

复杂背景下基于 LBP 纹理特征的 运动目标快速检测算法

裘莉娅^{1,2,3}, 陈玮琳^{1,2,3}, 李范鸣^{1,3*}, 刘士建^{1,3}, 李争^{1,3}, 谭畅^{1,2,3}

(1. 中国科学院上海技术物理研究所, 上海 200083;

2. 中国科学院大学, 北京 100049;

3. 中国科学院红外探测与成像技术重点实验室, 上海 200083)

摘要:在雨雪天气、树叶晃动、水面闪烁等有复杂背景的可见光与红外场景中,快速准确地提取完整目标一直是运动目标检测中的首要难题。为了满足实时性,并针对现有视频的前景提取算法依赖先验信息、召回率低、缺乏纹理和噪声较大等问题,提出了一种基于直方图统计和改进的局部二值模式(Local Binary Pattern, LBP)纹理特征相结合的背景建模方法。首先,使用各像素直方图的众数作为参考背景,无需先验知识,节省了大量存储空间,再采用邻域补偿策略提出了一种改进的S_MBLBP纹理直方图与参考背景进行背景建模,消除了大部分动态背景和光照变化影响,实现目标的精确提取。实验表明,所提的算法在红外和可见光的多种复杂场景下,能快速提取前景目标的同时,提高了准确率和召回率。

关键词:机器视觉;背景建模;LBP纹理特征;运动目标检测;复杂背景

中图分类号:TP391.9

文献标识码:A

Fast moving target detection algorithm based on LBP texture feature in complex background

QIU Li-Ya^{1,2,3}, CHEN Wei-Lin^{1,2,3}, LI Fan-Ming^{1,3*}, LIU Shi-Jian^{1,3}, LI Zheng^{1,3}, TAN Chang^{1,2,3}

(1. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China)

Abstract: In the visible and infrared scenes with complex background, such as rain and snow weather, leaf swaying, shimmering water, etc., fast and accurate extraction of a complete target has always been the primary problem in moving target detection. In order to be real time and aiming at the problems of existing video foreground extraction algorithms, such as dependence on prior information, low recall rate, lack of texture and large noise, a background modeling method based on histogram statistics and improved LBP (Local Binary Pattern) texture features is proposed. Firstly, the mode of each pixel histogram is used as the reference background without prior knowledge, which saves a lot of storage space. Then, an improved S_MBLBP texture histogram is proposed to model the background with the reference background by using neighborhood compensation strategy, which eliminates the most dynamic background and illumination changes, and realizes the accurate extraction of the target. Experimental results show that the proposed algorithm can quickly extract foreground targets in a variety of complex infrared and visible scenes, and can improve the accuracy and recall rate at the same time.

Key words: machine vision, background modeling, LBP textural features, moving target detection, complex background

收稿日期:2021-08-10,修回日期:2021-09-06

Received date:2021-08-10, Revised date:2021-09-06

基金项目:中国科学院青年创新促进会资助

Foundation items: Supported by the Youth Innovation Promotion Association CAS

作者简介(Biography):裘莉娅(1994—),女,浙江省嵊州市人,博士研究生,主要研究领域为复杂背景下的红外面目标搜索与跟踪. E-mail: qiu-liyad@163.com

*通讯作者(Corresponding author): E-mail: lfmjws@163.com

引言

运动目标检测(Moving Object Detection, MOD)一直是计算机视觉领域的一个热门研究点,在目标追踪、视频监控、异常行为分析等领域有着广泛应用^[1-4]。但运动目标检测一直是一个非常具有挑战性的工作,因为在实际运用场景中,光照的变化、树叶的晃动、行为者的不同、雨雪等复杂背景、物体遮挡、自遮挡等干扰都会使准确检测出运动目标的难度增加。为此,研究者提出了多种解决方法,如光流法、多特征融合法、背景差分法、三帧差分法^[5]等。光流法利用运动物体流向量的属性来检测图像序列中是否存在改变的区域以提取运动目标^[6-7],但其存在计算速度慢和抗噪声能力差等问题。多特征融合法主要通过融合运动像素的时空信息来提取前景目标^[8-10],但该方法对时间域变化比较敏感,且对空间对比度要求较高。但是近些年来,深度学习的方法迅速发展起来,相比较传统算法,深度学习算法具有较强的拟合能力和更高的准确度的优点,但性能的好坏依赖于数据集的大小。在红外场景^[11]、未知载具场景等一些特殊场景中,因缺乏样本不能满足随时部署的需求。对于这种少样本或零样本的场景依然没有有效的解决方案,而且深度学习算法对硬件的要求明显高于传统算法,在很大程度上会增加成本,因此传统方法仍然有着广泛的应用范围,但受到计算能力、实时性和准确度等因素的限制,在实际中,背景建模法因具有算法简单、检测精度高、运算速度快、可满足实时性等优点,逐渐成为广泛应用的目标检测方法。

背景建模法在许多视频分析中用于检测和分割前景的第一步,检测的结果质量在很大程度上取决于背景模型的准确性,一个理想的背景模型应该能够克服动态元素和噪声的影响,因此如何建立一个稳健的背景模型是背景差分法的研究热点^[12-14]。背景建模法主要分为基于像素的方法和基于局部区域特征的方法。基于像素的方法是将运动目标当作一系列独立的像素进行处理,如基于单高斯模型的处理方法:Wren等^[15]使用了单高斯模型描述背景,该模型适用于单一不变的背景,难以在复杂环境中完成目标检测任务。为了解决单高斯模型难以描述复杂背景的问题,Stauffer等^[16]使用了混合高斯模型(Gaussian Mixed Model, GMM)拟合背景,可以实现复杂背景的目标检测,但GMM模型需要大量的训练帧以确定模型参数,因此在当遇到简单背

景时,该模型在检测目标时容易出现拖影,目标由运动慢慢静止时容易产生误检和漏检,且忽略了像素间的局部关系^[17-18]。而基于局部区域特征的方法充分利用了像素的邻域信息进行背景建模,构建的背景模型更精确,如Elgammal等提出核密度估计法(Kernel Density Estimation, KDE),通过计算每个像素点成为背景的概率^[19],但是阈值设置困难,所以计算复杂度仍较大。Barnich等^[20]提出的基于视觉背景提取器(Visual Background Extractor, ViBe)模型的背景建模方法,该方法使用初始帧邻域像素并采取随机策略更新背景,能够有效适应各类环境,但要求更新速率稳定,更新太快会吸收目标产生空洞,反之则易产生鬼影,无法适应环境变化,尤其是在自然环境中受光照变化影响大。

为了消除光照变化的影响,研究者提出基于局部二值模式(LBP)的背景建模方法^[21-22]。LBP算子是纹理特征的描述算子,具有灰度不变的特点,但LBP对噪声较为敏感。因此,研究者将LBP与其它特征结合^[23-26]进行运用,虽能有效消除光照变化的影响,却大大增加了计算复杂度。针对这些问题,提出一种基于直方图统计和改进的LBP的建模算法,该算法通过统计各像素点直方图的众数来估计背景,使慢速或静止的前景出现时,背景灰度不受影响,并且可以有效地区分静态背景和动态背景;其次改进了LBP,将图像进行分块处理,与传统的LBP特征相比,在像素之间的局部3×3的邻域计算方面,多块的LBP特征可以捕捉大尺度结构,而这些结构可能是图像结构的主要特征;再通过和众数构建的参考背景进行邻域的比较,可以进一步地精确提取前景,减少噪声对LBP背景建模影响的同时,可以有效地提高复杂场景下的召回率和准确率。相比较Haar、HOG(Histogram of Oriented Gradient)等其他算子而言,存在着以下几个方面的优点:单纯的灰度变化不会导致LBP值发生改变,所以,光照引起的变化也不会改变整个图像的LBP值,LBP算法可以很好地抑制光照的变化以及相机的微小抖动;对噪声也有很好的抑制,不会出现高斯背景建模时出现的斑点噪声,因此没有必要滤波,计算速度提高,便可以更好地满足实时性;它对于摇摆的树木、起伏的湖面和闪烁的监视器这些复杂的情况,还有背景中有新的物体进入或者旧的物体移出这些情况都有很好的适应性;对于慢速目标,不会存在较大空洞,将其吸收为背景。

1 运动目标检测

所提算法的总流程图如图 1 所示,以下将结合该流程图对每一部分进行详细介绍。

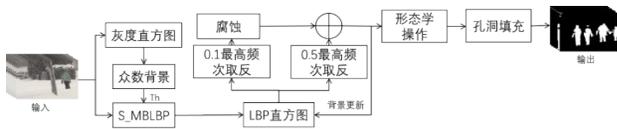


图 1 算法流程图
Fig. 1 The algorithm flow chart

1.1 众数背景建模法

文献[27]中, Lai A N, Yoon H, Lee G 将出现频率最高的灰度值作为真实背景值,但是这篇文章的直方图采用的是固定长度的图片序列,这样的方法在视频序列较短的情况下,容易错误地将移动速度慢的目标当成背景,而又会在视频序列长度过长时,造成内存的过度消耗。针对这些问题,我们对图像序列中的每一个像素点都建立灰度值的直方图,并且记录下不同的灰度值出现的次数,随着图

像序列在时间上的推移,读取的图像帧数增加,直方图也会随之更新,即可得到该像素点直方图数组,记为 $F_{(x,y),f} = \{f_1, f_2, \dots, f_f\}$ 。

图 2 中,(a)为输入的图像序列。红色的标记点为我们选取经常有运动目标经过观测点(200,200)展示众数背景建模过程;提取的灰度像素序列如(b)所示,可以看出有行人经过的时候,灰度产生了较大的波动;(c)中我们将所出现的灰度值进行统计得到灰度直方图,峰值即为出现频次最高的灰度值,在 3900 帧的图像序列中,(200,200)处出现频次最多的灰度值为 224,作为背景中该点的灰度值。因此,我们可以假设图像中某个像素点为 (x,y) ,当前读取的图像的帧数为 f ,直方图中某灰度值 n 出现的次数即为: $F_{(x,y),f}(n)$,将直方图中出现频率最高的灰度值作为当前背景灰度值 $B_{(x,y),f}$,则可以得到表达式:

$$B_{(x,y),f} = \text{Max} [F_{(x,y),f}(n)] \quad , \quad (1)$$

Max 的表达式为:

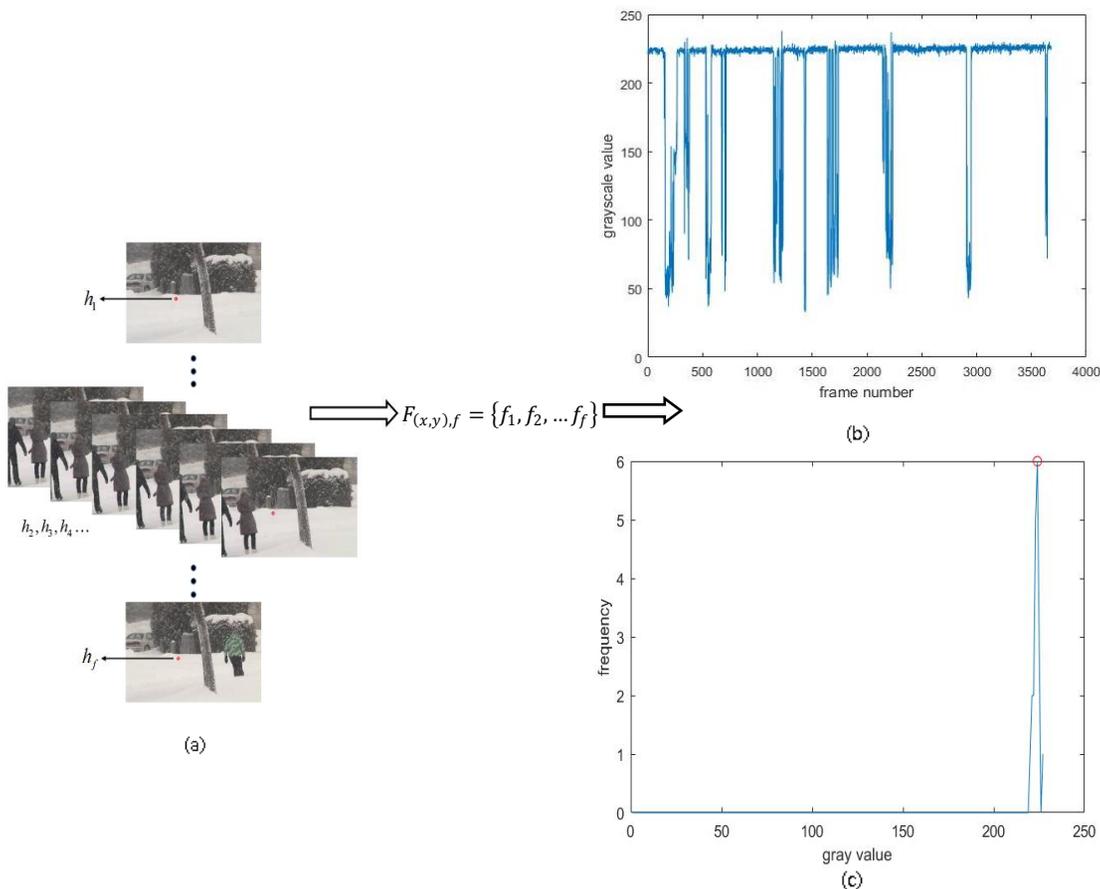


图 2 众数背景建模过程:(a)输入图像序列,(b)灰度变化曲线,(c)灰度直方图

Fig. 2 The mode background modeling process:(a) input image sequence, (b) grayscale change curve, (c) gray histogram

$$\text{Max}[F(n)] = \{n | \forall m \in [0, 255], F(m) \leq F(n)\}, \quad (2)$$

其中, f 为当前帧的帧数, $F(n)$ 、 $F(m)$ 分别灰度值 n 、 m 出现的频次, 由于使用的是 8 位的图像, 因此 n 的取值范围为 $[0, 255]$ (如果是 16 位图像, 则取值范围为 $[0, 65535]$)。为了降低算法的复杂度, 减少对内存的消耗, 前一帧的直方图中出现频率最高的灰度值为 $B_{(x,y),f-1}$, 出现次数为 $F_{(x,y),f-1}$, 当前帧 (x, y) 位置处的灰度值为 n , 出现次数为 $F_{(x,y),f}$, 则当前帧直方图中出现频率最高的灰度值 $B_{(x,y),f}$ 为:

$$B_{(x,y),f} = \begin{cases} B_{(x,y),f-1} & , F_{(x,y),f} \leq F_{(x,y),f-1} \\ n & , otherwise \end{cases} \quad (3)$$

此算法不需要存储历史图像序列, 并且 $\text{Max}(\cdot)$ 是非线性函数, 可以确保移速缓慢的前景出现时, 背景的灰度不会像 GMM 算法一样出现偏移, 同时也避免了 ViBe 算法会将前景吸收到背景中的问题, 并且随着时间的推移, 直方图的最大值也会越来越稳定, 保证了算法长时间运行的准确性和稳定性。

图 3 展示了不同场景下, 不同算法的建模结果。主要包含了复杂环境、红外两种不同的场景, 复杂环境分别为: 1、图 2 中图像序列的背景建模结果, 大雪场景中的滑雪者; 2、在树叶大幅度晃动中行走的

路人; 红外场景分别为: 3、在客厅取物的人; 4、长时间静止的读者。

可以看出, Vibe 算法在每一个场景中, 都在一定程度上将部分前景目标当成了背景; GMM 算法在第二个场景中, 由于树叶的长时间晃动, 背景建模效果并不理想, 尤其是在最后一个场景中, 由于前景长时间保持静止, 目标运动速度缓慢, GMM 和 Vibe 算法均把前景当成了背景, 进而影响后续的检测, 而众数背景建模在不同场景中得到的背景都最为准确干净。因此, 在不同的环境中, 众数背景建模算法对不同场景下不同运动状态的目标均有着良好的适应能力。

1.2 改进的 MBLBP 纹理特征

在彩色视频中, 有丰富的颜色特征和场景信息, 灰度视频相较彩色视频缺少很多信息, 并且基于灰度视频的建模存在精度低、缺乏纹理、易受光照变化影响的问题, 限制了它在复杂背景中的应用。因此, 需要更精确的阈值对前景和背景进行分割, 否则会导致召回率降低。

1.2.1 传统 LBP 纹理特征

LBP 是一种灰度级不变的纹理特征提取方法, 其计算简单, 具有旋转不变性, 并且对光照变化有



图 3 不同算法的背景建模结果

Fig. 3 Background modeling results of different algorithms

很好的鲁棒性,是纹理分析的有力工具。但是,由于最原始LBP算子采取的是相邻的8个像素点与中心像素点的灰度值作差分,如果某相邻像素大于中心像素的值,则该像素的值标记为1,反之则标记为0,再进行矩阵编码,将八位二进制数转换为十进制数作为此时中心像素的LBP值。LBP可用下式计算^[21]:

$$LBP_{N,R}(x_c,y_c) = \sum_{N=0}^{N-1} s(h_N - h_c) \times 2^N, \quad (4)$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad (5)$$

其中, (x_c, y_c) 表示中心像素, h_c 表示中心像素的灰度值,以像素点 (x_c, y_c) 圆心,半径 R 的圆区域称为中心像素点的邻域, N 表示邻域像素点的个数, h_N 为邻域像素点的灰度值。从上式可以看出,当中心像素的灰度值受到噪声影响之后,相邻8个像素的相对大小也会立马发生变化,对应的直方图会发生突变,十分不稳定,导致最终的计算结果对噪声极其敏感,错误识别像素也随之增多,并且由于数据集解码恢复后图像边界区域会出现的吉布斯震荡现象,图像经过LBP变换后边界会存在大量条状噪声,由于红外图像纹理相对于可见光图像更加简单,并且对比度低、模糊度高、受噪声影响更为严重。同时,传统LBP算子需要求取每一个像素的邻域直方图,选定的邻域较小,会造成求得的邻域直方图过于稀

疏而失去统计意义,且对每个像素求取直方图会导致计算量很大。

1.2.2 改进的S-MBLBP纹理特征

为了克服传统LBP算子易受噪声影响的缺点,采用分块取均值的方式计算参考差值点,充分利用邻域特征信息,避免直接采用中心单像素计算抗噪声性能差的缺点。原始的多尺度块LBP(Multiscale Block LBP, MBLBP)将图像分成诸多小块,再将每个小块划分成一个个小区域,对小区域内所有小块的灰度值求平均,然后将这个平均灰度值作为当前小区域的灰度值,再与相邻的小区域灰度进行比较形成LBP特征,生成的特征称为MBLBP,下图的小块大小为 9×9 ,小区域的大小为 3×3 。MBLBP有类似于先将图像进行平滑处理,然后再求LBP特征。

如图4所示为众数背景和266帧相同分块处在无前景的情况下,该位置得到的LBP值不同,会造成前景提取时背景突变错误检测为前景。为了抑制动态背景、像素值突变,提出了一种改进的稳定的多尺度块LBP(Stable_Multiscale Block LBP, S_MBLBP)算子。

S_MBLBP与MBLBP第一步相似,将得到的众数背景与每一帧在相同 9×9 的小块中,将小块平均分成 3×3 个小区域,求取灰度平均值。背景平均值 \bar{h}_b ,作为参考值;第 i 帧平均值 \bar{h}_i ,与 \bar{h}_b 进行比较,得

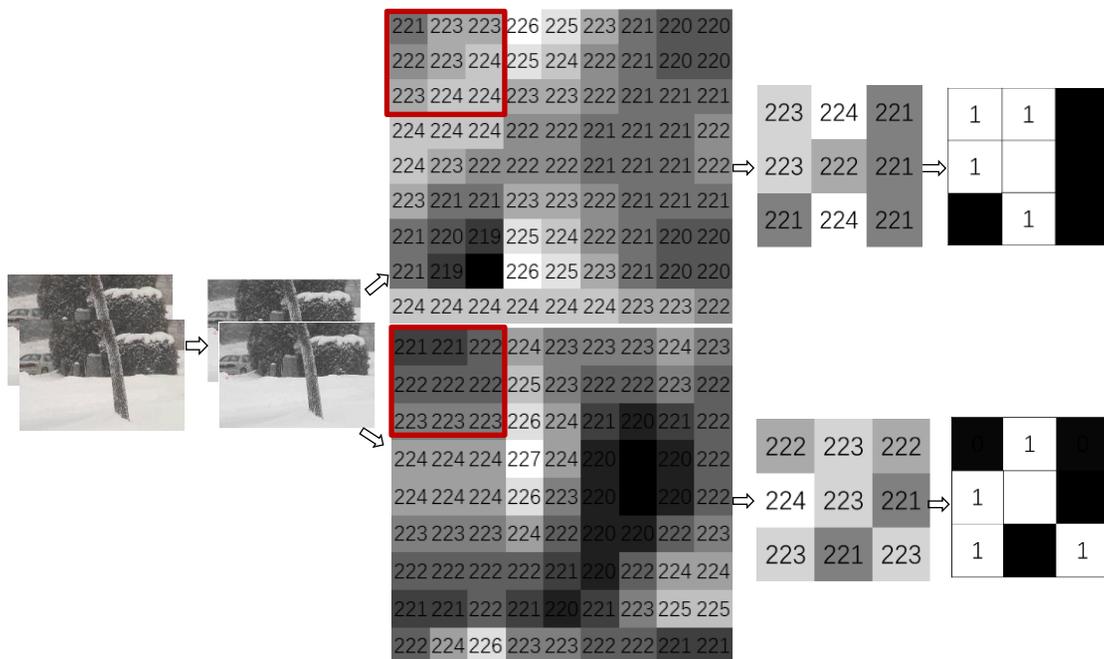


图4 众数背景和第266帧原始的MBLBP
Fig. 4 Original MBLBP of frame 265 and frame 266

到差值如果小于阈值 Th , 则将 \bar{h}_b 值给 \bar{h}_i 值, 否则不变。最后, 与相邻区域灰度平均值进行比较形成 LBP 特征, 得到 S_MBLBP 值。

如图 5 所示, 5(a) 为 266 帧 S_MBLBP 流程图, 与图 4 相比, 没有前景出现时 S_MBLBP 值与背景相同, 图 5(b) 为有目标前景时, 该位置处的 S_MBLBP 改变。

$$S_MLBP_{N,R}(X,Y) = \sum_{N=0}^{N-1} S(\bar{h}'_N - \bar{h}_c) \times 2^N, \quad (6)$$

$$\bar{h}_i = ARGh, \quad (7)$$

$$\bar{h}'_N = \begin{cases} \bar{h}'_N, & |\bar{h}_i - \bar{h}_b| > T_h \\ \bar{h}_b, & |\bar{h}_i - \bar{h}_b| \leq T_h \end{cases}, \quad (8)$$

$$S(H) = \begin{cases} 1, & H \geq 0 \\ 0, & H < 0 \end{cases}, \quad (9)$$

\bar{h}_c 表示中心小区域的灰度值, 以小区域 (X, Y) 圆心, 半径 R 的圆区域称为邻域, N 表示领域像素点的个数, \bar{h}'_N 为改进后的邻域灰度值, \bar{h}_b 为参考背景模

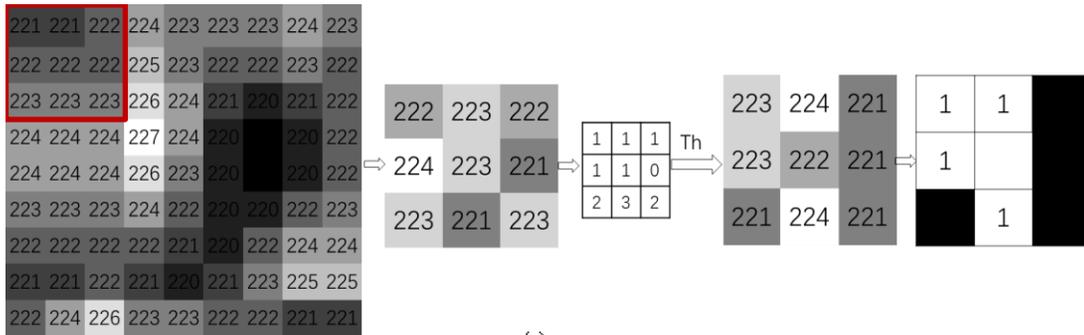
型帧的均值灰度值, h 为小区域内各点的像素值。

1.3 前景提取

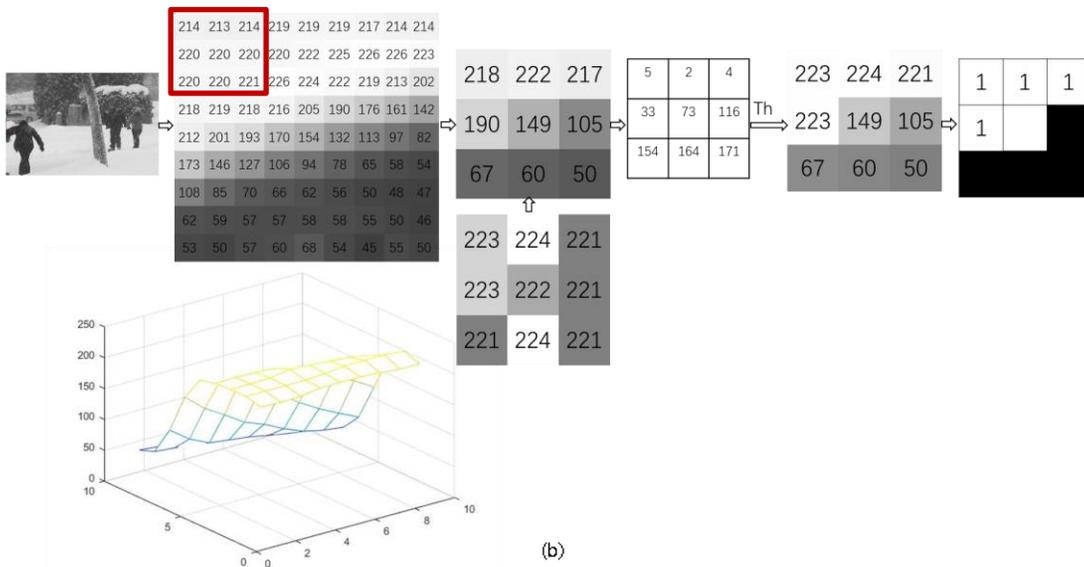
我们使用众数进行背景建模解决了有限的存储空间问题, 并且又保证了准确性, 而使用参数模型难以准确地刻画背景的实际情况。Vibe 算法^[30]和 KDE 算法^[29]均是无参数的方法, Vibe 采用先入先出的方式进行随机更新, 但是使用固定阈值 (经验值为 20), 需要反复的实验才能得到较好的效果; KDE 也同样需要人为设置参数, 并且参数也需要大量的实验才能得到理想的数值, 不具有普遍性, 因此提出了一种自适应阈值的前景判断方法, 能在复杂的环境中检测出前景区域:

$$T_1 = 1 - \begin{cases} 1, & F_{(x,y)} > 0.1F_B \\ 0, & otherwise \end{cases}, \quad (10)$$

$$T_2 = 1 - \begin{cases} 1, & F_{(x,y)} > 0.5F_B \\ 0, & otherwise \end{cases}, \quad (11)$$



(a)



(b)

图 5 S_MBLBP 示意图: (a) 改进的 S_MBLBP , (b) 有目标时的 S_MBLBP

Fig. 5 The schematic diagram of S_MBLBP : (a) improved S_MBLBP , (b) S_MBLBP with a target

$$T = (T_1 \ominus M_{\text{dilate}}) \cap T_2 \quad , \quad (12)$$

式中: T_1 为第一级前景掩模; T_2 为第二级前景掩模; T 为最终前景的掩模; \ominus 为腐蚀符号; M_{dilate} 为 3×3 单位矩阵。

首先使用2.1节中的算法对LBP值的频次统计得到LBP直方图,并且得到基于LBP值的众数背景建模更新背景,结果如图6(d)所示。从图6(b)我们可知,在有目标经过该点的时候,它的LBP值会发生较明显的变化,而经过S_MBLBP处理后得到LBP值减少了很多灰度值图像中因树叶晃动造成的变化,如图6(a)所示,因此式(10)经过多次实验使用最高频次的LBP值频次的0.1作为阈值,然后取反来计算第一级前景,较低的阈值可以在保证多数

前景被识别的同时,排除掉一些偶然出现的噪声杂点,但是很多背景也以散点的形式出现在了前景中;式(11)使用最高LBP值频次的0.5作为阈值,再进行取反操作,得到第二级前景。较高的阈值可以减少背景被误检到前景,但是也将少部分的前景识别到了背景中。因此,式(12)通过对 T_1 腐蚀操作然后来判断第一级前景是否存在第二级前景,只有存在才会被视为最终的前景。

经过上述操作后,进行形态学操作,记每帧处理得到的二值图中每个连通域为 R_i ,所含像素 $(x, y) \in R_i$,以及连通域中孔洞为 H_i ,所含像素点 $(x, y) \in H_i$,设操作Area可以得到连通域内包含像素点数量。前景消除小于10 pixels的连通域式(13)并

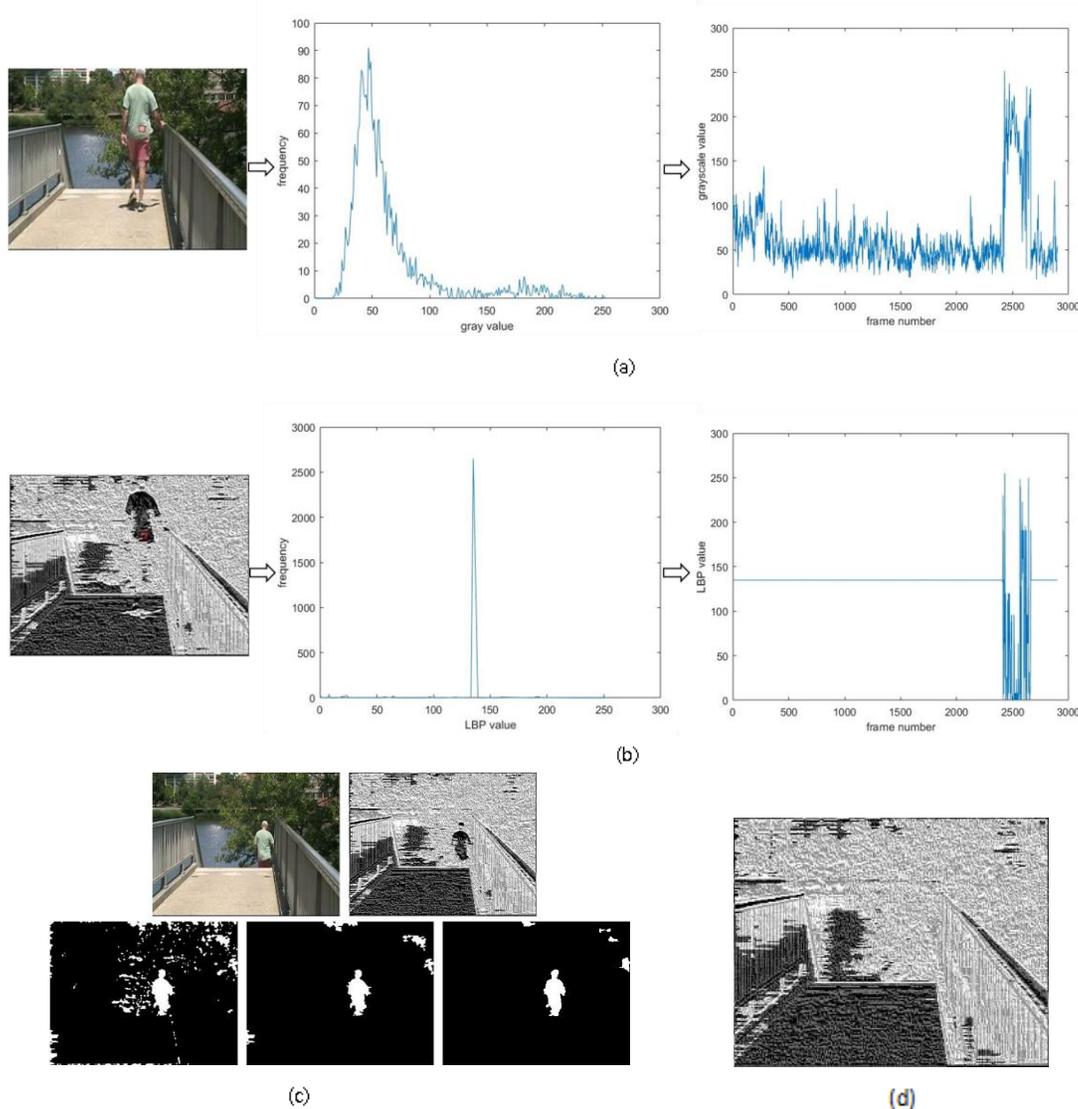


图6 前景目标的提取:(a)灰度值变化图和灰度直方图,(b)LBP值变化图和LBP直方图,(c)前景提取过程,(d)LBP背景建模

Fig. 6 The foreground target extraction:(a) gray value change diagram and gray histogram,(b) LBP value change diagram and LBP histogram,(c) foreground extraction process,(d) LBP background modeling

填充前景内部小于 50 pixels 的孔洞式(14),最终可以得到干净的结果。从图 6(b)可以看到,第一阶段的前景中包含了很多闪烁背景的散点,第二阶段的人体腿部存在错误的孔洞;经过式(12)处理及过形态学运算后,人体的孔洞被填充,而目标的轮廓也可以比较接近真实的形状,从而证明该方法有效。

$$R_i(x,y) = \begin{cases} 0, & Area < 10 \\ 255, & else \end{cases}, \quad (13)$$

$$H_i(x,y) = \begin{cases} 255, & Area < 50 \\ 0, & else \end{cases}. \quad (14)$$

2 实验结果分析

为了能够对所提算法进行定量分析及性能的验证,使用了 CDnet2014 数据集^[32]中的 4 组复杂可见光场景图像和 5 组红外图像进行对比实验,并与帧差法^[31]、ViBe^[30]、KDE^[29]、GMM^[29]进行对比,对比算法均使用原文默认参数。实验平台的硬件为 i7-9750H 2.6 GHz CPU、16 GB 内存,软件为 Win10、vs2019 和 OpenCV4.2.0,该算法处理不同分辨率图像和平均时长如表 1 所示,红外图像处理平均时长如表 2 所示,可以满足目标检测实时性的需求。

2.1 LBP 结果对比图

图 7 为同一帧图像的各种不同的 LBP 图和前景提取结果图,其中 Skating 为可见光数据集, lakeside 为红外数据集。

通过图 7 可知,LBP 结果与 MBLBP 结果相似,受噪声、动态背景等影响较大,扩展局部二进制模

表 1 可见光数据集平均处理时长

Table 1 Average processing time of visible light datasets

数据集名称	blizzard	skating	snowFall	canoe	overpass
尺寸	720*480	540*360	720*480	320*240	320*240
平均用时	70.9077ms	33.7383ms	59.4381ms	48.2896ms	50.9535ms

表 2 红外数据集平均处理时长

Table 2 Average processing time of infrared datasets

数据集名称	corridor	diningRoom	lakeSide	library	park
尺寸	320*240	320*240	320*240	320*240	320*240
平均用时	27.8161ms	18.246ms	23.6167ms	25.8494ms	54.8535ms

式 (Extended Local Binary Patterns, ELBP) 与 uniform MBLBP 由于得到 LBP 较暗,对比度太低无法提取准确的前景。所提算法与前面几种算法比较,图像稳定效果明显,前景图像误检像素点较少,前景目标更准确完整。

2.2 复杂场景的对比实验

由于复杂场景为可见光数据集,图像对比度较高,且存在大面积运动背景,通过大量实验选取 $Th=50$ 时,能有效抑制动态背景、噪声等影响。

图 8 中复杂场景所选取的数据集为 Cdnet2014 数据集中的 Bad weather 的 Skating 图 8(a)、Snowfall 图 8(b)、Dynamic background 中的 Overpass 图 8(c)

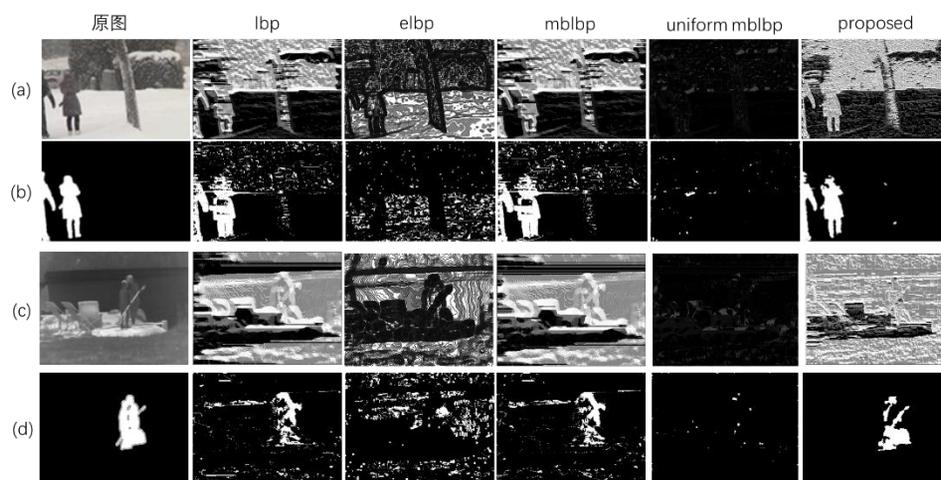


图 7 所提算法与各种 LBP 算法对比结果图:(a) Skating 中不同 LBP 算法对比,(b) Skating 中不同 LBP 算法运动目标提取结果对比,(c) lakeside 中不同 LBP 算法对比,(d) lakeside 中不同 LBP 算法运动目标提取结果对比

Fig. 7 Comparison results between the proposed algorithm and various LBP algorithms: (a) comparison of different LBP algorithms in Skating, (b) comparison of moving object extraction results from different LBP algorithms in Skating, (c) comparison of different LBP algorithms in Lakeside, (d) comparison of moving object extraction results from different LBP algorithms in Lakeside

和 Canoe 图 8(d)。不同算法的主观对比结果如图 8 所示。

图 8(a)所示的数据集为恶劣天气下的室外场景,主要挑战在于大片飘动的雪花和树叶上的雪花作为动态背景与前景的区分,以及前景目标的间歇运动和停留;图 8(b)所示的数据集同为室外场景,主要的挑战为空中持续性出现密集的雪花以及车灯容易造成像素值突变等干扰;图 8(c)所示的数据集的挑战主要为风吹动树叶造成大幅度持续晃动的动态背景中的目标检测;图 8(d)所示的数据集的挑战为提取湖面大区域不停闪烁动态背景下的前景。

从图 8 可以看到,在这四个复杂的室外场景中,相较其他算法,所提算法检测得到目标前景具有更好的完整性和准确性,尤其是对天气的影响、树叶的抖动和湖面的闪烁等动态背景有很好的抑制作用,得到的前景目标误检像素点少,并且前景的轮廓更符合实际情况。图 8(a)中除了 Vibe 算法和所提算法,其余算法都将大部分飘动的雪花以及树叶上的积雪当作了前景,但是所提算法中仍然有比较明显的误检点;图 8(b)中只有 Vibe 算法和所提算法将前景目标比较完整地提取出来了,但是所提算法的前景具有更好的完整性;图 8(c)中帧差法和 KDE 算法受动态背景影响较大,误将晃动的树叶当作了前景,Vibe 算法和 GMM 算法对动态背景有一定的抑制作用,但是 Vibe 算法的前景有比较明显的空洞,GMM 算法提取的前景比较完整,但是所提算法在较完整检测出了前景的同时也很好地抑制了动

态背景的影响;在图 8(d)中能够明显地看出本算法对动态背景有很好的抑制作用,Vibe 算法效果虽然比其他三个算法都好,但是对湖面大区域闪烁的动态背景抑制仍然不如所提算法,不过所提算法对船上人的提取并不完整。

2.3 红外场景的对比实验

由于红外数据集中的动态背景影响较小,在对比度低的情况下,还需抑制图像解压编码时产生的噪声,通过多次实验选取 $Th=20$ 。

图 9 中所选取的数据集为 Cdnet2014 数据集中的 thermal 中的 dining room 图 9(a)、lakeside 图 9(b)、library 图 9(c)、park 图 9(d)、corridor 图 9(e)。不同算法的主观对比结果如图 9 所示。

图 9(a)所示的数据集主要挑战为室内复杂环境中前景的检测,以及前景目标的间歇运动和停留;图 9(b)所示的数据集主要难点是大面积动态背景下的前景检测;图 9(c)所示数据集的难点为前景目标长时间的停留,读者在图书馆看书为慢速目标;图 8(d)所示的数据集中图像场景为公园并且本身存在大量噪点,难点为信噪比低、对比度低的红外场景中的目标检测;图 9(e)所示的数据集的主要挑战是室内环境中地面以及墙面影子的干扰。

从图 9 中可以看出,在各类红外场景中,所提的算法都能较好地提取出完整的前景目标。图 9(a)中帧差法和 Vibe 算法的前景目标都出现了较大的空洞,KDE 和 GMM 算法出现了较多明显的误检,所提算法检测到的前景目标比较符合实际情况;图 9(b)和图 9(d)中背景和前景的灰度差异比较小,其



图 8 复杂场景下不同算法处理结果:(a) skating, (b) snowfall, (c) overpass, (d) canoe

Fig. 8 Processing results of different algorithms in complex scenarios:(a) skating, (b) snowfall, (c) overpass, (d) canoe

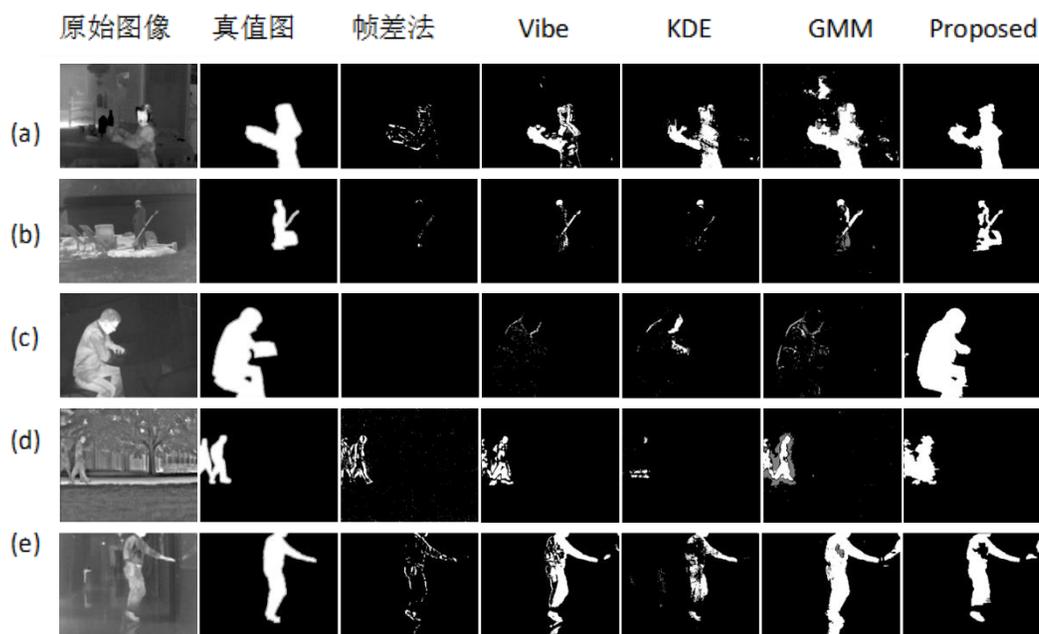


图9 红外场景下不同算法处理结果:(a) dining room, (b) lake side, (c) library, (d) park, (e) corridor

Fig. 9 The processing results of different algorithms in the infrared scene: (a) dining room, (b) lake side, (c) library, (d) park, (e) corridor

他算法提取的前景都在一定程度上不完整,所提算法目标提取较为完整,但是图9(d)中所提算法目标边缘处还是存在部分条状噪声;图9(c)中,帧差法已经将前景目标完全吸收为了背景,除了所提的算法,其他算法在不同程度上都将大部分前景目标当成了背景;图9(e)中,帧差法仅提取出了人物的边缘,Vibe和KDE的前景图像有很大的空洞,而GMM算法对前景的提取比较完整,但是把地面的倒影和墙面的影子也当成了前景,所提算法没有将地面倒影误检为前景,但是有比较明显的空洞。

2.4 定量分析

为了对所提算法进行定量分析,采用了CDet2014中使用召回率Re、准确率Pr、F测度、误检百分比PPWC四个评价指标来衡量。Re用来衡量结果的完整性;Pr和PPWC用来衡量结果的正确性;F测度用来衡量算法的综合性能。其公式分别为:

$$Re = \frac{f_{TP}}{f_{TP} + f_{FN}} \quad (15)$$

$$Pr = \frac{f_{TP}}{f_{FP} + f_{TP}} \quad (16)$$

$$F = \frac{2 \times Re \times Pr}{Re + Pr} \quad (17)$$

$$P_{PWC} = \frac{f_{FP} + f_{FN}}{f_{TP} + f_{TN} + f_{FP} + f_{FN}} \times 100\% \quad (18)$$

f_{TP} 为正确检测为前景的点, f_{TN} 为正确检测为背

景的点, f_{FP} 为误检为前景的点, f_{FN} 为误检为背景的点。一般情况下, R_e 、 P_r 、 F 测度均越大越好, P_{PWC} 越小越好。5种算法的分析结果如图10所示。

图10中绿色的点线为所提算法,纵坐标范围为0到1,横坐标依次为 R_e 、 P_r 、 F 、 P_{PWC} 。可以看到,所提算法在绝大部分场景中都有着较好的指标,9个场景中,所提算法的F测度均为最高值,证明所提算法综合性能优秀。

3 结论

首先,为了解决目标长时间不动或者缓慢运动时,目标被吸收成为背景、由于灰度值微弱变化引起LBP值跳变和动态背景等问题,提出了一种基于众数的背景建模方法,得到参考背景。然后提出了一种对MBLBP的改进算法——S_MBLBP算法,该算法对比背景和数据集分成一个个小区域得到的平均值来稳定图像,若差值小于阈值则数据集取参考背景的平均值,否则保持原来的值不变。最后,提出一种前景提取方法,用所有数据集的S_MBLBP直方图划分不同梯度进行前景提取。实验结果表明,所提算法可以极大地提高多种复杂背景环境下和红外场景中运动目标检测的召回率、准确率,并且降低了误检百分比,综合指标均优于其他算法,在目标静止不动后也不会将其吸收为背景。但是,

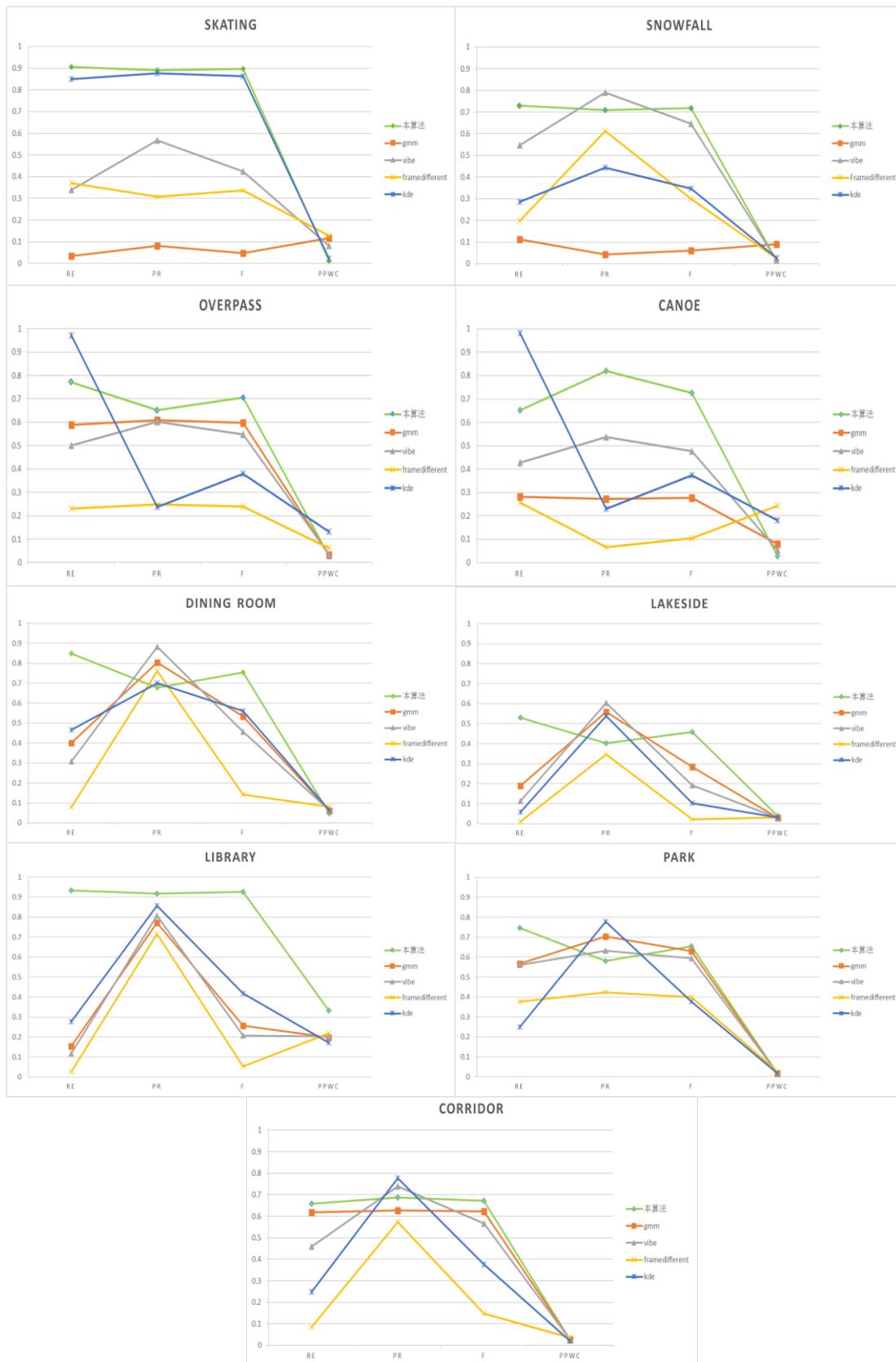


图 10 不同算法在 9 个场景下的对比
Fig. 10 Comparison of different algorithms in 9 scenarios

表 3 不同算法的评价指标结果

Table 3 Evaluation index results of different algorithms

场景	skating				snowfall				overpass			
算法	Re	Pr	F	Ppwc	Re	Pr	F	Ppwc	Re	Pr	F	Ppwc
本算法	0.9059	0.8900	0.8978	0.0126	0.7286	0.7084	0.7183	0.0147	0.7712	0.6504	0.7056	0.0269
gmm	0.0351	0.0805	0.0489	0.1173	0.1127	0.0416	0.0608	0.0898	0.5877	0.6088	0.5981	0.0330
vibe	0.3396	0.5675	0.4249	0.0810	0.5464	0.7908	0.6463	0.0154	0.5011	0.6031	0.5474	0.0346
framedifferent	0.3701	0.3074	0.3359	0.1289	0.1982	0.6124	0.2995	0.0239	0.2300	0.2490	0.2391	0.0611
kde	0.8493	0.8762	0.8625	0.0238	0.2854	0.4444	0.3476	0.0276	0.9704	0.2364	0.3801	0.1321

场景	canoe				Dining room				lakeside			
算法	Re	Pr	F	Ppwc	Re	Pr	F	Ppwc	Re	Pr	F	Ppwc
本算法	0.6518	0.8211	0.7267	0.0270	0.8479	0.6783	0.7537	0.0472	0.5302	0.4022	0.4574	0.0377
gmm	0.2810	0.2729	0.2769	0.0807	0.3998	0.8040	0.5340	0.0595	0.1906	0.5587	0.2842	0.0287
vibe	0.4273	0.5376	0.4761	0.0517	0.3083	0.8821	0.4569	0.0625	0.1142	0.6035	0.1920	0.0288
framedifferent	0.2579	0.0654	0.1044	0.2434	0.0781	0.7603	0.1417	0.0804	0.0114	0.3456	0.0221	0.0302
kde	0.9808	0.2307	0.3736	0.1809	0.4661	0.6991	0.5593	0.0624	0.0564	0.5400	0.1021	0.0297

场景	library				park				corridor			
算法	Re	Pr	F	Ppwc	Re	Pr	F	Ppwc	Re	Pr	F	Ppwc
本算法	0.9334	0.9177	0.9255	0.334	0.7458	0.5816	0.6535	0.0185	0.6581	0.6868	0.6722	0.0235
gmm	0.1540	0.7717	0.2568	0.1983	0.5678	0.7032	0.6288	0.0158	0.6189	0.6270	0.6229	0.0275
vibe	0.1188	0.8084	0.2071	0.2022	0.5607	0.6328	0.5946	0.0180	0.4598	0.7396	0.5671	0.0258
framedifferent	0.0273	0.7169	0.0526	0.2187	0.3766	0.4245	0.3991	0.0267	0.0848	0.5742	0.1478	0.0359
kde	0.2772	0.8574	0.4190	0.1710	0.2480	0.7766	0.3759	0.0194	0.2480	0.7766	0.3759	0.0194

所提算法在对比度过低的场景中效果较差,会有目标检测不完整的问题,背景建模的鬼影问题还未解决。接下来将进一步对算法进行优化,主要研究如何在所提算法中快速消除鬼影,以及对检测到的目标进行识别和跟踪。

References

- [1] JAVED S, MAHMOOD A, ALMAADEED S, *et al.* Moving object detection in complex scene using spatiotemporal structured-sparse RPCA [J]. *IEEE Transactions on Image Processing*, 2019, **28**(2): 1007-1022.
- [2] Eltantawy A, Shehata M S. An accelerated sequential PCP-based method for groundmoving objects detection from aerial videos [J]. *IEEE Transactions on Image Processing*, 2019, **28**(12): 5991-6006.
- [3] REZAEI B, OSTADABBAS S. Moving object detection through robust matrix completion augmented with objectness [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2018, **12**(6): 1313-1323.
- [4] ZHENG P, BAI H Y, LI Z M, *et al.* Design of accurate detection and tracking algorithm for moving target under jitter interference [J]. *Chinese Journal of Scientific Instrument*. 2019, **40**(11): 90-98
- [5] Wang Cong, Liu Mingguang, QI Fei. Review of dynamic target detection and recognition algorithms for intelligent video surveillance system [J]. *Electric technology*, 2018, **19**(09): 6-11. (王聪, 刘明光, 齐飞. 智能视频监控动态目标检测与识别算法综述 [J]. *电气技术*), 2018, **19**(09): 6-11.
- [6] Liu Pengpeng, Lyu Michael R, King Irwin, *et al.* Learning by Distillation: A Self-Supervised Learning Framework for Optical Flow Estimation. [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2021, PP.
- [7] Sengar S S, Mukhopadhyay S. Moving object area detection using normalized self-adaptive optical flow [J]. *Journal for Light and Electronoptic*, 2016, **127**(16): 6258 - 6267.
- [8] GUO Y, LI Z, LIU Y, *et al.* Video object extraction based on spatiotemporal consistency saliency detection [J]. *IEEE Access*, 2018, **6**: 35171-35181.
- [9] ABBASIFARD M R, NADERI H, ALAMDARI O I, *et al.* Efficient indexing for past and current position of moving objects on road networks [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, **19**(9): 2789-2800.
- [10] LU J, WANG Z, ZHU J. Space-time multiscale based moving object detection method [J]. *Xibe Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University*, 2017, **35**(1): 98-102.
- [11] Gao Shibo, Cheng Yongmei, Zhao Yongqiang, *et al.* Journal of infrared and millimeter waves, 2014, **33**(05): 498-506. (程咏梅, 赵永强, 等. 数据驱动二次相关滤波器红外目标检测 [J]. *红外与毫米波学报*), 2014, **33**(05): 498-506.
- [12] PANDO A G, MURGUIA M I. Analysis and trends on moving object detection algorithm techniques [J]. *IEEE Latin America Transactions*, 2019, **17**(11): 1771-1783.
- [13] Kong JUN, Tang Xinyi, Jiang Min, *et al.* Kalman filter tracking based on multi-scale feature extraction [J]. *Journal of infrared and millimeter waves*, 2011, **30**(05): 446-450. (孔军, 汤心溢, 蒋敏, 等. 基于多尺度特征提取的 Kalman 滤波跟踪 [J]. *红外与毫米波学报*), 2011, **30**(05): 446-450.
- [14] Romero J D, Lado M J, Mendez A J, *et al.* A background modeling and foreground detection algorithm using scaling coefficients defined with a color model called lightnessred-green-blue [J]. *IEEE Transactions on Image Processing*, 2018, **27**(3): 1243-1258.

- [15] Wren C R, Azarbayejani A, Darrell T, *et al.* Pfinder: real-time tracking of the human body [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, **19**(7): 780–785.
- [16] Stauffer C, Grimson W. Adaptive Background Mixture Models for Real-Time Tracking [C]// *Computer Vision and Pattern Recognition*, 1999. IEEE Computer Society Conference on. IEEE, 1999.
- [17] Niranjil KA, Sureshkumar C. Background subtraction in dynamic environment based on modified adaptive GMM with TTD for moving object detection [J]. *Journal of Electrical Engineering & Technology*, 2015, **10**(1): 372–378.
- [18] Zhang Xiwei, ZHANG Yanning, Liang Jun. Multi-source collaborative Detection of moving Target based on Foreground and background Discriminability Evaluation Factor [J]. *Journal of infrared and millimeter waves*, 2015, **34**(05): 619–629. (张秀伟, 张艳宁, 梁君. 基于前景—背景可区分性评价因子的运动目标多源协同检测 [J]. *红外与毫米波学报*), 2015, **34**(05): 619–629.
- [19] ELGAMMAL A, HARWOOD D, DAVIS L S, *et al.* Non-parametric model for background subtraction [C]// *European Conference on Computer Vision*, 2000: 751–767.
- [20] 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.
- [21] DING Y, WENHUI L, JINGTAO F, *et al.* Robust moving object detection under complex background [J]. *Computer Science and Information Systems*, 2010, **7**(1): 201–210
- [22] HEIKKILA M, PIETIKAINEN M. A texturebased method for modeling the background and detecting moving objects [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(4): 657–662.
- [23] Kumar P, Singhal A, Mehta S, *et al.* Real-time moving object detection algorithm on high-resolution videos using GPUs [J]. *Journal of Real-Time Image Processing*, 2016, **11**(1): 93–109.
- [24] Dong Lili, Zhang Tong, Ma Dongdong, *et al.* Infrared image Classification of Sea Background based on Directional Gradient Histogram and Local Contrast Feature [J]. *Journal of infrared and millimeter waves*, 2020, **39**(05): 650–658. (董丽丽, 张彤, 马冬冬, 等. 基于方向梯度直方图和局部对比度特征的海面背景红外图像分类 [J]. *红外与毫米波学报*), 2020, **39**(05): 650–658.
- [25] Chua T W, Leman K, Yue W. Fuzzy rule-based system for dynamic texture and color based background subtraction [J]. *IEEE*, 2012.
- [26] Song T, Ou L I, Cui H L. A Moving Object Detection Method Based on Scene Perception [J]. *Acta Electronica Sinica*, 2016.
- [27] Lai A N, Yoon H, Lee G. Robust background extraction scheme using histogram-wise for real-time tracking in urban traffic video [C]// *Computer and Information Technology*, 2008. CIT 2008. 8th IEEE International Conference on. IEEE, 2008: 845–850.
- [28] Liao S, Zhu X, Zhen L, *et al.* Learning Multi-scale Block Local Binary Patterns for Face Recognition [C]// *Proceedings of the 2007 international conference on Advances in Biometrics*. DBLP, 2007.
- [29] Wu Jia-wen, Wang Shi-yong. Adaptive background modeling algorithm for gray video based on statistics [J]. *Chinese Journal of Lasers*, 2011, **48**(03): 113–125. (武加文, 王世勇. 基于统计的灰度视频自适应背景建模算法 [J]. *中国激光*), 2011, **48**(03): 113–125.
- [30] Min W D, Guo X G, Han Q. An improved ViBe algorithm and its application in traffic video processing [J]. *Optics and Precision Engineering*, 2017, **25**(3): 806–811.
- [31] Gao L H. Research on Moving target Tracking Algorithm [D]. Hebei University of Science and Technology, 2019. (高丽慧. 移动目标跟踪算法研究 [D]. 河北科技大学, 2019.
- [32] Yi W, Jodoin PM, Porikli F, *et al.* CDnet 2014: An expanded change Detection Benchmark Dataset [J]. *Artificial Intelligence & Image Processing*.