

文章编号 1004-924X(2014)09-2545-08

动态背景下基于改进视觉背景提取的前景检测

陈星明*, 廖娟, 李勃, 陈启美

(南京大学电子科学与工程学院, 江苏南京 210046)

摘要: 由于视觉背景提取算法(ViBe)对存在动态背景的户外视频的前景检测结果依然不够精确,故提出了一种改进的 ViBe 算法。文中描述了经典 ViBe 算法及其特点;介绍了改进的 ViBe 算法针对动态背景的改进措施。该算法采用多帧连续图像初始化背景模型,降低了单帧图像初始化所产生的“鬼影”对前景检测精度的影响;在匹配过程中,引入自适应的匹配阈值,克服了单个的全局阈值对动态背景适应能力差的问题;最后,在更新过程引入空间一致性判断与模糊准则来减少算法的误检,提高了算法的鲁棒性。实验结果表明,该算法可以有效地检测动态背景下的运动目标,检测准确率比经典 ViBe 算法提高了 20% 以上。

关键词: 前景检测;视觉背景提取(ViBe);动态背景;自适应阈值;空间一致性;模糊准则

中图分类号: TP391.4 **文献标识码:** A **doi:** 10.3788/OPE.20142209.2545

Foreground detection based on modified ViBe in dynamic background

CHEN Xing-ming*, LIAO Juan, LI Bo, CHEN Qi-mei

(College of Electronic Science and Engineering, Nanjing University, Nanjing 210046, China)

* Corresponding author, E-mail: chris_cxm@163.com

Abstract: As Visual Background Extractor(ViBe) can not implement foreground detection precisely for a particular scene with dynamic backgrounds, This paper proposes a modified ViBe algorithm. It describes the original ViBe algorithm and its characteristics and discusses several modification schemes for the original ViBe in dynamic background scenes. Firstly, model initialization is conducted with several continuous frames instead of one single frame to handle ghosts. Then, self-adaptive threshold is adopted in the process of model matching so that background models is better suitable for the dynamic background. Finally, a spatial coherence estimation and a fuzzy rule in model maintenance are proposed to reduce false detections and to improve the robustness of the algorithm. Experiments demonstrate that the algorithm proposed detects effectively the movement targets in dynamic background scenes and its precision is improved by 20 percent as compared with that of the original ViBe algorithm.

Key words: foreground detection; Visual Background Extractor(ViBe); dynamic background; self-adaptive threshold; spatial coherence; fuzzy rule

收稿日期:2014-01-17;修订日期:2014-03-18.

基金项目:交通运输部科技项目(No. 2012-364-641-209),国家自然科学基金资助项目(No. 61105015),江苏省自然科学基金资助项目(No. BK2011511)

1 引言

运动目标检测是计算机视觉应用中的关键技术,在智能视频监控、图像压缩等领域有重要的研究价值。目前常见的运动目标检测算法有:光流法^[1]、背景差分法^[2-4]、帧间差分法^[5-6]等,其中:背景差分法比较常用,且实时性较好,但其检测性能很大程度上依赖于背景模型的准确性。影响背景模型准确性的因素有很多,包括动态背景、光线渐变、相机抖动、阴影等,其中动态背景是最常见且影响最大的因素。

为了建立适应动态背景的有效背景模型,研究人员提出了不同的背景建模方法。Stauffer 等人^[7-8]提出了混合高斯函数算法(Mixed of Gaussian Function, MOG),该算法用多个高斯模态描述背景模型,以克服单高斯模型^[9]的缺点,从而提高算法对动态背景的适应能力,但选择学习速率时,无法兼顾动态背景抑制和正确前景提取。Maddalena 等人^[10-12]提出了基于人工神经网络的自组织背景减法模型(Self Organizing Background Subtraction, SOBS),即通过自组织的方式学习运动信息,能够处理光线变化、遮挡、动态背景等复杂场景,但是运算代价比较大。Barnich 等人^[13]提出了基于像素的非参数化随机样本模型(Visual Background Extractor, ViBe),它采用像素样本值建立背景模型,先将检测帧的像素值与对应模型进行匹配,再通过固定阈值判断其是属于前景还是背景;对于匹配上的像素,采用随机更新机制更新该像素及其邻域的背景模型;该方法运算简单,在静态背景下有不错的检测效果,但其参数固定,这限制了算法对于动态背景(水面波纹、树叶晃动等)的自适应能力,其邻域扩散更新策略会造成运动较慢的前景目标过快地融入背景,从而增加了错误检测;其单帧输入图像初始化策略在输入图像含有前景目标的情况下会产生“鬼影”空洞,会影响背景模型的准确性。针对 ViBe 的不足, Droogenbroeck 等人在文献^[14]中提出了改进措施,包括抑制邻域扩散、引入距离函数以及检测闪烁像素。这些改进措施解决了前景融入过快的问题,降低了误检率,但各个改进措施的关联性和针对性不强,对动态背景的适应性较差。因此,文中对 ViBe 初始化、匹配及更新过程

进行了相关改进,提出了一种改进的 ViBe 算法,极大地提高了算法对动态背景的适应能力。

本文的后续框架构建如下:第 2 节描述了经典 ViBe 算法及其特点;第 3 节详细介绍了针对动态背景的改进措施,包括初始化、自适应阈值、空间一致性以及模糊准则;第 4 节给出了本文算法与经典算法的测试结果比较;第 5 节是全文总结。

2 Vibe 算法及其背景模型

Vibe 算法采用背景帧的像素样本构成像素的背景模型。对于任意一像素点 x , 设 $v(x)$ 为当前帧像素值, v_i 为背景样本值, 则背景模型定义为: $M(x) = \{v_1, \dots, v_i, \dots, v_N\}$ 。如图 1 所示, 定义一个以 $v(x)$ 为中心, 以 R 为半径的球体 $S_R(v(x))$ 。 $S_R(v(x))$ 表示所有与 $v(x)$ 距离小于 R 的点的集合, 用 $M(x)$ 落在球体 $S_R(v(x))$ 内的样本个数 $\#$ 来描述 $v(x)$ 与背景模型 $M(x)$ 的相似度。对于给定阈值 $\#_{\min}$, 如果 $\# < \#_{\min}$, 则 $v(x)$ 为前景; 如果 $\# > \#_{\min}$, 则 $v(x)$ 为背景。

Vibe 算法具有以下特点:

(1) 单帧图像初始化。对第一帧图像进行瞬间初始化, 从第二帧开始算法就可以有效检测运动目标, 但它从每个像素点的邻域采集背景样本, 对于初始帧存在运动前景的情况, 它会将前景像素加入到背景模型中, 从而会产生“鬼影”。

(2) 保守更新机制。传统的更新方式有盲更新和保守更新。盲更新是将新进样本都加入背景模型, 而保守更新只利用已判断为背景的像素点来更新背景模型, 从而降低了前景被误检为背景的可能性。但如果噪声影响使前景像素被误检为背景, 且用来更新背景模型, 则会造成后续误检。

(3) 随机更新机制。采用随机更新机制更新背景模型, 使匹配上的像素值以 $1/\varphi$ 的概率替代背景样本中的任一样本值, φ 为二次抽样时间因子。该机制保证了背景模型中样本生命期的平滑指数衰减, 从而使一个包含有限样本的背景模型能有效适应场景的变化, 但随机更新也可能替换掉匹配上的背景样本, 从而影响背景模型的准确性。

(4) 样本邻域传播。采用空间信息传播机制将背景信息扩散到相邻像素, 如果当前像素与背景模型可以匹配上, 则对其某一邻域像素的背景

模型也进行随机更新。这种样本邻域传播机制有效保证了空间一致性,但对于前景、背景交界处的像素会造成信息的错误传播,产生误检。

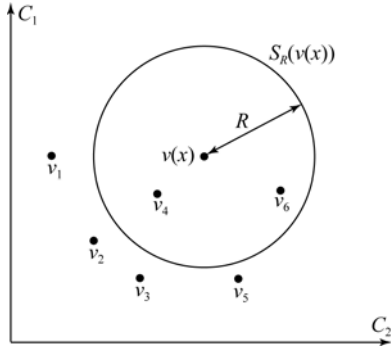


图1 二维欧式空间匹配示意图

Fig.1 Diagram of model matching system in two-dimensional European space

3 改进的 ViBe 算法

3.1 模型初始化

ViBe 算法假设相邻像素点具有相似的时空分布。其首先采用第一帧图像初始化背景模型,接着,从每个像素点的邻域中随机选择 N 个像素值作为该像素点的背景样本,若第一帧图像中存在前景,则前景像素值也会作为背景样本点加入到初始背景模型中,当运动前景移走后,就会导致“鬼影”,即前景覆盖处背景像素集合被误检为前景目标。为了降低前景点对初始背景模型的影响,文中采用时间一致性原则,即对于任一像素点 x ,采用其前 N 帧图像的像素值初始化背景模型 $M(x)$:

$$M(x) = \{v_1, \dots, v_N\} = \{I_{t-(N-1)}(x), \dots, I_t(x)\}, \quad (1)$$

式中: $I_t(x)$ 为 t 时刻的像素值。对于第一帧存在运动前景的情况,初始化的样本集中将会存入更多前景移动走后的背景信息,使得“鬼影”能够迅速地消除。

3.2 模型匹配

对于任一像素点 x ,假设 $I_t(x)$ 为当前像素值, $M(x) = \{v_1, \dots, v_i, \dots, v_N\}$ 为其对应的背景模型,则前景二值图 $F(x)$ 可表示为:

$$F(x) = \begin{cases} 1 & \text{if } \# \{M(x) \cap S_R(I_t(x))\} < \#_{\min}; \\ 0 & \text{else} \end{cases}; \quad (2)$$

$$\# \{M(x) \cap S_R(I_t(x))\} = \begin{cases} 1 & \text{if } \text{dist}(v_i(x), I_t(x)) < R, \\ 0 & \text{else} \end{cases}, \quad (3)$$

式中: $\text{dist}(\cdot)$ 表示欧氏距离。

在 ViBe 中,采用全局固定的匹配阈值 R 。这对于静态背景是合适的,但实际场景并非是绝对静态的,对于高动态区域(如晃动的树叶、闪烁的水面),应增大阈值 R ,以使背景不被检测为前景;对于低动态区域,应降低阈值 R ,以能检测出细微变化的前景。因此,文中根据背景的动态程度自适应地设置阈值 R ,定义一个最小距离的集合 $D(x)$:

$$D(x) = \{D_1(x), \dots, D_k(x), \dots, D_N(x)\}, \quad (4)$$

式中: $D_k(x) = \min \{\text{dist}(I_t(x), v_i(x))\}$, 表示当前像素 $I_t(x)$ 与其对应背景样本 $v_i(x)$ 的最小欧氏距离; \min 表示最小距离对应的样本编号 i ; 用 N 个 $D_k(x)$ 的平均值 $d_{\min}(x)$ 表征背景的动态运动情况,定义为:

$$d_{\min}(x) = \frac{1}{N} \sum_k D_k(x), \quad (5)$$

即每次 $I_t(x)$ 与 $v_i(x)$ 匹配成功以后都记录一个 $d_{\min}(x)$ 。对于静态背景, $d_{\min}(x)$ 趋于稳定;对于动态背景, $d_{\min}(x)$ 则较大。通过 $d_{\min}(x)$ 的值实现 $R(x)$ 的自适应更新,如下式:

$$R(x) = \begin{cases} R(x) \cdot (1 - \alpha_{\text{dec}}) & \text{if } R(x) > d_{\min}(x) \cdot \zeta, \\ R(x) \cdot (1 + \alpha_{\text{inc}}) & \text{else} \end{cases}, \quad (6)$$

式中: α_{dec} 、 α_{inc} 和 ζ 都是固定的参数。对于不变的 $d_{\min}(x)$, $R(x)$ 最终将趋向于稳定值 $d_{\min}(x) \cdot \zeta$; 当出现动态背景扰动时, $R(x)$ 会慢慢增加,来适应这一扰动。

如图 2 所示,图 2(a) 为 Canoe 视频第 20 帧的输入图像,图 2(b) 为 ViBe 算法检测后的二值图。由图 2(b) 可知,由于水面动态变化很明显,水波纹被误检为前景,极大地影响了前景目标检测的准确性。因此,文中对动态的水面区域使用较大的阈值 R ,来抑制水波纹的误检。图 2(c) ~ 2(e) 为引入自适应阈值 R 后第 30、40、50 帧的二值图,通过与图 2(b) 比较可以看出,引自适应阈值 R 提高了算法对水面动态变化的适应性,降低了水波纹对前景检测的影响。

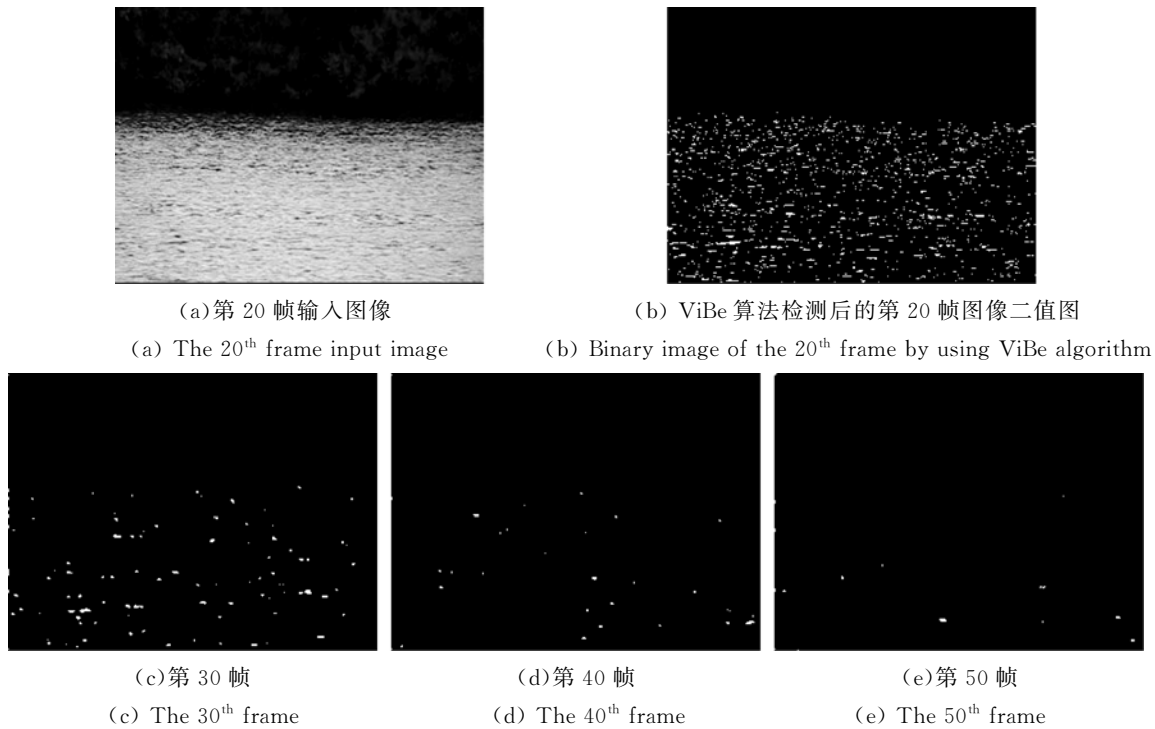


图 2 使用 ViBe 算法及加入自适应 R 后 Canoe 视频的检测二值图

Fig. 2 The 20th frame Canoe sequence and binary images by using ViBe algorithm and added adaptive R

3.3 模型更新

如第 2 节所述,当背景模型受外在因素干扰(如噪声)时,前景会被误检为背景,此时若采用保守更新机制更新背景模型,则会使下一帧前景像素更容易与背景模型匹配上,如此下去就会形成死循环,造成误检率增大。为了降低算法的误检率,提高算法的鲁棒性,本文在匹配更新过程中加入空间一致性判断,即利用邻域像素的匹配情况协助判断中心像素背景模型描述是否准确,从而降低前景像素被误检的可能性。

对于视频帧 I 中的任一像素点 $x(x_i, x_j)$,定义其 $k \times k$ 邻域为:

$$N_x = \{y = (y_i, y_j) \in I : |x_i - y_i| \leq k, |x_j - y_j| \leq k\}. \quad (7)$$

定义集合 Ω_x 为 N_x 中与背景模型匹配上的像素:

$$\Omega_x = \{y \in N_x : \#\{M(y) \cap S_R(I(y))\} < \#_{\min}\}. \quad (8)$$

依据文献[15]可知,定义邻域一致性因子为:

$$NCF(x) = \frac{|\Omega_x|}{|N_x|}, \quad (9)$$

式中: $|\cdot|$ 表示集合基数, $NCF(x)$ 可以作为衡量背景模型正确性的参数,如果 $NCF(x) > 0.5$, 则

说明 x 的邻域中大多数像素点都能够被背景模型描述,则像素点 x 也能够被该背景模型描述,故应该更新其背景模型。

3.4 模糊准则

在背景更新过程中,一般遵循这样的准则:如果满足条件 A,那么执行操作 B,但不能完全满足条件 A 时,是否执行操作 B 就要视条件 A 的满足程度而定,其越趋向于满足条件 A,那么越趋向于执行操作 B。例如,对于条件“一个城市天空比较蓝,河水比较清,空气比较清新”,得到结论“这座城市环境比较好”。这里,“比较蓝”、“比较清”、“比较清新”就是语言标签,模糊集理论提供了一个有效表达这些标签之间固有模糊性的机制。模糊集理论的灵活性和有效性使得基于模糊的系统较传统系统在信息表征方面有更大的吸引力。

在 ViBe 中,背景是否更新取决于像素 x 是否与背景模型 $M(x)$ 匹配,但是如果噪声使得前景像素被误检为背景,并且背景模型被更新了,就会降低模型的准确性,产生更多误检。因此,在背景更新过程中引入模糊理论,定义模糊系统如下:

如果“ x 与 $M(x)$ 匹配”并且“ $NCF(x)$ 比较大”,则更新 $M(x)$, 且二次抽样时间因子 $\varphi =$

$F_1(x) \times \varphi$, 其中 $F_1(x)$ 为模糊函数, 定义为:

$$F_1(x) = \begin{cases} 1/(2 \times NCF(x)) & NCF(x) \geq 0.5 \\ 0 & \text{else} \end{cases} \quad (10)$$

根据式(9)可知, $NCF(x)$ 越大, 则 x 邻域中能被背景模型准确描述的像素点越多, 则用 x 来更新背景模型的概率就应适当增大, 如式(10)所示。依据 $NCF(p)$ 自适应调整背景更新概率, 以提高背景模型对动态背景的鲁棒性。

4 测量实验与结果

测试硬件平台的 CPU 为 Intel Core i5, 内存为 4 GB, 操作系统为 Linux Suse 12.1。实验数据来源于 www.changedetection.net 网站提供的 5 个动态背景场景的视频序列: Canoe、Fall、Fountain01、Fountain02 和 Overpass。所有实验中, 样本个数 $N=20$; 最小匹配个数 $\#_{\min}=2$; 初始距离阈值 $R=20$; 二次时间采样因子 $\varphi=20$; 自增适应参数 $\alpha_{inc}=0.05$; 尺度因子 $\zeta=5$; 自减适应参数 $\alpha_{dec}=0.5$ 。这里, 在动态背景区域出现真实运

动前景时, 为了使距离阈值 R 能够快速减小以减少漏检的概率, 本文设置自减适应参数 α_{dec} 为自增适应参数 α_{inc} 的 10 倍。另外, 为了防止 R 过低产生不必要的误检, 设置距离阈值下限 $R_{\text{bottom}}=15$ 。

图 3 为 Canoe 视频序列的检测结果。以列为单位, 从左至右分别为: 当前帧、真实值、ViBe 和本文改进算法的检测结果。由图 3 可以看出, ViBe 算法一直将水面波纹检测为前景, 而改进 ViBe 算法随着 R 的自适应, 在 20~50 frame 的过程中, 渐渐适应了动态背景的扰动。另外, 空间一致性的判断也大大减少了前景的误检。当出现运动前景时(968 帧), 阈值 R 迅速自适应减小, 使得正确的前景不会被漏检, 与原始 ViBe 算法相比, 改进的 ViBe 算法取得了更好的检测效果。

另一组场景(Fountain02)下的对比结果如图 4 所示。由图 4 可以看出, ViBe 算法将喷泉四散的水花误检为运动前景, 而改进后的算法很好地去除了错误前景的干扰, 并在第 744 帧和第 1 275 帧完整地提取出了真实的运动前景, 获得了更优的检测效果。

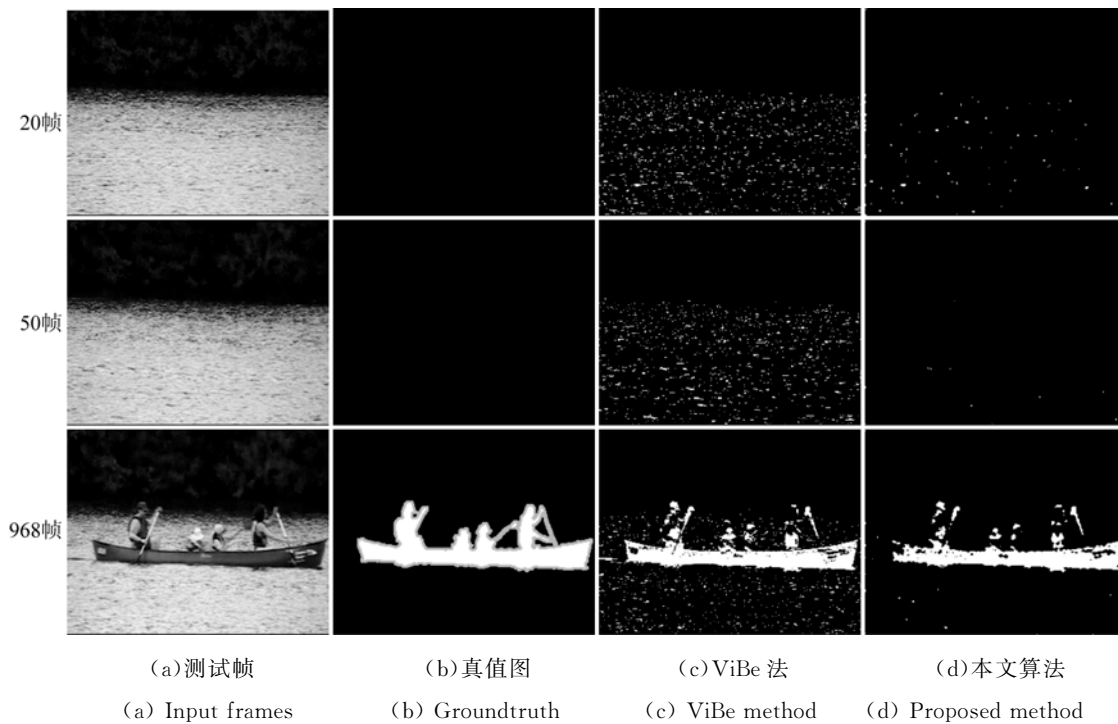


图 3 Canoe 视频序列的测试结果比较

Fig. 3 Comparison of test results of ViBe and proposed method in Canoe sequence

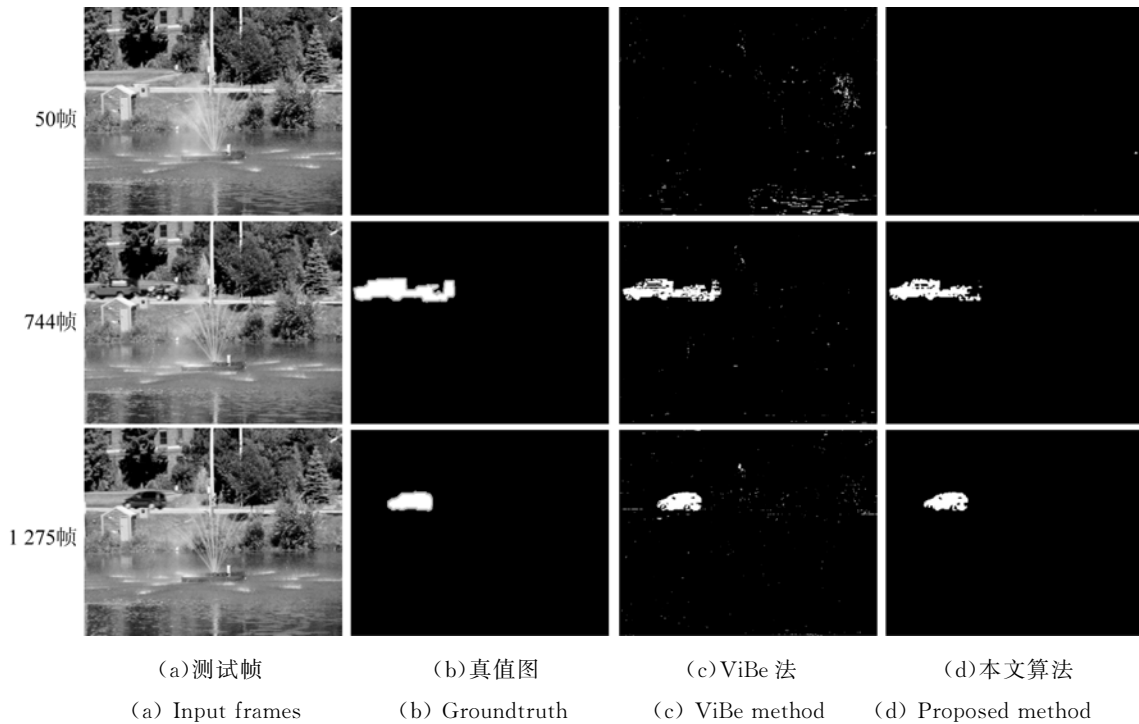


图 4 Fountain02 视频序列的测试结果比较

Fig. 4 Comparison of test results of ViBe and proposed method in Fountain02 sequence

图 5 给出了 MOG 算法、SOBS 算法、ViBe 算法和本文算法在其余 3 组场景 (Fall, Fountain01, Overpass) 下的检测效果。可以发现,在 Fall 场景下,MOG、SOBS 及 ViBe 算法都将剧烈晃动的树枝树叶误检为运动前景,而本文算法则只将运动车辆检测为前景;在 Fountain01 场景下,前 3 种算法都将喷泉检测为前景,而本文算法却很好地抑制了喷泉的干扰;在 Overpass 场景下,由于树叶晃动幅度不大,多模态的混合高斯算法 MOG 与本文算法都能较好地提取出正确的运动前景,SOBS 和 ViBe 算法的检测结果则差一些,有较多的错误前景点。

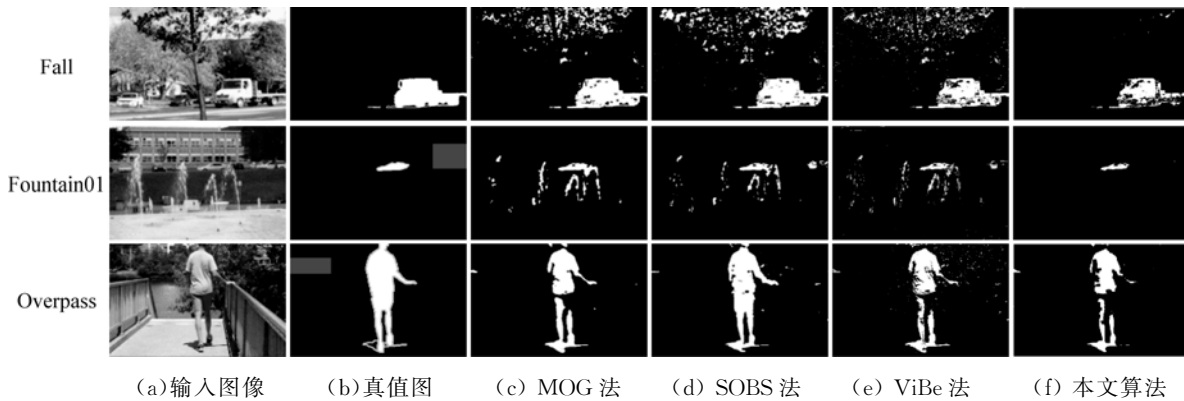
为了定量比较几种算法的性能,采用准确率 (Precision) 和召回率 (Recall) 作为量化指标。它们定义如下:

$$\text{Recall} = \frac{TP}{TP + FN}; \quad (11)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (12)$$

其中: TP 表示正确检测的前景数; FP 表示错误检测的前景数; FN 表示错误检测的背景数。Recall 反映了正确检测的像素占 Groundtruth 中正确前景像素的百分比; Precision 反映了正确检测的像素占检测出的前景总像素的百分比。

图 6 给出了 Fall 场景下的 4 种算法的 Precision & Recall 直方图。由于本文算法大大降低了动态背景引起的错误前景检测数 (FP), 因此在 Precision 指标上明显高出其他 3 种算法, 比第二名的 MOG 算法高出将近 20 个百分点。在 Recall 指标上, 本文算法也高于原始的 ViBe 算法, 与 SOBS 以及 MOG 基本持平。综合两项指标可以得出以下结论: 本文算法在动态背景场景下相对其他算法具有明显的优势。



(a) 输入图像 (b) 真值图 (c) MOG 法 (d) SOBS 法 (e) ViBe 法 (f) 本文算法
 (a) Input frames (b) Groundtruth (c) MOG method (d) SOBS method (e) ViBe method (f) Proposed method

图 5 4 种背景建模算法在动态背景场景下的检测结果比较

Fig. 5 Comparison of test results of four background extraction algorithms in dynamic scenes

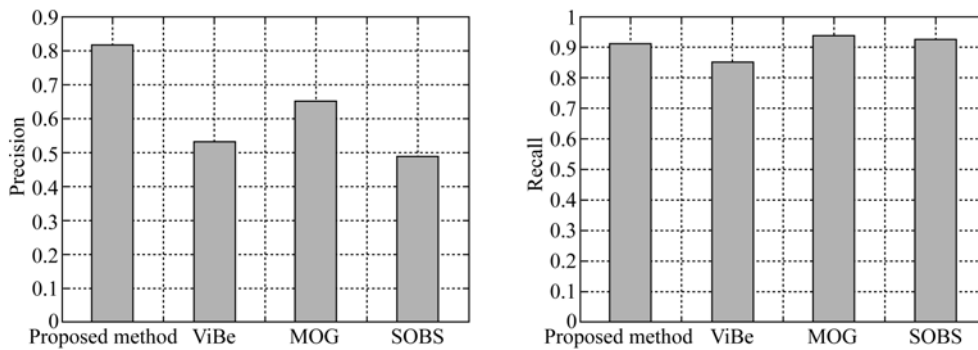


图 6 Fall 场景下 4 种算法的 Precision & Recall 直方图

Fig. 6 Precision & Recall histogram of four algorithms in Fall sequence

5 结 论

本文提出了一种动态背景下基于 ViBe 的改进算法。该算法利用多帧连续图像代替单帧图像进行初始化,从而有效降低了“鬼影”对背景模型

的影响;通过引入自适应阈值 R 大大改善了模型对动态背景的适应能力;同时引入空间一致性判断,并结合模糊准则,明显降低了误检率,提高了算法的鲁棒性。实验证明,在动态背景下本文算法较原始 ViBe 算法的检测准确率提高了 20% 以上。

参考文献:

[1] KERANER J K, THOMPSON W B, BOLEY D L. Optical flow estimation: an error analysis of gradient-based methods with local optimization[J]. *IEEE Transaction on Patten and Analysis Machine Intelligence*, 1987, 9(2):229-244.

[2] BRUTZER S, HOFERLIN B, HEIDEMANN G. Evaluation of background subtraction techniques for video surveillance [C]. *Computer Vision and Pattern Recognition*, 2011:1937-1944.

[3] BOUWMANS T. Recent advanced statistical background modeling for foreground detection: a systematic survey[J]. *Recent Patents on Computer Science*, 2011, 4(3):146-176.

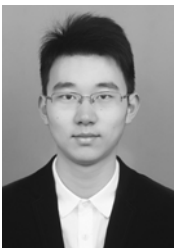
[4] 朱娟娟, 郭宝龙. 复杂场景中基于变块差分的运动目标检测[J]. *光学精密工程*, 2011, 19(1): 184-191.

ZHU J J, GUO B L. Moving object detection based on variant block difference in complex scenes [J]. *Opt. Precision Eng.*, 2011, 19(1):184-191. (in Chinese)

[5] 李刚, 曾锐利, 林凌, 等. 基于帧间颜色梯度的背景建模[J]. *光学精密工程*, 2007, 15(8):

- 1258-1262.
- LI G, ZENG R L, LIN L, *et al.*. A novel background modeling method based on color grade of different frames [J]. *Opt. Precision Eng.*, 2007, 15(8):1258-1262. (in Chinese)
- [6] 张红颖, 胡正. CenSurE 特征和时空信息相结合的运动目标检测[J]. *光学精密工程*, 2013, 21(9): 2453-2463.
- ZHANG H Y, HU ZH. Moving object detection in combination of CenSurE and spatial-temporal information [J]. *Opt. Precision Eng.*, 2013, 21(9):2453-2463. (in Chinese)
- [7] STAUFFER C, GRIMSON W E L. Learning patterns of activity using real-time tracking [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2000, 22(8):747-757.
- [8] ZIVKOVIC Z, VAN D H F. Efficient adaptive density estimation per image pixel for the task of background subtraction [J]. *Pattern Recognition Letters*, 2006: 773-780.
- [9] WREN C R, AZARBAYEJANI A, DARRELL T, *et al.*. Pfunder: Real-time tracking of the human body [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1997, 19(7):780-785.
- [10] MADDALENA L, PETROSINO A. A self-organizing approach to background subtraction for visual surveillance applications [J]. *IEEE Transactions on Image Processing*, 2008, 17(7):1168-1177.
- [11] MADDALENA L, PETROSINO A. A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection[J]. *Neural Computer Application*, 2010, 19(2): 179-186.
- [12] MADDALENA L, PETROSINO A. The SOBS algorithm: what are the Limits? [C]. *Computer Vision and Pattern Recognition Workshops*, 2012:21-26.
- [13] BARNICH O, DROOGENBROECK M V. ViBe: A universal background subtraction algorithm for video sequences [J]. *IEEE Transaction on Image Processing*, 2011, 20(6):1709-1724.
- [14] DROOGENBROECK M V, PAQUOT O. Background subtraction: experiments and improvements for ViBe [C]. *Computer Vision and Pattern Recognition Workshops*, 2012:32-37.
- [15] DING J, MA R, CHEN S. A scale-based connected coherence tree algorithm for image segmentation [J]. *IEEE Transaction on Image Process*, 2008, 17(2):204-216.

作者简介:



陈星明(1989—),男,江苏江阴人,硕士,2012年于南京大学获得学士学位,主要从事智能图像处理的研究。E-mail:chris_cxm@163.com

导师简介:



李勃(1982—),男,湖南湘潭人,博士,副教授,硕士生导师,2009年于南京大学获得博士学位,主要从事机器视觉、人工智能等方面的研究。E-mail:liboee@nju.edu.cn

导师简介:



陈启美(1947—),男,江苏南京人,教授,博士生导师,1982年于清华大学获得硕士学位,主要从事网络通信、智能交通等方面的研究。E-mail:chenqimei@nju.edu.cn