

## 时空相关多通道聚类的运动目标检测

徐艳 王培光 杨青 董江涛

### Moving target detection algorithm based on spatiotemporal correlation multi-channel clustering

Xu Yan Wang Pei-Guang Yang Qing Dong Jiang-Tao

引用信息 Citation: *Acta Physica Sinica*, 68, 164203 (2019) DOI: 10.7498/aps.68.20190161

在线阅读 View online: <https://doi.org/10.7498/aps.68.20190161>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

动态背景下基于光流场分析的运动目标检测算法

Moving object detection based on optical flow field analysis in dynamic scenes

物理学报. 2017, 66(8): 084203 <https://doi.org/10.7498/aps.66.084203>

基于两正交互耦1550 nm垂直腔面发射激光器获取多路随机数

Multi-channel physical random number generation based on two orthogonally mutually coupled 1550 nm vertical-cavity surface-emitting lasers

物理学报. 2018, 67(2): 024204 <https://doi.org/10.7498/aps.67.20171902>

基于随机聚类的复杂背景建模与前景检测算法

Complex background model and foreground detection based on random aggregation

物理学报. 2015, 64(15): 150701 <https://doi.org/10.7498/aps.64.150701>

基于偏振旋转耦合1550 nm垂直腔面发射激光器环形系统产生多路高质量混沌信号

Generations of multi-channel high-quality chaotic signals based on a ring system composed of polarization rotated coupled 1550 nm vertical-cavity surface-emitting lasers

物理学报. 2016, 65(19): 194207 <https://doi.org/10.7498/aps.65.194207>

基于相关光子多模式相关性的InSb模拟探测器定标方法

Absolute calibration of an analog InSb detector based on multimode spatial correlation of correlated photons

物理学报. 2015, 64(24): 240601 <https://doi.org/10.7498/aps.64.240601>

通过放慢钠通道开闭控制心脏中的螺旋波和时空混沌

Controlling spiral wave and spatiotemporal chaos in cardiac tissues by slowing sodium channel activation and inactivation

物理学报. 2016, 65(19): 198201 <https://doi.org/10.7498/aps.65.198201>

# 时空相关多通道聚类的运动目标检测\*

徐艳<sup>1)2)</sup> 王培光<sup>1)†</sup> 杨青<sup>2)3)</sup> 董江涛<sup>4)</sup>

1) (河北大学电子信息工程学院, 保定 071002)

2) (陆军工程大学石家庄校区电子与光学工程系, 石家庄 050000)

3) (河北工业大学电子信息工程学院, 天津 300401)

4) (中国电子科技集团公司第五十四研究所, 石家庄 050000)

(2019年1月28日收到; 2019年5月20日收到修改稿)

针对某些光照变化、噪声不稳定等多模态场景不适合离线训练背景模型来提取目标信息的问题, 在基于混合高斯的背景建模的基础上, 利用帧间差分与邻域相似性实现模型初始参数的选取; 提出将随机子采样与邻域空间传播理论相结合改进参数更新过程; 在时间维度上建立观测向量, 实现模型参数的优化, 加快模型收敛速度; 并将颜色信息和梯度相融合实现基于多特征的多通道背景模型的建立, 采用背景点的随机采样策略简化多通道模型建立的计算量, 最终实现复杂环境下的运动目标的检测. 实验表明, 算法在抑制鬼影、动态背景和遮挡等方面有良好的检测性能, 且执行效率能够满足实时计算的需求.

**关键词:** 混合高斯, 随机子采样, 邻域相关, 多通道

**PACS:** 42.30.Tz, 42.30.Va, 42.66.Lc, 47.50.Cd

**DOI:** 10.7498/aps.68.20190161

## 1 引言

运动目标检测和跟踪受到场景变换、遮挡和噪声等的影响, 检测和跟踪精度一直是计算机视觉领域研究的热点. 在 CVPR2015-2018 中, 大部分目标跟踪算法都是基于检测的跟踪, 且目标跟踪算法是基于大量离线数据的训练下得到的, 这就限制了此类算法在某些不便于采集数据或者不便于离线训练的特殊场景的应用, 再加之当前视频环境中背景多是多模态场景噪声多, 且运动目标表现出不同特征, 这就需要有一个稳定、有较强鲁棒性的运动目标检测算法进行处理.

目前主流的运动目标检测算法有光流法<sup>[1-3]</sup>、帧间差分法、背景减除法<sup>[4-6]</sup>. 光流法是通过分析运动场的变化对目标进行检测, 这种方法不需要预

先知道场景的任何信息, 但噪声、多光源等因素会对光流场分布的计算结果造成严重影响; 帧间差分法原理简单, 易于实现, 针对简单背景可以取得较好的效果, 但对环境噪声较敏感; 背景减除法是利用视频序列对背景建模, 根据背景图像与当前帧图像的差异获得运动目标区域, 具有检测准确且精度高的优点, 但其性能很大程度上依赖于背景建模的准确性. 实际复杂场景中无法预测和不确定因素的存在、动态背景、光照<sup>[7]</sup>的变化等, 给背景的准确建模带来很大挑战.

学者们对于背景建模方法进行了很多研究<sup>[5-9]</sup>, 由于高斯函数在样本数据点足够大时, 可以表征任何一种事物的规律, 文献<sup>[4]</sup>提出的 GMM 算法是使用混合高斯分布模型表征图像帧中每一个像素点的特征, 通过在线的学习与训练得到背景模型并完成运动目标的检测. 文献<sup>[10]</sup>提出了采用单像素

\* 国家自然科学基金 (批准号: 11771115) 资助的课题.

† 通信作者. E-mail: [pgwang@hbu.edu.cn](mailto:pgwang@hbu.edu.cn)

时间序列直方图进行背景建模,得到的前景目标区域更干净,误差像素较少,但该方法存储开销大,背景生成缓慢;针对学习速率 $\alpha$ 无法适应运动目标的速度时,参考背景图像中有可能出现目标的鬼影,文献[11]提出了基于三帧差分的学习率自适应混合高斯模型背景更新的方法,减少了光照突变和目标快速运动带来的误检情况;文献[12]引入引导滤波、空间邻域更新及时空平滑方法,较好地抑制了背景扰动引起的虚假运动目标. ViBe算法[13]则采用随机聚类技术,采用随机策略更新样本点,避免了大量的样本建模.

从上述分析可知,目前的背景建模都是通过不同聚类方法完成背景聚类,进而完成对运动目标的检测,但都存在以下不足:

- 1) 当初始帧中存在运动目标时,检测结果可能出现“ghost”现象;
- 2) 逐像素运算,忽略了像素点间的相关性,运算效率低下;
- 3) 在样本或模型更新过程中阈值或参数固定不变,导致不同场景中的目标检测精准度差距较大.

本文在混合高斯建模的基础上,采用帧间差分法获取模型初始参数,减少模型建立初始时间的同时使背景模型的初始参数更加贴近真实值,加快模型的收敛速度;将文献[13]的邻域传播理论引入到混合高斯模型的更新过程,提出邻域相关度的近似计算方法,在匹配过程中,将像素级的模型建立与邻域随机相关相结合,在多模态背景模型建立过程中又兼顾执行效率;引入像素点的活跃级别对学习率进行调整,加快整体模型的收敛;最后,采用颜色-梯度相结合的方式实现多通道的混合高斯模型的建立及运动目标的检测.

## 2 基于混合高斯模型 GMM 的目标检测

图像序列中同一位置的像素点序列可以用 $K$ 个高斯分布综合描述,当前像素值的概率密度由下式计算得到:

$$p(x_t) = \sum_{k=1}^K \varpi_{k,t} \times \eta\left(x_t, \mu_{k,t}, \sum_{k,t}\right), \quad (1)$$

其中 $K$ 是高斯分布的数量, $\varpi_{k,t}$ 是 $t$ 时刻第 $k$ 个高

斯分布的权重, $\mu_{k,t}$ 和 $\sum_{k,t}$ 分别是 $t$ 时刻第 $k$ 个高斯分布的期望和协方差矩阵,该点高斯分布为

$$\eta\left(x_t, \mu_{k,t}, \sum_{k,t}\right) = \frac{1}{(2\pi)^{\frac{n}{2}} \left|\sum_{k,t}\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t - \mu_{k,t})^T \sum_{k,t}^{-1} (x_t - \mu_{k,t})}, \quad (2)$$

其中协方差矩阵为 $\sum_{k,t} = \sigma_{k,t}^2 I$ .

### 1) 高斯模型参数的更新

在背景更新过程中,将当前帧中像素点与所建立的 $K$ 个高斯分布按照下式进行匹配:

$$|x_t - \mu_{k,t-1}| < D \times \sigma_{k,t-1}, \quad (3)$$

式中, $\sigma_{k,t-1}$ 为第 $k$ 个高斯模型在 $t-1$ 时刻的标准差, $D$ 为置信参数.如果像素点匹配成功,即当前像素点满足第 $k$ 个高斯模型的分布,则按照下式对匹配的高斯分布进行参数更新,且 $M_{k,t-1} = 1$ :

$$\begin{aligned} \varpi_{k,t} &= (1 - \alpha) \times \varpi_{k,t-1} + \alpha M_{k,t-1}, \\ \mu_{k,t} &= (1 - \beta) \times \mu_{k,t-1} + \beta \times x_t, \\ \sigma_{k,t}^2 &= (1 - \beta) \times \sigma_{k,t-1}^2 + (x_t - \mu_{k,t-1})^2, \end{aligned} \quad (4)$$

其中 $\varpi_{k,t}$ 为第 $k$ 个高斯函数在 $t$ 时刻的归一化权重, $\sum_{k=1}^K \varpi_{k,t} = 1$ ;  $\alpha$ 是参数的学习率, $\beta = \alpha \times \eta(x_t, \mu_{k,t}, \sum_{k,t})$ ,  $\eta(x_t, \mu_{k,t}, \sum_{k,t})$ 为第 $k$ 个高斯函数的概率密度.

### 2) 生成背景分布

像素点的高斯混合模型参数更新完毕后,需要对 $K$ 个高斯分布进行估计,得到能够代表背景模型的高斯分布[5].将 $K$ 个高斯分布重新进行权重归一化处理,按照 $\varpi_{k,t}/\sigma_{k,t}$ 从大到小的顺序排序,将前 $B$ 个高斯分布作为背景模型:

$$B = \arg \min_b \left( \sum_{i=1}^b \varpi_{i,t} > T \right). \quad (5)$$

### 3) 检测运动目标

如果当前像素点与任意一个高斯分布匹配成功,则该点为背景点,否则判断为运动前景.

## 3 改进的 GMM 背景模型建立方法

### 3.1 基于五帧差分法的空间邻域模型参数初始化

由于没有先验知识,高斯模型参数通常是初始化为固定值,在随后的更新过程中逐步调整,如果

模型参数初值选取不当将导致模型的更新较慢.

本文采用五帧差分与帧内邻域平均的方法, 实现混合高斯模型的参数初始化选取. 首先提取连续五帧图像, 利用中间帧对前后两帧进行差分运算, 对差分结果进行二值化处理后再进行与运算, 用以检测初始视频中是否含有目标信息, 随后计算除去运动区域后背景图像中的五帧均值  $\mu(x, y)$ , 作为像素点  $(x, y)$  的初始均值, 同时计一个累加矩阵 CA, CA 的大小与图像大小一致, CA  $(x, y)$  记录满足规则的邻域像素点的个数, 累加器各元素的更新方式为

$$\begin{aligned} CA(x, y) &= CA(x, y) + 1, \\ |f(x, y) - \mu(x, y)| &< \alpha\lambda. \end{aligned} \quad (6)$$

同时, 还要对均值矩阵  $\mu(x, y)$  进行更新, 即

$$\begin{aligned} \mu(x, y) &= (f(x, y) + \mu(x, y))/2, \\ |f(x, y) - \mu(x, y)| &< \alpha\lambda, \end{aligned} \quad (7)$$

计算相应的标准差  $\sigma$ :

$$\sigma = \sqrt{\frac{1}{5} \sum_{t=1}^5 (f_t(x, y) - \mu(x, y))^2}$$

并以此构建混合高斯模型中的其中一个高斯分布:

$$\begin{cases} \mu_1(x, y) = \mu(x, y), \\ \sigma_1(x, y) = \sigma(x, y), \\ \varpi_1(x, y) = \varpi_M. \end{cases} \quad (8)$$

而其他高斯模型的参数按照下式获取:

$$\begin{cases} \mu_k(x, y) = f(x, y), \\ \sigma_k(x, y) = \sigma_M, \\ \varpi_k(x, y) = \varepsilon. \end{cases} \quad k = 2, \dots, K. \quad (9)$$

由于初始化阶段考虑了帧间数据的关联性, 去除了运动目标在初始阶段对建模的影响, 更加不易产生鬼影现象; 将邻域数据的一致性加入到模型参数的初始化环节, 使得构建的混合高斯模型更加符合背景像素点的特征.

### 3.2 随机子采样与邻域空间传播理论相结合的模型参数更新策略

原 GMM 在建模过程中像素点间模型信息相互独立, 具体体现在:

1) 对像素点孤立单独建模, 忽略了空间邻域信息的相关性, 对背景像素的整体描述不够, 因此受到噪声或背景变化时效果下降;

2) 对当前时刻  $t$  的每一个像素点都进行模型参数训练, 导致计算量很大.

图像空间信息的连续性, 像素点间既有互异性又有相似性<sup>[13]</sup>, 利用互异性建模能保持像素点的变化规律, 利用邻域相似性可以减少像素间的冗余信息<sup>[14]</sup>, 避免相似点的重复建模. 由以上两点, 本文提出利用相邻像素相关度来衡量邻域范围内像素的相关性, 当像素相关性在某阈值范围内时, 认为邻域像素与当前像素属于同一类别 (同属于背景或运动前景), 对于同一类别的邻域则不进行高斯分布的参数更新计算, 由于背景的背景像素相关性很大, 因此这种基于邻域像素相关性度量的参数更新可以在很大程度上缩减计算量.

为了得到任意时刻的均值和方差, 本文引入文献<sup>[15]</sup>在均值和方差的在线估计中用到的滑动窗口的概念, 对任意  $t$  时刻计算邻域相关度, 需要用到前  $t - N + 1$  帧图像.  $t$  时刻滑动窗口  $N$  内的图像序列如图 1 所示.

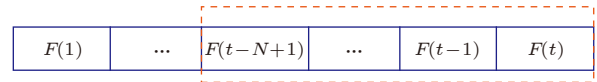


图 1  $t$  时刻滑动窗口  $N$  内的图像序列

Fig. 1. The image sequence in the sliding window  $N$  at time  $t$ .

设  $x_i, y_i$  分别为相邻的两个像素点的灰度值,  $D(x) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2$  为  $x_i$  的方差,  $\text{cov}(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))(y_i - E(y))$  为两点的协方差,  $E(x) = \frac{1}{N} \sum_{i=1}^N x_i$  为  $x_i$  的均值,  $N$  为参与邻域相关计算的图像序列的帧数.

为简化计算和节省存储空间, 本文对  $E(x)$  和  $D(x)$  进行了简化, 将其用该像素点混合高斯模型的均值和方差的加权平均代替, 从而得到

$$\hat{E}(x_t) = \sum_{k=1}^K \omega_{k,t} \mu_{k,t-1}, \quad \hat{D}(x_t) = \sum_{k=1}^K \omega_{k,t} \sigma_{k,t-1}, \quad (10)$$

其中  $K$  为高斯分布的数量,  $\sum_{k=1}^K \omega_{k,t} = 1$ , 进而得到

$$\text{cov} \hat{v}(x_t, y_t) = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{E}(x_t)) (y_i - \hat{E}(y_t)), \quad (11)$$

(一般滑动窗口  $N = 50$ ), 最终邻域像素相关性

度量:

$$\hat{\gamma}_{xy} = \frac{\text{co} \hat{v}(x_t, y_t)}{\sqrt{\hat{D}(x_t)} \sqrt{\hat{D}(y_t)}}. \quad (12)$$

在进行参数更新前, 首先计算在当前像素点 8 邻域范围内相对当前像素点的相关度, 如果像素相关度大于某一阈值  $T_\gamma$ , 则邻域点与当前像素点有很大概率属于同一类别, 那么在进行模型更新时邻域值可以直接沿用当前像素的模型, 从而不进行模型更新计算; 而对于不属于同一类别的像素点则仍然采用模型更新策略.

文献 [16] 指出, 当不能准确地判断像素变化的模型时, 采用随机更新策略, 能够在一定程度上模拟像素变化的不确定性. 由于图像中背景的更新缓慢, 即使是动态背景, 在相邻帧之间背景也不会有太大变化, 因此对前面帧中已确定为背景的像素点没有必要逐帧更新, 因此论文采用随机子采样策略, 对上一帧已经确定为背景的像素点随机选取  $1/\varphi$  ( $\varphi = 16$ ) 进行模型更新计算, 更新点邻域范围内的相似像素进行模型拷贝, 其他背景点则维持原有高斯分布不变.

具体的更新过程如图 2 所示.

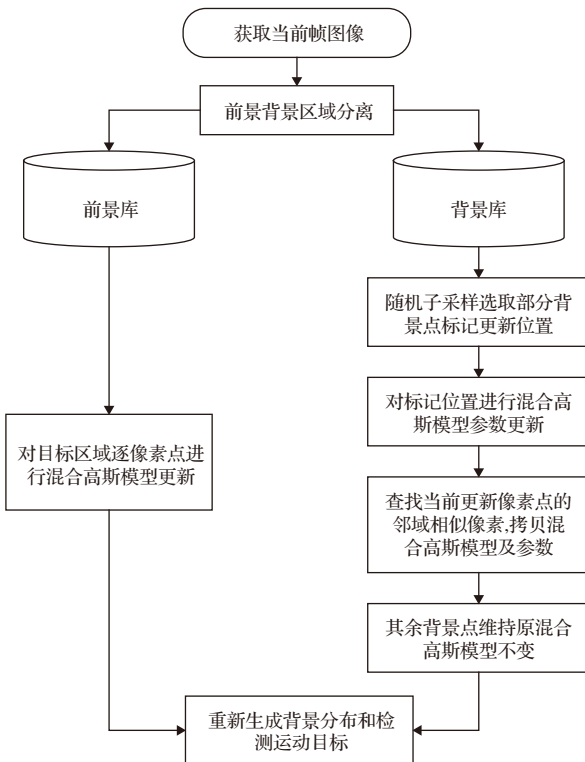


图 2 模型参数更新流程

Fig. 2. Model parameter update process.

### 3.3 基于观测向量的参数更新策略

在原 GMM 算法中, 匹配的高斯分布按照 (4) 式进行更新, 而其他没有匹配的则只更新权值, 且权值的更新速度较慢, 不妨假设第  $k$  个高斯分布在前  $n$  个样本点都没有得到匹配, 则在  $t$  时刻其权值更新为  $\varpi_{k,t} = (1 - \alpha)^n \times \varpi_{k,0}$ , 通常  $\alpha$  取值较小, 这就使得  $\varpi_{k,t}$  的减小速度比较慢, 而实际情况是这一分布应该是不符合背景模型的, 我们希望这样的分布能有更快的收敛速度.

为此, 在每一个像素位置引入一个含有  $K$  个元素的计数向量, 该向量元素  $c_k$  负责记录第  $k$  个高斯分布的观测值:

首先计算像素与第  $k$  个高斯分布匹配的后验概率:

$$q_k(x_t) = p_k(x_t) / \sum_{j=1}^K p_j(x_t), \quad (13)$$

令  $c_k(x_t) = c_k(x_{t-1}) + q_k(x_t)$ ,

$\hat{\eta}(x_t) = q_k(x_t) \left( \frac{1 - \alpha}{c_k(x_t)} + \alpha \right)$ , 则得到新的均值和方差的更新公式为:

$$\mu_{k,t} = (1 - \hat{\eta}(x_t)) \times \mu_{k,t-1} + \hat{\eta}(x_t) \times x_t,$$

$$\sigma_{k,t}^2 = (1 - \hat{\eta}(x_t)) \times \sigma_{k,t-1}^2 + \hat{\eta}(x_t) \times (x_t - \mu_{k,t-1})^2, \quad (14)$$

$$\varpi_{k,t} = (1 - \alpha) \times \varpi_{k,t-1} + \alpha q_k(x_t). \quad (15)$$

如果像素点没有匹配成功,  $\sum_{j=1}^K p_j(x_t) = 0$ , 则所有高斯分量都进行退化.

同时, 加入像素点的活跃级别来限制  $\alpha$ :

$$\alpha = \begin{cases} \frac{\alpha}{5}, & A_{k,t} > T_A \\ \alpha, & A_{k,t} \leq T_A, \end{cases} \quad (16)$$

其中  $A_{k,t} = (1 - \lambda) A_{k,t-1} + \lambda |x_t - x_{t-1}|$ ,  $\lambda = 0.09$ ,  $A_{k,t}$  的初值设为 0, 当像素点较为活跃时, 代表该像素点的值变化较大, 这种点成为背景的可能性就较小, 此时减小  $\alpha$  以增大  $\varpi_{k,t}$  的衰减速度, 从而加快整体模型的收敛.

## 4 基于颜色-梯度的运动目标检测

通过上述混合高斯模型的建立, 实现了背景和运动目标的聚类, 但是单纯基于灰度一个通道的建模方法不能充分利用图像中的像素信息, 因此对于

运动目标的检测结果含有较多空洞区域.

通过分析可知, 目标通常是由若干点组成的非连续目标, 相对背景而言, 运动目标具有如下特点:

1) 即使是相对颜色一致的运动物体也会比背景产生更多变化, 况且一般情况下物体都是带有不同颜色的;

2) 目标在局部仍然具有特征连续性的特点, 且运动目标与周围背景应有比较明显的差别.

文献 [17] 提出可以采用多信息融合的方法实现目标的描述, 文献 [18,19] 引出基于像素颜色信息的背景建模方法, 而常用的 RGB 颜色空间将颜色和亮度信息混杂, 当亮度改变, 三个颜色通道的值都会发生改变, 从而降低了颜色的判别能力. HSI 空间包含色度 (hue)、饱和度 (saturation) 和亮度 (intensity) 信息, 是一种更符合人类视觉特点的颜色表达, 文献 [20] 通过实验验证了其在目标检测和跟踪方面的性能优于 RGB, 同时, 颜色对目标形状、尺度等具有较好的不变性, 而梯度对光照等具有较好的不变性 [21]. 据此, 本文提出通过颜色和梯度特征相结合的方法完成背景模型的建立, 首先对原始灰度图像求 1 阶梯度, 随后将图像从 RGB 空间转换到 HSI 空间, 分别对各颜色分量和梯度分量进行背景模型的初始化和训练, 最终得到 HSG(hue, saturation, gradient) 三通道的混合高斯建模结果, 最终完成运动目标的检测: 如果当前像素点的三个通道值与通道内任意一个高斯分布匹配成功, 则该点为背景点, 否则判断为运动前景.

为验证 HSG-GMM 算法的执行效率, 选取 Intel Core i5-5300U CPU, 8G 内存的硬件平台运行算法程序, 视频帧图像大小为  $352 \times 228$  时, RGB-GMM 算法、HSG-GMM 算法和 ViBe 算法的运行时间对比如图 3 所示. ViBe 可以用于实时目标检测, GMM 的均值一般在 0.07 s 左右, 适用

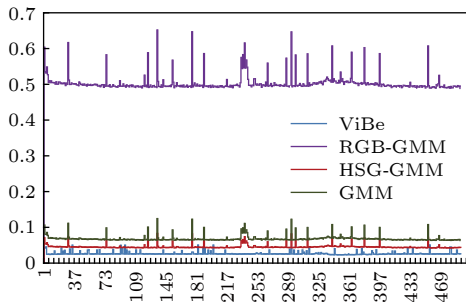


图 3 算法执行效率对比

Fig. 3. Algorithm execution efficiency comparison.

于实时性要求不高的目标检测领域, RGB-GMM 运行效率最低, 只能用于离线检测; HSG-GMM 算法采用三通道的高斯建模方法, 虽然较原始 GMM 算法增加了通道数目, 由于引入了随机子采样和邻域相关度量, 导致最后的总体计算量与 GMM 相比不增反减, 其实际运行时间虽然不能和 ViBe 相比, 但是在满足精度要求下仍能达到实时处理的效果.

```

for  $t = 1$  to  $T$  ( $T$  为序列的总帧数)
    读取图像, 完成 H, S 和 G(梯度) 三个不同通道的特征采集
    If  $t < 5$  then
        初始化: 采用帧差法完成不同通道的高斯模型的初始化得到  $\omega_{k,t}, \mu_{k,t}, \sigma_{k,t}^2$  初值
    Else
        for 通道 = 1 to 3(分别对应 HSG 三个通道)
            步骤 1 随机子采样: 根据  $t - 1$  时刻的背景检测结果, 随机选取 1/16 背景点作为匹配位置
            匹配
            if 当前点是需要匹配的背景点或目标点
                步骤 2 根据式 (3), 判断当前点与高斯分布是否匹配
            训练/更新
            步骤 3 更新当前点的高斯混合模型参数, 式 (15-17)
            步骤 4 查找并标记当前点邻域范围内的邻域相关度, 式 (14)
            步骤 5 根据步骤 3 的更新结果直接对邻域相关像素进行模型替换
            步骤 6 生成背景分布, 式 (5), 结合三个通道检测运动目标
        Else(没有被随机子采样选中的点)
            维持当前像素的背景模型不变
    End
End
End
End
    
```

## 5 实验与结果分析

采用 Universidad Politecnica de Madrid(UPM)

的 Grupo de Tratamiento de Imagenes 的 LASIESTA Database<sup>[22]</sup>, 进行实验.

### 5.1 性能分析

根据准确率和检测精度指标<sup>[23]</sup>来考察算法的性能:

$$\text{精度度(precision)} : P = TP / (TP + FP), \quad (17)$$

$$\text{准确率(accuracy)} : A = TP / (TP + FN) = (TP + TN) / (TP + FN + FP + TN). \quad (18)$$

将数据库中的 O\_CL\_01 数据集完成检测准确度和精度的检测, 分别将传统的 GMM, 基于 RGB 三通道的 GMM(RGB-GMM)、ViBe 算法和本文 HSG 算法 (HSG-GMM) 进行比较.

对于 O\_CL\_01 数据集中, 该图像序列共有 216 帧, 其中第 135—206 帧存在有树遮挡的运动目标, 图 4 为四种算法的精度度和准确率分布曲线, 图 4(a) 和图 4(b) 为整个视频序列的检测结果, 图 4(c) 和图 4(d) 为 135—206 帧存在运动目标时分布曲线的放大结果. 可以看出四种算法静态背景

的检测结果基本一致, 但对于运动目标检测的准确率和精度度就有所差别. GMM 算法对某些干扰图像中的检测结果要明显低于其他算法, HSG-GMM 算法与 RGB-GMM 算法比较, 在准确率和精度度上都提高; 由于 ViBe 算法的基本原理与其他三种不同, 检测结果的分布也呈现不同规律: 在某些区域 ViBe 可以获得很高的精度, 但在运动目标区域又呈现大范围波动.

图 5 为对第 161 帧视频图像的检测结果, 对于纯背景帧的检测, 由于不存在运动目标, 因此  $P$  基本为零, 四种算法结果基本一致, 只是在某些帧中含有少量噪声; 而对于运动目标的检测, GMM 和 RGB-GMM 产生目标空洞较多, 虽然 ViBe 对背景噪声的抑制效果更好, 但是目标区域仍然存在较大空洞.

### 5.2 效果比较

#### 1) 静态背景测试

采用视频序列 signal.avi 检验初始静态背景中存在运动目标时算法的有效性, 如图 6 所示

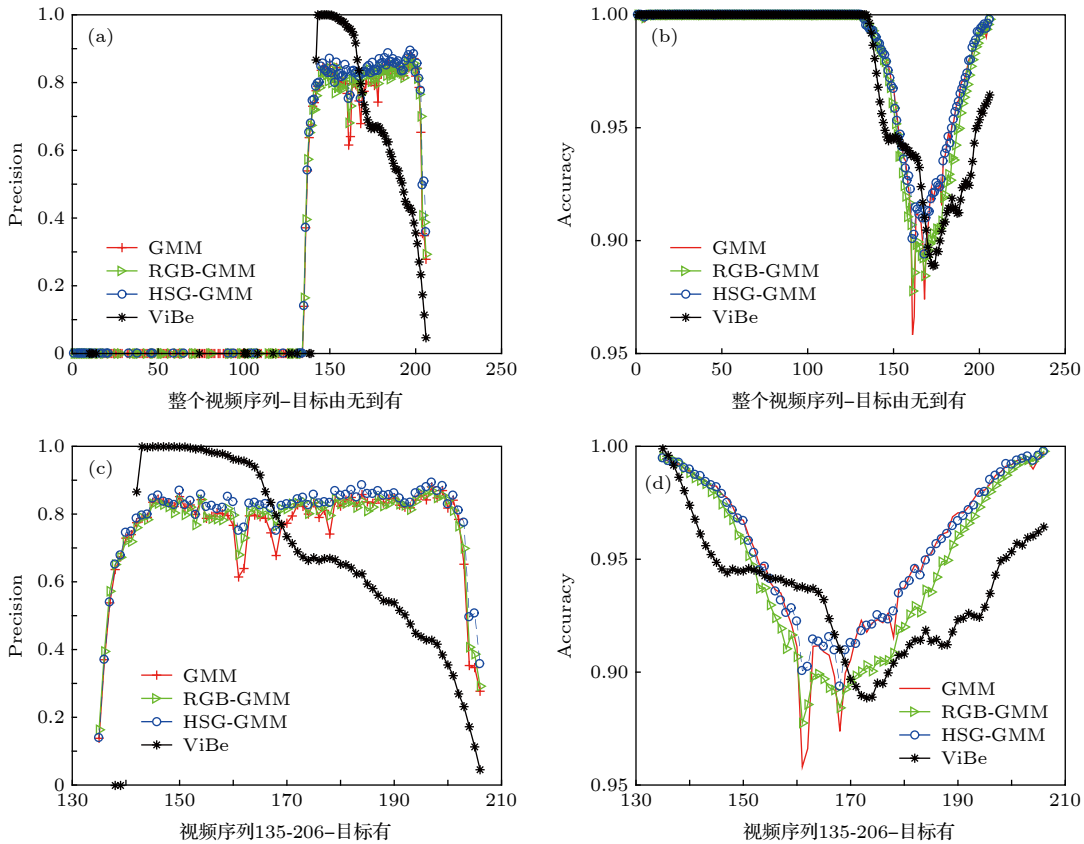


图 4 算法精度度和准确率验证 (a)、(c) 精度度; (b)、(d) 准确率

Fig. 4. The accuracy and accuracy of the algorithm verify the results: (a)、(c) precision; (b)、(d) accuracy.

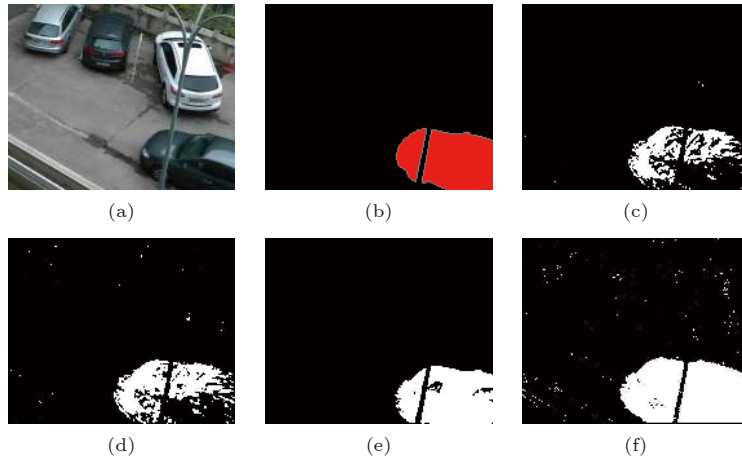


图 5 O\_CL\_01 数据集第 161 帧的处理结果 (a) 原始图像; (b) 标注真值; (c) GMM; (d) RGB\_GMM; (e) ViBe; (f) HSG\_GMM

Fig. 5. The result of the frame 161 in O\_CL\_01 data set: (a) Original image; (b) true Value image; (c) GMM result; (d) RGB\_GMM result; (e) ViBe result; (f) HSG\_GMM result.

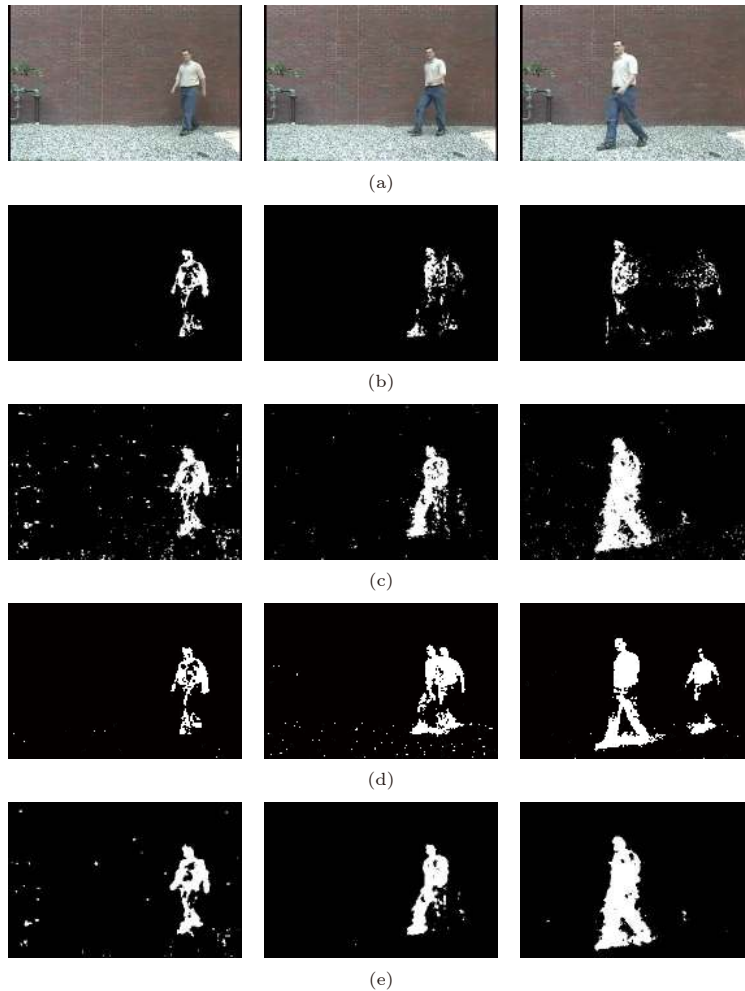


图 6 初始视频中存在动目标的静态背景检验结果 (a) 原始图像; (b) GMM 算法结果; (c) RGB-GMM 结果; (d) ViBe 结果; (e) HSG-GMM 结果

Fig. 6. Detection result of moving target in initial video in static background: (a) Initial image; (b) GMM result; (c) RGB-GMM result; (d) ViBe result; (e) HSG-GMM result.

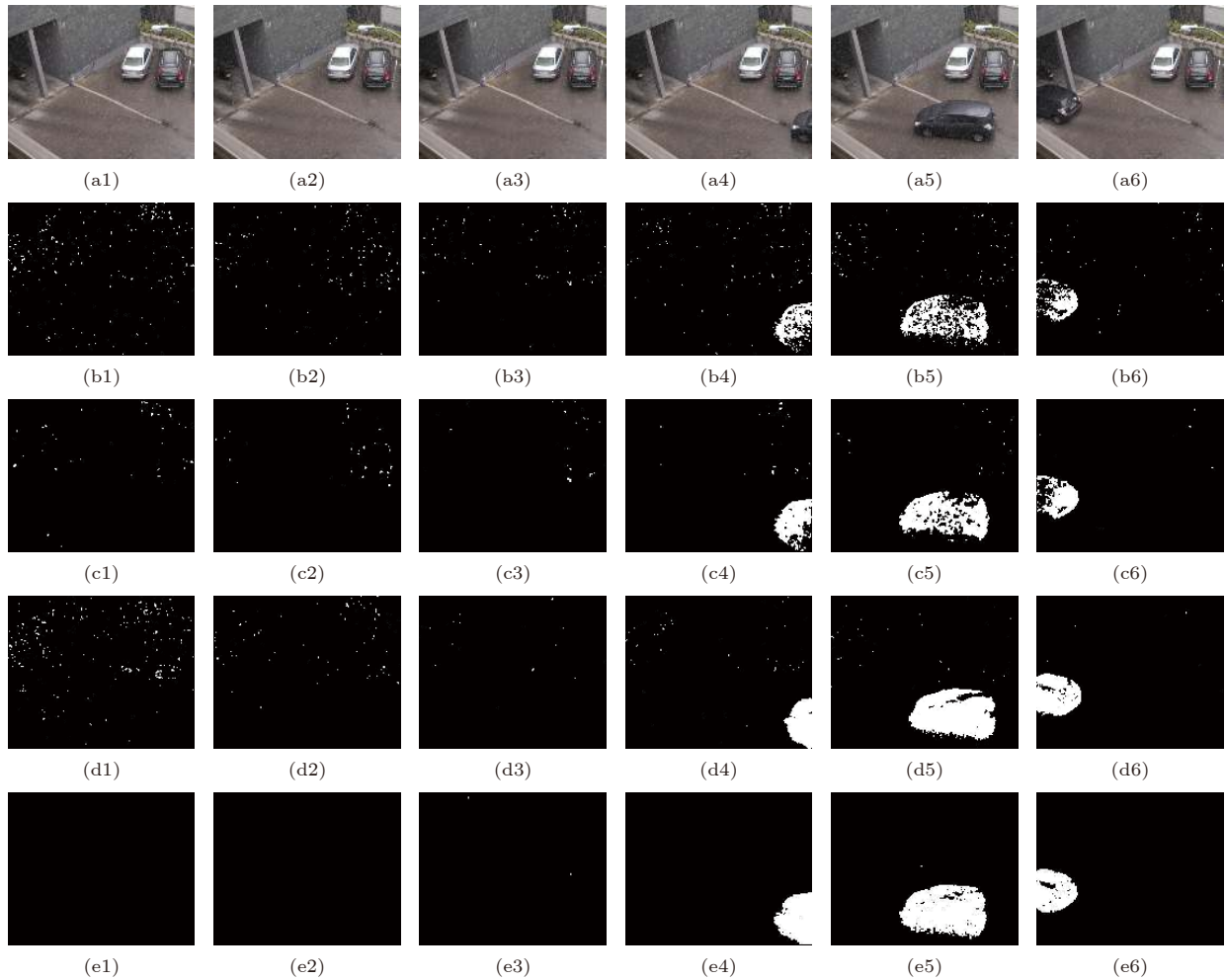


图 7 动态背景环境中的运动目标检测 (a) 原始图像;(b) GMM; (c) RGB-GMM; (d) ViBe; (e) HSG-GMM

Fig. 7. Moving target detection in dynamic background environment: (a) Initial image;(b) GMM; (c) RGB-GMM; (d) ViBe; (e) HSG-GMM.

图 6(a) 分别为第 6 帧、14 帧、58 帧原始图像, 图 6(b) 为 GMM 算法的检测结果, 图 6(c) 为 RGB-GMM 的检测结果, 图 6(d) 为 ViBe 的检测结果, 图 6(e) 为本文 HSG-GMM 的检测结果. GMM 算法和 ViBe 算法中对初始视频序列中存在运动目标的处理都会导致拖影问题; RGB-GMM 由于颜色特征又可能使某些场景中存留的噪声得以保留.

HSG-GMM 特征信息更加完备, 背景模型的建立更加准确, 运动目标的检测更加完整, 并且克服了 RGB 值不稳定导致的噪声干扰; 由于初始化阶段考虑了帧间数据的关联性, 将邻域数据的一致性加入到模型参数的初始化环节, 去除了运动目标在初始阶段对建模的影响; 采用的基于像素空间传播特性的更新策略, 使背景模型逐渐向邻域扩散, 有利于加快 Ghost 区域的识别速度, 拖影的抑制效果也比较明显.

## 2) 动态复杂背景测试

图 7 中是对数据集 O\_SN\_01 中下雪场景中运动目标的检测, 1—6 列分别为视频序列的第 10, 48, 308, 320, 348, 361 帧的处理结果.

GMM, RGB-GMM 算法对雪花类的明显动态背景有一定的抑制, 但不能消除, 且获得的目标空洞较严重; 本文的 HSG-GMM 采用观测向量的方法, 抑制动态噪声的能力要高于 ViBe, 获得参数优化后的背景模型更加接近实际, 最终检测的运动目标区域也较为完整.

## 6 结 论

通过改变高斯模型的初始化方法和参数的计算方法, 加快了模型的收敛速度, 将随机子采样与邻域空间传播理论相结合, 充分利用像素的邻域相

关性简化模型更新过程,降低运算量,并将 HSI 空间的色度信息与一阶梯度信息相结合,构建多通道的模型建立和更新机制,使背景模型的建立更加完善. 实验表明,算法在抑制动态背景的影响、检测目标完整性方面有很大提升,目标空洞现象更少,但是对于摄像机运动条件下的背景建模效果却不是很理想.

## 参考文献

- [1] Barron J L, Fleet D J, Beauchemin S S 1994 *Computer Vision and Pattern Recognition Manufactured*, Netherlands, February, 1994 p43
- [2] Lai L J, Xu Z Y, Zhang X Y 2016 *Infrared and Laser Engineering* **45** 273 (in Chinese) [赖丽君, 徐智勇, 张翎铤 2016 红外与激光工程 **45** 273]
- [3] Cui Z G, Wang H, Li A H 2017 *Acta Phys. Sin.* **66** 084203 (in Chinese) [崔智高, 王华, 李艾华 2017 物理学报 **66** 084203]
- [4] Li W, Yao J G, Dong T Z, Li H 2016 *International Congress on Image & Signal Processing Shenyang, China* October 14—16, 2015 p969
- [5] Staffer C, Grimson W L 1999 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* Fort Collins, Colorado, June 23—25, 1999 p2246
- [6] Wren C R, Azarbayejani A, Darrell T, Pentland P A 1997 *IEEE Transactions on Pattern Analysis and Machine Intelligence* Washington, DC, USA 1997 p780
- [7] Zhao X D, Liu P, Tang J L 2011 *Acta Automatic Sinica* **37** 915 (in Chinese) [赵旭东, 刘鹏, 唐降龙 2011 自动化学报 **37** 915]
- [8] Thierry B, Baf F E, Vachon B 2008 *Recent Patents on Computer Science* **1** 219
- [9] Kaewtrakulpong P, Bowden R 2002 *Kluwer Academic Publishers* 135
- [10] Kang K, Cao Y, Zhang J, Wang Z F 2016 *Multimedia Tools & Applications* **75** 1443
- [11] Li X Y, Ma D Z, Fu Y J 2018 *Journal of Jilin University (Information Science Edition)* **36** 61 (in Chinese) [李晓瑜, 马大中, 付英杰 2018 吉林大学学报(信息科学版) **36** 61]
- [12] Zhu W J, Wang G L, Tian J, Qiao Z T, Gao F Q 2018 *Journal of Beijing Institute of Technology* **38** 165 (in Chinese) [朱文杰, 王广龙, 田杰, 乔中涛, 高凤岐 2018 北京理工大学学报 **38** 165]
- [13] Chen Y Y, Wang J Q, Lu H Q 2015 *IEEE International Conference on Multimedia and Expo Turin, Italy, June 2 9*
- [14] Jeon M, Noh S, Noh S J, Jeon M 2012 *Asian Conference on Computer Vision* Daejeon, Korea, November 05—09, 2012 p493
- [15] Choi M, Sweetman B 2013 *Structural Health Monitoring* **9** 13
- [16] Barnich O, Droogenbroeck M V 2011 *IEEE Trans. Image Process.* **20** 1709
- [17] Xu Y, Dong J T, Wang S H 2010 *Acta Phys. Sin.* **59** 7535 (in Chinese) [徐艳, 董江涛, 王少华 2010 物理学报 **59** 7535]
- [18] Maha M A, Shedeed H A, Hussein A S 2010 *International Conference on Image Processing* Hong Kong September 26—29, 2010 p3453
- [19] Li Y D, Xu X P, Chen J, Wang H C 2017 *Chinese Journal of Scientific Instrument* **38** 445 (in Chinese) [李艳获, 徐熙平, 陈江, 王鹤程 2017 仪器仪表学报 **38** 445]
- [20] Martin D, Fahad S K, Michael F, Weijer J V D 2014 *IEEE Conference on Computer Vision and Pattern Recognition* Columbus, USA, June 23—28, 2014 p1090
- [21] Jiang Y S, Ma J W 2015 *IEEE Conference on Computer Vision and Pattern Recognition* Boston, USA, June 7—12, 2015 p240
- [22] Cuevas C, Yáñez E M, García N 2016 *Comput. Vision and Image Understanding* **152** p103
- [23] Goyette N, Pierre M J, Porikli F, Konrad J, Ishwar P 2012 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* Providence, Rhode Island, June 16—21, 2012 p1

# Moving target detection algorithm based on spatiotemporal correlation multi-channel clustering\*

Xu Yan<sup>1)2)</sup> Wang Pei-Guang<sup>1)†</sup> Yang Qing<sup>2)3)</sup> Dong Jiang-Tao<sup>4)</sup>

1) (*College of Electronic Information Engineering, Hebei University, Baoding 071002, China*)

2) (*Department of Electronic and Optical Engineering, Army Engineering University Shijiazhuang Campus, Shijiazhuang 050000, China*)

3) (*College of Electronic Information Engineering, Hebei University of Technology, Tianjin 300401, China*)

4) (*The 54th Research Institute of China Electronics Science and Technology Corporation, Shijiazhuang 050000, China*)

( Received 28 January 2019; revised manuscript received 20 May 2019 )

## Abstract

In the process of tracking target, certain multi-modal background scenes are not suitable for the off-line training model, and moving target detection is affected as background in the current video environment is mostly multi-modal scene with much noise, and the characters of moving targets irregularly change, which, therefore, requires a more stable and robust moving target detection algorithm. To solve this problem, taking advantage of spatiotemporal relationship learning, the mixed Gaussian model (GMM) is improved in three aspects.

First, the initialization method combining five-frame difference and intra-frame neighborhood average is proposed to obtain the initial parameters of the mixed Gaussian model. The five-frame difference method is introduced to obtain the initial parameters of the model, so that the background model is closer to the real scene. The intra-frame neighborhood average value is introduced, and an accumulation matrix CA is proposed to record the number of neighboring pixel points, then to enhance the information relevant to the neighborhood. This process can reduce the discontinuity of the target.

Second, the calculation method of the neighborhood correlation is introduced to update the parameter of Gaussian model. Since the single pixel feature is related to the neighborhood random correlation, the random subsampling technology and neighborhood spatial propagation theory are combined together, and the execution efficiency is taken into account to simplify the process of updating model. To speed up the model convergence, an observation vector is built in the time dimension to optimize the model parameters, and the weight  $\omega$  is gained based on the posterior probability.

Then, the color-gradient method incorporated with the color HSI space and gradient information is adopted in this paper to complete the multi-channel Gaussian mixture model. The initial and the updated parameters of the Gaussian model in each channel can be acquired via the above steps. To simplify the computation of three channels, the random sampling of background pixels is introduced. Finally the detection of moving targets in complex environments is realized.

The experiments show that the proposed algorithm has a great improvement in suppressing the influence of complex background and detecting target integrity, and the influence of the moving target in the initial stage is eliminated.

**Keywords:** Gaussian mixture, random subsampling, neighborhood correlation, multi-channel

**PACS:** 42.30.Tz, 42.30.Va, 42.66.Lc, 47.50.Cd

**DOI:** [10.7498/aps.68.20190161](https://doi.org/10.7498/aps.68.20190161)

\* Project supported by the National Natural Science Foundation of China (Grant No. 11771115).

† Corresponding author. E-mail: [pgwang@hbu.edu.cn](mailto:pgwang@hbu.edu.cn)