

一种基于多层背景模型的前景检测算法

杨涛, 李静, 潘泉, 程咏梅

(西北工业大学自动化学院, 西安, 710072)

摘要: 动态场景中的前景检测是后继处理的基础和制约整个智能视频监控系统稳定性、可靠性的关键。为了保证运动目标检测的基础上, 进一步检测出前景中的静止目标并消除“鬼影(Ghost)”, 提出一种基于多层背景模型的前景检测算法。将背景分为参考背景和动态背景两层, 分别采用单高斯和混合高斯模型进行背景建模。在线检测时采用动态背景提取变化前景, 用动态背景与参考背景之间高斯分布的差异提取静止前景, 同时, 通过逐层分析、比较输入像素与两层背景模型分布的相互关系, 快速消除 Ghost, 降低虚警。实验表明, 多层背景模型具有良好的检测性能和实时性, 为后继跟踪、分类等处理提供了坚实的基础。目前, 以该算法为核心构建了一个实时目标检测、跟踪系统 GreatWall system, 对图像大小为 320x240 的视频序列的平均处理速度达到 15 帧/秒。

关键词: 多层背景模型; 背景建模; 混合高斯; 静止前景检测

中图分类号: TP391 **文献标识码:** A

A Novel Multiple Layer Background Model For Foreground Detection

YANG Tao, LI Jing, PAN Quan, CHENG Yongmei

College of Automation, Northwestern Polytechnical University, Xi'an, 710072, China

yangtaonwpu@163.com, jinglinwpu@163.com

Abstract: Foreground detection is an important research problem in visual surveillance. In this paper, we present a novel multiple layer background model to detect and classify foreground into three classes, moving object, static object and ghost. The background is divided into two layers, reference background and dynamic background. Single Gaussian model and Gaussian mixture model are used respectively. Compared with many existing background models, a unique characteristic of the proposed algorithm is that through analyzing the Gaussian distributions of the two layers, static object and ghost are correctly labeled. A real-time object detection and tracking system named as GreatWall system is developed and tested under indoor and outdoor scenes with various scenarios. Extensive experimental results demonstrate that the proposed algorithm is effective and efficient and the processing speed of the system reaches 15fps for the image size of 320x240.

Keywords: Multiple background model; background modeling; Gaussian mixture model; static foreground detection;

1 引言

随着图像传感器的普及和广泛使用, 智能视频监控已经成为计算视觉与模式识别领域研究的热点和前沿课题。智能视频监控[1,2]的主要任务是对动态场景内感兴趣的区域和目标自动进行检测、跟踪、

分类和识别。作为后继处理的基础和制约整个智能视频监控系统稳定性、可靠性的关键点, 前景检测得到了国内外众多学者的高度关注。

基于背景差分的前景检测算法通过分析输入图像和估计背景的差异提取变化的前景区域, 对目标

类型、观测角度、距离等因素不敏感，因此在很多监控场合得到广泛应用。这类方法的关键是如何建立有效的背景模型。

考虑到真实场景的复杂性，我们认为理想的背景模型应具有解决下列问题的能力：

- 1) 实时性：实时性是决定模型是否能够实际用的关键；
- 2) 光照变化：不受控环境中(如室外)的亮度变化；
- 3) 运动干扰：枝叶晃动、摄像机抖动、雨、雪等混乱运动的干扰；
- 4) 目标移入：新目标移入时的自动背景更新；
- 5) 目标移出：场景内静止目标重新运动时 Ghost 的消除；
- 6) 静止目标：目标在场境内静止，如遗留物品，停止的车辆或行人；
- 7) 运动阴影：运动目标的阴影、倒影、玻璃橱窗的反光；
- 8) 伪装：运动目标与背景灰度相似；
- 9) 部分遮挡：目标被电线杆、路灯、树木等景物部分遮挡产生割裂；
- 10) 小目标：低信噪比环境中的小目标检测；

实际上由于精度、鲁棒性和实时性之间的矛盾，完全理想化的模型难以建立。目前的研究主要是针对特定场景实际需求建立对应的背景模型。

前景可分为变化前景（运动前景和 Ghost）[3]和静止前景[4]。前景检测的研究重点是解决实时性、光照变化、运动干扰、目标移入、目标移出、静止目标这六类问题。对运动阴影的研究是在提取运动前景的基础上，通过建立阴影模型对前景进行分类判断[5]。对伪装、部分遮挡以及小目标等问题多采用融合目标检测、跟踪、识别结果的思路，将后继处理结果反馈到前端进行综合决策。

为了解决前景检测问题，学者们提出了许多背景建模算法。时间平均模型是早期采用的背景模型

之一，它将图像上某一像素点在一段时间内的灰度平均值作为背景，算法具有容易实现、实时性强等特点，但易受噪声干扰。一种改进算法是采用中值滤波[6]，将像素灰度在时间窗内的中值作为背景灰度，这种算法对灰度突变产生的噪声有较强的抑制作用，但是由于需要对时间窗内的灰度值进行排序，计算量大，很难对整幅图像进行实时处理。为了克服光线变化的影响，J.Heikkila [7]提出一种加权更新算法，由于该算法具有很高的处理速度，且对环境光线变化有一定的适应能力，可用于运动干扰较少的场景中。

当图像上某一点灰度值存在规律性变化时，如枝叶晃动、水面波光、闪动的电脑屏幕等，像素的灰度分布将出现多个峰值，为解决多峰问题，学者们开始采用灰度统计特性进行背景建模。W4[8]用一定范围代替传统的单一灰度值作为背景，通过计算像素灰度可能的最大、最小值，以及相邻两帧的最大灰度差来构造背景模。Pfinder[9]率先提出用像素灰度的统计特性设计背景模型，对每一个像素采用单个高斯模型进行背景建模，C.Stauffer [10]用混合高斯模型(Gaussian mixture model, GMM)描述像素灰度值变化，用多个高斯模型可以描述灰度的多峰分布，可以有效的抑制运动干扰，但是过多增加高斯模型的个数会增加计算量。Elgammal[11]和 D. Comaniciu[12]用基于核密度估计的非参数方法描述任意分布的背景模型，为了降低计算量，典型的核函数值需要离线计算。

基于像素灰度统计特性的建模算法较好的解决了光照变化、运动干扰等条件下的背景更新问题，在许多监控场景中得到广泛的应用，但当目标静止时，上述算法均会将目标更新到背景，从而无法有效检测静止目标。此外，当静止目标重新开始运动时，目标所在位置会产生 Ghost。

本文针对动态场景内的前景检测进行研究。由于在实际的智能监控系统中，场景内新增的静止目

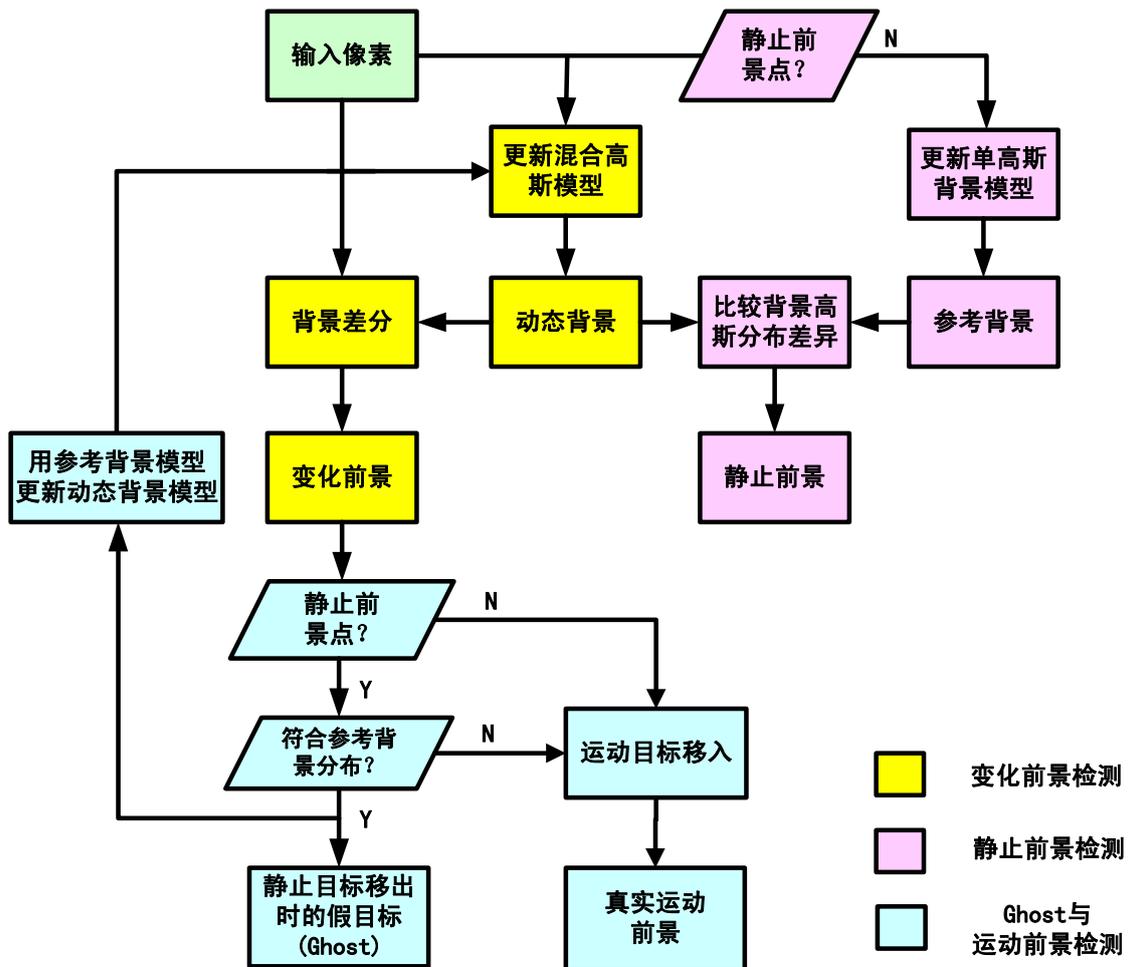


图 1 基于多层背景模型的前景检测算法框图

标，如遗留物 (Abandoned Object) [4],长时间滞留的行人、车辆等恰恰是监控人员应该特别关注的对象，为了在保证运动目标检测的基础上，进一步检测出前景中的静止目标并消除“鬼影 (Ghost)”，提出一种新的基于多层背景模型 (Multiple Layer Background Model, MLBM) 的前景检测算法，将背景分为参考背景和动态背景两层，分别采用单高斯模型和混合高斯模型进行建模。与传统背景建模算法相比，该算法不仅继承了混合高斯模型的优点，对光照变化、运动干扰不敏感，而且通过层间交互，有效地解决了静止目标的检测和目标重新移动时的 Ghost 快速消除问题。

第二节介绍了算法流程及框图，第三节详细讨论了基于多层背景模型的前景检测算法，第四节给

出了对实际场景的视频数据 (包括, PETS 数据库 Performance Evaluation of Tracking and Surveillance) 的实验结果及分析。

2 算法框图

算法包括三部分：变化前景检测，其中包括运动前景、Ghost 检测两部分以及静止前景检测，算法流程图见图 1。

变化前景检测模块用混合高斯模型进行背景建模，并将权值较高的几个高斯作为动态背景，不符合动态背景的像素标记为变化前景。

静止前景检测模块采用多个单高斯模型进行背景建模。为了适应场景的缓慢变化 (如光照变化)，在线检测时用输入图像有选择的对参考背景不断进

行更新，同时通过比较动态背景与参考背景高斯分布的差异来检测静止目标。

算法将变化前景分为运动前景和 Ghost 两部分。只有当静止前景处的像素发生变化时，分析该变化像素灰度值与参考背景的关系，如果像素灰度服从参考背景的高斯分布，则判断该处的像素变化是由于静止物体移出造成的 Ghost，同时用参考背景模型快速更新混合高斯模型参数；如果不符合参考背景的高斯分布，则判断像素变化是由于真实运动前景造成的。

3 基于多层背景模型的前景检测算法

3.1 变化前景检测

变化前景检测模块用混合高斯模型(Gaussian mixture model, GMM)进行背景建模，GMM 简述如下：

C.Stauffer [10]提出用 GMM 描述像素灰度值变化，每一个像素的灰度值用 K 个高斯分布描述，通常 K 值取 3~5， K 值的大小取决于计算机内存及对算法的速度要求， K 值越大，处理灰度变化的能力越强，相应所需的处理时间也就越长。定义像素点灰度值用变量 X_t 表示，其概率密度函数可用如下 K 个三维高斯函数表示 (1)。

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

上式中 $\omega_{i,t}$ 为第 i 个高斯分布在 t 时刻的权重，且有

$$\sum_{i=1}^K \omega_{i,t} = 1, \eta(X_t, \mu_{i,t}, \sum_{i,t}) \text{ 是 } t \text{ 时刻的第 } i \text{ 个高斯}$$

分布 (2)，其均值为 $\mu_{i,t}$ ，协方差矩阵为 $\sum_{i,t}$ 。

$$\eta(X_t, \mu_{i,t}, \sum_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\sum_{i,t}|^{1/2}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T \sum_{i,t}^{-1} (X_t - \mu_{i,t})}, i = 1, \dots, K \quad (2)$$

上式中 n 表示 X_t 的维数，为了降低计算量，通常认

为 R, G, B 三个通道相互独立，并有相同的方差，则有 $\sum_{i,t} = \sigma_i^2 I$ 。

本文将 GMM 中权值最大的几个高斯分布作为动态背景，如果某一像素灰度值不符合所有可能的动态背景分布，则判别为变化前景。此外，为了降低变量 X_t 维数，提高算法的实时性，在预处理时先将彩色图像灰度化，然后用混合高斯对灰度图像进行建模。

3.2 静止前景检测

GMM 在一定程度上解决了亮度变化、运动干扰等条件下的背景灰度多峰分布问题，但与大多数背景模型一样，这种算法会将静止不变的像素更新到背景，从而无法检测静止目标。为此，本算法采用多层背景思想，用参考背景与 GMM 生成的动态背景模型之间的分布差异检测静止目标。考虑到背景像素灰度的多峰分布，采用多个单高斯模型建立参考背景。

初始化时将 GMM 中权值最大的前 m 个高斯分布作为参考背景(3)。

$$\eta_{REF}(X_t, \mu_{k,t}, \sum_{k,t}) = \eta_{GMM}(X_t, \mu_{k,t}, \sum_{k,t}), k = 1, \dots, m \quad (3)$$

上式中 $\eta_{GMM}(X_t, \mu_{k,t}, \sum_{k,t})$ 是 t 时刻 GMM 的权值最大的第 k 个高斯分布， $\eta_{REF}(X_t, \mu_{k,t}, \sum_{k,t})$ 是 t 时刻参考背景的第 k 个高斯分布。

考虑到光照变化引起的背景本身灰度的缓慢改变，需要对参考背景进行动态更新。为了避免将变化前景误更新到参考背景中，与 GMM 不同，参考背景中多个单高斯模型在更新时没有权重区别，仅当输入像素灰度值符合某一高斯分布时才进行更新 (4)(5)(6)。

$$\mu_{k,t+1} = (1-\alpha) \cdot \mu_{k,t} + \alpha \cdot X_t \quad (4)$$

$$\sum_{k,t+1} = (1-\alpha) \cdot \sum_{k,t} + \alpha \cdot (X_t - \mu_t)^T (X_t - \mu_t) \quad (5)$$

$$\alpha = \beta \cdot \eta_{REF}(X_t | \mu_{k,t}, \sum_{k,t}) \quad (6)$$

其中 α 是更新率, 它由一个控制更新快慢的常数 β 以及概率 $\eta_{REF}(X_t | \mu_{k,t}, \sum_{k,t})$ 得到。

当目标静止时, 会在 GMM 中引入一个新的高斯分布, 随着静止时间的增加, 新引入的高斯分布权值逐渐增大, 并成为动态背景的一部分。而参考背景的更新机制保证了静止目标不会改变原有的模型参数, 因此可以通过比较动态背景与参考背景的高斯分布差异来检测静止前景。

当目标静止时间足够长时, 其引入的高斯分布权值会达到最大, 定义 $\eta_{GMM}(X_t, \mu_{\max,t}, \sum_{\max,t})$ 是点 (x, y) 在 t 时刻 GMM 权值最大的高斯分布, 均值为 $\mu_{\max,t}$ 。 $\eta_{REF}(X_t, \mu_{k,t}, \sum_{k,t})$ 是点 (x, y) 在 t 时刻的参考背景分布, $k = 1, \dots, m$, 通过判断均值 $\mu_{\max,t}$ 是否符合参考背景的高斯分布来检测静止目标。像素 (x, y) 的状态可用 $FS_t(x, y)$ (7) 表示。

$$FS_t(x, y) = \begin{cases} 1 & \eta_{REF}(\mu_{\max,t}, \mu_{k,t}, \sum_{k,t}) < T_p, k = 1, \dots, m \\ 0 & otherwise \end{cases} \quad (7)$$

上式中 $FS_t(x, y)$ 为 1 表示静止前景, T_p 是概率阈值。如果 $\mu_{\max,t}$ 不满足参考背景中全部 m 个高斯分布, 则该点被判定为静止前景。为降低计算量, 在实际计算时可用等价阈值 T 代替概率阈值 T_p , 则(7)式可改写为(8)。

$$FS_t(x, y) = \begin{cases} 1 & |\mu_{\max,t} - \mu_{k,t}| \cdot \sum_{k,t}^{-1/2} < T, k = 1, \dots, m \\ 0 & otherwise \end{cases} \quad (8)$$

3.3 真实运动前景与 Ghost 检测

如前所述, 静止在场景中的物体会融入背景中, 当物体重新开始移动时, 移出的背景会被检测为变化前景, 如果背景模型不能迅速跟上实际背景的变化, 检测结果中就会出现不合理的长时间静止前景物体。包括 GMM 在内的许多背景模型都存在不能立刻消除 Ghost 的问题。为此, 本算法将变化前景

分为两类: Ghost 和真实运动前景, 并通过分析变化像素灰度值与参考背景的关系进行分类判断。

Ghost 检测的基本思想是判断像素变化的原因是新物体移入还是静止物体移出。

定义 FM_t, FG_t, FT_t 分别是 t 时刻的变化前景, Ghost 和真实运动前景, I_t 是 t 时刻的输入图像。对于变化的前景像素点 (x, y) , 首先通过 $FS_t(x, y)$ 的取值判断该点是否属于静止前景, 如果是静止前景, 则判断输入像素灰度 $I_t(x, y)$ 是否符合参考背景的高斯分布, 如果符合则判断点像素点 (x, y) 的变化是由于静止物体移出造成的 Ghost, FG_t, FT_t 的计算见(9)(10)。

$$FG_t(x, y) = \begin{cases} 1 & \begin{cases} FM_t(x, y) = 1 \\ FS_t(x, y) = 1 \\ \eta_{REF}(I_t(x, y), \mu_{k,t}, \sum_{k,t}) < T_p, k = 1, \dots, m \end{cases} \\ 0 & otherwise \end{cases} \quad (9)$$

$$FT_t(x, y) = \begin{cases} 1 & \text{if } FM_t(x, y) \neq FG_t(x, y) \\ 0 & otherwise \end{cases} \quad (10)$$

对于判断为 Ghost 的像素点, 计算 $I_t(x, y)$ 在参考背景中每个单高斯模型中的概率, 用概率最大的第 k_{\max} 个高斯模型更新 GMM 动态背景中权值最大的高斯分布(11), 快速恢复背景。

$$\eta_{GMM}(X_t, \mu_{\max,t}, \sum_{\max,t}) = \eta_{REF}(X_t, \mu_{k_{\max,t}}, \sum_{k_{\max,t}}) \quad (11)$$

$$k_{\max} = \arg \max_k \eta_{REF}(I_t(x, y), \mu_{k,t}, \sum_{k,t}), k = 1, \dots, m \quad (12)$$

4 实验分析

以本文算法为核心, 用 VC++ 编程语言开发了一个实时目标检测、跟踪系统 GreatWall system。系统配置为 P4 处理器, 512M 内存, 输入视频序列分辨率为 320x240, 平均处理速度达到 15 帧/秒。文中分别给出该系统对 PETS 2002 (图 2)、PETS 2006 (图 3) 以及自建数据库 (图 4) 的实验结果。

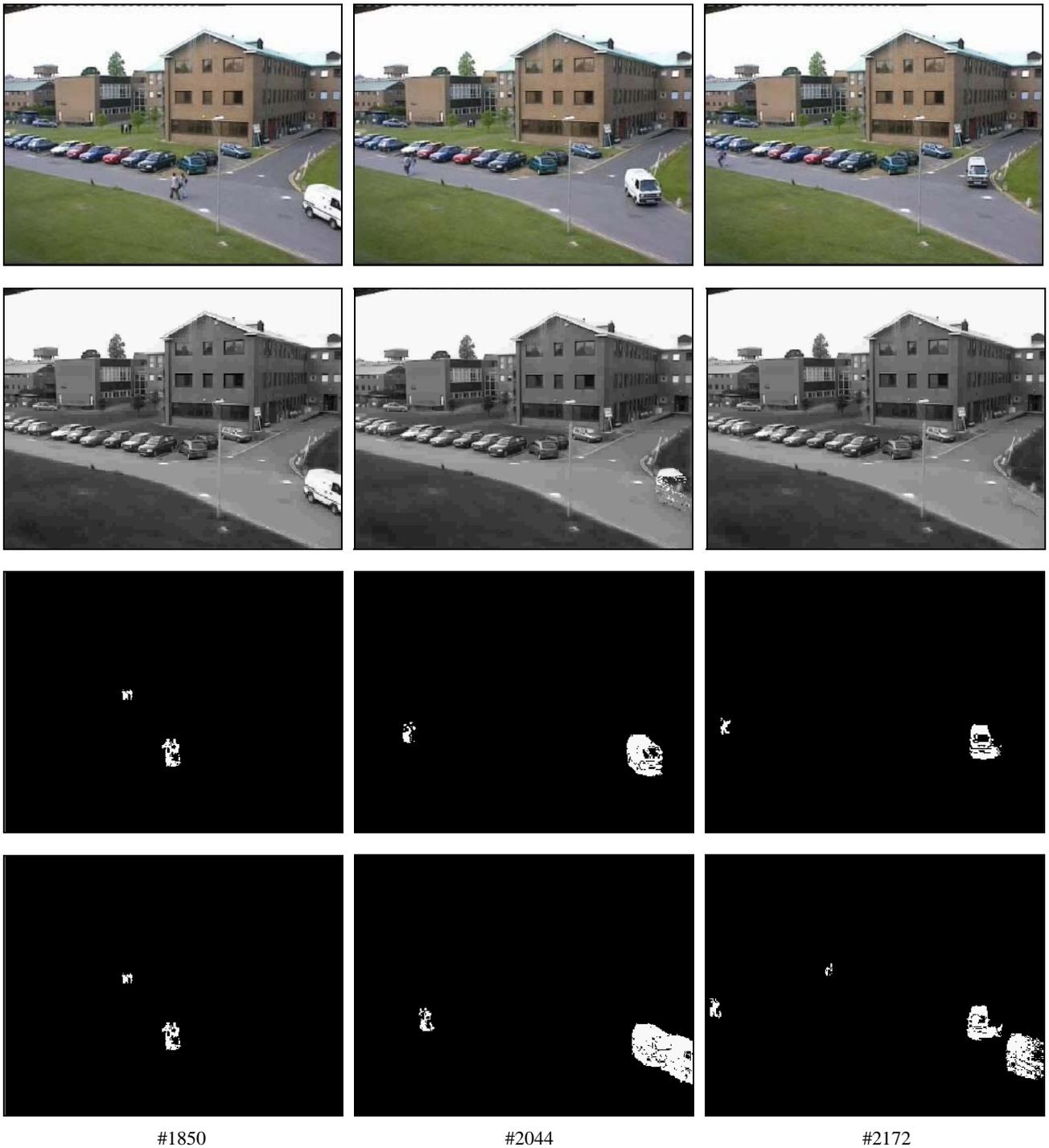


图 2 PETS 2002 数据库运动前景检测结果比较。第一行：输入视频序列；第二行：MLBM 的动态背景；第三行：MLBM 的前景检测结果；第四行：C.Stauffer 的混合高斯背景模型 [10] 的前景检测结果。

图 2 给出静止物体重新移动时算法对运动前景的检测结果，测试数据采用 PETS2002 数据库。图 2 中第一行是输入视频序列，第二、三行是本文算法 (MLBM) 得到的动态背景图和运动前景检测结果。为了对比不同算法在景物移出时的性能，第四行给出 C.Stauffer 的混合高斯背景模型 (GMM) [10]

对该序列的前景检测结果。

在白色面包车静止时(图 2,#1850)，MLBM 和 GMM 对于运动前景检测的性能相当 (图 2,#1805, 第三、四行)。当白色面包车开始移动时(图2,#2044)，动态背景发生变化，由于 GMM 需要一段时间的延迟才能完成背景更新，因此在它的前景检测结果中，

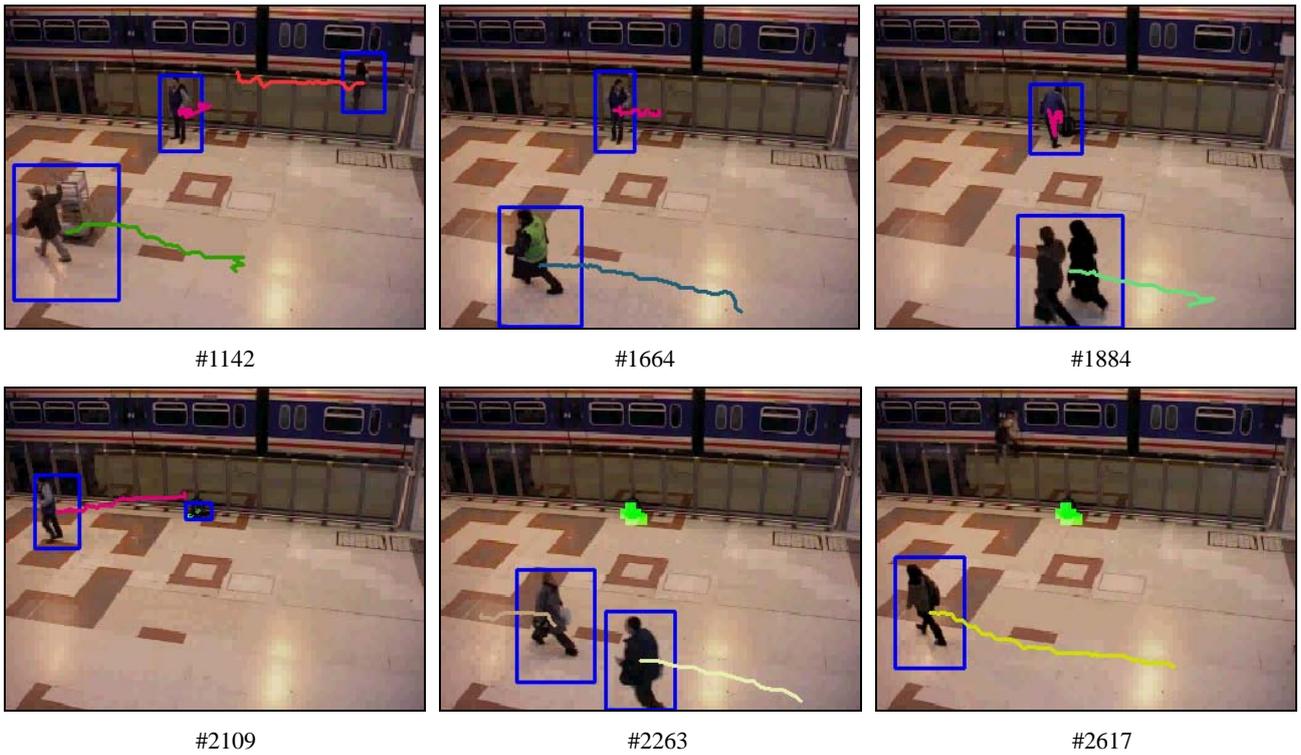


图 3 PETS 2006 数据库遗留物品检测与运动目标跟踪测试结果。检测到的遗留物品在图中用绿色显示 (#2263, #2617)，蓝色矩形框表示检测到的运动目标，彩色曲线表示目标轨迹。

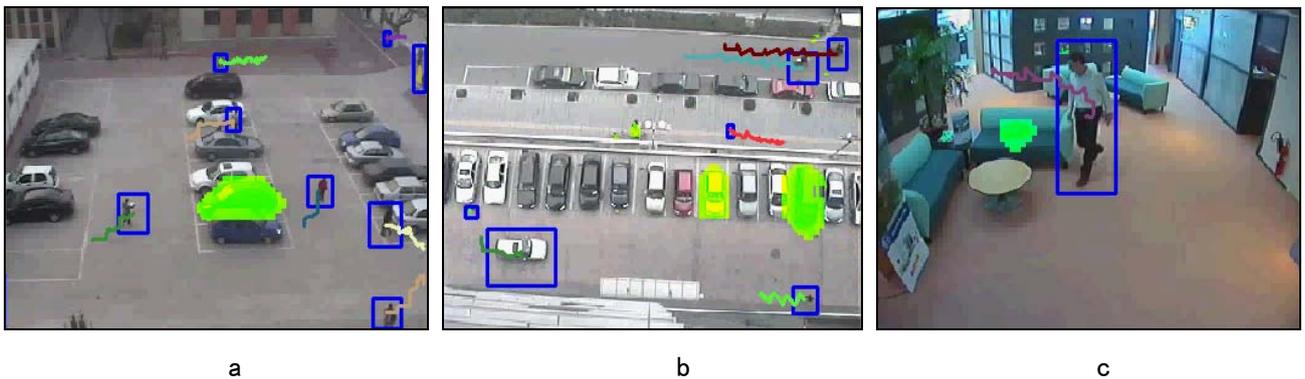


图 4 室内、外监控场景目标检测与跟踪结果。静止目标用绿色显示。

出现了大量 Ghost(图 2, #2044, #2172, 第四行)。而 MLBM 准确的判断出该处像素变化是景物移出背景造成的 Ghost, 并用参考背景对动态背景模型进行快速更新(见(11)式)。由图中可以看出, 车辆停留处的动态背景被实时恢复(图 2, #2044, #2172, 第二行), 同时在 MLBM 检测的真实运动前景中有效的消除了 Ghost(图 2, #2044, #2172, 第三行)。

图 3 给出一组在地铁站内遗留包裹检测的实验结果, 测试数据采用 PETS2006 数据库。图 3 中包裹的主人在徘徊一段时间后(图 3,#1142,#1664) 将

背包放在栅栏旁边(图 3,#1884), 然后离开(图 3,#2109)。传统的背景建模方法会慢慢将静止包裹更新到背景中, 从而在随后的前景检测中难以准确的提取该物品。而 MLBM 通过比较参考背景与动态背景之间高斯分布的差异, 快速检测出静止前景(图 3, #2263, #2617 绿色区域)。

GreatWall system 采用本文算法进行前景检测, 并用作者在文献[13]中提出的算法对检测到的运动前景进行跟踪。图 3, 4 中蓝色矩形框表示检测到的运动目标, 彩色曲线表示目标轨迹, 静止前景用绿

色标识。实验表明 MLBM 在室外(图 4.a, 图 4.b)、室内(图 4.c)均具有良好的实时性和鲁棒性,能够准确分割出静止和运动目标,为后继的目标跟踪、目标分类等处理提供可靠的检测结果。

5 结论

提出一种基于多层背景模型的前景检测算法,将背景模型分为参考背景和动态背景两层,分别采用单高斯模型和混合高斯模型建模。通过层间交互将前景分为运动前景, Ghost, 静止前景三类进行检测,在保留 GMM 模型原有优点的同时,还能对动态场景中的静止目标(如遗留物品,长期滞留的行人、车辆)进行检测,并快速消除目标重新移动时的 Ghost,大幅度提高前景提取的准确性,降低虚警。良好的检测性能和实时性为后继跟踪、分类等处理提供了坚实的基础。

算法提出的多层背景模型框架很好的解决了动态场景背景建模面临的实时性、光照变化、运动干扰、目标移入、目标移出、静止物体这六类问题。以该算法为核心的实时目标检测与跟踪系统 GreatWall system 对图像大小为 320x240 的视频序列的平均处理速度达到 15 帧/秒。

参考文献

- [1] R. T. Collins et al. A system for video surveillance and monitoring: VSAM final report. Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.
- [2] Arun Hampapur, Lisa M. Brown, Jonathan Connell, Max Lu, Hans Merkl, S. Pankanti, Andrew W. Senior, Chiao-fe Shu, and Ying-li Tian ,Multi-scale Tracking for Smart Video Surveillance, IEEE Transactions on Signal Processing, Vol. 22, No. 2, March 2005.
- [3] Ying-li Tian, Max Lu, and Arun Hampapur, Robust and Efficient Foreground Analysis for Real-time Video Surveillance, IEEE Computer Vision and Pattern Recognition, San Diego, June, 2005.
- [4] K. Smith, P. Quelhas, D. Gatica-Perez, Detecting Abandoned Luggage Items in a Public Space, Performance Evaluation of Tracking and Surveillance (PETS) Workshop at CVPR'06, New York, NY, June 18,2006.
- [5] Andrea Prati, Ivana Mikic, Mohan M. Trivedi, Rita Cucchiara, Detecting Moving Shadows: Algorithms and Evaluation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 7, pp.918-923, July 2003.
- [6] R. Cutler, L. Davis. View-based detection and analysis of periodic motion. International Conference on Pattern Recognition, pp. 495-500, 1998.
- [7] J. Heikkila, O. Silven. A real-time system for monitoring of cyclists and pedestrians.2nd IEEE Workshop on Visual Surveillance, pp.74-81,1999.
- [8] I. Haritaoglu, D. Harwood and L.S.Davis, W4: real-time surveillance of people and their activities, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, 2000: 809-830.
- [9] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfunder: Real-time Tracking of the Human Body. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 19, no.7, pp. 780-785, 1997.
- [10] C. Stauffer, W. Eric L. Grimson. Learning Patterns of Activity Using Real-Time Tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22 , Issue 8 , pp. 747 -757. August 2000.
- [11] A.Elgammal,et al., Background and Foreground Modeling using Non-parametric Kernel Density Estimation for Video Surveillance, Proceedings of the IEEE.90(7):1151-1163,2002.
- [12] B. Han, D. Comaniciu, and L. Davis. Sequential kernel density approximation through mode propagation: applications to background modeling. Proc. of Asian Conf. on Computer Vision, Jeju Island, Korea, 2004.
- [13] T.Yang, Stan Z.Li, Q.Pan, J.Li. Real-time Multiple Object Tracking with Occlusion Handling in Dynamic Scenes. IEEE Computer Vision and Pattern Recognition Conference (CVPR 2005) Vol I, pp.970-975, San Diego, CA, USA, June 20-26, 2005.