

## 基于统计的灰度视频自适应背景建模算法

武加文<sup>1,2</sup>, 王世勇<sup>1\*</sup>

<sup>1</sup>中国科学院上海技术物理研究所红外探测与成像技术重点实验室, 上海 200083;

<sup>2</sup>中国科学院大学, 北京 100049

**摘要** 针对灰度视频的目标检测依赖先验知识、召回率低以及单一算法无法同时兼顾静态与动态背景等问题, 提出一种基于统计的背景建模算法。该算法无需先验知识, 根据统计信息可以准确区分静态背景和动态背景, 并采取不同的检测策略提取目标。对于静态背景, 采用改进的三帧差分法自适应设置阈值, 可以保证较高的召回率。对于动态背景, 采用改进的概率密度估计法可以有效降低虚警率。采用所提算法对光照变化以及阴影进行处理, 可以进一步提升算法的性能。在公开数据集与实际采集红外数据进行验证实验。实验结果表明, 所提算法在多种场景中处理灰度视频的结果比其他传统算法好, 在保证准确率的同时可以极大地提升召回率, 并且有效提高目标的完整性。

**关键词** 机器视觉; 运动目标检测; 背景建模; 自适应阈值; 光照变化

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/CJL202148.0309001

### 1 引言

目标检测一直是计算机视觉领域的一个研究重点, 而且应用场景十分广泛, 如安防告警、家庭监控、道路交通、行车导航和视频会议等<sup>[1]</sup>。国内外科学家们提出了许多目标检测算法, 如光流法、帧差法和背景差分法等传统算法, 以及近年来发展迅速的深度学习算法。相比于传统算法, 深度学习算法具有拟合能力强和准确度高的优点, 但性能的好坏依赖于数据集的大小。部分特殊场景(红外场景和未知载具场景)因缺乏样本不能满足随时部署的需求, 对于少样本或零样本依然没有有效的解决方案, 并且深度学习算法对于硬件的要求显著高于传统算法, 这在很大程度上会增加成本。鉴于此, 传统方法依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, 实际中普遍采用基于背景建模的目标检测方法。基于灰度视频的背景建模存在视频精度较低、纹理缺乏和噪声较大的问题, 解决该问题具有更大的挑战性<sup>[2]</sup>。Wren 等<sup>[3]</sup>使用了单高

斯模型描述背景, 该模型适用于静态稳定的背景, 难以胜任复杂环境的目标检测任务。为了解决单高斯模型难以描述复杂背景的问题, Stauffer 等<sup>[4]</sup>使用了混合高斯模型(GMM)拟合背景, 可以实现复杂背景的目标检测, 但当遇到简单背景时, 该模型会错误地学习前景灰度, 导致目标丢失, 并且训练复杂度高, 阈值选取困难。除了上述参数模型外, 还有无参数模型, 如码本模型<sup>[5]</sup>为每个像素建立多个码字并计算每个码字的上下界, 能够有效实现不同场景的分割, 但该模型一旦在建立过程中出现前景, 则会导致码字错误, 直到下次更新码字前都会一直误检, 且无法适应较剧烈的环境变化。Barnich 等<sup>[6]</sup>提出的无参数模型 ViBe 能够快速建模, 使用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。闵卫东等<sup>[7]</sup>使用了改进的 ViBe 模型对鬼影问题进行了优化, 但仍无法解决慢速前景图像存在空洞的问题。Elgammal 等<sup>[8]</sup>提出了核密度估计(KDE)法, 通过统

收稿日期: 2020-06-29; 修回日期: 2020-07-27; 录用日期: 2020-09-09

基金项目: 上海市现场物证重点实验室基金(2017xcwzk08)

\* E-mail: wangshiyong@mail.sitp.ac.cn

计像素点的直方图可以得到背景分布,并采用 KDE 法计算每个像素点为背景的概率,但仅能得到概率,因阈值的设置困难,所以计算概率的复杂度较高。

上述算法都需要针对不同的场景设置不同的阈值,实用性差,而且难以兼顾动态背景、静态背景、高速目标和慢速目标,并且处理灰度视频易丢失部分前景,无法判断阴影。除了 ViBe 模型以外,上述算法都难以在光照发生变化时可以快速适应或重建。针对这些问题,本文提出一种基于直方图统计的建模算法,该算法通过直方图的众数估计背景,避免慢速或静止目标导致背景飘移。首先使用直方图可以有效地区分静态背景和动态背景,并对不同背景使用不同的判断策略,接着采用改进的三帧差分法计算自适应阈值,可以有效地提高复杂场景下的召回率和准确率,最后使用差分图像估计亮度变化与阴影区域,避免阴影与亮度变化对灰度视频目标检测的影响。

## 2 背景建模算法

采用在线更新的方法对常见的 8 bit 灰度视频中的每个像素点建立灰度直方图,相比于基于高斯分布和泊松分布模型的方法,该方法更能反映像素灰度的真实分布,而基于样本的建模方法<sup>[6-10]</sup>在使用相同数量级内存的情况下可以记录更长时间范围内的样本分布,并且复杂度更低。

### 2.1 基于直方图众数的背景估计

文献[11]将出现频率最高的灰度值作为真实背景值,而且该实验使用固定长度的图片序列构建直方图。若图片序列的长度较短,则慢速目标会被错误当作背景,若长度增加,则会导致内存消耗太多。鉴于此,对图像中的每一个像素点 $(x, y)$ 建立灰度直方图并统计不同灰度值出现的次数,该直方图会随着输入图像帧数 $p$ 的增加不断更新,记为 $H_{(x,y),p}$ 。 $H_{(x,y),p}$ 中灰度 $i$ 出现的次数记为 $H_{(x,y),p}(i)$ ,由于实验使用 8 bit 深度图像,因此 $i$ 的取值范围为 $[0, 255]$ 。将直方图中出现次数最多的灰度值 $B_{(x,y),t}$ 作为该点的背景灰度值,其中 $t$ 为时刻,表达式为

$$\operatorname{argmax} [f(i)] = \{i \mid \forall j \in [0, 255], f(j) \leq f(i)\}, \quad (1)$$

$$B_{(x,y),t} = \operatorname{argmax} [H_{(x,y),p}(i)], \quad (2)$$

式中: $f(i)$ 和 $f(j)$ 为灰度 $i$ 和 $j$ 出现的频次。

使用直方图统计灰度值出现的频率,无需保存历史图像,为此可以减少内存消耗。此外 $\operatorname{argmax}(\cdot)$ 为非线性函数,可以保证在出现慢速前

景的同时背景灰度不会出现偏移,同时可以避免建模初期 ViBe 算法存在前景而导致鬼影的产生。随着算法的运行,背景越来越稳定,为此可以保证算法长时间运行的准确性和稳定性。通常在直方图中寻找最大灰度值的复杂度为 $O(n)$ ,其中 $n$ 为最大灰度值,因传统方法的复杂度过高,所以根据直方图连续变化的特点,提出一种快速在直方图中寻找最大灰度值的算法。首先记录上一帧直方图最大值对应的灰度为 $a$ ,并统计灰度为 $a$ 的次数 $k$ ;然后在当前帧的像素点 $(x, y)$ 输入灰度为 $b$ ,并统计灰度为 $b$ 的次数 $l$ ;最后比较 $k$ 和 $l$ ,若 $k > l$ 则当前帧直方图最大值对应灰度为 $a$ ,否则为 $b$ 。

按照上述步骤,得到复杂度为 $O(1)$ ,这可以极大地提高背景估计的速度,而且比 GMM 算法更快。图 1 为不同算法针对不同运动状态场景的背景建模结果,其中 image 1 为行驶的车辆, image 2 为小范围内踱步的行人, image 3 为静止的读者, image 4 为缓慢航行的船只。从图 1 可以看到,所提算法在不同场景中估计得到的背景都最为准确干净,而其他算法均在一定程度上将前景当作背景,进而影响后续的检测能力。在不同的环境中,说明所提算法对不同运动状态的目标均有着良好的适应能力。

### 2.2 基于改进三帧差分的噪声估计

灰度视频的信息比彩色视频少,所以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。多数无参数模型采用固定阈值或全局自适应阈值的方法,高斯模型通过固定速率来更新均值和方差,但均无法适应不同的场景,所以如何准确估计视频内的噪声就成了一个重要的问题。对于数码相机在拍摄视频的过程中,噪声来源有许多,主要有非均匀响应噪声、随机噪声和固有噪声(热噪声、读出电路噪声和放大电路噪声等)三类。除了这些噪声以外,还可能存在视频流传输的过程中由编码引入的噪声等,其中除了随机噪声与亮度平方根呈正相关外,其他噪声均与像素灰度无关。由于任意像素的大部分灰度均为背景灰度,因此假设单一像素点在不同时刻的噪声水平是大致相同的。文献[8]认为同一像素在相邻帧之间通常服从正态分布 $N(\mu, \sigma^2)$ ,则两帧差分结果服从 $N(0, 2\sigma^2)$ ,其中 $\mu$ 为均值, $\sigma$ 为标准差。两帧差分后,除了前景与背景在更替时刻切换帧以外,其余时刻连续两帧的 $\mu$ 值相同,所以两帧差分序列无论当前是背景还是前景均服从同一分布,因此只需计算像素两帧差分序列的方差,就可以估算出 $2\sigma^2$ 的值,进而得到 $\sigma$ 。

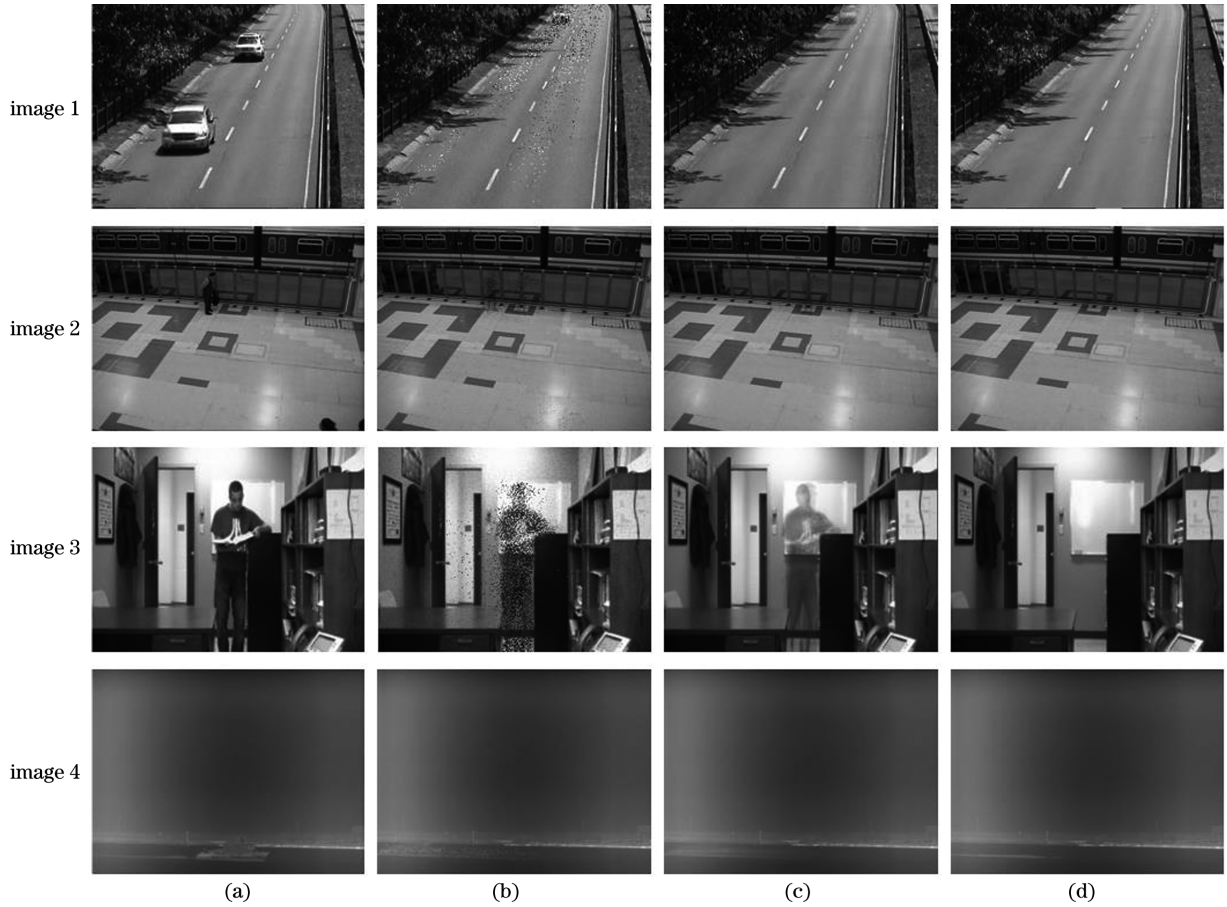


图 1 不同算法的背景建模结果。(a)原始图像;(b)ViBe 算法;(c)GMM 算法;(d)所提算法

Fig. 1 Background modeling results of different algorithms. (a) Original images; (b) ViBe algorithm; (c) GMM algorithm; (d) proposed algorithm

对于简单场景,采用两帧差分法可以有效估计噪声方差,但当像素频繁在前景和背景中切换或者背景为频繁变化的动态背景时,由切换帧引起的误差会导致该方法失效。受到该方法的启发,提出一种基于改进三帧差分估计噪声分布的算法,采用该算法为每个像素计算单独的自适应阈值,进一步提高准确率与召回率。对于时刻  $t$ ,首先对前三帧分别进行两帧差分,并选取绝对值较小的值作为噪声样本  $N_{(x,y),t}$ ,表达式为

$$N_{(x,y),t} = \min(|I_{(x,y),t} - I_{(x,y),t-1}|, |I_{(x,y),t-1} - I_{(x,y),t-2}|), \quad (3)$$

式中: $I_{(x,y),t}$  为  $t$  时刻在  $(x,y)$  处的输入图像。由于绝对值操作以及噪声分布具有对称性,方差计算公式为

$$\sigma_{(x,y),t}^2 = \sigma_{(x,y),t-1}^2 \cdot \frac{t-1}{t} + N_{(x,y),t}^2 \cdot \frac{1}{t}, \quad (4)$$

$$\sigma_{(x,y),t} = \sqrt{\sigma_{(x,y),t}^2}, \quad (5)$$

式中: $\sigma^2$  为方差; $\sigma$  为标准差。

为了验证算法的可行性,使用 CDnet/PETS2006 数据集<sup>[12]</sup> 对其进行验证。图 2(a)为输入的图像序列,方框为观察的像素点位置。图 2(b)为 1~300 frame 图像有前景出现的灰度变化曲线,记为序列 A。图 2(c)为 301~600 frame 图像持续为背景的灰度变化曲线,记为序列 B。图 2(d)为序列 A 和序列 B 经过(3)式变换后的频率变化曲线。从图 2(d)可以看到,两条曲线非常接近,说明该算法不受场景变化的干扰,能够在复杂环境中得到真实的噪声特性。

经过(4)式和(5)式计算得到序列 A 的标准差为 1.99,而序列 B 的标准差为 1.84,两者非常接近。采用蒙特卡洛法验证服从高斯分布的加性噪声序列经过(3)式变化后得到的方差恰为真实方差,因此改进的三帧差分法不会受到前景和动态背景的干扰,而且更能准确地估计噪声方差。

### 2.3 动静态背景划分

单高斯模型适合处理简单背景,而 GMM 处理复杂背景的效果更佳。针对动态背景与静态背景,应当采用不同的算法进行处理。郭治成等<sup>[13]</sup> 提出

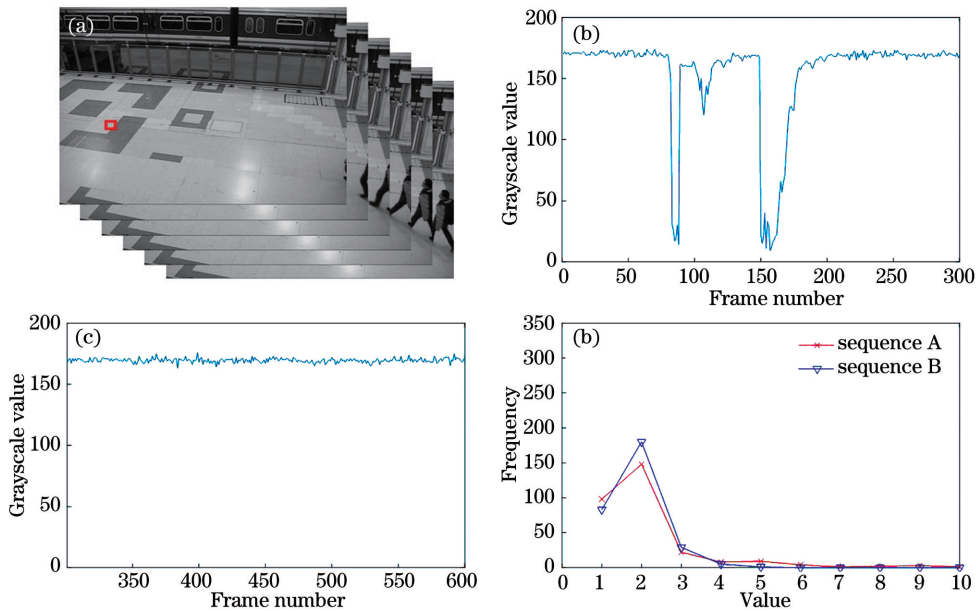


图 2 三帧差分变换算法估计噪声的流程。(a)输入的图像序列;(b)序列 A 的灰度变化曲线;  
(c)序列 B 的灰度变化曲线;(d)序列 A 和序列 B 的频率变化曲线

Fig. 2 Process of noise estimation by three-frame differential transformation algorithm. (a) Input image sequence; (b) grayscale change curve of sequence A; (c) grayscale change curve of sequence B; (d) frequency change curves of sequence A and sequence B

了动态指数对不同的动态区域采用不同阈值,刘伟等<sup>[14-15]</sup>使用了固定阈值来去除 GMM 算法中权重值较小的灰度,从而适应不同的复杂度。上述方法可以在一定程度上解决不同背景下的适应性,但会增加计算量,且阈值的选取会影响结果的准确性。为了解决上

述问题,提出一种快速有效的动静态背景划分方法。

静态背景按 2.2 节的结论可以近似表示为  $N(\mu_{(x,y),t}, \sigma_{(x,y),t}^2)$ ,但动态背景则不能这样来表示,因为动态背景的灰度更为分散。动静态像素的灰度直方图如图 3 所示。从图 3 高速公路场景可以看

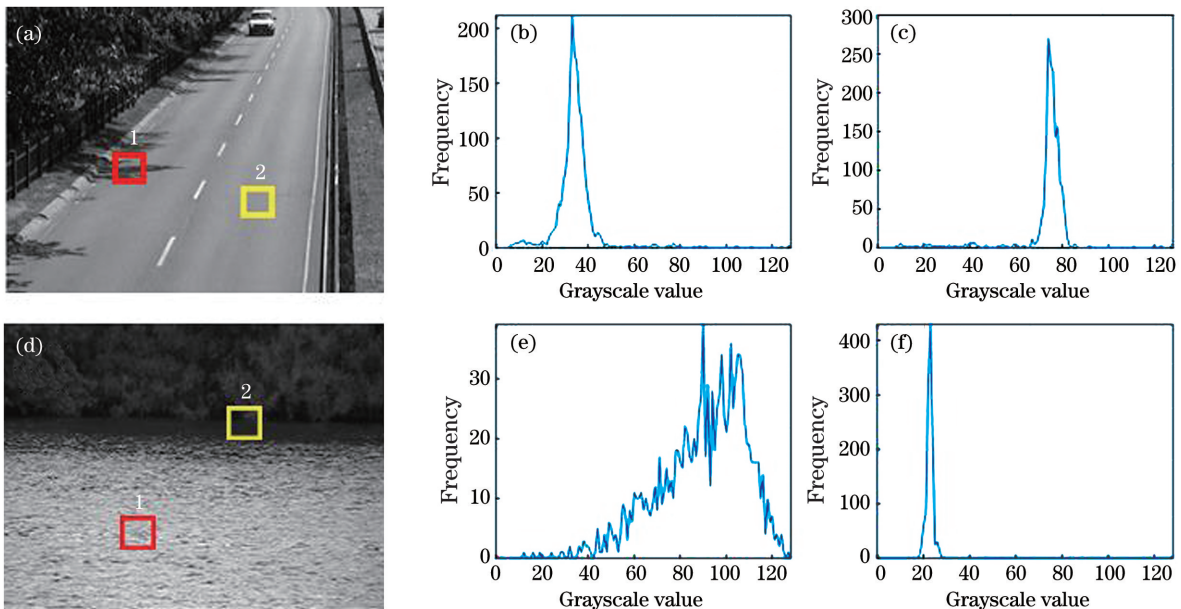


图 3 不同场景中动、静态背景的实验结果。(a)静态背景的原图;(b)图(a)方框 1 的灰度直方图;(c)图(a)方框 2 的灰度直方图;(d)动态背景的原图;(e)图(d)方框 1 的灰度直方图;(f)图(d)方框 2 的灰度直方图

Fig. 3 Experimental results of dynamic or static background under different scenes. (a) Original image of static background; (b) gray histogram of Fig. (a) box 1; (c) gray histogram of Fig. (a) box 2; (d) original image of dynamic background; (e) gray histogram of Fig. (d) box 1; (f) Figure (d) gray histogram of Fig. (d) box 2

到,树枝晃动阴影处的灰度分布在 20~50 pixel 的范围内,而路面处的灰度分布在 70~85 pixel 的范围内,由此可见静态背景的灰度分布更为集中,在湖面场景中同样可以观察到这种特性。

针对动静态背景的特性并结合 2.1 节与 2.2 节得到的结论,提出一种基于直方图的动静态背景划分方法。静态背景的灰度直方图服从  $N(\mu_{(x,y),t}, \sigma_{(x,y),t}^2)$ , 动态背景的灰度直方图则是以  $\sigma_{(x,y),t}^2$  分布在多个背景灰度的周围。若像素属于静态背景,则

其灰度直方图应当服从正态分布。已知均值为  $\mu$  (即直方图最大值对应的灰度),当灰度为  $\mu_{(x,y),t} \pm \sigma_{(x,y),t}$  时,按照正态分布的特性,查表可知直方图在对应灰度的值应满足

$$H_{(x,y),t}(\mu_{(x,y),t} \pm \sigma_{(x,y),t}) \approx 0.6H_{(x,y),t}(\mu_{(x,y),t}). \quad (6)$$

动态背景的灰度直方图分布更分散且平缓,所以不满足这种特征。多次实验后,采用 0.7 倍的峰值来区分动静态背景,表达式为

$$M_{(x,y),t} = \begin{cases} 255, & \min [H_{(x,y),t}(\mu_{(x,y),t} \pm \sigma_{(x,y),t})] > 0.7H_{(x,y),t}(\mu_{(x,y),t}), \\ 0, & \text{otherwise} \end{cases}, \quad (7)$$

式中: $M$  为动静态背景掩模板;255 为动态背景;0 为静态背景。图 4 为基于直方图的动静态背景划分

方法在不同场景中的区分结果。

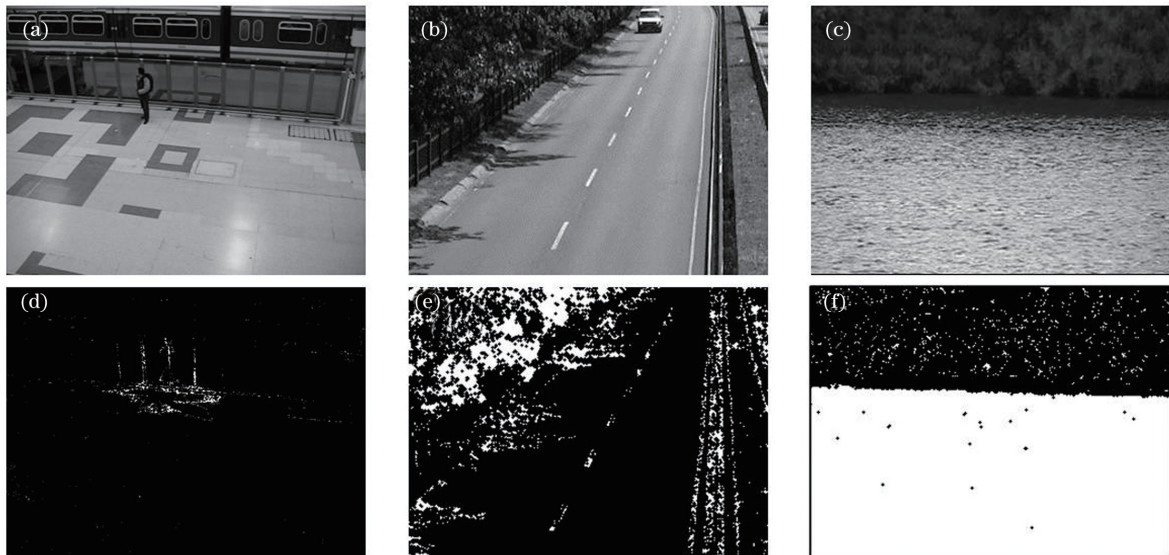


图 4 不同场景下的动、静态背景划分结果。(a)PETS2006 图像;(b)高速公路场景;(c)湖面场景;(d)图(a)的划分结果;(e)图(b)的划分结果;(f)图(c)的划分结果

Fig. 4 Dynamic and static background partition results under different scenes. (a) PETS2006 image; (b) freeway scene; (c) lake scene; (d) division result of Fig. (a); (e) division result of Fig. (b); (f) division result of Fig. (c)

从图 4 可以看到,除了 PETS2006 图像中静态目标边缘有细微运动而被误识别为动态区域以外,该方法的区分结果十分准确;值得注意的是,除了准确提取树木随风摆动的动态背景区域外,道路标线和路沿同样被标记为动态背景区域,这是因为传感器受像元尺寸的限制,图像灰度梯度较大的边沿区域往往会在相邻像素间抖动,所以表现出动态背景的特征。

#### 2.4 前景提取

静态背景的灰度符合单高斯模型分布的特征,提取前景的表达式为

$$D_{\text{sta}(x,y),t} = \begin{cases} 255, & |I_{(x,y),t} - \mu_{(x,y),t}| > 2.5\sigma_{(x,y),t}, \\ 0, & \text{otherwise} \end{cases}, \quad (8)$$

式中: $D_{\text{sta}(x,y),t}$  为静态背景下的输出。

动态背景的灰度分布为动态背景与噪声的叠加,使用参数模型难以准确描述背景的灰度分布。KDE 法无法平衡准确性与高效性,而且需要人为设定阈值来判断像素对应的概率以判断其是否属于前景,故提出一种两级自适应阈值的前景判断方法。该方法利用有限的储存空间,使用全部的历史灰度数据构建直方图,大量的样本既可以保证准确性,又无

需使用核函数平滑概率密度函数,从而大大加快计算速度。两级自适应阈值的前景判断方法的表达式为

$$D_{\text{dyn}1/2(x,y),t} = \begin{cases} 255, & H_{(x,y),t}(I_{(x,y),t}) > H_{(x,y),t}(\mu_{(x,y),t})/2, \\ 0, & \text{otherwise} \end{cases}, \quad (9)$$

$$D_{\text{dyn}1/8(x,y),t} = \begin{cases} 255, & H_{(x,y),t}(I_{(x,y),t}) > H_{(x,y),t}(\mu_{(x,y),t})/8, \\ 0, & \text{otherwise} \end{cases}, \quad (10)$$

$$D_{\text{dyn}} = 255 - (D_{\text{dyn}1/2} \oplus \mathbf{E}_{\text{dilate}}) \cap D_{\text{dyn}1/8}, \quad (11)$$

式中: $D_{\text{dyn}1/2}$  为第一级背景掩模; $D_{\text{dyn}1/8}$  为第二级背景掩模; $D_{\text{dyn}}$  为最终前景的掩模; $\oplus$ 为膨胀符号; $\mathbf{E}_{\text{dilate}}$  为  $3 \times 3$  单位矩阵。

(9)式使用直方图最高频次的  $1/2$  作为阈值以计算出第一级背景,较高的阈值可以在保证低前景误检率的同时,也可以保证大多数背景点被识别。(10)式使用直方图最高频次的  $1/8$  作为阈值,较低的阈值可以识别出大量的背景,但会导致前景内部出现被误判为背景的噪点。由于前景被误判为背景的点通常出现在前景内部,即为“孤点”,因此(11)式通过膨胀操作来判断第二级背景邻域是否存在第一级背景,只有存在才会被视为最终的背景,对结果取

反即可得到前景。

经过上述操作后, $D_{\text{dyn}}$  中存在少量的噪点,通过形态学操作来断开不连续的前景,消除小于  $10$  pixel 的连通域并填充前景内部小于  $50$  pixel 的孔洞,最终可以得到干净的结果。图 5 为两级自适应阈值的前景判断方法在不同阶段的结果。从图 5 可以看到,第一阶段的背景出现较多漏检的现象,第二阶段的船身存在一些错误的孔洞;经过(11)式处理后,船身上的孔洞被抑制,背景的漏检情况未发生(部分静态背景噪点未在本步骤消除);经过形态学运算后,船身的孔洞被填充,而目标的轮廓也可以较好的保留,从而证明该方法的有效性。

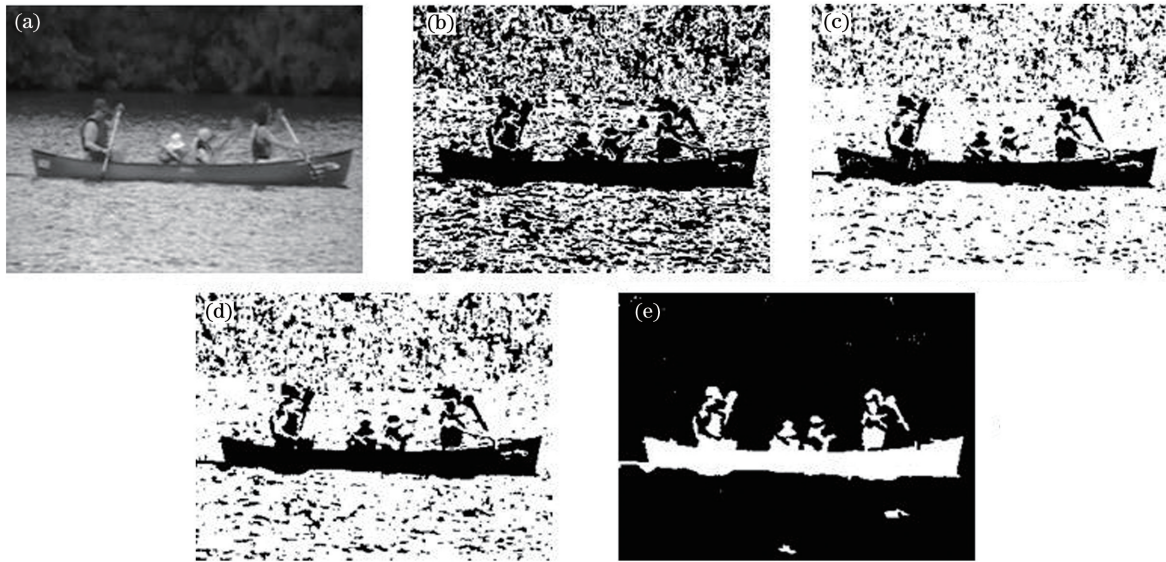


图 5 动态背景下的目标提取流程。(a)原始图像;(b)第一级阈值;(c)第二级阈值;(d)消除误检;(e)形态学操作  
Fig. 5 Target extraction process under dynamic background. (a) Original image; (b) 1<sup>st</sup> level threshold; (c) 2<sup>nd</sup> level threshold; (d) elimination of false detection; (e) morphological operation

前景检测最终的输出可表示为

$$D_{(x,y),t} = \begin{cases} D_{\text{sta}(x,y),t}, & M_{(x,y)} = 0 \\ D_{\text{dyn}(x,y),t}, & M_{(x,y)} = 255 \end{cases} \quad (12)$$

## 2.5 全局光照变化检测

全局光照变化是目标检测中的一个常见问题,这关系到算法的稳定性。光照产生变化的原因分为两种:一是短时间内由设备的动态调整,感

光度或曝光时间的改变会导致灰度发生变化;二是室外场景的环境光变化。当光照发生变化时,所提算法将不再适用,因为该算法假设  $\mu_{(x,y),t}$  值是稳定的,所以需要在光照产生变化时对直方图进行更新。丁洁等<sup>[16-17]</sup>采用了两帧差分法来检测光照突变以重新建模,但当相机频繁动态调整时会反复更新模型,导致检测能力降低。Xu 等<sup>[18]</sup>利

用了两帧差分法的中位数估计光照变化幅度以调整画面灰度,当前景灰度较大时会存在较大误差。为了解决上述问题,在 Xu 等<sup>[18]</sup>工作的基础上对算法进行改进。

首先得到背景图像  $B$  的直方图  $H_B$ , 计算  $H_B$  中出现最多的灰度  $G_B$ , 表达式为

$$G_B = \operatorname{argmax}(H_B). \quad (13)$$

然后找到  $B$  中所有灰度等于  $G_B$  的像素点坐标, 表达式为

$$Q = \{(x, y) \mid B_{(x, y), t} = G_B\}. \quad (14)$$

将剔除当前输入图像  $I$  中所有不属于集合  $Q$

的像素以得到剩余像素的图像  $I_R$ , 表达式为

$$I_{R(x, y)} = \begin{cases} I_{(x, y), t}, & (x, y) \in Q \\ -1, & (x, y) \notin Q \end{cases}. \quad (15)$$

得到  $I_R$  的直方图  $H_I$ , 计算  $H_I$  中出现次数最多的灰度  $G_I$ , 表达式为

$$G_I = \operatorname{argmax}(H_I). \quad (16)$$

最后使用  $G_B$  和  $G_I$  对当前帧的灰度进行调整, 表达式为

$$I_{\text{adj}} = I \cdot G_B / G_I, \quad (17)$$

式中:  $I_{\text{adj}}$  为当前帧调整亮度后的图像。图 6 为光照调整算法处理静态背景第 297 frame 图像的处理

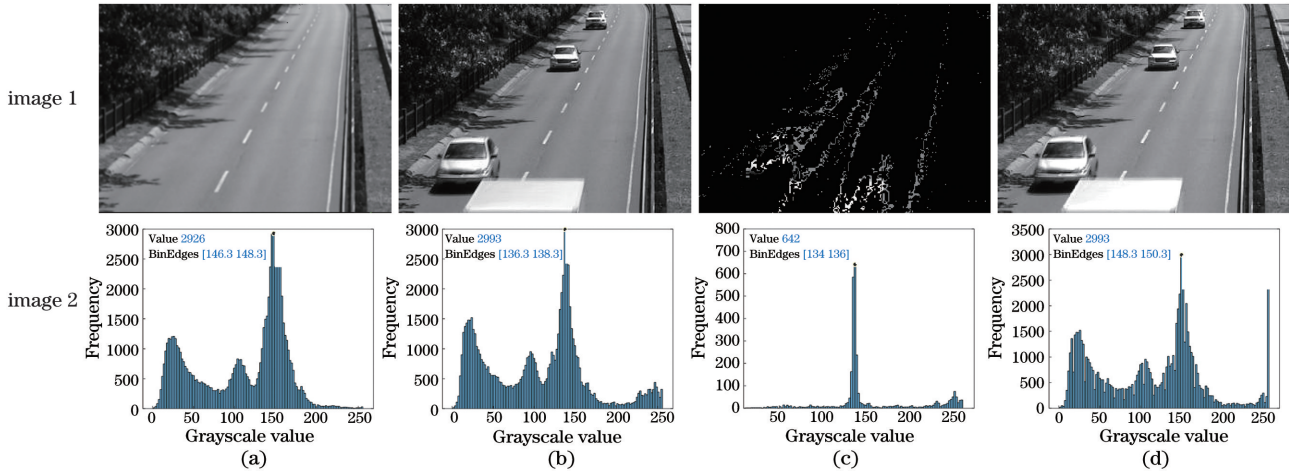


图 6 光照调整算法的过程及对应的灰度直图。(a)第 297 frame 的背景图像;(b)第 297 frame 的原始图像;(c)灰度值等于  $G_B$  的像素集合;(d)最终结果

Fig. 6 Process of light adjustment algorithm and corresponding grayscale straight maps. (a) Background image of 297<sup>th</sup> frame; (b) original image of 297<sup>th</sup> frame; (c) set of pixels whose gray value is equal to  $G_B$ ; (d) final results

结果。

从图 6(a) 可以看到, 第 297 frame 背景出现最频繁的灰度为 147; 图 6(c) 最高频次灰度则为 135。二者显然不匹配, 说明此时亮度发生了变化。经过调整后得到图 6(d), 此时最高频次灰度重新回到 147。说明光照调整算法可以避免由亮度变化导致计算结果存在误差。

因为前景灰度复杂, 使用直方图最大值对应灰度估计亮度变化可以有效消除误差, 避免由相机动态调整造成画面亮度发生变化。此外, 若  $|G_B - G_I| \leq 1$ , 说明当前光照没有发生变化, 否则说明环境光照发生较大改变, 此时应保存当前直方图并重新建模。若新模型连续  $N$  帧满足  $|G_B - G_I| \leq 1$  ( $N$  值大于上一段模型稳定帧数的一半), 便舍弃旧模型使用新模型, 否则仍然使用旧模型进行判断, 这可以解决长期光照缓慢变化的问题。在少数情况下, 旧模型调整亮度后, 仍存在大量前景, 此时则直

接舍弃旧模型。因为该模型可通过单帧建立, 故能快速从突变中恢复。

## 2.6 阴影消除

将彩色图像从色彩空间转换到另一个亮度与色度相分离的色彩空间进行阴影检测<sup>[19-21]</sup>, 但单色图像无法进行色度计算, 所以常常不对阴影进行处理或使用较大的阈值消除灰度变化较小的阴影区域。

文献[22]将图像视为照度分量与反射分量的乘积, 通过抑制照度分量来达到光照不变性。实际上, 阴影与光照变化类似, 都呈现为反射分量不变性, 局部区域照度分量按相同比值下降。将前景与背景图像相除, 则阴影区域会表现出整体平坦的特性, 此外比值大于 1 的部分均可视为真实前景, 而阴影部分的比值一定在  $0 \sim 1$  区间内, 将比值小于 1 的灰度乘 255 取整拉伸后并使用拉普拉斯算子计算梯度, 此时阴影区域的梯度几乎为 0, 对结果进行二值化处理<sup>[23]</sup>即可消除阴影区域, 求得剩余区域的凸包并与

前景相交即可得到最终结果。图 7 为前景阴影消除

法处理 PETS2006 第 83 frame 图像的各阶段结果。

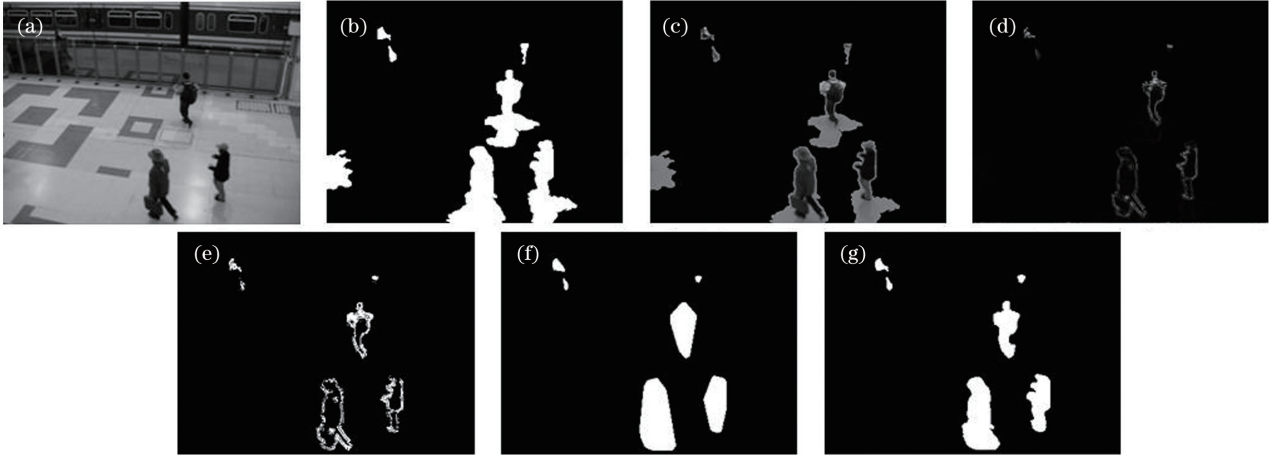


图 7 前景阴影消除法的各阶段结果。(a)原始图像;(b)前景掩模;(c)前景与背景比值;(d)比值梯度;(e)梯度图二值化;(f)点集凸包;(g)凸包与前景交集

Fig. 7 Results of each stage of foreground shadow elimination method. (a) Original image; (b) foreground mask; (c) foreground to background ratio; (d) ratio gradient; (e) gradient graph binarization; (f) convex hull of point set; (g) convex hull intersects foreground

从图 7(b)可以看到,所提算法的检测能力很好,可以完整检测出目标、阴影与倒影,而且其他区域均未出现虚警的现象。从图 7(c)可以看到,阴影区域的梯度显著低于前景,可以证明前景阴影消除法的有效性,从图 7(g)可以看到,阴影消除后的结果更优。

### 3 实验与结果分析

#### 3.1 CDnet2014 数据集的对比实验

为了检验所提建模算法的性能,使用 CDnet2014 数据集<sup>[12]</sup>中的 5 组场景可见光转灰度图像以及一组红外图像进行验证,并与文献[5]的 CodeBook、文献[7]的 ViBe、文献[8]的 KDE、文献[14]的 GMM、文献[24]的 PBAS 算法进行对比,对比算法均使用原文默认的参数。

实验平台的硬件为 i7-7700HQ 2.8 GHz CPU, 16 GB 内存,软件为 Win10、vs2017 和 OpenCV 4.0.1,所提算法在平台上处理分辨率为 320 pixel×240 pixel 的图像的平均时长为 13 ms,可以满足目标检测实时性的需求。

不同算法的主观对比结果如图 8 所示。图 8(a)挑战室内灯光的阴影干扰与长时间停留的前景;图 8(b)考验动静态背景混合场景下算法的适应能力;图 8(c)挑战低质量视频与户外阴影;图 8(d)挑战前景的长时间逗留和相机动态调整;图 8(e)挑战大面积动态背景下的前景;图 8(f)挑战低动态范围与低信噪比红外视频的目标检测。

从图 8 可以看到,在各类场景中,所提算法都能较好地保证目标的完整性;除了 CodeBook 和所提算法以外,其他算法均将图 8(a)和图 8(d)的静止目标当作背景;图 8(e)的前景较大且颜色均匀,KDE、PBAS 和 ViBe 算法错将船头当作背景,导致船尾检测不完整;图 8(a)和图 8(c)均存在较大阴影,所提算法在完整保留前景的同时也在一定程度上抑制阴影区域;图 8(b)和图 8(e)存在大量的动态背景区域,所提算法可以很好地控制虚警的出现;图 8(d)中人的出现导致相机出现多次曝光调整,当相机调整曝光时,CodeBook 算法和 PBAS 算法由于无法快速适应光照变化而出现较多的虚警现象;图 8(f)中前景与背景的灰度差异较小,其他算法处理后的前景均出现不完整的现象,而所提算法还是可以较好地保留真实前景,但出现较多的误检测点。

为了对不同算法进行定量分析,采用较为显著的 4 个指标进行评价,包含召回率( $R_e$ )、准确率( $P_r$ )、误检百分比( $P_{PWC}$ )和  $F$  数,计算公式分别为

$$R_e = x_{TP} / (x_{TP} + x_{FN}), \quad (18)$$

$$P_r = x_{TP} / (x_{FP} + x_{TP}), \quad (19)$$

$$P_{PWC} = (x_{FN} + x_{FP}) / (x_{TP} + x_{FN} + x_{FP} + x_{TN}) \times 100\%, \quad (20)$$

$$F = 2 \times P_r \times R_e / (P_r + R_e), \quad (21)$$

式中: $x_{TP}$  为正确检测为前景的点; $x_{TN}$  为正确检测为背景的点; $x_{FP}$  为错误检测为前景的点, $x_{FN}$  为错误检测为背景的点。对于图 8 的各评价指标结果如



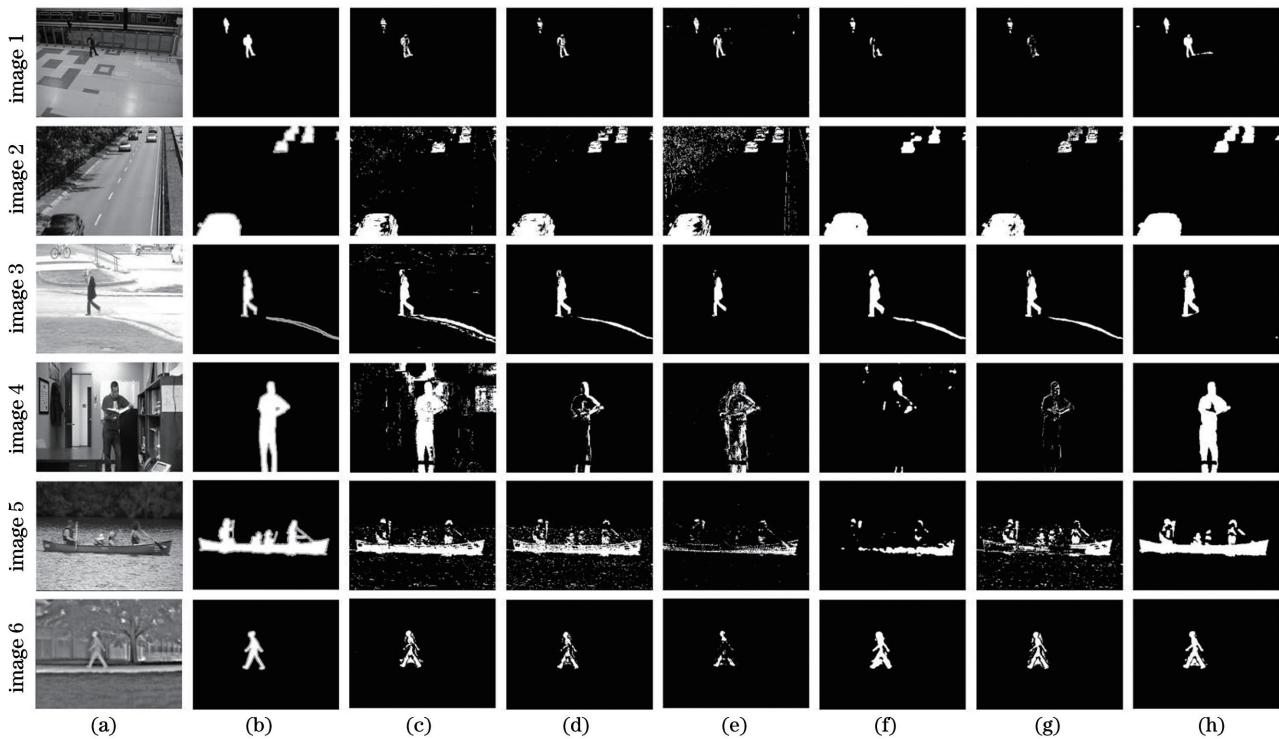


图 8 不同算法的处理结果。(a)原始图像;(b)真值图;(c)CodeBook 算法;(d)GMM 算法;(e)KDE 算法;(f)PBAS 算法;(g)ViBe 算法;(h)所提算法

Fig. 8 Processing results of different algorithms. (a) Original images; (b) Truth graphs; (c) CodeBook algorithm; (d) GMM algorithm; (e) KDE algorithm; (f) PBAS algorithm; (g) ViBe algorithm; (h) proposed algorithm

表 1 所示,其中一个下划线代表指标最好的算法,双下划线代表指标次好的算法。从表 1 可以看到,所提算法在绝大多数场景下都有极低的误检率、极高

的召回率与  $F$  数,准确率较好,其中  $F$  数最能反映算法的综合能力,可以看到所提算法的整体性能最优。

表 1 不同算法的评价指标结果

Table 1 Evaluation index results of different algorithms

Algorithm	Image 1				Image 2				Image 3			
	$R_e$	$P_r$	$P_{pwc}$	$F$	$R_e$	$P_r$	$P_{pwc}$	$F$	$R_e$	$P_r$	$P_{pwc}$	$F$
CodeBook	0.740	<b>0.709</b>	0.729	0.724	0.773	0.639	3.835	0.700	0.877	0.287	2.817	0.433
GMM	0.786	0.693	0.726	0.737	0.762	<b>0.722</b>	3.068	<b>0.741</b>	0.905	<b>0.548</b>	1.031	0.683
KDE	<b>0.870</b>	0.695	<b>0.661</b>	<b>0.773</b>	0.699	0.466	6.363	0.559	0.659	<b>0.804</b>	<b>0.615</b>	<b>0.724</b>
PBAS	0.651	0.667	0.872	0.659	<b>0.812</b>	<b>0.714</b>	<b>2.967</b>	0.760	<b>0.963</b>	0.500	1.223	0.658
ViBe	0.600	0.662	0.914	0.630	0.747	0.712	3.204	0.729	0.925	0.539	1.062	0.681
Proposed algorithm	<b>0.968</b>	<b>0.698</b>	<b>0.583</b>	<b>0.811</b>	<b>0.989</b>	0.697	<b>2.549</b>	<b>0.817</b>	<b>0.960</b>	0.538	<b>1.057</b>	<b>0.690</b>

Algorithm	Image 4				Image 5				Image 6			
	$R_e$	$P_r$	$P_{pwc}$	$F$	$R_e$	$P_r$	$P_{pwc}$	$F$	$R_e$	$P_r$	$P_{pwc}$	$F$
CodeBook	<b>0.932</b>	0.412	9.586	<b>0.571</b>	<b>0.728</b>	0.652	<b>3.637</b>	<b>0.688</b>	0.519	0.701	1.639	0.597
GMM	0.445	<b>0.637</b>	<b>5.537</b>	0.524	0.695	0.584	4.400	0.635	0.507	0.725	1.600	0.597
KDE	0.414	0.585	6.028	0.485	0.294	0.644	4.776	0.403	0.245	<b>0.777</b>	1.928	0.372
PBAS	0.336	0.513	6.734	0.406	0.312	<b>0.866</b>	4.047	0.459	<b>0.641</b>	0.672	<b>1.571</b>	<b>0.656</b>
ViBe	0.162	0.613	6.444	0.256	0.441	0.555	5.022	0.491	0.605	0.683	1.578	0.642
Proposed algorithm	<b>0.928</b>	<b>0.805</b>	<b>2.037</b>	<b>0.862</b>	<b>0.858</b>	<b>0.775</b>	<b>2.150</b>	<b>0.815</b>	<b>0.671</b>	<b>0.729</b>	<b>1.353</b>	<b>0.699</b>

### 3.2 中波红外外面阵探测器数据的对比实验

实验使用的设备为中波红外外面阵探测器,采样频率为 100 frame/s,使用该设备对所提算法的可用性进行验证。实验场景包含一组室内场景和两组室外海面场景,实验参数与实验环境与 3.1 节相同。由于无法对所有数据进行标注,本节仅展示不同算法主观处理的结果。图 9(a)为室内走廊行人,图 9(b)和图 9(c)为海面游艇。从图 9 可以看到,随

着更新速率的提高,KDE、PBAS 和 ViBe 算法均将前景当作背景,而且存在严重的前景空洞;CodeBook 算法和 GMM 算法出现明显的鬼影现象,并因设定固定的阈值导致无法区分低动态范围下的前景与背景,使前景出现部分缺失;所提算法在不改变参数的情况下仍然保持良好的检测性能,但海面场景下的尾迹也被错误检测为前景,后续需要针对性地优化算法。

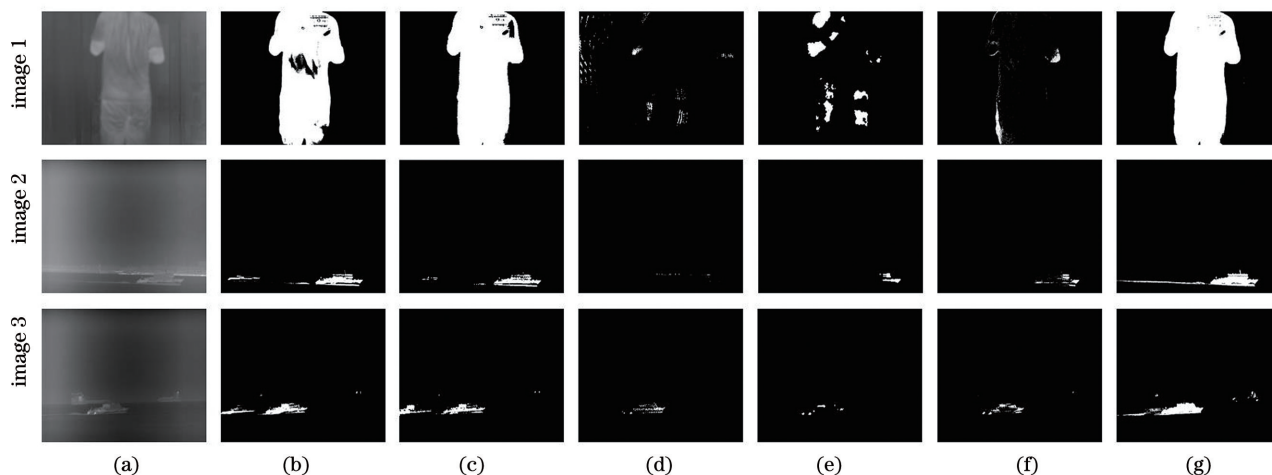


图 9 不同算法对中波红外图像的处理结果。(a)原始图像;(b)真值图;(c)CodeBook 算法;(d)GMM 算法;(e)KDE 算法;(f)PBAS 算法;(g)ViBe 算法;(h)所提算法

Fig. 9 Processing results of mid-wave infrared image by different algorithms. (a) Original images; (b) Truth graphs; (c) CodeBook algorithm; (d) GMM algorithm; (e) KDE algorithm; (f) PBAS algorithm; (g) ViBe algorithm; (h) proposed algorithm

## 4 结 论

针对灰度视频比彩色视频的动态范围小以及可用信息少的问题,提出一种基于直方图统计的众数建模算法。首先通过改进的三帧差分法计算噪声分布,接着利用获得的噪声分布将图像划分为动、静态区域并采取不同的策略提取目标,然后利用差分图像直方图检测全局光照的短时波动和长时间变动,最后通过前景和背景图像的比值梯度实现灰度图像的阴影抑制。实验结果表明,所提算法可以极大地提高灰度视频下目标检测的召回率,能够有效检测不同速度的目标,而且无需人为设定先验知识,适应性强;在多种不同场景下的综合指标均优于其他算法,可有效提高黑白监控和红外监控等应用场景的目标检测效率。但所提算法对阴影抑制的效果较差,部分场景的准确率也有待进一步加强。接下来的研究工作是进一步优化算法,提升算法的检测准确率,并将程序改写为并行算法,提高算法的运行速度。

## 参 考 文 献

- [1] Jain N K, Saini R K, Mittal P. A review on traffic monitoring system techniques[M]//Ray K, Sharma T, Rawat S, et al. Soft computing: theories and applications. Advances in intelligent systems and computing. Singapore: Springer, 2019, 742: 569-577.
- [2] Garcia-Garcia B, Bouwmans T, Rosales Silva A J. Background subtraction in real applications: challenges, current models and future directions[J]. Computer Science Review, 2020, 35: 100204.
- [3] Wren C R, Azarbayejani A, Darrell T, et al. Pfunder: real-time tracking of the human body[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.
- [4] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [C] // Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 23-25, 1999, Fort Collins, CO, USA. New York: IEEE, 1999: 246-252.

- [5] Wang Z D, Xie E, Zhu B. A background segmentation algorithm based on improved codebook model [J]. *Science Technology and Engineering*, 2017, 17(32): 148-153.  
汪兆栋, 谢恩, 朱斌. 一种基于改进码本模型的背景分割算法[J]. *科学技术与工程*, 2017, 17(32): 148-153.
- [6] 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.
- [7] 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.  
闵卫东, 郭晓光, 韩清. 改进的 ViBe 算法及其在交通视频处理中的应用[J]. *光学精密工程*, 2017, 25(3): 806-811.
- [8] Elgammal A, Duraiswami R, Harwood D, et al. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance [J]. *Proceedings of the IEEE*, 2002, 90(7): 1151-1163.
- [9] Wang H Z, Suter D. Background subtraction based on a robust consensus method [C] // 18th International Conference on Pattern Recognition (ICPR'06), August 20-24, 2006, Hong Kong, China. New York: IEEE, 2006: 223-226.
- [10] Maity S, Chakrabarti A, Bhattacharjee D. Block-based quantized histogram (BBQH) for efficient background modeling and foreground extraction in video [C] // 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), February 24-26, 2017, Pune, India. New York: IEEE, 2017: 224-229.
- [11] Lai A N, Yoon H, Lee G. Robust background extraction scheme using histogram-wise for real-time tracking in urban traffic video [C] // 2008 8th IEEE International Conference on Computer and Information Technology, July 8-11, 2008, Sydney, NSW, Australia. New York: IEEE, 2008: 845-850.
- [12] Wang Y, Jodoin P M, Porikli F, et al. CDnet 2014: an expanded change detection benchmark dataset [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 393-400.
- [13] Guo Z C, Dang J W, Wang Y P, et al. Multi-feature background modeling algorithm based on improved census transform [J]. *Acta Optica Sinica*, 2019, 39(8): 0815003.  
郭治成, 党建武, 王阳萍, 等. 基于改进 Census 变换的多特征背景建模算法[J]. *光学学报*, 2019, 39(8): 0815003.
- [14] Liu W, Hao X L, Lü J L. Efficient moving targets detection based on adaptive Gaussian mixture modelling [J]. *Journal of Image and Graphics*, 2020, 25(1): 113-125.  
刘伟, 郝晓丽, 吕进来. 自适应混合高斯建模的高效运动目标检测[J]. *中国图象图形学报*, 2020, 25(1): 113-125.
- [15] Yu X M, Li S Y, Shi S N. An improved algorithm for moving target detection using a Gaussian mixture with three-frame difference [J]. *Infrared Technology*, 2019, 41(3): 256-261.  
于晓明, 李思颖, 史胜楠. 混合高斯融合三帧差的运动目标检测改进算法[J]. *红外技术*, 2019, 41(3): 256-261.
- [16] Ding J, Kuang L Q, Han X. Improved moving object detection method based on ViBe [J]. *Computer Engineering and Design*, 2017, 38(2): 374-378.  
丁洁, 况立群, 韩燮. 基于 ViBe 的运动目标检测改进方法[J]. *计算机工程与设计*, 2017, 38(2): 374-378.
- [17] Qu Z, Chai G H, Liu Y. Improved algorithm of visual background extraction for illumination robustness [J]. *Computer Engineering and Design*, 2019, 40(4): 1046-1051, 1189.  
瞿中, 柴国华, 刘妍. 改进光照鲁棒的视觉背景提取算法[J]. *计算机工程与设计*, 2019, 40(4): 1046-1051, 1189.
- [18] Xu Y P, Ji H B, Zhang W B. Coarse-to-fine sample-based background subtraction for moving object detection [J]. *Optik*, 2020, 207: 164195.
- [19] Horprasert T, Harwood D, Davis L S. A statistical approach for real-time robust background subtraction and shadow detection [EB/OL]. [2020-06-28]. [https://www.researchgate.net/publication/228609042\\_A\\_statistical\\_approach\\_for\\_real-time\\_robust\\_background\\_subtraction\\_and\\_shadow\\_detection](https://www.researchgate.net/publication/228609042_A_statistical_approach_for_real-time_robust_background_subtraction_and_shadow_detection).
- [20] Xie S R, Ye S B, Yang B H, et al. Moving target detection based on improved YUV\_ViBe fusion algorithm [J]. *Laser & Optoelectronics Progress*, 2018, 55(11): 111002.  
谢申汝, 叶生波, 杨宝华, 等. 基于改进的 YUV\_ViBe 融合算法的运动目标检测[J]. *激光与光电子学进展*, 2018, 55(11): 111002.
- [21] Fang L, Yu F Q. Moving object detection algorithm based on removed ghost and shadow visual background extractor [J]. *Laser & Optoelectronics*

- Progress, 2019, 56(13): 131002.
- 方岚, 于凤芹. 去除鬼影及阴影的视觉背景提取运动目标检测算法[J]. 激光与光电子学进展, 2019, 56(13): 131002.
- [22] Lou J G, Yang H, Hu W M, et al. An illumination invariant change detection algorithm [C] // The 5th Asian Conference on Computer Vision, January 23-25, 2002, Melbourne, Australia. New York: IEEE, 2002.
- [23] Otsu N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66.
- [24] Hofmann M, Tiefenbacher P, Rigoll G. Background segmentation with feedback: the pixel-based adaptive segmenter [C] // 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 38-43.

## Statistical-Based Adaptive Background Modeling Algorithm for Grayscale Video

Wu Jiawen<sup>1,2</sup>, Wang Shiyong<sup>1\*</sup>

<sup>1</sup> Key Laboratory of Infrared System Detection and Imaging Technology, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China

### Abstract

**Objective** Target detection is an active research field of computer vision and forms the basis of applications such as video surveillance and live streaming. Numerous algorithms are used in target detection including traditional methods such as optical flow, frame difference, background modeling, and deep learning, the latter of which has developed rapidly in recent years. However, with the exception of background modeling, most of these algorithms are limited by factors such as budget constraints as well as low processor performance, real-time capability, and accuracy. Thus, background modeling is widely used in real applications. This method also has limitations, however. For example, background modeling will lead to poor video quality and texture and a lot of noise when processing grayscale images such as infrared video. In the present study, we address these limitations by proposing a novel statistical-based adaptive background modeling algorithm that shows high accuracy and an improved recall rate.

**Methods** We build a grayscale histogram for each pixel in an 8-bit grayscale video and update it on each frame. Then, we estimate the true background from the histogram mode and obtain the noise threshold by using an improved three-frame difference method. From the results, we divide each frame into two regions: a static region and a dynamic region. For the static region, we simply apply single Gaussian modeling to the obtained noise threshold to determine the foreground. For the dynamic region, we utilize an algorithm based on kernel density estimation (KDE) for the detection. Moreover, we resolve the illumination shift and shadow by calculating the grayscale difference between the background and the input frame.

**Results and Discussions** In the proposed algorithm, a period of 13 ms is required to handle one frame with 320 pixel × 240 pixel on an i7-7700HQ platform, which fits the real-time requirement. To prove the high performance of the proposed algorithm, we compare the results of other traditional background modeling algorithms including CodeBook (CB), ViBe, KDE, the Gaussian mixture model (GMM), and the pixel-based adaptive segmenter (PBAS) on public test dataset CDnet2014 (**Fig. 8**) with the infrared dataset obtained in the present study (**Fig. 9**). These datasets contain various scenes including mall, highway, avenue, office, river, infrared park, and infrared sea. Target detection under these scenes faces several common challenges such as shadow, static target, complex background, dynamic background, and low-quality images. The results (**Table 1**) show some of the advantages of the proposed algorithm. In particular, (1) the new algorithm shows the highest recall rate, at 90% in most scenes, which efficiently ensures the integrity of the target. (2) In the mall and office scenes, the previous algorithms except for CB absorb the static foreground into the background. (3) In the river scene, a slow, large, and homogeneous color foreground caused by a canoe led KDE, PBAS, and ViBe to incorrectly consider the front part of the canoe as the background, thus causing incomplete detection of the back part. (4) In the mall and avenue scenes, abundant shadows are present that interfere with the detection result. The proposed algorithm not only

preserves the intact foreground but also eliminates the shadow to some extent. (5) Numerous dynamic background areas are present in the highway and river scenes that result in huge amounts false positives. The proposed algorithm effectively decreases the false alarm rate. (6) In the office scene, foreground motion causes the camera exposure to adjust several few times. At the moment of exposure adjustment, CB and PBAS are unable to rapidly adapt to the new illumination, which causes false positives that last for long periods. (7) The images in the infrared park and infrared sea scenes are low in contrast, resulting in extremely poor image quality. Only the proposed algorithm can effectively distinguish the foreground from background. The maximum recall rate of the previous method is 0.671, and the false positives are abundant; that of the proposed algorithm is 0.900, and the false alarm rate is significantly decreased. All these results show that proposed algorithm can significantly increase the detection recall rate and integrity of the target with a low false alarm rate and considerably rapid processing speed.

**Conclusions** According to the experiment results, the proposed algorithm adopts various detection strategies for dynamic and static regions and combines the advantages of various algorithms to significantly improve the recall rate of target detection in grayscale video. Moreover, it can also detect targets with complex motion pattern with no prior knowledge, which makes this method highly robust in many applications. This robustness results in an overall better index compared with that in common background modeling algorithms. Further, the proposed algorithm can effectively enhance the target detection performance in monochromatic and infrared video. However, its shadow elimination is inadequate, and the accuracy of some scenes requires further optimization.

**Key words** machine vision; motion target detection; background modeling; adaptive threshold; illumination variation

**OCIS codes** 150.0155; 330.4150; 100.2000