集越大, 表示新像素点与样本集越相关. 因此可以通过改变阈值大小来改变模型的灵敏度. 由图 1 可知, 当低动态背景和阈值为固定值 R 时, v_p 为背景点. 而当背景变为高动态背景时, 如果依然取固定阈值 R , 则 v_p 不再为背景点; 如果取自适应阈值 $R(x, y)$, 则 v_p 再次变回背景点.

2.3 自适应更新因子

当背景模型受到外界因素干扰时, 如果采用盲目更新机制, 前景可能被误认为是背景. 如果采用保守的更新机制对模型进行更新, 则很可能使下一帧的前景像素与背景模型相匹配. 如果这种情况继续下去, 将导致一个死循环, 从而增大误检率.

为了降低误检率, 增强算法的鲁棒性, 本文在匹配更新阶段引入空间一致性判断, 即首先定义像素点的邻域集, 然后判断该邻域集合与其背景模型的空间信息一致性, 最后根据该一致性来决定背景模型的更新概率.

设 $P(x_i, y_i)$ 表示任一个视频帧 Z 的任一个像素点, 定义其 δ 邻域如下:

$$LY_p = \{p(x, y) | p(x, y) \in Z, |x - x_i| \leq \delta, |y - y_i| \leq \delta\}, \quad (7)$$

式中, δ 为一正整数.

设 $JJ_p = LY_p \cap B(x, y)$, 则空间一致系数定义如下:

$$SCC(P) = \frac{|JJ_p|}{|LY_p|}, \quad (8)$$

式中, $|\cdot|$ 表示集合元素个数.

$SCC(P)$ 可以作为像素点是背景点可能性的度量. $SCC(P)$ 越大说明 P 的邻域中的像素点是背景模型中的像素点的越多, 则 P 越有可能是背景, 从而 P 的背景模型更新概率就越高. 因此可以利用 $SCC(P)$ 定义如下自适应更新因子:

$$\phi(P) = SCC(P) / \varphi. \quad (9)$$

$\varphi(P)$ 可以自适应的动态调整更新概率, 这可以有效减少前景被误判为背景, 从而增强算法的鲁棒性.

3 实验结果与分析

为了验证本文改进的 ViBe 算法的有效性, 分别选取来自 Change detection 数据集^[9]的 4 个不同视频场景, 包括含有大量动态背景的 Caampus 视频, 前景目标和背景相似、含有光照变化及初始帧含有运动目标的 Escalator 视频、含有动态背景及初始帧含有运动目标的 Pedestrian 视频, 前景目标和背景相似及初始帧含有运动目标的 Hall 视频, 对 ViBe 算法、发表于 2019 年的文献[9]和[11]的改进 ViBe 算法及本文的改进 ViBe 算法 4 种算法进行测试对比. 之所以选择这 4 组视频, 是因为它们情况复杂, 如复杂的动态背景、运动目标和背景相似、初始帧含有多个运动目标、多密集目标、光照变化等, 运动目标检测难度大, 适合测试算法的有效性.

本文以查准率(P)、查全率(R)和 F 度量 (查准率和查全率的调和平均) 作为算法的定量评价指标, 其计算公式如下:

$$P = \frac{TP}{TP + FP}, \quad (10)$$

$$R = \frac{TP}{TP + FN}, \quad (11)$$

$$F = \frac{2 \times P \times R}{P + R}, \quad (12)$$

式中, TP 表示正确检测出的运动目标数目, FP 表示错误检测出的非运动目标数目, FN 表示未检测出的运动目标数目.

实验所涉及的参数分别取值为: $\delta = 2, T = 2, N = 20, n = 8, R = 10, \phi = 8$.

首先, 对本文提出的 3 个改进技术方法分别进行消融实验, 如图 2~4.

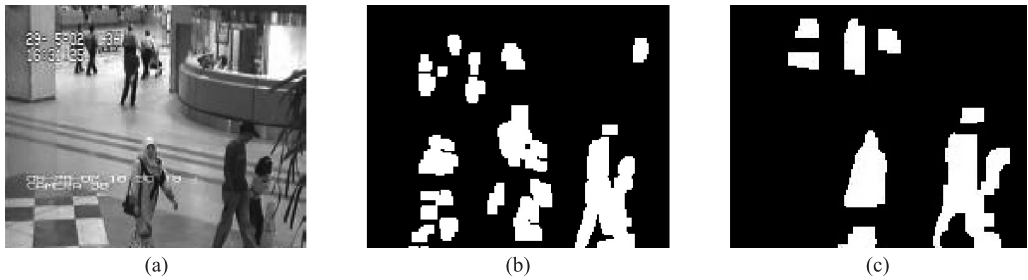


图 2 Hall 视频多帧中位数初始化背景的消融实验结果

Fig. 2 Experimental results of Hall video with background initialization using median of multiframe images

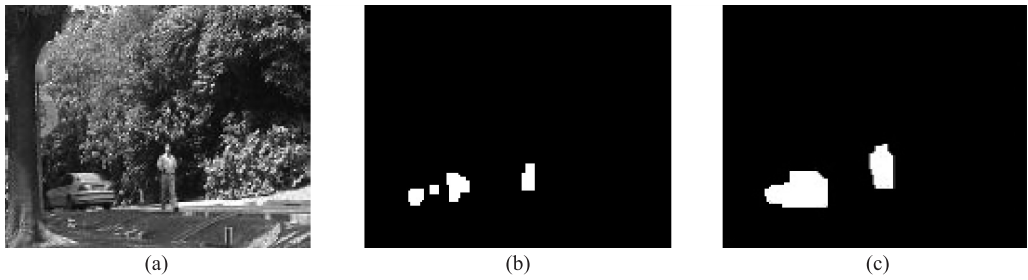


图 3 Caampus 视频自适应阈值的消融实验结果

Fig. 3 Experimental results of Caampus video with adaptive threshold

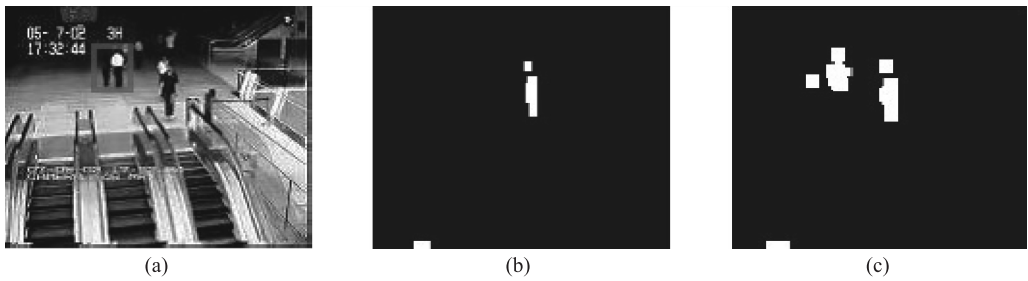


图 4 Escalator 视频自适应更新因子的消融实验结果

Fig. 4 Experimental results of Escalator video with adaptive updating factor

图 2(a)为 Hall 视频的第 19 帧,图 2(b)为原始 ViBe 算法的结果,图 2(c)为将原始 ViBe 算法的背景初始化由单帧改为多帧中位数的结果. 由图 2(b)可以看出,原始 ViBe 算法提取的运动目标出现了鬼影,这是由于 Hall 视频初始帧含有运动目标,原始 ViBe 算法用单帧初始化背景无法处理这种情况. 由图 2(c)可以看出,当背景初始化由单帧改为多帧中位数后,鬼影消失了,这说明用多帧中位数初始化背景对初始帧含有运动目标的视频是有效的.

图 3(a)为 Caampus 视频的第 828 帧,图 3(b)为原始 ViBe 算法的结果,图 3(c)为将原始 ViBe 算法的固定阈值改为自适应阈值的结果. 由图 3(b)可以看出,原始 ViBe 算法提取的运动目标很不完整,这是由于 Caampus 视频含有大量动态背景,原始 ViBe 算法采用固定阈值对动态背景适应性差. 由图 3(c)可以看出,ViBe 算法固定阈值变为自适应阈值后,提取的运动目标相对完整,这说明本文设计的自适应阈值对含有动态背景的视频是有效的.

图 4(a)为 Escalator 视频的第 466 帧,图 4(b)为原始 ViBe 算法的结果,图 4(c)为将原始 ViBe 算法的固定更新因子改为自适应更新因子的结果. 由图 4(b)可以看出,原始 ViBe 算法未能提取红框内的运动目标,这是由于红框内的运动目标为缓慢移动目标,原始 ViBe 算法采用固定更新因子不能处理缓慢移动目标. 由图 4(c)可以看出,固定更新因子变为自适应更新因子后,成功提取了红框内缓慢移动目标的一部分,这说明本文设计的自适应更新因子对含有缓慢移动目标的视频是有效的.

图 5 和表 1 给出了 4 种算法在 4 个不同场景视频测试实验中的 3 种评价指标的平均值.

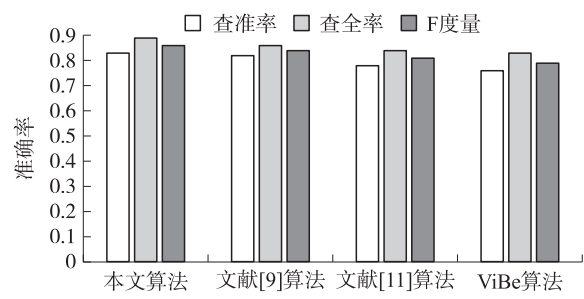


图 5 4 种算法 3 种评价指标对比

表 1 不同算法不同视频场景 3 种评价指标的平均值
Table 1 Average values of three evaluation indexes for different video scenes with different algorithms

算法	查准率 P	查全率 R	F 度量
ViBe	0.76	0.83	0.79
文献[9]算法	0.82	0.86	0.84
文献[11]算法	0.78	0.84	0.81
本文算法	0.83	0.89	0.86

Fig. 5 Comparison of three evaluation indexes of four algorithms

由图 5 和表 1 可以看出,本文、文献[9]及[11]的改进 ViBe 算法均提高了 ViBe 算法的查准率、查全率及 F 度量,但本文算法提高幅度最大,在各项指标均是最高. 因此,本文算法优于原始 ViBe 算法及文献[9]和[11]的改进 ViBe 算法,明显提高了 ViBe 算法的综合性能.

图 6~9 给出了 4 种算法对 4 个视频部分帧的检测结果.

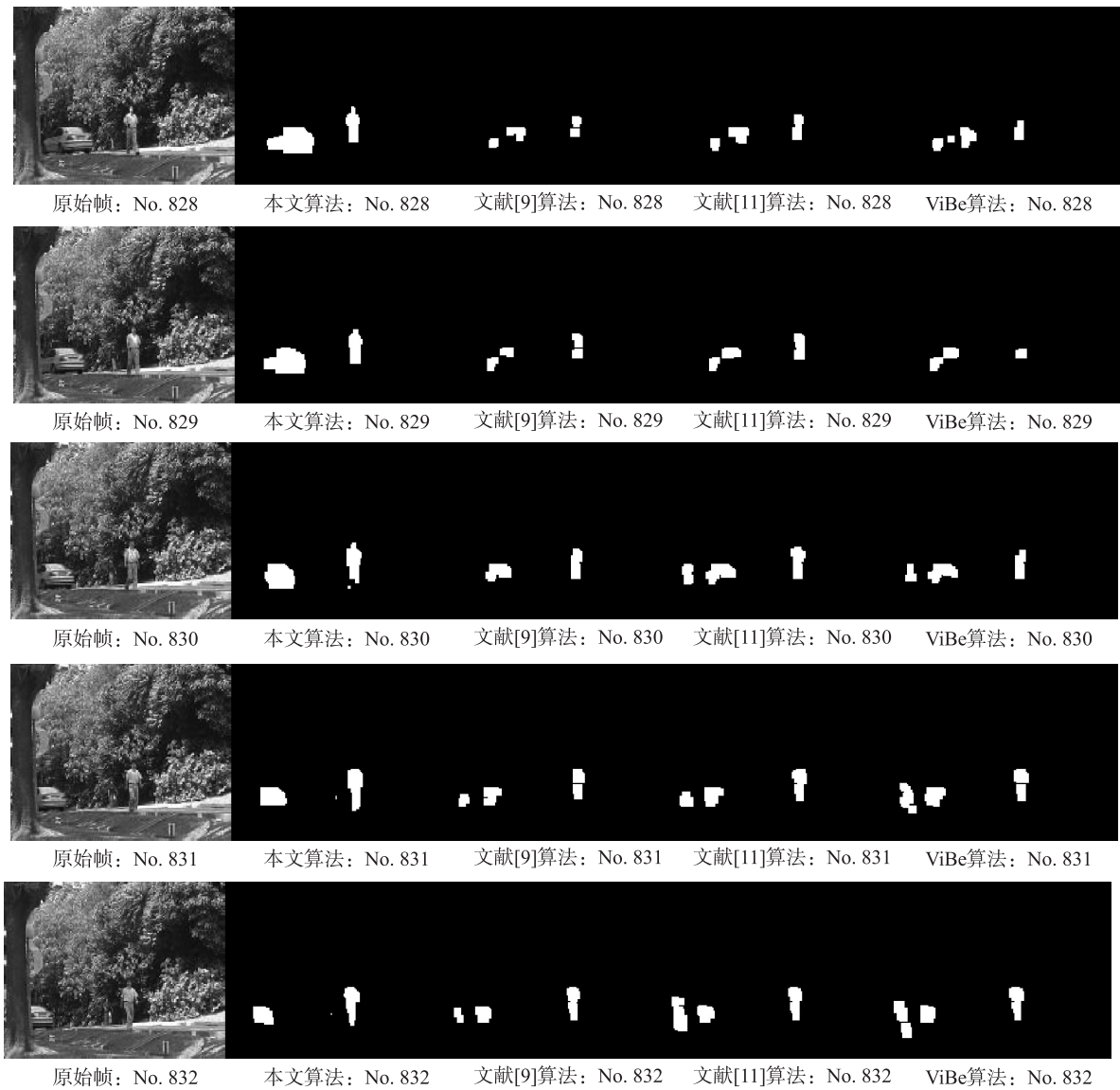


图 6 4 种算法对 Campus 视频中部分帧的检测结果

Fig. 6 Detection results of some frames in campus video by four algorithms

图 6 中,视频 campus 含有大量动态背景且初始帧不含运动目标. 对比各算法的检测结果可以看出,ViBe 算法、文献[9]和[11]的算法的检测结果不准确不完整,而本文改进的 ViBe 算法可以较完整地检测出运动目标,噪声少,且没有出现多检和漏检现象,但也出现了一处运动目标没有分开的情况.

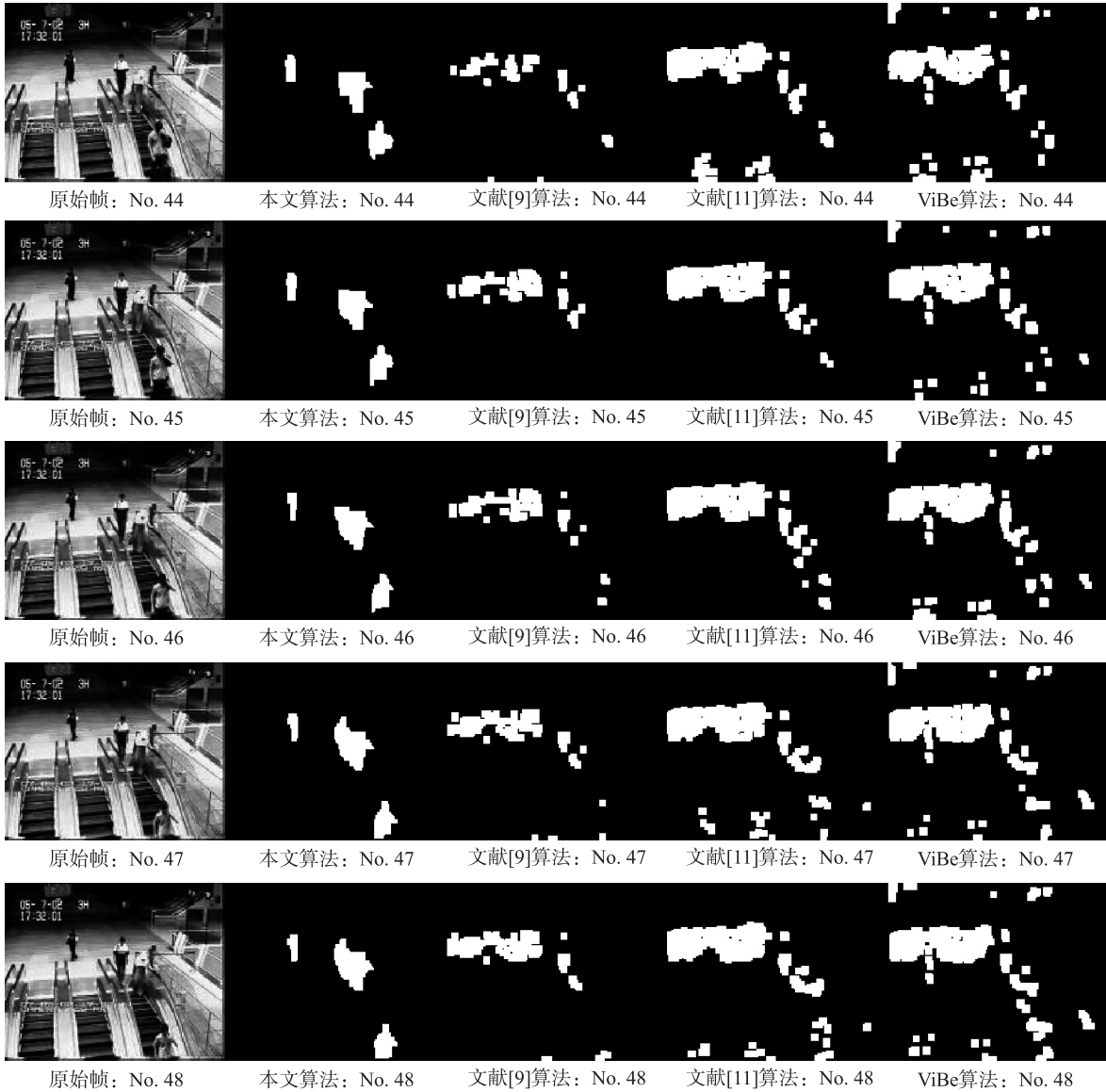


图 7 4 种算法对 Escalator 视频中部分帧的检测结果

Fig. 7 Detection results of some frames in escalator video by four algorithms

图 7 中,视频 Escalator 初始帧含有运动目标,前景目标与背景相似,光线不好,存在光照变化且整个视频前景目标比较小. 文献[9]和[11]的算法及 ViBe 算法存在很多噪音且几乎检测不到运动目标. 对比而言,本文算法没有噪音,比较准确地检测到了运动目标,只是存在一处未能完全分离的运动目标.

图 8 中,视频 Pedestrian 初始帧含有运动目标,有动态背景,运动目标很多且很小,有的地方运动目标密集. 文献[9]和[11]的算法及 ViBe 算法存在很严重的鬼影、误检、漏检及密集目标无法分开的问题. 相比而言,本文算法无鬼影现象,几乎没有漏检和误检,只是存在两处未能分开的密集目标.

图 9 中,视频 Hall 初始帧含有运动目标,前景目标与背景相似,运动目标比较小,有重叠,有一处比较密集. 文献[9]和[11]的算法及 ViBe 算法存在大量噪音且有鬼影现象. 相比而言,本文算法无鬼影且几乎无噪音,但也未能分清密集目标和重叠目标.

综合以上分析可知本文算法有效抑制了鬼影,大幅降低了噪音,对动态背景及光照变化具有很好的鲁棒性.

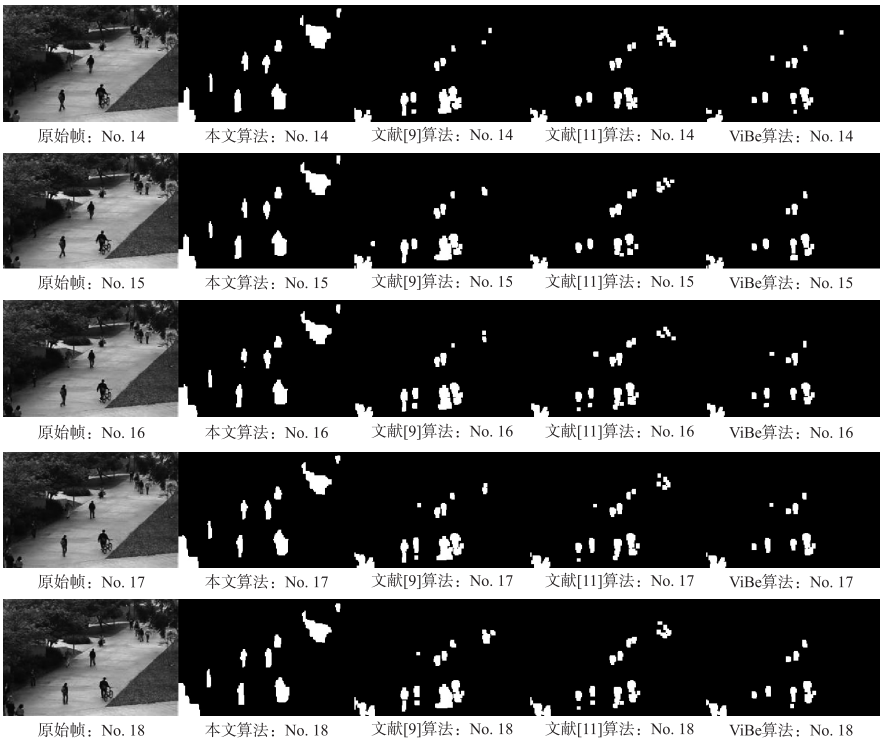


图 8 4 种算法对 Pedestrian 视频中部分帧的检测结果

Fig. 8 Detection results of some frames in pedestrian video by four algorithms

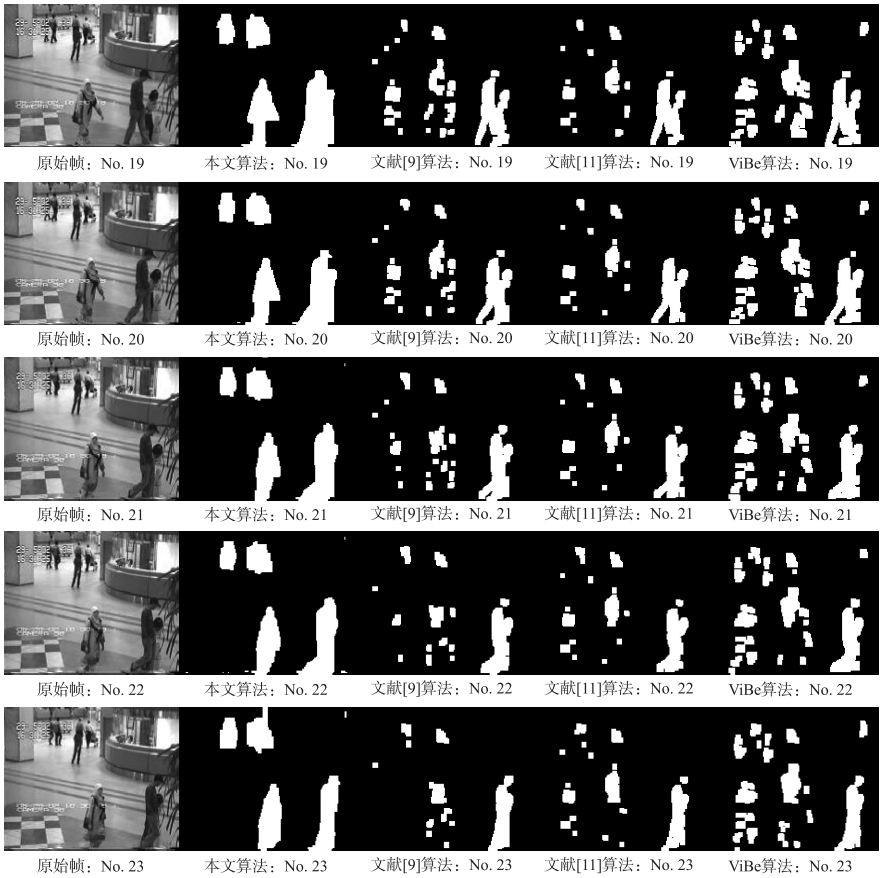


图 9 4 种算法对 Hall 视频中部分帧的检测结果

Fig. 9 Detection results of some frames in hall video by four algorithms

4 结论

本文提出了一种基于中位数、分散度系数及空间一致系数的 ViBe 改进算法. 首先,采用多帧连续图像的中位数代替单帧图像对背景进行初始化,减少了鬼影对背景模型的干扰;然后,利用分散度系数建立自适应分割阈值,提高了模型对动态背景的适应性. 最后,加入了空间一致性判断,降低了误判率. 总之,本文改进的 ViBe 算法抗干扰性强鲁棒性好,检测结果更准确,具有很好的适用性.

[参考文献]

- [1] 邱联奎,刘启亮,雷文龙. 基于背景减除与三帧差分相融合的运动检测[J]. 合肥工业大学学报(自然科学版),2014,37(5):572-577.
- [2] BARRON J L, FLEET D J, BEAUCHEMIN S. Performance of optical flow techniques[J]. International journal of computer vision, 1994, 12(1):43-77.
- [3] 刘仲民,何胜蛟,胡文瑾,等. 基于背景减除法的视频序列运动目标检测[J]. 计算机应用,2017,37(6):1777-1781.
- [4] 李清瑶,邹皓,赵群,等. 基于帧间差分自适应法的车辆抛洒物检测[J]. 长春理工大学学报(自然科学版),2018,41(4):108-113.
- [5] 李秋宇,张玉明,杨福猛,等. 一种视频微表情检测的改进光流算法[J]. 图学学报,2018,39(3):448-452.
- [6] BHARTI, THIND T. Background subtraction techniques-review[J]. International journal of innovative technology & exploring engineering, 2013, 2(3):3099-3104.
- [7] EVANGELIO R H, PATZOLD M, KELLER I, et al. Adaptively splitted gmm with feedback improvement for the task of background subtraction[J]. IEEE transactions on information forensics & security, 2014, 9(5):863-874.
- [8] 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.
- [9] 瞿中,刘帅,刘妍. 融合时域信息的自适应 ViBe 算法[J]. 计算机工程与设计,2019,40(3):189-194.
- [10] 杨丹,戴芳. 运动目标检测的 ViBe 算法改进[J]. 中国图象图形学报,2018,23(12):41-56.
- [11] 张磊,孟亮. 基于最小误差阈值自适应的 ViBe 改进算法[J]. 现代电子技术,2019,42(4):169-173.
- [12] ZHAO W, WANG L, ZHANG Z. A novel atom search optimization for dispersion coefficient estimation in groundwater[J]. Future generation computer systems, 2019, 91(2):601-610.

[责任编辑:顾晓天]