

中图法分类号: TP391 文献标志码: A 文章编号: 1006-8961(2011)06-1002-06

论文索引信息: 陈明生, 梁光明, 孙即祥, 刘东华, 赵键. 利用时空背景模型的快速运动目标检测方法 [J]. 中国图象图形学报, 2011, 16(6): 1002-1007

利用时空背景模型的快速运动目标检测方法

陈明生, 梁光明, 孙即祥, 刘东华, 赵键

(国防科技大学电子科学与工程学院, 长沙 410073)

摘要: 为了弥补运动目标检测中传统混合高斯背景模型仅对单个像素建模、运算耗时的不足, 通过提取背景时间统计特征和空间区域特征建立模型, 针对模型中的高斯分量采用一种改进的分量个数自适应算法, 并在此模型基础上, 提出一种自适应迭代分块目标检测方法。通过包含区域信息的背景模型检测目标, 减少在同一背景区域中目标的误判和漏判。将自适应迭代分块检测算法与背景的区域信息结合, 可以在不降低检测精度的前提下大大提高算法执行速度。实验结果表明, 相对于传统算法, 本文检测法检测结果信噪比更高, 目标更加完整, 运行速度平均提高了22%。

关键词: 混合高斯; 背景区域; 自适应; 分块检测

Fast moving object detection method using temporal-spatial background model

Chen Mingsheng, Liang Guangming, Sun Jixiang, Liu Donghua, Zhao Jian

(Electronic Engineering and science technology college, national university of defense technology, changsha 410073 China)

Abstract: Moving objects extraction is a key part of video surveillance system. To improve the performance of moving objects detection method based on the Gaussian Mixture Model (GMM), an iterative detection algorithm with adaptive partitioning block of pixels is proposed. It is based on the temporal-spatial background that the number of components is improved adaptively and the feature of areas extracted spatially is combined. With the spatial areas information, the algorithm decreases the number of small fake objects and reduces the fragmentation of objects that caused by all kinds of noise. Comparing with detection method based on single pixel, the proposed method would not almost impact the detected results when it reduces the algorithm computation obviously. The results show that the objects extracted by the proposed method with higher SNR and the processing time decreases 22% contrasting to traditional algorithm.

Keywords: GMM; areas of background; adaptive; partition block detection

0 引言

运动目标检测是视频处理的关键技术。背景差法是其中研究较为广泛的一种方法^[1], 其关键是背景的构建。Wren 等人研究像素灰度值在时间轴上的统计特性提出了单高斯的背景模型^[2], 对每一个像素建立高斯模型来表征像素在时间轴上的变化。

Stauffer 和 Grimson 在单高斯模型基础上, 针对复杂场景中像素统计特征的多峰情况, 提出对单个像素的灰度值采用混合高斯建模^[3]。混合高斯模型以其对多模态的良好描述能力, 较好地克服了背景中的扰动因素(水面的波纹、摇晃的树枝等)给运动目标检测带来的困扰。当在颜色空间前景和背景差异较小时, 背景差法易产生较高的漏检率致使完全漏检目标或检测的目标不完整, Zhang 等人通过前几

收稿日期: 2010-03-26; 修回日期: 2010-06-22

第一作者简介: 陈明生(1983—), 男, 博士研究生。主要研究方向为图像处理、视频处理、模式识别。E-mail: ms-chen20@163.com。

帧的检测结果得到当前时刻的预测目标,融合该预测目标和通过背景差法检测出的目标进行综合判决,得到最终的检测结果^[4]。杨涛等人在混合高斯背景模型基础上,引入一个由多个独立单高斯模型组成的参考背景用来检测静止目标^[5]。

基于像素的混合高斯背景模型目标检测算法及改进算法^[6]都是以像素为单元进行处理,仅仅分析利用了单个像素在时间轴上的统计特性,而没有考虑像素之间的关联信息,影响了检测的精度。针对这一问题,朱碧婷等人在利用背景差法判决当前帧的前景目标时,融合当前像素点周围的像素特征综合判决,提高了目标检测的鲁棒性^[7]。Dickinson 等人提出在5维空间 $[x, y, Y, U, V]$ 中进行目标的检测^[8],该算法虽然结合了像素的空间位置信息,但在5维空间中进行检测复杂度、运算量大,不利于算法的实际应用,特别是目前许多视频应用都在朝高分辨率发展,这一矛盾更为突出。

本文针对以下两个问题研究分析已有的基于混合高斯背景的目标检测算法:

1) 检测到的目标容易出现破碎现象^[9]——目标内部出现不规则的非目标空洞或者目标由一些不连通的部分组成,影响了目标的完整性,不利于后续对这些目标提取特征。而在没有目标的区域容易出现碎片目标,引起误判。

2) 在进行处理分析时,对包含目标的区域和很少出现目标的背景区域进行同样的处理——遍历每一个像素,并作判决更新处理,使得算法占用较多的计算时间,尤其是在高分辨率画面下距离实时相差较远。

提出以下方法改进:1) 在传统背景模型基础上,提取背景的空间结构信息,对背景中的区域建模,使得背景模型中同一区域内的像素不再是相互孤立地进行判决,避免了出现在一个背景区域内的同一个目标一部分被判为前景一部分,被误判为背景造成的目标的不完整。2) 研究分析已有的高斯分量数目自适应算法,同时利用对背景已经提取出的区域信息,在不包含目标的区域内,采用基于像素块的检测方法,提高了检测速度。

1 融合空间信息的混合高斯背景模型

1.1 时空混合高斯背景模型表示

在 Stauffer 的混合高斯背景模型中,对每个像素

建立含 K 个分量的混合高斯模型,对宽 W 高 H 的图像序列,实际用了 $W \times H \times K$ 个分量来表征背景模型。实际场景中背景是由一些物体表面所组成,同一个表面所成像的相邻像素点在图像上构成区域,这些像素点的时域和空域特性都有很强的相似性,表征这些模态的分量在参数空间具有聚类特征。对这些模态在参数空间利用 K 均值算法进行聚类,对聚类后属于同一个类别的模态用一个高斯分量描述,简化背景模型的同时反映了背景的区域结构特征,实现对背景的像素级和区域级混合建模,在检测时则可以避免对同一区域内的同质像素(同为背景或前景)做出不同的判决。

相邻像素的混合高斯模型分量(包含 K 个分量)经过聚类后,可以表示为

$$p(\mathbf{X}_t) = \sum_{j=1}^{K_1} \omega_j^R \eta(\mathbf{X}_t; \theta_j^R) + \sum_{j=K_1+1}^K \omega_j \eta(\mathbf{X}_t; \theta_j) \quad (1)$$

式中, \mathbf{X}_t 为具有 R、G、B 3 个颜色分量的矢量, $\eta(\mathbf{X}; \theta_j)$ 为第 j 个分量服从的高斯分布, $\theta_j = [\boldsymbol{\mu}_j, \sigma_j^2]$, $\theta_j^R = [\boldsymbol{\mu}_j^R, (\sigma^R)_j^2]$, $\boldsymbol{\mu}_j, \boldsymbol{\mu}_j^R$ 为高斯模型的均值,方差为 $\sigma_j^2 \mathbf{I}$ (或 $(\sigma^R)_j^2 \mathbf{I}$, \mathbf{I} 为单位矩阵,假设 R、G、B 3 个颜色分量相互独立且具有相同方差), ω_j^R, ω_j 均表示高斯分量的权重,且 $\sum_{j=1}^{K_1} \omega_j^R + \sum_{j=K_1+1}^K \omega_j = 1$ 。

式(1)的等号右侧第1部分(即包含上标 R 的分量)表示对 \mathbf{X}_t 所属背景区域建立的模型,是 \mathbf{X}_t 与相邻像素的聚类结果,为 \mathbf{X}_t 与相邻像素共有的分量,称其为背景模型的区域分量;第2部分表示对应 \mathbf{X}_t 独立于相邻像素变化建立的分量。在多数情况下 K_1 为1或0,等于0表示该像素位置不构成区域。

1.2 新背景模型的在线参数更新

KaewTraKulPong 采用改进的在线 EM 算法估计混合高斯模型的参数^[10],主要针对原始 GMM 模型的更新率参数 α 进行了改进:

$$\alpha = \begin{cases} 1/n & n < N \\ N & n \geq N \end{cases} \quad (2)$$

式中, n 为模型的更新帧数, N 为窗口宽度(为常数)。其在模型初始阶段的动态更新率提高了模型的学习速度和缩短了检测系统的反应时间。但它是针对基于单像素处理的运动目标检测算法而提出的,由于本文在时域上对背景建模外,还需要对背景在空间上进行区域聚类,所以该参数更新方法在这里不能完全适用。本文在文献[10]更新方法基础

上,利用在线连通区域检测方法^[11]判断区域,在线 K 均值算法更新区域,实现对背景空间区域特征的建模。本文的背景更新方法如下。

1) 像素 X_i 被判决为前景,将其作为新的高斯分量,赋予初始权重和方差加入模型中。

2) 像素 X_i 被判决为背景,设 D_i 为与 X_i 连通的区域, D_i 当前包含 N 个元素,表示为 D_i^N , μ_i^N 、 \sum_i^N 分别为区域 D_i^N 的均值和方差, $\sum_i^N = (\sigma_i^N)^2 I$ 。

如果 $\|X_i - \mu_i^N\| < 2.5\sigma_i^N$,则以 X_i 对区域 D_i 更新:

$$\mu_i^{N+1} = (1 - \alpha)\mu_i^N + \alpha X_i \quad (3)$$

$$(\sigma_i^{N+1})^2 = (1 - \alpha)(\sigma_i^N)^2 + \alpha \|X_i - \mu_i^{N+1}\|^2 \quad (4)$$

式中, $\alpha = 1/(N+1)$ 。

如果不存在与 X_i 连通的背景区域或 X_i 没有被判决属于任何一个与其连通的背景区域,则仍采用文献[10]的方法更新。

2 高斯分量个数的自适应选取

实际背景中各个像素的模态数目并非都相等,较稳定的区域可能是单模态,用一个高斯分量即可建模,有扰动的区域则需要多个分量来建模,可以通过自适应方法减少冗余的高斯分量以减少计算量。

Suo 等人提出一种利用分量权重之间的梯度来减少高斯分量的方法^[12]。在经过排序的高斯分量队列中,删除权重与前一个分量相差小于阈值的分量,减少分量数目。如果两个模态近似等概率交替出现,这两个模态对应的高斯分量权重将会较大且接近,对应的梯度较小可能会遭到误删。同时新加入背景的分量在队列尾部,权重较小从而与前一分量权重差值也较小,容易遭到误删,不利于背景的更新。在研究文献[12]算法基础上,在背景更新过程中判断已经排序的分量权重队列中倒数第 2 个分量的权重,如果该权重小于阈值 T_ω 则删除(T_ω 小于分量的权重初始值),最后一个是最新加入的分量暂时予以保留。 K 个高斯模型表示为

$$p(X_i) = \sum_{j=1}^{K_2} \omega_j \eta(X_i; \theta_j) + \sum_{j=K_2+1}^{K-1} \omega_j \eta(X_i; \theta_j) + \omega_K \eta(X_i; \theta_K) \quad (5)$$

式中,等号右侧第 1 部分表示稳定的背景(K_2 大于

或等于式(1)中 K_1),中间部分表示前一时间段内不同的前景、噪声所产生的高斯分量,权重较小,最后一个表示最新加入到模型的高斯分量(可能是前景,也可能是发生变化的背景)。在对每一帧的迭代更新过程中,删除权重小于阈值 T_ω 的高斯分量,背景模型简化为 K_2+1 个分量:

$$p(X_i) = \sum_{j=1}^{K_2} \omega_j \eta(X_i; \theta_j) + \omega_{K_2+1} \eta(X_i; \theta_{K_2+1}) \quad (6)$$

在删除权重小于 T_ω 的分量时,模型中仅权重参数 ω_j 需要更新(即剩下分量的权重重新归一化)。

通过与式(5)比较可以发现,新的背景模型抑制了前景或噪声产生的高斯分量 $\sum_{j=K_2+1}^{K-1} \omega_j \eta(X_i; \theta_j)$ 。模型最大分量个数为 K ,阈值 T_ω 取 $1/K$ 。

3 自适应的迭代分块检测

在视频监控场景中,通常目标所占画面的比例较小,背景显露区域占比例较大^[13]。背景相对于前景,结构中很多区域,这些区域内部在像素颜色空间上具有平坦的特征。第 1 节已经提取出背景中的这些平坦区域,在线处理时如果利用背景区域内部的平坦性特征,在平坦区域内利用多个相邻像素构成的像素块进行检测,可以在保持检测器检测精度的前提下,显著提高检测速度。

本文检测方法的思路:根据领域的相关性,在前一像素判决属于背景中的区域时(图像的平坦部分),对于当前像素选取 4×4 的像素块(当前像素位于左上角),计算其样本均值并与背景模型中的区域分量(式(1)中的 $\eta(X_i; \theta_j^R)$)匹配(即 X_i 落在该高斯分量的置信区间,置信度通常取 98.76%),如匹配成功,则以该样本均值更新对应高斯分量,否则退化为 2×2 像素块进行处理,如仍不匹配则以传统的单像素模式判决。

算法处理步骤:

1) 判断处理的前一像素是否被判决属于模型中的区域分量,如不是,则转步骤 5)。

2) 计算 4×4 像素块的样本均值 $\overline{X_i^{16}}$,同时保留计算过程中 2×2 像素块的样本均值 $\overline{X_i^4}$ 。

3) 将 $\overline{X_i^{16}}$ 与所在区域 D_i 对应模型 $\eta(X_i; \theta_j^R)$ 匹配。如匹配,则该像素块对应的 16 个像素判决为背景,且参照式(3)(4)以 $\overline{X_i^{16}}$ 更新该区域模型 $\eta(X_i; \theta_j^R)$,

当前位置处理完毕,转至下一待处理像素。否则转步骤 4)。

4) 将 X_i^r 与区域 D_i 的对应模型 $\eta(X_i; \theta_j^r)$ 匹配。如匹配,该像素块的 4 个像素判决为背景,以式(3)(4)更新该模型,当前位置处理完毕,转至下一待处理像素。否则转步骤 5)。

5) 以传统的单像素模式处理,处理完毕转至下一待处理像素。

分块检测流程图如图 1 所示。

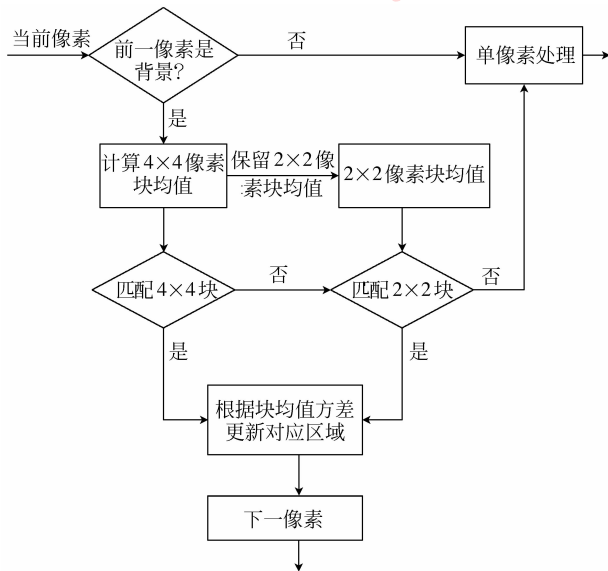


图 1 分块检测流程图

Fig. 1 Detection based on adaptive partitioning block of pixels

4 实验

实验采用 ImageLab 实验室提供的测试视频 Camera1_070605、video6_long 和 tunnelvideo^[14], 它们的大小分别为 384×288 、 384×240 和 720×576 , 实验系统在 VC++ 环境下基于 OpenCV 开发, 与本文算法进行对比的是经典混合高斯背景模型结合文献[12]的高斯分量个数自适应的算法(以下简称 GMM 改进算法), 本文算法与 GMM 改进算法的高斯分量个数最大值为 5(能满足绝大多数场景的建模), 初始方差随前景背景对比度的高低而不同, 在此取 30, 方差系数 2.5(对置信度为 98.76%), 初始权重 0.05, 窗口宽度^[7] 200 帧(随场景变化剧烈程度不同而不同), 即系统稳定之后的背景更新率为 0.005。

图 2—3 是对上述视频 Camera1_070605 和 tunnelvideo 在本文算法和 GMM 改进算法下的检测结果。从图可以看出, 本文算法相对于 GMM 改进算法, 检测结果噪声较少, 检测到的目标较为完整, 目标内部空洞减少, 有利用后续的进一步处理。而在图 2(e) 和图 3(e) 中, 前景区域明显超出实际的前景区域, 这是由于 GMM 改进算法中的分量个数自适应算法误删了图 2 实验中前几帧出现的前景和图 3 实验中光线变化在模型中产生的高斯分量, 从

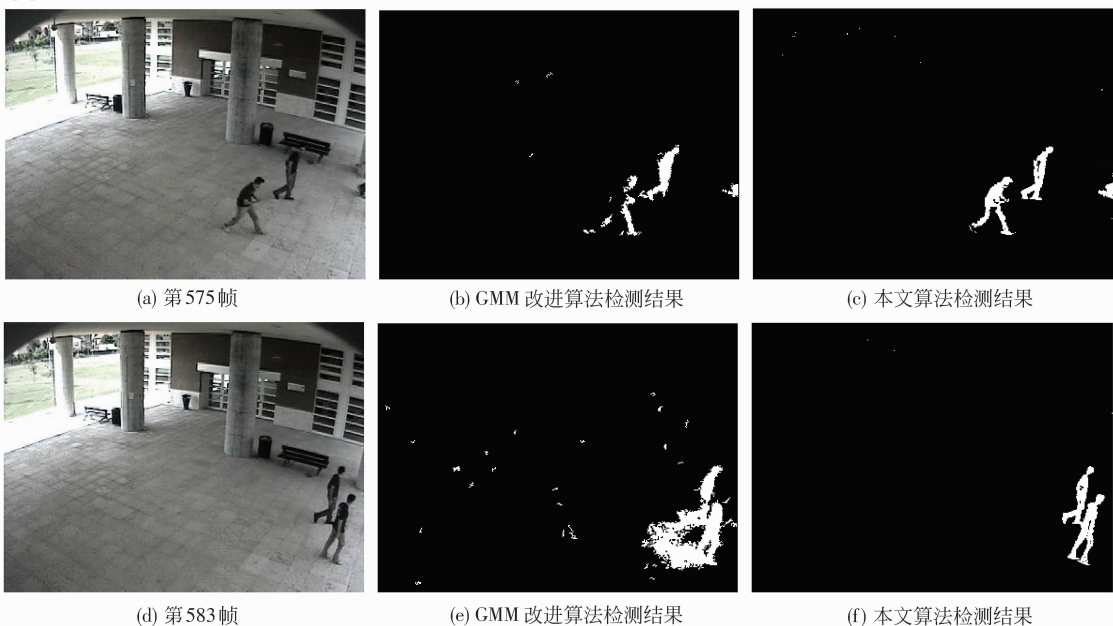


图 2 针对视频 Camera1_070605 的实验结果
Fig. 2 Results of experiment on video Camera1_070605

而产生了较多的误判,本文算法则避免了此种情况的产生。

图 4 是对视频 video6_long 的检测结果。针对该视频 ImageLab 提供了理想的检测结果 (video6_groundTruth_long), 利用该理想检测结果可以精确计算出待评价算法的虚警像素和漏警像素, 用来对不同的检测算法进行评价。本文实验的评价结果如图 5 所示。

图 6 是对上述视频进行目标检测, 算法所消耗

时间的统计结果。

由图 6 可以得出, 针对视频 video6_long、Camera1_070605 和 tunnelvideo, 算法运行时间分别下降了 21%、11% 和 32%。处理 Camera1_070605 的时间消耗降低不明显, 主要是由于该视频背景中区域较小而纹理较多, 算法更多的是在单像素模式下进行处理; 处理视频 tunnelvideo 的时间减少最多, 则是由于该视频尺寸最大且视频背景中区域的面积占了大部分, 利用本文算法显著减少了处理时间。

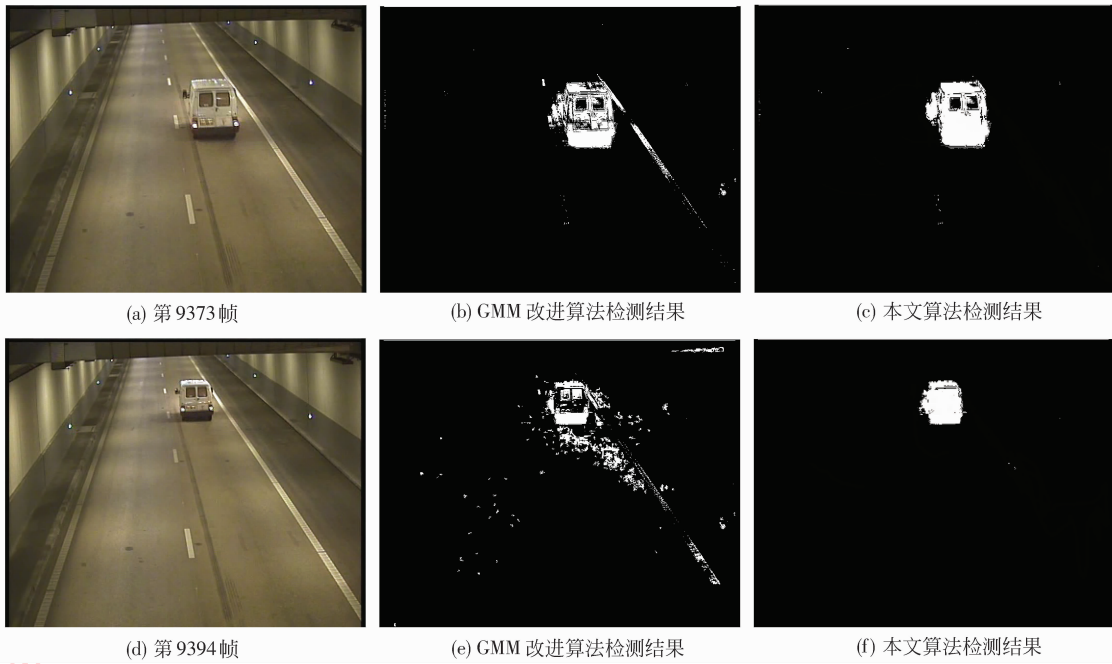


图 3 针对视频 tunnelvideo 的实验结果
Fig. 3 Results of experiment on video tunnelvideo

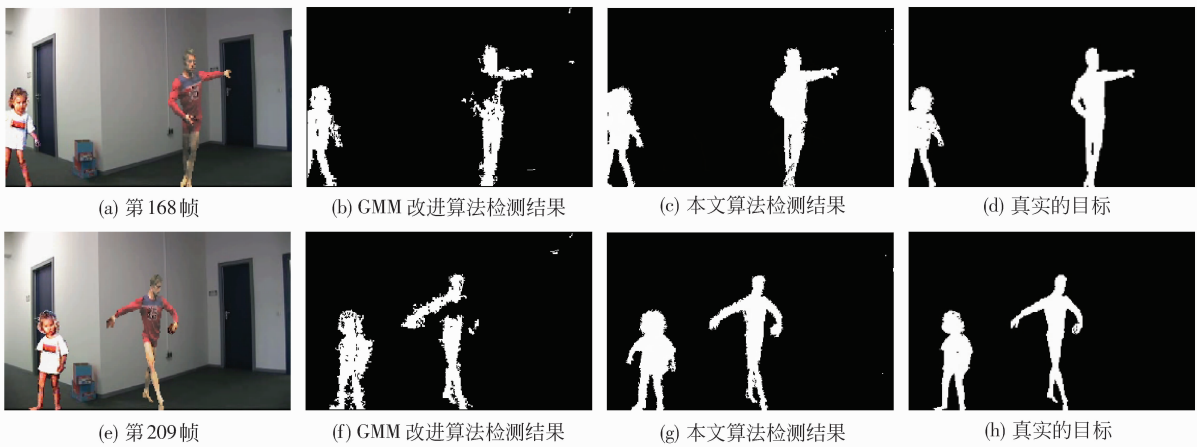


图 4 针对视频 video6_long 的实验结果
Fig. 4 Results of experiment on video video6_long

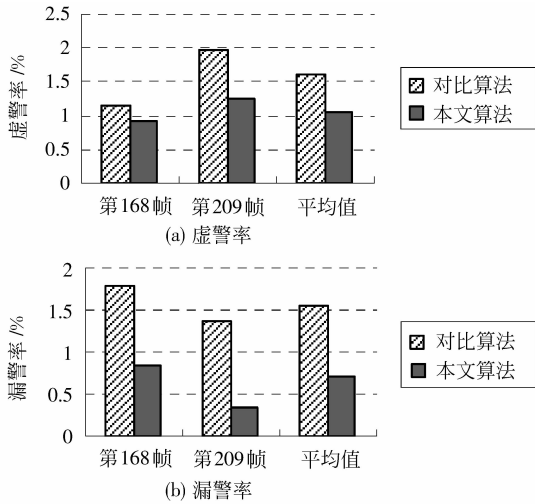


图 5 两种算法的对比统计结果
Fig. 5 Stat. of the two algorithms

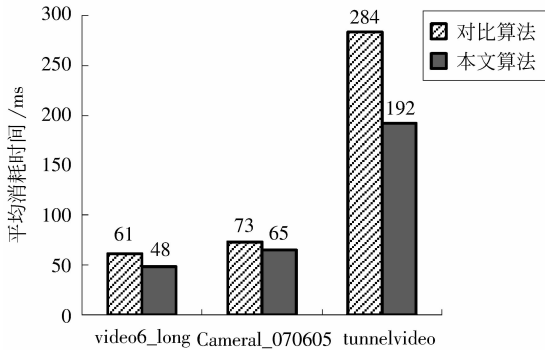


图 6 算法运行时间统计结果
Fig. 6 Stat. of the two algorithms' processing time

5 结 论

提出融合像素空间信息的混合高斯背景模型,该背景模型在经典单像素模型基础上通过区域聚类提取背景的空间区域信息以改善背景模型的性能。利用该背景模型中的区域信息减少了检测噪声,而且在该区域信息基础上实现了自适应的迭代分块检测,在保持目标检测精度的前提下提高了检测速度。同时提出一种混合高斯模型分量个数自适应的改进方法,以实现减少模型中冗余信息的目的。最后实验结果表明,本文算法相对传统算法,提高了检测目标的完整性和检测速度,说明空间信息能够显著增强目标检测算法的鲁棒性。

参考文献 (References)

[1] Gan Xinsheng, Zhao Shubin. Comparison on background

subtraction algorithms for moving target detection [J]. Command Control & Simulation, 2008, 30(3):45-50. [甘新胜, 赵书斌. 基于背景差的运动目标检测方法比较分析 [J]. 指挥控制与仿真, 2008, 30(3): 45-50.]

[2] 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.

[3] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking [C] // CVPR '99. Fort Collins, Colorado, USA: IEEE, 1999: 246-252.

[4] Zhang Xiang, Yang Jie. A novel algorithm to segment foreground from a similarly colored background [J]. International Journal of Electronics and Communications, 2009, 63(10): 831-840.

[5] Yang Tao, Li Jing, Pan Quan, et al. A multiple layer background model for foreground detection [J]. Journal of Image and Graphics, 2008, 13(7): 1303-1308. [杨涛, 李静, 潘泉, 等. 一种基于多层背景模型的前景检测算法 [J]. 中国图象图形学报, 2008, 13(7): 1303-1308.]

[6] Liu Xin, Liu Hui, Qiang Zhenping, et al. Adaptive background modeling based on mixture Gaussian model and frame subtraction [J]. Journal of Image and Graphics, 2008, 13(4): 729-734. [刘鑫, 刘辉, 强振平, 等. 混合高斯模型和帧间差分相融合的自适应背景模型 [J]. 中国图象图形学报, 2008, 13(4): 729-734.]

[7] Zhu Biting, Zheng Shibao. Space-domain background subtraction and shadow elimination based on Gaussian mixture model [J]. Journal of Image and Graphics, 2008, 13(10): 1906-1909. [朱碧婷, 郑世宝. 基于高斯混合模型的空间域背景分离法及阴影消除法 [J]. 中国图象图形学报, 2008, 13(10): 1906-1909.]

[8] Patrick Dickinson, Andrew Hunter, Kofi Appiah. A spatially distributed model for foreground segmentation [J]. Image and Vision Computing, 2009, 27(9): 1326-1335.

[9] Tian Yingli, Lu Max, Hampapur Arun. Robust and efficient foreground analysis for real-time video surveillance [C] // CVPR '05. San Diego, CA, USA: IEEE, 2005: 1182-1187.

[10] KaewTraKulPong P, Bowden R. An improved adaptive background mixture model for real-time tracking with shadow detection [C] // Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems. Kingston, UK: University of Kingston, 2001: 149-158.

[11] Kong Bin. A fast connected component analysis algorithm and its implementation [J]. PR & AI, 2003, 16(1): 110-115. [孔斌. 快速连通域分析算法及其实现 [J]. 模式识别与人工智能, 2003, 16(1): 110-115.]

[12] Suo Peng, Wang Yanjiang. An improved adaptive background modeling algorithm based on Gaussian mixture model [C] // ICSP2008. Leipsig, Germany: Springer-Verlag, 2008: 1436-1439.

[13] Yang Guanglin, Kong Lingfu. Approach of constructing background model based on image blocks [J]. ROBOT, 2007, 29(1): 29-34. [杨广林, 孔令富. 基于图像分块的背景模型构建方法 [J]. 机器人, 2007, 29(1): 29-34.]

[14] ImageLab. University of Modena and Reggio Emilia, Italia [EB/OL] [2010-01-05]. <http://imagelab.ing.unimore.it/imagelab/index.asp>.