

自适应核密度估计运动检测方法

徐东彬^{1,2} 黄磊¹ 刘昌平¹

摘要 提出了一种自适应的核密度估计 (Kernel density estimation, KDE) 运动检测算法. 算法首先提出一种自适应前景、背景阈值的双阈值选择方法, 用于像素分类. 该方法用双阈值克服了单阈值分类存在的不足, 阈值的选择能自适应进行, 且能适应不同的场景. 在此基础上, 本文提出了基于概率的背景更新模型, 按照像素的概率来更新背景, 并利用帧间差分背景模型和 KDE 分类结果解决背景更新中的死锁问题, 同时检测背景的突然变化. 实验证明了所提出方法的适应性和可靠性.

关键词 核密度估计, 运动检测, 自适应背景/前景阈值, 突变背景
中图分类号 TP391.4

Adaptive Kernel Density Estimation for Motion Detection

XU Dong-Bin^{1,2} HUANG Lei¹ LIU Chang-Ping¹

Abstract This paper proposed a method of adaptive kernel density estimation (KDE) for motion detection. To begin with, an approach for adaptive selecting thresholds of foreground and background was proposed. By using the two thresholds, the approach can overcome defects of using only one threshold. More importantly, these two thresholds can be selected automatically and they are independent of scenes. Meanwhile, a background model updated according to probability was also provided. The background model of inter-frame difference incorporated with results of KDE can solve deadlock situations in background model. It can also be used to detect suddenly changed background. Experimental results were given to demonstrate that the proposed algorithms are suitable and effective for motion detection.

Key words Kernel density estimation (KDE), motion detection, adaptive background/foreground threshold, suddenly changed background

从视频序列中分离出运动目标, 是计算机视觉中的重要研究内容, 可以应用到交通监测、行为识别和人机交互等领域. 背景减除技术 (Background subtraction) 是摄像机静止条件下的一种应用广泛的运动检测技术, 为此研究人员提出了不同的方法^[1]. 其中时间平均 (Time average)^[2] 是最简单的方法, 它对内存要求低, 处理速度快; Lo 和 Velastin 采用最近 n 帧的中值作为背景模型^[3], 这种方法在场景中存在多个运动目标或目标运动缓慢时, 前景会污染背景; 帧间差分 (Inter-frame difference) 信息^[4-6] 可以用来生成背景, 这类方法能够应用于多个运动目标的场合, 但更新率不容易选择; Colombari 等把像素分布看作高斯分布, 对背景建模^[7]; Stauffer 和 Grimson 用 k 个高斯混合分布 (Gaussian mixture model, GMM) 对每个像素进行建模^[8], 解决了像素的多峰分布和缓慢变化背景的问题, 这类方法假设像素符合高斯分布, 但实际像素分布可能不满足该假设, 对于频繁变化的像素, 需

要多个高斯混合分布才能反映这种变化; 左军毅等假设背景满足空间一致性, 将背景复杂度用熵来描述, 以此为纽带建立背景模型^[9]; Oliver 等提出在整幅图像上进行特征值分解, 充分利用空间相关性^[10], 取得了较好效果, 但需耗费大量内存, 且检测精度低于高斯混合模型^[1]; Elgammal 等提出了基于核密度估计的非参数背景模型^[11], 该方法能够适应不同的场景, 不同于 GMM, 它充分利用最近的历史帧信息来表示背景模型, 能够适应复杂的像素分布密度, 克服像素值在短时间内发生的频繁变化, 因此能得到较准确的估计, 但是文中通过给定的假正率 (Percentage of false positives) 来选择阈值, 需要有场景的先验知识, 对不同的场景需要人工干预进行重新选择; Mittal 和 Paragios 根据设定的误报率 (False alarm rate) 和漏报率 (Miss probability), 通过训练样本调整阈值^[12], 使用该阈值在给定的条件下能够获得相对好的分类结果, 但采用单阈值很难解决误报率和漏报率之间的矛盾, 不可避免带来分类误差, 特别是当场景改变时, 需要重新选择样本进行训练, 才能得到适用于该场景的阈值. 本文对概率直方图进行分析, 提出了一种自适应前景阈值和背景阈值的双阈值选择方法. 该方法能够适应不同场景, 克服采用单阈值引起的分类误差. 在此基础上, 本文采用概率来更新背景模型, 能够得到更为合理的背景, 有利于运动目标的检测. 此外, 本文还提出了利

收稿日期 2007-11-16 收修改稿日期 2008-03-25
Received November 16, 2007; in revised form March 25, 2008
1. 中国科学院自动化研究所文字识别工程中心 北京 100190 2. 中国科学院研究生院 北京 100049
1. Character Recognition Engineering Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190
2. Graduate University, Chinese Academy of Sciences, Beijing 100049
DOI: 10.3724/SP.J.1004.2009.00379

用帧间差分背景模型和核密度估计 (Kernel density estimation, KDE) 分类结果, 解决背景更新中的死锁问题^[11, 13], 并用于突变背景的检测, 较好解决了背景突变时的运动目标误检测问题。

本文第 1 节介绍文献 [11] 提出的 KDE 估计背景的运动检测方法; 第 2 节提出自适应的前景和背景阈值的选择方法; 第 3 节提出以概率更新背景、背景死锁问题的解决及背景突变时的检测、更新方法; 第 4 节给出实验结果; 最后给出结论. 整个系统框图如图 1 所示。

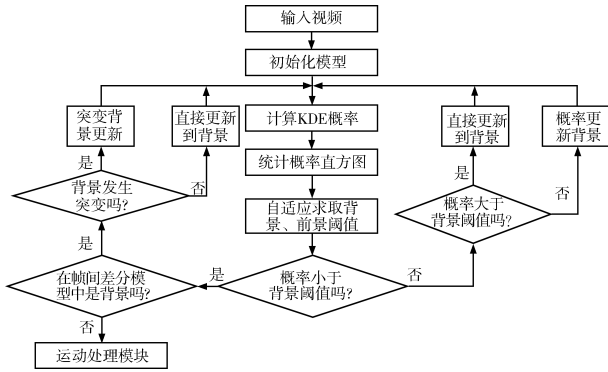


图 1 系统框图

Fig. 1 System architecture

1 核密度估计 (KDE) 运动检测方法

设视频帧中有 M 个像素点, 每个像素点有 N 个背景样本, 则在时刻 t 视频帧中第 i 个像素的像素值为 $\mathbf{x}(t)_i$, 该像素对应的第 j 个背景样本的像素值为 $\mathbf{x}(t)_{i,j}$, 则时刻 t 像素 i 的概率 $P(\mathbf{x}(t)_i)$ 可以通过下式进行估计

$$P(\mathbf{x}(t)_i) = \frac{1}{N} \sum_{j=1}^N K(\mathbf{x}(t)_i - \mathbf{x}(t)_{i,j}) \quad (1)$$

其中 K 为核估计子, 假设 K 服从正态分布, 对于 RGB 图像可以取 R、G、B 分量做为特征值, 如果它们相互独立, 则 N 个样本的概率之和为

$$P(\mathbf{x}(t)_i) = \frac{1}{N} \sum_{j=1}^N \prod_{m=1}^d \frac{1}{\sqrt{2\pi}\sigma_{i,m}} e^{-\frac{(\mathbf{x}(t)_{im} - \mathbf{x}(t)_{i,j})^2}{2\sigma_{i,m}^2}} \quad (2)$$

其中, d 为像素的特征维数, $\mathbf{x}(t)_{im}$ 为第 i 个像素的第 m 个特征值, $\sigma_{i,m}$ 为第 m 个特征核宽. 采用式 (2) 得到的概率如果满足

$$P(\mathbf{x}(t)_i) > T_f \quad (3)$$

则 $\mathbf{x}(t)_i$ 为前景, 其中 T_f 为整幅图像下的全局阈值.

如式 (3) 所示, 采用单阈值进行分类会带来如下问题: 如果要降低误报率必定会导致漏报率的增加;

反之, 如果要降低漏报率则会导致误报率的增加, 因此用单阈值很难协调这对矛盾. 根据误报率调整阈值, 需要场景的先验知识, 而当场景改变时, 需要人工干预. 本文从分析像素概率直方图出发, 给出了一种自适应前景、背景阈值的双阈值选择方法, 用于像素的分类.

文献 [11] 在更新背景模型时采用 Long-term 模型和 Short-term 模型, 在 Long-term 模型中采用盲目更新机制, 即不管分类的结果, 以速率 W/N 把采样直接更新到该模型中去, W 为时间窗口的大小, N 为背景样本的采样数目; 在 Short-term 模型中采用选择更新机制, 即只有分类为背景的样本才包括在该模型中. 采用 Short-term 模型, 如果检测结果不正确, 会导致负样本被模型预测为正样本 (False positives); 而采用 Long-term 模型, 由于不是最近的背景模型, 会导致负样本被模型预测为正样本, 由于目标像素可能出现在模型中, 会导致正样本被模型预测为负样本 (False negatives). 为了克服上述缺点, 文章采用来自于两个模型的交集 $M(\mathbf{x}(t)_i)$

$$M(\mathbf{x}(t)_i) = \begin{cases} 1, & \text{若 } \mathbf{x}(t)_i \in M_s \text{ 且 } \mathbf{x}(t)_i \in M_l \\ 0, & \text{否则} \end{cases} \quad (4)$$

其中, M_s 为 Short-term 背景模型, M_l 为 Long-term 背景模型, 当式 (4) 取 1 时才进行更新. 尽管采用式 (4) 能够克服上述模型各自的缺点, 但是会抑制 Short-term 模型中的被模型预测为正的样本 (True positives) 和 Long-term 模型中的 False negatives. 文章采用在两者交集中像素邻域内的像素, 如果同样包括在 Short-term 模型中, 则这些邻近的像素也包括在最终的背景模型中. 此背景更新机制, 对每个像素点建立两个模型, 需要耗费大量的内存, 而消除式 (4) 所带来的不利影响也会增加运算量.

2 自适应核密度估计 (KDE) 运动检测算法

2.1 概率直方图分析

按照文献 [11] 中的方法估计核宽 σ_i , 当核宽的最大可能取值 σ_{\max} 和最小可能取值 σ_{\min} 确定后, 在时刻 t 每个像素的 N 个采样的概率之和的可能最大值 P_{\max} 可由下式确定

$$P_{\max} = \frac{N}{(\sqrt{2\pi}\sigma_{\min})^d} \quad (5)$$

N 个采样的概率之和的可能最小值 P_{\min} , 不仅取决于 σ_{\max} , 还由式 (2) 中的指数项决定, 可以将其值取为 0. 将概率乘以比例因子 β 量化为 $[0, \beta P_{\max}]$, 其中 $\beta > 1$, 然后, 遍历所有像素的概率, 形成概率

分布直方图 H_p

$$H_p(j) = \begin{cases} H_p(j) + 1, & \text{若 } \beta P(\mathbf{x}(t)_i) = j \\ H_p(j), & \text{否则} \end{cases} \quad (6)$$

图 2(a) 为 ATON^[14] 中的视频 “Campus” 的第 402 帧, 场景中的运动目标为行人; 图 2(b) 为图 2(a) 的概率分布; 图 2(c) 为估计得到的概率图像; 其中 X 、 Y 轴分别对应图像的宽度和高度, Z 轴表示像素的概率; 图 2(d) 为图 2(a) 的概率直方图, 其中 U 轴表示像素的概率值取值, V 轴表示概率等于该取值的像素数目。

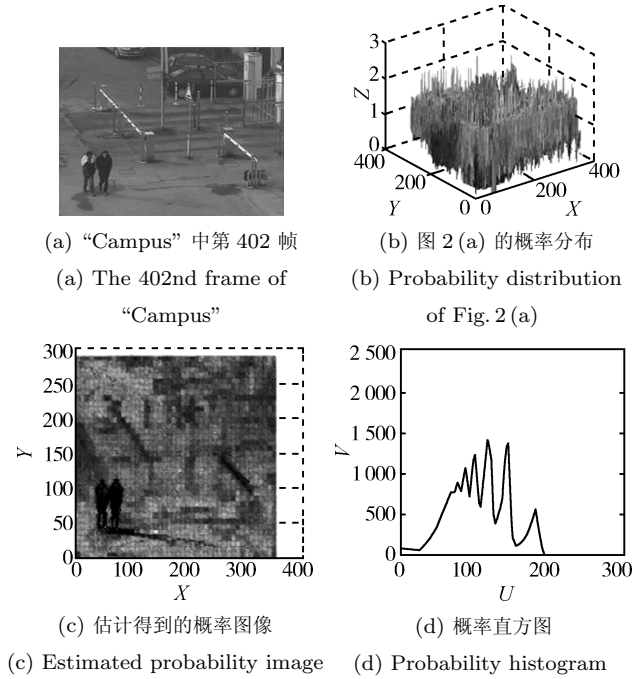


图 2 像素的概率分布及其直方图

Fig. 2 Probability distribution and its histogram

对照图 2(a)~(c) 不难发现: 运动目标和背景的差异较大, 概率值较小, 且属于该区域的像素极为相似; 而背景区域像素概率值较高, 且大部分背景概率介于同一范围。属于背景像素的概率分布存在一定的差异性, 这种差异性反映了背景的复杂程度, 背景越复杂, 差异越明显, 如图 2(d) 中的局部峰值数目反映了背景的复杂性, 峰值的大小反映了属于该值的像素数目的差异性。

2.2 自适应前景阈值的选择

为了便于分析, 用三线性插值对 H_p 进行平滑得到 H_{ps} 。对相邻的 bin 按照下式进行差分, 在此用小阈值 T_s 来进一步去除干扰, 得到差分直方图 H_d

$$H_d(j) = \begin{cases} d, & \text{若 } d > T_s, d = |H_{ps}(i) - H_{ps}(i+1)| \\ 0, & \text{否则} \end{cases} \quad (7)$$

如图 3 所示, 图 3(a) 为图 2(d) 的局部放大; 图 3(b) 为图 2(d) 对应的差分直方图; 图 3(c) 为图 3(b) 的局部放大, 其中水平轴 U 表示概率的取值, 垂直轴 V 表示属于该概率的像素的数目; 图 3(d) 为选取差分直方图的第一个零点 (图 3(c) 中的 P_L) 作为阈值的分割结果。如图 3(a) 所示, 概率直方图根据变化趋势, 被分成 3 段: A 段出现峰值并迅速下降, 直方图 bin 数目相差很大; B 段直方图变化不明显, 直方图 bin 数目相差很小; C 段直方图曲线开始上升, 达到局部峰值后出现多个局部峰值。

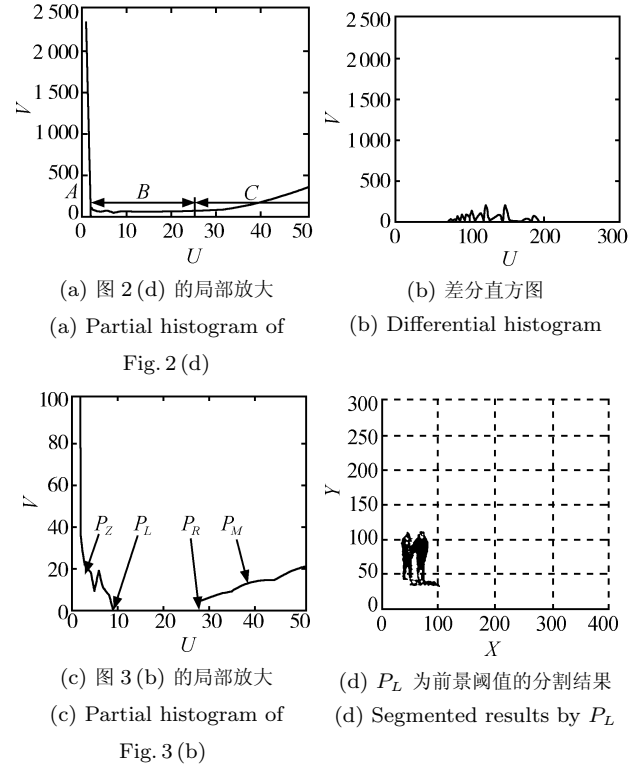


图 3 概率直方图、差分直方图和分割结果

Fig. 3 Probability and differential histogram and segmented results

从第 2.1 节中的分析可知, 属于 A 段的像素概率低, 应当属于前景。 C 段像素概率较高, 属于背景像素, B 段介于两者之间, 处于由一种状态向另一种状态的过渡, 这里对应于由前景到背景之间的过渡过程, 如原来背景改变为运动目标, 或运动目标变为背景的一部分。为此如下选取前景阈值 T_f :

1) 通过差分直方图 H_d 寻找直方图由陡变缓的转折点 P_Z , 满足下式

$$P_Z = \arg \max_i \left| \frac{H_d(i-1) - H_d(i)}{H_d(i) - H_d(i+1)} \right| \quad (8)$$

2) 在差分直方图 H_d 上从 P_Z 开始向右侧寻找变化趋势变缓的起点 P_L ;

3) 在差分直方图 H_d 上从 P_L 开始向右侧寻找缓慢变化的终点 P_R , 直到 P_M (为前景的概率最大值, 如可由目标在场景中的最大比例得到) 止;

4) 当 T_f 取 P_L 时会漏掉部分前景目标, 当取 P_R 时会把部分背景误检测为前景, 为此做如下处理

$$T_f = P_L + \frac{P_R - P_L}{\alpha} \quad (9)$$

其中 α 是大于 1 的系数, 引入 $(P_R - P_L)/\alpha$ 是为了降低漏报率, 同时减少检测误差.

图 4 为取不同概率范围的分割结果, 图 4(a) 为概率满足 $P(\mathbf{x}(t)_i) < T_f$ 时的分类的结果, 运动目标(行人)漏检测像素很少, 有少量属于背景的像素被当作前景来检测; 图 4(b) 为 $T_f \leq P(\mathbf{x}(t)_i) < P_R$ 时的结果, 原来由运动到静止的目标(栏杆)也被检测出来, 且存在较多噪声, 即对应上述分析的 B 段; 图 4(c) 为 $P(\mathbf{x}(t)_i) > P_R$ 时的分割结果, 从图中可以看出只有少量属于前景的像素被检测为背景; 图 4(d) 为 $P(\mathbf{x}(t)_i) \leq P_R$ 时的分割结果, 运动目标(行人)、原来静止的栏杆、部分属于背景的像素被当作前景来检测.

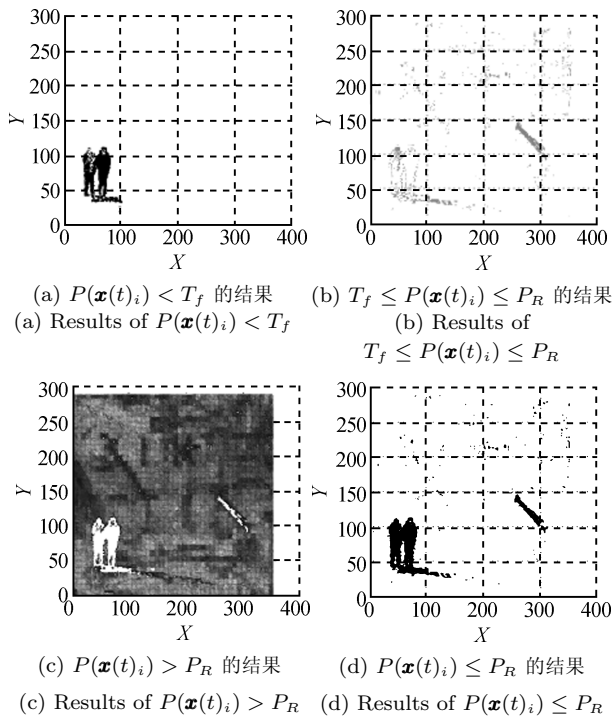


图 4 像素的概率分布及其直方图

Fig. 4 Probability distribution and histogram of pixels' probability

2.3 自适应背景阈值的选择

自适应背景阈值的选取是为了减少把前景错分为背景的概率, 从而有利于合理更新背景. 根据第 2.2 节的分析, 可知道 B 段的像素对应于由前景到

背景之间的过渡过程, 因此把这部分像素立即作为背景更新是不合适的, 为此本文将 P_R 作为背景阈值

$$T_b = P_R \quad (10)$$

图 4(c) 为取背景阈值时的分割结果, 其中除白色区域外的其他区域均为背景, 可以看出判断为背景的区域更接近于真实背景.

3 背景模型的更新

3.1 基于概率的背景更新模型

由式 (2) 得到的像素概率, 实际上反映了该像素和背景的相似程度. 因此, 可以把概率值作为背景更新的依据. 考虑在时刻 t 概率 $P(\mathbf{x}(t)_i)$, 显然如果 $P(\mathbf{x}(t)_i)$ 小于前景阈值 T_f , 像素 $\mathbf{x}(t)_i$ 为前景, 在背景更新时不作考虑; 如果 $P(\mathbf{x}(t)_i)$ 大于背景阈值 T_b , 像素为背景, 可以直接把该像素更新到背景模型中去; 概率介于两者之间的像素, 以新概率进行更新

$$P_b(\mathbf{x}(t)_i) = \begin{cases} 1, & \text{若 } P(\mathbf{x}(t)_i) > T_b \\ 0, & \text{若 } P(\mathbf{x}(t)_i) < T_f \\ \frac{P(\mathbf{x}(t)_i) - T_f}{T_b - T_f}, & \text{否则} \end{cases} \quad (11)$$

在时刻 t 背景模型中第 i 个像素的第 j 个样本按照先入先出的原则进行更新

$$\mathbf{x}(t)_{i,j} = P_b(\mathbf{x}(t)_i) \cdot (\mathbf{x}(t)_i) + (1 - P_b(\mathbf{x}(t)_i)) \cdot (\mathbf{x}(t-1)_{i,j}) \quad (12)$$

采用上式更新可以对属于背景的像素进行可靠更新, 对于属于前景而误检测为背景的像素以概率进行更新, 而属于前景的像素不作更新, 从而避免了污染背景, 并且能够反映背景的缓慢变化.

如图 4 所示, 图 4(a) 中的像素为前景, 因此不会更新到背景中去; 图 4(c) 中除白色区域以外的像素, 直接以概率 1 进行更新; 图 4(b) 中黑色像素对应于由运动到静止的栏杆和部分阴影及噪声像素, 这部分像素以概率 $(P(\mathbf{x}(t)_i) - T_f)/(T_b - T_f)$ 更新到背景中去, 因此能够得到合理的背景.

3.2 背景死锁的解决

设 $I_C(i)$ 、 $I_L(i)$ 、 $I_B(i)$ 分别为像素 i 的当前帧、前一帧和背景的像素值, 不同于文献 [4], 本文除了利用相邻帧间差分的信息, 还利用了当前帧和背景帧之间的信息, 采用如下公式来统计像素未发生变化的连续帧数

$$C(i) = \begin{cases} C(i) + 1, & \text{若 } |I_C(i) - I_L(i)| < T_1 \\ & \text{且 } |I_C(i) - I_B(i)| \geq T_1 \\ 0, & \text{否则} \end{cases} \quad (13)$$

其中 $C(i)$ 为第 i 个像素没有发生变化的帧数, T_1 为一个小阈值, 由图像的噪声决定. 当帧计数 $C(i)$ 满足式 (14), 就认为是背景, 需要以当前像素值来更新对应位置的背景像素值, 其中 T_2 为更新率

$$I_B(i) = \begin{cases} I_C(i), & \text{若 } C(i) > T_2 \\ I_B(i), & \text{否则} \end{cases} \quad (14)$$

当背景发生局部改变或由于分类错误时, 将导致在 KDE 中把属于背景的像素误检测为前景. 但在帧间差分模型, 对应的像素为背景, 可以利用这种特性来解决背景更新中的死锁, 如式 (15) 所示

$$M_B(\mathbf{x}(t)_i) = \begin{cases} 1, & \text{若 } \mathbf{x}(t)_i \in I_B \text{ 且 } P(\mathbf{x}(t)_i) < T_f \\ 0, & \text{否则} \end{cases} \quad (15)$$

其中 $M_B(\mathbf{x}(t)_i)$ 为 1, 表示 $\mathbf{x}(t)_i$ 像素在帧间差分模型中为背景, 在 KDE 检测中为前景, 此时, 可以按照式 (12) 进行更新, 其中更新的概率 $P_b(\mathbf{x}(t)_i)$ 取 1. 采用帧间差分模型只是利用帧间的差分、当前帧和背景帧的信息, 相比较文献 [11] 中的方法, 耗费内存要少, 运算量相对要小. 由于只有属于背景的像素才会更新到 KDE 的背景样本中, 能够解决背景更新中的死锁问题. 由于前景目标不会出现在模型中, 因此能减少 False negatives 的数目.

3.3 突变背景的检测

应用第 3.2 节的方法虽能更新背景模型, 但是当背景发生突变时, 需要快速地响应这种变化, 本文仍利用第 3.2 节的结论来检测背景是否发生了突变. 当背景发生突变时, 通过式 (15) 检测到的像素, 表现为区域性, 面积较大. 对所有突变像素进行连通域分析, 提取连通域的面积、个数、最大长度、最大宽度等特征. 当这些特征满足预先设定的背景突变的条件时, 才认为背景确实发生了突变, 用当前帧开始后的 N 帧图像更新背景模型中的样本, 保证背景模型得到及时更新.

4 实验结果及分析

实验环境为 CPU P4/3.0 GHz、内存 512 MB. 选择 ATON^[14] 中的视频数据为测试样本, 图像宽为 320 像素, 高为 240 像素的灰度图像, 背景样本为 50 帧, 处理帧率大于 15 fps.

实验中采用如下指标: 假正率 (False positive rate, FPR) 和假负率 (False negative rate, FNR)^[15] 来评价本文算法的性能, 用 FP 表示被模型预测为正的负样本, FN 表示被模型预测为负的正样本, TP 表示被模型预测为正的样本, TN 表示被模型预测为负的背景样本, 则 $FPR = FP/(FP + TN)$,

$$FNR = FN/(TP + FN).$$

实验中从 “Laboratory” 视频序列中随机选择有运动目标的 20 帧视频, 为了增加可比性, 把本文方法得到的假正率 $FPR_a(m)$ 的平均值取作采用单分类阈值得到的 $FPR_f(m)$, 即固定 $FPR_f(m)$, 统计采用单阈值方法得到的 $FNR_f(m)$, 其中 $m = 1, \dots, 20$ 随机选择的视频帧.

图 5 为采用单阈值和自适应阈值的 FNR 和 FPR 曲线, 其中水平轴表示帧序号, 纵轴表示 FNR 和 FPR 的百分率. 从图中的曲线可以看出当 FPR 平均值相同时, 本文算法得到的 FNR 均小于采用单阈值进行分类的 FNR , 这是因为本文算法采用双阈值较好地解决了误报率和漏报率之间的矛盾, 从而取得较理想的检测性能. 本文算法把背景误检测为前景的像素数目要低于固定阈值下像素数目, 说明背景阈值选择不会给背景更新带来较大误差. 同样对于 ATON 中的其他视频序列, 实验都能得到类似的结果.

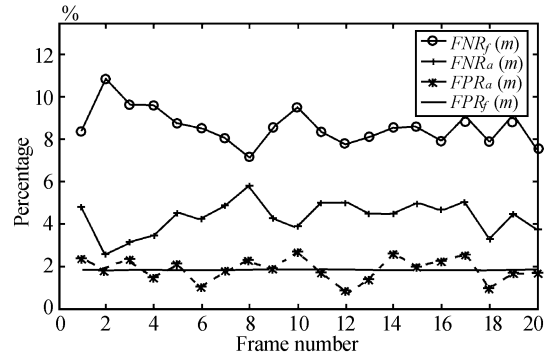


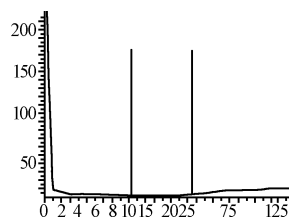
图 5 采用单阈值和自适应阈值的 FNR 和 FPR 曲线
Fig. 5 Curves of FNR and FPR at fixed and adaptive thresholds

图 6 (见下页) 给出了背景、前景检测结果及背景更新的结果. 图 6(a) 为 “Laboratory” 中第 152 帧, 图 6(b) 为图 6(a) 的概率直方图, 图 6(c) 中除去黑色像素, 其余的均为背景, 图 6(d) 为前景检测结果, 白色像素为前景, 图 6(e) 为第 152 帧时采用本文算法得到的背景. 图 6(b) 中从左往右第一条垂线与水平轴的交点为前景分割阈值, 第二条垂线与水平轴的交点为背景分割阈值. 结合图 6(c) 和图 6(d), 在图 6(c) 中有少量属于背景的像素没有判断成背景, 但从图 6(d) 中可以看出这些像素也没有被当作前景, 而是用概率更新到背景中去, 由于这些像素和背景更接近, 由式 (11) 计算得到的背景更新概率较大, 能够较快更新到背景中去, 从而得到较为理想的背景. 在图 6(d) 中有些属于前景的像素未被检测成前景, 但由于这些像素和前景的概率更接近, 会以较小的概率更新到背景中去, 即使存在误检测, 也需要较长时间, 才会更新到背景中去, 从而最大程度

地避免了背景被前景污染,如图 6(e) 所示,背景更接近于真实背景.



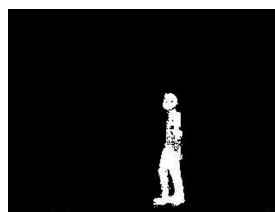
(a) “Laboratory” 中第 152 帧
(a) The 152nd frame of
“Laboratory”



(b) 图 6(a) 的概率分布
(b) Probability distribution of
Fig. 6(a)



(c) 背景分割结果
(c) Segmented background



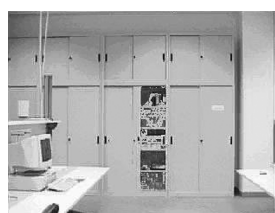
(d) 前景分割结果
(d) Segmented foreground



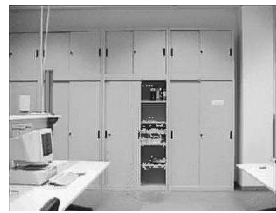
(e) 第 152 帧对应的背景
(e) Background of 152nd
frame



(f) “Laboratory” 中第 360 帧
(f) The 360th frame of
“Laboratory”



(g) 在 500 帧用文献 [11] 得到的背景
(g) Background of using [11]
at 500th frame



(h) 在 500 帧用本文算法的背景
(h) Background of using
proposed method at 500th
frame

图 6 前景和背景阈值的分割结果及背景更新中死锁的解决
Fig. 6 Results of segmentation using foreground and
background thresholds and the solution of deadlock
situations in background model

图 6(f)~(h) 用来说明本文的算法能够解决背景更新中的死锁问题. 其中图 6(f) 为 “Laboratory” 中第 360 帧, 图 6(g) 和图 6(h) 分别为 500 帧时用文献 [11] 和本文算法得到的背景. 从第 360 帧开始, 原来关闭的门被打开, 此时如果只根据分类结果进

行更新, 则门打开后留下的空洞区域会一直当作前景, 造成误检测. 图 6(g) 为加入盲目更新后得到的背景, 实验中更新率选为 $1/2$, 从图中看出部分前景像素会被作为背景像素更新到背景中去. 图 6(h) 为采用帧间差分模型和 KDE 检测结果相结合后得到的背景更新结果, 用帧间差分模型很容易得到场景变化后的背景, 当像素被 KDE 检测为前景, 在帧差模型中为背景时, 该像素直接被更新到背景模型中, 因此可以解决背景更新中的死锁问题. 从图 6(g) 中看出采用盲目更新会把前景更新到背景中去.

为了验证本文提出的突变背景检测、更新算法, 实验选用的视频为高速公路场景, 拍摄时摄像机角度调整频繁、调整范围较大. 图 7(a) 为第 3053 视频帧, 镜头未偏移, 图 7(b) 和图 7(c) 分别为第 3124, 3280 视频帧, 此时镜头已经发生偏移, 图 7(d) 为在 3280 帧时刻用帧间差分模型得到的背景, 图 7(e) 中的白色区域表示在 KDE 中检测为前景, 而在图 7(d) 中被认为是背景, 并经过连通域分析得到背景



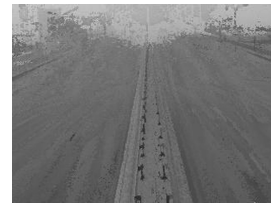
(a) 第 3053 帧
(a) The 3053rd frame



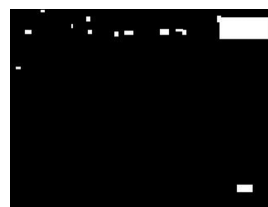
(b) 第 3124 帧
(b) The 3124th frame



(c) 第 3280 帧
(c) The 3280th frame



(d) 第 3280 帧时帧间差分背景
(d) Background of inter-frame
difference at 3280th frame



(e) 第 3280 帧时在 KDE 中为前景、帧间差分模型为背景的像素
(e) Pixels belong to background
in KDE, but in inter-frame
difference model are foreground



(f) 本文算法得到的第 3400 帧时的背景
(f) Background at 3400th
frame by the proposed
method

图 7 突变背景的检测与更新
Fig. 7 Detection and updating of suddenly
changed background

突变区域. 图 7 (f) 为背景突变检测后, 重新初始化 KDE 背景样本, 更新后得到的背景. 在此只是利用了帧间差分模型中的背景, 因此能够可靠检测背景突变的发生.

5 结论

本文提出了一种自适应的核密度估计运动检测方法. 该方法能够得到更为合理的背景模型, 有利于运动目标的检测. 此外, 本文还利用帧间差分的快速性, 结合 KDE 检测结果, 较好地解决了背景突变的检测、更新问题. 实验证明了本文算法的可靠性和适应性. 但是 KDE 算法计算相对复杂, 本文在计算核宽时, 利用了上次的计算结果, 降低了时间复杂度; 另外帧间差分背景模型的引入, 也使得运算时间有所降低.

References

- Piccardi M. Background subtraction techniques: a review. In: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics. Hague, Netherlands: IEEE, 2004. 3099–3104
- Wren C R, Azarhayejani A, Darrell T, Pentland A P. Pfunder: real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, **19**(7): 780–785
- Lo B P L, Velastin S A. Automatic congestion detection system for underground platforms. In: Proceedings of International Symposium on Intelligent Multimedia, Video, and Speech Processing. Hong Kong, China: IEEE, 2001. 158–161
- Chien S Y, Ma S Y, Chen L G. Efficient moving object segmentation algorithm using background registration technique. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, **12**(7): 577–586
- Li Gang, Qiu Shang-Bin, Lin Ling, Zeng Rui-Li. New moving target detection method based on background differencing and coterminous frames differencing. *Chinese Journal of Scientific Instrument*, 2006, **27**(8): 961–964
(李刚, 邱尚斌, 林凌, 曾锐利. 基于背景差法和帧间差法的运动目标检测方法. 仪器仪表学报, 2006, **27**(8): 961–964)
- Wang Xiao-Mei, Wang Yang-Li, Niu Ping-Hong. Gait detection and recognition based on adaptive background model. *Application Research of Computers*, 2006, **23**(11): 258–261
(王晓梅, 王养利, 牛平宏. 基于自适应背景模型的步态检测与识别. 计算机应用研究, 2006, **23**(11): 258–261)
- Colombari A, Fusiello A, Murino V. Segmentation and tracking of multiple video objects. *Pattern Recognition*, 2007, **40**(4): 1307–1317
- Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking. In: Proceedings of the Computer Society on Computer Vision and Pattern Recognition. Fort Collins, USA: IEEE, 1999. 246–252
- Zuo Jun-Yi, Pan Quan, Liang Yan, Zhang Hong-Cai, Cheng Yong-Mei. Model switching based adaptive background modeling approach. *Acta Automatica Sinica*, 2007, **33**(5): 467–473
(左军毅, 潘泉, 梁彦, 张洪才, 程咏梅. 基于模型切换的自适应背景建模方法. 自动化学报, 2007, **33**(5): 467–473)
- Oliver N M, Rosario B, Pentland A P. A Bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 831–843
- Elgammal A M, Hanwood D, Davis L S. Non-parametric model for background subtraction. In: Proceedings of the 6th European Conference on Computer Vision. Dublin, Ireland: Springer-Verlag, 2000. 751–767
- Mittal A, Paragios N. Motion-based background subtraction using adaptive kernel density estimation. In: Proceedings of the Computer Society on Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2004. 302–309
- Wang H, Suter D. A consensus-based method for tracking: modeling background scenario and foreground appearance. *Pattern Recognition*, 2007, **40**(3): 1091–1105
- Autonomous agents for on-scene networked incident management (ATON). Shadow detection [Online], available: <http://cvrr.ucsd.edu/aton/shadow/index.html>, June 1, 2007
- Georis B, Brémond F, Thonnat M, Macq B. Use of an evaluation and diagnosis method to improve tracking performances [Online], available: <http://www-sop.inria.fr/orion/Publications/Articles/viip03.pdf>, June 1, 2007



徐东彬 中国科学院自动化研究所博士研究生. 主要研究方向为视频行为识别和智能监控. 本文通信作者.

E-mail: dongbin.xu@ia.ac.cn

(**XU Dong-Bin** Ph.D. candidate at the Institute of Automation, Chinese Academy of Sciences. His research interest covers behavior recognition and

intelligent surveillance for security. Corresponding author of this paper.)



黄磊 中国科学院自动化研究所副研究员. 主要研究方向为文字识别, 人脸识别和智能监控.

E-mail: lei.huang@mail.ia.ac.cn

(**HUANG Lei** Associate professor at the Institute of Automation, Chinese Academy of Sciences. His research interest covers character recognition, face recognition, and intelligent surveillance for security.)



刘昌平 中国科学院自动化研究所研究员. 主要研究方向为文字识别, 人脸识别和智能监控.

E-mail: changping.liu@mail.ia.ac.cn

(**LIU Chang-Ping** Professor at the Institute of Automation, Chinese Academy of Sciences. His research interest covers character recognition, face recognition, and intelligent surveillance for security.)