

◎博士论坛◎

核密度估计与高斯模型联级运动目标检测

芮挺¹, 周游², 马光彦¹, 廖明¹RUI Ting¹, ZHOU You², MA Guangyan¹, LIAO Ming¹

1. 解放军理工大学 工程兵工程学院, 南京 210007

2. 江苏经贸职业技术学院, 南京 211168

1. Engineering Institute of Engineering Corps, PLA University of Science & Technology, Nanjing 210007, China

2. Jiangsu Institute of Economic and Trade Technology, Nanjing 211168, China

RUI Ting, ZHOU You, MA Guangyan, et al. Target detection using kernel density estimation and Gaussian model cascade mechanism. *Computer Engineering and Applications*, 2011, 47(18): 1-3.

Abstract: Gaussian model and kernel density estimation model are effective ways for background modeling. Calculation of Gaussian model is simple, however, it suffers from low robustness when there are dynamic scenes and/or sudden lighting changes. Kernel density estimation is robustness but it is too complex to calculate in real-time. A cascade detection mechanism is proposed. Most of the stable pixels are segmented by Gaussian model. After that, for a small part of the pixels that the Gaussian model can not accurately describe are segmented by kernel density estimation model. Experiments confirm that the proposed method is effective to deal with dynamic backgrounds and fast in computation.

Key words: target detection; kernel density estimation; Gaussian model; cascade mechanism

摘要: 高斯模型与核密度估计模型是两种有效的背景建模及目标检测方法。高斯模型运算简单, 但对复杂背景的描述能力差; 核密度估计模型对背景描述能力强, 但运算复杂, 难以实现实时性检测。提出了一种分层联级检测机制, 由高斯模型对大部分相对稳定的像素进行分割与检测, 对于高斯模型无法精确描述的小部分像素通过核密度估计模型完成分割与检测。实验证实了该方法在适应动态背景扰动与运行效率方面的有效性。

关键词: 目标检测; 核密度估计; 高斯模型; 联级机制

DOI: 10.3778/j.issn.1002-8331.2011.18.001 文章编号: 1002-8331(2011)18-0001-03 文献标识码: A 中图分类号: TP391

1 引言

智能视觉监控(Intelligent Visual Surveillance, IVS)技术源自计算机视觉(Computer Vision, CV)与人工智能(Artificial Intelligent, AI)的研究, 其目的在于将图像与事件描述之间建立一种映射关系, 利用人工智能技术自动分析、抽取视频源中的有用信息, 使计算机从纷繁的视频图像中分辨、识别出关键目标对象, 并对其行为进行分析与判断。

典型的智能视觉分析过程包括: 场景建模(Environment Modeling)、运动目标分割(Motion Segmentation)、目标跟踪(Object Tracking)、目标分类(Object Classification)、行为分析和描述(Behavior Understanding and Description)。场景建模作为大多数视觉分析任务的第一步是实现目标跟踪、目标分类、行为分析等视觉任务的重要基础。

场景建模的困难主要在于存在背景扰动, 通常背景可能产生如下几种典型扰动^[1]: 场景光线变化, 包括逐渐的变化和

突然的变化; 状态变化, 包括摄像机抖动和背景中物体的闪动; 背景物体变化, 包括物体融入和移出背景。

背景的变化特性对背景模型的动态更新提出了很高的要求。时间平均模型^[2](Temporal Average Model)和高斯背景模型^[3](Single Gaussian Model)计算量小, 但当背景像素灰度呈多峰分布时检测效果不佳; 基于卡尔曼滤波的背景估计模型^[4]能够跟踪背景的光线变化, 但对于背景物体的变化不能及时跟踪, 会导致目标的误检和漏检; Stauffer 和 Grimson 提出了一种高斯混合模型(Mixed Gaussian Model)用于背景提取^[5-6], 通过多个高斯分布拟合单个像素值变化, 能够及时地跟踪背景的变化, 但是这种算法的计算量很大, 必须采用专门的硬件才能用于实时监控; Elgammal 和 Duraiswami 提出了一种非参数的核密度估计(Kernel Density Estimation)方法^[7], 不用事先假定像素值的分布, 就可以直接从像素的连续变化值中估计出概率密度函数, 其检测效果优于高斯混合模型, 计算量

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.61070104, 61070108); 解放军理工大学工程兵工程学院基金。

作者简介: 芮挺(1972—), 男, 博士, 副教授, IEEE 高级会员, 研究方向: 模式识别, 图像视频分析, 智能计算; 周游, 副教授; 马光彦, 博士, 讲师; 廖明, 博士, 讲师。E-mail: rtinguu@sohu.com

收稿日期: 2011-02-28; **修回日期:** 2011-04-18

却进一步增大。

高斯模型运算简单,核密度估计模型对随机序列分布描述精确。本文提出了一种综合利用两者的优势与特点,建立联级机制实现运动目标检测的方法,有效地降低了运算量,提高了检测的鲁棒性。通过实验证实了本文方法的有效性。

2 高斯模型背景建模

高斯模型背景建模的基本原理是对每一个像素利用高斯模型建模,每一像素点都被认为是服从均值 μ 和标准方差 σ 的高斯分布,且每一点的高斯分布是独立的,从而用一个高斯函数描述一个像素序列的变化。

在室内及稳定的室外环境下,高斯模型是建立背景模型的有效方法,具体步骤如下:

(1)初始化高斯模型。将视频序列图像中每个像素的灰度变化看作是一个随机变量序列,计算一段时间内这个随机序列的均值(即平均灰度 μ_0)与方差 σ_0^2 ,由此构造了一组用来描述各个像素灰度随机序列分布的高斯模型 $\sum(\mu_0, \sigma_0^2)$ 。其中:

$$\mu_0(x, y) = \frac{1}{T} \sum_{i=0}^{T-1} f_i(x, y) \quad (1)$$

$$\sigma_0^2(x, y) = \frac{1}{T} \sum_{i=0}^{T-1} [f_i(x, y) - \mu_0(x, y)]^2 \quad (2)$$

$f_i(x, y)$ 是第 i 帧中像素 (x, y) 的灰度值。

(2)前/背景分割。根据所建立的高斯模型分布,确定当前帧中的像素属于背景还是属于前景。对于当前帧中的一个像素来说,如果其满足该像素序列所描述的背景高斯模型,则其被确定为背景,否则为前景,即按式(3)进行分割。

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{[F(x, y) - \mu]^2}{2\sigma^2}\right) \begin{cases} > \text{threshold} & \text{背景} \\ \leq \text{threshold} & \text{前景} \end{cases} \quad (3)$$

其中, $F(x, y)$ 表示当前帧图像像素 (x, y) 的灰度值; threshold 为阈值。

(3)更新高斯模型。在视频序列当前帧更新时,模型需要根据背景灰度变化或前景变化进行更新,对于前景与背景像素,采用不同的更新策略。

对于确定为背景的像素,高斯模型需要对两个参数 (μ_0, σ_0^2) 按照式(4)、(5)进行更新。

$$\mu_i(x, y) = \mu_{i-1}(x, y) + \alpha[F(x, y) - \mu] \quad (4)$$

$$\sigma_i(x, y) = \sigma_{i-1}(x, y) + \alpha\{[F(x, y) - \mu_i(x, y)]^2 - \sigma_{i-1}(x, y)\} \quad (5)$$

其中, α 表示模型更新速率,是一个 $[0, 1]$ 的常数,用来反映当前帧对模型的影响权重。

高斯模型在稳定环境下具有较好的背景描述能力和检测效果,此时像素序列分布的方差较小,同时基本符合高斯分布的假设。但当环境出现较明显的扰动时,像素序列分布往往呈多峰形式,并且由于原高斯模型的方差较小,很容易将扰动错误地分割成前景目标。

3 核密度估计背景建模

核密度估计是数理统计及相关应用领域的重要工具。它运用解析化的手段,对给定样本及分布规律未知的数据集求解随机变量的分布密度函数,从而构造出这些数据样本在特定空间的分布模型。

核密度估计方法不需要假定背景特征概率分布形式,也

不需要设置模型参数和对参数优化,而是从样本中直接得到概率密度估计函数。只要有充分的样本,核密度估计就能渐近收敛于任何一个概率密度函数。

核密度估计通过加权平均中心点位于采样值的局部函数来估计未知的密度分布。给定某像素特征空间的一个样本集 $S = \{x_1, x_2, \dots, x_N\}$, 观测值 x_i 的密度分布用 $p(x_i)$ 来估计

$$p(x_i) = \sum_{i=1}^N \alpha_i K_h(x_i - x_i) \quad (6)$$

式中权值 $\alpha_i = 1/N$; $K_h = (1/h)K(t/h)$ 是核函数; h 是核函数的宽度。如果一个像素考虑 m 个特征,核函数 K_h 选择高斯函数, σ_m 表示各个特征的核函数带宽,则对于式(6)可以进一步表示为

$$p(x_i) = \alpha_i \sum_{i=1}^N \prod_{m=1}^d \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-\frac{1}{2} \left[\frac{(x_{im} - x_{im})^2}{\sigma_m^2}\right]\right) \quad (7)$$

在核密度估计中,核函数带宽 σ 的选取远比核函数本身更为重要。带宽选择的过大会使概率密度估计出现欠平滑的现象,反之则会出现过平滑,这些都将对估计结果产生不利影响。假设 x_i 服从正态分布 $N(\mu, \sigma^2)$, 则 $(x_i - x_{i+m})$ 服从正态分布 $N(0, 2\sigma^2)$, m 是相邻两帧像素特征值差的绝对值中值,即 $m = \text{median}(|x_i - x_{i+m}|)$, 由正态分布的对称性和中值性质得到

$$P[N(0, 2\sigma^2) > m] = 0.25 \quad (8)$$

由正态分布知:

$$\sigma = \frac{m}{0.68\sqrt{2}} \quad (9)$$

在实际使用中发现,当用来进行核密度估计的样本帧序列较长时,计算样本帧序列像素的中值将耗费大量的时间,为了提高实时性,可以将样本序列的中值用均值近似,对估计结果的影响并不显著。

核密度估计背景建模本质上是一种扩展的混合高斯背景建模方法,相对于混合高斯模型中只用3~5个高斯分布描述背景模型的方法而言,可以更准确地描述样本的密度分布,因而具有更强的鲁棒性。由于核密度估计方法理论上的先进性,在实际使用中也取得了较好的效果,但它也存在比较明显的不足,主要是计算复杂,无法满足视频检测的实时性要求。核密度估计计算复杂性问题自算法提出时起,就引起了研究人员的广泛关注,特别是在视频分析这类对实时性要求较高的应用中,提高运算效率已成为使核密度估计方法能够有效应用的重要问题。

4 高斯模型与核密度估计联级目标检测

单高斯模型在各类目标检测方法中,运算最为简单,在环境稳定、背景单一的条件下,像素序列分布基本符合单峰形态,可以用高斯分布近似描述,因而高斯模型的分割能获得较好的效果。在实际应用中,无论室内室外,环境条件和背景变化的扰动都是不可避免的,对于这些复杂的变化,单高斯模型无力进行精确的描述,因此当出现较明显的快速扰动时,单高斯模型容易出现错误分割,并在短时间内崩溃。而核密度估计作为一种有效的非参数估计方法,可以比较好地描述复杂多峰的概率分布,但其运算复杂,难以满足视频分析实时性的要求。

本文提出采用高斯模型与核密度估计联级的方法实现背景建模以及对运动目标的检测。对于同一场景,分别建立描

述背景的高斯模型与核密度估计模型;首先利用高斯模型对当前帧进行前/背景分割;然后对于高斯模型不能精确描述的像素,利用核密度估计方法进行二次分割,确定这类像素前/背景类别。

由于在一个场景中出现目标的时间和区域是有限的,采用这样的机制,大部分的检测工作是由运算量较小的高斯模型完成的,通过高斯模型分割筛选后,少量高斯模型无法精确描述的像素由核密度估计模型完成检测工作。这样的机制使算法的实时性和鲁棒性得到了较好的统一。本文算法的实现过程如图1所示。

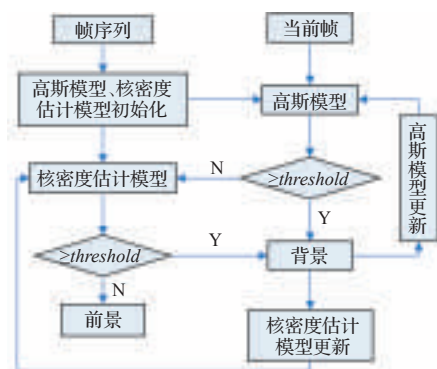


图1 算法实现流程图

首先,由若干初始帧分别建立高斯背景模型与核密度估计模型;然后,将当前帧中各像素用高斯模型进行分割,如果某像素以大概率($\geq threshold$,这里的 $threshold$ 大于传统高斯模型分割的阈值)符合高斯背景模型,则划分为背景,如果像素符合高斯模型的概率小于阈值,则表明用高斯模型分割像素可能出现误分割,此时利用核密度估计方法对该像素进行前/背景分割;最后,对于背景像素更新高斯背景模型与核密度估计背景模型,前景部分不参与背景模型的更新。

为了应对可能出现的原背景物体出现运动,转化成前景目标,或原前景目标长期静止而转化成背景物体,可设置一个计数器 $counter$,如当前帧像素为背景时,置 $counter=0$,如当前帧像素为前景时,置 $counter=counter+1$ 。当 $counter$ 大于计数器阈值 $counterthresh$ 时,认为该像素的状态长期未作改变,将该像素点记为背景,参与高斯模型与核密度估计模型的更新。

5 测试实验与结果

为了验证算法对自然条件扰动的适应性,测试实验选取室外自然环境,设置了不同天气与背景的监控场景,主要检验算法对目标检测的有效性和对环境扰动的适应型。测试所用设备为2.4 GHz主频CPU的笔记本,算法由VC++6.0基于OpenCV1.0实现,视频分辨率640×480像素。

核密度估计模型由30帧图像序列和高斯核函数构成,分割阈值 $threshold=10^{-6}$;对于高斯模型,分割阈值 $threshold=10^{-3}$ 。

在环境稳定的条件下,高斯模型完成了对绝大多数像素的分割,当出现环境扰动时,核密度估计模型克服了扰动的影响,将受到扰动的像素作为背景,并更新高斯模型与核密度估计模型。更新后的高斯模型方差变大,抗扰动的能力也相应增强,在后续的处理中又可以接管大部分核密度估计模型的分割工作。因此在目标检测过程中,高斯模型与核密度估计

模型进行联级的协同工作,运算简单的高斯模型完成大部分像素的分割处理,核密度估计模型提供对扰动抵抗,同时为高斯模型提供方差的修正,来提高高斯模型的抗扰动能力。

图2中是两处不同监控场景对于行人运动目标检测的结果。可以看出,在不同的环境中,本文算法完整地提取出了小尺度目标,未出现误检区域;同时该测试在长达数小时的检测中运行稳定,不受环境光线与其他扰动的影响。测试结果证明系统采用的方法是有效的。



图2 目标检测实验分割结果图

6 结论

核密度估计方法对背景模型有更精确的描述,但存在运算量巨大无法满足实时检测的缺陷。本文提出了一种分层联级目标检测机制,由高斯模型完成对大部分像素的检测工作,运算量大为降低;由核密度估计模型提供对高斯模型无法表达像素的精确描述,只有在某些像素以小概率符合高斯模型时,才由核密度估计对这些像素进行前/背景分割。这样一种机制合理分配了简单高斯模型与复杂核密度估计模型的检测对象,综合利用了高斯模型与核密度估计模型各自的优点,在显著降低运算量的同时,使鲁棒性得到明显提高。长时间的实验测试证明了本文方法的有效性。

参考文献:

- [1] Huang S S, Fu L C, Hsiao P Y. Region-level motion-based background modeling and subtraction using MRFs[J]. IEEE Transactions on Image Processing, 2007, 16(5): 1446-1456.
- [2] Friedman N, Russell S. Image segmentation in video sequences: a probabilistic approach[C]//Proceedings of 15th Conference on Uncertainty in Artificial Intelligence. Providence, Rhode Island, USA: Morgan Kaufmann Publishers, 1997: 175-181.
- [3] Rider C, Munkelt O, Kirchner H. Adaptive background estimation and foreground detection using Kalman-filtering[C]//Proceedings of International Conf on Recent advances in Mechatronics(ICRAM'95), Istanbul, Turkey, 1995: 193-195.
- [4] Koller D, Weber J, Huang T, et al. Towards robust automatic traffic scene analysis in real-time[C]//Proceedings of International Conference on Pattern Recognition, Jerusalem, Palestine, 1994: 1262-1311.
- [5] Stauffer C, Grimson W E L. Learning patterns of activity using real-time tracking[J]. IEEE Trans PAMI, 2000, 22(8): 474-757.
- [6] 李红波, 唐培竣, 吴渝. Kalman滤波器对混合高斯背景建模的改进[J]. 计算机工程与应用, 2009, 45(24): 162-165.
- [7] Elgammal A, Duraiswami R, Harwood D, et al. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance[C]//Proceedings of IEEE Conference, New York, USA, 2002: 11512-11631.