

# 智能监控系统中的运动目标检测算法

田 峥, 徐 成, 杨志邦, 冯 堃

(湖南大学计算机与通信学院, 长沙 410082)

**摘要:**为解决现有视频监控系统中目标检测算法无法应付复杂的室外环境且计算量和存储量较大等问题,将像素从 RGB 空间转换到 YUV 空间建立基于码本的背景模型,并单独对每个码字中的亮度分量进行高斯建模,提取运动目标的轮廓后,用连通区域算法对图像进行形态学处理。典型测试序列和 ROC 数据的对比实验结果证明该算法是高效和实用的,且易于在 DSP 或 FPGA 等嵌入式系统上实时实现。

**关键词:**运动目标检测;智能监控系统;码本模型;高斯混合模型;嵌入式系统

## Moving Objects Detection Algorithm in Intelligent Surveillance System

TIAN Zheng, XU Cheng, YANG Zhi-bang, FENG Kun

(School of Computer and Communication, Hunan University, Changsha, 410082, China)

**【Abstract】**In order to solve the problems that the existing motion detection algorithm in surveillance system can not work well in complex outdoor scene, and needs too much computation and memory, this paper proposes a moving objects detection algorithm based on improved codebook model. Pixels are converted from RGB space to YUV space to build the Codebook Model(CBM), and then the luminance component of each codeword is modeled by Gaussian model. The image is morphological processed by connected components algorithm. A test with the typical video sequences and the analysis of ROC data prove that the algorithm is effective and practical, and it can easily be implemented in embedded system such as DSP and FPGA.

**【Key words】**moving objects detection; intelligent surveillance system; Codebook Model(CBM); GMM; embedded system

DOI: 10.3969/j.issn.1000-3428.2011.04.0001

### 1 概述

智能视频监控系统是计算机视觉领域的一个重要应用,而运动目标的实时检测是智能监控系统中一个基础和关键的任务。目前在智能监控系统中通常采用帧间差分法和背景差分的方法来检测运动目标<sup>[1]</sup>,但是对于室外复杂的动态环境,由于存在光照变化、物体的周期性运动(如树枝晃动)、自然界中的动态纹理(如水波),因此现有的算法仍然无法精确提取出感兴趣的目标对象。目前效果较好的自适应背景建模方法由于计算量和内存占用太大,不适合在实时性要求较高的(特别是基于嵌入式系统的)视频监控系统中使用。

国内外很多研究者对运动检测算法进行了研究。例如,文献[2]利用统计方法为每个像素建立一个单模的分布模型,但是不能处理像树枝晃动这样的多模情况。文献[3]提出为复杂非静态背景建立高斯混合模型(GMM),取得了很好的效果,并且在许多应用中被广泛使用。但是该算法运算量大且更新参数很难调试。文献[4]提出了一种码本建模的思想,它为每个像素建立一个由多个码字组成的码本模型(Codebook Model, CBM),然后对新输入的像素值按照亮度范围和色度失真度进行聚类,从而提取出前景目标。该算法保持了背景的多模特征,而且内存占用和计算量都很小,具有较强的实用性。

考虑到 YUV 颜色空间更符合 CBM 算法亮度和色度分开计算的特性,本文提出在 YUV 空间下建立基于码本的背景模型;然后单独对每个码字中的亮度分量进行单高斯建模。在提取运动目标的轮廓后,用连通区域算法<sup>[5]</sup>对图像进行形态学处理,以消除噪点并提取出运动目标。实验结果表明,

本文算法是高效和实用的,且易于在 DSP 或 FPGA 等嵌入式系统上实时实现。

### 2 YUV 下基于码本的背景建模

视频序列中同一位置的像素点在不同时刻的观察值组成的序列  $X = \{X_1, X_2, \dots, X_N\}$  [tz1] 可以由一个码本模型  $C = \{c_1, c_2, \dots, c_L\}$  来描述;每个码本含有  $L$  个码字  $c_i$ , 每个码字是一个八元组:  $c_i = \langle \check{Y}_i, \hat{Y}_i, \bar{U}_i, \bar{V}_i, f_i, \lambda_i, p_i, q_i \rangle$ , 其中,  $\check{Y}_i, \hat{Y}_i$  表示该码字中所有像素的最小和最大  $Y$  分量;  $\bar{U}_i, \bar{V}_i$  表示该码字中所有像素的  $U$  和  $V$  分量的平均值;  $f_i$  代表属于该类码字的像素个数,也可理解为匹配成功的次数;  $\lambda_i$  代表没有被匹配的最大时间间隔;  $p_i, q_i$  表示第 1 次和最后一次匹配成功的时间。

算法的主要思想是将每个码字看作一个分类,然后对输入的像素值进行聚类,以判断该像素是否匹配码本模型。

图 1 显示了 YUV 空间下的码本匹配模型,如果输入的像素值  $X_i = \langle Y_i, U_i, V_i \rangle$  落在判别边界内部,那么  $X_i$  匹配这个码字;否则不匹配。匹配公式由如下 2 个公式给出为:

$$\text{brightness} = Y_i \in \langle Y_{\text{low}}, Y_{\text{hi}} \rangle \quad (1)$$

**基金项目:**国家“863”计划基金资助项目(2007AA01Z104)

**作者简介:**田 峥(1983 - ),男,博士研究生,主研方向:模式识别,计算机视觉;徐 成,教授、博士;杨志邦,博士研究生;冯 堃,硕士研究生

**收稿日期:**2010-07-11 **E-mail:** jt.tianzheng@gmail.com



### 5.1 实时性和与内存占用率

表 1、表 2 中的数据是通过对 3 种算法在 2 段视频序列上进行多次测试后得出的平均值。从结果可以看出, GMM 是实时性最差且内存占用最多的算法,这基本符合其算法特性。而 CB\_GMM 是实时性最好且内存占用最少的算法,其速度比 CBM 算法快了 1 倍多,这也基本符合上文的分析,由此证明,本文算法可以满足大多数智能视频监控系统对实时性的要求。

表 1 3 种算法的处理速度比较 f/s

算法)	视频序列 1	视频序列 2
GMM	16.23	3.86
CBM	31.71	9.41
CB_GMM	64.17	21.30

表 2 3 种算法的内存占用量比较 MB

算法	视频序列 1	视频序列 2
GMM	31.180	145.260
CBM	20.332	67.804
CB_GMM	15.652	59.684

### 5.2 ROC 数据分析

将视频序列的 Ground Truth Data 与算法的处理结果进行比较,并计算出算法的真阳性检测率  $TPR$  和假阳性检测率  $FPR$  等指标是目前对运动目标检测算法进行客观评价最常用的方法,也是最能准确反映算法真实效果的一种方法<sup>[6]</sup>。图 2 列出了 3 种算法对视频序列 2 中的每一帧处理的 ROC 数据。从  $FPR$  和  $TPR$  的含义可以得出, $FPR$  越低且  $TPR$  越高的算法,其计算结果的精确度越高,误检率也越小;及图中的即 ROC 数据越靠近坐标图的左上角,它所代表的算法的效果就越好。

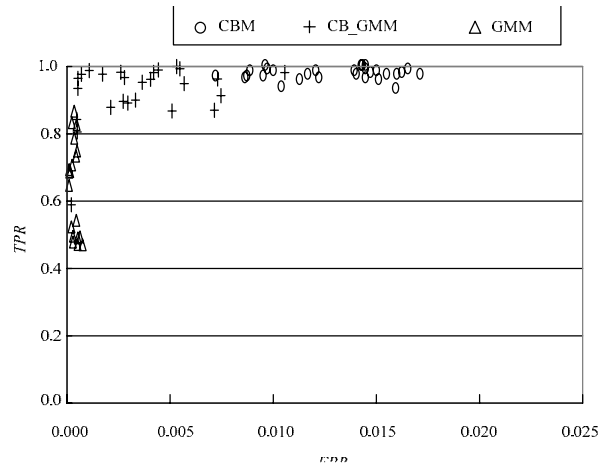


图 2 3 种算法对视频序列 2 处理的 ROC 数据

由图 2 可以看出, CB\_GMM 的 ROC 数据大多聚集在坐标图的左上角,而其他 2 种算法的 ROC 数据分布比较发散, GMM 算法虽然整体  $FPR$  值比较小,但是  $TPR$  的值偏低;而 CBM 正好相反,其  $TPR$  和  $FPR$  的值都很高。因此, CB\_GMM 是整体效果最好的算法,而在要求低误报率的情况下, GMM 的效果好于 CBM。

### 5.3 分割效果

视频序列 1 是一段典型的存在全局和局部光线变化的测试序列,图 3 显示了 3 种算法针对其中第 954 帧的背景分割结果(未经滤波处理);从左至右从上至下依次是:原图、CBM、GMM 和 CB\_GMM。从图中可以看出,效果最差的是 CBM 算法,目标提取不精确且误判点过多,对于与背景颜色相近的前景点无法准确提取,这都是由 CBM 算法的先验参数没有设置好所导致的。效果较好的是 GMM 算法,但是仍有一些噪声被判为前景,而且前者对于与背景颜色相近的前景点无法准确提取。效果最好的是 CB\_GMM 算法,噪声点很少,而且运动目标提取很精确。

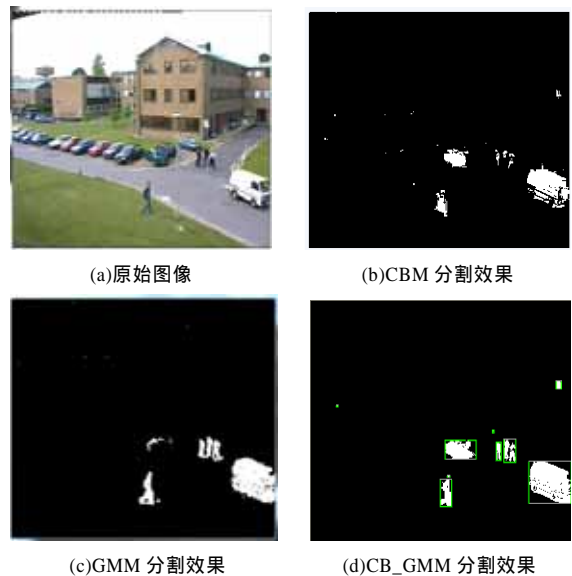
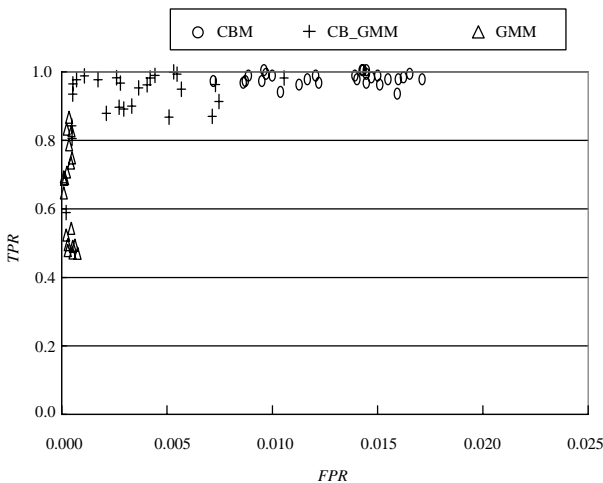


图 3 3 种算法对视频序列 1 的分割效果(未滤波)

## 6 结束语

本文提出了一种基于背景建模的运动目标实时检测算法。该算法主要有三大特点:

(1)在 YUV 空间下构建码本模型,利用其亮度与色度分离的特性来降低计算复杂度,使算法具有很好的实时性,易

于在嵌入式系统上实现。

(2)对每个码字中的  $Y$  分量建立单高斯模型,使整个码本具备高斯混合模型的特性,既保持了 CBM 算法原有的背景高度压缩、训练过程中允许前景存在、能适应全局或局部光线变化的特性,又结合了 GMM 算法鲁棒性和自适应性较好的优势。

(3)采用连通区域算法对分割后的二值图像进行形态学处理,提高了算法的抗噪性和运动目标提取的精度。

下一步将研究自适应的参数机制以应对算法先验参数过多的缺陷,同时考虑将该算法在嵌入式系统上实现。

#### 参考文献

- [1] 李广伦, 妥伟群. 视频监控系统中运动目标的实时检测[J]. 计算机工程, 2009, 35(17): 217-218, 222.
- [2] Wren C, Azabayejani A, Darrel T, et al. Pfunder: Real-time Tracking of the Human Body[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780-785.
- [3] Stauffer C, Grimson W. Adaptive Background Mixture Models for Real-time Tracking[C]//Proc. of IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE Press, 1999.
- [4] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time Foreground-background Segmentation Using Codebook Model[J]. Real-time Imaging, 2005, 11(3): 16772-256185.
- [5] 徐正光, 鲍东来, 张利欣. 基于递归的二值图像连通域像素标记算法[J]. 计算机工程, 2006, 32(24): 186-188.
- [6] Ellis T. Performance Metrics and Methods for Tracking in Surveillance[C]//Proc. of the 3rd IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. Copenhagen, Denmark: IEEE Press, 2002: 26-31.

编辑 张 帆