

基于滑动窗的混合高斯模型运动目标检测方法

周建英 吴小培* 张超 吕钊

(安徽大学计算智能与信号处理教育部重点实验室 合肥 230039)

摘要: 在复杂场景下,传统混合高斯模型能较好地检测出运动目标,但随着时间的推移,模型参数收敛缓慢且难以适应场景中真实背景的实时变化,从而导致运动目标的错误检测率增加。该文利用滑动窗技术的短时历史记忆特性,提出一种新颖的基于滑动窗的混合高斯模型运动目标检测方法,该方法弥补了传统混合高斯背景模型不能及时形成新背景的缺点,提高了运动检测的完整性,并进一步降低了算法对场景光照变化的敏感性。多场景下的对比实验结果表明,该方法能更准确、完整地检测出运动目标并具有更好的环境适应性。

关键词: 运动目标检测; 滑动窗; 混合高斯模型; 背景模型

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2013)07-1650-07

DOI: 10.3724/SP.J.1146.2012.01449

A Moving Object Detection Method Based on Sliding Window Gaussian Mixture Model

Zhou Jian-ying Wu Xiao-pei Zhang Chao Lü Zhao

(Key Laboratory of Intelligent Computing and Signal Processing Anhui University, Hefei 230039, China)

Abstract: In complex scenes, the traditional Gaussian Mixture Model (GMM) algorithm is an effective way to extract moving objects. However, after a period of modeling, the model's updating speed is difficult to keep up with the changes of true background. Consequently, false alarm rate will be increased in moving object detection. This paper proposed a new moving object detection method, which utilizes sliding window technology to retain short-term historical information. To a certain extent, it remedies that the traditional mixture Gaussian background model can not form new background in time, and it improves the integrality of motion detection. Furthermore, it reduces the algorithm sensitivity for the scene illumination change. The experimental results show that the proposed algorithm can more accurately, perfectly detect moving targets, and adapts rapidly to variations in the environment.

Key words: Moving object detection; Sliding window; Gaussian Mixture Model (GMM); Background model

1 引言

运动检测是智能监控的关键技术,也越来越受众多研究者的关注。常见的运动检测方法主要有光流法、相邻帧差法、背景减法、基于统计学习的方法^[1]。光流法^[2]能在不需要背景的任何先验知识的条件下实现运动目标的检测和跟踪,在运动的背景环境也能检测出独立的运动目标,但大多数的光流法计算都相当复杂,且抗干扰能力弱,所以适用性较低。帧差法^[3]是利用相邻帧相减得到运动目标,计算比较简单,但检测结果容易出现空洞现象。背景减法^[4]是通过提取特征数据来对场景中的背景进行建模,算法在一定程度上能较好地检测出运动目标,比较有代表性的背景减法如:简单自适应背景差分

法^[5]和基于混合高斯模型(Gaussian Mixture Model, GMM)^[6-9]的背景减法。简单自适应背景差分法能动态描述背景变化,但建立的是单一背景且包含大量环境噪声,在复杂场景下的运动检测效果较差;GMM运用多个高斯分布描述监控场景背景存在的不同状态,在一定程度上增强了模型的抗干扰能力,使建模得到的背景更具稳定性,但算法经过一段时间的建模,模型的部分参数逐渐达到某种稳定状态。如果此时光照发生突变或事物运动状态发生改变,GMM参数的更新难以跟上场景中真实背景的变化,从而导致运动检测结果中出现伪目标。基于统计学习方法^[10]抗干扰能力较强,检测到的运动目标比较准确,但算法难以采用统一的模型描述目标的运动特征,且计算复杂,所以算法不适用于实时处理。

针对上述运动检测方法存在的不足,本文综合考虑算法的复杂性和实用性,在混合高斯背景建模的基础上借助滑动窗技术具有短暂历史记忆特点,

2012-11-12 收到, 2013-02-18 改回

安徽省科技攻关强警专项(1101b0403030), 国家自然科学基金(61271352)和中国科学院上海微系统与信息技术研究所横向项目资助课题

*通信作者: 吴小培 IIP_HCIAHU@163.com

提出一种新颖的滑动窗式混合高斯模型参数更新思想,将运动检测分为滑动窗背景累积阶段和滑动窗混合高斯模型运动检测阶段。在运动检测阶段,算法通过设置不同的滑动窗长和调整窗口的移动步长来灵活地控制高斯模型对历史信息的遗忘速度,使建模得到的背景模型能更准确地反映场景中真实背景的实时动态。经过多次实验验证,算法提高了运动目标检测的准确性和完整性,同时弥补了场景光照突然发生变化而导致高斯背景模型来不及更新的缺点。

2 滑动窗背景模型

滑动窗技术在很多领域都得到了应用,也常用作信号的动态分析处理。在时域上,滑动窗口实际上可以看作信号的有限历史记录。因此,采用滑动窗来描述运动检测中的背景具有良好的实时性,但由于进行滑动窗数据分析的运算量都比较大,同时视频图像数据本身就很庞大,所以建立滑动窗在线递归算法是必要的。

在运动检测中,令 $\{I(t)\}_{t=1,\dots,T}$ 是给定的视频序列, $\{B(i)\}_{i=1,\dots,t-1}$ 是 t 时刻时所有的候选背景序列, N 是滑动窗的大小,滑动步长设为 1。滑动窗背景模型认为 t 时刻帧 $I(t)$ 的背景只与它最近的 N 帧候选背景序列有关。 t 时刻基于滑动窗口 N 的候选背景序列如图 1 虚线框所示。

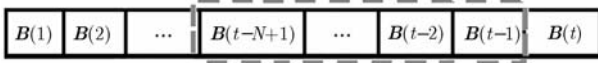


图 1 t 时刻滑动窗背景序列

基于现有均值和方差在线估计算法^[11],递推出 t 时刻连续背景序列的均值和方差矩阵:

$$\mu_t = \mu_{t-1} + \frac{1}{t} (B(t) - \mu_{t-1}) \quad (1)$$

$$\sigma_t^2 = \frac{t-1}{t} \sigma_{t-1}^2 + \frac{1}{t} (B(t) - \mu_{t-1}) \cdot (B(t) - \mu_t) \quad (2)$$

从图 1 可以看出, $t(t > N)$ 时刻时连续候选背景序列由 $t-N$ 时刻时连续候选背景序列和 t 时刻滑动窗内的候选背景序列组成。结合式(1)和式(2)不难得出 t 时刻连续候选背景序列对应的均值矩阵 μ_t 和方差矩阵 σ_t^2 的递推公式:

$$\mu_t = \frac{t-N}{t} \mu_{t-N} + \frac{N}{t} \mu_{st} \quad (3)$$

$$s_t^2 = s_{t-N}^2 + s_{st}^2 + \frac{(t-N)N}{t} (\mu_{t-N} - \mu_{st})^2 \quad (4)$$

其中 μ_{st} 是 t 时刻滑动窗内候选背景序列对应的均值矩阵, $\sigma_t^2 = s_t^2/t$, s_{st}^2 是 t 时刻滑动窗内候选背景序列

对应的方差和矩阵。从式(3)和式(4)推导,我们得到 t 时刻滑动窗背景模型对应的均值矩阵 μ_{st} 和方差矩阵 σ_{st}^2 的在线更新公式:

$$\mu_{st} = \frac{t}{N} \mu_t - \frac{t-N}{N} \mu_{t-N} \quad (5)$$

$$\sigma_{st}^2 = \frac{t}{N} \sigma_t^2 - \frac{t-N}{N} \sigma_{t-N}^2 - \frac{t-N}{t} (\mu_{t-N} - \mu_{st})^2 \quad (6)$$

3 混合高斯背景建模

观测一段视频序列,特定像素点 (x_0, y_0) 在任意 t 时刻的灰度值是随机的,该点在 t 时间段的灰度值集合为: $X = \{X_1, X_2, \dots, X_t\}$,且该点在大部分时间表示的是背景信息。因此,可以用混合高斯分布^[12]来近似建模。图 2 给出了考察某一像素点在一段时间内灰度取值的高斯分布情况,图 2(a)为单高斯分布示意图,图 2(b)为 3 高斯模型混合统计得到的概率分布图。

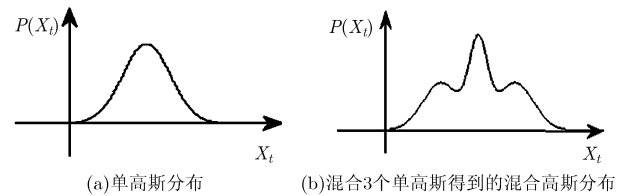


图 2 高斯概率分布示意图

GMM 算法是用 K 个高斯分布混合表示该随机过程,各高斯分布代表像素点不同状态。 $\omega_{i,t}$ 是 t 时刻像素点 (x_0, y_0) 第 i 个分布采样值占总采样值的权重, $\omega_{i,t} \in [0,1]$ 且 K 个模型权重和为 1, $\mu_{i,t}$ 和 $\sigma_{i,t}^2$ 是 t 时刻像素点 (x_0, y_0) 第 i 个模型的均值和方差, $\eta(X_t | \mu_{i,t}, \sigma_{i,t}^2)$ 是 t 时刻像素点 (x_0, y_0) 第 i 个分布的概率密度。于是, t 时刻像素点 (x_0, y_0) 的混合高斯概率密度 $P(X_t)$ 为

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t | \mu_{i,t}, \sigma_{i,t}^2) \quad (7)$$

采用期望最大化算法 (Expectation Maximization, EM) 迭代估计^[13],以寻求最优状态 $\hat{\theta}$ 。

$$\hat{\theta} = \arg \max_i \left\{ \omega_{i,t} * \eta(X_t | \mu_{i,t}, \sigma_{i,t}^2) \right\} \quad (8)$$

4 滑动窗混合高斯背景建模

传统 GMM 算法对场景进行一段时间的建模后,模型会因为场景光照变化、静止物体突然移动或移除而难以及时更新且模型抗干扰能力极差。本文引入滑动窗技术对传统算法进行改进,算法通过设定固定的滑动窗窗长和移动步长,从而有意向地加快

模型参数对历史信息的遗忘速度，使场景中新的背景信息能在较短的时间段内占据更新模型参数的优势，并实现模型对真实背景进行较准确地描述。

图3为传统GMM方法和基于滑动窗的GMM方法对某视频序列背景状态进行模拟的效果图。图3(a)是从序列中抽取出的5帧原图；图3(b), 3(c)分别为传统GMM方法和基于滑动窗的GMM方法在选定帧时刻模拟的灰度背景图；图3(d)为传统GMM和加滑动窗的GMM方法模拟选定像素点(图3(a)中圆圈标示)背景状态变化的曲线图，纵坐标代表像素灰度值，横坐标代表帧数，可以看出通过设置固定长度的滑动窗，模型可以在较短的时间段内完成新旧背景的更替，并能较准确地拟合出当前场景中该点的背景状态。图4给出了滑动窗混合高斯背景建模的示意图。图4(a)为距离 t 时刻最近的候选背景序列图像；

图4(b)为序列对应的灰度像素值矩阵；图4(c)为计算当前帧属于背景信息的概率矩阵。

在滑动窗GMM建模初期，我们采用混合高斯背景建模思想实现各模型参数的初始化。根据实际场景情况，设定滑动窗GMM参数初始化所需帧数 G_N ，一般取 $G_N = 200-300$ 帧便可获得较理想的初始化效果。模型初始化结束后，开始滑动窗GMM运动检测阶段。在滑动窗混合高斯模型运动检测过程中，将 t 时刻到来的帧图像像素点 (x, y) 的颜色灰度值 X_t 按照式(9)与 K 个滑动窗背景模型逐个实现匹配，式中 c 是经验常数，一般取 $c = 2.0 \sim 3.0$ 较合适。并按照匹配结果更新各个模型的权重： $\omega_{si,t} = (1 - \alpha) \cdot \omega_{si,t-1} + M_{si,t} \alpha$ ，匹配时 $M_{si,t} = 1$ ，否则 $M_{si,t} = 0$ 。

$$|X_t - \mu_{si,t-1}| \leq c \cdot \sigma_{si,t-1} \tag{9}$$

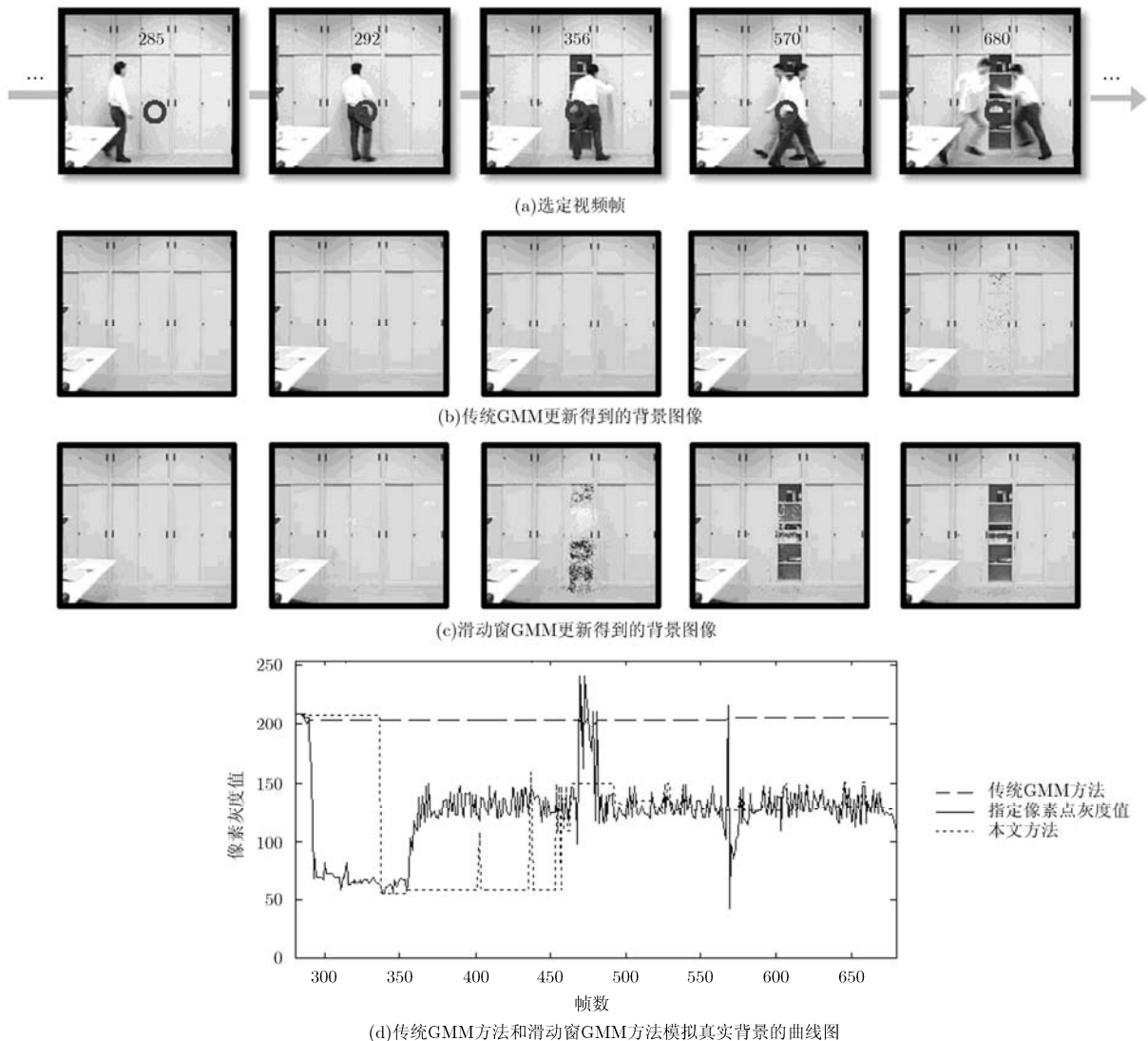


图3 传统GMM方法和滑动窗GMM方法模拟真实背景的效果图

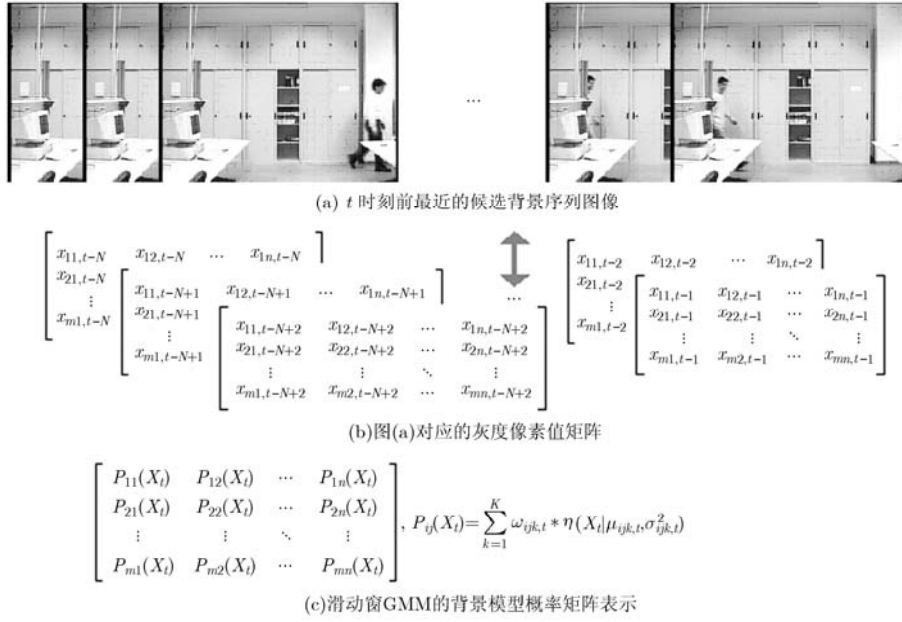


图4 滑动窗混合高斯建模示意图

若 X_t 与模型 i 满足式(9)，则按式(10)和式(11)更新该模型的均值 $\mu_{si,t}$ 和方差 $\sigma_{si,t}^2$ ， N 是窗长：

$$\mu_{si,t} = \frac{t}{N} \mu_{i,t} - \frac{t-N}{N} \mu_{i,t-N} \quad (10)$$

$$\sigma_{si,t}^2 = \frac{t}{N} \sigma_{i,t}^2 - \frac{t-N}{N} \sigma_{i,t-N}^2 - \frac{t-N}{t} (\mu_{i,t-N} - \mu_{si,t})^2 \quad (11)$$

若未找到模型与之匹配，则用式(10)和式(11)更新最后一个滑动窗模型的均值和方差，并按该模型的匹配次数提升对应的权重。模型参数更新结束后，对 $\omega_{si,t}$ 作归一化处理，并按照 $\omega_{si,t}/\sigma_{si,t}^2$ 优先顺序对所有模型排序，选取前 B 个满足式(12)的模型作为背景分布，其中 τ 是权重阈值。

$$B = \arg \min_i \left(\sum_{i=1} \omega_{si,t} \geq \tau \right) \quad (12)$$

若 X_t 与背景模型匹配，则认为该点是背景点；否则，判断为前景。算法流程如图 5 所示。

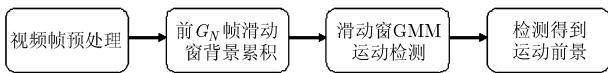


图5 滑动窗 GMM 运动检测流程

5 实验结果与分析

为了证实本文所提方法的有效性，我们采用室内和室外两段不同场景的视频在 matlab 平台上进行仿真。每段视频都分别使用简单自适应背景差分法、传统 GMM 方法和滑动窗 GMM 方法作对比分析，说明本文提出方法较其他方法更优越。实验参数设置如表 1 所示。

表 1 试验中各参数设置

混合高斯的模型数 $K=3$	与模型匹配的标准差系数 $c=2.9$
混合高斯建模背景更新率 $\beta_1=0.2, \beta_2=0.02$	B 个背景模型的权值门限 $\tau=0.8$
滑动窗背景累积帧数 $G_N=200$	各滑动窗模型的窗长 $N=50$
自适应背景减阈值 $Th=20$	自适应背景减更新率 0.004

为了更好地体现滑动窗 GMM 方法的优势，我们在传统 GMM 方法实验中设计了两种背景更新率：一种方法为整个检测过程采用一个固定较大的背景更新率 $\beta_1 = \beta_2 = 0.2$ ；一种方法为建模初期采用较大更新率 $\beta_1 = 0.2$ ，在背景模型稳定后采用较小更新率 $\beta_2 = 0.02$ 。

实验的第 1 段视频是来源于 VS-PETS 2001 dataset 的一段小区监控视频，视频的分辨率为 320×240 。为了更好突显本文方法背景更新较快的优势，我们选取视频中绿色小车由静止到运动离开停车位的一段视频进行实验。实验对比结果如图 6。图 6(a)为视频帧原图；图 6(b)是手工提取运动前景；图 6(c)为简单自适应背景差分法的运动检测结果，该方法在每帧运动检测结束后总是利用当前帧乘以一定的更新率来更新当前的背景模型，因此，形成的背景模型中含有环境噪声以及运动缓慢的目标信息，从图 6(c)中第 2 列、第 3 列都可看到缓慢运动的小车出现拖尾，且第 4 列图中仍包含绿色小车静止时的停车位以及早已静止的银灰色小车；图 6(d)

是模型更新率 β_2 为 0.2 的传统 GMM 方法得到的检测结果，该模型虽然能及时更新背景但同时易将部分运动缓慢的目标融入模型中，从第 1 列和第 2 列的检测结果能看出运动缓慢的小车被漏检了，且检测出的前景欠缺完整性；图 6(e)是模型更新率 β_2 为 0.02 的传统混合高斯建模得到的检测结果，该模型在检测运动缓慢的目标方面比模型更新率 β_2 为 0.2 的传统 GMM 稍显优势，但模型的更新速度缓慢导致检测结果中存在一些伪前景；图 6(f)为本文方法的检测结果，从图中可以看出该方法对噪声的抑

制优于简单自适应背景差分法，模型参数的更新也能较好地适应场景的变化，且能够获得较理想的检测结果。

实验的第 2 段视频选取的是一段室内视频，视频分辨率为 320×240 ，该段视频由于室内灯光的影响，运动阴影比较严重。为了更好地说明各方法对场景的适应效果，我们选取了柜门未拉开和完全拉开后各方法的检测结果，如图 7 所示。图 7(a)为选定视频帧原图；图 7(b)为手工提取运动前景；图 7(c)为简单自适应背景差分法的运动检测结果；图 7(d)

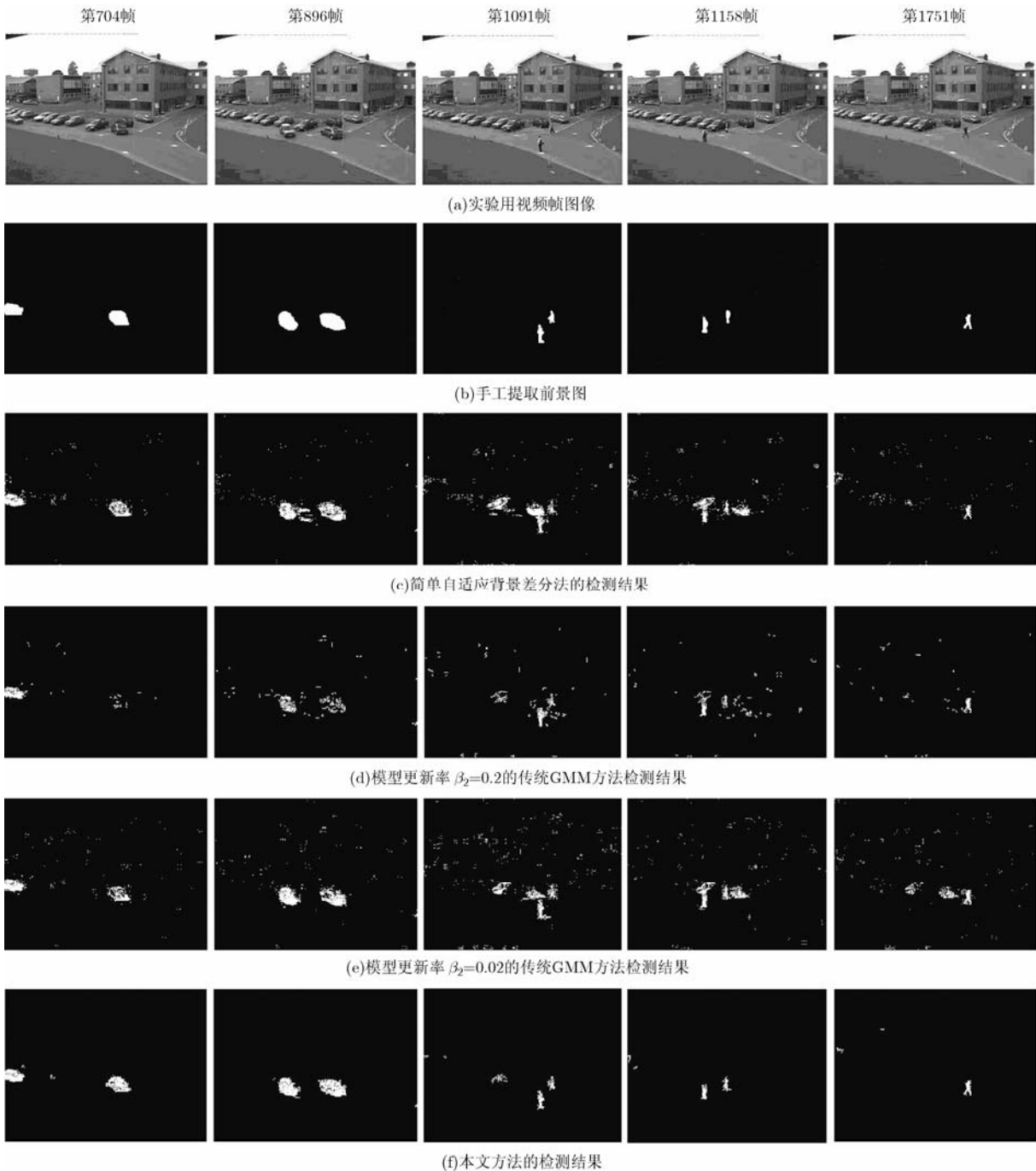


图 6 实验 1 中运动目标检测结果对比图

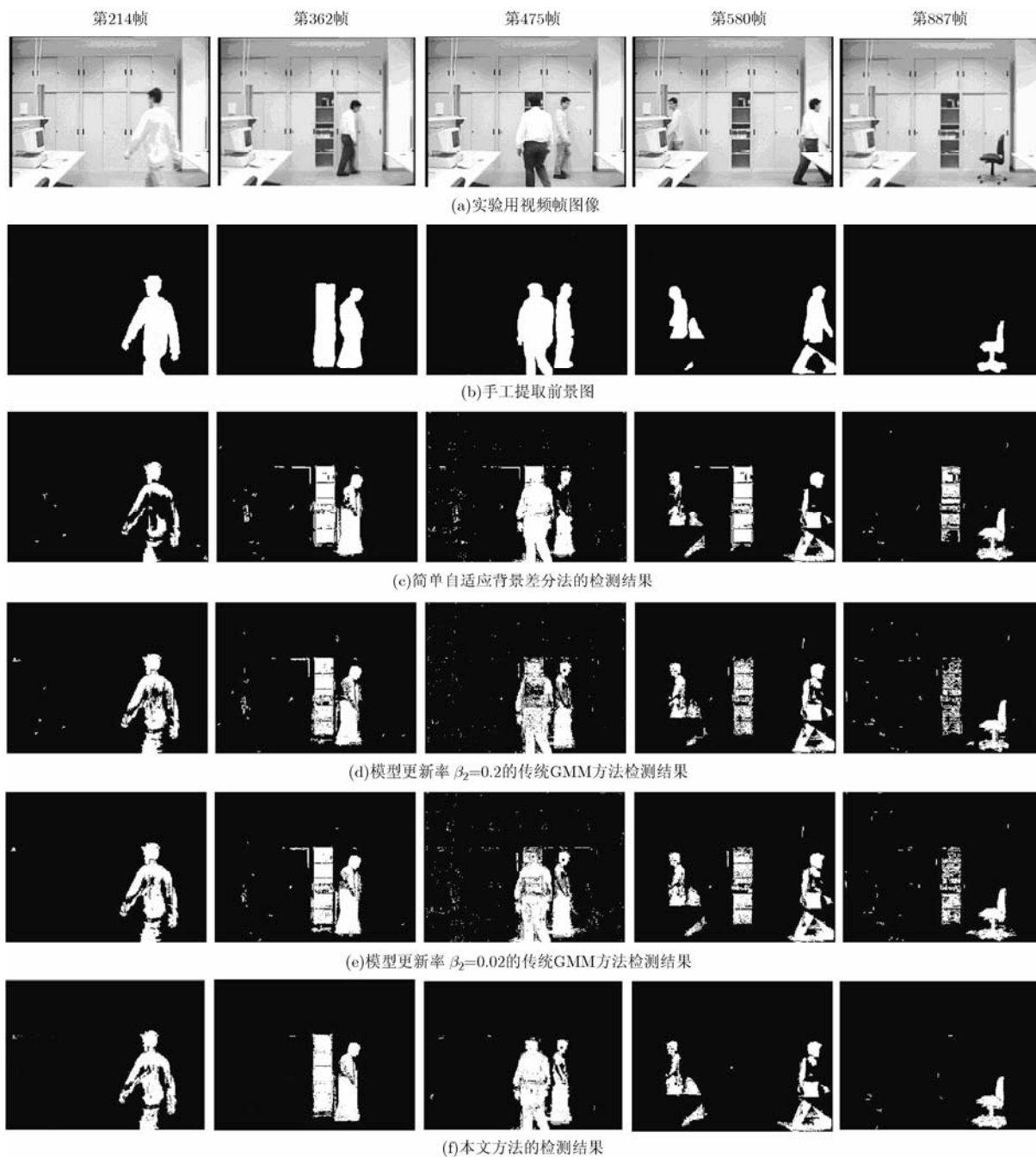


图7 实验2中运动目标检测结果对比图

是模型更新率 β_2 为 0.2 的传统 GMM 的检测结果；图 7(e) 是模型更新率 β_2 为 0.02 的传统 GMM 的检测结果；图 7(f) 为本文方法的检测结果。对比分图 7(c), 7(d), 7(e), 7(f) 可以看出 7(f) 中各列图噪声较少，且从第 3, 4, 5 列图可以看出其他方法因受到实验者甲、乙在拉开的柜门前来回交错行走而干扰其背景模型的更新，使得检测出的前景中始终包含裸露的储物柜；而本文方法很快将拉开的储物柜更新到背景中，从第 475 帧前景图可以看出分图 7(f) 已经没有裸露的储物柜干扰检测结果了，且不论实验者甲、

乙多次来回行走干扰都未影响运动检测结果，说明本方法在复杂场景下抗噪声能力强且背景更新速度较快，也能得到较完整的运动目标。

为了进一步描述各方法检测结果的准确性和完整性，我们使用衡量目标形状完整性的两个性能指标^[14]：检测率(DR)和误报率(FAR)，并对实验结果进行量化分析：

$$DR = \frac{TP}{TP+FN} \quad (13)$$

$$FAR = \frac{FP}{FP+TP} \quad (14)$$

其中 TP 为检测出属于真实前景的像素点个数, FN 为未被检测出的前景像素点个数, FP 为被误当作前景而检测出的属于背景区域的像素点个数。实验对比结果如表 2 所示。由于真实运动区域需要手工统计, 所以为了减少不必要的误差, 我们从各实验中随机抽取 50 帧图像, 并对各运动检测结果做 3×3 模板的形态学开运算。表 2 中给出的结果是我们统计每帧各运动检测方法的 DR 和 FAR 值后求得的平均值。通过对表中实验数据进行分析, 可以看出由于简单自适应背景差分法的背景单一, 在复杂场景下的检测率较低且误报率高; 更新率 $\beta_2=0.02$ 的传统 GMM 方法的检测率较高但误报率也较高, 该方法的误报率仅低于简单自适应背景差分法; 更新率 $\beta_2=0.2$ 的传统 GMM 方法的误报率相对较低, 但该方法的检测率比简单自适应背景差分法的检测率还低; 本文方法在保证低的误报率的同时, 方法的平均检测率接近 94%。

表 2 两个实验中 4 种运动检测方法的检测率和误报率对比

运动检测方法	实验 1		实验 2	
	检测率 (DR)	误报率 (FAR)	检测率 (DR)	误报率 (FAR)
简单自适应背景 差分法	0.6413	0.1924	0.7856	0.3826
传统 GMM ($\beta_2=0.2$)	0.6341	0.0672	0.7606	0.2912
传统 GMM ($\beta_2=0.02$)	0.7977	0.1608	0.9210	0.3608
本文方法	0.9347	0.0602	0.9376	0.1825

6 结束语

本文在传统的混合高斯建模基础上提出滑动窗的混合高斯模型来实现运动目标检测, 在复杂场景下该方法建立的背景模型有较好的抗噪声能力且对环境有很好的适应能力。实验验证表明, 该方法在一定程度上克服了自适应背景减法和传统混合高斯模型背景减法对环境噪声的敏感和适应环境变化迟缓的缺点, 且能得到较完整、准确的运动目标。然而, 由于在混合高斯建模过程中采用了滑动窗技术, 因此需要开辟较大的内存空间来记录一段时间内各模型的均值和方差, 这种做法将会增加运算量。在未来的工作中, 我们将综合考虑算法的复杂性和运动检测效果, 并尝试使用可变滑动窗长和可变移动步长来降低计算开销。

参 考 文 献

[1] 杨学超, 刘文萍. 视频图像序列中的运动目标检测技术[J]. 计

算机应用与软件, 2008, 25(1): 215-217.

- [2] Md Mosharrof Hossain Sarker and Andy Sloane. TGSE/TLoG filter with optical flow technique for large motion detection [J]. *Machine Graphics & Vision International Journal*, 2007, 16(3): 207-219.
- [3] 薛丽霞, 罗艳丽, 王佐成. 基于帧间差分的自适应运动目标检测方法[J]. 计算机应用研究, 2011, 28(4): 1551-1560.
- [4] Piccardi M. Background subtraction techniques: a review[C]. IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands, 2004: 3099-3104.
- [5] Xiong Wei-hua, Xiang Lei, Li Jun-feng, et al. Moving object detection algorithm based on background subtraction and frame differencing[C]. Chinese Control Conference, Yantai, China, 2011: 3273-3276.
- [6] Stauffer C and Grimson W E L. Adaptive background mixture models for real-time tracking[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, 1999: 246-252.
- [7] Greggio N, Bernardino A, Laschi C, et al. Self-adaptive Gaussian mixture models for real-time video segmentation and background subtraction[C]. 10th International Conference on Intelligent Systems Design and Applications, Cairo, 2010: 983-989.
- [8] Chen Ze-zhi and Ellis T. Self-adaptive Gaussian Mixture Model for urban traffic monitoring system[C]. IEEE International Conference on Computer Vision Workshops, Barcelona, 2011: 1769-1776.
- [9] 李拥军, 曾标, 徐克付, 等. 复杂背景下基于贝叶斯-全概率联合估计的前景检测[J]. 电子与信息学报, 2012, 34(2): 388-392.
- [10] 孙显, 王宏琦, 杨志峰. 基于形状统计模型的多类目标自动识别方法[J]. 电子与信息学报, 2009, 31(11): 2626-2631.
- [11] Myoungkeun Choi and Bert Sweetman. Efficient calculation of statistical moments for structural health monitoring[J]. *Structural Health Monitoring*, 2010, 9(1): 13-24.
- [12] Bailo G, Bariani M, Ijas P, et al. Background estimation with Gaussian distribution for image segmentation, a fast approach[C]. IEEE International Workshop on Measurement Systems for Homeland Security, Contraband Detection and Personal Safety Workshop, USA, 2005: 2-5.
- [13] Samé A, Ambroise C, and Govaert G. An online classification EM algorithm based on the mixture model[J]. *Statistics and Computing*, 2007, 17(3): 209-218.
- [14] Fakharian A, Hosseini S, and Gustafsson T. Hybrid object detection using improved gaussian mixture model[C]. International Conference on Control, Automation and Systems, Gyeonggi-do, 2011: 1475-1479.

周建英: 女, 1986年生, 硕士, 研究方向为图像处理、运动目标检测.

吴小培: 男, 1966年生, 教授, 博士生导师, 研究方向为智能信息处理与人机交互、生物医学信号处理.

张超: 男, 1983年生, 博士生, 研究方向为图像处理.