

基于混合高斯模型的运动目标检测

王 健, 王晓东*, 郭 磊

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 针对运动目标检测中的背景复杂度高、视频数据计算量大等问题, 且为避免计算不同复杂程度的视频背景, 并能够准确地获取所需要的运动目标, 提出了一种基于混合高斯模型的运动目标检测方法. 首先采用混合高斯模型获取运动目标特征; 然后利用中值滤波方法去除视频目标运动特征中的背景噪声; 最后依据形态学运算方法对通过统计直方图得到的运动显著图进行处理, 从而获取最终的运动目标. 对标准视频序列集的检测表明, 利用该算法获取的运动目标不仅能抑制背景噪声, 而且精准度和误差都优于普通的视频运动目标检测算法.

关键词: 运动目标; 混合高斯模型; 中值滤波; 全局显著性

中图分类号: TP301.6

文献标志码: A

文章编号: 1001-5132 (2019) 03-0035-05

运动目标检测的精确性直接影响随后识别效果的好坏. 国内外众多对运动目标检测的算法都是针对一幅图像背景, 在摄像头固定的情况下进行检测. 而视频运动目标检测的处理方式复杂, 需要对多组连续图像进行处理, 背景复杂度高, 同时巨大的视频数据量也会带来更高的计算复杂度, 对此很多学者进行了深入的研究. 在众多的算法中, 有因为前期处理图像的算法考虑不够全面, 导致目标检测不理想, 如陆军等^[1]根据最小绝对差和匹配准则及三步搜索法对当前图像块做处理, 再对背景图像块进行运动补偿抵消, 因背景的变化, 图像块在相邻帧之间的匹配不理想. 同样, 在结合帧差法和背景减法的基础上, 熊卫华等^[2]有效地消除了光线突变和阴影, 但对灰度值接近背景的前景目标检测效果不理想. 除此之外, 视频中的光照变化也是一大影响因素, 如谢之宇等^[3]改进和优化了传统帧差法, 但是没有解决突变光照的影响. 刘静等^[4]虽然对背景和目标区域采用了不同的更新策略, 提高了模型的收敛速度, 但同样没有消除突变光照对模型的影响. 在多场景处理问题上, Xia等^[5]对运动的车辆进行检测, 通过融合最大期望算法(expectation Maximization Algorithm, EM)和高

斯混合模型算法(Gaussian Mixed Model, GMM)得到运动目标, 并加以填充空洞部分, 最后利用颜色等空间特征解决遮挡物等问题. 但算法只针对单一场景, 不具备普遍性. 而光流法^[6]和帧差法^[7]的效率低下, 单高斯模型算法实时处理效果差, 最终结果不理想.

大多数的视频背景都是移动的^[8], 而混合高斯模型可以针对背景变化的视频加以处理. 虽然利用核密度估计^[9]也可以对视频进行运动目标检测, 但是目标模型粗略, 计算量大, 对计算机的配置要求较高. 黄东军等^[10]通过将像素的领域特性与中值滤波相结合, 改进了混合高斯算法, 使其有较好的可适应性和鲁棒性. GMM 算法在有背景复杂干扰的情况下, 要准确检测出运动目标, 对背景的变化处理尤为重要^[11].

本文主要采用混合高斯模型算法对运动目标进行处理, 并利用图像的显著性能提升对运动目标检测的准确性, 并与多种同类算法进行比较, 以检测结果是否具有优越性.

1 基于混合高斯模型的运动目标检测

本文提出了基于混合高斯模型背景建模的运

收稿日期: 2018-06-01.

宁波大学学报(理工版)网址: <http://journal.g.nbu.edu.cn/>

基金项目: 国家科技支撑计划(2012BAH67F00).

第一作者: 王健(1993-), 男, 安徽肥东人, 在读硕士研究生, 主要研究方向: 图像图形处理. E-mail: 1668254476@qq.com

*通信作者: 王晓东(1970-), 男, 浙江上虞人, 教授, 主要研究方向: 视频图像处理. E-mail: wangxiaodong@nbu.edu.cn

动目标检测方法,其算法流程如图1所示.具体过程为:首先采用混合高斯模型得到运动目标特征,针对得到的图像用中值滤波方法去除视频目标运动特征中的背景噪声;然后利用形态学运算方法对根据统计直方图得到的运动显著图进行处理.

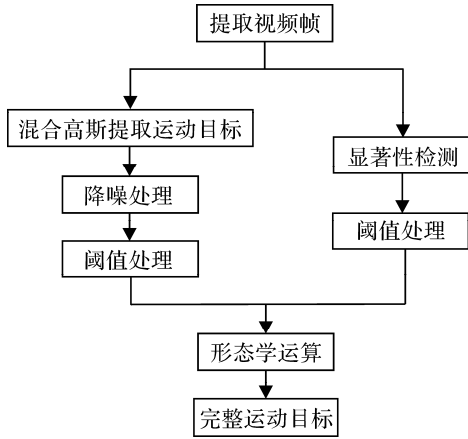


图1 本文算法流程

1.1 基于混合高斯模型的目标特征提取

混合高斯模型法通常采用视频序列图像的第一帧来构建背景模型,但是如果第一帧图像中存在运动物体时,很大程度上会导致确定的背景模型不准确和检测结果错误.为了避免上述问题,我们提取前 N 帧图像,并将这 N 帧图像相同位置上的像素点求和,取平均;然后将均值赋给初始背景对应的像素点,完成对初始背景的确定.

首先将一段视频进行分帧处理,并将其以时间序列 $\{X_1, X_2, \dots, X_t\}$ 表示,对其中 t 时刻那一帧图像,用3~5个高斯分布的高斯密度函数表示其像素点特征,计算公式为:

$$P(X_t) = \sum_{i=1}^k w_{i,t} \cdot \eta(X_t, \mu_{i,t}, \sum_{i,t}), \quad (1)$$

$$\eta(X_t, \mu_{i,t}, \sum_{i,t}) = \frac{\exp[-1/2(X_t - \mu_{i,t})^T \sum_{i,t}^{-1} (X_t - \mu_{i,t})]}{2\pi^{1/2} |\sum_{i,t}|^{1/2}}, \quad (2)$$

式中: $\eta(X_t, \mu_{i,t}, \sum_{i,t})$ 为高斯概率密度函数; $w_{i,t}$, $\mu_{i,t}$ 和 $\sum_{i,t}$ 分别为该时刻的权重、均值和协方差矩阵; $\sum_{i,t} = \sigma_{i,t}^2$, $\sigma_{i,t}^2$ 为方差; 本文取 $k=3$.

这时所提算法将当前帧的像素 X_t 按 $X_t - \mu_{i,t-1} \leq$

$2.5\sigma_{i,t-1}$ 匹配3个高斯分布,其表示形式为:

$$J_{GM} = \begin{cases} 1, & |X_t - \mu_{i,t-1}| \leq 2.5\sigma_{i,t-1}, \\ 0, & |X_t - \mu_{i,t-1}| > 2.5\sigma_{i,t-1}, \end{cases} \quad (3)$$

式中: J_{GM} 表示匹配结果,若为1,表示匹配成功;为0,表示匹配不成功.

当被检测的像素与式(3)范围一致时,则自动分配到该背景模型中.定义权值为:

$$w_{i,t} = (1 - \alpha)w_{i,t-1} + \alpha M_{i,t}, \quad (4)$$

式中: α 是对背景的学习率,为加速鬼影的消除,经多次实验,学习率选取0.01. 如果计算结果一致,即 $J_{GM} = 1$, $M_{i,t} = 1$, 相反 $M_{i,t} = 0$. 从中可看出,像素值的均值和方差是否更新由权重所决定.

$$\mu_t = (1 - \beta)\mu_{t-1} + \beta X_t, \quad (5)$$

$$\sigma_t^2 = (1 - \beta)\sigma_{t-1}^2 + \beta(X_t - \mu_t)^T (X_t - \mu_t), \quad (6)$$

式中: $\beta = \alpha / w_{i,t}$ 表示更新速率.

最后进行归一化处理:

$$w_{i,t} = w_{i,t} / \sum_{j=1}^k w_{j,t}, \quad (7)$$

背景图像可表征为:

$$B = \arg \min_n \left(\sum_{k=1}^n w_{i,k} > T \right), \quad (8)$$

$$X_{bg} = \sum_{i=1}^B w_{i,t} \mu_{i,t}, \quad (9)$$

式中: $T=0.25$; X_{bg} 为背景图像.

通过上述处理,得到的结果如图2所示.其中,图2(a)为视频帧的第305帧,即显示的为当前帧;图2(b)为背景帧;图2(c)是通过混合高斯模型得到的结果.对结果进行初步检测表明,虽能大致检测到运动目标,但整体轮廓稍有缺失,噪声点较多.

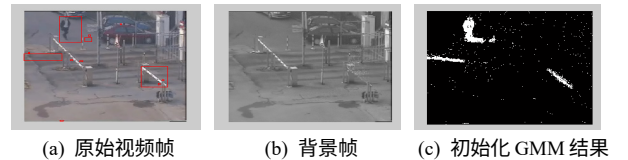


图2 基于混合高斯模型提取得到的运动目标

1.2 背景噪声处理

去噪声处理中,常用的有利用二维高斯卷积、均值滤波器或自适应维纳滤波器等手段.经过综合比较,本文选用了中值滤波法^[8]去噪.

中值滤波法是通过非线性处理,根据图像局

部区域的像素, 按灰度值的大小, 取该区域像素灰度中值作灰度值, 通过邻域运算, 得到最终的像素值. 该方法能较好地保护边界.

以 1.1 节的结果图为例, 分别使用 3×3 , 5×5 , 7×7 , 9×9 模板进行中值滤波^[10], 结果如图 3 所示.

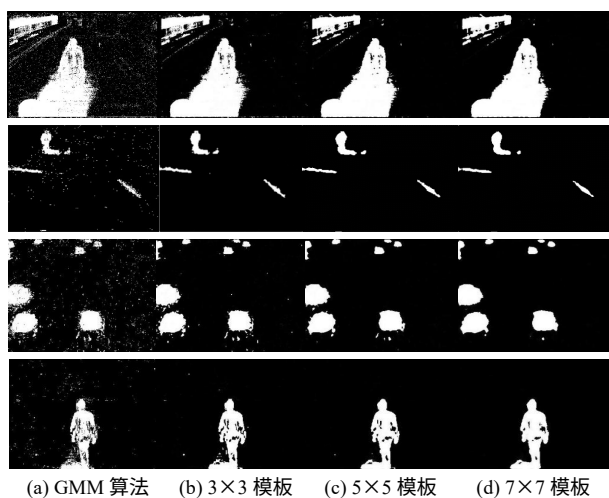


图 3 各模板中值滤波结果

从图 3 对多个标准视频进行实验结果比较可以发现, 3×3 模板(图 3(b))对噪声处理效果较差, 噪点不能完全去除; 而 7×7 模板(图 3(d))对于运动目标边缘部分同化严重, 细节有损. 为更好地去噪, 我们选用了 5×5 模板(图 3(c)).

1.3 多特征处理

中值滤波处理噪声时也对部分前景目标作处理, 使得到的目标特征不完整, 需要提取运动目标的全局显著性^[13]. 因此, 采用了直方图统计方法进行逐帧处理, 根据直方图对比, 利用像素间的色彩差异得到所需的显著图^[14]. 同样, 对随机视频帧的第 54 帧和第 426 帧进行处理, 得到基于直方图统计衡量运动特征的全局显著(图 4).



图 4 全局对比显著图

图 4 中因为背景较为复杂, 利用普通显著性算法得到的结果不理想. 考虑到多个空间特征因素, 显著性算法对图像的部分背景也识别为人眼关注部分, 利用这一点可弥补 GMM 算法的完整性.

基于图 4, 我们对得到的图像先进行腐蚀处理, 再进行膨胀处理后与全局显著图进行相与处理, 结果发现, 在图像空洞不多的情况下, 有一定的弥补作用. 得到最终的运动目标显著图如图 5 所示.

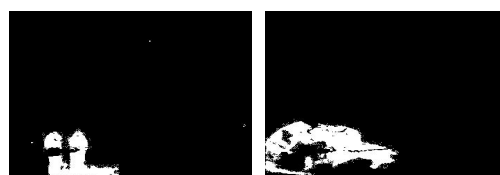
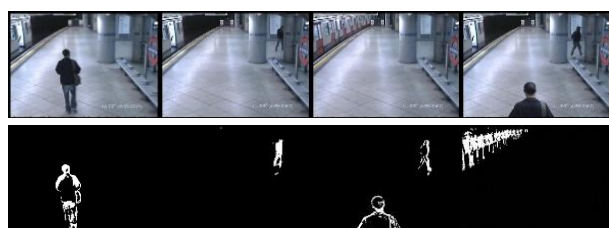


图 5 形态学运算处理后图像

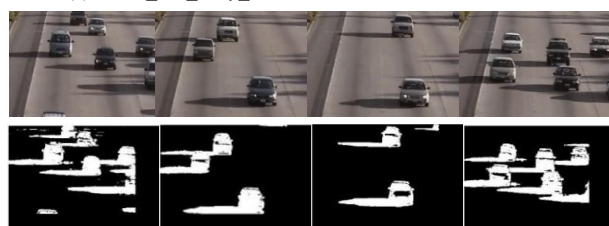
2 实验结果与分析

为验证本文算法的有效性, 使用比较典型的视频进行测试. 所选视频背景复杂, 有车辆和多个行人行动, 行人与车辆皆为被检测的运动目标. 实验使用 matlab 2014a 进行处理, 通过视频帧得到的结果与数据库中配套的 Ground Truth 图进行对比^[15], 以验证本文算法的有效性. 对多个视频序列进行测试^[16], 包括 highwayI、highwayII、Laboratory、AVSS_AB_Easy_Divx、campus 等, 结果如图 6 所示. 图 6 中白色部分为本文算法需要检测的运动目标, 黑色为背景, 通过二值化方法更直观地表现. 其中, 图 6(a)为 AVSS_AB_Easy_Divx 序列随机挑选的某几帧结果, 该视频包含高速移动的列车和走动的人流, 由此可见本文算法针对行人有很好的检测效果, 而针对列车高速移动, 因其阴影较大, 阴影的检测效果不明显. 图 6(b)和图 6(c)是 highwayI 和 highwayII 序列随机挑选的某几帧分别在有阴影和无阴影下的检测结果. 从图 6(b)和图 6(c)看, 对于运动目标较小, 无论是否有阴影, 运动中的车辆都有很好的检测结果(轮廓完整清晰). 图 6(d)是 Laboratory 序列随机挑选的某几帧结果, 因人在行动中有遮挡物, 而检测结果可以只检测有运动的部分, 被遮挡部分即无运动目标, 因此虽然轮廓不完整, 结果却是准确的. 图 6(e)是 campus 序列随机挑选的某几帧结果, 包含较多的行人, 背景复杂, 但同样有很好的检测结果, 能较好地解决背景问

题被误检的情况.



(a) AVSS_AB_Easy_Divx 序列随机挑选帧原图及结果



(b) highway I 序列随机挑选帧原图及结果



(c) highway II 序列随机挑选帧原图及结果



(d) Laboratory 序列随机挑选帧原图及结果



(e) campus 序列随机挑选帧原图及结果

图 6 多个序列运动目标检测结果

本文通过识别率 D 和误检率 $F^{[17]}$ 来进行验证, 计算公式为:

$$D = P_1 / (P_1 + P_2), \quad F = P_2 / (P_1 + P_2), \quad (10)$$

其中, P_1 为检测结果; P_2 为真图与结果不符部分.

以 campus 序列为例, 在视频帧数和其余条件均相同的情况下, 与帧差法^[18]、四帧差法^[19]进行比较, 结果见表 1.

表 1 不同算法的误检率和识别率

算法	识别率	误检率
GMM	0.642	0.124
帧差法	0.788	0.098
四帧差法	0.876	0.084
本文算法	0.894	0.082

从表 1 结果可以看出, 本文算法在针对复杂的背景图像上不会因为背景变化等原因而降低运动目标的精准率, 并且能完整地将运动目标检测出来. 识别率比其他同类算法高(比混合高斯模型增加 39.3%), 同时降低了算法的误检率. 帧差法^[18]在检测运动目标时效果不理想, 易出现“空洞”和“重影”现象. 四帧差法^[19]是在帧差法的基础上改进而来, 目标运动时无法检测出完整轮廓. 结合实验结果图像和实验数据可知, 本文算法能有效地解决此类现象, 可以较好地检测出轮廓完整的运动目标.

利用 highwayI 和 AVSS_AB_Easy_Divx 两个具有代表性的视频, 将本文算法与 GMM 算法和核密度算法的运行时间进行比较, 结果见表 2.

表 2 不同算法运行时间比较 ms

算法	highwayI			AVSS_AB_Easy_Divx		
	第 304 帧	第 515 帧	第 671 帧	第 304 帧	第 515 帧	第 671 帧
GMM	1 055	1 049	1 048	660	642	670
核密度法	1 005	998	995	510	505	501
本文算法	845	833	837	471	459	465

从表 2 可知, highwayI 的运动目标较多, 算法运算时间相对较慢. AVSS_AB_Easy_Divx 视频中运动目标少, 运算时间相对较少. 对比 GMM 和核密度算法, 本文算法的运行效率优于上述 2 种算法.

3 结论

为获得视频中理想的运动目标, 文中以混合高斯模型获得初步运动目标, 获取显著性目标后经过形态学运算获得最终结果. 通过图 6 可以看出, 该算法对于背景是否移动、运动速度的快慢以及目标是否单一等情况均能很好地适应. 同时该方法能很好地抑制背景噪声, 但是对于分辨率较高的视频, 实时性上还需要进一步优化. 综上所述, 基于混合高斯模型的运动目标检测对提取运动目标具有很好的效果, 不仅能够抑制背景噪声,

同时在误差率和精准率上具有较好的表现。

参考文献:

- [1] 陆军, 李凤玲, 姜迈. 摄像机运动下的动态目标检测与跟踪[J]. 哈尔滨工程大学学报, 2008, 29(8):831-835.
- [2] 熊卫华, 向磊, 李俊峰, 等. 背景减除与帧间差分相结合的运动目标检测方法[EB/OL]. [2018-10-12]. <http://cpfd.cnki.com.cn/Article/CPFDTOTAL-KZLL201107001151.htm>.
- [3] 谢之宇, 蒋晓瑜, 汪熙, 等. 改进的帧差法在目标匹配中的应用[J]. 计算机工程与应用, 2011, 47(34):208-211.
- [4] 刘静, 王玲. 混合高斯模型背景法的一种改进算法[J]. 计算机工程与应用, 2010, 46(13):168-170.
- [5] Xia Y, Shi X, Song G, et al. Towards improving quality of video-based vehicle counting method for traffic flow estimation[J]. Signal Processing, 2016, 120:672-681.
- [6] 桂本桦, 钱徽, 朱森良. 一种优化梯度计算的改进HS光流算法[J]. 中国图象图形学报, 2005, 10(8):1052-1058.
- [7] Meyer D, Denzler J, Niemann H. Model based extraction of articulated objects in image sequences for gait analysis[C]//Proceedings of IEEE International Conference on Image Processing, Santa Barbara, 1997:78-81.
- [8] 汪巍. 动态场景下运动目标检测与跟踪算法的研究[D]. 合肥: 合肥工业大学, 2012.
- [9] 陈栋喜. 基于改进自适应混合高斯模型的背景建模及对象计数方法研究[D]. 南京: 南京航空航天大学, 2013.
- [10] 黄东军, 杨颖华. 基于改进混合高斯模型的运动物体检测研究[J]. 计算机应用研究, 2017, 34(6):1862-1866.
- [11] 孙炜烨. 动态场景下运动目标检测算法的研究[D]. 天津: 河北工业大学, 2012.
- [12] 高智勇, 唐文峰, 贺良杰. 基于运动显著性的移动镜头下的运动目标检测[J]. 计算机应用, 2016, 36(6):1692-1698.
- [13] 李晓悦. 基于改进 CV 模型的工业 CT 图像分割算法研究[D]. 大连: 大连理工大学, 2014.
- [14] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection[EB/OL]. [2018-10-12]. http://vecg.cs.ucl.ac.uk/Projects/SmartGeometry/contrast_saliency/paper_docs/contrast_saliency_small_cvpr11.pdf.
- [15] 高秀丽. 融合深度信息的视觉注意模型研究及其应用[D]. 杭州: 杭州电子科技大学, 2013.
- [16] 胡燕, 王慧琴, 马宗方. 改进混合高斯模型的自适应烟雾图像分割算法[J]. 计算机辅助设计与图形学学报, 2016, 28(7):1138-1145.
- [17] 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.
- [18] Choi Y J, Lee J S, Cho W D. Robust extraction of moving objects based on hue and hue gradient[C]//Universal Access in Human-computer Interaction, Ambient Interaction. Berlin: Springer, 2007:784-791.
- [19] 郭伟, 高媛媛, 刘鑫焱. 改进的基于混合高斯模型的运动目标检测算法[J]. 计算机工程与应用, 2016, 52(13):195-200.

Moving target detection based on Gaussian mixture model and salience

WANG Jian, WANG Xiao-dong*, GUO Lei

(Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China)

Abstract: Aiming at tackling the problems of high background complexity and large amount of video data in moving target detection, and in order to avoid computing the background of different complexity, a moving target detection method based on Gaussian mixture model is proposed. First, a Gaussian mixture model is used to acquire the features of the moving target, and then median filtering is applied to remove the background noise in the motion feature of the video target. Finally, by taking advantage of the morphological operation method, the motion image obtained by the statistical histogram is processed to obtain the final moving target. Tested on the standard video sequence set, the moving target obtained using the proposed algorithm can not only suppress the background noise, but also be better than the ordinary video moving target detection algorithm both in the accuracy improvement and error attenuation.

Key words: moving target; Gaussian mixed model; median filter; global saliency

(责任编辑 史小丽)