

文章编号: 1006-2475(2020)06-0007-07

结合 HED 网络和双阈值分割的 GMM 目标检测算法

李睿, 王德忠

(兰州理工大学计算机与通信学院, 甘肃 兰州 730000)

摘要: GMM 在目标检测过程中容易受到灯光、目标颜色与背景颜色相似、目标阴影和拍摄高度等因素的干扰。针对以上问题, 本文提出一种结合改进 HED 网络和 OTSU 双阈值分割的 GMM 算法。首先, 改进模型针对视频帧的背景、噪声、前景目标进行双阈值分割, 合理选取高斯模型个数。其次, 利用 HED 网络对输入图片进行边缘检测, 将 HED 网络检测的边缘结果和双阈值分割的 GMM 检测结果进行“与”运算, 得到最终目标检测结果。通过实验验证, 改进算法的检测率更高, 目标较小时检测轮廓更加完整, 检测效果更好。

关键词: OTSU 双阈值; HED 网络; 边缘检测; 与运算

中图分类号: TP391.4 文献标识码: A doi: 10.3969/j.issn.1006-2475.2020.06.002

Target Detection Algorithm of Gaussian Mixture Model Combined with HED Network and Double Threshold Segmentation

LI Rui, WANG De-zhong

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730000, China)

Abstract: In the process of target detection, the GMM is easily interfered by lighting, the similarity of target color and background color, target shadow and shooting height. Aiming at the above problems, a GMM algorithm is combined with improved HED network and OTSU double threshold segmentation is proposed. First, the improved model divides the background, noise and foreground targets of video frame by double threshold, and reasonably selects the number of GMM. Secondly, HED network is used for edge detection of the input images. The “and” operation of edge result detected by the HED network and the GMM detection result of the double threshold segmentation is completed to obtain the final target detection result. Experimental results show that the improved algorithm has a higher detection rate, a more complete detection profile “and” a better detection effect.

Key words: OTSU double threshold; HED network; edge detection; “and” operation

0 引言

运动目标检测是计算机视觉领域中非常重要的研究内容, 主要目的是提取出视频序列中的运动目标^[1]。经典的目标检测方法有光流法、帧差法、背景建模法^[2-3]。光流法是找到图像中每个像素点的速度矢量, 通过各个像素点的速度矢量特征对图像进行动态分析, 其缺点就是计算量较大、无法保证实时性和实用性^[4-5]。帧差法是对时间上连续的 2 帧或者多帧进行差分运算, 通过连续帧图像的相同位置像素点相减后的灰度差的绝对值与设定阈值的比较判断运动目标^[6], 其最大缺点是检测的目标轮廓不完整、内部

含有空洞^[7]。背景建模法的基本思想是对图像的背景进行建模, 建立的背景模型与当前检测的视频帧进行比较, 通过阈值确定前景目标^[8]。常用的背景建模方法有: 平均背景模型、非参数化背景模型、背景模型、VIBE 算法、核密度估计方法、GMM 算法^[9-11]。

GMM 算法是最常用的背景建模方法, 可以获得更加理想的检测结果, 也是当前目标检测研究的热点^[12]。针对 GMM 算法的不足之处, 研究人员对混合高斯模型算法提出了许多改进算法。文献[13]针对模型在背景更新时收敛性差、易受环境噪声、光照突变影响和易产生虚假目标等问题, 提出了将帧差法引入 GMM 算法, 快速区分背景区域和运动目标区域,

收稿日期: 2019-11-08; 修回日期: 2019-12-10

基金项目: 国家自然科学基金资助项目(61761028); 甘肃省重点研发计划资助项目(18YF1GA060)

作者简介: 李睿(1971-), 女, 甘肃秦安人, 教授, 硕士生导师, 硕士, 研究方向: 模式识别, 数字图像处理, 智能信息处理, 数字水印, E-mail: 927692061@qq.com; 王德忠(1993-), 男, 甘肃兰州人, 硕士研究生, 研究方向: 智能信息, 多媒体信号处理, E-mail: 1594814035@qq.com。

消除虚假目标。文献[14]针对复杂背景下运动目标检测的问题,将视频帧的空间信息引入建模过程中,研究了邻域更新、二维联合直方图信息熵判别,提高了目标检测的准确性和完整性。文献[15]针对传统GMM算法检测耗时长、复杂度高、突发运动和光照敏感等因素的影响,结合GMM算法和六帧差分算法获得前景运动目标的轮廓,并且使用不同更新率自适应环境变化,该算法提高了检测目标的清晰度,有效地解决了噪声对检测结果的影响。

以上文献中的算法对GMM算法做出了改进,但是在以下环境中检测结果仍不理想:1)视频存在灯光,运动目标产生影子;2)运动目标较小,距离较远;3)运动目标颜色与背景颜色相似。以上外界因素的影响会导致误检率和漏检率较高。针对GMM算法中的上述问题,本文提出结合改进HED网络和三维OTSU双阈值分割的GMM算法。针对HED网络在目标边缘检测中的问题,提出一种改进的HED网络模型,通过三维OTSU算法对视频帧进行双阈值分割,将视频帧分割为背景、噪声、前景运动目标3个部分,不同区域选取不同模型个数。改进算法提高了检测准确性。

1 GMM 算法

GMM算法是基于像素统计信息的概率密度模型^[16-17],针对视频帧中的每一个像素点,计算该像素点的平均灰度值 μ_0 和方差 σ_0 ,每个像素点定义 K 个高斯模型。 K 一般取值为3~5,通过加权和描述像素点的状态。各项表达式如下:

$$\mu_0 = \frac{1}{N} \sum_{i=0}^{N-1} x_i \quad (1)$$

$$\sigma_0^2 = \frac{1}{N} \sum_{i=0}^{N-1} (x_i - \mu_0)^2 \quad (2)$$

$$p(x_i) = \sum_{i=1}^K \omega_{i,t} \eta(x_i, \mu_{i,t}, \Sigma_{i,t}) \quad (3)$$

$$\eta(x_i, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{i,t}|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_i - \mu_{i,t})^T \Sigma_{i,t}^{-1} (x_i - \mu_{i,t})} \quad (4)$$

其中 x_i 是视频帧中任意一个像素点的观测值, $p(x_i)$ 表示观测值 x_i 出现的概率, $\omega_{i,t}$ 是 t 时刻混合高斯模型中第 i 个高斯分布的权重, $\sum_{i=1}^K \omega_{i,t} = 1$, $\mu_{i,t}$ 是 t 时刻第 i 个高斯分布的均值, $\Sigma_{i,t}^{-1}$ 为 t 时刻第 i 个高斯分布的协方差矩阵的逆矩阵,当观测值 x_i 相互独立时 $\Sigma_{i,t} = \sigma_{i,t}^2 I$, I 为单位矩阵。 n 为 x_i 的维数, $\sigma_{i,t}^2$ 是像素点第 i 个高斯分布在时刻 t 的方差, $\eta(x_i, \mu_{i,t}, \Sigma_{i,t})$ 是第 i 个高斯分布在 t 时刻的概率密度函数。

随着检测视频帧的变化,模型参数相应地进行更新,新观测值 x_i 需要与当前存在的 k ($1 \leq k \leq K$)个高斯模型进行匹配,当 $|x_i - \mu_{i,t-1}| \leq 2.5\sigma_{i,t-1}$,判定该像素点与模型匹配。对于没有匹配成功的模型,均值 μ_0 和方差 σ_0 不变,匹配成功的第 i ($1 \leq i \leq k$)个高斯分布,权值 $\omega_{i,t}$ 、均值 μ_0 和方差 σ_0 通过以下公式更新:

$$\omega_{i,t} = (1 - \alpha) \omega_{i,t-1} + \alpha M_{i,t} \quad (5)$$

$$\mu_{i,t} = (1 - \beta) \mu_{i,t-1} + \beta x_i \quad (6)$$

$$\sigma_{i,t}^2 = (1 - \beta) \sigma_{i,t-1}^2 + \beta (x_i - \mu_{i,t})^T (x_i - \mu_{i,t}) \quad (7)$$

其中 α ($0 < \alpha < 1$)为用户自定义的学习率, α 的大小决定了背景更新速度和抗噪声性能。 $\beta = \alpha \eta(x_i, \mu_{i,t}, \Sigma_{i,t})$ 为参数学习率。

参数更新完成后,把 K 个高斯分布按 $\rho_{i,t}$ 降序排列。选取前 B 个高斯分布作为背景像素的最佳描述模型。若像素值满足以下公式,则判定像素点为背景点,否则为前景点。其中 T 为背景选取的阈值。

$$B = \arg \min_b \left(\sum_{i=1}^b \omega_{i,t} > T \right) \quad (8)$$

2 HED 网络

HED网络是基于VGG网络修改而来,设计了5个侧边输出层,输出层的感受野在不断增大,语义信息也渐渐丰富^[18-19],去除了VGG网络的全连接层。去除全连接层的HED网络可以输入任意大小的图片,降低了网络参数数目和计算复杂度,提升了运行效率。在训练过程中,HED网络改进了训练代价函数,由于边缘提取中轮廓线与非轮廓线的像素数有着巨大的差异,使得使用正常的损失函数进行训练会造成模型的不稳定,HED网络对边缘位置损失函数的系数进行了放大,损失函数如下:

$$L_{\text{side}}(W, \mu) = \sum_{m=1}^M \alpha_m l_{\text{side}}^{(m)}(W, \mu^{(m)}) \quad (9)$$

$l_{\text{side}}^{(m)}$ 表示侧输出的图像级损失函数:

$$l_{\text{side}}^{(m)}(W, \mu^{(m)}) = -\beta \sum_{j \in Y_+} \log p_r(y_j = 1 | X; W, \mu^{(m)}) - (1 - \beta) \sum_{j \in Y_-} \log p_r(y_j = 0 | X; W, \mu^{(m)}) \quad (10)$$

其中 $\beta = |Y_-|/|Y|$, $1 - \beta = |Y_+|/|Y|$ 。 $|Y_-|$ 、 $|Y_+|$ 分别表示边缘和非边缘的像素由侧边输出的响应值,通过逻辑回归函数计算可以得到:

$p_r(y_j = 1 | X; W, \mu^{(m)}) = \sigma(a_j^{(m)}) \in [0, 1]$ 。在每一个侧边输出层通过逻辑回归得到其响应值,对 m 个侧边响应输出:

$$L_{\text{fuse}}(W, \mu, h) = \text{Dist}(Y, Y_{\text{fuse}}) \quad (11)$$

其中 $Y_{\text{fuse}} = \sigma\left(\sum_{m=1}^M h_m A_{\text{side}}^{(m)}\right)$, h 是融合权重, $\text{Dist}(\cdot)$ 是融合预测结果与标签之间的距离。最终网络根据融

合代价与侧边代价的和利用梯度下降法对各卷积核内的参数以及融合权重进行迭代优化,如式(12):

$$(W, \mu, h)^* = \arg \min (L_{\text{side}}(W, \mu) + (W, \mu, h)) \quad (12)$$

3 OTSU 阈值分割

OTSU 阈值分割的核心思想是通过阈值将图像分为前景目标和背景图像^[20-21],并计算两者的最大方差为最佳分割阈值^[22-23]。假设一幅图像大小为 $m \times n$ 图像灰度级的取值范围为 $[0, L - 1]$ 。像素点 (x, y) 的灰度值为 $h(x, y)$, 在 $k \times k$ 的邻域内该像素点的平均灰度值为 $f(x, y)$ 。假设 $h(x, y) = i, f(x, y) = j, q_{ij}$ 为像素点 (x, y) 出现的次数, 则 q_{ij} 的联合概率密度为:

$$p_{ij} = \frac{q_{ij}}{m \times n} \quad (13)$$

假设 D_0 和 D_1 分别表示背景区域和目标区域, ω_0 和 ω_1 分别表示 D_0 和 D_1 的概率, μ_0 和 μ_1 分别表示 D_0 和 D_1 的均值矢量:

$$\omega_0 = \sum_{i=0}^{s-1} \sum_{j=0}^{t-1} p_{ij} \quad (14)$$

$$\omega_1 = \sum_{i=s}^{L-1} \sum_{j=t}^{L-1} p_{ij} \quad (15)$$

$$\mu_0 = (\mu_{0i}, \mu_{0j})^T = \left[\sum_{i=0}^{s-1} \sum_{j=0}^{t-1} \frac{ip_{ij}}{\omega_0}, \sum_{i=0}^{s-1} \sum_{j=0}^{t-1} \frac{jp_{ij}}{\omega_0} \right]^T \quad (16)$$

$$\mu_1 = (\mu_{1i}, \mu_{1j})^T = \left[\sum_{i=s}^{L-1} \sum_{j=t}^{L-1} \frac{ip_{ij}}{\omega_1}, \sum_{i=s}^{L-1} \sum_{j=t}^{L-1} \frac{jp_{ij}}{\omega_1} \right]^T \quad (17)$$

图像的总均值矢量为:

$$\mu = (\mu_i, \mu_j)^T = \left[\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ip_{ij}, \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} jp_{ij} \right]^T \quad (18)$$

类间离散度矩阵为:

$$S_B(s, t) = \omega_0(\mu_0 - \mu)^T(\mu_0 - \mu) + \omega_1(\mu_1 - \mu)^T(\mu_1 - \mu) \quad (19)$$

当 $T_r(S_B(s, t))$ 取最大值时,可获得最佳阈值:

$$T_r(S_B(s, t)) = \omega_0 [(\mu_{0i} - \mu_i)^2 + (\mu_{0j} - \mu_j)^2] + \omega_1 [(\mu_{1i} - \mu_i)^2 + (\mu_{1j} - \mu_j)^2] \quad (20)$$

4 结合 HED 网络和双阈值分割的 GMM 算法

HED 网络在卷积主干网络上增加多个输出层, 在输出层进行深度监督, 从而得到不同尺度对应的边缘。由于高层图分辨率过低, 很容易丢失目标的部分轮廓信息。过多的池化层导致侧边输出边缘检测结果宽度过大, 与低层融合时会引入较多的内部纹理信息。针对这个问题, 去掉第 4 层池化层, 针对轮廓信息丢失严重的问题, 选取了局部能量函数计算出中心点与边缘点重合时局部能量的最大值。改进的 HED 网络结构如图 1 所示。

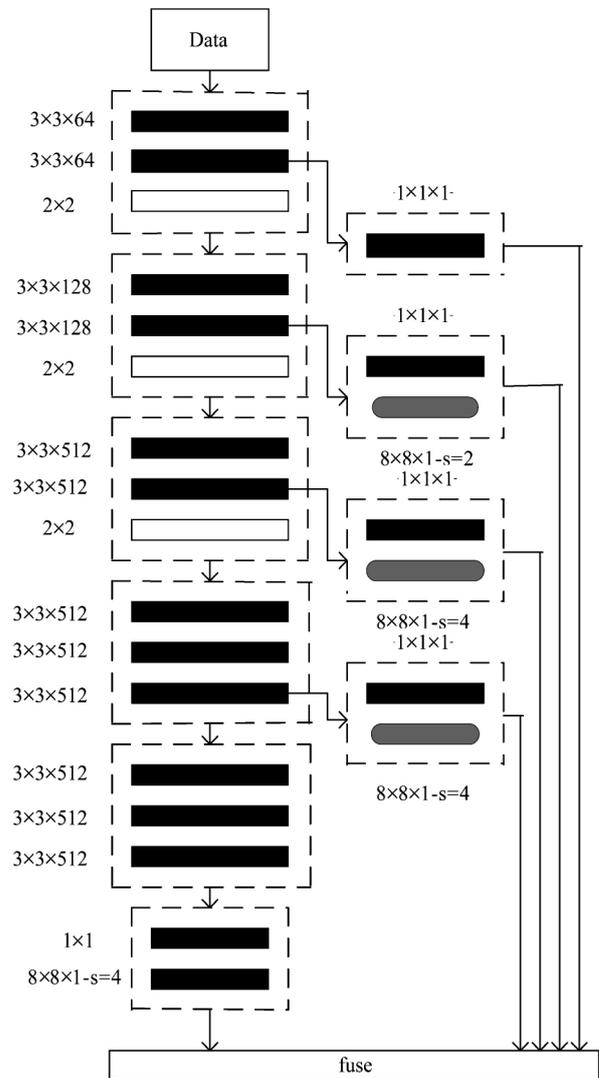


图 1 改进的 HED 网络结构

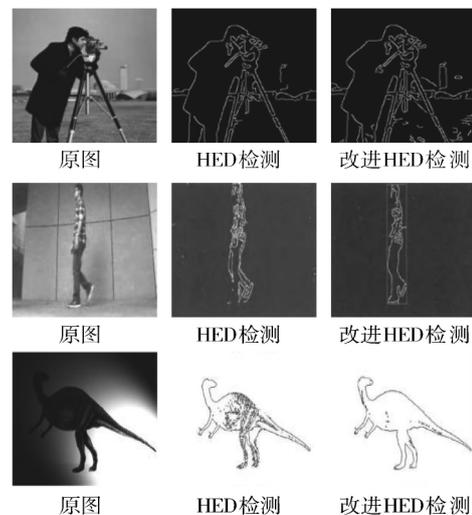


图 2 边缘检测对比图

减少了池化层使得降采样次数也降低了, 侧边输出层的图像分辨率更高, 通过局部能量函数的最大值快速准确地确定边缘位置, 解决了 HED 网络在边缘

检测中轮廓模糊的问题,从而得到了更好的检测效果。边缘检测图如图2所示。

在传统二维OTSU算法中,只考虑了背景和运动目标的概率,将噪声和边缘的概率之和近似为0,忽略了噪声和边缘的影响。然而,噪声是影响图像分割的关键因素,如果不能对噪声进行有效的分割和抑制,就会出现将背景检测为前景运动目标的现象。根据背景、噪声、运动目标将视频帧 H 分为3个部分。其中,背景图像为 X 、噪声为 Y 、前景运动目标为 Z 。视频每一帧图像大小为 $m \times n$,灰度级为 L ,其中某一像素点 (x, y) 的灰度值为 $h(x, y)$,在 $k \times k$ 的邻域内该像素点的平均灰度值为 $f(x, y)$,灰度中值为 $g(x, y)$,灰度取值均为 $0 \sim L - 1$ 。以 $h(x, y)$ 、 $f(x, y)$ 和 $g(x, y)$ 组成的三元组 (i, j, k) 出现的概率为 p_{ijk} 。各项定义如下:

$$H = X + Y + Z \quad (21)$$

$$f(x, y) = \frac{1}{k^2} \sum_{m=-\frac{k}{2}}^{\frac{k}{2}} \sum_{n=-\frac{k}{2}}^{\frac{k}{2}} h(x+m, y+n) \quad (22)$$

$$g(x, y) = \text{med}\{h(x+m, y+n)\} \quad (23)$$

$$(m = -\frac{k}{2}, \dots, \frac{k}{2}; n = -\frac{k}{2}, \dots, \frac{k}{2})$$

$$p_{ijk} = \frac{q_{ijk}}{m \times n} \quad (24)$$

其中 $q_{ijk} (0 \leq i, j, k \leq L - 1)$ 代表频数。

背景 X 、噪声 Y 和运动目标 Z 发生的概率分别为 p_X, p_Y, p_Z ,定义为:

$$p_X = \sum_{(i, j, k \in C_X(s, t, q))} p_{ijk} \quad (25)$$

$$p_Y = \sum_{(i, j, k \in C_Y(s, t, q))} p_{ijk} \quad (26)$$

$$p_Z = \sum_{(i, j, k \in C_Z(s, t, q))} p_{ijk} \quad (27)$$

三者的均值矢量为:

$$\mu_X = (\mu_{X_i}, \mu_{X_j}, \mu_{X_k})^T$$

$$= \left\{ \frac{\sum_{(i, j, k \in C_X(s, t, q))} ip_{ijk}}{p_X}, \frac{\sum_{(i, j, k \in C_X(s, t, q))} jp_{ijk}}{p_X}, \frac{\sum_{(i, j, k \in C_X(s, t, q))} kp_{ijk}}{p_X} \right\}^T \quad (28)$$

$$\mu_Y = (\mu_{Y_i}, \mu_{Y_j}, \mu_{Y_k})^T$$

$$= \left\{ \frac{\sum_{(i, j, k \in C_Y(s, t, q))} ip_{ijk}}{p_Y}, \frac{\sum_{(i, j, k \in C_Y(s, t, q))} jp_{ijk}}{p_Y}, \frac{\sum_{(i, j, k \in C_Y(s, t, q))} kp_{ijk}}{p_Y} \right\}^T \quad (29)$$

$$\mu_Z = (\mu_{Z_i}, \mu_{Z_j}, \mu_{Z_k})^T$$

$$= \left\{ \frac{\sum_{(i, j, k \in C_Z(s, t, q))} ip_{ijk}}{p_Z}, \frac{\sum_{(i, j, k \in C_Z(s, t, q))} jp_{ijk}}{p_Z}, \frac{\sum_{(i, j, k \in C_Z(s, t, q))} kp_{ijk}}{p_Z} \right\}^T \quad (30)$$

总体均值为 $\mu_T = (\mu_{T_i}, \mu_{T_j}, \mu_{T_k})^T$:

$$\mu_T = \left(\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} ip_{ijk}, \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} jp_{ijk}, \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} kp_{ijk} \right)^T \quad (31)$$

背景 X 与噪声 Y 、噪声 Y 与运动目标 Z 之间的类间离差定义为:

$$S_1 = p_X [(\mu_X - \mu_T)(\mu_X - \mu_T)^T] + p_Y [(\mu_Y - \mu_T)(\mu_Y - \mu_T)^T] \quad (32)$$

$$S_2 = p_Y [(\mu_Y - \mu_T)(\mu_Y - \mu_T)^T] + p_Z [(\mu_Z - \mu_T)(\mu_Z - \mu_T)^T] \quad (33)$$

$$t_r S_1 = p_X [(\mu_{X_i} - \mu_{T_i})^2 + (\mu_{X_j} - \mu_{T_j})^2 + (\mu_{X_k} - \mu_{T_k})^2] + p_Y [(\mu_{Y_i} - \mu_{T_i})^2 + (\mu_{Y_j} - \mu_{T_j})^2 + (\mu_{Y_k} - \mu_{T_k})^2] \quad (34)$$

$$t_r S_2 = p_Y [(\mu_{Y_i} - \mu_{T_i})^2 + (\mu_{Y_j} - \mu_{T_j})^2 + (\mu_{Y_k} - \mu_{T_k})^2] + p_Z [(\mu_{Z_i} - \mu_{T_i})^2 + (\mu_{Z_j} - \mu_{T_j})^2 + (\mu_{Z_k} - \mu_{T_k})^2] \quad (35)$$

最优阈值为 $t_r S_1$ 和 $t_r S_2$ 取得最大值, $t_r S_1$ 为背景 X 与噪声 Y 分割的最佳阈值, $t_r S_2$ 为噪声 Y 和运动目标 Z 分割的最佳阈值。定义如下:

$$(s', t', q') = \underset{\substack{0 \leq s \leq L-1 \\ 0 \leq t \leq L-1 \\ 0 \leq q \leq L-1}}{\text{Arg max}} t_r S_1(s, t, q) \quad (36)$$

$$(s'', t'', q'') = \underset{\substack{0 \leq s \leq L-1 \\ 0 \leq t \leq L-1 \\ 0 \leq q \leq L-1}}{\text{Arg max}} t_r S_2(s, t, q) \quad (37)$$

基于此,提出了结合改进的HED网络和三维OTSU双阈值分割的GMM算法。改进的HED网络结合局部能量对视频帧进行边缘检测,能够有效地抑制边缘断裂和毛刺,引入双阈值作为判断背景、噪声和运动目标3个部分的条件,根据不同的分割类型调整高斯分布的个数。将两者检测的结果融合后进行形态学“与”处理,从而最终得到检测结果。模型流程图如图3所示。

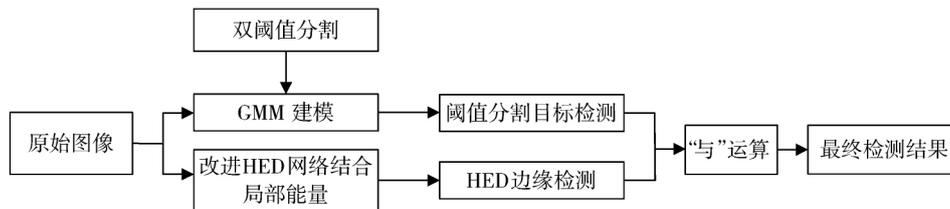


图3 模型流程图

算法描述如下:

步骤1 模型初建时,计算出视频前 N 帧的像素点的平均值以及方差,并且对视频帧中的每一个像素点构造出 K 个高斯模型。

步骤2 计算某一像素点的灰度值、平均灰度值、灰度中值。通过二维OTSU双阈值分割确定背景与噪声最佳的分割阈值、噪声和运动目标的最佳分割阈值,通过分割阈值确定出检测序列中每一个像素点

所需要的高斯模型个数,对已经建立的高斯背景模型进行更新。

步骤3 改进的HED网络结合局部能量对视频帧进行运动目标的边缘检测,提取了包含高层语义信息的显著性边缘,引入局部能量能够有效地抑制边缘断裂和毛刺,消除细毛刺对边缘检测结果的影响。

步骤4 通过双阈值分割的GMM检测结果和HED网络的边缘检测进行“与”运算,得到比较完整的运动目标。经过填充处理完善后得到最终检测结果。

5 实验及结果

实验环境: Windows7 Inter(R) core(TM) i3 CPU M380@2.53 GHz、内存为2 GB的PC,编程软件为: Matlab(2015b),本文算法中参数学习率为 $\beta = 0.0062$,高斯模型的最大个数为 $K = 5$ 。

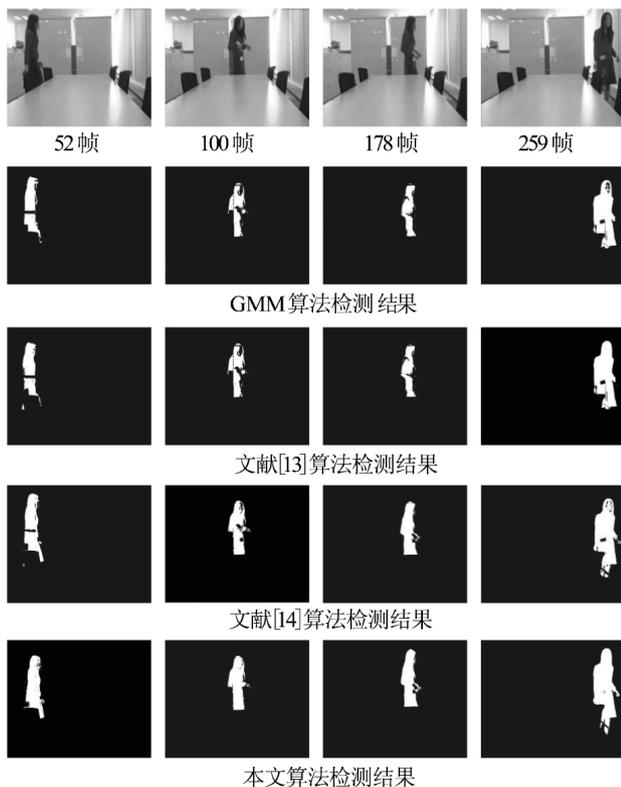


图4 视频1检测结果对比

在视频1中,目标由静止到缓慢移动的过程中,运动过程受到室内灯光和光滑平面对光线反射的影响。通过图4所示的对比结果可以看出,文献中的算法较GMM算法检测效果较好,但是检测结果中均出现空洞,目标检测不完整的现象,本文算法检测的目标更完整,漏检率更低。

视频2中有多个运动目标,而且运动目标是快速

地通过监控区域,在运动的过程中,运动目标受到太阳光照的影响,在路面上形成影子。通过图5中实验结果对比发现,GMM算法在检测过程中,不能有效地抑制影子的干扰,还将车辆中的玻璃部分检测为背景。相比于GMM算法,文献中的算法更好地抑制了误检现象,但是更容易受到地面运动目标影子的干扰,目标检测的误检率较高。通过本文算法的检测结果发现,不仅有效地检测出了运动目标,并且抑制了阴影的干扰,降低了误检率。

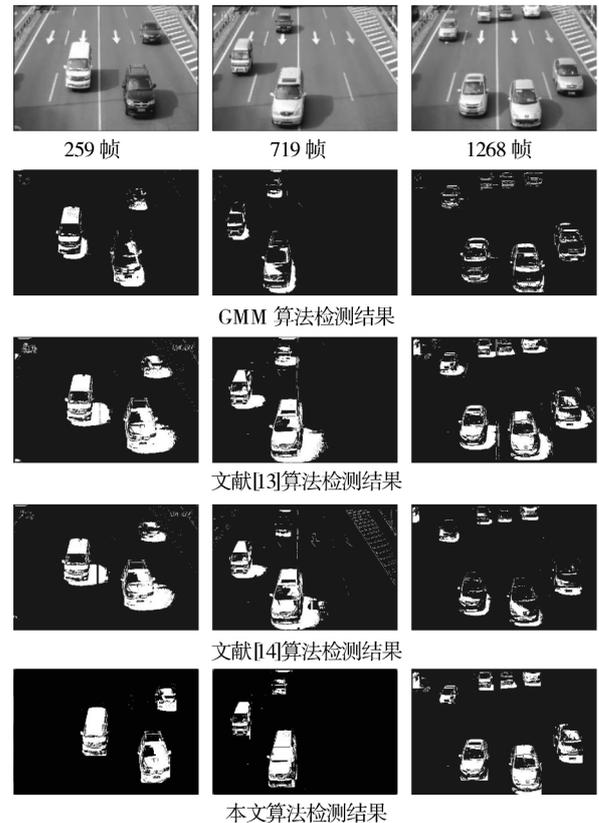


图5 视频2检测结果对比

视频3序列是从上方拍摄的,由于拍摄高度较高,视频中的运动目标较小,视频的第100帧和第215帧是单独一个人行走时截取的,第383帧是2个人握手时截取的视频画面。原视频中显示2个人的胳膊不明显。通过图6中视频序列的对比结果可以发现,第100帧画面中行人的衬衣的颜色和地面颜色相似,GMM算法和文献中的算法将运动目标误检为背景,GMM算法对行人的胳膊没有有效地检测出来。文献[13]和文献[14]中的算法虽然检测出来胳膊,但是检测的轮廓不完整,漏检率较高。本文算法有效地分割了目标和背景,抑制了颜色相似的干扰,在行人行走和2个人握手时,完整地检测出运动目标的胳膊,提高了检测的准确率。

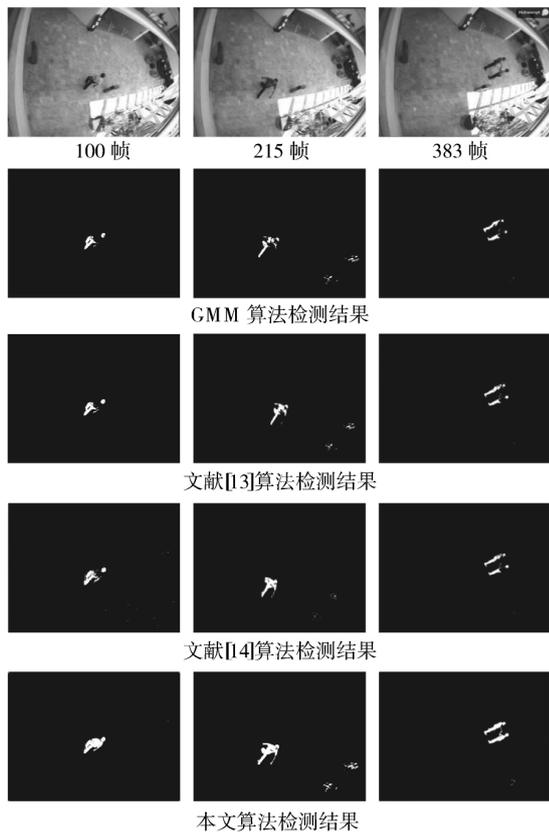


图6 视频3检测结果对比

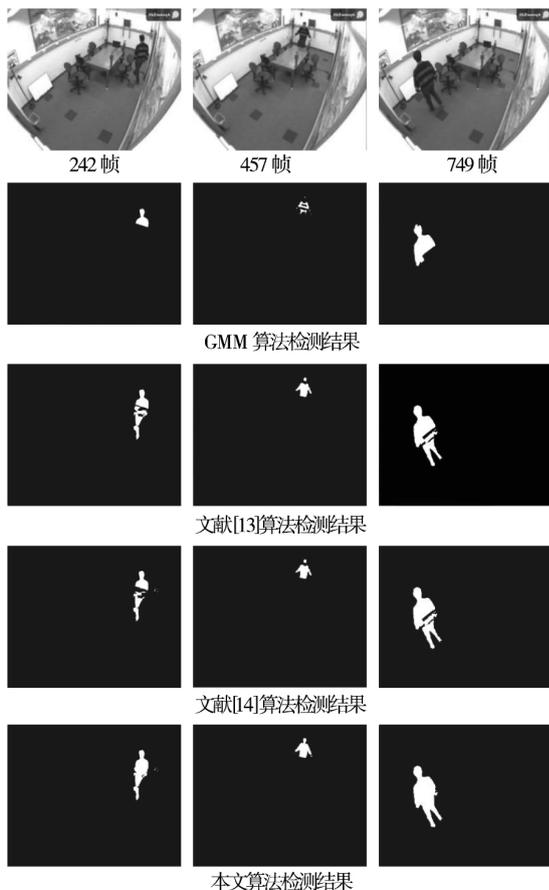


图7 视频4检测结果对比

视频4中,运动目标(人)缓慢运动,截取第242帧、457帧、749帧分别用混合高斯模型算法、文献[13]算法和[文献]14算法与本文算法进行检测比较。通过图7中视频序列的对比结果来看,GMM算法只检测出运动目标颜色和背景颜色相差较大时的目标轮廓,而完全漏检了目标的整个下半部分。文献[13]算法和文献[14]算法虽然完全检测出了目标轮廓,但是受到目标与运动背景颜色相似的影响较大,检测结果出现空洞。本文算法完整地检测出运动目标的轮廓,基本没有受到颜色相似的干扰。

评价本文算法的标准有很多,为了直观地进行分析,本文对检测结果进行量化处理,采用衡量目标检测的性能指标:检测耗时、识别率 DR 和误检率 FAR 。表达式如下:

$$DR = \frac{TP}{TP + FN} \tag{38}$$

$$FAR = \frac{FP}{TP + FP} \tag{39}$$

其中, TP 为检测出的属于真实前景的像素数, FN 和 FP 分别为未检测出的和错误检测出的前景像素数。对每一帧图像进行多次检测后,不同算法在不同视频上的各项平均指标如表1所示。

表1 不同算法检测性能指标

视频	算法	DR	FAR	耗时/ms
视频1	GMM	0.706	0.135	31.462
	文献[13]	0.749	0.072	30.091
	文献[14]	0.757	0.063	29.901
	本文算法	0.846	0.032	25.956
视频2	GMM	0.671	0.115	29.718
	文献[13]	0.743	0.103	29.460
	文献[14]	0.737	0.106	28.877
	本文算法	0.809	0.062	24.389
视频3	GMM	0.614	0.096	34.964
	文献[13]	0.678	0.083	34.469
	文献[14]	0.699	0.032	33.096
	本文算法	0.808	0.009	28.591
视频4	GMM	0.509	0.099	33.987
	文献[13]	0.771	0.052	32.138
	文献[14]	0.784	0.036	33.722
	本文算法	0.867	0.008	27.201

6 结束语

从表1可以发现,在选取的每一个视频的每一帧中,相比于GMM算法、文献[13]算法、文献[14]算法,本文算法的识别率、误检率均有所改进,检测耗时

更低。本文算法比 GMM 算法和文献 [13-14] 中的算法更加有效的原因在于:

1) 传统 GMM 算法在描述每一个像素点状态时均采用 3~5 个高斯模型,变化较大的区域应该需要较多的高斯分布描述,而变化较小的区域需要较少的高斯分布个数,GMM 算法中固定的模型个数不仅浪费了计算机运算资源还导致算法实时性较差。通过改进三维 OTSU 双阈值分割有效地抑制了噪声干扰,根据不同的变化区域阈值选取不同的高斯模型个数,缩短了算法的检测时间,提高了实时性。

2) 改进的 HED 网络结合局部能量边缘检测有效地检测出运动目标的边缘,能够准确地识别小目标运动状态,抑制了由于颜色相似引起的漏检现象。改进算法有效地解决了 GMM 算法中常出现的问题,提高了检测准确率、降低了检测耗时。

3) 三维 OTSU 进行双阈值分割,克服了 GMM 算法单阈值分割无法抑制背景噪声的干扰;解决了背景与前景运动目标误检问题;克服了检测轮廓不完整、检测出现空洞的问题。

接下来的工作重点是算法如何克服摄像头拍摄时受到恶劣天气和摄像机抖动等因素的影响,在恶劣环境下提高算法鲁棒性、实时性、检测准确率,同时降低算法复杂度、误检率和漏检率。

参考文献:

- [1] 慈文彦. 运动目标检测方法综述[J]. 信息技术, 2016 (12): 93-96.
- [2] 王梦菊, 吴小龙, 杜海涛. 基于背景差分与帧间差分的目标检测改进算法[J]. 自动化技术与应用, 2018, 37 (10): 89-92.
- [3] WANG Z H, LI C Y, LIU Q P. Improved global optical flow estimation with mean shift algorithm for target detection[J]. International Journal of Wireless and Mobile Computing, 2015, 8(4): 323-328.
- [4] 范长军, 文凌艳, 毛泉涌, 等. 结合单高斯与光流法的无人机运动目标检测[J]. 计算机系统应用, 2019, 28 (2): 184-189.
- [5] 于雯越, 安博文, 赵明. 基于光流法与 RPCA 的红外运动目标检测[J]. 现代计算机, 2018(23): 66-71.
- [6] 张应辉, 刘养硕. 基于帧差法和背景差法的运动目标检测[J]. 计算机技术与发展, 2017, 27(2): 25-28.
- [7] 刘青芳. 基于帧差法的运动目标检测方法研究[J]. 数

字通信世界, 2015(11): 54.

- [8] LI X X, CHEN L L, ZHU S H. Statistical background model-based target detection[J]. Pattern Analysis and Applications, 2016, 19(3): 783-791.
- [9] 乐应英, 胡静波. 融合 LBP 与背景建模的自适应目标检测混合算法[J]. 计算机与数字工程, 2013, 41(7): 1081-1084.
- [10] 丁哲, 陆文总. 基于 Vibe 背景建模的运动目标检测算法[J]. 计算机系统应用, 2019, 28(4): 183-187.
- [11] 董俊宁, 杨词慧. 空间约束混合高斯运动目标检测[J]. 中国图象图形学报, 2016, 21(5): 588-594.
- [12] 高海壮, 段先华. 基于帧差法和混合高斯的海上运动目标检测[J]. 计算机与数字工程, 2019, 47(5): 1140-1144.
- [13] 苏剑臣, 李策, 杨峰. 基于边缘帧差和高斯混合模型的行人目标检测[J]. 计算机应用研究, 2018, 35(4): 292-295.
- [14] 朱文杰, 王广龙, 田杰, 等. 空时自适应混合高斯模型复杂背景运动目标检测[J]. 北京理工大学学报, 2018, 38(2): 165-172.
- [15] 段锁林, 严翔, 朱方, 等. 基于改进混合高斯模型与六帧差分的目标检测算法[J]. 计算机工程, 2017, 43(7): 234-238.
- [16] 李睿, 盛超. 基于熵和相关接近度的混合高斯目标检测算法[J]. 计算机科学, 2017, 44(12): 304-309.
- [17] KE Z X, JIANG H H, ZHANG C L. Moving objects detection based on three frames differencing and GMM[J]. Advanced Materials Research, 2013, 694-697: 1974-1977.
- [18] XIE S N, TU Z W. Holistically-nested edge detection[J]. International Journal of Computer Vision, 2017, 125(1/3): 3-18.
- [19] 焦安波, 何淼, 罗海波. 一种改进的 HED 网络及其在边缘检测中的应用[J]. 红外技术, 2019, 41(1): 72-77.
- [20] 孙力. 基于 ROI 与自适应 OSTU 相结合的图像分割算法[J]. 现代电子技术, 2011, 34(6): 1-3.
- [21] WANG J B, XIONG H D, JIANG N, et al. An improved flocc image segmentation algorithm based on Otsu and particle swarm optimisation[J]. International Journal of Computational Science and Engineering, 2017, 15(1/2): 49-56.
- [22] 郭永芳, 于明, 黄凯. 基于细菌趋药性的 OSTU 双阈值图像分割算法[J]. 计算机工程, 2011, 37(22): 8-11.
- [23] 宋森森, 贾振红, 杨杰, 等. 结合 OSTU 阈值法的最小生成树图像分割算法[J]. 计算机工程与应用, 2019, 55 (9): 178-183.