

一种自适应学习的混合高斯模型 视频目标检测算法

张 恒 胡文龙 丁赤飏

(中国科学院空间信息处理与应用系统重点实验室, 北京 100190)

摘 要 为解决背景模型的更新问题,提高视频运动目标检测性能,通过定义像素样本对模型更新的有效因子,提出一种自适应学习的混合高斯模型检测算法。用样本有效因子的历史累加量反映背景模型的质量,并用于动态调整模型更新速度。同时,对检测出的前景区域进行目标分析,由分析结果间接控制模型更新,保证更新的准确性和模型的稳定性。实验结果表明,该算法可以快速适应背景变化,同时保证目标检测的完整性。算法性能已在不同监控场景中得到验证。

关键词 混合高斯模型 智能视频监控 自适应学习

中图法分类号: TP242.6+2 文献标志码: A 文章编号: 1006-8961(2010)04-631-06

Adaptive Learning Gaussian Mixture Models for Video Target Detection

ZHANG Heng, HU Wenlong, DING Chibiao

(Key Laboratory of Spatial Information Processing and Application System Technology,
Chinese Academy of Sciences, Beijing 100190)

Abstract Background subtraction is a widely used method for video object detection and its performance is dependent on the quality of background model. In this paper, an algorithm for video target detection based on adaptive learning GMM was proposed by defining an efficiency factor between pixel samples and their background models. The accumulation of efficiency factor(AEF) shows how well the models can represent the background and was used to adjust the learning-rate dynamically. At the same time, how to update the models was dependent on the changes of the background after the foreground image analysis. The performance and robustness of the algorithm has been verified experimentally.

Keywords Gaussian mixture models(GMM), intelligence video surveillance, adaptive learning

0 引言

在智能视频监控领域,混合高斯模型^[1]是一种重要的背景建模方法,如何根据背景变化对模型作出及时而准确的更新是影响背景差运动目标检测的关键。很多文献对此提出了改进方法。文献[2]将模型更新分训练和检测两个阶段并采用不同的更新速度,尽管可以缩短背景模型的训练时间,但由于检

测阶段的更新速度较小且恒定,模型仍不能很好地适应背景变化。文献[3-4]则采用调整模式个数和学习速率的方法提高模型对背景变化的适应能力,但容易将运动较慢、面积较大的目标学习为背景,导致目标检测不完整。为此,文献[5-6]引入样本时域变化因子控制学习速率,但其有效性严重依赖于目标的纹理特征。

本文引入像素样本有效因子概念,用有效因子的历史累加量动态调整模型更新速度,以提高模型

收稿日期:2008-12-08;改回日期:2009-02-27

第一作者简介:张 恒(1982—),男。2006年于西北工业大学获学士学位,现为中国科学院电子学研究所硕士研究生。主要研究方向为智能视频监控、计算机视觉等。E-mail: zhh0408@yahoo.com.cn

对背景变化的适应能力。为保证模型更新的正确性,利用空域信息对前景区域进行目标分析,并用分析信息间接控制模型更新。

1 自适应学习混合高斯模型

针对复杂多变的背景,Stauffer 提出了混合高斯模型^[1],认为背景的每种变化形式均符合高斯分布,并对应于一个模式,用权重、均值和方差描述,各模式权重之和为 1,模型中权重较大、方差较小的模式为背景模式。为适应背景变化,模型需要不断更新。更新速度过慢,则模型不能及时反映出背景的变化而导致误检。而更新速度过快,又往往使运动较慢、面积较大或纹理较均匀的目标区域融入背景,造成目标检测不完整。

本节给出一种自适应学习的混合高斯模型,基于模型质量调整模型更新速度,提高模型对背景变化的适应能力,并利用上一帧的目标分析信息控制模型更新,保证模型的准确性和目标检测的完整性。

1.1 基于样本有效因子的模型更新方式

用高斯核函数描述样本与背景模型的匹配度:

$$K(d) = \begin{cases} \exp[0.5\lambda(\alpha^2 - d^2)] & d \leq T \\ 0 & \text{其他} \end{cases} \quad (1)$$

其中, α ($0 \leq \alpha < T$)、 λ ($\lambda \geq 0$) 控制核函数的形状, d 为样本与模式中心的标准差归一化距离, T 为距离阈值(取 3.0)。如果 $d \leq T$, 则称样本与模式匹配。

在一段时间内,如果背景模型与真实背景样本的匹配度很高,则说明模型质量比较好,应该减小模型更新速度,反之则应加快更新。此外,样本越新越能反映背景的变化,故在匹配度的定义中引入记忆系数 β (一般 $0.5 \leq \beta < 1$), 并将引入记忆系数的匹配度称为样本的有效因子。设 t' 时刻的样本为 $s(t')$, 其在 t' 时刻的匹配模式 M 在 t 时刻的均值为 $\mu_M(t)$, 标准差为 $\sigma_M(t)$, 定义 t 时刻该样本对背景模型的有效因子为

$$f_{M,t}(s(t')) = \beta^{(t-t')} \times K(d(s(t'), \mu_M(t), \sigma_M(t))) \quad (2)$$

$$t' = 1, 2, \dots, t$$

应用中,将历史样本有效因子的累加量 $\tau(t)$ 初始化为 $K(0)$, 并按下式更新:

$$\tau(t) = \beta \times \tau(t-1) + K(d) \quad (3)$$

易知, $\tau \in [K(T)/(1-\beta), K(0)/(1-\beta)]$ 。若模型中不存在与当前样本匹配的模式,则为该样本

添加一个匹配模式,匹配度为 $K(0)$ 。

历史样本有效因子的累加量反映了一定时间内背景模型的质量,可用于控制背景模型的更新速度。模型更新需要首先根据式(1)、式(2)计算出当前样本与背景模型的匹配度 $K(d)$, 然后按式(3)更新有效因子累加量,最后更新各模式权重、均值、方差如下:

$$w_k(t) = w_k(t-1) \times \left[1 - \frac{K(d)}{\tau(t)} \right] + \frac{L_k(t) \times K(d)}{\tau(t)} \quad (4)$$

$$\mu_{k,i}(t) = \mu_{k,i}(t-1) + \frac{L_k(t) \times K(d)}{\tau_k(t)} (s_i(t) - \mu_{k,i}(t-1)) \quad (5)$$

$$\sigma_k^2(t) = \sigma_k^2(t-1) + L_k(t) \times \frac{K(d)}{\tau_k(t)} \times \left[\left(1 - \frac{K(d)}{\tau_k(t)} \right) (s_i(t) - \mu_{k,i}(t-1))^2 - \sigma_k^2(t-1) \right] \quad (6)$$

其中, $s_i(t)$ 为样本 $s(t)$ 的第 i 个分量, $\tau_k(t) = \tau(t) \times w_k(t)$, k 为模式序号。若模式 k 为匹配模式,则 L_k 取 1, 否则取 0。

由式(4)~(6)可以看出,模型更新速度与有效因子的累加量成反比,累加量越小,模型更新越快,反之则更新越慢。若能够区分出背景样本与非背景样本,则样本有效因子累加量就能够反映背景模型的质量,实现模型质量对模型更新速度的控制。

1.2 前景分析及目标判决

基于样本时域分布的模型更新方式没能充分利用空域信息,不能区分出真正的背景变化和由目标引起的背景变化,即不能区分出背景样本与非背景样本,导致背景模型错误更新。我们注意到,在任何应用场景中,背景与兴趣目标是两个相对的概念,背景模型的更新应与兴趣目标的定义相关,故需要利用目标分析信息。检测框图如图 1 所示。

目标分析包括特征提取、目标判决及前景标记 3 个部分。首先结合目标先验知识提取前景区域特征(主要指空域特征,如面积、形状等),其次利用这些特征结合先验信息对前景区域进行目标判决。采用连通域标记算法^[7]标记前景像素,使前景像素可以根据其标号直接访问所属区域的特征及其判决结果,这样就可将目标分析信息送到前景检测模块。

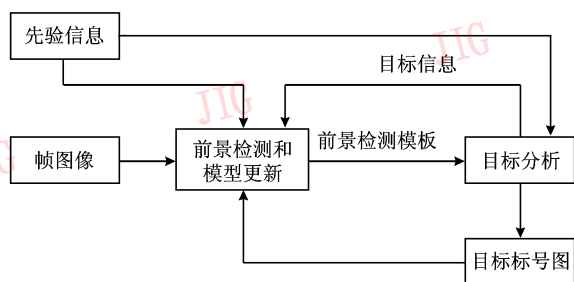


图1 融合目标分析信息的检测框图

Fig. 1 A feedback framework for motion detection integrating object analysis information

设每个像素的背景模型对应一个模式数组,并按权重由大到小的顺序进行存储。根据模式权重将这些模式分为有效模式和无效模式两类,如果模式权重不小于新模式的权重,则为有效模式,否则为无效模式。按下式^[1]将有效模式进一步分为背景模式和潜在背景模式:

$$b(t) = \arg \min_{b(t)} \left(\sum_{k=0}^{b(t)} w_k(t) > T_w \right) \quad (7)$$

$$0 < b(t) < n(t)$$

式中, T_w 为背景模式所应占的总比例, $n(t)$ 为 t 时刻有效模式的个数,则前 $b(t)$ 个有效模式即为背景模式。

在前景检测时,如果当前像素在上一帧被判断为目标,则样本只与背景模型中的背景模式进行匹配,否则与所有有效模式进行匹配,但模型更新策略相同。如果上一帧的前景像素被分析为目标像素并对应一个潜在背景模式,那么这种匹配策略将不会给该模式成长为背景模式的机会,从而阻止目标融入背景;相反,如果被分析为非目标,则将允许该模式成长为背景模式,因此不会影响非目标样本对应的潜在背景模式成长为背景模式,并适应不断变化的背景。

由匹配及匹配度定义可知,目标样本与背景模型的匹配度为0,变化剧烈的非目标样本与背景模型的匹配度也很小,故可将它们统一看做样本的剧烈变化。根据有效因子定义,如果样本的剧烈变化持续了一段时间,则会使有效因子的累加量减小,从而加快模型更新速度。但目标分析信息的利用可以阻止目标样本对应的潜在背景模式成长为背景模式,也就是说,累加量的减小只会加快对应于真实背景变化的有效模式的更新。

因此,利用上层目标分析信息对模型更新的间

接控制,可以在提高模型自适应力的同时,保证背景模型反映且只反映背景的变化,通过不断更新提高背景模型的质量。

2 混合高斯模型运动检测及模型更新

2.1 模型初始化

在系统开始运行时,用第1帧图像初始化背景模型(图像中可以包含目标),为每个像素添加一个新模式。新模式的均值为像素样本值,标准差设为 Var ,因此样本与该模式的匹配度为 $K(0)$ 。历史样本有效因子的累加量 $\tau(0)$ 为 $K(0)$,新模式权重为1。根据式(7)有效模式和背景模式的定义,新模式为像素的背景模式。

2.2 运动检测

如果某像素历史样本有效因子的累加量小于 $K(T)/(1-\beta)$,则该像素处于模型训练期。在模型训练期,模式匹配及模型的更新不利用目标分析信息。模型初始化的开始阶段,有效因子的累加量很小,像素样本与各有效模式进行匹配。如果存在匹配模式且该模式属于背景模式,则检测为背景;否则如果匹配模式为潜在背景模式,或没有匹配模式,则检测为前景。根据匹配情况按式(3)~(6)更新背景模型。如果不存在匹配模式,则为之添加一个新模式,添加方式同模型初始化,新模式权重初始化为 $K(0)/\tau(t)$ 。

在模型更新中,如果需要添加新模式且此时有效模式的个数已经达到模式个数上限 N ,则替换权重最小的有效模式 $b(t)$ 并从 $\tau(t)$ 中减去 $\tau_{b(t)}(t)$ 。如果有效模式 k 的权重在更新后小于 $K(0)/\tau(t)$,则有效模式个数减小1,同样地从 $\tau(t)$ 中减去 $\tau_k(t)$ 。对于匹配模式,其权重在模式更新后会增加,为保持权重由大到小的排列顺序,对匹配模式向前排序。模型更新完后,对更新后的有效模式进行权重归一。

如果历史样本的有效因子累加量不小于 $K(T)/(1-\beta)$,说明该像素的背景模型已经训练完成,并利用上一帧的目标分析信息对样本进行模式匹配。模型的更新方式与模型训练期相同。

3 实验结果与分析

用不同场景的视频样本对算法进行分析。视

频 1(380 × 284, 12.5 帧/s, 47 s) 为交通繁忙的街区, 摄像机存在明显抖动。视频 2(704 × 576, 30 帧/s, 59 s) 记录了一火车从进入场景到离开场景的整个过程。由于光照、抖动等原因, 背景有较大变化。

3.1 模型更新速度分析

模型更新速度的调整是通过更新样本有效因子的累加量实现的。不同位置的像素具有不同的有效因子累加量, 也就有不同的更新速度。取视频 1 中两像素 a(15, 109)、b(126, 315), 如图 2(a) 所示。像素 a 为建筑物上一点, b 为路面上一点, 不断有目标经过。在 350 ~ 450 帧之间, 有目标逐渐停在 b 处, 然后离开。

图 2(b) 显示了有效因子累加量的变化情况, 虚线为其动态范围。两像素样本值及相应的背景值如图 2(c)(d) 所示, 图中点线表示检测为前景。这里用最近背景模式估计背景值, 即如果当前样本被检测为前景, 则背景值保持不变, 否则用匹配模式的均值表示背景值, 但不要求该模式的权重最大^[8]。

通过大量类似实验发现:

- 1) 基于样本有效因子的模型更新方式可以大大缩短模型的训练时间, 如图 2(c)(d) 所示。随着系统运行, 有效因子累加量逐渐增加使模型更新速度变慢而趋于稳定。
- 2) 背景的剧烈变化会使有效因子的累加量减小, 从而使模型更新速度加快。
- 3) 在一段时间内, 像素变化越平稳, 有效因子的累加量越大, 反之越小。

可见基于样本有效因子的模型更新方式能够根据最近样本与背景模型的匹配情况, 动态调整模型更新速度, 提高了模型对背景变化的适应能力。

同时, 我们也应该看到, b 处目标在停止很短时间后, 就融入了背景而不能再检测出来(见图 2(d)), 这是我们不希望的; 但对于虚警目标, 这则是

有利的。这也说明了前景检测及模型更新中目标分析信息的必要性。

3.2 目标检测性能分析

针对两种截然不同的场景, 用本文算法检测其中的兴趣目标, 检测参数均设置为 $T_w = 0.85, \beta = 0.99, N = 3, Var = 4$ 并认为面积大于 9(视频 1) 和 350(视频 2) 的前景区域对应目标。在相同参数设置下, 对比文献[3]($\alpha = 1 - \beta$) 的检测结果(图 3)。两算法都只显示了大于面积阈值的前景区域。图 4 给出了系统初始阶段, 视频 1 的检测结果以及用最近背景模式估计的背景图。

图 3 视频 1 重点对比了小目标(图中红色椭圆区域)由运动到逐渐停止, 后又重新启动的整个过程的检测情况。当目标色彩与背景相近时, 文献[3]检测出的目标中会出现孔洞。当目标较大或运动速度较小时, 目标尾部会检测得不完整, 调高 β 会有所改善, 但同时会降低模型对场景变化的适应能力。文献[3]对视频 2 的检测, 在背景发生变化时, 会存在很长时间的虚警, 这主要是由模型更新较慢造成的。另外目标检测得也不完整(这在后续帧中会更加严重), 这是由目标较大、纹理较均匀造成的, 纹理均匀的部分对应的模式会成长为背景模式而导致检测不出来。

本文算法避免了将停留目标融入背景(视频 1), 也保证了目标检测的完整性同时能够较快地适应背景变化(视频 2), 且检测效果对检测参数的设置不敏感, 具有较强的鲁棒性。

由图 4 可以看出, 背景模型在数十帧内即可完成训练, 且随着系统运行, 模型质量会得到明显改善。整个过程中, 目标检测受到的影响不是很大。

3.3 时间耗费分析

目标分析是智能视频监控系统中不可或缺的部分, 将目标分析信息应用在前景检测及模型更新

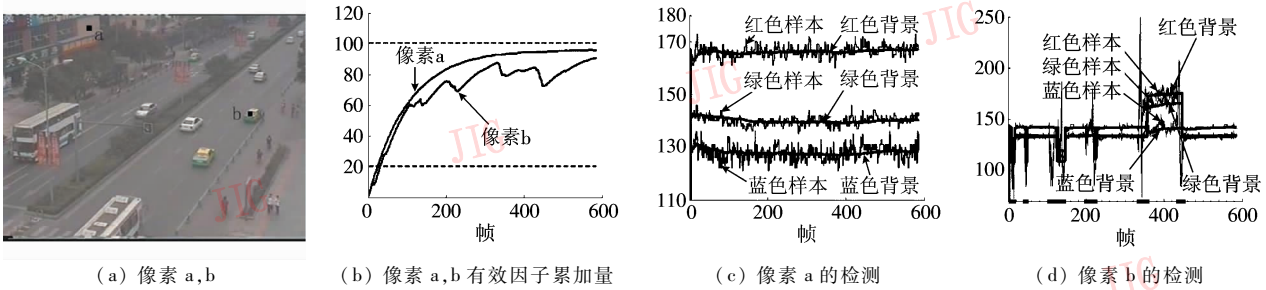


图 2 不同位置像素的检测情况
Fig. 2 Detection results of different pixels

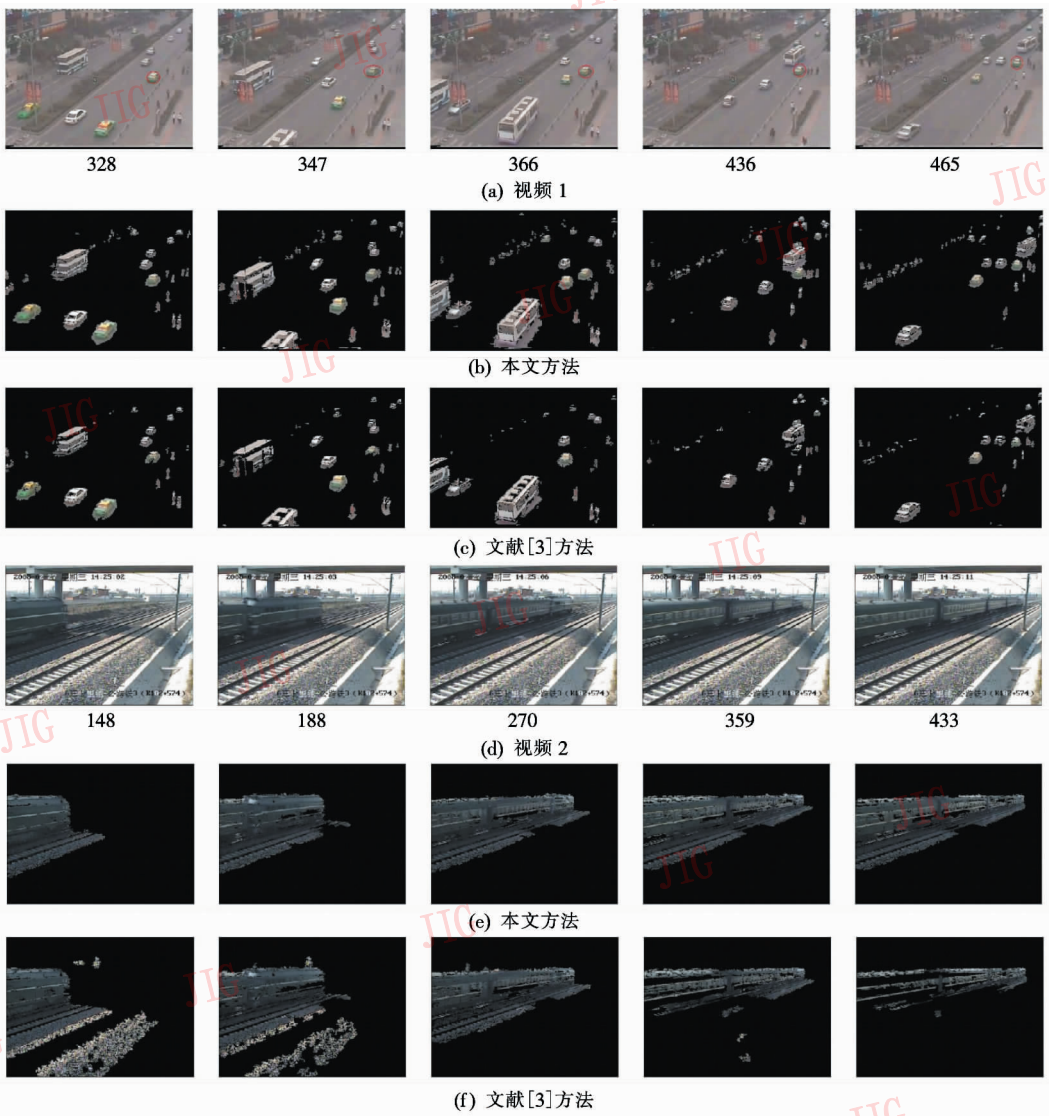


图 3 视频 1、2 的部分检测结果

Fig.3 Objects detection for video 1 and video 2 with the proposed method and the method in [3]

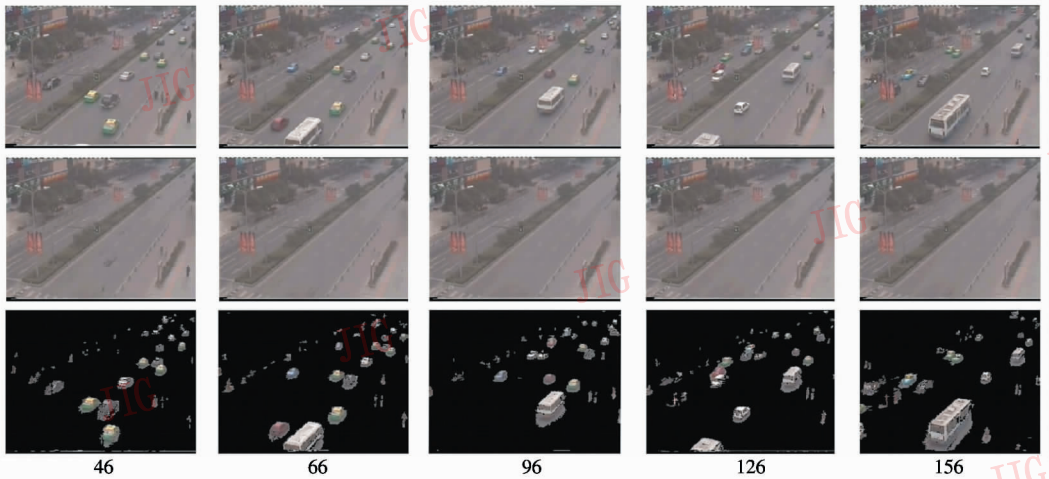


图 4 视频 1 中, 部分帧的背景图及相应检测结果

Fig.4 Background images and objects detection results for some frames in video 1 with the proposed method

部分,不仅提高了信息利用率,更重要的是显著改善了检测效果。与传统混合高斯模型相比,因此而增加的时间耗费主要在目标特征提取及标记部分。在主频为 3G 的奔 4 机器上统计视频 1 前 500 帧的总时间,进而得出每帧平均时间耗费约为 90 ms/帧(特征包括矩形边框和面积),每帧前景区域的平均个数(包括单像素前景点)近 300 个。

4 结 论

本文提出了一种自适应学习的混合高斯模型视频目标检测算法,引入背景更新的有效因子来调整背景更新的快慢,提高模型对背景变化的适应能力,并利用目标判决信息间接控制模型更新。在保证背景模型稳定性的前提下,大大缩短了模型的训练时间,提高了模型对背景变化的响应速度,保证了目标检测的完整性。其检测效果的鲁棒性和准确性已在实验中得到证实。

本文算法的有效性有赖于兴趣目标的正确判决。研究中只利用了前景的面积特征,且没有作阴影去除处理,这在一定程度上影响了算法效果。如能结合目标轨迹特征进一步提高兴趣目标判决的准确性,效果可能会更好。这些正是后续研究的内容。

参考文献 (References)

[1] Chris Stauffer, Eric W, Grimson L. Learning patterns of activity

using real-time tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 747-757.

[2] Kaewtrakulpong Pakorn, Bowden Richard. An improved adaptive background mixture model for real-time tracking with shadow detection[C]//Proceedings of the 2nd European Workshop Based on Advanced Video-based Surveillance Systems Providence, USA: kluwer Academic Publishers, 2001: 149-158.

[3] Zoran Zivkovic. Recursive unsupervised learning of finite mixture models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(5): 651-656.

[4] Lee Dar-Shyang. Effective gaussian mixture learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(5): 827-832.

[5] Michael Harville, Gaile Gordon, John Woodfill. Foreground segmentation using adaptive mixture models in color and depth [C]//IEEE Workshop on Detection and Recognition of Events in Video. New York: Computer Science Press, 2001: 3-11.

[6] Thongkamwitoon T, Aramvith S, Chalidabhongse T H. An adaptive real-time background subtraction and moving shadows detection[C]//IEEE International Conference on Multimedia and Expo. New York: IEEE, 2004, 2: 1459-1462.

[7] Chao Yuyan, Kenji Suzuki. A run-based two-scan labeling algorithm[J]. IEEE Intelligent Transaction on Image Processing, 2008, 17(5): 749-756.

[8] Han Hongzhe, Wang Zhiling. Adaptive background modeling with shadow suppression [EB/OL]. [2008-12-04]. <http://ieeexplore.ieee.org/stamp.jsp?arnumber=01252045>.