

一种对阴影鲁棒的背景建模算法

范志辉, 卢朝阳, 李 静, 姜维, 姚超

(西安电子科技大学 综合业务网理论与关键技术国家重点实验室, 陕西西安 710071)

摘要:为消除前景图像的阴影,提出了一种基于尺度不变局部三元模式(SILTP)的视频图像背景建模算法。该算法利用图像帧级、图像块级和像素级三级信息进行设计。在图像帧级,利用全局灰度均值处理亮度突变;在图像块级,利用SILTP纹理图像基于图像块进行背景建模,快速定位前景目标大致轮廓;在像素级,用类ViBe算法检测前景目标精确边界。实验表明该算法不仅可有效消除前景图像的阴影,对复杂场景亦具有较好适应性。

关键词:运动分割;目标检测;背景差;阴影消除

中图分类号:TP391.41 **文献标识码:**A **doi:**10.3969/j.issn.0253-2778.2015.07.005

引用格式: FAN Zhihui, LU Zhaoyang, LI Jing, et al. A robust background modeling algorithm for shadows[J]. Journal of University of Science and Technology of China, 2015, 45(7): 561-567.
范志辉, 卢朝阳, 李 静, 等. 一种对阴影鲁棒的背景建模算法[J]. 中国科学技术大学学报, 2015, 45(7): 561-567.

A robust background modeling algorithm for shadows

FAN Zhihui, LU Zhaoyang, LI Jing, JIANG Wei, YAO Chao

(State Key Lab. of Integrated Service Networks, Xidian Univ., Xi'an 710071, China)

Abstract: To remove shadows from foreground objects, an algorithm for background subtraction based on Scale-invariant model of local pattern (SILTP) was proposed. The algorithm was designed by fusing three levels of information from the frame, region and pixel levels. The frame level detects sudden, global changes between frames by the global mean gray value. At the region level, the SILTP operator was utilized to get the texture images. Then the texture histograms were used to model background, so as to quickly find contours of moving objects in the frame sequence. The pixel-level obtains object contours accurately using similar-ViBe algorithm. Exhaustive experimental evaluations on complex scenes show efficiency of the proposed method.

Key words: Motion segmentation; Object detection; Background subtraction; Shadow removal

0 引言

运动目标检测是视频监控、智能交通、机器人视觉等领域非常关键的步骤^[1-2]。精确的运动目标检

测能极大地提高目标跟踪、识别、分类和行为分析的性能。运动目标自身投影容易改变目标的形状和大小,从而影响对目标动作和行为的进一步判断。构建对阴影鲁棒的背景模型是当前的研究热点。背景建

收稿日期:2014-09-12;修回日期:2015-03-06

基金项目:中央高校基本科研业务费专项基金(K50510010007)资助。

作者简介:范志辉,男,1982年生,博士生。研究方向:计算机视觉。E-mail: fanzhihui0379@163.com

通讯作者:卢朝阳,男,博士/教授。E-mail: zhylu@xidian.edu.cn

模除了要解决前景目标自身投影问题,还需要处理以下几种情况^[3]:①光照变化;②背景干扰③自动初始化;④识别伪装;⑤对场景改变敏感:快速吸收停止运动的目标,快速消除鬼影。

近年来,针对固定摄像机场景的背景建模国内外学者提出了大量有效方法^[4-7],用到的图像信息包括颜色、纹理、时空关联信息等,有的利用单个特征有的融合多个特征进行背景建模.建模方法主要分为参数化和非参数化两类.参数化的方法假设图像像素值的变化符合某种概率分布,其中最具代表性的是混合高斯(mix of Gaussian, MoG).文献[8]利用 MoG 为每个像素建模.该方法对环境光线变化和干扰敏感,阴影和鬼影消除效果不好.文献[9]提出基于图像块的背景建模算法,对干扰有一定适应性,可快速得到目标大致轮廓,但是无法得到运动目标的细节信息.非参数化的方法认为图像像素值的变化可能是任意形式的概率,其中利用核密度估计数据的概率分布是最具代表性的方法.文献[10]分析了阴影和干扰对图像灰度值的影响,提出 SILTP (scale invariant local ternary pattern) 纹理算子,设计基于模式核密度估计的背景建模算法,算法对光照变化具有较好的鲁棒性.文献[11]设计一种纹理基元,基于核密度估计进行背景建模适应水下场景,表现出较好的对水下运动目标检测性能.核密度估计方法处理速度欠佳.上述背景建模方法只关注图像帧内信息,忽视了图像序列间的关联信息,而且大多仅从图像某一方面解决问题.文献[12]联合时空信息进行建模背景,取得较好效果.文献[13]提出综合利用图像帧间信息、图像帧内信息和图像像素信息进行背景建模,算法在室内场景取得了较好的效果,但不能消除室外运动目标阴影.文献[14]提出 ViBe 算法,利用邻域相关性和随机性做背景建模;并通过大量实验证实算法不仅快速有效而且对于干扰有较好的适应性.由于 ViBe 仍然是一种基于像素的方法,难以克服基于像素背景建模方法的固有缺点,如不能适应光照突变、无法消除阴影等;因此文献[15]将 HSV 空间与 ViBe 算法结合试图消除目标阴影,但该算法需要根据不同场景和光照条件调整门限参数,适应性较差.如何有效消除阴影、抗环境及成像设备的干扰、提高处理速度仍然是背景建模算法需要解决的问题。

基于以上分析,本文设计了一种对运动目标自身阴影和环境干扰鲁棒的背景建模算法.算法利用

图像帧、图像块和图像像素三级信息进行背景建模,快速准确检测运动目标.实验数据验证了本文算法的性能。

1 算法设计

1.1 设计动机

设计对于干扰鲁棒适应光照变化且可有效消除前景目标投影的背景建模算法的一个有效途径是多特征多方法的融合,其中的关键是选择可互补长短的图像特征、互不抵触的融合方法,以达到取长补短,同时不大幅降低处理速度的目的.上述算法分别用到图像某一方面信息建模背景,各有优缺点,又有优势互补的可能.本文利用图像的帧间相关性、像素的区域相关性和像素的时域变化特征设计算法,在图像帧级利用全局灰度均值应对亮度突变;在图像块级,利用 SILTP 算子得到纹理图,基于图像块背景建模,快速得到目标轮廓;最后,在像素级利用类 ViBe 算法进行细节检测,实现对以上算法的有效融合,取长补短,达到快速检测运动目标,有效消除阴影的目的。

1.2 算法设计

1.2.1 算法处理流程

算法处理过程分为三步,如图 1 所示.第一步,处理环境光照突变.统计图像序列灰度均值的变化幅值,当两帧之间平均灰度值差大于预设阈值 T ,认为存在环境光照突变,对新帧中每个像素的灰度值增加或减少 T ;第二步,初步定位目标轮廓.即利用 SILTP 纹理对环境光照变化的鲁棒性消除前景目标的阴影,基于图像块背景建模,快速检测运动目标轮廓;第三步,在轮廓区域利用类 ViBe 算法获取目标的精确轮廓.最后,将第三步和第二步的检测结果融合得到最终前景图像.图 1 表示算法整体处理流程,图 2 和图 3 分别表示第二、三步的算法流程。

1.2.2 SILTP 模型

SILTP 模型^[10]是消除前景目标阴影的关键,SILTP 算子的灰度值尺度不变性可处理图像局部光照变化,消除阴影.SILTP 算子计算公式如下:

$$\text{SILTP}_{N,R}^{\tau}(X_C, Y_C) = \bigoplus_{k=0}^{N-1} S_{\tau}(I_C, I_k) \quad (1)$$

式中, I_C 表示中心点像素的灰度值, I_k 是 I_C 的半径为 R 的邻域中 N 个像素的灰度值, \bigoplus 表示二进制字符的串接运算, τ 是尺度变化因子决定阈值区间,变换函数 S_{τ} 定义如下:

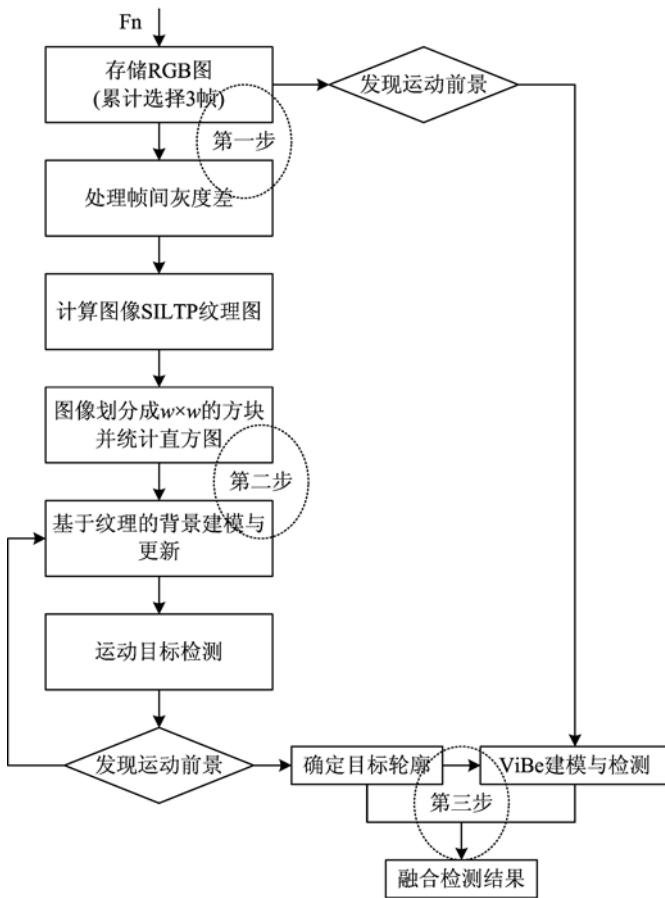


图 1 算法整体处理流程

Fig. 1 Scheme of the proposed method for motion segmentation

$$S_{\tau}(I_C, I_K) = \begin{cases} 10, & \text{if } I_K < (1 + \tau) I_C \\ 01, & \text{if } I_K > (1 + \tau) I_C \\ 00, & \text{其他} \end{cases} \quad (2)$$

式中, S_{τ} 用 3 个 2 位的二进制数 (10、01、00) 表征纹理特性. 文献[10]分析并验证了 SILTP 纹理算子对干扰和阴影的鲁棒性.

1.2.3 基于图像块背景建模算法设计

如图 2 所示, 对每一帧图像, 首先利用 SILTP 模型得到图像的纹理图, 采用部分重叠方式将图像划分为 $\omega \times \omega$ 大小的方块, 每个方块用一组纹理直方图 $\{x_1, \dots, x_t\}$ 表征, 每个纹理直方图对应一个权重, t 时刻的纹理直方图 k^{th} 对应权重 $\omega_{k,t}$. 每个块的初始化和更新过程为: 初始化阶段利用 N (不大于 5) 帧图像对每个方块进行背景建模, 每个块对应一组由各纹理直方图组成的纹理直方图组, 随机分配初始权重, 所有 N 个权重的和等于 1, 对 N 个直方图按权重从大到小排序, 前 K 个权重和累加达到门限 T_B 的纹理直方图代表该块的纹理特征. 对新一

帧图像, 新块与所有 N 个块利用

$$s(h_1, h_2, B) = \frac{\sum_i^B h_{1,i} h_{2,i} - \sum_i^B h_{1,i} \sum_i^B h_{2,i} \frac{1}{B}}{\sqrt{(\sum_i^B h_{1,i}^2 - (\sum_i^B h_{1,i})^2 \cdot B) (\sum_i^B h_{2,i}^2 - (\sum_i^B h_{2,i})^2 \cdot B)}} \quad (3)$$

比较纹理直方图相关 (CORREL) 性, 得到相关性值 S . 当 S 大于门限 T_D , 则利用

$$m_{k,t} = V_b h + (1 - V_b) m_{k,t-1}, \quad V_b \in [0, 1] \quad (4)$$

更新纹理直方图的 bin 值, 利用

$$\omega_{k,t} = V_w M_k + (1 - V_w) \omega_{k,t-1}, \quad V_w \in [0, 1] \quad (5)$$

更新直方图权重. 如果不匹配, 仅更新权重. 前景和背景的判断: 采用保守的策略, K 个纹理直方图中只要有一个纹理直方图与新纹理直方图相似, 则认为新块属于背景, 否则认为新块属于前景.

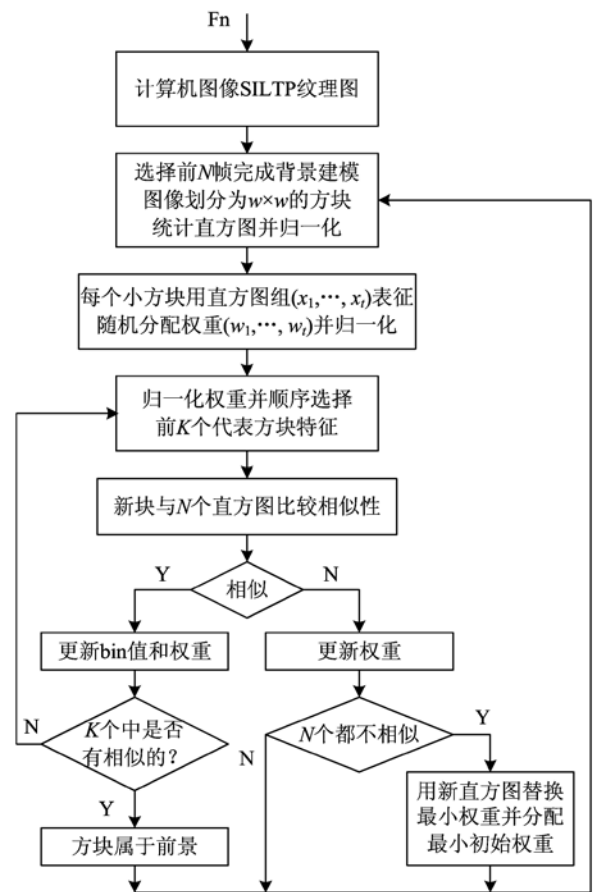


图 2 基于图像块的背景建模算法流程

Fig. 2 Steps of the process of the block histogram-based motion segmentation

如果 N 个直方图没有 1 个与新纹理直方图相似, 则用新纹理直方图替换权重最小的纹理直方图, 并分

配一个较小的初始化权重. 最后对直方图权重归一化, 重新按照权重从大到小排序, 选择前 K 个权重和累加达到门限 T_B 的纹理直方图代表块的特征, 经过以上过程完成直方图及权重更新和运动目标检测.

公式(3)中, h_1 和 h_2 表示两个直方图, B 表示直方图 bin 的数量. 公式(4)和(5)中, $m_{k,t}$ 和 $\omega_{k,t}$ 分别表示 t 时刻第 k 个直方图的 bin 值和权重, 最匹配的直方图 M_k 取 1, 其他直方图 M_k 取 0, \vec{h} 表示新直方图. V_b 和 V_w 是算法运行参数, 值越小算法更新策略越保守.

1.2.4 类 ViBe 算法设计

ViBe 随机选择像素周围邻域像素值和像素的历史值进行背景建模, 具有计算速度快、检测精确及单帧建模等优点. 这里对 ViBe 算法做修改, 使其符合本文算法整体需要, 如图 3 所示. 在算法起始阶段不进行背景建模而是始终保存三帧图像, 直到以上算法检测到前景运动目标后, 再启动 ViBe 算法. ViBe 算法先利用存储的三帧图像建立背景模型, 然后在轮廓区域内进行前景检测. 前景检测过程: 首先记录包含目标轮廓的最小矩形区域的位置和大小; 然后在该矩形区域内检测运动目标的精确轮廓; 最后将各矩形区域内检测到的目标按位置合并. 该修改仍然保持了 ViBe 算法的设计思想, 保证了算法处理速度快和检测精确的优点.

2 实验设计及结果分析

2.1 实验设计

利用标准视频库 CDM'12 和 PETS2006 中 8 段代表性视频比较了 MoG 算法、LTP 算法、ViBe 算法及本文算法的前景检测效果、正确分类比 (percentage of correct classification, PCC) 和秒平均帧率 (frames per second, FPS). 程序运行环境: PC Intel Core 2 Duo CPU E7400 2.8 GHz, 内存 1G, 操作系统 XP, 开发平台 VC++6.0, 开源图像函数库 OpenCV1.0. 为保证比较的公平性, 所有算法均未做形态学处理, 仅过滤小目标点 (小于 20 个像素).

通过两种方式比较算法效果, 一种是主观评价, 观察前景运动目标的提取效果, 如图 4 和 5 所示. 另一种是客观评价, 用 PCC 和 FPS 两个量化指标比较算法性能. 单帧图像的 PCC 指标用表示如下:

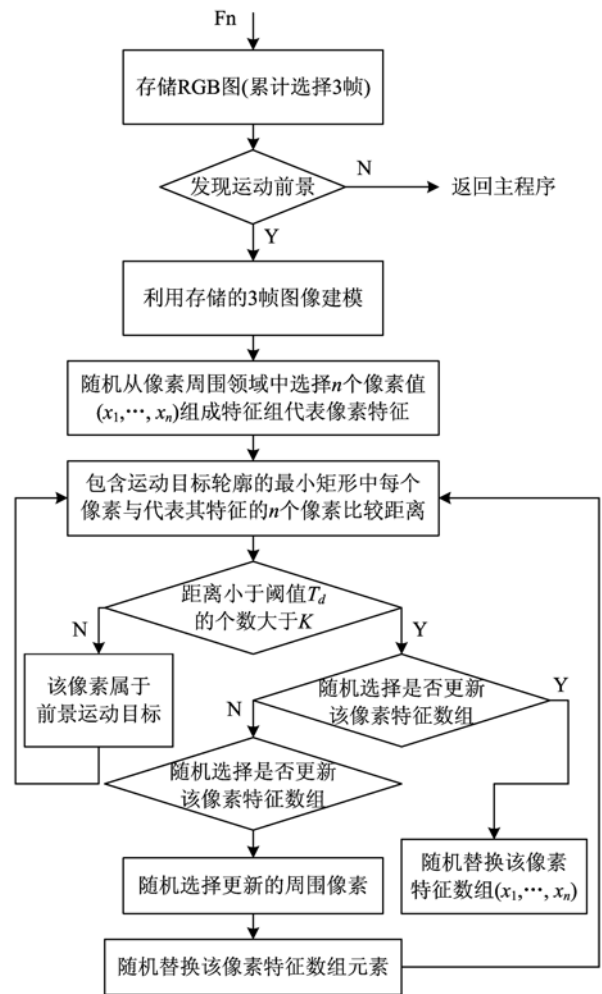


图 3 类 ViBe 算法流程

Fig. 3 Steps of the process of the similar-ViBe

$$PCC = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

式中, TP(true positives)代表前景点被正确判断出来的次数, TN(true negatives)代表背景点被正确判断出来的次数, FP(false positives)代表背景点被错误判断为前景的次数, FN(false negatives)代表前景点被错误判断为背景的次数. PCC 值越大, 说明匹配越准确, 算法性能越好. 从场景中截取 100 帧连续图像计算算法的 PCC 均值, 客观评价结果如表 1 所示.

2.2 实验结果及分析

图 4 是算法的前景检测效果. 对于普通场景“people”, 四种算法均能准确分离出前景. LTP 算法基于图像块设计, 目标中间部分纹理相似出现空洞. 对于包含阴影的室内场景“PETS2006”: MoG 算法不能消除阴影; ViBe 算法对阴影消除效果不好; LTP

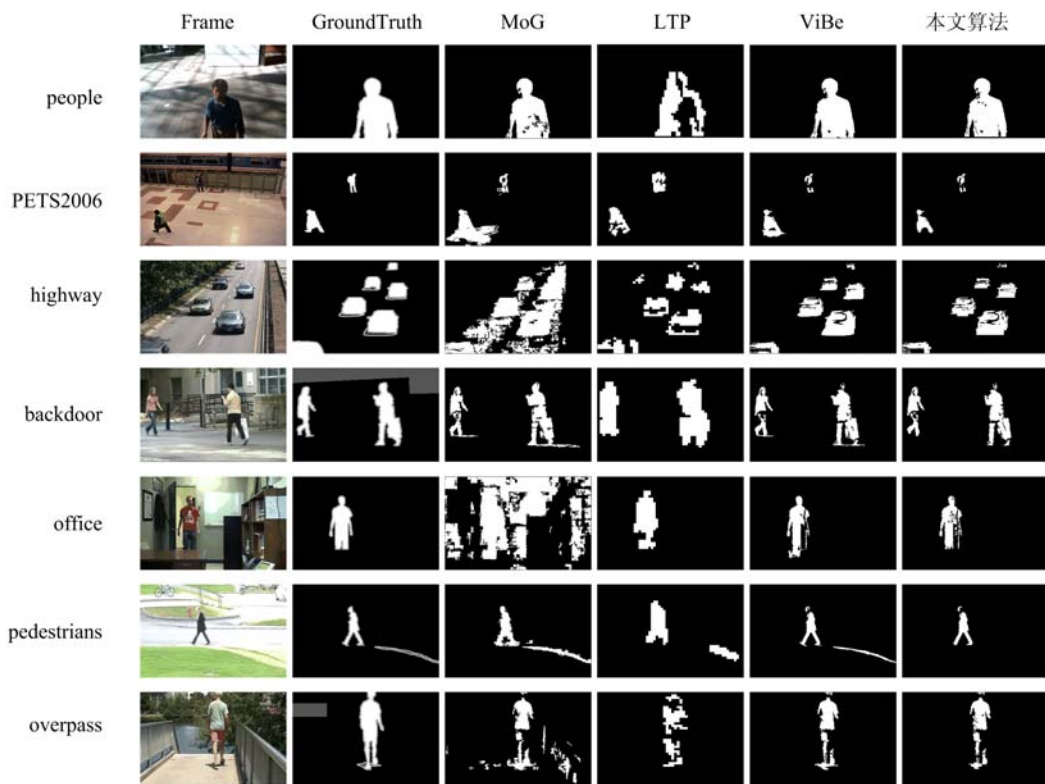


图 4 不同场景前景分离效果对比图

Fig. 4 Segmentation results on seven datasets

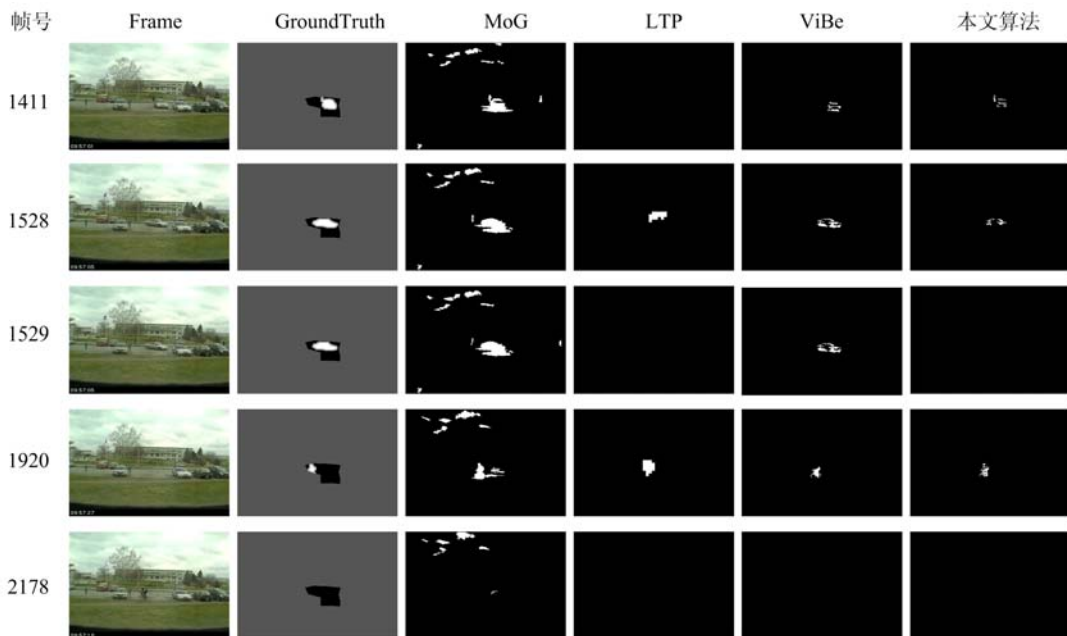


图 5 算法快速适应性效果对比图

Fig. 5 Sensitivity evaluation to “sleeping person” and ghost

算法可有效消除阴影,但是不能得到目标精确轮廓;本文算法不仅可有效滤除阴影,而且可以得到目标较精确轮廓.对于场景“highway”,存在树枝摆动及

由此引起的周围亮度变化,ViBe 算法和本文算法均可以有效排除干扰,精确分离前景目标;MoG 算法受影响较大,几乎不能分离出前景目标;LTP 算法

不能够过滤树枝的大范围摆动. 对于室外场景“backdoor”: 由于存在光照变化, MoG 算法和 ViBe 算法均不能有效消除阴影; LTP 算法和本文算法可以有效排除光照影响. 对于室内场景“office”: 目标进入房间后引起室内光照变化, MoG 算法受影响较大, 几乎不能分离前景和背景; ViBe 算法误将目标自身阴影作为前景的一部分; LTP 算法和本文算法不受此影响. 对于包含阴影的室外场景“pedestrians”, 目标出现较长阴影, 各部分深浅不一. 由图 4 可以明显看出, MoG 算法和 ViBe 算法几乎将全部阴影检测为前景; LTP 算法可消除部分阴影; 本文算法表现最好, 可有效消除全部阴影, 精确分离前景和背景. 对于包含背景干扰的室外场景“overpass”, 场景中有树枝轻微摆动. 只有 MoG 算法仍然不能适应目标出现引起的周围光照变化, 其他算法均可以排除干扰. 综上, 本文算法的综合视觉效果优于其他三种.

表 1 各算法不同场景 PCC(%) 及 FPS 对比表
Tab. 1 Performance of PCC(%) on the complex video sequences and the running framerate(fps) in average

| 场景\算法名称 | MoG | LTP | ViBe | 本文算法 |
|-----------------------------|-------|-------|-------|-------|
| outdoor\people | 91.47 | 84.32 | 97.27 | 96.73 |
| indoor\PETS2006 | 93.22 | 95.54 | 94.36 | 98.25 |
| outdoor\highway | 79.17 | 89.51 | 95.46 | 96.18 |
| outdoor\backdoor | 91.59 | 92.61 | 94.25 | 96.36 |
| indoor\office | 59.07 | 93.28 | 95.62 | 97.36 |
| outdoor\pedestrians | 81.22 | 90.6 | 83.69 | 99.56 |
| dynamic background\overpass | 92.22 | 94.4 | 96.41 | 98.48 |
| PCC-average | 83.99 | 91.47 | 93.87 | 97.56 |
| FPS-outdoor\pedestrians | 31.8 | 41.3 | 65.4 | 38.6 |

图 5 是算法快速适应性和灵敏性效果对比. 场景中目标存在间歇性运动且包含小目标. 本文算法、MoG 算法及 ViBe 算法均能及时检测到运动目标(1 411 帧、1 920 帧), LTP 算法在 1 411 帧未能检测到小目标. 本文算法和 LTP 算法比 ViBe 有更好的鬼影消除效果(1 528 帧、1 529 帧), 也能更快吸收静止下来的目标, MoG 算法鬼影消除和吸收静止目标效果最差, 直到 2 178 帧仍没有消除 1 411 帧检测到的车辆的鬼影.

由表 1 可知, 本文算法 PCC 均值最大, MoG 算法 PCC 均值最小, ViBe 算法 PCC 均值波动较大, 基本与图像效果一致, 量化指标进一步验证了本文

算法性能. 采用“pedestrians”场景大小为 360×240 的图像序列比较各个算法的 FPS 性能. 从表 1 的数据可以看出, ViBe 算法速度最快, 这与 ViBe 算法主要为比较、选择和累加运算相符; LTP 算法基于图像块设计速度也较快; 本文算法处理速度优于 MoG 算法. 图 4 和图 5 的视觉直观效果和表 1 的量化数据显示, 本文算法综合性能优于其他三种算法.

2.3 算法性能的进一步验证

为充分验证算法性能, 采用文献[10]所用视频库比较本文算法和 PKDE^[10] 算法的 F-score^[10] 和 FPS 性能. F-score 如下:

$$\frac{2TP}{2TP+FP+FN} \quad (7)$$

式中, TP、FN 及 FP 的含义与式(6)相同. 采用文献[10]中三个版本算法: PKDE_{sltp} (表 2 中对应 I)、PKDE_{mb-siltp} ^{$\omega=3$} (表 2 中对应 II)、PKDE_{mb-siltp} ^{$\omega=1+2+3$} (表 2 中对应 III), 与本文算法比较, 结果如表 2 所示.

表 2 本文算法与文献[10]算法的 F(%) 和 FPS 性能对比
Tab. 2 Performance of F(%) and FPS

| 场景\算法名称 | 算法I | 算法II | 算法III | 本文算法 |
|-------------------------------|-------|-------|-------|-------|
| AirportHall | 62.26 | 58.31 | 63.07 | 65.27 |
| Bootstrap | 69.89 | 60.27 | 68.71 | 74.89 |
| Curtain | 79.20 | 84.58 | 87.39 | 88.01 |
| Escalator | 56.73 | 58.46 | 63.11 | 68.56 |
| Fountain | 79.33 | 75.21 | 79.63 | 84.14 |
| ShoppingMall | 77.45 | 75.16 | 77.81 | 75.71 |
| Lobby | 75.21 | 67.23 | 75.06 | 81.19 |
| Trees | 40.34 | 57.15 | 62.92 | 64.26 |
| WaterSurface | 71.39 | 75.48 | 77.91 | 86.34 |
| Total | 71.35 | 69.67 | 70.24 | 77.22 |
| FPS-ShoppingMall (320×256) | 8.5 | 62.6 | 6.2 | 40.9 |

由表 2 可知, 文献[10]算法三个版本中, PKDE_{mb-siltp} ^{$\omega=3$} 综合性能最好, F-score 的性能虽然与 PKDE_{sltp} 和 PKDE_{mb-siltp} ^{$\omega=1+2+3$} 分别相差 1.68 和 0.57 个百分点, FPS 速度却提升了几近十倍, 体现了基于图像块的设计方法的优点. 本文算法与 PKDE_{mb-siltp} ^{$\omega=3$} 算法相比, F-score 性能提升了约 10 个百分点, 速度略有下降, 但仍然有较高的处理速度, 达到 40.9 帧/秒.

3 结论

本文针对视频图像序列运动目标检测中目标自

身阴影消除问题,利用多个图像特征的互补特性设计算法,提出综合利用图像帧级、图像块级和像素级三级信息进行背景建模消除前景目标的阴影.算法利用帧间信息应对图像光照突变,利用 SILTP 算子提取图像纹理信息,克服阴影和干扰,采用 ViBe 算法获取前景目标精确轮廓.最后用多种视频场景比较了本文算法与 MoG 背景建模算法、LTP 背景建模算法、ViBe 算法和 PKDE 背景建模算法的性能.大量实验数据表明,本文算法能有效消除室内阴影,抑制室外阴影,对干扰具有较好的适应性,能精确及时地检测到运动目标轮廓,处理速度较好,算法综合性能优于其他代表性算法.

参考文献(References)

- [1] Elgammal A. Background Subtraction: Theory and Practice[M]. Berlin, Germany: Springer, 2014.
- [2] 李维维, 张陈斌, 陈宗海, 等. 基于特征学习与特征记忆模板更新机制的粒子滤波跟踪[J]. 中国科学技术大学学报, 2014, 44(4): 292-302.
LI Weiwei, ZHANG Chenbin, CHEN Zonghai, et al. Particle filter tracking based on feature-learning and feature-memory template update mechanism[J]. Journal of University of Science and Technology of China, 2014, 44(4): 292-302.
- [3] 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: 4-21.
- [4] 赵旭东, 刘鹏, 唐降龙, 等. 一种适应户外光照变化的背景建模及目标检测方法[J]. 自动化学报, 2011, 37(8): 915-922.
ZHAO Xudong, LIU Peng, TANG Jianglong, et al. Background modeling adaptive to outdoor illumination variation and foreground detection approach[J]. Acta Automatica Sinica, 2011, 37(8): 915-922.
- [5] HUANG, Jiahua, YIN Dong, ZHANG Rong, et al. Multi-view multi-pose object detection using splitted cascade model[J]. Journal of University of Science and Technology of China, 2013, 43(7): 540-546.
黄江华, 尹东, 张荣, 等. 基于分裂级联模型的多视角多姿态目标检测[J]. 中国科学技术大学学报, 2013, 43(7): 540-546.
- [6] Choi J M, Chang H J, Yoo Y J, et al. Robust moving object detection against fast illumination change [J]. Computer Vision and Image Understanding, 2012, 116(2): 179-193.
- [7] Zhou H, Chen Y R, Feng R. A novel background subtraction method based on color invariants [J]. Computer Vision and Image Understanding, 2013, 117(11): 1589-1597.
- [8] KaewTraKulPong P, Bowden R. An improved adaptive background mixture model for real-time tracking with shadow detection [A]// Video-Based Surveillance Systems. Kluwer Academic Publishers, 2002: 135-144.
- [9] Heikkilä M, Pietikäinen M, Heikkilä J. A texture-based method for detecting moving objects [A]// British Machine Vision Conference, 2004: 187-196.
- [10] Liao S C, Zhao G Y, Kellokumpu V, et al. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes[C]// IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE Press, 2010: 1301-1306.
- [11] Spampinato C, Palazzo S, Kvasidis I. A texton-based kernel density estimation approach for background modeling under extreme conditions [J]. Computer Vision and Image Understanding, 2014, 122: 74-83.
- [12] Yoshinaga S, Shimada A, Nagahara H, et al. Object detection based on spatiotemporal background models [J]. Computer Vision and Image Understanding, 2014, 122: 84-91.
- [13] Varcheie P D Z, Sills-Lavoie M, Bilodeau G A. A multiscale region-based motion detection and background subtraction algorithm[J]. Sensors, 2010, 10(2): 1041-1061.
- [14] Van Droogenbroeck M, Paquot O. Background subtraction: Experiments and improvements for ViBe [C]// IEEE Computer Vision and Pattern Recognition Workshops. Providence, USA: IEEE Press, 2012: 32-37.
- [15] Qin Y S, Sun S F, Ma X B, et al. A shadow removal algorithm for ViBe in HSV color space [C]// 3rd International Conference on Multimedia Technology. Beijing: Atlantis Press, 2013: 966-973.