

基于运动行为的复杂监控事件探测

梁浩哲¹, 王振华², 徐树奎¹, 李国辉¹

(1. 国防科技大学 系统工程系, 湖南 长沙 410073; 2. 上海海洋大学 信息学院, 上海 201306)

摘要: 针对视频监控过程, 使用运动目标的状态特征描述场景中存在的语义内容. 基于 DBSCAN 聚类模型学习特征集的潜在结构, 生成了运动行为模式集. 使用高级 Petri 网刻画模式间的连续、并发等时序关系, 构成复杂语义事件探测模型. 无监督式的模式学习过程对低层噪声有较强的鲁棒性, 而定性的事件描述模型对于高层事件的推理具有更强的灵活性. 在实验中, 通过聚类学习得到的行为模式, 给出了事件 Petri 网的具体建模过程, 并演示了“停留”与“偷车”两个感兴趣事件的探测结果.

关键词: 视频监控; 行为理解; DBSCAN 聚类; Petri 网

中图分类号: TP391

文献标识码: A

Complex Surveillance Events Detection Based on Motion Activity

LIANG Haozhe¹, WANG Zhenhua², XU Shukui¹, LI Guohui¹

(1. Department of Engineering System, National University of Defense Technology, Changsha 410073, China; 2. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China)

Abstract: For visual surveillance, the semantic content of video was modelled by the states of motion targets. Feature vectors were clustered based on DBSCAN to obtain activity patterns which represent potential structure of training set. Then a complex events detection model was investigated by utilizing high-level Petri nets to model the temporal dependence relationship of activity patterns, like continuum and concurrence. The unsupervised learning process of activity patterns was robust to low-level noise. By combining Petri-framework representation, inference of semantic events could be more flexibly. In the experiments, Petri nets modeling process was demonstrated based on the results of clustering and the validity was given by the detection of two interesting semantic events ‘staying’ and ‘stealing’.

Key words: video surveillance; activity understanding; DBSCAN; Petri net

面向监控视频的语义事件探测技术是目前的研究热点^[1-2]. Hamid^[3] 使用 n -gram 模型表示事件序列, 利用图聚类序列集而得到潜在事件结构; Nevatia 提出了监控视频领域的本体模型^[4]; Lavee 等^[4] 将本体语义转译为 Petri 网模型, 完成事件识别与推理. 目前研究多采用手动标注形成场景事件训练集, 通过事件的低层特征描述探测场景中的简单事件. 与新闻视频与体育视频不同, 监控视频事件不具备明显的单元结构性, 很难做到精细的事件标注, 因此需要一种主动事件学习模型.

监控场景中, 运动目标的行为状态是描述语义事件的重要信息线索, 通过分析状态特征间的关联可实现对事件要素的自学习. Xiang 等^[5-6] 使用 Pixel Change History 模板描述行为, 基于 Gaussian 混合模型聚类得到场景中的基本事件集; Cedric 等^[7] 使用决策树对轮廓特征量进行聚合分析, 实现场景事件识别. 在上述研究中, 通过特征挖掘事件属于一种低层数据驱动的方式, 对噪声影响具有较强的鲁棒性, 但是相应的“自下而上”的方式在事件时空语义描述方面存在较差的伸缩性, 而透过“专家”的领域知识灵活建立事件描述模型可以弥补以上缺陷.

针对以上问题, 本文提出一种监控视频的复杂事件探测模型, 通过聚类学习得到场景的行为模式, 以其作为语义事件的组成要素; 其后使用 Petri 网构建基于运动行为的语义事件时序关系模型. 在实验中, 给出了测试视频中的基本行为模式及其语义表达, 并通过“停留”与“偷车”两个具体的事件案例演示了探测过程的建模有效性与可扩展性, 探测结果表明了模型的有效性.

收稿日期: 2011-03-11

基金项目: 国家自然科学基金项目(61170158)

第一作者: 梁浩哲(1983—), 男, 工学博士, 主要研究方向为机器视觉信息处理工程. E-mail: lianghaozhe@hotmail.com

通信作者: 李国辉(1963—), 男, 教授, 工学博士, 博士生导师, 主要研究方向为计算机视觉集成与应用. E-mail: guohli@nudt.edu.cn

1 复杂监控事件探测模型

针对监控上下文,可使用三个层次刻画其潜在语义结构(图1):运动行为集 $a=\{a_i\}$;行为模式 $A=\{A_i\}$;复杂语义事件 $E=\{E_i\}$.

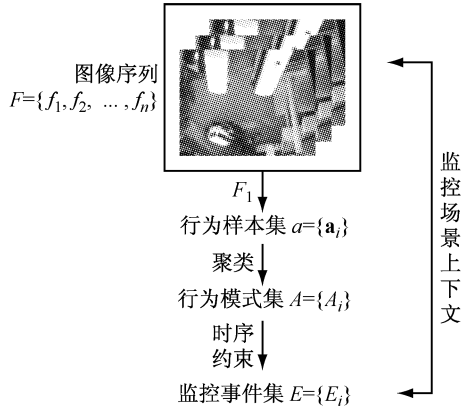


图1 事件探测模型

Fig.1 Events detection model

本文通过 blob 特征描述行为矢量^[7]:

$$a_i := (x_c, y_c, s_h, s_w, V, u, v)$$

其中: x_c, y_c 为 blob 质心坐标; s_h, s_w 为检测框尺寸; V 为速率; u, v 为方向. 在环境约束条件下,行为按一定规则运动,因此 a_i 间存在较强的特征关联. 基于此假设,事件探测模型由双层语义关系构成(图1).

对行为样本集 a 聚类:

$$a_i \rightarrow A_i \quad (1)$$

式中: $a_i \in a, A_i \in A$, 构成行为模式集 $A = \{A_i\}$, 回答监控场景中“有什么”的问题.

以 A 为组成要素,按照特定时序关系构成语义事件集 $E=\{E_i\}$,存在模式间的时序约束:

$$A' = \{A_i\} \rightarrow E_i \quad (2)$$

式中: $A' \subset A$, 为模式子集; $E_i \in E$, 解决模式间“关联”问题. 对应式(1), (2), 事件探测过程分为两步: ① 使用 DBSCAN 聚类模型学习行为样本集 $a=\{a_i\}$ 的潜在结构, 形成行为模式表; ② 使用 Petri 构建模式间的连续、并行等时序关系, 构成语义事件集 $E=\{E_i\}$, 并完成事件探测.

2 基于 DBSCAN 聚类的行为模式学习

2.1 行为特征提取

对于视频序列,使用平均背景建模加光流场运动估计^[8]完成行为特征提取. 目标矢量为

$$a_i := (x_c, y_c, s_h, s_w, \sqrt{u^2 + v^2}, u, v)$$

步骤1: blob 位置与尺寸提取. 前景对象条件:

$$f(x, y, t) \in [B(x, y, t) - 2.5 \sqrt{\sigma(x, y, t)}, B(x, y, t) + 2.5 \sqrt{\sigma(x, y, t)}]$$

其中: t 为当前时刻; B 为累积背景均值; f 为当前帧; σ 为背景累积方差. 经过形态学运算并四连通域分析,得 blob 集: $b_i = \{b_{i1}, b_{i2}, \dots, b_{in}\}$ (如图2,第2列所示). 针对 b_{in} , 尺寸框为

$$x_c = [\max(x) + \min(x)]/2 \quad (3)$$

$$y_c = [\max(y) + \min(y)]/2 \quad (4)$$

$$s_h = \max(y) - \min(y) \quad (5)$$

$$s_w = \max(x) - \min(x) \quad (6)$$

式(3)~(6)中, $\forall x, y \in b_{in}$. 图2,第3列为尺寸检测框;图2,第4列为绿色质心点.

步骤2: 速度估计. 使用 h 帧平均速度:

$$V = \frac{1}{h} \| \text{centriod}(b_{in}, t+h) - \text{centriod}(b_{in}, t) \| \quad (7)$$

式中, $\text{centriod}(\cdot, t)$ 为 t 时刻 blob 质心矢量, $\|\cdot\|$ 为欧氏范数,实验中使用 $h=8$.

步骤3: 方向估计. 对 b_{in} , 计算其光流值:

$$\text{flow}: b_{in} \rightarrow \{[u, v]_{in}^i\}$$

flow 为 Lukas-kanade 算子, $[u, v]_{in}^i$ 为光流集, 平均方向角为

$$\theta = \arctan \left[\frac{\sum v_{in}}{\sum u_{in}} \right], \quad \theta \in [0, \frac{\pi}{2}) \quad (8)$$

得归一化平均光流方向(如图2,第5列):

$$[u_{\text{mean}}^i, v_{\text{mean}}^i] := \cos \theta + \sin(\theta i) \quad (9)$$

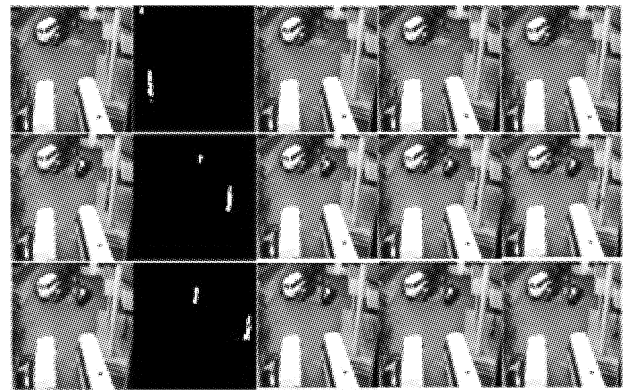


图2 基于 blob 状态矢量提取运动行为特征

Fig.2 Extraction of foreground blob

使用式(3)~(9),计算视频序列 F 得行为样本集 $a=\{a_i\}$.

2.2 运动行为 DBSCAN 聚类

DBSCAN 是一种基于密度的特征矢量聚类模

型^[9],可以解决行为样本集 a 聚类中存在的以下问题:① 模式集 $A=\{A_i\}$ 内部结构无先验知识,类数 n 难以提前确定,DBSCAN 仅需定义样本间距离;② 行为特征的时空分布并不规则,DBSCAN 聚类可形成任意形状分布,而 K-means, Gaussian 混合等聚类法的类圆型聚类形状存在局限性;③ DBSCAN 的孤立点识别具有较强的抗噪能力。

使用 DBSCAN 聚类模型给出行为样本学习模型如下^[9]:

(1) 定义行为 a_i 的状态值 ϵ 邻接域:

$$N_\epsilon(a_i) = \{a_j \in D \mid \text{dist}(a_i, a_j) \leq \epsilon\}$$

其中, $\text{dist}(\cdot)$ 为行为矢量间距离。

(2) 行为 a_i 到 a_j 的直接密度可达(本文称为直接状态可达):

如果 $a_j \in N_\epsilon(a_i) \quad \forall \quad |N_\epsilon(a_i)| \geq \text{MinPts}$

MinPts 为孤立点阈值,与样本点密集程度相关。

(3) 通过 a_i 到 a_j 的状态直接可达的传递性定义状态可达。

(4) a_i 到 a_j 的密度联通(称为状态联通):如果存在 a_k ,使得 a_i 与 a_j 到 a_k 都状态可达。

(5) $A=\{A_i\}$ 模式集定义:

$A_i: \forall a_i, a_j$ 如果 $a_i \in A_i$ 满足 a_i, a_j 的直接可达,则 $a_j \in A_i$ 。

$A_i: \forall a_i, a_j \in A_i$ 满足, a_i 到 a_j 是状态联通。

由于特征分量间存在量纲差异,模型使用层次性 DBSCAN 聚类:

步骤 1: 对分量 x_c, y_c, s_h, s_w, V , 使用 Mahalanobis 距离:

$$d_M(a_i, a_j) = (a_i - a_j)^T \Sigma^{-1} (a_i - a_j) \quad (10)$$

其中, Σ 为协方差阵。

步骤 2: 对 u, v 分量,使用矢量距离, $<, >$ 为内积运算:

$$d_{\text{dir}}(a_i, a_j) = \langle [u_{\text{mean}}, v_{\text{mean}}]_i, [u_{\text{mean}}, v_{\text{mean}}]_j \rangle \quad (11)$$

使用式(10), (11)为距离度量,聚类形成行为模式集 $A=\{A_i\}$ 。

3 基于 Petri 网的复杂监控事件探测模型

模式集 $A=\{A_i\}$ 形成场景基本语义单元,描述了行为样本集的潜在结构,以此为基础,通过定义 A_i 间的时序关系可形成语义事件模型 $E=\{E_i\}$ 。

Petri 网是一种数学建模语言^[10],文献^[11]提出 7 种事件时序关系。如图 3 所示,本文提出一种面向

监控视频的 Petri 事件探测模型(surveillance Petri detection model, SDM-Petri),建立模式间的时序关系。

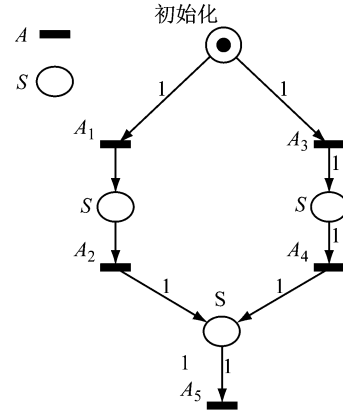


图3 SDM-Petri 模型结构

Fig.3 Structure of SDM-Petri model

构建一张 SDM-Petri 网描述场景中存在的多个语义事件 E_i 。基本要素包括模式节点集 $A=\{A_i\}$, 状态位置 S , 以及模式转移弧: 定义为有向对 (A_i, S) 。在模型构成中, S 代表了当前运动目标集 $O=\{O_i\}$ 的状态值, O_i 等价于 Petri 令牌, 如图 3 中黑点表示。 A 代表模型随 O_i 状态值而产生的变迁。SDM-Petri 的演化规则^[11]简单定义为:

当匹配映射 $\text{fullfill}(\text{attribute}(O_i), A_j) \rightarrow \text{true}$, 则 $\text{enable}(A_i) \rightarrow \text{true}$ 。其中, 函数 $\text{attribute}(O_i)$ 为目标瞬时状态矢量, $\text{fullfill}(a, b)$ 定义为匹配关系, 即 a 满足于 b ; 处于 enable 状态后 A_i 可 fire(点火), O_i 由 A_i 的前置状态转移至后置状态。

SDM-Petri 模型在表示串行模式的同时, 也可以刻画非对称时序关系。如图 3 所示, 模式 A_1, A_2 为顺序关系, A_1 可理解为 A_2 的发生条件; A_1, A_3 存在竞争关系, 当仅为单一目标时, A_1, A_3 表现为非对称式的条件互斥, 而多目标的不同变迁使 A_1, A_3 刻画模式并行; A_2, A_4 实现模式同步; A_5 为终点模式, 表示语义事件 E_i 发生或 O_i 消失。在实验中将结合具体实例演示 SDM-Petri 的构造过程。

4 事件探测模型测试实验

本文实验环境为户外远景视频监控: 帧速为 15 帧 $\cdot \text{s}^{-1}$; 图像分辨率为 320×240 ; 实验的硬件环境是 Pentium(R) 4 CPU 2.40 GHz, 内存 1 GB; 使用平均背景模型收集行为样本集 $a=\{a_i\}$, 时长为 5 h^[12](图 2)。

4.1 行为模式学习

针对行为样本 $a = \{a_i\}$, 图 4 为 DBSCAN 聚类结果. 将数据点映射到 (x_c, y_c) 分量空间进行可视化显示, 行为模式具有明显的区域性, 存在 5 个区域, 其边缘间存在相互重叠, 为直观显示, 使用 5 条分离的闭合曲线显示(图 4a). 表 1 中为区域性模式集 $A = \{A_i\}$ 的简要语义. 图 4b 为各模式的关键帧.

表 1 模式语义说明

Tab.1 Semantic description of activity patterns

模式名	语义短描述
A_1	目标在路径 E 上来回运动
A_2	目标沿路径 D 进出建筑物
A_3	目标在停车区 A 停留
A_4	目标在停车区域 B 停留
A_5	目标在在入口区 C 停留

图 4a 左为式(10)聚类结果. 对 x_c, y_c, s_h, s_w 分量, 由于目标运动的连续性, 以及远景监控下行人的

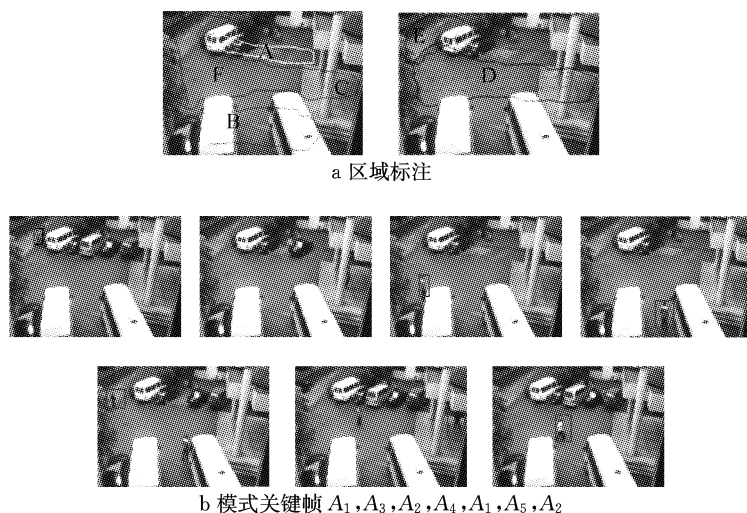


图 4 模式区域标注及模式关键帧

Fig.4 Patterns labeling and key-frame of patterns

表 2 模式集规范化表示

Tab.2 Regularized semantic descriptions of motion patterns

模式名	语义短描述
A_1	目标出现在场景中.
A_2	目标沿路径 D 进出建筑物的匀速运动
A_3	目标进入建筑物
A_4	目标进入停车区
A_5	目标在停车区边缘产生运动
A_6	目标在停车区的停留过程
A_7	目标在停车区的过度停留, 表示为终点模式

针对每一视频帧: ① 定义运动目标集 $O = \{O_i\}$ 的属性值, 见表 3; ② 以及各模式的使能条件, O_i 代表各运动对象的状态值, 见表 4.

近似刚体运动, 样本集 $a = \{a_i\}$ 的 x_c, y_c, s_h, s_w 分量相互状态联通或状态可达; 速度 V 的强分离性将场景划分为两部分. 其中, A, B, C 为停留区; 黑色闭合曲线 F 覆盖了运动区.

图 4a 右使用式(11)进一步分割模式集, u, v 分量将黑色闭合曲线 F 分裂为 E, D 区域(图 4a 右), 各包含两个运动主方向(图 4a 右中直线运动域所示), 近似代表了场景中的运动路径.

4.2 SDM-Petri 事件探测

4.1 表明模式集 $A = \{A_i\}$ 代表了监控场景运动区域语义信息. 针对测试场景(图 4), 基于 SDM-Petri 模型定义目标从道路 D 进入场景发生“偷车”事件(图 5 所示, 以下简称 SDM-Petri 1.1). SDM-Petri 模型规范化模式集 $A = \{A_i\}$, 包括添加起始运动模式, 以及终点运动模式, 见表 2.

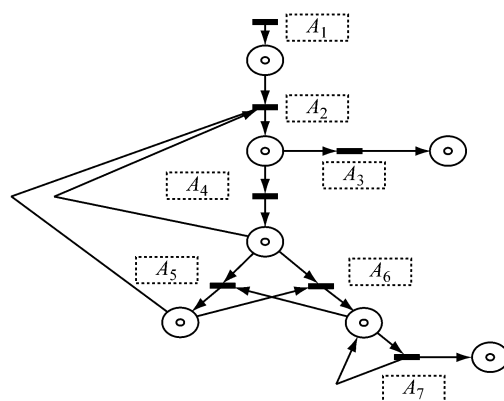


图 5 SDM-Petri 实例 1.1

Fig.5 SDM-Petri 1.1 for complex event detection

表 3 SDM-Petri 目标属性描述

属性名	属性语义说明
IDNew	判断是否为新目标进入路径 D 或消失, 使用跟踪器的 ID 号判别.
ParkZone	判断目标的位置是否进入停车区 $P = A \cup B$ 区域, 通过计算目标 x_c, y_c 分量与 P 边界距离判别.
Velocity	判断目标为静止模式或者运动行模式, 计算式式(7): $V > T_1$
T_1	目标速度阈值.
T_2	目标停车区停留阈值.
T_3	目标非停车区速度阈值.
T_4	目标非停车区停留阈值.

表4 SDM-Petri 1.1 使能条件

模式名	SDM-Petri 1.1 使能条件
A_1	$O_i, IDNew = True.$
A_2	$O_i, ParkZone = False, O_i, V > T_1.$
A_3	$O_i, IDNew = False.$
A_4	$O_i, ParkZone = True.$
A_5	$O_i, ParkZone = True, O_i, V > T_1.$
A_6	$O_i, ParkZone = True, O_i, V < T_1.$
A_7	$O_i, ParkZone = True, O_i, V < T_1, Transition(A_7), Delay > T_2.$

对于模型 SDM-Petri 1.1 的描述:当目标 O_i 出现在道路 D 中,定义属性 IDNew 为 true,触发初始化模式 A_1 ;模型 1.1 同时定义了偷车终点模式 A_7 .通过检测跟踪计算 x_c, y_c 分量,更新 ParkZone 及 IDNew;若 IDNew 为假,触发模式 A_3 ,目标进入建筑物;ParkZone 为真,则触发 A_4 ,目标进入停车场区,此时预警开始,提示目标越界;当目标进入区域 P 后,存在 3 种状态变迁:① 停留对应模式 A_6 ;② 沿停车区与路径 D 重叠边缘区运动,对应模式 A_5 ;③ 沿路径 D 运动,对应模式 A_2 ;通过对 ParkZone 值与 Velocity 值的循环更新;三模式相互可为发生条件.而模式 A_6 触发后, A_7 进入自循环触发,与其他模式不同, A_7 会延迟触发,延迟时长 $t > T_2$ 时,触发代表“偷车”事件发生,否则触发 A_5 ;同时注意到 A_7 具有自反性质,触发后的目标数量代表事件的严重程度;终点模式 A_3 与 A_7 ,触发 A_3 后的状态目标数量可进行人流统计, A_7 预警“偷车”模式.

SDM-Petri 模型可扩展多语义事件,SDM-Petri 1.2 增加了“运动区停留”事件探测,如图 6 所示,添加 A_8, A_9 模式,其语义与使能条件见表 5. A_9 与 A_7 类似,同样为延迟触发,以自循环触发表示事件重要程度.

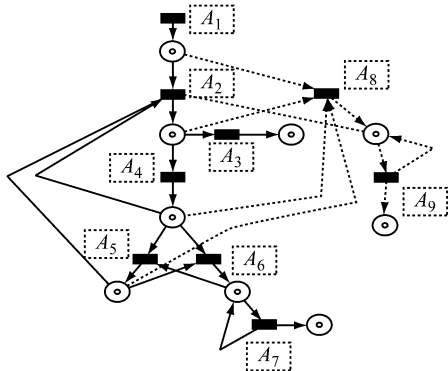


图6 扩展 SDM-Petri 实例 1.2

Fig.6 Extended SDM-Petri 1.2

4.3 实时监控事件探测结果

使用 SDM-Petri 1.2,图 7 为“运动区停留”事件

与“偷车”事件的探测结果.

表5 模式 A_8 与 A_9 描述与使能条件

模式名	语义描述	SDM-Petri 1.2 使能条件
A_8	目标在运动区的停留.	$O_i, ParkZone = False, O_i, V < T_3$
A_9	目标在运动区域的过度停留,为事件终点模式.	$O_i, ParkZone = False, O_i, V < T_3, Transition(A_9), Delay > T_4$



图7 事件探测结果

Fig.7 Experiment result for event detection of SDM-Petri 1.2

图 7 第 2 行,为“停留”事件探测,左图提示目标运动速度低于阈值,检测框的颜色从左至右逐渐变深,指示停留时间增长,可自动探测到运动区的静止目标.图 7 第 1,3 行为两个停车区的“偷车”事件探测,左图提示目标进入停车区,相似的,检测框的颜色渐变指示停留.图 7 第 3 行右图表示目标速度大于阈值,停止触发 A_7 同时转向模式 A_5 ,结束预警.

在表 6 中,给出了“偷车”事件模型的实时探测结果.为了验证性能,首先将停留阈值设定为 2 s,模型预警了全部进入停车区的运动对象,正确率接近 95%;其次,将阈值设为 5 s,在实验过程中加入异常停留行为,模型预警到全部异常行为;SDM-Petri 模型可满足实时监控预警需要.

表6 “偷车事件”探测结果

停留阈值 T_3/s	实际预警数/个	正确预警数/个
2	58	55
5	6	6

5 结语

本文研究了一种监控场景中的高层语义事件探测技术,针对运动行为样本,给出一种 Dbscan 行为模式自动学习方法;以此为基础,使用 SDM-Petri 建模其时序关系,完成事件探测.通过实时监控视频演

示了模型的探测结果.

强鲁棒性的 SDM-Petri 建模已经成为目前的研究热点,例如表 6 中停留阈值的分布学习、模型组件的自适应构建. 从行为模式学习结果可看出模式与上下文间存在较强依赖性,因此,面向复杂场景上下文的先验信息融合^[13]也是未来的重要研究方向.

参考文献:

- [1] 徐光佑, 曹媛媛. 动作识别与行为理解综述[J]. 中国图象图形学报, 2009, 14(2): 189.
XU Guangyou, CAO Yuanyuan. Action recognition and activity understanding: a review[J]. Journal of Image and Graphics, 2009, 14(2): 189.
- [2] 王素玉, 沈兰荪. 智能视觉监控技术研究进展[J], 中国图象图形学报, 2009, 12(9): 1505.
WANG Suyu, SHEN Lansun. Intelligent visual surveillance technology: a survey [J]. Journal of Image and Graphics, 2009, 12(9): 1505.
- [3] Hamid R, Maddi S, Bobick A, et al. Unsupervised analysis of activity sequences using event-motifs[C]//Proceeding of 4th Acm International Workshop on Video Surveillance and Sensor Networks. California: ACM Press, 2006: 71—78.
- [4] Lavee G, Rudzsky M, Rivlin E, et al. Video event modeling and recognition in generalized stochastic petri nets[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(1): 102.
- [5] XIANG Tao, GONG Shaogang. Incremental and adaptive abnormal behaviour detection[J]. Computer Vision and Image Understanding, 2008, 111(1): 59.
- [6] XIANG Tao, GONG Shaogang. Beyond tracking: modeling activity and understanding behavior[J]. International journal of Computer Vision, 2006, 67(1): 21.
- [7] Cedric S, Jerome M, De Christophe. Using decision trees to recognize visual events[C]//Proceeding of 1st Acm Workshop on Analysis and Retrieval of Events/Actions and Workflows in Video Streams. Vancouver: ACM Press, 2008: 41—48.
- [8] 于起峰, 尚洋. 摄像测量学原理与应用研究[M]. 北京: 科学出版社. 2009.
YU Qifeng, SHANG Yang. Video metrics: principles and research[M]. Beijing: Science Press, 2009.
- [9] Ester M, Kriegel H, Xu X W, et al. A density-based algorithm for discovering clusters in large spatial databases with noise [C]//Proceeding of 2nd International Conference on Knowledge Discovery and Data Mining. Oregon: AAAI Press, 1996: 226—231.
- [10] Borzin A, Rivlin E, Rudzsky M. Surveillance interpretation using generalized stochastic petri nets[C]//8th International Workshop on Image Analysis for Multimedia Interactive Services. Santorini: IET, 2007: 102—118.
- [11] Lavee G, Borzin A, Rivlin E. Building petri nets from video event ontology [C]//Proceeding of 3rd International Conference on Advances in Visual Computing. Nevada: Springer-Verlag, 2007: 442—451.
- [12] 焦波. 面向智能视频监控的运动目标检测与跟踪方法研究[D]. 长沙: 国防科学技术大学系统工程系, 2009.
JIAO Bo. Research on moving object detection and tracking method for intelligent video surveillance [D]. Changsha: National University of Defense Technology. Department of Engineering System, 2009.
- [13] Wang Xiaogang, Ma Xiaoxu, Grimson Eric. Unsupervised activity perception in crowded and complicated scenes using hierarchical bayesian models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(3): 539.