

基于运动轨迹的监控场景分析模型

梁浩哲 李国辉 张 军

(国防科技大学信息系统工程重点实验室 长沙 410073)

摘要 所提出的模型通过层次分类聚类过程学习运动轨迹信息结构。用拓扑先验分类轨迹的相似性运动空间,然后结合混合模型拟合各运动特征的统计分布,得到潜在运动规则;最后基于规则模型检测监控场景中的异常运动行为。轨迹多维运动线索的统计模型对噪声具有较强的鲁棒性,同时利用先验分类使运动规则具有较明显的语义结构。实验结果验证了模型的有效性。

关键词 视觉监控,混合分布估计,场景理解,异常检测

中图分类号 TP391 文献标识码 A

Surveillance Scene Analysis Model Based on Motion Trajectory

LIANG Hao-zhe LI Guo-hui ZHANG Jun

(National University of Defense Technology, Information System Engineering Key Labortory, Changsha 410073, China)

Abstract The proposed model learns the potential structure of motion trajectory within a hierachical process composing of classification and clustering. Topology prior was used to classify the spatial similarity of trajectory. After that by utilizing Mixture Model to estimate distribution of motion features, the potential motion rules of the scene were obtained, based on which abnormal motion can be detected. The model is robust to low-level noises because of the statistic learning of multi-dimentional motion clues, and by combining prior the motion rules have obvious semantic structures. The real surveillance experiment results verified the efficiency of the proposed model.

Keywords Visual surveillance, Mixture distribution estimation, Scene understanding, Abnormal detection

1 引言

基于检测和跟踪结果进行监控视频语义内容提取是目前研究的热点^[1-4],即分析视频帧的底层特征得到合乎人逻辑的结构化信息。大量研究使用统计分析监控场景知识,利用特定分布描述特征样本集,对其潜在模式给出概率化的解释。Xiang^[5]等利用 HMM 分析运动轮廓集,给出其时序模式的统计结构;Wang^[6,7]等使用概率主题模型分析场景中光流场的语义格式。通过估计样本的统计分布,统计分析对底层噪声有较好的鲁棒性,但分析结果往往难以理解并依赖于参数的设定^[8]。

本文提出一种分类聚类模型分析轨迹运动特征。首先利用拓扑先验分类轨迹的空间信息,使统计分析结果具备整齐的语义结构;给定空间拓扑,以多维运动线索为底层特征,使用高斯混合(Gaussian Mixture Model,以下简称 GMM)与 Von Mises 混合(Von Mises Mixture Model,以下简称 VMM)估计其潜在分布,得到监控场景中合理的运动规则;最后通过规则模型实现异常运动检测。

2 运动轨迹层次分类聚类模型

监控场景中的运动对象及其附加属性是刻画场景语义的

重要运动线索。轨迹表示为时空序列:

$$T = \{(x_1, y_1, t_1, o_1), (x_2, y_2, t_2, o_2), \dots, (x_l, y_l, t_l, o_l)\}$$

式中, l 为运动对象持续帧长。轨迹运动线索包括: x, y :空间分量; t :时间戳; o :运动对象属性,包括尺寸,运动方向,速率等分量。多维运动线索矢量 $(x_1, y_1, o_1)^T$ 间的相似性反映了场景中存在的潜在运动规则。

2.1 运动线索分析框架

文献[5]中,直接利用 GMM 模型估计整体运动线索矢量的空间分布。GMM 虽然可以近似拟合任意形式的分布,但拟合结果呈现超球形特点^[7]。如图 1(a)中所示,直接对空间 (x, y) 分量进行球形拟合,并不能代表直观的高层语义信息。

图 1(a)中,点集 $\{(x, y) | (x_i, y_i) \in T_i\}$ 反映轨迹的运动空间,具备拓扑相似性,在帧中表现为运动路径约束,此约束在视觉上有较直观的体现(道路规划),可建立场景中的拓扑先验。

图 1(b)中存在两条运动方向相异轨迹, V_x, V_y 表示纵横速度分量,其分布具备多峰性的特点(如图 1(c),图 1(d)所示)。定义运动矢量:

$$o = (x, y, s = \sqrt{v_x^2 + v_y^2}, c = h \cdot w, a = \arctan(v_y/v_x))$$

式中, s 为速率, c 为轮廓尺寸, a 为运动方向。首先利用拓扑先验分类轨迹,使分析结果具备较强的语义结构,利用 GMM

到稿日期:2011-04-13 返修日期:2011-07-03 本文受国家自然科学基金项目(60902093)资助。

梁浩哲(1983-),男,博士,主要研究方向为机器视觉信息处理工程,E-mail:lianghaozhe@hotmail.com;李国辉(1963-),男,博士,教授,博士生导师,主要研究方向为计算机视觉集成与应用。

拟合 s, c 分布, VMM 拟合 a 分布。表 1 给出了本文分析框架。

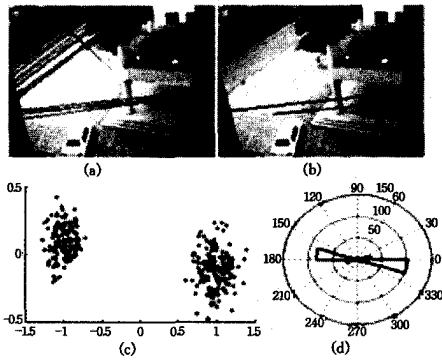


图 1 轨迹信息分布

表 1 分析框架

框架描述	运动规则学习分析与异常运动检测
	步骤 1: 利用拓扑先验提取拓扑相似的轨迹集, 形成拓扑模式集 TPS。
	步骤 2: 建立运动线索统计模型 $p(O TP_k)$, O 为运动线索随机量。利用 GMM 与 VMM 模型拟合 p , 通过均值方差估计得到运动线索的统计模型。
	步骤 3: 基于规则模型的异常检测, 利用似然检测运动目标状态。

2.2 运动线索的拓扑分类与统计聚类模型

轨迹 T_i 的拓扑模式刻画对象运动空间的相似性。本文利用拓扑先验(一般表现为路径规划)确定需分类的训练样本中心。如图 2(a)所示, 手动建立拓扑模式集: $TPS = \{TP_1, TP_2, \dots, TP_k\}$ 。

对轨迹 T_i 进行拓扑分割, 样本点 $P_j = (x_j, y_j, t_j) \in T_i$, 定义 T_i 到模式 TP_k 中的最近邻点 $P_k = (x_k', y_k')$ [9]:

$$(x_k', y_k') = \underset{(x_k', y_k') \in TP_k}{\operatorname{argmin}} d(P_k, P_j) = \|(x_k' - x_j, y_k' - y_j)\|$$

式中, $\|\cdot\|$ 为欧式距离。 P_j 的拓扑标号:

$$k = \underset{TP_k \in TPS}{\operatorname{argmin}} d(P_{TP_k}, P_j)$$

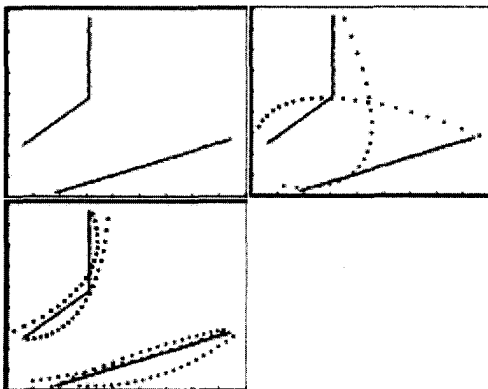


图 2 拓扑模式类

图 2(a)给出了两种拓扑模式, 图 2(b)给出了两条轨迹拓扑分类分割结果, 图 2(c)给出了大量轨迹的分割结果, 由此可得到符合人视觉感知的轨迹空间分类。

给定拓扑模式, 可估计其上运动线索的统计分布。如表 1 所列, 分别使用 GMM 估计尺寸与速率分布, VMM 估计方向分布。估计式为:

$$p(x | TP_k) = \sum_{j=1}^n \eta(\mu_j, \gamma_j | TP_k) \pi(z=j | TP_k) \quad (1)$$

式中, μ_j 均值向量, γ_j 为协方差阵, π 为分布权重系数, 针对 GMM, η 为高斯分布函数; 针对 VMM, η 为 Von-Mises 分布函数。在混合分布估计中, 需定义模型结构即 n 的大小, n 过大导致过拟合。贝叶斯信息准则(Bayesian Information Criteria)[5]是一种优良的模型选择方法, 基于样本可估计 n 的最优值。相比于 AIC, BIC 对参数数量的增长有较高的惩罚, 以避免过拟合问题。

对于备选模型集 $M = \{M_i\}$, BIC 选择估计式为:

$$m_i = \operatorname{argmin}(-2 \log(\text{likelihood} | m_i) + D_i * \log N) \quad (2)$$

式中, D_i 为备选模型“有效”参数量, N 为训练样本数, likelihood 为样本似然。

对 GMM, 给定组件个数 n , 监控过程存在由于遮挡、阴影等造成的轨迹片段, 估计过程假设高斯混合比为 $1/n$, 有效参数[8]包括:

- 1) $n * p$, 代表 μ 个数。
- 2) 1, 代表有效 π 数目。
- 3) γ_j 为对称矩阵, 因此其参数量为 $(p^2 + p)/2$ 。

得参数量为:

$$D_i = n * p + \frac{p^2 + p}{2} * n + 1 \quad (3)$$

将式(1)、式(3)代入式(2), BIC 估计式可得最优 GMM 结构。

类似地, 对于 VMM, 特征维度 $p=1$, 得参数量为:

$$D_i = n * 3 + 1$$

通过空间拓扑先验及其上的运动线索分布模型, 在具备结构化的语义的同时可以对不同的运动规则给出概率解释。如图 3 所示, 场景包括了拓扑分类 $S = \{TP_1, TP_2, \dots, TP_k\}$ 及各拓扑分类具有的统计模型 $TP_i = \{G_n, V_n\}$ 。

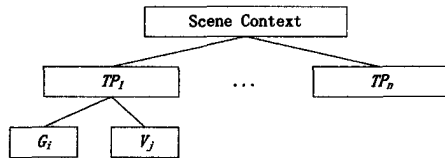


图 3 监控场景层次规则模型

2.3 帧序列中的异常运动检测

利用运动规则模型可检测监控场景中的异常运动。当前帧运动目标集 $O_t = \{o_i\}$, 其中, $o_i = (id_i, x_i, y_i, s_i, c_i, a_i, t)$, 帧间检测跟踪会造成运动线索的较大误差, 本文使用 h 帧均值描述速度:

$$V = \frac{1}{h} \|(centriod(o_i, t+h) - centriod(o_i, t))\| \quad (4)$$

式中, $centriod(\cdot, t)$ 为 t 时刻 (x, y) 分量, 实验过程中使用 $h=6$ 。

给定 o_i , 计算拓扑模式标号 k , 得到运动对象状态变量 $S_k = (o_i, k)$, 使用最大似然分类 S_k :

$$m = \underset{(G_m' \text{ or } V_m') \in TP_k}{\operatorname{argmax}} (\log \text{likelihood}(S_k | G_m' \text{ or } V_m')) \quad (5)$$

定义异常运动条件:

$$\log \text{likelihood}(S_k | G_m' \text{ or } V_m') < T \quad (6)$$

阈值 T 为 95% 置信区间。

3 实验结果及分析

构建的实验环境为远景监控视频, CPU 为 Pentium(R) Dual 1.8GHz, 2G 内存, 使用减背景加 Kalman 滤波器进行多

目标跟踪^[9],时间为上午8:00左右,此时区对象的运动具有潜在规律性,轨迹量为500,视频分辨率为320*240,帧速为15帧/s。

3.1 运动规则识别

图4(a)给出了3种实验场景的预定义拓扑模式,对应场景的3条路径,分别为a,b,c,包括一条主干道与两条进出建筑物路线。图4(b)给出了两条轨迹拓扑分类分割结果,图4(c)给出了大量轨迹的分割结果。在实际监控场景中,监控人员对这种先验知识的把握能力是现实存在的。

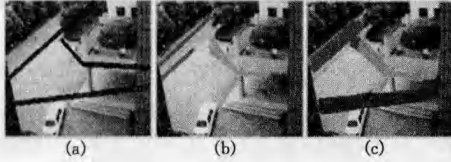


图4 拓扑先验分类

给定拓扑模式,表2为场景潜在运动规则及其语义内容描述,图5中二维高斯为(s,c)GMM拟合,曲线为VMM各组件的概率密度函数。针对b,c拓扑模式,图5(a)右为GMM拟合结果,图5(c)分别对应了其上运动规则A5,A6的VMM分布函数;针对分量(s,c)与方向a,表现为单组件混合,从GMM均值可以看出,路径上仅仅有行人单向运动。

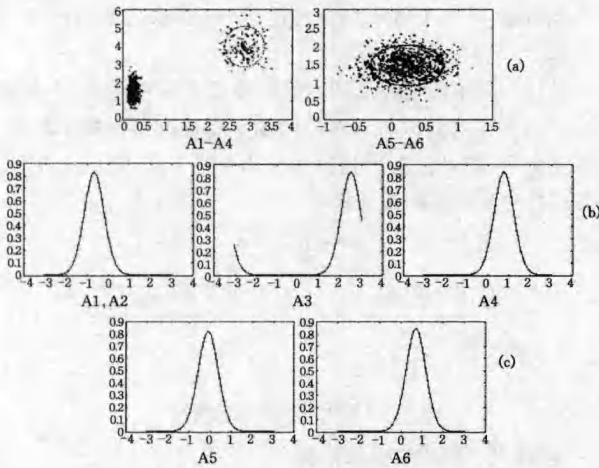


图5 运动线索统计聚类

表2 运动规则及其语义描述

拓扑模式	规则	描述	GMM均值(s&c)	VMM均值
a	A1	车辆上行	2.8689&-3.9798(组件1)	-0.6741(单组件)
a	A2	行人上行	0.2331&-1.5720(组件2)	-0.6691(组件1)
a	A3	行人下行	0.2331&-1.5720(组件2)	2.5095(组件2)
a	A4	穿越路径a	0.2331&-1.5720(组件2)	0.7625(组件3)
b	A5	行人右行	0.2455&-1.4733(单组件)	-0.0910(单组件)
c	A6	行人下行	0.2258&-1.4665(单组件)	0.7155(单组件)

针对拓扑模式a,其上存在4种不同的运动规则。图5(a)左为GMM拟合结果,结合速率尺寸矢量(s,c),不同的运动对象在此分量下存在较强的分离性,分别表现为车辆与行人两种运动对象;对运动方向a,车辆表现为单组件的VMM分布,即存在单向运动,针对行人,则存在3组件VMM分布,其概率密度函数如图5(b)所示,分别对应3种不同的行人运动方向。图6按照运动规则编号给出了各运动规则的关键帧,可以看出,规则集A={A1-A6}刻画了此“上班时区”监控场景中存在的合理的潜在的视觉运动。

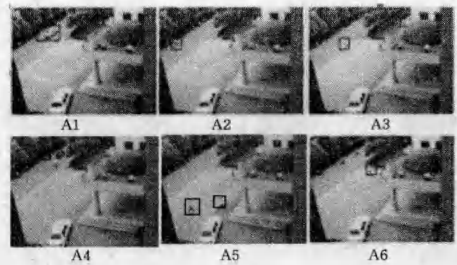


图6 运动规则关键帧

3.2 异常运动检测

通过建立6种此时区下的运动规则,可根据2.3节中的最大似然模型检测场景中的异常运动。图7为行人在路径b中跑步,可视为一种典型的“非规则”运动。在整个过程中,跑动者状态演变为“加速”→“减速”→“稳定行进”。



图7 异常运动关键帧

异常检测过程:首先使用拓扑分类为模式b;其次使用式(5)似然分类为运动A5;图8(a)、图(b)为前后两“规则”平稳运动对象在规则A5下的似然曲线,其演化过程较为平稳;图7(c)为跑动者的似然曲线,在加速过程中似然值较低,在稳定行进过程中似然曲线与(a),(b)情况下相似;图7前4帧“跑动”关键帧,可对异常运动显示为红色并进行报警,后两帧在监测到运动对象状态值的演变后,报警结束。

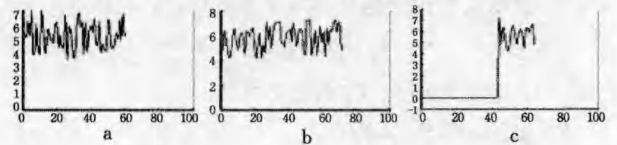


图8 高斯似然曲线

由于一般实验环境下的异常运动较少,为了进一步验证本文提出的模型,在实验视频中手动加入异常运动,添加方法如图9所示:使用刚性矩形框代表运动对象,保持常量速率,随机采样视频序列帧号,插入类似于图9的目标运动,图9前3帧符合正常运动规则,图9后3帧运动规则变为异常或再次转入正常,使用持续时间t表示违反运动规则的程度。表3为对跑动与随机来回运动的检测结果(将持续报警帧数定义为一个检测单位),“跑动”事件正确报警率达到80%，“来回运动”报警率为97%,因此规则模型可以有效地对场景异常运动进行报警,降低用户的实时监控压力。

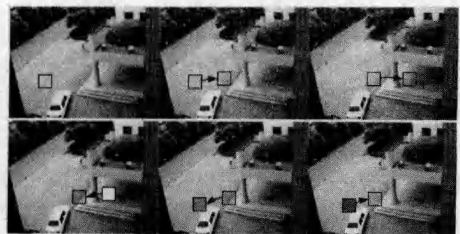


图9 自定义异常运动关键帧

图8中绘出了对视屏解码应用随机映射得到的能耗/延时点,并且分别绘出了由CDPSO和遗传算法得到映射结果的Pareto-front。从图9局部放大的结果中可以清楚地看到,CDPSO所搜索到的映射结果较之遗传算法和随机映射都有较大的优势。图中显示的结果表明,CDPSO得到的映射方案比被随机映射得到的最好结果在能耗上低33.7%,延时上低31%,比遗传算法得到的结果在能量和延时上分别低了26.2%和10.1%。

结束语 本文设计了混沌离散粒子群优化算法,针对NoC映射的两阶段分别提出了能耗/延时粗略估算模型和精确计算模型,对异构多核协作系统进行以能耗和延时为优化目标的分阶段NoC映射求解。在对复杂多媒体系统进行仿真的实验中,本文提出的映射算法与随机映射和基本遗传算法相比表现出更良好的性能。

参考文献

[1] Marculescu R, Ogras U Y, Peh L-S, et al. Outstanding Research Problems in NoC Design: System, Microarchitecture, and Circuit Perspectives[J]. IEEE Transaction on Computer-aided Design of Integrated Circuits and systems, 2009, 28(1): 3-21

[2] Kahng A B, Li Bin, Peh L-S, et al. ORION 2.0: A fast and accurate NoC power and area model for early-stage design space exploration[C]// Proceedings of Design, Automation and Test in Europe. 2009: 423-428

[3] Singh A K, Wu Ji-gang, Prakash A, et al. Mapping algorithms for NoC-based heterogeneous MPSoC platforms[C]// 12th Eu-

romicro Conference on Digital System Design: Architectures, Methods and Tools. 2009: 133-140

[4] Ghosh P, Sen A. Efficient mapping and voltage islanding technique for energy minimization in NoC under design constraints [C]// Proceedings of the 2010 ACM Symposium on Applied Computing. 2010: 535-541

[5] 常政威, 熊光泽, 桑楠, 等. 基于电压岛的能量和可靠性感知NoC映射[J]. 计算机辅助设计与图形学学报, 2009, 21(1): 19-26

[6] 杨盛光, 李丽, 高明伦, 等. 面向能耗和延时的NoC映射方法[J]. 电子学报, 2008, 36(5): 937-942

[7] Hu J, Marculescu R. Energy- and performance-aware mapping for regular NoC architectures[J]. IEEE Transaction on Computer-aided Design of Integrated Circuits and Systems, 2005, 24(4): 551-562

[8] Sotiriadis P P. Interconnect modeling and optimization in deep sub-micron technologies [D]. Cambridge, USA: Massachusetts Institute of Technology, 2002: 23-202

[9] 沈林成, 霍霄华, 牛铁峰. 离散粒子群优化算法研究现状综述[J]. 系统工程与电子技术, 2008, 30(10): 1986-1990

[10] Afshinmanesh F, Marandi A, RahimisKian A. A Novel Binary Particle Swarm Optimization Method Using Artificial Immune System[C]// EUROCON 2005 - The International Conference on Computer as a Tool. 2005, 1: 217-220

[11] Van der Tol E B, Jaspers E G T. Mapping of MPEG-4 decoding on a flexible architecture platform[C]// Proceedings of SPIE-Media Processors. 2002, 4674: 1-13

(上接第266页)

表3 异常运动检测结果

异常行为	预警数	正确预警数
跑动	19	15
来回运动	31	30

本文为视觉监控场景语义分析提供了一个一般化的框架,用于推理监控场景中结构化的运动规则。对于更为复杂的监控视频,例如购物商场,场景中运动对象的数量巨大,不能进行有效的检测跟踪,因此基于轨迹的场景理解技术应用具有局限性,目前直接利用视觉运动线索或使用知识模型进行场景分析已经有很多的研究成果^[11,12],未来的研究会尝试将其纳入分析框架中,使自动分析方法对监控环境限制具有更强的适应性。

参考文献

[1] Wang Su-yu, Shen Lan-sun. Intelligent Visual Surveillance Technology, A Survey[J]. Journal of Image and Graphics, 2009, 12(9): 1505-1514

[2] Gu Jun-xia, Ding Xiao-qing. A Survey of Activity Analysis Algorithms[J]. Journal of Image and Graphics, 2009, 14(3): 377-387

[3] 王昊冉, 白亮, 老松杨. 基于图模型的足球视频语义建模方法[J]. 计算机科学, 2011, 38(6): 266-269

[4] Turaga. Machine Recognition of Human Activities: A Survey [J]. IEEE Transactions on Circuits and Systems for Video

Technology, 2008, 18(11): 1473-1488

[5] Xiang, Gong. Beyond Tracking: Modelling Activity and Understanding Behaviour[J]. International Journal of Computer Vision, 2006, 67(1): 21-51

[6] Wang Xiao-gang. Trajectory Analysis and Semantic Region Modeling Using A Nonparametric Bayesian Model[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Patter Recognition. Anchorage, USA, 2008: 1-8

[7] Wang Xiao-gang. Unsupervised Activity Perception in Crowded and Complicated Scenes Using Hierarchical Bayesian Models [J]. IEEE Transactions on PAMI, 2009, 31(3): 539-555

[8] Mitchell. Machine Learning[M]. China Machine Press, 2003

[9] Wang Xiao-gang. Learning Semantic Scene Models by Trajectory Analysis[C]// ECCV 2006. Graz, Austria, 2006: 110-123

[10] Jiao Bo. Research on Moving Object Detection and Tracking Methods for Intelligent Video Surveillance[D]. Changsha: National University of Defense Technology, 2009

[11] Davis J W, Bobick A F. The Representation and Recognition of Action Using Temporal Templates[C]// IEEE Conference on Computer Vision and Pattern Recognition. Santa Barbara, USA, 1997: 928-934

[12] Lavee G. Video Event Modeling and Recognition in Generalized Stochastic Petri Nets [J]. IEEE Transaction on Circuits and Systems for video technology, 2010, 20(1): 102-118