

密级:_____



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

视频监控中运动目标轨迹分析研究

作者姓名: _____ 李 策 _____

指导教师: _____ 焦建彬 教授 中国科学院大学 _____

学位类别: _____ 工学博士 _____

学科专业: _____ 计算机应用技术 _____

研究所: _____ 中国科学院大学电子电气与通信工程学院 _____

二零一五年 五月

Research on Trajectory Analysis Methods of Moving Objects
in Video Surveillance

By
Ce Li

A Dissertation Submitted to
The University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Doctor of Computer Application Technology

School of Electronic, Electrical and Communication Engineering

May, 2015

中国科学院大学直属院系 研究生学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学直属院系 学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密的学位论文在解密后适用本声明。

作者签名：

日 期：

导师签名：

日 期：

摘 要

目标轨迹分析是计算机视觉的一个重要问题，涉及到模式识别与智能系统、图像处理、统计学、机器学习等多个领域内容。经过近二十年深入的研究和发展，目标轨迹分析技术已经在安防监控、交通分析以及视频检索等方面广泛应用。

传统的视频监控中运动目标轨迹分析研究包括轨迹分类、预测和异常检测其中关键的核心问题都是根据学习样本的轨迹数据，来获取参考的分类模型，用于判断视频序列中的目标运动行为是否异常。

尽管视频监控中运动目标轨迹分析技术已经获得了长足的发展，但其可靠性与稳定性并不是十分令人满意。轨迹的形状、方向和长度多样、局部噪声、学习样本标记少等问题，给目标轨迹分析算法带来了挑战，造成了算法性能不稳定。针对上述问题，本文逐步提出了非监督的基于稀疏重构的目标轨迹分析算法。本文的主要研究内容和贡献如下：

(1) 提出一种基于最小二乘三次 B 样条近似 (Least-squares Cubic Spline Curves Approximation, LCSCA) 的轨迹表示方法。该方法通过对目标轨迹的三维时变序列数据的变换求解，为多样轨迹曲线提供了一种统一数学表示，其中控制顶点和权值因子为精确表示和解析各种形状的曲线提供了充分的灵活性。本文的算法还对控制点参数做了优化选择，从而保证对原始轨迹数据的较好拟合。

(2) 提出一种基于稀疏重构 (Sparse Reconstruction Analysis, SRA) 的目标轨迹分析算法。监控场景中轨迹模式 (即样本集构造的模式) 的多样性会导致使用样本集重构新样本时系数的稀疏性，因此，本文提出通过稀疏重构分析获取样本稀疏重构系数和重构误差的方法。该方法通过重构系数和重构误差进行评估，从样本集中选择出能够揭示新样本轨迹主要特征和内在结构的代表性样本来对新样本进行重构，从而使得异常轨迹检测具有较强的鲁棒性。

(3) 提出一种基于局部约束下稀疏重构 (Locality-constrained Sparse Reconstruction, LSR) 的目标轨迹分析算法。在实际应用中，大部分要处理的轨迹数据是不完整的而且常常带有噪声和局部变化，因此，本文首先需要对不完整轨迹进行预测处理，然后使用轨迹划分的方法引入局部性约束。在局部稀疏

理论的基础上，新样本轨迹可以由局部样本子集来稀疏重构，这样有助于提高分类准确度。本文提出了一系列采用局部约束下的轨迹划分、判别编码、损失加权解码和最小化能量函数解码算法，在一定程度上解决了局部不完整轨迹对轨迹分析效果的影响。

(4) 提出一种基于聚类的局部稀疏重构 (Cluster-based Locality-constrained Sparse Reconstruction, CLSR) 目标轨迹分析算法。轨迹分析方法多数采用监督学习，一般需要获取较多的标记样本，导致所需成本通常很高。另外，实际应用中获取有利于分类的数据比较困难，即往往缺少已学习模型中需要的类别先验。结合增量字典学习的特性，本文提出了基于 K-Means 聚类的增量学习轨迹特征的局部字典方法。该方法还使用最小化能量函数解码作为局部字典的算法，不仅解决了学习模型中缺少类别先验的问题，而且还能更好地选择片段内局部相似度高的样本集合来表示目标轨迹，从而提高了目标轨迹分类的准确度，解决了现有技术在实际应用中训练样本的类别标注非常难以获取的问题。

关键词：视频监控，目标轨迹分析，轨迹分类，异常检测，轨迹表示，稀疏重构

Abstract

Trajectory analysis of moving objects, which combines advanced technologies and research achievements in pattern recognition and intelligent system, image processing, statistics, machine learning and other relative fields, is one of the most important branches in computer vision. After nearly 20 years of R&D, it has been widely applied in video surveillance, traffic analysis, video retrieval, etc.

The conventional study on trajectory analysis includes trajectory classification, prediction and abnormal detection. The key elements of such trajectory-based approaches are how to represent a motion trajectory and get the model of patterns for distinguishing abnormal behaviors in surveillant videos.

Although the research has obtained great development, there is still not enough reliability and stability. Due to various shapes, directions, and lengths of the trajectories, the noises and the lack of the labels of samples, the failure of trajectory analysis often happens. In this paper, we propose an unsupervised sparse reconstruction based method in order to solve those above mentioned problems. The contributions of this paper are summarized as follows.

(1) A novel trajectory representation based on LCSCA (Least-squares Cubic Spline Curves Approximation). It is a parametric vector constructed by three dimensional time series data transformation, which gives a unified mathematical description for various curves, and its control vertices and weights of factors can provide sufficient flexibility. In addition, the proposed LCSCA optimizes parameters of the control points have been optimized to get a better fitting for the original data.

(2) A novel trajectory model based on SRA (Sparse Reconstruction Analysis). Due to the variety of the trajectory patterns in surveillant videos (samples in the set), sparse coefficients are often obtained when constructing a new sample. Therefore, we propose a method to get the reconstruction coefficients and reconstruction errors by sparse reconstruction. Because these samples can reveal the main characteristic of the new sample and its inner structure, we make robust model for classification and detection through the reconstruction of testing trajectories.

(3) A novel trajectory model based on LSR (Locality-constrained Sparse Reconstruction). In practice, most of the trajectory is incomplete with noises and partial changes. Therefore, it is very important to predict the incomplete trajectory and employ the local information to divide a trajectory, and then reconstruct it with subset of local samples, in order to improve the classification accuracy. This paper puts forward a series of algorithms about trajectory partition, discriminate coding, loss weighted and energy minimization decoding. It can solve the problem of local incomplete trajectory in behavior analysis.

(4) A novel unsupervised model on CLSR (Cluster-based Locality-constrained Sparse Reconstruction). Nowadays, existing methods are mostly supervised algorithms that need to collect many tagged samples with high cost and difficulty. Based on the incremental dictionary learning, we propose a method that combined K-means clustering method in the incremental learning local dictionary. During the procedure, the energy minimization decoding strategy is also used. The algorithm can improve the accuracy of the trajectory classification and struggle with limits of labeled training set.

Key Words: Visual Surveillance, Trajectory Analysis, Trajectory Classification, Abnormal Detection, Trajectory Representation, Sparse Reconstruction Analysis

目录

摘 要	I
Abstract.....	III
目录	V
图目录	IX
表目录	XI
第一章 绪论	1
1.1 引言	1
1.2 课题研究背景与意义	1
1.1.1 课题的来源	1
1.1.2 课题的背景	1
1.1.3 课题的应用领域	3
1.3 国内外研究现状	4
1.4 轨迹分析主要分类和方法	5
1.4.1 轨迹分析的研究内容	5
1.4.2 轨迹分析的分类	6
1.5 难点问题分析	8
1.6 本文的研究内容与主要贡献	9
1.7 本文的组织结构	10
第二章 目标轨迹分析综述	13
2.1 轨迹获取	13
2.2 轨迹表示	14
2.2.1 引言	14
2.2.2 基于轨迹分量的表示方法	14
2.2.3 基于曲线近似的轨迹表示方法	15
2.3 轨迹分类	18
2.3.1 引言	18
2.3.2 有监督的轨迹学习	18
2.3.3 非监督的轨迹学习	20
2.3.4 半监督的轨迹学习	23
2.4 异常检测	24
2.5 轨迹预测	25
2.6 本章小结	26
第三章 基于稀疏重构的目标轨迹分析	27
3.1 最小二乘三次 B 样条近似的轨迹表示	28
3.1.1 三次 B 样条基函数	28

3.1.2	B 样条曲线	30
3.1.3	函数近似	33
3.1.4	最小二乘三次 B 样条近似的表示算法	33
3.2	稀疏重构研究综述	36
3.2.1	L_1 范数研究综述	36
3.2.2	稀疏重构的基本原理	37
3.2.3	稀疏重构在相关领域的应用	39
3.3	基于稀疏重构分析的模型	40
3.3.1	目标轨迹的样本集构造	41
3.3.2	目标轨迹的稀疏重构分析	41
3.4	分类器的构建	42
3.4.1	轨迹分类模型	42
3.4.2	异常检测模型	43
3.5	实验结果分析	44
3.5.1	本文实验环境配置	44
3.5.2	实验结果比较与算法分析	45
3.6	本章小结	51
第四章	基于局部约束下稀疏重构的目标轨迹分析	53
4.1	局部约束下稀疏重构分析的思路	54
4.1.1	局部约束的定义	54
4.1.2	局部字典与全局字典	55
4.2	轨迹划分	55
4.2.1	样本集的划分	56
4.2.2	局部性分析	56
4.3	判别编码矩阵	58
4.3.1	判别编码的来源	58
4.3.2	判别编码矩阵的定义	59
4.4	损失加权解码和最小化能量函数解码	60
4.4.1	损失加权解码算法	60
4.4.2	最小化能量函数解码	61
4.5	实验结果分析	62
4.5.1	本章实验的数据描述	62
4.5.2	基于局部约束下稀疏重构模型的轨迹分析实验	63
4.6	本章小结	67
第五章	基于聚类的局部稀疏重构的目标轨迹分析	69
5.1	基于聚类的局部稀疏重构的思路	70
5.2	基于 K-Means 的轨迹聚类	71
5.2.1	K-Means 聚类	71
5.2.2	学习样本的特征选择	71

5.3 增量字典学习模型	72
5.3.1 字典初始化	72
5.3.2 字典学习	73
5.3.3 字典优化	75
5.4 实验结果分析	78
5.4.1 算法的分类性能验证	78
5.4.2 增量学习的有效性验证	81
5.5 本章小结	83
第六章 总结与展望	85
6.1 本文工作总结	85
6.2 未来工作展望	86
参考文献	88
附录：主要符号对照表	96
作者简历与论文发表	98
致 谢	100

图目录

图 1-1 视频监控中运动目标轨迹分析的实现过程.....	3
图 1-2 本文研究内容框图	9
图 3-1 基于稀疏重构的目标轨迹分析框架图.....	27
图 3-2 B 样条基函数的递归定义示意图	29
图 3-3 B 样条曲线的凸包性示意图	31
图 3-4 三次 B 样条曲线中基函数求解示意图	31
图 3-5 LCSCA 轨迹表示的控制点示意图.....	34
图 3-6 稀疏重构示意图	38
图 3-7 部分正常/异常行为的轨迹样本示意图	44
图 3-8 基于 B 样条曲线拟合误差的控制点数的选择示意图.....	45
图 3-9 CAVIAR 样本集中轨迹类别的示意图.....	46
图 3-10 CAVIAR 样本集中正常/异常轨迹的 LCSCA 表示对比图	47
图 3-11 目标轨迹基于稀疏重构的轨迹分析实例.....	48
图 3-12 $R1$ 表示时轨迹分类和异常检测性能对比图	49
图 3-13 $R2$ 表示时轨迹分类和异常检测性能对比图.....	50
图 3-14 CAVIAR 和 NGSIM 数据集上检测性能对比图.....	51
图 4-1 基于局部约束下稀疏重构的流程图.....	53
图 4-2 轨迹的局部约束示意图	54
图 4-3 控制点的局部支撑示意图.....	57
图 4-4 轨迹划分的局部约束示意图.....	57
图 4-5 纠错输出编码框架示意图	59
图 4-6 CAVIAR 和 Carpark 两个场景中待分类的轨迹示例	63
图 4-7 CAVIAR 数据集来自不同类别各一条轨迹的示例	64
图 4-8 Carpark 数据集中不同类别的各一条轨迹的示例.....	64
图 4-9 局部约束下稀疏重构模型的轨迹分类性能对比图.....	65
图 4-10 CAVIAR 和 Carpark 数据集中待分类轨迹与其重构样本示意图	66
图 5-1 基于聚类的局部稀疏重构的目标轨迹分析的流程图.....	69
图 5-2 CAVIAR 数据集上四种方法的轨迹分类结果对比图.....	80
图 5-3 Carpark 数据集上四种方法的轨迹分类结果对比图.....	81
图 5-4 CAVIAR 和 Carpark 数据集上增量学习模型的分类和耗时性能对比	82

表目录

表 1-1	运动目标轨迹分析的研究方法对比.....	7
表 1-2	轨迹学习方法对比	8
表 3-1	基于 B 样条曲线拟合误差的控制点数的选择算法	35
表 3-2	基于稀疏重构的异常行为检测算法.....	43
表 3-3	视频监控轨迹分析的轨迹数据集列表.....	44
表 4-1	一个判别编码矩阵的样例	59
表 4-2	基于损失加权解码的分类算法	60
表 4-3	基于最小化能量函数解码的分类算法.....	62
表 4-4	基于局部约束下稀疏重构的目标轨迹分析的实验数据集列表	63
表 4-5	局部约束下稀疏重构模型的异常轨迹检测性能对比.....	67
表 5-1	基于聚类的字典学习算法	76
表 5-2	增量的字典学习算法	77
表 5-3	CAVIAR 数据集上四种方法的分类性能对比.....	79
表 5-4	Carpark 数据集上四种方法的分类性能对比.....	81

第一章 绪论

1.1 引言

随着大数据时代的到来，网络信息空间急剧膨胀，为图像及视频分析提供了海量的媒体数据，超过了人类所能手动处理的极限。这使得大规模视频数据分析发展成为日益迫切的重大需求，例如海量的交通管理监控图像、安全敏感场合的联网监控视频等，大数据视频资源已成为国家经济、科技和国防等的建设和发展中必须重点关注的问题。因此，如何快速高效地挖掘这些海量视频监控数据的语义，自动分析其中的运动目标行为，已成为当下最活跃的前沿研究领域之一，也是当今知识经济时代中一个具有重大理论意义与应用前景的科学问题。

1.2 课题研究背景与意义

1.1.1 课题的来源

本文研究工作受到国家重点基础研究发展计划（973 计划）、国家自然科学基金以及中国科学院“百人计划”课题资助。作者在论文工作期间，参与了以下课题：

(1) “基于多源数据的飞行器进近威胁目标检测跟踪及行为预测”，国家自然科学基金重点项目（课题编号：61039003），2011.01-2014.12，已结题。

(2) “飞行器威胁目标识别与图像鲁棒匹配理论与方法”，国家 973 计划子课题（课题编号：2010CB731804-2），2010.01-2014.12，已结题。

(3) “基于移动平台的威胁目标跟踪与异常行为检测研究”，中国科学院研究生院院长基金 A 类，2012.05-2014.12，已结题。

1.1.2 课题的背景

图像与视频在人类获知信息的途径中占据着非常重要的地位。计算机视觉正是为了通过计算机模拟人体视觉系统的信息处理过程而产生的学科，目前在

科学研究领域及实际应用中备受瞩目。几十年来，人们不断探索高效的图像、视频处理与分析技术，计算机视觉研究取得了很大进展。由于图像传感器技术的发展和人们对于社会公共安全的日益关注，计算机视觉基础上的运动目标轨迹分析已成为当前一个流行的研究领域，其主要任务是在视频序列中对运动目标进行检测、跟踪和行为分析。如何能通过计算机分析并模仿人类对视频实现监控，是当今视频工程的努力方向。

近几十年以来，视频监控的重要支撑技术一共经历了三个飞跃阶段的发展^[1]。上世纪 30 年代，电视广播等视频技术代替传统的静止图像，实现活动图像的传输成为了连接电视台和众多家庭的媒体的主要技术；80 年代以后，电视会议、视频监控等技术摒弃模拟视频通信，采用数字视频通信，逐渐进入了电视会议和安防系统中；如今，智能视频监控技术越来越成熟，逐步实现视频信息处理从机械化向智能化的转变，将要应用到各行各业的智能视频系统中。

视频监控中的重要支撑技术在近几十年的发展已经逐渐形成了一系列较为完整的流程框架，可以实现对特定场景的计算机自动监控。针对场景中出现的移动目标，通过记录和分析运动目标的轨迹，实现检测目标入侵、徘徊、滞留和移走等功能。一般地，智能视频监控系统会涉及到四个过程及相应的技术^[2,3]，包括运动目标检测、目标分类、目标跟踪和目标行为分析（图 1-1）。运动目标的检测、跟踪是后续处理的基础，而行为分析才是视频监控中的难点。基于轨迹的目标行为分析是行为分析领域的重要研究方法，所以轨迹分析受到越来越多的关注。在智能视频监控领域，视频监控系统的最终目标往往都是要解释所监控的场景里的目标行为进而检测异常，因而，视频序列中目标异常行为的分析识别占据重要地位。智能视频监控迫切需要计算机能够对摄像机获得的视频序列自动分析，识别出场景中运动目标的异常行为，这个系统极大地促进人们的工作和生活全面提升，这也是该学科一直努力的方向。

当今世界上，图像传感器技术飞速发展，人们日益关注社会公共安全，导致大规模的监控摄像头网络被广泛应用在机场、校园、医院、地铁站等特殊的公共场景。基于上述应用，视频监控中运动目标轨迹分析可以为公安部门提供刑事案件的线索，为交通部门提供实时交通路况信息，为个人和商户提供维护财产权益的手段，具有非常重要应用意义。

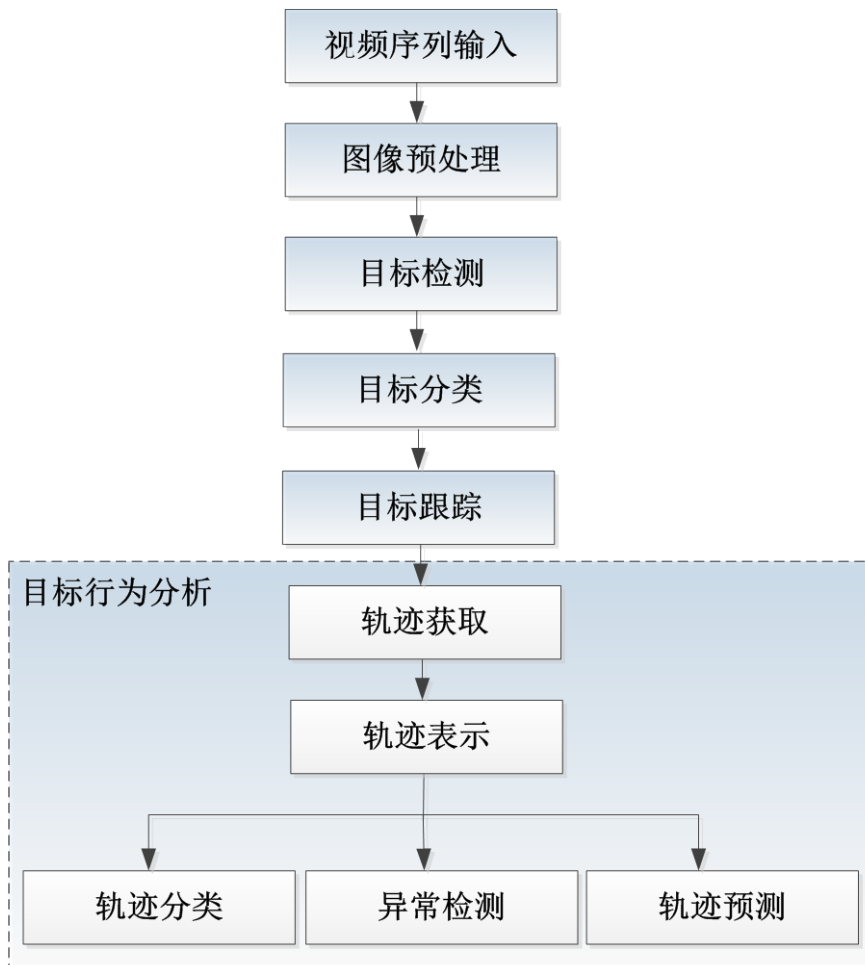


图 1-1 视频监控中运动目标轨迹分析的实现过程

1.1.3 课题的应用领域

视频监控中的运动目标轨迹分析涉及到模式识别与智能系统、图像处理、统计学、机器学习等多个方面，是一门跨学科难度较高的课题。它的研究目的是以视频图像为基础，完成对运动目标的检测、跟踪及轨迹分析，其典型的应用主要有以下几个方面^[4]：

(1) 安防监控

视频监控系统可以应用在机场、地铁站、校园、停车场、银行、医院、ATM取款机等对安全比较敏感的场所，通过实时监视场景中的运动目标，判断目标行为是否异常并自动报警。这不仅能高效地保障公共场所的安全，而且能节省

人力物力资源，产生更大的社会经济收益。

(2) 交通分析

视频监控系统可以应用在交通路口的车辆轨迹分析，通过判断车辆轨迹不同的运动方向和转弯动作，能够判断车辆的违规行驶，预测车辆的行进方向，从而分析交通路口不同动作的流量，对应不同的交通改善方案。

(3) 基于内容的视频检索

当前，通过文字检索视频中感兴趣的内容非常热门，而视频的主要语义信息包括了视频中大量的目标行为。视频监控系统可以分析其中多个对象的交互行为，如碰撞等。如果使用计算机直接检索视频轨迹分析的结果，那么不仅便于检索，而且会提高检索正确率。

1.3 国内外研究现状

通过调查发现，国内外研究人员为了满足智能视频监控系统在民用、商业、国防安全等领域中巨大的应用需求，多年来陆续投入了大量的科研经费，逐步对视频监控中的轨迹分析展开了广泛研究。目前，许多重要的国际学术会议和权威期刊都针对该领域的最新理论研究进展做出了专门讨论，如：International Conference on CVPR (Computer Vision and Pattern Recognition)、ICCV (International Conference on Computer Vision)、ECCV (European Conference on Computer Vision)、ACCV (Asian Conference on Computer Vision)、ICPR (International Conference on Pattern Recognition)、IJCV (International Journal of Computer Vision)、PAMI (IEEE Transaction on Pattern Analysis and Machine Intelligence)、TIP (IEEE Transaction on Image Processing)、PR (Pattern Recognition) 等将运动目标的轨迹分析研究作为主题内容之一，为该领域的研究人员提供了更多的交流机会。

国外在这方面的研究起步较早，国外非常著名的几个视频监控原型系统都包括了场景中运动目标的行为分析研究。例如，美国国防高级研究项目署资助的 HID (Human Identification at a Distance) 研究能够快速发现和分析监控视频中人的活动；Maryland 大学的实时监控系统 W4^[1]能够有效地实时监控出入楼宇、停车场、银行等公共区域可疑人员的简单行为；欧洲的 AVS-PV (Advanced Video-based Surveillance Prevention of Vandalism in metro stations) 项目^[5]可以实

现对地铁站内人的活动行为进行监控；英国雷丁大学的 VIEWS 项目组^[6]研究了视频中车辆和行人的跟踪及其交互行为识别，所提出的一些新技术可以通过分析车辆和行人的运动轨迹来检测异常行为并快速发出警报。

随着图像传感器技术和图像处理技术的快速发展，国家越来越重视智能化监控技术在全国范围内的推广工作。公安部为此建设全国城市报警与监控系统建设“3111”试点工程，863 项目也针对该领域开展了许多重大课题项目，促使国内的一些研究团队在视频监控基于轨迹的行为分析研究上卓有成效。例如，中国科学院自动化所模式识别国家重点实验室图像和视频分析研究组^[7]开发了人体运动视觉分析系统、交通场景监控系统、交通行为事件分析系统等；清华大学智能技术与系统国家重点实验室智能视觉研究组^[3,8]的多项国家重点项目成果，主要有视频中路口交通流分析、高速公路出口的车辆行为监控、公共场合的“天视”数字监控；北京大学数字媒体所^[9]将运动人体轨迹分析成果应用于智能视频监控、体育视频分析和高级人机交互系统；西安交通大学^[10]在自适应巡航控制的车载系统中加入了动态背景下的车辆行为分析技术；北京航空航天大学电子信息工程学院^[11]深入研究了航空交通领域中精确定位航迹的监控的核心理论和关键技术；北京理工大学智能信息技术重点实验室^[12]在智能视频监控中的场景检测分析研究中涉及了人体目标行为分析和异常检测等技术。

1.4 轨迹分析主要分类和方法

1.4.1 轨迹分析的研究内容

目前，关于目标轨迹分析，近年来研究人员开发了大量新的分析算法，这些算法所采用的运动特征表述方法大致分为三类，包括：运动轨迹、时空图和光流。以往的视频监控系统常常依赖于场景信息，即由操作人员来预先定义场景中的运动模式，导致这类视频监控系统对环境的适应性不强。然而，在实际应用中，某一场景中大量的目标运动轨迹往往可以反映出该场景中目标的各种行为模式，因此，本文主要研究基于运动轨迹的目标行为分析。

视频监控系统首先通过目标跟踪获得轨迹数据，即包含目标位置的横坐标、纵坐标、时间的三维时变序列，然后收集场景中大量的运动目标轨迹，通过轨迹分析来识别人或车辆的运动行为，最后建立一个检测模型用于判断场景中的异常行为。这项研究能够针对不同的场景构造出运动目标有效的行为模式，在

智能监控领域具有非常深刻的理论价值和不容忽视的重要意义。

如图 1-1 所示,视频监控中运动目标轨迹分析研究主要内容包含轨迹获取、轨迹表示、轨迹分类、异常检测和轨迹预测。其中,轨迹获取作为目标跟踪模型的输出,是轨迹分析和理解的基础;轨迹表示、轨迹分类和异常检测是轨迹分析的核心内容;轨迹预测是对场景中出现的轨迹的行为语义判断。

1.4.2 轨迹分析的分类

视频监控中运动目标轨迹分析算法^[13-18]依据不同的分类标准可以进行不同的分类(详细对比如表 1-1)。经过我们的调研和总结,基于轨迹的目标行为分析研究的分类如下:

(1) 按照轨迹表示的方法不同,可以分为两类:基于轨迹分量的表示方法和基于曲线近似的表示方法。基于轨迹分量的表示方法是将目标轨迹的分量(位置、速度、加速度等信息)作为表示序列,较为直接。例如 Johnson 等^[41]将运动目标的轨迹表示成连续的流向量,该向量包含了目标位置和速度分量信息。Owens、Stauffer 等^[18,42]对此加以改进,引入加速度信息。这些方法共同的缺点是,轨迹的表示序列长度不一会带来学习建模的困难。于是, Hu 等^[43-45]提出在流向量中加入轨迹长度信息,采用等距离间隔采样来线性插值近似获得固定长度的表示序列,但这种方法仍然不能准确地描述轨迹的形状和方向信息。因此,近几年的研究重点偏向基于目标轨迹的曲线近似的表示方法。Naftel 等^[49]首次将函数近似的方法用于轨迹表示,提出三种分别用最小二乘多项式、Chebyshev 多项式、DFT (Discrete Fourier Transform, 离散傅里叶变换)的方法计算系数特征向量作为轨迹曲线的表示。在此基础上, Sillito 等^[19,20]提出 Haar 小波变换、DFT 系数、Chebyshev 多项式函数的系数和三次 B 样条控制点来求轨迹曲线的特征向量。

(2) 按照轨迹学习的方法不同,可以分为三类:有监督的轨迹分析方法^[2]、非监督的轨迹分析方法^[19]和半监督的轨迹分析^[20]方法。其中,有监督的学习方法主要有:FSM (Finite State Machine, 有限状态机)^[53]、HMM (Hidden Markov Model, 隐马尔科夫模型)^[2,54]、One-Class SVM (One-Class Support Vector Machine, 一类支撑向量机)^[47]、主题词模型^[21]、词袋法^[21]、贝叶斯模型^[22,23,115]等。这类方法在检测过程中利用得到的模型,可以较快地得到新样本的学习结果,而且

适用于对简单的轨迹模式学习。常用的贝叶斯模型有分层狄利克雷分布模型^[22]和分层狄利克雷过程模型^[23]，可以捕捉到数据内部有效的统计类别信息，同时

表 1-1 运动目标轨迹分析的研究方法对比

方法	轨迹表示		固定长度 的表示序 列	轨迹学习			无需预 定义聚 类数目
	分量 表示	曲线 近似		非监督	半监督	有监督	
Johnson 等 ^[41]	√			√			
Owens 等 ^[18]	√			√			
Stauffer 等 ^[42]	√			√			
Hu 等 ^[43]	√		√	√			√
Piciarelli 等 ^[46]	√			√			
Piciarelli 等 ^[47]	√		√			√	√
Naftel 等 ^[49] DFT 变换 +SOM		√	√	√			
Naftel 等 ^[49] Cheybshev 多 项式+SOM		√		√			
Sillito 等 ^[19] Haar 小波系数 + SI-GMMs		√	√		√		√
Sillito 等 ^[20] 三次 B 样条控 制点+ SI-GMMs		√	√		√		√

上下层关系通过统计条件关系相互关联。非监督的学习方法有：K-Means（K均值）聚类^[43]、FCM(Fuzzy C-Means, 模糊 C-均值)聚类^[24]、SOM(Self-Organizing Map, 自组织映射)神经网络^[21,25]、FSOM（Fuzzy SOM, 模糊自组织映射）神经网络^[26]、相似阈值模型^[21]。这类方法直接对训练样本进行学习，自动得到目标的运动行为模式。为了解决学习模型中缺少类别先验的问题，研究人员将监督学习与非监督学习相结合，提出了一种半监督的轨迹学习方法，旨在解决如何利用少量的标注样本和大量的未标注样本进行训练和分类的问题。目前监督

学习的运动目标轨迹分析经过发展，仍然存在需要人工标定数据样本的缺陷，也是其结合实际应用推广时的最大阻碍。然而，半监督的轨迹分析方法，可以减少大量的人工标注代价，具有更广泛的实际应用意义。此外，半监督学习在提高机器学习性能和算法适应能力方面还具有重要的科学意义。典型的半监督轨迹分析方法示例是 SI-GMMs (Semi-supervised Incremental Gauss Mixture Models, 半监督的增量高斯混合模型)^[19]。表 1-2 对比了上述主要的轨迹学习方法的优缺点。

表 1-2 轨迹学习方法对比

类型	方法	优点	缺点
迭代优化	K-Means ^[43] FCM ^[24] GMMs ^[19]	简单有效，易于更新算式	产生类的大小相差不大会很大，对于噪声数据和离群数据很敏感
在线自适应	相似阈值 ^[21]	无需训练数据集，适于实时应用	聚类数目初始值难以确定，对于离群数据不能保证优化
分层	分层 FCMs ^[24]	智能选择聚类数目，可以用图论算法求解	对数据集如何划分敏感
神经网络	SOM ^[25] FSOM ^[26]	可以把复杂的非线性数据描述成有相似结构的节点，降维	参数多，收敛慢
同现分解	词袋模型 ^[21] 主题词模型 ^[21] 贝叶斯模型 ^[22,23]	可以捕捉到数据内部有效的统计类别信息	词表集的大小影响聚类效果，不能保持数据时序不变

1.5 难点问题分析

然而，就目前来讲，视觉目标轨迹分析技术在一般意义上并不完善。传统的视频监控中运动目标轨迹分析研究包括轨迹分类、预测和异常检测，其中关

键的核心问题都是根据学习样本的轨迹数据，建立轨迹分类模型，用于判断视频序列中的目标行为是否异常。但是，由于建立轨迹学习模型时受到监控场景的多样性、目标运动方向的不确定性、目标运动轨迹的长度差异性影响，运动目标轨迹分析研究还存在一些需要解决的问题：（1）轨迹本身形状、方向多样和长度不一等；（2）轨迹不完整、存在噪声和局部变化；（3）学习样本数目少；（4）学习模型中缺少类别先验。上述亟需解决的问题可以概括为：（1）如何更好地提取轨迹的特征表示，从而使得对目标的描述更加准确；（2）如何建立轨迹特征的完备集合以更好地区分和判别轨迹；（3）如何建立轨迹特征的局部字典能够更好地适应轨迹变化；（4）如何更好地增量学习轨迹特征的局部字典，从而能够场景中非监督的轨迹学习。

1.6 本文的研究内容与主要贡献

结合上文所介绍的运动目标轨迹分析研究所存在的难点问题，本文从这些问题出发寻求相应的解决方法（图 1-2），分别提出了四种算法，旨在获取目标轨迹准确有效的特征表示，建立完备的重构样本集合对轨迹的线性重构，提出一种新的轨迹分类和异常行为检测框架，提高视频监控中异常行为轨迹的检测正确率，并使其在上述问题情况下仍能鲁棒地分辨出场景中的轨迹类别，为智能视频监控研究提供一定的理论基础。本文主要研究内容及贡献如下：

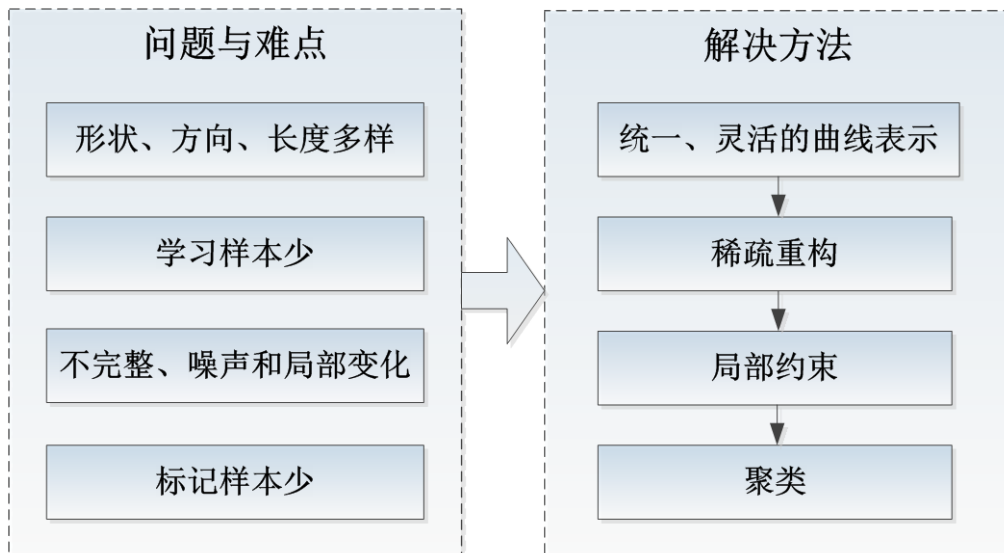


图 1-2 本文研究内容框图

(1) 提出一种基于最小二乘三次 B 样条近似 (Least-squares Cubic Spline Curves Approximation, LCSCA) 的轨迹表示方法。该方法通过对目标轨迹的三维时变序列数据的变换求解, 为各种曲线提供了一个公共的数学表示, 其中控制顶点和权值因子为精确表示和解析各种形状的曲线提供了充分灵活性。本文的算法中对控制点参数做了优化, 从而保证对原始轨迹数据的最好拟合。

(2) 提出一种基于稀疏重构 (Sparse Reconstruction Analysis, SRA) 的目标轨迹分析算法。监控场景中轨迹模式 (样本集构造的模式) 的多样性导致使用样本集重构新样本时系数稀疏, 因此提出通过稀疏重构分析获取样本稀疏重构系数和重构误差的方法。该方法通过重构系数和重构误差进行评估, 从样本集中选择出能够揭示新样本轨迹主要特征和内在结构的代表性样本来对新样本进行重构, 从而使得异常轨迹检测具有较强的鲁棒性。

(3) 提出一种基于局部约束下稀疏重构 (Locality-constrained Sparse Reconstruction, LSR) 的目标轨迹分析算法。在实际应用中, 大部分要处理的轨迹数据是不完整的且带有噪声和局部变化, 因此, 需要对不完整轨迹进行轨迹预测处理, 然后用轨迹划分的方法引入局部性约束。在局部稀疏的理论基础上, 新样本轨迹可以由局部样本子集来重构, 这样有助于提高分类准确度。本文提出了一系列采用局部约束下的轨迹划分、判别编码、损失加权解码和最小化能量函数解码算法, 在一定程度上解决了局部不完整轨迹对轨迹分析效果的影响。

(4) 提出一种基于聚类的局部稀疏重构 (Cluster-based Locality-constrained Sparse Reconstruction, CLSR) 的目标轨迹分析算法。轨迹分析大多采用监督学习的算法, 一般需要获取较多的标记样本, 这样做成本通常很高。另外, 实际应用中获取有利于分类的数据比较困难, 即往往缺少已学习模型中需要的类别先验。结合增量字典学习的特性, 本文提出了基于 K-Means 聚类的增量学习轨迹特征的局部字典方法。该方法与局部字典的最小化能量函数解码算法共同使用, 不仅解决了学习模型中缺少类别先验的问题, 而且还能更好地选择片段内相似度高的局部样本集合对目标轨迹进行表示, 提高了目标轨迹分类的准确度, 解决了现有技术在实际应用中训练样本的类别标注非常难以获取的问题。

1.7 本文的组织结构

第一章, 绪论。主要论述了视频监控中运动目标轨迹分析算法的研究背景

和研究意义，总结了国内外轨迹分析的研究现状，归纳了当前轨迹分析的主要分类及方法，分析了当前研究存在的难点问题，明确了本文的主要研究目的和研究内容，列出了本文的主要贡献。

第二章，目标轨迹分析综述。详细介绍基于轨迹分析中的轨迹获取、轨迹表示、轨迹分类、异常检测和轨迹预测的理论和相关方法基础。

第三章，基于稀疏重构的目标轨迹分析。首先，论述基于最小二乘三次 B 样条近似的轨迹表示的原理方法。然后，介绍稀疏重构分析在计算机视觉领域的应用优势和求解方法，阐述了如何基于稀疏重构分析的模型来构建轨迹分类器和异常检测器，并在轨迹数据集上验证了所提出的算法性能。

第四章，基于局部约束下稀疏重构的目标轨迹分析。本章系统地论述局部约束下如何对轨迹完成划分、编码和解码，详细介绍轨迹划分、判别编码矩阵和损失加权解码的算法来源、关系和解法。然后，建立一种适用于不完整轨迹的局部约束下稀疏重构的分类模型，对不完整、有噪声和局部有变化等情况下的鲁棒性轨迹分类进行了验证和对比分析。

第五章，基于聚类的局部稀疏重构的目标轨迹分析。介绍字典学习的基本思想及其常用方法等，结合 K-Means 聚类和 LC-KSVD (Label Consistent K-Singular Value Decomposition, 标签一致的 K-奇异值分解) 字典学习的方法，引入增量更新机制，设计出基于聚类的局部稀疏重构的算法，并将局部字典的损失加权解码算法改进为最小化能量函数解码算法。最后分析了基于聚类的局部稀疏重构的目标轨迹分析与其他算法的区别，并在轨迹数据集上进行了实验验证。

最后一章是总结了现有的工作和展望了未来工作，还探讨了如何进一步提高目标轨迹分析的鲁棒性和准确性等研究热点和难点。

第二章 目标轨迹分析综述

视频监控中的运动目标轨迹分析是计算机视觉、模式识别等领域的热点研究问题之一。国内外研究学者们都在关注如何提高轨迹分析算法的鲁棒性和准确性。基于轨迹的运动目标行为分析的研究内容包含轨迹获取、轨迹表示、轨迹分类、异常检测和轨迹预测。其中，轨迹获取作为目标跟踪模型的输出，是轨迹分析和理解的基础；轨迹表示、轨迹分类和异常检测是轨迹分析的核心内容；轨迹预测是对场景中出现的轨迹的行为语义判断。本章将按照运动目标轨迹分析流程中的轨迹获取、轨迹表示、轨迹分类、异常检测和轨迹预测，重点介绍上述五个方面的一些经典方法。

2.1 轨迹获取

为了获取目标的运动轨迹，视频监控系统需要对检测出来的目标进行跟踪。跟踪就是将当前帧检测出的运动目标匹配到目标模板的过程，也就是在视频的帧序列中为目标进行定位，而运动目标的轨迹数据则记录的是每帧图像中运动目标跟踪位置的中心点坐标。

监控场景下经典的多目标跟踪算法有两类：一类是预测跟踪位置，另一类是优化搜索方向。典型的预测跟踪算法有 Kalman 滤波和粒子滤波。Kalman 滤波算法^[27,28]用来描述运动系统的模型包含了一个状态方程和一个观测方程，该运动系统模型利用线性最小方差估计来预测运动目标位置。Kalman 滤波具有计算量小、能比较准确的预测目标的运动位置等优点，但是它只适用于线性且呈高斯分布的系统。在目标形式多样化且运动情况复杂多变的场景中，运动系统呈现非线性、非高斯分布，粒子滤波表现更好。粒子滤波^[27]是一种基于贝叶斯最优估计理论的跟踪算法，核心思想是利用一个被称作粒子的样本集合来近似后验概率，集合中每个粒子有一个相应的权重表示其重要性，然后每个粒子通过动态模型向下一个时刻进行传播，最后粒子的权重会经过观察模型再调整。Mean-Shift 算法和 Cam-Shift 算法是优化搜索方向的跟踪方法中较为广泛使用的算法。Mean-Shift 算法^[29,30]是一种非参数的概率密度估计方法，目标的尺度和方向值通过零阶矩、一阶矩及二阶矩来估计，可以较好地适应目标在运动中可

能产生的形变，但会因光照变化等情况导致跟踪目标的丢失。Cam-Shift 算法^[31]在跟踪之前需要人工设定所要跟踪的目标，它无法实现自主识别和跟踪目标的要求，而且该算法可能会收敛于局部最优值而造成跟踪算法失效。

近几年，为了解决背景变化较小、视频亮度较小、跟踪目标尺度伸缩变化大、物体局部遮挡等问题，运动目标跟踪研究人员为了尽可能多的得到较为可靠的运动目标跟踪轨迹，做出了很大的努力，得到了一些较成熟鲁棒的算法。例如，韩^[32-34]的基于自适应样例稀疏表示的目标跟踪算法和基于自适应特征选择的目标跟踪算法，李^[35,36]的基于 One-Class SVM 目标重构的跟踪算法和基于稀疏表示目标重构的跟踪算法，Kalal^[37]的基于改进的在线学习机制的自适应 TLD (Tracking-Learning-Detection, 检测学习跟踪) 算法。本文重点在于研究轨迹分析，因此分别选用了室内大厅监控人的数据集 CAVIAR^[38]、路口监控车辆的数据集 NGSIM^[39]、和停车场监控车辆数据集 Carpark^[40]的跟踪标注轨迹数据来做研究。

2.2 轨迹表示

2.2.1 引言

在一个固定的场景中，我们希望计算机能够自动学习场景中运动目标的行为模式，然后准确判断该场景中的行为模式类别。由于大多数监控场景中的行为模式往往种类很多，而且运动目标的轨迹形状、长度、方向非常多样，所以目标轨迹在模式学习之前需要做轨迹特征提取，即目标轨迹的三维时变序列数据需要变换为轨迹表示序列。以下介绍目标轨迹分析中的两大类轨迹表示方法，它们分别是：基于轨迹分量的表示方法和基于曲线近似的轨迹表示方法。

2.2.2 基于轨迹分量的表示方法

这一类方法将目标轨迹的分量（位置、速度、加速度等信息）作为表示序列，是轨迹表示和分析研究中的较为直接的一类方法。Johnson 等^[41]最早于 1996 年开始研究基于轨迹的异常行为检测，将跟踪得到的目标运动轨迹描述成连续的流向量（包含轨迹的目标位置和速度分量信息），学习出一个竞争学习神经网络模型，该模型需要预先设定场景中行为模式的个数，利用概率密度实现轨迹

检测。Johnson 模型的不足是没有经过迭代学习，即各个输入样本对模型的各个参数只做了一次调整。2000 年时，Owens 等^[18]改进了上述目标轨迹的流向量表示，在其中引入加速度信息，然后提出使用 SOM 神经网络来学习流向量的分布的方法。然而，由于模型的局限性，这种方法只能判断轨迹中的某个点是否异常，并不能识别整条轨迹所代表的运动行为。Stauffer 等^[42]通过将流向量用在线 K-Means 聚类的方法构造多条目标跟踪的轨迹的联合共生统计量表，并以此来进行轨迹分类。

为了解决以上方法中目标轨迹的表示序列长度不一致的问题，Hu 等^[43-45]提出在流向量中加入轨迹长度信息，采用等距离间隔采样来线性插值近似获得固定长度的表示序列用于模糊 K-Means 聚类学习运动模式，并通过计算运动模式的最大后验概率阈值来检测异常行为。

上述方法都是针对一次获得全部训练样本的轨迹数据研究，而 2006 年 Piciarelli 和 Foresti^[46]提出在线聚类算法则是针对逐次获得的训练样本。文中对逐个获得的训练轨迹计算与聚类原型间归一化的欧式距离的平均值，以此为依据更新树状聚类结构上的节点（包括添加操作和分裂操作），并且采用累计频率来判断目标行为是否异常。随后，Piciarelli 等^[47]使用 One-class SVM 来学习重采样的相同长度的轨迹向量的分类模型，用于检测场景中的异常行为。在此之后，Dizan 等^[48]提出了基于 GHMM 的增量学习重采样的目标运动轨迹向量的算法，然而，该方法的限制在于学习过程中每增加一条轨迹样本，就要更新一次模型中离散状态数目、状态矩阵和概率参数。

2.2.3 基于曲线近似的轨迹表示方法

由于轨迹分析关注的不仅仅是轨迹数据的分量，更多的是目标轨迹的形状、方向等信息，所以近些年这方面的研究产生了另一类基于目标轨迹的曲线近似的表示方法。Naftel 等^[49]2006 年首次将曲线近似的方法用于轨迹表示，分别使用最小二乘多项式、Chebyshev 多项式、DFT 三种方法计算轨迹曲线的系数特征向量，并按照系数特征向量间的欧氏距离聚类分析。在此基础上，Sillito 等^[50]对 Haar 小波变换、DFT 变换、Chebyshev 多项式函数近似的系数特征向量表示和三次 B 样条控制点特征向量表示在五个不同目标轨迹数据集上作了对比，三次 B 样条控制点的轨迹表示方法在对比实验定义的评价指标中表现最好。而

且, Sillito 等^[19,20]还将三次 B 样条控制点的轨迹表示方法用在了其提出的 SI-GMMs 学习框架中, 取得了检测异常行为轨迹的较好实验结果。

基于函数近似的目标轨迹曲线表示方法实质就是把轨迹数据通过某种变换求得表示矩阵的方法, 即将轨迹 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_{N-1}, y_{N-1}), (x_N, y_N)\}$ 表示为 $F = \{C_1^X, C_2^X, \dots, C_p^X, C_1^Y, C_2^Y, \dots, C_p^Y\}$ 。以下是三种典型求解方法的介绍。

(1) Haar 小波系数法。基于 Haar 小波变换的尺度函数和小波函数^[50,51]如公式 2-1 和公式 2-2, 将参数向量 $s = \{0, s_2, \dots, s_{N-1}, s_N\}$ 和轨迹序列 T 通过该算法中求得 Haar 小波系数, 作为表示序列 F 。该算法要求轨迹上的轨迹点数为 2 的整数次幂, 否则需要通过线性插值来达到这一条件。

$$\phi(s_n) = \begin{cases} 1 & 0 \leq s_n < 1 \\ 0 & \text{其他} \end{cases} \quad (2-1)$$

$$\psi_{j,k}(s_n) = 2^{j/2} \psi(2^j s_n - k),$$

其中, $j \in \{0, 1, \dots, J-1\}, k \in \{0, 1, \dots, 2^j - 1\}$, 且

$$\psi(x) = \begin{cases} 1 & 0 \leq x \leq \frac{1}{2} \\ -1 & \frac{1}{2} < x \leq 1 \\ 0 & \text{其他} \end{cases} \quad (2-2)$$

(2) DFT 变换法。DFT 变换通常就是将一堆数据变成其他表示空间中的另一堆数据。DFT 轨迹表示法^[50,52]就是对长度为 N 的信号定义个 $\frac{N}{2}$ 不同的频率变换, 如公式 2-3, 然后按公式 2-4 利用每种傅里叶变换的结果中的实部和虚部值来构造表示序列 F 。

$$\begin{aligned} f_k(\bar{X}) &= \frac{1}{N} \sum_{n=1}^N x_n \exp\left(-\frac{2\pi i(k-1)(n-1)}{N}\right) \\ f_k(\bar{Y}) &= \frac{1}{N} \sum_{n=1}^N y_n \exp\left(-\frac{2\pi i(k-1)(n-1)}{N}\right) \end{aligned} \quad (2-3)$$

其中, $i = \sqrt{-1}, k \in \left\{1, 2, \dots, \frac{N}{2}\right\}$

$$\begin{bmatrix} \Re(f_1(\bar{X})) & \Re(f_1(\bar{Y})) \\ \Im(f_1(\bar{X})) & \Im(f_1(\bar{Y})) \\ \vdots & \vdots \\ \Re(f_{N/2}(\bar{X})) & \Re(f_{N/2}(\bar{Y})) \\ \Im(f_{N/2}(\bar{X})) & \Im(f_{N/2}(\bar{Y})) \end{bmatrix} \quad (2-4)$$

(3) Cheybshev 多项式系数法。由于 Cheybshev 节点（第一类 Cheybshev 多项式的根）可以用于多项式插值，所以 Cheybshev 多项式系数^[50]可以通过第 n 个 Cheybshev 多项式函数（公式 2-5）来实现该多项式在连续函数上的最优一致近似。该方法得到的表示序列 F 由所求解的 Cheybshev 多项式函数^[52]中系数组成，求解过程见公式 2-6。由于 Cheybshev 多项式是定义在对偶区间 $[-1,1]$ 上的正交化多项式，该算法在求解时首先需要将参数向量 s 的定义域由 $[0,1]$ 扩展为 $[-1,1]$ ，轨迹数据需要与之相应地通过线性插值扩展到长度为 $2N-2$ ，即 $x_{0,k}$ 、 $y_{0,k}$ 和 $s_{0,k}$ 逐一对应。该算法具有一定的局限性，要求计算的参数向量的定义域必须变换到对偶的闭区间上。

$$\begin{aligned} s_{0,k} &= \cos\left(\frac{1}{M}\pi\left(k + \frac{1}{2}\right)\right) \\ T_j(s_{0,k}) &= \cos\left(n \arccos(s_{0,k})\right) \\ \text{其中, } k, j &\in \{1, 2, \dots, 2N-2\} \end{aligned} \quad (2-5)$$

$$\begin{aligned} c_j(x) &= \frac{2}{2N-2} \sum_{k=0}^{2N-2} x_{0,k} T_j(s_{0,k}) \\ c_j(y) &= \frac{2}{2N-2} \sum_{k=0}^{2N-2} y_{0,k} T_j(s_{0,k}) \\ \{c_0(x), c_1(x), \dots, c_{2N-2}(x), c_0(y), c_1(y), \dots, c_{2N-2}(y)\} \end{aligned} \quad (2-6)$$

近年来，基于曲线近似的目标轨迹表示方法相比基于轨迹分量的目标轨迹表示方法越来越多地受到研究人员的青睐，这是因为基于曲线近似的目标轨迹表示方法具有以下优势。首先，该类方法更多地关注轨迹曲线的趋势信息，得到的表示序列包含了目标轨迹的形状、方向等，这种表示对由于视频监控中遮挡原因所产生的目标跟踪结果中不完整的噪声轨迹具有鲁棒性。其次，基于曲线近似的一些目标轨迹表示方法可以得到固定长度的表示序列，使得这种目标行为模式的轨迹表示方法很容易和其他模式分析的方法结合使用。

2.3 轨迹分类

2.3.1 引言

对于轨迹模式的学习的研究越来越多集中于算法的鲁棒性上，其中轨迹学习模型主要用于轨迹分类。经过这些年的发展，轨迹分类的学习方法主要有三类：有监督的学习方法、非监督的学习方法和半监督的学习方法。

2.3.2 有监督的轨迹学习

有监督的模式学习方法是预先确定存在轨迹行为模式的模型，通过有监督的训练和学习样本集合来求解该模型。这类方法利用所求解的模型可以快速检测新样本，得出轨迹模式的学习结果。但是，这类方法一般对模型的合适程度要求很高，当训练模型出现“欠拟合”和“过拟合”时，往往很可能在训练集以外的数据集上得不到有效的结果。现有的轨迹学习模式方法包括以下几种：

(1) FSM。FSM 是一种根据多种模型提供的检测结果来存储并更新已经检测到和最新检测到的目标之间、同一目标在不同时刻之间的关系的结构^[53]。它由一个用于描述有限数目的状态和状态之间相互转移关系的状态转移函数构成，然后采用该函数确定最佳状态，用于确定测试样本和参考样本是否匹配。FSM 可以分为确定有限态自动机和非确定有限态自动机，两者区别在于有着不同的转移函数。确定有限态自动机的每个输入只有一个状态转移；而非确定有限态自动机的每个输入会有多个状态转移，因此其最佳状态也是随机选择的。刘^[53]在轨迹分析中应用了有限状态机，实现了辅助驾驶中危险事件的检测。

(2) HMM。HMM 是一种特殊的马尔科夫链，该模型的状态是由一个具有相应概率密度分布的状态序列（即观测向量）产生，而不是直接观察得到。HMM 的学习过程分为训练和分类两个阶段^[2]，在训练阶段，HMM 模型需要指定其中隐藏的状态个数，通过优化各个状态转换概率和输出概率确保输出结果与所在类别的特征匹配。在分类阶段，HMM 模型需要计算出产生所观测的目标轨迹类别的概率。因此，HMM 可以应用在行为识别领域。Oliver 等^[54]提出了两种不同的基于状态的人的行为建模学习方法：HMM 和 CHMM（Coupled Hidden Markov Model，耦合隐马尔可夫模型）。后来，为了学习变长时间量度的行为轨迹，Dizan 等^[48]提出了基于 GHMM（Growing Hidden Markov Model，增长的隐

马尔科夫)模型的增量学习重采样的目标运动轨迹向量的算法,该方法随着观察序列的到达,依次逐步更新离散状态数目、状态矩阵和概率参数。

(3) **One-Class SVM**。**SVM**方法是近年来热门的机器学习算法,因为它较好的数据处理能力及对样本的宽松要求,广泛的应用于轨迹分析和识别中。其中,**One-Class SVM**方法由 Scholkopf 等^[55]在 **SVM**的基础上提出,但是它与 **SVM**有很大的不同,主要优势在于:**One-Class SVM**不要求大量的正例样本和反例样本,主要应用在多数样本包含在一类的学习过程。**One-Class SVM**的基本思想是首先假设在样本的特征空间中某一类样本的数据服从某个概率分布,通过构建描述样本概率分布的二值模型,判断待测试的样本在特征空间是否服从该模型分布,即该待测数据属于该类或不属于该类。实际中,**One-Class SVM**是在特征空间中估计一块体积最小的部分,需要包含尽可能多的相关样本,同时保持较大的分类间隔。因此,它可以被应用到异常轨迹检测中,Piciarelli 等^[46]就利用对重采样的固定长度的轨迹向量使用 **One-class SVM**分类方法来训练模型,取得了较好的异常行为检测结果。

(4) 统计模型。统计模型方法不同于上述的判别模型方法,不仅可以统计有效的描述数据,而且可以应用概率准则做分类和预测。监控视频分析中常见的非贝叶斯模型主要是通过将数据矩阵化,然后使用矩阵分解^[43]来获取能量集中的特征向量,继而将数据降维到能量集中的子空间或者在数据原有的空间(如欧氏空间),为数据构造距离测度算子^[56],这样的距离测度可以最大程度的描述数据之间的类别关系。常见的贝叶斯模型有分层狄利克雷分布模型^[22]和分层狄利克雷过程模型^[23]。这些分层贝叶斯模型,可以有效地捕捉到数据内部的统计类别信息;同时上下层关系通过统计条件关系相互关联,可以有效地降维和抽象数据;而且经过模型计算,可以应用概率准则做分类。

上述方法都是模式识别领域中目标识别的经典方法,近些年来陆续被引入到轨迹分析和异常轨迹检测中加以应用。其中,**FSM**中预先定义的状态十分有限,适合监控场景中的较简单的模式学习。**HMM**的训练方法要求大量的轨迹点,否则将出现过拟合的现象。**One-Class SVM**通常需要大量的训练数据,而监控场景中异常行为轨迹的数据搜集相对比较困难,往往很难达到训练的要求。统计模型做 0-1 判断比较困难,通常在分类性能上统计模型比较难以胜过判别模型。

2.3.3 非监督的轨迹学习

非监督的学习方法是指在训练样本集上直接学习，自动得到目标的运动行为模式。在多数实际应用中，监控场景中目标的运动模式只能通过自动地学习其连续运动序列来检测异常和识别行为，因此，这类方法比有监督的学习方法更为适用，这方面的研究典型的有 SOM 神经网络和一些轨迹聚类分析方法 [24,26,42,57,58]。

SOM 神经网络^[25]是一种非监督的学习方法。这种神经网络的拓扑结构是由一个输入层与一个输出层构成。输入层的节点数是输入样本的维数，其中每一节点代表输入样本中的一个分量。输出层节点排列结构是二维矩阵。输入层 X 中的每个节点均与输出层 Y 每个神经元节点通过一个权值(权向量为 W)相连接，这样每个输出层节点均对应于一个连接权向量。自组织特征映射的基本原理是，当某类模式输入时，其输出层某一节点得到最大响应而获胜，获胜节点周围的一些节点因侧向作用也受到较大响应。此时网络会做一次学习操作，获胜节点及其周围节点的连接权向量朝着输入模式的方向作相应的修正。当输入模式类别发生变化时，二维平面上的获胜节点也从原来节点转移到其它节点。通过这样的自组织方式，网络模型用大量训练样本数据来调整网络的连接权值，最终使得网络输出层的特征图可以反映样本数据的分布情况。根据 SOM 网络的输出状况，不仅能判断输入模式所属的类别，使输出节点代表某类模式，而且能够得到整个数据区域的分布情况，即从样本数据得出所有数据的分布特征。Owens 等^[18]在 2000 年在目标轨迹的流向量中引入加速度信息，提出了通过使用 SOM 神经网络来学习流向量的分布，然而，该方法只能决定运动目标的一个轨迹点是否异常，并不能识别整条轨迹所代表的运动行为。

轨迹聚类的方法需要计算轨迹相似性的度量距离，然后将距离相近的轨迹归为一类。例如，Stauffer 等^[42]通过将流向量用在线 K-均值聚类的方法构造多条目标跟踪的轨迹的联合共生统计量表来进行轨迹分类。Liu 等^[24]提出了基于马氏距离的模糊 C-均值轨迹聚类方法。郝等^[58]提出利用轨迹的空间特征和速度方向特征来改进轨迹相似性度量方法，实现轨迹间距离的计算，最终采用层次聚类算法对同一场景区域内的轨迹聚类分析。

这些轨迹聚类方法的核心是对轨迹相似性度量方法的定义，就是轨迹聚类首先需要定义一种衡量轨迹样本间相似度的方法。对此，Hu^[21]的实验室曾对经

典的轨迹聚类中的不同轨迹相似性度量方法作了深入研究, Zhang^[44,45]的对比实验结果说明, Euclidean 距离、PCA (Principle Components Analysis, 主成分分析)+Euclidean 距离、Hausdorff 距离、LCSS (Longest Common Subsequence, 最长公共子序列) 距离、DTW (Dynamic Time Warping, 动态时间规整) 距离。传统的 Euclidean 距离要求两条轨迹的轨迹点个数相同, PCA+Euclidean 距离的方法丢失了轨迹点的运动方向信息, Hausdorff 距离可以对不同长度的轨迹进行编码, 然而该编码没有考虑轨迹方向信息, 时间消耗也比较高。LCSS 和 DTW 方法也可以有效地度量轨迹, 但其理论描述更适于非刚性物体形状变换的轨迹。文献[44]中权衡了识别率和时间消耗, 认为 PCA+Euclidean 距离的方法结果较好, 更适于监控视频中的目标轨迹聚类。

以下是这几种轨迹距离度量的定义介绍^[44,117]。一般地, 一条目标轨迹被表示为二维空间的坐标序列, 如 $\{(a_1^x, a_1^y), (a_2^x, a_2^y), \dots, (a_n^x, a_n^y)\}$, 假设 A 和 B 分别是长度为 N 和 M 的轨迹。

(1) Euclidean 距离。传统的 Euclidean 距离要求两条轨迹的长度相等, 即 $N=M$ 。 A 和 B 之间的 Euclidean 距离为相应坐标向量的平方差之和的开方, 公式如 2-7 所示。

$$D_1(A, B) = \frac{1}{N} \sum_{n=1}^N \left[(a_n^x - b_n^x)^2 + (a_n^y - b_n^y)^2 \right]^{\frac{1}{2}} \quad (2-7)$$

(2) PCA+Euclidean 距离。首先, 轨迹的原始序列的 $x-y$ 坐标值被投影成一维向量, 于是轨迹原始序列转换为一系列 PCA 系数, 如 $\{a_1^c, a_2^c, \dots, a_k^c\}$ 。求解 A 和 B 之间的 PCA+Euclidean 距离定义为计算 PCA 系数的 Euclidean 距离, 公式如 2-8 所示。

$$D_2(A, B) = \left[\sum_{k=1}^K (a_k^c - b_k^c)^2 \right]^{\frac{1}{2}} \quad (2-8)$$

其中, a_k^c 和 b_k^c 分别表示 A 和 B 的第 k 个 PCA 系数, 且 $K \ll 2N$, N 的定义同上, 表示 A 和 B 的轨迹的长度。

(3) Hausdorff 距离。Hausdorff 距离是有限像素点集合之间的一种距离, 度量的是两个特征点集之间的相似度。已知两组有限像素点集合 $A = \{a_1, a_2, \dots, a_n\}$ 和 $B = \{b_1, b_2, \dots, b_m\}$, 则 Hausdorff 距离定义如 2-9 所示。

$$D_3(A, B) = \max\{d(A, B), d(B, A)\} \quad (2-9)$$

其中， $d(A, B) = \max_{a_i \in A} \min_{b_j \in B} \|a_i - b_j\|$ 是点集 $A = \{a_1, a_2, \dots, a_n\}$ 到点集 $B = \{b_1, b_2, \dots, b_m\}$ 的有向 Hausdorff 距离，表示点集 A 中所有点到点集 B 的距离 $d_B(a_i)$ 最大值。定义 $d_B(a_i) = \min_{b_j \in B} \|a_i - b_j\|$ ，表示点 a_i 到集合 B 的距离，是点 a_i 到集合 B 中每一点距离的最小值。这里 $\|\cdot\|$ 表示某种范式，根据使用特征的不同，Hausdorff 距离可以选择使用不同的范式形式，轨迹特征的 Hausdorff 距离主要使用均方误差函数、绝对差累加和函数、归一化互相关函数等。

(4) LCSS 距离。最长公共子序列距离是基于最长公共子序列的概念提出的。最长公共子序列可以用来描述两段序列之间的相似度，它的定义是，一个序列，如果分别是两个已知序列的子序列，而且是所有符合此条件序列中最长的，则其成为已知序列的最长公共子序列。假定 $Head(A)$ 是 A 的前 $N-1$ 个轨迹点， $Head(B)$ 是 B 的前 $M-1$ 个轨迹点， δ 和 ε 是较小的阈值，那么 $LCSS(A, B)$ 最长公共子序列的长度如公式 2-10。

$$\begin{cases} 0 & \text{当 } A \text{ 或 } B \text{ 为空} \\ 1 + LCSS_{\delta, \varepsilon}(Head(A), Head(B)) & \text{当 } \|a_N - b_M\| < \varepsilon \text{ 且 } |N - M| \leq \delta \\ \max(LCSS_{\delta, \varepsilon}(Head(A), B), LCSS_{\delta, \varepsilon}(A, Head(B))) & \text{其他} \end{cases} \quad (2-10)$$

在此基础上， A 和 B 的最长公共子序列距离定义如 2-11 所示。

$$D_4(\delta, \varepsilon, A, B) = 1 - \frac{LCSS_{\delta, \varepsilon}(A, B)}{\max(N, M)} \quad (2-11)$$

其中， $\|a_N - b_M\|$ 是点集 A 中第 N 个点与点集 B 中第 M 个点的某种范式。

在 R^n 上向量 $x = (x_1, x_2, \dots, x_n)^T \in R^n$ 有三种常用的范数形式，1-范数： $\|x\|_1 = \sum_{i=1}^n |x_i|$ ；

2-范数： $\|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}$ ； ∞ -范数： $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ 。

(5) DTW 距离。与 LCSS 距离类似，DTW 距离也是一种基于序列对齐的距离。DTW 算法是从动态规划的思想出发，把时间规整和距离计算结合起来，通

过伸长或缩短测试序列，使其与参考模板长度一致，以便其特征量与标准模式对应。DTW 首先一个满足一定条件的时时间规整函数，用于描述测试序列和参考序列的时间对应关系，然后求解两模板匹配时累积距离最小所对应的规整函数。具体地，在一个二维直角坐标系中，横轴上标出测试序列的 N 个点，纵轴上标出参考序列的 M 个点，通过这些表示帧号的整数坐标画出纵横线形成一个网络，网络中的每一个交叉点 (n, m) 表示测试模式中某一帧的交汇点。那么，动态规划算法就是为了寻找一条通过此网络中若干格点的路径 W ，该路径从左下角出发在右上角结束且具有最小的累积距离，并且路径所通过的格点对应为测试和参考模板用于计算的帧号。在此基础上，轨迹的 DTW 距离公式定义如 2-12 所示。

$$D_5(A, B) = \min \left\{ \frac{1}{K} \left[\sum_{k=1}^K w_k \right]^{\frac{1}{2}} \right\} \quad (2-12)$$

其中， w_k 是所求的规整路径 W 上的第 k 个值。DTW 距离的算法优势在于，即使测试序列模式与参考序列模式的时间尺度不能完全一致，只要时间次序约束存在，它仍能较好地完成测试序列和参考序列之间的模式匹配。这种方法也存在一些问题，例如，识别性能往往过分依赖于端点检测，不能对样本做动态训练，没有充分利用信号的时序动态特性等。

2.3.4 半监督的轨迹学习

文献[19,20]提出了一种对初始训练样本通过核密度估计方法来构建 GMMs 模型的半监督学习方法。当训练样本有新进的增加时，通过更新该 GMMs 中的部分权重、均值、方差参数，来逐步得到分类模型。具体的方法流程：

初始化设定 N_{\max} ，则需要求解 N_{\max} 个高斯分布。假设有 n 个训练样本 $\{x_1, x_2, \dots, x_n\}$ ， $\sigma_1 = 0.001$ ，对于每个训练样本通过核密度估计求解对应的高斯分布，如第 i 个训练样本 $G_i = \{\mu_i, \Sigma_i, \omega_i\}, i = 1, \dots, N_{\max}$ ，其中， $\Sigma_i = I^d \cdot \sigma_i^2$ ， d 是特征向量的维数。当 $n > N_{\max}$ 时，将模型 $\{G_1, G_2, \dots, G_n\}$ 通过混合模型合并的方法更新为 $\{G_1, G_2, \dots, G_{\max}\}$ ，合并如公式 2-13 所示。

随着训练样本个数 n 的增加，逐次合并更新 GMMs 模型，优化模型是 $\arg \min_{G_i, G_j} (\text{cost}(G_i, G_j))$ ，代价函数如公式 2-14。最后，经过期望最大化算法迭代得到的 GMMs 模型即为学习得到的模型，其中各个高斯分布的匹配相似度

来判断轨迹类别。

$$\begin{aligned}\mu_{merge}(i, j) &= \frac{\omega_i}{\omega_i + \omega_j} \mu_i + \frac{\omega_j}{\omega_i + \omega_j} \mu_j \\ \Sigma_{merge}(i, j) &= \frac{\omega_i}{\omega_i + \omega_j} \left[\Sigma_i + (\mu_i - \mu_{merge}(i, j))(\mu_i - \mu_{merge}(i, j))^T \right] \\ &\quad + \frac{\omega_j}{\omega_i + \omega_j} \left[\Sigma_j + (\mu_j - \mu_{merge}(i, j))(\mu_j - \mu_{merge}(i, j))^T \right] \\ \omega_{merge}(i, j) &= \omega_i + \omega_j\end{aligned}\quad (2-13)$$

$$\begin{aligned}\text{cost}(G_i, G_j) &= \omega_i f(G_i \| G_{merge}(i, j)) + \omega_j f(G_j \| G_{merge}(i, j)) \\ f(G_p \| G_q) &= \frac{1}{2} \left(\log \frac{|\Sigma_p|}{|\Sigma_q|} + \text{Tr}(\Sigma_q^{-1} \Sigma_p) + (\mu_p - \mu_q) \Sigma_q^{-1} (\mu_p - \mu_q)^T - d \right)\end{aligned}\quad (2-14)$$

2.4 异常检测

如何判断目标行为是否异常，挖掘大量的轨迹数据中的异常信息已逐渐成为一个备受关注的研究问题。Knoor 等^[59]将轨迹表示成一个由位置、方向和速度组成的三元组，然后运用传统距离函数来计算轨迹间的距离，实现异常轨迹的检测。Li 等^[60]提出基于主题的异常检测算法，它使用 K-Means 算法从滑动窗口中收集 K 个主题，然后使用分类器来检测异常轨迹。Lee 等^[61]提出 TRAOD (Trajectory Outlier Detection, 轨迹异常检测) 算法，先将轨迹划分成线段，用划分的线段表示轨迹局部特征，然后使用 Hausdorff 距离计算每两个线段之间的距离以衡量各个区段的不匹配性，从而确定异常轨迹。在此基础上，Liu 等^[62]提出基于局部异常度的轨迹异常点检测算法，即通过确定每个轨迹点的局部异常程度来确定轨迹是否异常。然而，在实际应用中，上述算法因忽略了其整体特性而导致长轨迹的异常检测准确度较低，而且局部划分的线段特征表示往往难以满足复杂轨迹的要求。Sillito 等^[19]提出用 GMMs 模型来解释每类轨迹不同的位置和速度变化趋势，通过对每类轨迹建立高斯回归模型来将这种区别量化，模型如前文公式 2-14 所介绍。最终，将需要测试的运动轨迹按公式 2-15 计算与该模型的匹配度，从而计算出整个运动轨迹是否异常。其中， $D_M(z)$ 是轨迹的特征向量 z 到模型 G 的 Mahalanobis 距离， N 是用于估计模型 G 的观测轨迹的数量。

$$P(D_M(z)|N) = \exp \left(- \exp \left(- \frac{D_M(z) - \left((2 \ln N)^{\frac{1}{2}} - \frac{\ln \ln N + \ln 2\pi}{2(2 \ln N)^{\frac{1}{2}}} \right)}{(2 \ln N)^{\frac{1}{2}}} \right) \right) \quad (2-15)$$

2.5 轨迹预测

监控场景轨迹分析研究通过对场景历史轨迹数据处理, 学习场景的行为模式, 依据学习得到的行为模式, 判断测试轨迹的行为模式类别, 并计算其属于某种行为模式的概率, 预测目标将来的行为, 这样用机器学习的方法对观测得到的历史轨迹数据建立的分类模型, 不仅能得到目标最可能的运动方向预测, 还能预测目标轨迹将来行进的具体位置及概率分布。

当前的轨迹预测研究主要是基于一种对场景历史数据行为模式学习的预测。Dizan 等^[48]对场景训练轨迹用聚类的方法得到场景的运动行为模式, 计算测试轨迹数据与每一类行为模式的似然函数, 选取似然值最大的行为模式为测试轨迹将来的运动行为。Ellis 等^[5]提出一种非参数的长轨迹预测研究框架, 该研究中用高斯过程对训练轨迹建立一个概率模型, 用得到的概率模型估计目标下一时刻的运动位置, 如此递归计算, 预测出长轨迹。Hermes 等^[63]提出四元旋转不变最长公共子序列方法作为轨迹间相似性度量, 使用粒子滤波算法跟踪运动目标获取轨迹数据, 在跟踪数据内计算概率密度函数预测目标位置。Hermes 等^[63]在 2011 年使用历史轨迹数据获取场景的行为模式来预测运动目标行为, 具体做法先用 Chebyshev 多项式表示目标轨迹, 将轨迹数据映射到一个参数空间中, 利用 UKR (Unsupervised Kernel Regression, 非监督的核回归) 方法将参数向量投射到低维的流形场中, 在该流形场中用粒子滤波跟踪历史轨迹数据并预测运动目标的位置。Bashir 等^[64]将训练轨迹划分为片段轨迹, 使用 K-means 聚类方法得到片段轨迹的行为模式类别, 这些片段轨迹的行为模式类别作为隐马尔科夫的状态, 利用状态之间的转移来预测其将来的行为; Picciarelli 等^[46,47]利用树结构来表示轨迹, 根据树节点之间的转移关系来预测目标的行为; Hu 等^[21]用两层 FCM 聚类来进行模式学习, 第二层的聚类结果用高斯分布建模, 根据贝叶斯准则计算概率最大的运动方向。

2.6 本章小结

本章详细地描述了运动目标轨迹分析中的轨迹获取、轨迹表示、轨迹分类、异常检测和轨迹预测的理论基础和相关方法。首先，本章介绍了两类经典的多目标跟踪算法和近年跟踪算法的前沿研究成果，还说明了如何获取本文研究的数据集中轨迹。然后，重点介绍了两类轨迹表示方法的数学描述，接着给出了轨迹分类三种学习方法的详细对比和应用分析，之后阐述了几种异常轨迹检测算法。最后，简单说明了轨迹预测的现有方法原理。本章内容主要为后续章节所提出算法等相关内容的技术综述。

第三章 基于稀疏重构的目标轨迹分析

为了解决监控场景中目标运动轨迹的形状、方向、长度多样和学习样本数目少的问题，本章主要是针对轨迹表示、轨迹分类和异常检测研究，首先提出了一种最小二乘三次 B 样条近似（Least-squares Cubic Spline Curves Approximation, LCSCA）的轨迹表示方法，然后提出了一种基于稀疏重构（Sparse Reconstruction Analysis, SRA）的目标轨迹分析算法，设计出基于稀疏重构分析的模型，并构建了轨迹分类和异常检测的分类器。图 3-1 显示了 LCSCA 的轨迹表示和稀疏重构用于目标轨迹分析的流程图。

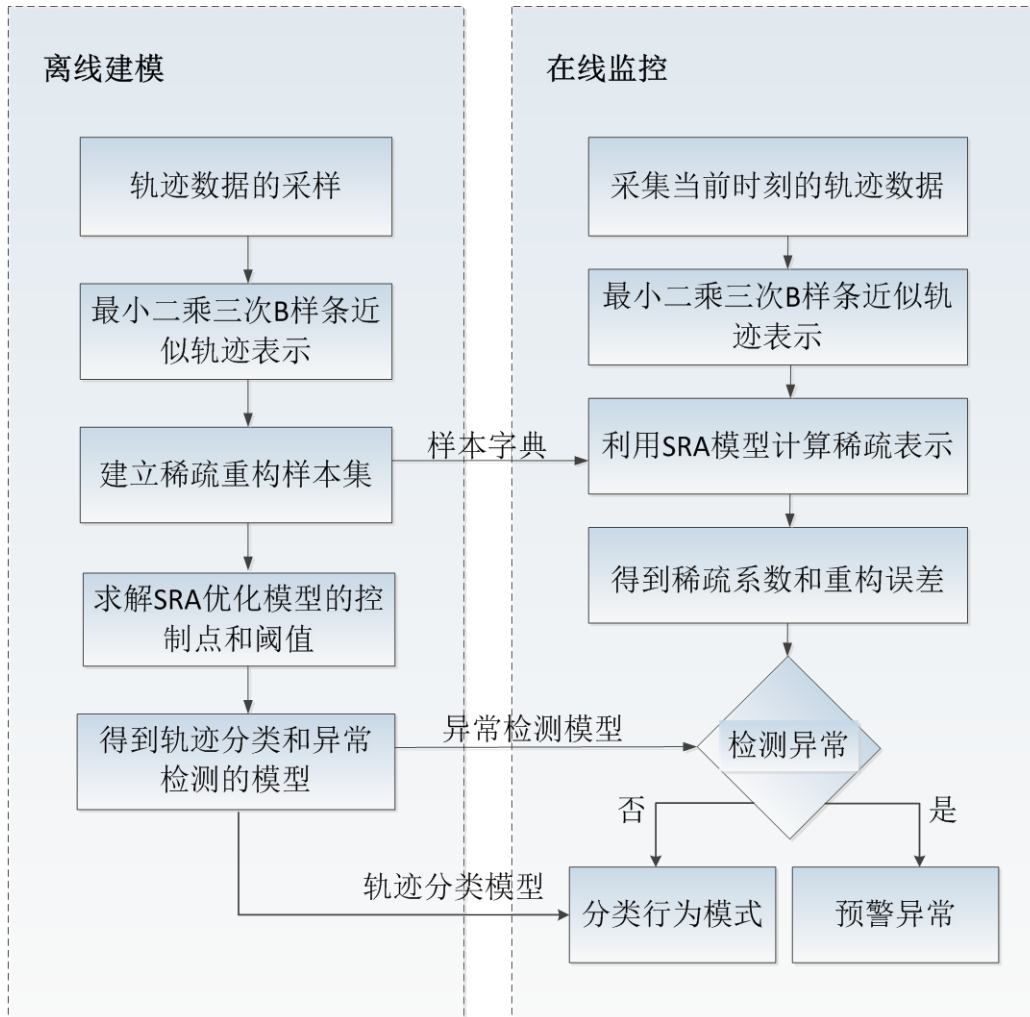


图 3-1 基于稀疏重构的目标轨迹分析框架图

3.1 最小二乘三次 B 样条近似的轨迹表示

为了将形状、方向多样和长度不一的轨迹统一表示，本文提出了最小二乘三次 B 样条近似(LCSCA)轨迹表示方法。最小二乘三次 B 样条近似(LCSCA)轨迹表示意在通过对目标轨迹的三维时变序列数据的变换求解，通过数学函数解析各种轨迹曲线，实现轨迹特征的提取。该方法利用了 B 样条基函数在曲线近似方面的优良性能，克服了贝兹曲线在局部修改上的缺陷，提取得到相同长度的准确描述形状、方向信息的轨迹特征。

3.1.1 三次 B 样条基函数

B 样条是数值分析领域里样条曲线的一种特殊表示形式，是 B 样条基曲线的线性组合^[65, 66]。B 样条基函数也称为基本样条 (Basic spline)，是样条函数空间中具有最小支撑的一组基。B 样条基函数有多种定义形式，例如，积分、差分、卷积、截断幂、递推等。理论上较多采用截断幂函数的差商定义，但实际计算中则更多采用 Deboor-Cox 递推公式^[64]定义 B 样条基函数，这样不仅便于计算分析基函数的性质，而且具有明显的几何形状特征，简明直观。

定义 1 (B 样条基函数) 设 $T = \{t_0, \dots, t_m\}$ 是单调递增的实数序列，即 $t_i < t_{i+1} (i=0, 1, \dots, m-1)$ ， t_i 是节点， T 是节点向量， $N_{i,p}(t)$ 是第 i 段次(或 $p+1$ 阶)

B 样条基函数，定义如下：

$$N_{i,0}(t) = \begin{cases} 1 & \text{若 } t_i \leq t \leq t_{i+1} \\ 0 & \text{否则} \end{cases}, \quad (3-1)$$

$$N_{i,p}(t) = \frac{t-t_i}{t_{i+p}-t_i} N_{i,p-1}(t) + \frac{t_{i+p+1}-t}{t_{i+p+1}-t_{i+1}} N_{i+1,p-1}(t)$$

并且规定： $\frac{0}{0} = 0$ 。如图 3-2 所示，上述定义的 B 样条基函数具有以下性质：

(1) 局部支撑性和非负性。根据 B 样条基的微分差分公式，有：

$$N_{i,p}(t) = \begin{cases} \geq 0 & \text{若 } t \in [t_i, t_{i+p+1}] \\ = 0 & \text{其他} \end{cases}. \quad (3-2)$$

(2) 任意给定节点区间 $[t_i, t_{i+1}]$ 最多只有 $p+1$ 个 $N_{i,p}$ 非零，即

$N_{i-p,p}(t), \dots, N_{i,p}(t)$ 非零。

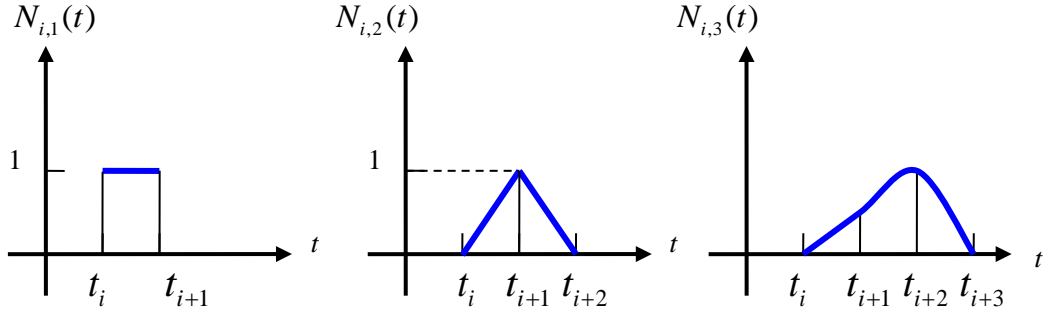


图 3-2 B 样条基函数的递归定义示意图

(3) 规范性。对于任意 $t \in [t_i, t_{i+1}]$, $\sum_j N_{j,p}(t) = \sum_{j=i+1-p}^i N_{j,p}(t) = 1$ 。

(4) 最大值唯一性。除 $p=0$ 外, $N_{i,p}(t)$ 存在一个最大值。

(5) 连续性。在节点区间内部, $N_{i,p}(t)$ 任意阶可导, 而在节点处 $N_{i,p}(t)$ 是 $p-k$ 次可导, 其中 k 是该节点的重复个数, 降低节点重复个数可增加连续性。

而且, $N_{i,p}(t)$ 的存在导函数可表示为:

$$N_{i,p}^{(k)}(t) = (p-1) \left(\frac{N_{i,p-1}^{(k-1)}(t)}{t_{i+p-1} - t_i} - \frac{N_{i+1,p-1}^{(k-1)}(t)}{t_{i+p} - t_{i+1}} \right). \quad (3-3)$$

(6) 对于非周期节点向量 $T = \left\{ \underbrace{a, \dots, a}_{p+1}, t_{p+1}, \dots, t_{m-p-1}, \underbrace{b, \dots, b}_{p+1} \right\}$, B 样条基函数还

具有以下特殊性质: 设节点个数为 $m+1$, 则有 n 个基函数, 其中 $n = m - p$,

$N_{0,p}(a) = 1$ 和 $N_{n,p}(b) = 1$ 。对于特殊节点向量 $T = \left\{ \underbrace{0, \dots, 0}_{p+1}, 1, \dots, 1 \right\}$, B 样条基函数

变成了 p 次贝兹基函数。

B 样条基函数的上述性质导致 B 样条具有对一般的几何曲线建造精确模型的优势。而且, B 样条基函数既保留了贝兹基函数的一切优点, 又克服了贝兹基函数不具有局部修改性的缺点。所以, Sillito 将三次 B 样条与函数近似的理

论结合用于轨迹曲线的表示^[50]。

3.1.2 B 样条曲线

B 样条曲线是一种形状描述的数学方法，它由 Gordon 和 Riesenfeld^[67]对贝兹曲线的一般化，它的进一步推广非均匀有理 B 样条是 CAGD (Computer Aided Geometric Design, 计算机辅助几何设计) 的核心技术，在计算机的其他领域也得到大量应用和推广。

定义 2(B 样条曲线) 设 P_0, P_1, \dots, P_n ($n > k$) 为给定空间中的 $n+1$ 个点，则 $k+1$ 阶 (k 次) 的 B 样条曲线定义为如下参数曲线：

$$P(t) = \sum_{i=0}^n P_i N_{i,k+1}(t) \quad t_k \leq t \leq t_{n+1}, \quad (3-4)$$

其中， $P(t)$ 的控制多边形是折线 (P_0, P_1, \dots, P_n) ， $\{P_i\}_0^n$ 是 $P(t)$ 的控制顶点， $N_{i,k+1}(t)$ 是定义在分割空间 $\{t_i\}_{i=0}^{n+k+1}$ 上的 k 次 B 样条基函数，由公式 3-1 确定。

分析上述数学定义，B 样条曲线具有以下性质：

(1) 可导性。由 B 样条基的微分差分公式，有：

$$\begin{aligned} P'(t) &= \left(\sum_{i=0}^n P_i N_{i,k+1}(t) \right)' \\ &= \sum_{i=0}^n P_i N_{i,k+1}'(t) \\ &= k \sum_{i=0}^n \left(\frac{P_i - P_{i-1}}{t_{i+k} - t_i} \right) N_{i,k}(t) \end{aligned} \quad (3-5)$$

(2) 唯一性。以 $\{P_i\}_0^n$ 为控制顶点的 k 阶 B 样条曲线可以唯一表示为公式 3-5。

(3) 规范性。当且仅当 $\{P_i\}_0^n$ 退化为一 点 (即 $P_i \equiv A$) 时，由其所确定的 B 样条曲线也退化为一 点。这是因为，若 $P_i \equiv A$ ，则 $P(t) = \sum A N_{i,k+1}(t) = A \sum N_{i,k+1}(t) = A$ 。反之，若 $P(t) = A$ ，即 $\sum P_i N_{i,k+1}(t) = A$ ，对此两边求导得出 $k \sum \left(\frac{P_i - P_{i-1}}{t_{i+k} - t_i} \right) N_{i,k}(t) = 0$ 。那么，根据 B 样条基 $N_{i,k}(t)$ 的线

性无关性可得 $P_i \equiv P_{i-1}$ ，从而 $P_i \equiv A$ 。

(4) 连续性。 $k+1$ 阶 B 样条是分段的 k 次参数曲线，在节点处不低于 $k-m$ 次连续， m 为节点的重复个数。

(5) 凸包性。 $k+1$ 阶 B 样条的任意一段，都落在相应的 $k+1$ 个顶点构成的凸包内。如图 3-3 所示，以三阶 B 样条为例，二次曲线的各段分别落在每三个顶点构成的凸包 (P_0, P_1, P_2) 、 (P_1, P_2, P_3) 、 (P_2, P_3, P_4) 、 (P_3, P_4, P_5) 内。

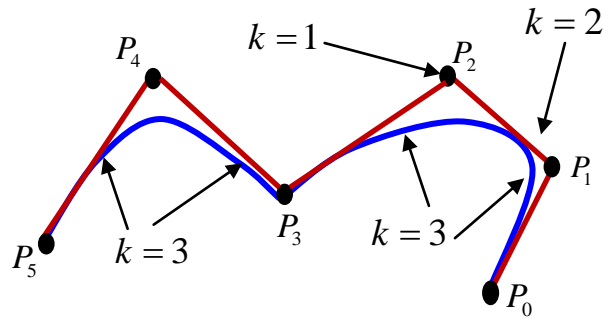


图 3-3 B 样条曲线的凸包性示意图

(6) 局部性。改变 $k+1$ 阶 B 样条的一个顶点，至多影响以此点为中心的相邻 $k+1$ 段曲线。

(7) 直线保持性。控制多边形退化为一条直线时，曲线也退化为一条直线。

(8) 几何不变性。B 样条曲线的形状和位置与坐标系的选择无关。

(9) 缩减局部性。设平面内 $n+1$ 个控制顶点 (P_0, P_1, \dots, P_n) 构成 B 样条曲线 $P(t)$ 的特征多边形，则在该平面内的任意一条直线与 $P(t)$ 的交点个数不多于该直线和特征多边形的交点个数。

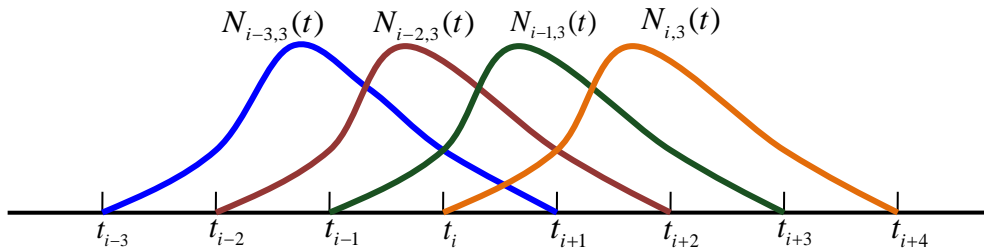


图 3-4 三次 B 样条曲线中基函数求解示意图

B 样条曲线具有很好的几何特性。以四阶三次 B 样条曲线^[63,64,68]为例，曲线段起止点的位置、一阶导数、二阶导数可直接由控制多边形的顶点位置确定。如图 3-4 示意，计算三次 B 样条函数在 $[t_i, t_{i+1}]$ 上的一点，要用到除 t_i 、 t_{i+1} 的前续 3 个和后续 3 个，包括了 t_{i-3} 、 t_{i-2} 、 t_{i-1} 、 t_i 、 t_{i+1} 、 t_{i+2} 、 t_{i+3} 和 t_{i+4} 共 8 个节点的参数值。正是因为基函数是这样局部支撑的，仅在有限区间上有值，才使得 B 样条曲线具有很好的局部修改性。由此可见，B 样条基函数的计算求解是 B 样条曲线方法的关键，以下是求解算法 Deboor-Cox 的导出定义。

定义 3 (Deboor-Cox 算法) B 样条基函数公式 3-1 代入参数曲线公式 3-4 可以转化为:

$$\begin{aligned}
 P(t) &= \sum_{i=0}^n P_i N_{i,k+1}(t) \\
 &= \sum_{i=j-k+2}^j P_i N_{i,k+1}(t) \\
 &= \sum_{i=j-k+2}^j P_i \left[\frac{t-t_i}{t_{i+k-1}-t_i} N_{i,k}(t) + \frac{t_{i+k}-t}{t_{i+k}-t_{i+1}} N_{i+1,k}(t) \right] . \\
 &= \sum_{i=j-k+2}^j \left[\frac{t-t_i}{t_{i+k}-t_i} P_i + \frac{t_{i+k+1}-t}{t_{i+k+1}-t_{i+1}} P_{i-1} \right] N_{i,k}(t)
 \end{aligned} \tag{3-6}$$

其中，当 $r=0$ 且 $i=j-k+2, j-k+3, \dots, j$ 时，令 $P_i^{(r)}(t) = P_i$ ；当 $r=1, 2, \dots, k$ 且 $i=j-k+r+2, j-k+r+3, \dots, j$ 时，令

$$P_i^{(r)}(t) = \frac{t-t_i}{t_{i+k-r+1}-t_i} P_i^{(r-1)}(t) + \frac{t_{i+k-r+1}-t}{t_{i+k-r+1}-t_i} P_{i-1}^{(r-1)}(t) . \tag{3-7}$$

那么，公式 3-6 可以表示为:

$$P(t) = \sum_{i=j-k+2}^j P_i N_{i,k+1}(t) = \sum_{i=j-k+3}^j P_i^{(1)}(t) N_{i,k}(t) . \tag{3-8}$$

上式是同一条曲线 $P(t)$ 从 $k+1$ 阶 B 样条表示到可 k 阶 B 样条表示的递推公式，以此类推， $P(t)$ 最终可以通过以下公式递推求得

$$P(t) = P_j^{(k)}(t) . \tag{3-9}$$

3.1.3 函数近似

在实际应用中，我们常常要面对用简单函数近似代替较复杂函数的问题。插值就是近似代替的方法之一，插值的近似标准是在插值点出误差为零，但实际应用中往往不是对某些点误差要求严格，而是要求考虑整体的误差限制，这就引出了曲线的函数近似^[69]的概念。

定义 4（函数近似）对函数类 A 中给定的函数 $f(x)$ ，记作 $f(x) \in A$ ，要求在另一类简单的便于计算的函数类 B 中求函数 $p(x) \in B$ ，使 $p(x)$ 与 $f(x)$ 的误差在某种意义下最小。函数类 A 通常是区间 $[a, b]$ 上的连续函数，记作 $C[a, b]$ ，称为函数近似空间；而函数 B 通常为 n 次多项式、有理函数或分段的低次多项式等。对连续的函数空间 $C[a, b]$ ，若 $f(x) \in C[a, b]$ 可以定义以下三种常用的函数范数：

$$\begin{aligned} \|f\|_1 &= \int_a^b |f(x)| dx && 1\text{-范数} \\ \|f\|_2 &= \left(\int_a^b |f(x)|^2 dx \right)^{\frac{1}{2}} && 2\text{-范数} \\ \|f\|_\infty &= \max_{a \leq x \leq b} |f(x)| && \infty\text{-范数} \end{aligned} \quad (3-10)$$

根据魏尔斯特拉斯定理^[70]，取任何 $\varepsilon > 0$ ，存在一个代数多项式 $p(x)$ 使公式

$$\|f(x) - p(x)\|_\infty < \varepsilon \quad (3-11)$$

在区间 $[a, b]$ 上一致成立。

3.1.4 最小二乘三次 B 样条近似的表示算法

最小二乘三次 B 样条近似(LCSCA)轨迹表示就是基于上述 B 样条基函数、B 样条曲线、函数近似理论所提出的，采用 B 样条函数来对轨迹曲线实现函数近似。LCSCA 轨迹表示的基本思想就是以轨迹点为二次曲线弧的型值点，按照距离加权插值，然后计算反求得轨迹曲线的控制点，对几何变换和投影变换保持不变性，同时形状描述因控制顶点和权因子而具有灵活性。

LCSCA 轨迹表示算法首先根据目标轨迹的分布趋势，将轨迹数据序列

$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_{N-1}, y_{N-1}), (x_N, y_N)\}$ ，以上一节定义的 **B** 样条基得出系数矩阵 Φ ，建立控制点序列求解方程（公式 3-12），得到轨迹的特征表序列示 $F = \{C_1^X, C_2^X, \dots, C_p^X, C_1^Y, C_2^Y, \dots, C_p^Y\}$ 。其中， p 表示所选取控制点个数， T^{XY} 表示轨迹序列， F^{XY} 表示控制点序列， Φ 表示反求控制点的系数矩阵，并且 $\Phi^\dagger = (\Phi^T \Phi)^{-1} \Phi^T$ 。图 3-5 是当 $p=7$ 时 LCSCA 轨迹表示的控制点示意。在获取监控视频中的目标轨迹后，本文将每条轨迹通过基于 LCSCA 的目标轨迹表示算法做出轨迹表示。具体算法流程如下：

$$F^{XY} = \Phi^\dagger T^{XY} \quad (3-12)$$

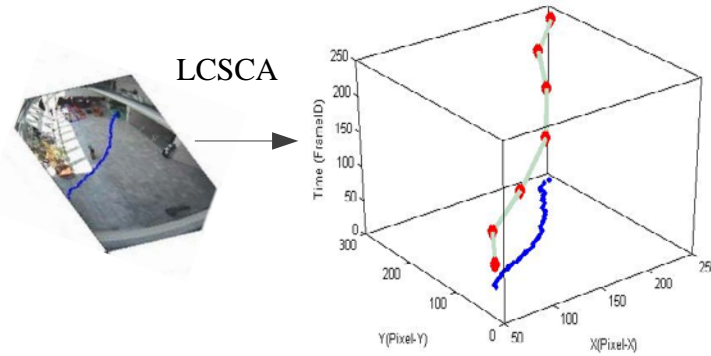


图 3-5 LCSCA 轨迹表示的控制点示意图

(1) 根据目标轨迹的分布趋势，构造参数向量和非均匀节点向量。选取与轨迹时间信息相关的弧长比值作为参数向量 $s = \{0, s_2, \dots, s_{N-1}, s_N\}$ ，求解公式：

$$s_n = \frac{\sum_{i=2}^n \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{\sum_{i=2}^N \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}, \quad (3-13)$$

其中， $n = 2, 3, \dots, N$ ， $s_n \in (0, 1]$ 。选取控制点的个数 p ，构造非均匀的节点向量：

$$\vec{\tau}_{1 \times (p+4)} = \left\{ \underbrace{0, 0, 0, 0}_{1, \dots, 4}, \underbrace{\frac{1}{p-3}, \frac{2}{p-3}, \dots, \frac{p-4}{p-3}}_{5, \dots, p}, \underbrace{1, 1, 1, 1}_{p+1, \dots, p+4} \right\}. \quad (3-14)$$

(2) 使用计算得到的节点向量构造非均匀 **B** 样条基函数。由于一段三次 **B** 样条曲线只受四个控制点的影响，所以四阶三次 **B** 样条的基函数^[71]的递推求解公式如下：

$$B_{p,1}(s_n) = \begin{cases} 1 & \text{当 } \tau_p \leq s_n < \tau_{p+1} \text{ 时} \\ 0 & \text{其他} \end{cases} \quad (3-15)$$

$$B_{p,4}(s_n) = \frac{s_n - \tau_p}{\tau_{p+3} - \tau_p} B_{p,3}(s_n) + \frac{\tau_{p+4} - s_n}{\tau_{p+4} - \tau_{p+1}} B_{p+1,3}(s_n)$$

(3) 构建求控制点的系数矩阵。控制顶点又称 Deboor (德布尔) 点, 可以根据 Deboor-Cox 的公式 3-9 的递推式求出反求控制点的系数矩阵:

$$\Phi = \begin{Bmatrix} B_{1,4}(s_1) & \cdots & B_{p,4}(s_1) \\ \vdots & \ddots & \vdots \\ B_{1,4}(s_N) & \cdots & B_{p,4}(s_N) \end{Bmatrix} \quad (3-16)$$

(4) 建立控制点的反算方程, 求解控制点序列。利用非均匀三次 B 样条基解析表示轨迹数据的过程, 可以由公式 3-12 计算。其中, 控制点序列和轨迹序列分别表示为:

$$F^{XY} = \begin{bmatrix} c_1^X & c_1^Y \\ \vdots & \vdots \\ c_p^X & c_p^Y \end{bmatrix}, \quad T^{XY} = \begin{bmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_N & y_N \end{bmatrix} \quad (3-17)$$

表 3-1 基于 B 样条曲线拟合误差的控制点数的选择算法

算法 1 基于 B 样条曲线拟合误差的控制点数的选择算法

1. **初始化** 读入原始轨迹的数据集 trajectories (假设有 N 个), 枚举所有的控制点数的可能值构成集合 ControlPoints (假设有 M 个)。
2. **对集合 ControlPoints 中的一个 p 值**
 - 2.1 计算归一化后的轨迹, 即在图像中的目标水平、垂直方向坐标序列 $[x \ y]$,
 - 2.2 对集合 trajectories 中的每条轨迹序列 T ,
 - (a) 计算 LCSCA 表示序列 F , 由控制点 $[C_x, C_y]$ 构成,
 - (b) 计算表示序列 F 中各个控制点与轨迹序列 T 的平均距离, 记作 average_distance, 其中, 一个控制点与轨迹的距离定义为该拟合曲线点与所有轨迹点的最小距离,
 - 2.3 结束 2.2 循环。计算集合 trajectories 上的误差 deviation, 其中 deviation 定义为 N 个 average_distance 值的中值, 即 deviation=median of {all average_distances}。
3. **计算误差** 结束 2 的循环, 得到 M 个 deviation 值,
4. **选择控制点个数** 选择 deviation 值的二阶变化率为零时所对应的最小 p 值。

此后，为了提高轨迹表示的准确性，本文还对上述所求的表示序列做了归一化处理，并在实验数据集上根据 B 样条曲线拟合误差来选择最优的控制点个数，具体算法见表 3-1。

3.2 稀疏重构研究综述

稀疏重构分析 (Sparse Reconstruction Analysis, SRA) 是近几年来信号处理机器学习和模式识别等领域中的热点研究问题之一。它起源于压缩感知理论，由斯坦福大学的 David Donoho，加州理工学院的 Emmanuel Candes 以及加州大学洛杉矶分校的陶喆轩等人提出的压缩感知理论^[72-76]，是对奈奎斯特采样定理的突破，理论中信号可以以低于奈奎斯特采样频率的频率进行采样，并且仍然能够精确重建原始信号。

3.2.1 L_1 范数研究综述

早在范数理论创建时，许多数学研究者曾经详细研究和分析了 L_1 范数的性质^[77]，但是在实际应用中，同样作为数学空间中某种距离度量的 L_2 范数则备受关注。在分析各个范数的应用范围前，首先简单介绍下 L_0 范数、 L_1 范数和 L_2 范数以及稀疏的概念。

定义 5 (L_0 范数) L_0 范数是指向量中各个元素非零项的个数，公式为：

$$L_0(w) = \sum_{i=1}^k \{1: w_i \neq 0\}. \quad (3-18)$$

定义 6 (L_1 范数) L_1 范数是指向量中各个元素绝对值之和，公式为：

$$L_1(w) = \sum_{i=1}^k |w_i|. \quad (3-19)$$

定义 7 (L_2 范数) L_2 范数也称欧几里德距离的平方根，公式为：

$$L_2(w) = \sum_{i=1}^k |w_i|^2. \quad (3-20)$$

定义 8 (稀疏) 当参数向量 w 的大部分元素是 0 时，这个向量就是稀疏的。

由上述定义可知，稀疏意味着仅使用少量的数据单位来有效表示典型的数据向量，可以实现特征自动选择和模型的可解释性。 L_0 范数和 L_1 范数的规则项

都可以实现稀疏,但 L_1 范数因具有比 L_0 范数更好的优化求解特性而被广泛应用。这是因为 L_1 范数可以看做 L_0 范数的最优凸近似,具有连续但非光滑的性质。换言之,当规则化算子在 $w_i = 0$ 处虽不可微却可以分解为一个求和形式时,这个规则化算子就可以实现稀疏。此外, L_2 范数的规则项 $|w_i|^2$ 可以用来限制模型空间,用于防止模型学习过程中的过拟合,从而提升模型的泛化能力。

在 L_1 范数的稀疏性被 Candes 等^[70]在理论上证明后,关于 L_1 范数的应用研究变得越来越多。Chen^[78]提出使用 L_1 范数最小化方法建立优化模型解决回归问题,并使用基追逐法去求解。接着, Tibshirani^[79]在统计学刊上发表 Lasso 算法,该算法针对回归问题,建立最小二乘误差和 L_1 范数正则化的约束模型,使用贪心算法求解。这项研究起初并未引起学者们的广泛关注,大多数研究人员也没有意识到 L_1 范数和稀疏性之间的联系。直到压缩感知理论在信号处理领域中的出现,稀疏表示的问题也应运产生,人们才考虑到使用 L_1 范数作为 L_0 范数的一个近似估计去重构信号。

3.2.2 稀疏重构的基本原理

随着信号处理领域压缩感知理论^[71]的提出,稀疏重构理论才越来越受到学者们的重视。压缩感知^[80]源自泛函分析和逼近论,由 Candès、Tao、Romberg、Donoho 等创立。具体是指,当信号在某个变换域是稀疏或可压缩的时,可以利用与变换矩阵非相干的测量矩阵将变换系数线性投影为低维观测向量,同时这种投影保持了信号重建时所需的信息,通过进一步求解稀疏优化问题就能够从低维观测向量精确地或高概率精确地重构原始信号。

信号的稀疏表示是指信号在一个过完备的基上所求得表示是稀疏的,这个集合可以是向量或矩阵,大部分元素为零。由于过完备字典中的信号元素具有冗余,各信号元素的表示向量之间并不是线性无关的,因此新的信号元素可以由多种由字典中的信号元素构成线性组合。稀疏重构就是找到一种由少数信号元素构成的线性组合,利用过完备的元素集合,将信号表示成这些少数元素的线性组合,稀疏表示可以使得信号只集中于较少的元素上,非零系数的元素则揭示了信号的主要特征和内在结构。因此,稀疏重构具有一定的抵抗噪声的能力,能够简洁而准确地表达信号。如图 3-6 所示,字典 D 的每一列为一个信号

元素，信号 x 由字典 D 中的信号表示产生稀疏向量。

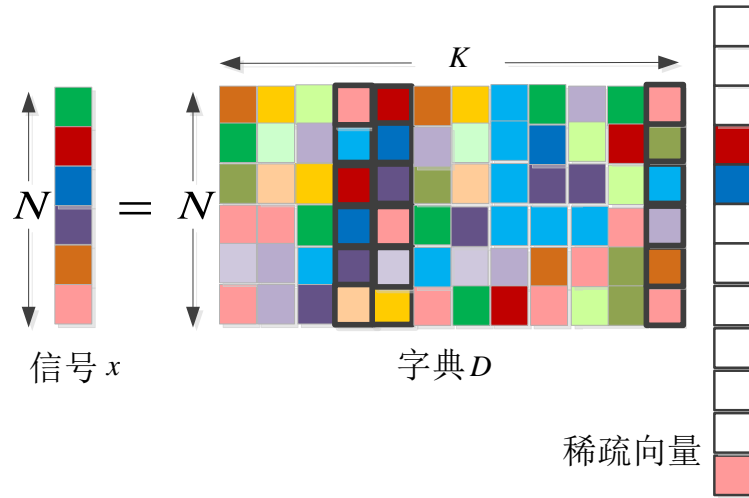


图 3-6 稀疏重构示意图

概括地讲，稀疏重构的模型需要确认三个问题：过完备字典、满足非相干性的测量矩阵和快速鲁棒的信号重建。针对这三个问题，设计压缩感知的 L_0 范数稀疏重构模型为：

$$\min_x \|x\|_0, \quad s.t. y = Ax. \quad (3-21)$$

其中， A 为观测矩阵， x 为原始信号经过稀疏变换后的表示系数， y 是接收端对信号进行重构之后的结果。 $\|\cdot\|_0$ 为 L_0 范数，表示系数向量非零项的个数。于是，信号的稀疏重构问题最终转化为和式的最优化问题。当 A 是 Hilbert 空间的一个正交基时，稀疏向量和优化问题都容易求解，但是这对于一个冗余字典却是 NP-hard 问题。为了求解这个问题，2003 年 Elad 和 Bruckstein^[81] 研究证明：如果矩阵 A 满足： $\sigma(A) \geq 2\|x\|_0$ ，那么 L_0 范数最优化问题具有唯一解。 $\sigma(A)$ 指的是矩阵 A 中最小的线性相关的列向量集所含的向量个数。虽然上述证明模型存在唯一解，但是其求解过程仍然是 NP-hard。2006 年，Tao 和 Candes^[82,83] 证明了在 RIP (Restricted Isometry Property, 限制等距性质) 条件下， L_0 范数优化问题与公式 3-19 中 L_1 范数优化问题具有相同的解^[78]。其中，RIP 条件要求稀疏信号在观测矩阵作用下必须保持几何性质一致。这意味着，如果 x 在 A 中有非常稀

疏的表示，那么该表示可以通过求解较为简单的 L_1 范数问题来精确重构。

$$\min_x \|x\|_1, \quad s.t. y = Ax \quad (3-22)$$

RIP 性质的提出才使稀疏表示和重构的理论和实践完成统一，随后很多研究学者将其用于模式识别等领域，促进了该理论求解的研究发展。现有稀疏重构的求解算法主要有：MP (Matching Pursuit, 匹配追踪) 方法^[84]、OMP (Orthogonal Matching Pursuit, 正交匹配追踪) 方法^[85]、LASSO 方法^[77]、BP (Basis Pursuit, 基追踪) 方法^[86]，GPSR (Gradient Projection for Sparse Reconstruction, 梯度投影稀疏重构) 方法^[87]和 Bregman 迭代法^[88]等。按照实现方式不同，这些算法可以分为三类：直接优化法、迭代收缩优化法和范数正则化优化法。

(1) 直接优化法。这类方法直接使用 L_0 范数最小化作为稀疏目标函数，优化模型与公式 3-18 的模型类似，使用贪心算法优化求解，典型的算法有匹配追踪法^[83]和正交匹配追踪法^[83]等。

(2) 迭代收缩优化法。最具代表性的 LASSO 算法是针对回归问题，建立最小二乘误差和 L_1 范数正则化的约束模型，使用贪心算法求解。它可以通过不断松弛以下模型（公式 3-23）中的 ε 因子，来实现迭代收缩求解。

$$\min_x \|Ax - y\|_2, \quad s.t. \|x\|_1 \leq \varepsilon \quad (3-23)$$

(3) 范数正则化优化法。这是目前研究比较热门的一类方法，常见的有基追踪方法^[84]，软阈值方法(如梯度投影稀疏重构方法^[85]和 Bregman 迭代法^[86])。基追踪方法是寻求从过完备字典集合中得到信号最稀疏的表示，即用尽可能少的基精确地表示原信号，从而获得信号的内在本质特性。基追踪方法采用表示系统的范数作为信号稀疏性的度量，通过最小化 L_1 范数将信号稀疏表示问题定义为一类有约束的极值问题，进而转化成线性规划问题进行求解。该算法的优点是重建精度高，需要的重建元素个数少，但是计算复杂度相对较高。

3.2.3 稀疏重构在相关领域的应用

随着稀疏重构理论的发展和不断突破，稀疏重构广泛地应用于图像处理的各个领域。近几年，信号处理、图像处理、模式识别、计算机视觉等领域的各种会议期刊上大量发表稀疏表示和重构分析的相关研究。稀疏重构在理论发展

和它对噪声的良好特性，逐渐引起相关领域研究人员的关注，在很多问题上得到了广泛的研究和应用。Mairal 等^[89]提出了基于稀疏表示的图像重建算法，它使用稀疏表示对图像进行去噪，同时对较小的区域进行修复，取得良好的结果。Wright 和 Ma^[90]等提出基于 SRC (Sparse Representation Classification, 稀疏表示分类)的人脸模式识别方法，对一张测试人脸样本由训练样本集进行 L_1 范数最小化的线性的稀疏表示，接着基于所得到一个稀疏的系数向量来求解最小稀疏重构误差，实现人脸识别。该算法在人脸有部分遮挡的时候具有很大优势，可以自动找出被遮挡的部分，并且对该部分进行重建。Yang 等^[91]提出了基于稀疏表示的超分辨率重建。他们在图像信号具有稀疏性的基础上，在高分辨率图像和低分辨率图像中建立两组对应的基。根据高低分辨率之间的对应性，从而实现了超分辨率重建。它的另一项进展在于，给定两组对应的基之后，算法不再需要当前待处理图片以外的信息，即可以用一张图片就完成超分辨率重建。在图像处理的其他方面，如图像融合、图像目标检测、图像恢复等领域也不断发展，基于稀疏表示提出了许多令人瞩目的算法。Han^[34]将 L_1 范数最小化得稀疏重构应用到多目标跟踪中，用搜索区域内的跟踪样本来线性重构需要跟踪目标，从而在视频帧的搜索区域中找到相应的跟踪目标，该方法在目标被遮挡的情况下非常有效。Zhu^[118]将卷积稀疏编码用于三维轨迹重构，该方法可以用来描述实际应用中运动的非刚性结构。

3.3 基于稀疏重构分析的模型

由于监控场景中的行为模式（样本集构造的模式）往往很多，导致使用样本集进行新样本重构的系数会很稀疏，样本集中的多数类别的样本对应重构系数趋近于零，因此在本节提出了通过稀疏重构分析获取样本稀疏重构系数和重构误差的方法。由于复杂条件的限制，在实际应用中采集到的样本轨迹通常较少，而且待识别的轨迹变化一般较大。然而，现有的稀疏重构分析方法中关于小样本、噪声问题的深入研究，正好为解决轨迹分析中的这些问题奠定了理论基础。本文提出的基于稀疏重构分析的方法意在通过学习得出一个多类的线性分类器，实现轨迹模式分类。这种方法通过重构系数和重构误差进行评估，从样本集中选择出最具有代表性的一些样本来重构新样本。这些样本揭示着新样本轨迹的主要特征和内在结构，通过对新样本的重构，从而使得轨迹分类和异

常检测具有较强的鲁棒性。

3.3.1 目标轨迹的样本集构造

监控视频中基于稀疏重构分析的异常行为检测的样本集构造，就是建立监控场景中运动目标轨迹表示的过完备集合。本文按照一定间隔获取视频序列中运动目标的位置坐标，得到了运动目标的原始轨迹数据。然后划分监控场景的出入口区域块，根据运动目标所经过的出入口区域块的不同来给原始轨迹数据分别标签。接着，监控场景中所有的训练轨迹经过 LCSCA 轨迹表示算法变换为的表示序列，最后，这些表示序列构成稀疏重构分析中的样本集合。其中，每一条轨迹的表示序列是样本集的一个列向量。具体地，监控场景中的每种典型正常行为分别建立样本集合 $A_j = \{a_j^1, a_j^2, \dots, a_j^K\}$ ，其中， K 表示第 j 种行为的轨迹样本个数。将所有行为的轨迹样本合并，形成监控区域轨迹样本集合，如下：

$$B = \cup\{A_j\} = \{a_1^1, a_1^2, \dots, a_1^K, a_2^1, a_2^2, \dots, a_2^K, \dots, a_J^K\}, \quad (3-24)$$

其中， $j = 1, \dots, J$ 。 J 表示监控区域中典型正常与异常行为的种类数目。因此，我们可以得到一个 $K \times J$ 条样本轨迹的监控区域行为集合 B 。

3.3.2 目标轨迹的稀疏重构分析

由于监控场景中往往有很多的行为模式，所以构造的样本集也很大，这便会导致使用样本集进行新样本重构的系数会很稀疏，样本集中的多数类别样本对应的重构系数趋近于零。这样的稀疏重构系数在实际中保证了样本集中只有少数和新目标轨迹相似的样本会用于新样本的重构。以目标轨迹的表示序列作为样本，它可以由上述所构造的样本集来线性重构得到，如公式：

$$B\psi \approx F, \quad (3-25)$$

其中， F 是轨迹样本的 LCSCA 轨迹表示， $\psi = \{\psi_j^k\}$ 是对应 B 中样本的系数向量，且 $j = 1, \dots, J, k = 1, \dots, K$ ， ψ_j^k 就是第 j 种行为的第 k 个样本的重构系数。

实际应用中， $\psi = \{\psi_j^k\}$ 中只有 l 个是非零的系数 ($l \ll k$)。在数学中，我们将这

样的矩阵称为样本的 l 稀疏表示。非零系数的个数可以用 $\|\psi\|_0$ 来表示，所以我们可以通过最小化 $\|\psi\|_0$ 获得样本的稀疏重构。这个 L_0 范数的问题是 NP-hard 的问题。根据 3.2 节的理论，可以通过 L_1 范数最小化可以用来解决稀疏表示问题：

$$\arg \min \|\psi\|_1, \quad s.t. \|B\psi - F\|_2 \leq \varepsilon, \quad (3-26)$$

其中， $\|\cdot\|_1$ 表示 L_1 范数， $\varepsilon > 0$ 表示一个较小值。

通过样本集 B 和稀疏系数向量 ψ ，目标轨迹就可以通过重构来得到，重构公式为：

$$F \approx B\psi = \sum_{k=1}^K B^k \psi^k. \quad (3-27)$$

基于 L_1 范数最小化的稀疏重构，保证了该重构是很紧密的，也就是说，样本集中样本代表着目标轨迹的某一部分或者与整条目标轨迹相似。我们从样本集中挑选出最有代表的样本来重构目标轨迹，这样可以有效地通过稀疏重构系数和重构误差来揭示目标轨迹的主要特征和内在结构与样本集中的哪些样本相似一致，使得稀疏重构系数和稀疏重构误差用作轨迹分类和异常检测时具有较强的鲁棒性。

3.4 分类器的构建

3.4.1 轨迹分类模型

当目标轨迹重构完成，我们由公式 3-26 计算求出其对所有样本集合 B 的稀疏系数 ψ 。随后，对于每一种行为 j ，本文定义其对应的特征函数 δ_j 。其中 δ_j 定义为只保留轨迹样本集合 B 中第 j 种轨迹所对应的稀疏系数，同时将其他种类样本所对应的稀疏系数都赋值为 0。基于上述定义，我们利用重构误差 $r_j(F)$ 来构建目标轨迹分类的模型函数如下：

$$r_j(F) = \|F - B\delta_j(\psi)\|_2, \quad j = 1 \dots J \quad (3-28)$$

$$\text{Classify}(F) = \underset{j}{\operatorname{argmin}} r_j(F). \quad (3-29)$$

其中 $r_j(F)$ 是用 ψ 的特征函数值来重构 F 时的第 j 个重构误差值，该模型取重

构误差最小时的 j 为 F 分类。

3.4.2 异常检测模型

同样地，基于上述定义 δ_j 与集合 B ，利用重构比值来判断目标轨迹是否属于异常行为。基于稀疏重构的异常行为检测算法见表 3-2，其中，异常检测模型函数如下：

$$\text{Detect}(F) = \text{sign}(\min_{j=1, \dots, J} H_j - \theta), \quad H_j = \frac{1/r_j(F)}{\sum_{i=1}^J 1/r_i(F)}. \quad (3-30)$$

其中， θ 表示判断阈值，实验中经验阈值 $\theta=0.03$ 。对于公式 3-30，若 $\text{Detect}(F) < 0$ ，则目标轨迹 F 为异常行为。这意味着，当且仅当所有的重构比值都大于阈值 θ 时，目标轨迹才被判断不属于异常行为。当目标轨迹被判断为不属于异常行为时，我们还可以通过公式 3-29 将目标轨迹样本 F 分类。

表 3-2 基于稀疏重构的异常行为检测算法

算法 2 基于稀疏重构的异常行为检测算法

1. **初始化** 初始化目标轨迹，并且计算其特征 F 。
 2. **构造样本集** 对目标轨迹 F 构建目标轨迹的样本集 S ，计算它们的特征集 B 。
 3. **获取目标轨迹稀疏重构**
 - 3.1 使用 B 计算目标的稀疏系数 ψ ，
 - 3.2 初始化 $j=1$ ，
 - (a) 计算稀疏系数 ψ 的 δ_j 的特征函数值，
 - (b) 基于样本 S 和 $\delta_j(\psi)$ 重构目标轨迹，
 - (c) 计算重构误差值 $r_j(F)$ ，
 - (d) 跳转到 3.2(a)，直至 $j=J$ 时结束循环。
 4. **检测异常** 计算重构比值 H_j 用于检测。
-

3.5 实验结果分析

3.5.1 本文实验环境配置

本文重点在于研究轨迹分析，因此分别选用了三个公共测试的数据集 CAVIAR^[38]、NGSIM^[39]、和 Carpark^[40]的跟踪标注轨迹数据来做研究。视频和轨迹数据包括室内大厅监控人、路口监控车辆、停车场监控车辆的数据集，其中包括了轨迹形状、方向、长度多样和轨迹不完整、有噪声和局部有变化等轨迹分析中的难点问题。表 3-3 介绍了监控视频中的目标轨迹数据和正常/异常行为区分的样本个数，一些正常行为和异常行为的轨迹样本示意如下图 3-7。其中，图 3-7(a)是 CAVIAR 中的部分正常行为轨迹示意图，图 3-7(b)是 CAVIAR 中的部分异常行为轨迹示意；与之类似，图 3-7(c)和图 3-7(d)分别对应 Carpark 的正常和异常行为轨迹示意。

表 3-3 视频监控轨迹分析的轨迹数据集列表

数据集	介绍
CAVIAR	视频分辨率 640×460，室内监控人的正常/异常行为轨迹样本共 2675 个，轨迹类别个数为 22。
NGSIM	视频分辨率 720×576，路口监控车的正常/异常行为轨迹样本共 1213 个，轨迹类别个数为 8。
Carpark	视频分辨率 330×288，停车场的正常/异常行为轨迹样本共 269 个，轨迹类别个数为 16。

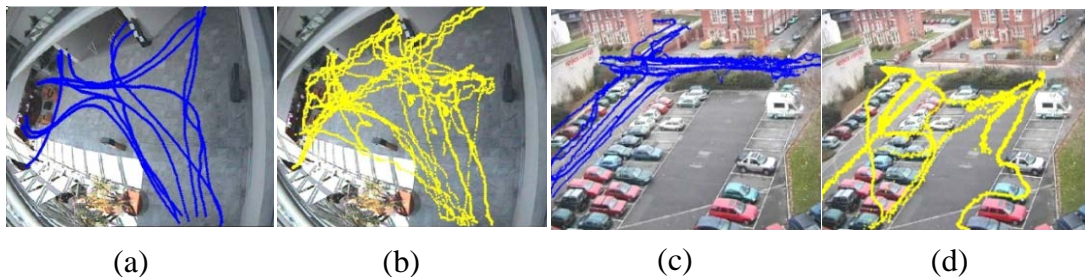


图 3-7 部分正常/异常行为的轨迹样本示意图

为了形象并且量化地评价本章所描述的算法性能，分类和检测异常的效果

采用检测精确度 $DACC$ (Detection ACCuracy)和正确分类率 CCR (Correct Classification Rate)来衡量，其定义如下：

$$DACC = \frac{TP+TN}{\text{测试轨迹数}} , \quad (3-31)$$

$$CCR = \frac{\text{分类正确的测试轨迹数}}{\text{正常行为的测试轨迹数}} , \quad (3-32)$$

其中, TP 是指根据算法判断为正常行为的正常轨迹数, 称作判断为真的正确率; TN 是指根据算法判断为异常的异常轨迹数, 称作判断为假的正确率。从定义可知: 算法的性能越好, 其 $DACC$ 便会越高; 随着选作重构样本集的样本数 K 的增加, $DACC$ 的尽可能在越小的样本集取得较高值则越好, 这样便可以直观地比较算法的性能。此外, CCR 越高说明算法在分类性能上也越好。本文是在 Intel Core(TM) i7 CPU(2.93GHz) 4GB RAM 的机器上完成实验的。

3.5.2 实验结果比较与算法分析

3.5.2.1 轨迹表示中控制点数选择的实验

根据第 3.1 节所提出的基于 B 样条曲线拟合误差的控制点数的选择算法, 本章节优化选择最小二乘三次 B 样条近似轨迹表示中控制点的个数 p 。

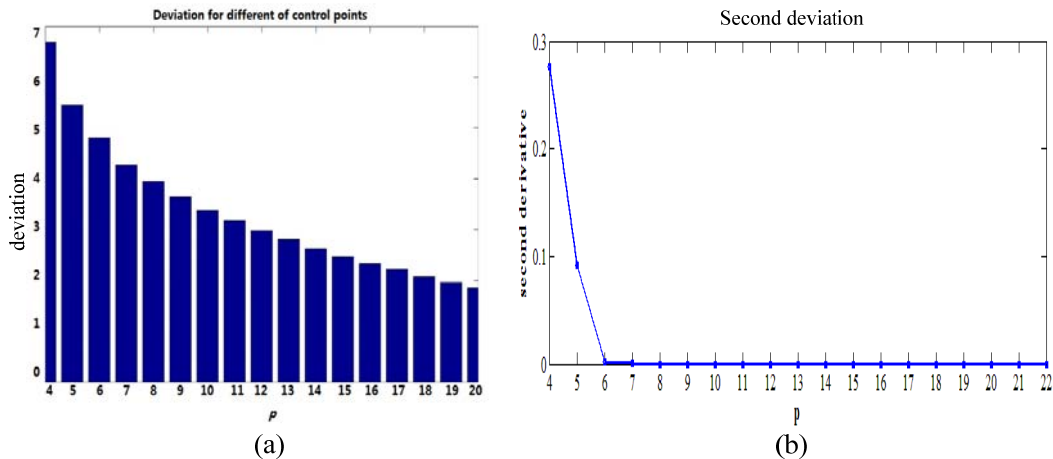


图 3-8 基于 B 样条曲线拟合误差的控制点数的选择示意图

图 3-8 示意了该算法的优化过程, 其中, 图 3-8(a)展示了不同控制点数时的

B 样条曲线拟合误差变化趋势，可以看出，随着控制点数的增大，拟合误差值的确会越来越小。图 3-8(b)显示了本文是如何根据二阶变化率选择控制点数。随着控制点数的增加，拟合误差的二阶变化率逐渐减小，本文选择下一个二阶变化率为零的相应控制点数值，即 $p = 7$ 。在此参数选择的基础上，本文还做了轨迹表示的研究对比。采用两种表示方法： $R1$ 直接用本文第 3.1 节的算法求出的表示序列 $F = \{C_1^x, C_2^x, \dots, C_p^x, C_1^y, C_2^y, \dots, C_p^y\}$ ； $R2$ 在表示序列 F 首尾加入轨迹的起始点和终止点的坐标。这组对比实验用于检验最小二乘三次 B 样条轨迹表示算法对轨迹分类和异常检测研究的影响。此外，后文实验中还选用了几个与本节所选最优控制点个数相近的数值 $p = \{5, 7, 9, 11\}$ 来作对比分析，这组对比可用于检验本节中所做的控制点个数 p 的最优化选择是否有效。

3.5.2.2 基于稀疏重构模型的轨迹分析实验

以室内大厅监控人的数据集 CAVIAR^[38]为例，场景中有 11 种由不同的出入口对构成的轨迹路线（如图 3-9 所示），再考虑到每条路线方向不同的区别，于是就有 22 种不同种类的轨迹。

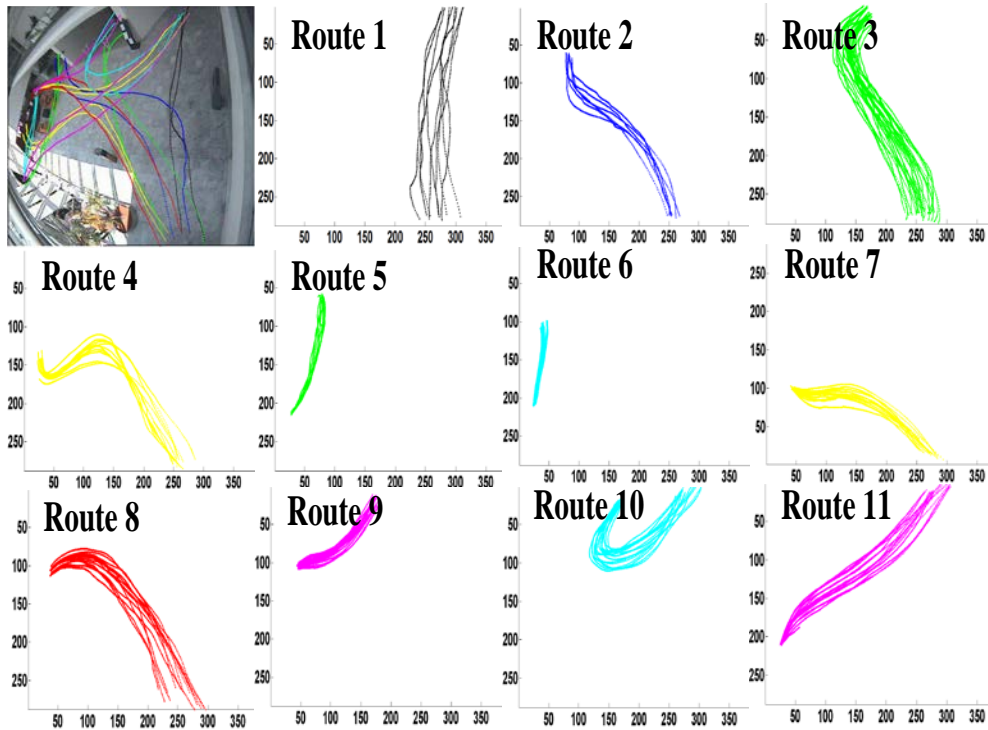


图 3-9 CAVIAR 样本集中轨迹类别的示意图

本文对监控场景中的大量目标正常行为的轨迹 2200 条（每种轨迹 100 条）做了样本集构造。利用这个样本集，我们做了 40 个测试数据的实验。其中，测试数据包括了 21 条场景中运动的人直接从一个入口走到另一个出口的运动轨迹数据，19 条场景中运动的人打斗、晕倒、停留、徘徊的运动轨迹数据。图 3-10 展示了该数据集中正常轨迹（图 3-10(a)）及其 LCSCA 表示（图 3-10(c)）和异常轨迹（图 3-10(b)）及其 LCSCA 表示（图 3-10(d)）的对比。

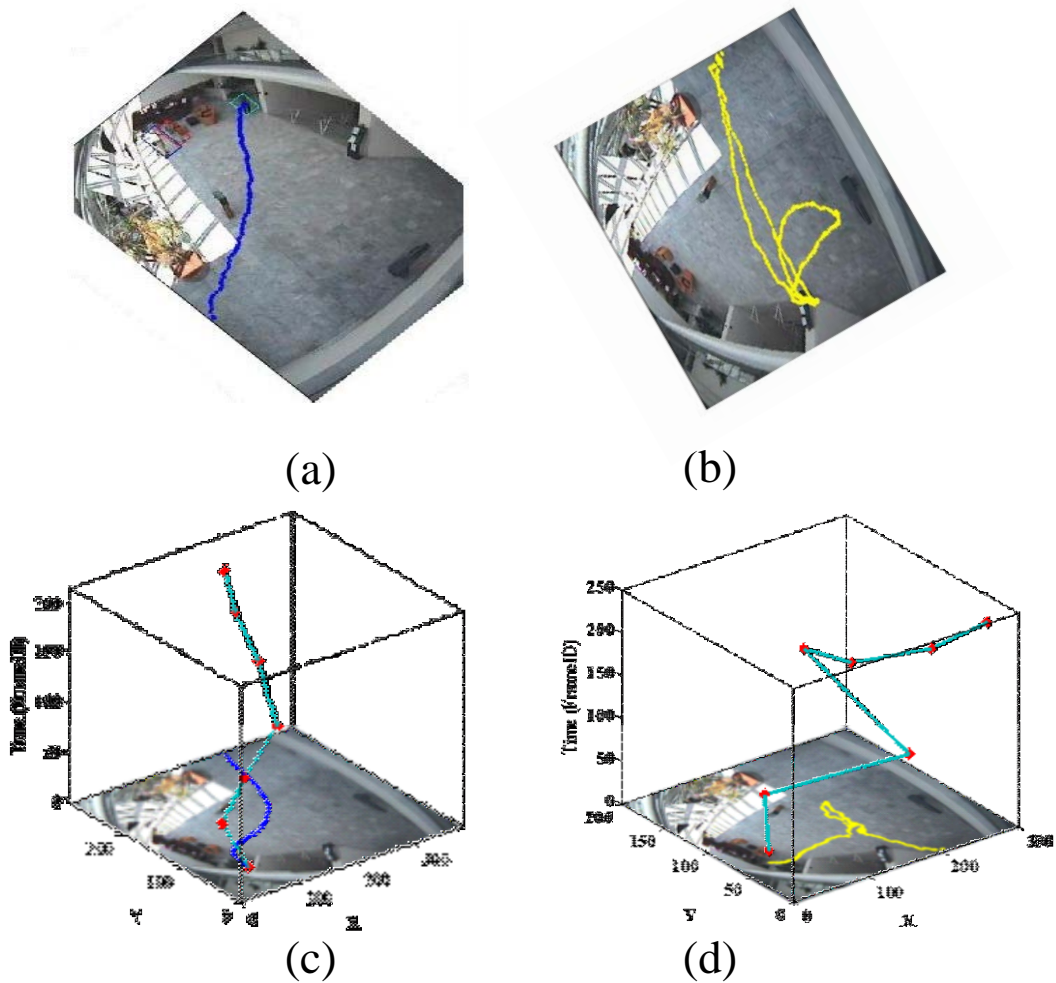


图 3-10 CAVIAR 样本集中正常/异常轨迹的 LCSCA 表示对比图

此外，图 3-11 给出了目标轨迹基于稀疏重构的目标轨迹分析实例。图 3-11(a)

红色标识是蓝色目标轨迹的控制点，该控制点表示在样本集上经过基于 L_1 范数最小化的稀疏重构后，得出稀疏重构系数如图 3-11(b)。可以看出，有很少的一些系数为非零，这意味着利用这些少数的样本重构的目标轨迹，重构错误率会比较低。与之相对应，图 3-11(c)中的稀疏重构误差最小时正好对应该轨迹的分类图 3-9 中的 Route7。最后，图 3-11(d)示例，仅当所有的重构比值都大于阈值 θ 时（红色虚线示意），目标轨迹才被判断不属于异常行为。

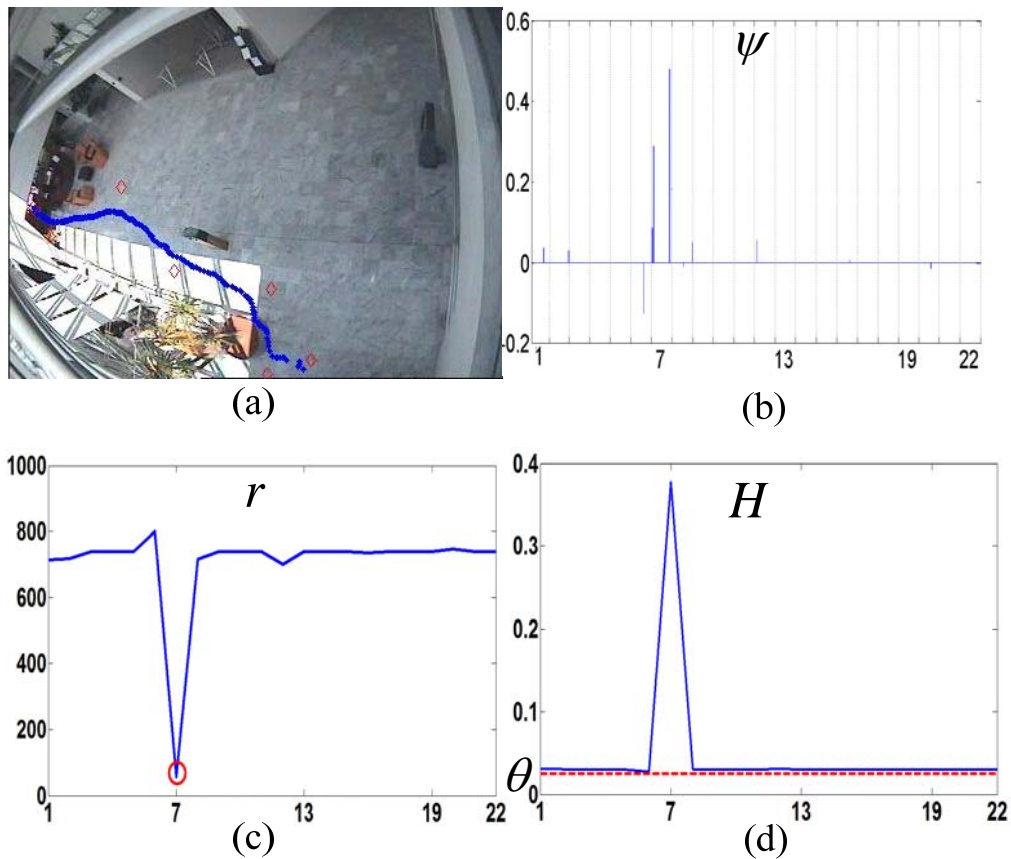


图 3-11 目标轨迹基于稀疏重构的轨迹分析实例

图 3-12 是对于 RI 表示，选取不同的控制点个数时，DACC 和 CCR 随着重构样本集的样本数 K 增加的变化曲线。图中所示结果是 10 次随机试验的平均结果。可以看出， $p=5$ 时，图 3-12(a) 的红色曲线上的检测精确度可以达到 $DACC=90.42\%(\pm 3.85\%)$ ，图 3-12(b) 的红色曲线的分类正确率可以达到 $CCR=70.09\%(\pm 6.13\%)$ 。尤其是当重构样本集的样本很少 ($K=5$) 时，算法就可以取得较好的效果，DACC 达到 84.75%，CCR 达到 62.38%。

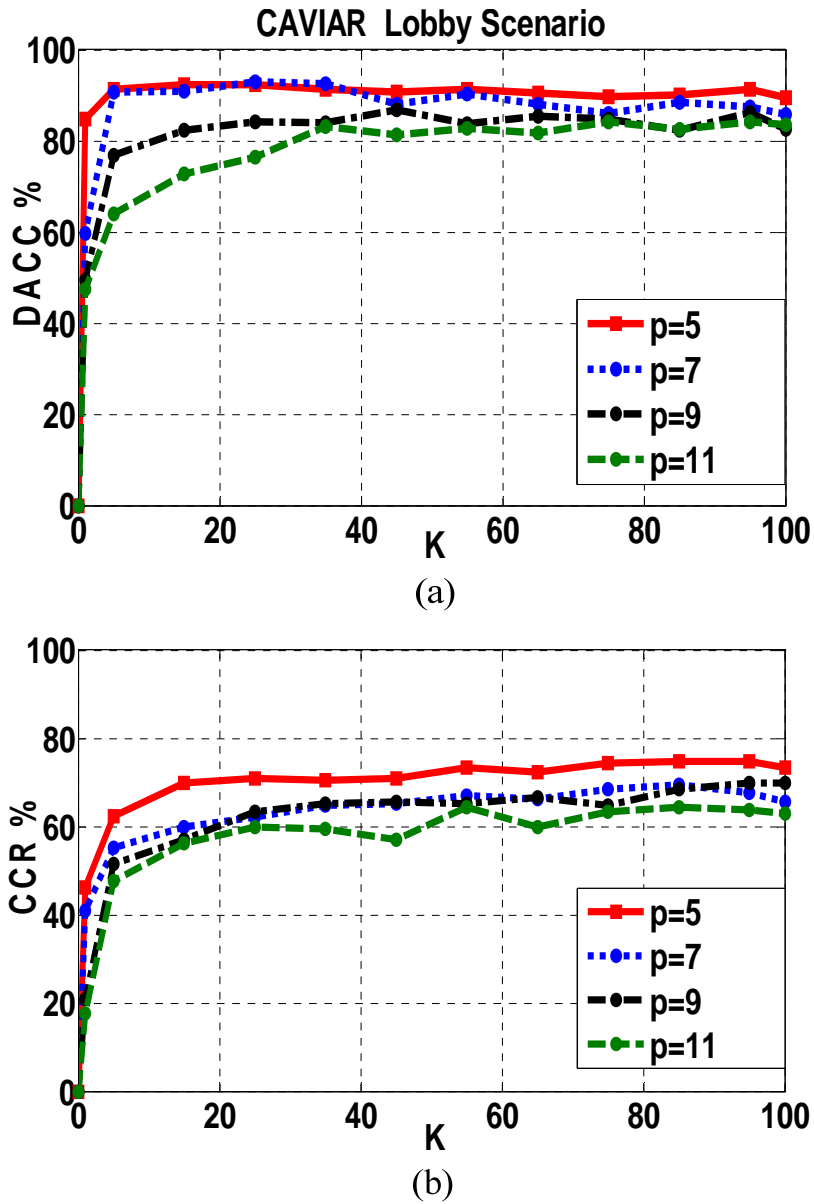


图 3-12 $R1$ 表示时轨迹分类和异常检测性能对比图

图 3-13 则对 $R2$ 表示, 选取不同的控制点个数时, DACC 和 CCR 随着重构样本集的样本数 K 增加的变化曲线。可以看出和图 3-12 所示相似, $p=5$ 时的算法性能最好, $p=7$ 时次之。图 3-13(a) 的红色曲线 $p=5$ 时检测精确度可以达到 $DACC=83.75\%(\pm 2.88\%)$, 图 3-13(b) 的红色曲线 $p=7$ 时分类正确率可以达到 $CCR=68.15\%(\pm 4.53\%)$ 。尤其是当重构样本集的样本很少 ($K=5$) 时, 算法就

可以取得较好的效果，DACC 达到 81.75%，CCR 达到 63.81%。对比图 3-12 和图 3-13 的结果， $R1$ 表示比 $R2$ 表示的结果略优。我们分析这可能是由于不同的运动目标轨迹往往有着相近的出入口坐标，使得表示序列中的这两维数据特征区分性不大，导致 $R2$ 表示对一些相类似的轨迹难以分辨。但是，这也说明了 LCSCA 轨迹表示算法所得到的表示序列用于样本模式分类的特征十分有效。

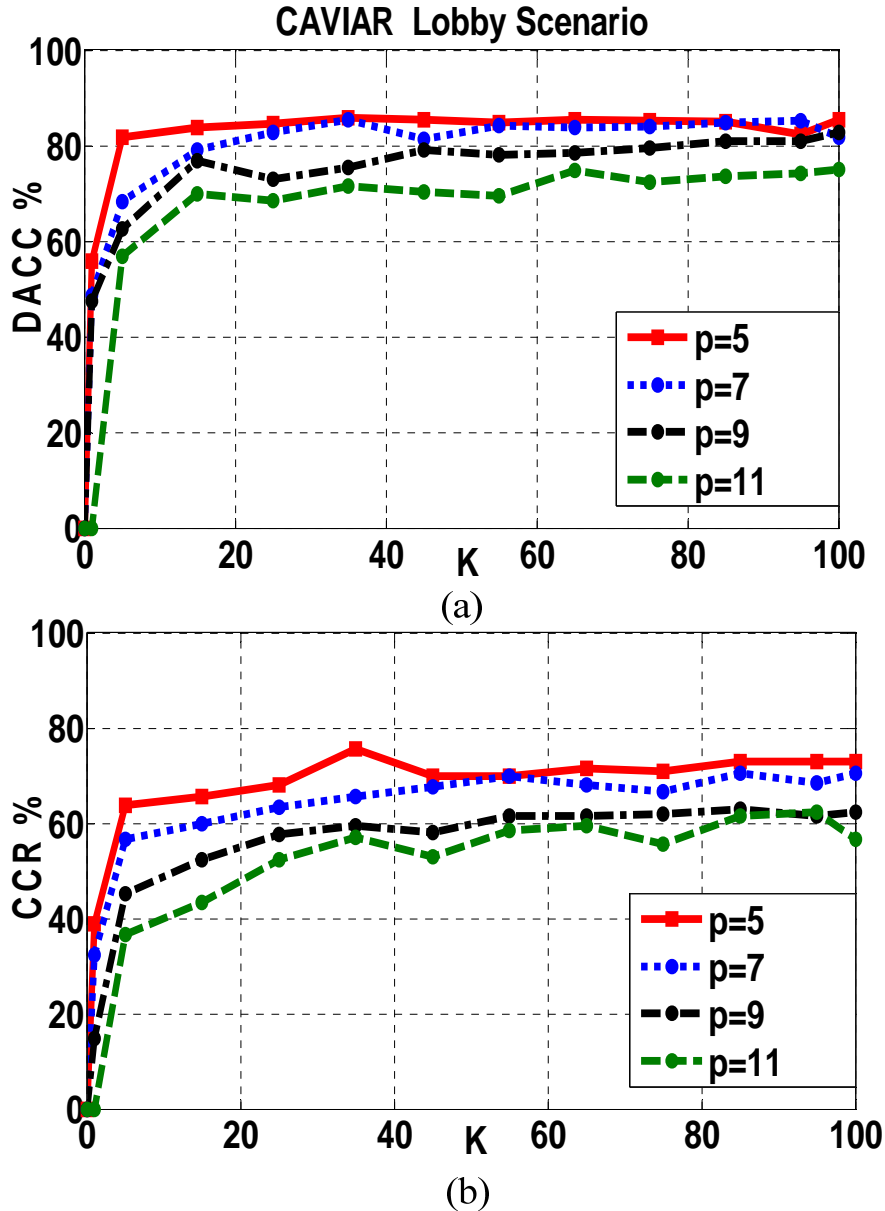


图 3-13 $R2$ 表示时轨迹分类和异常检测性能对比图

此后，我们还将基于目标轨迹分析的异常行为检测算法与本文 2.4 节介绍的

GMMs 算法做了更加深入的检测性能对比，得到了两种算法随着重构样本集的样本数 K 增加的 DACC 曲线对比图。图 3-14(a)给出 CAVIAR 数据集的大厅场景中行人轨迹的异常行为检测性能，图 3-14(b)给出 NGSIM 数据集中 Lankershim 街道上车辆的异常行为检测性能。

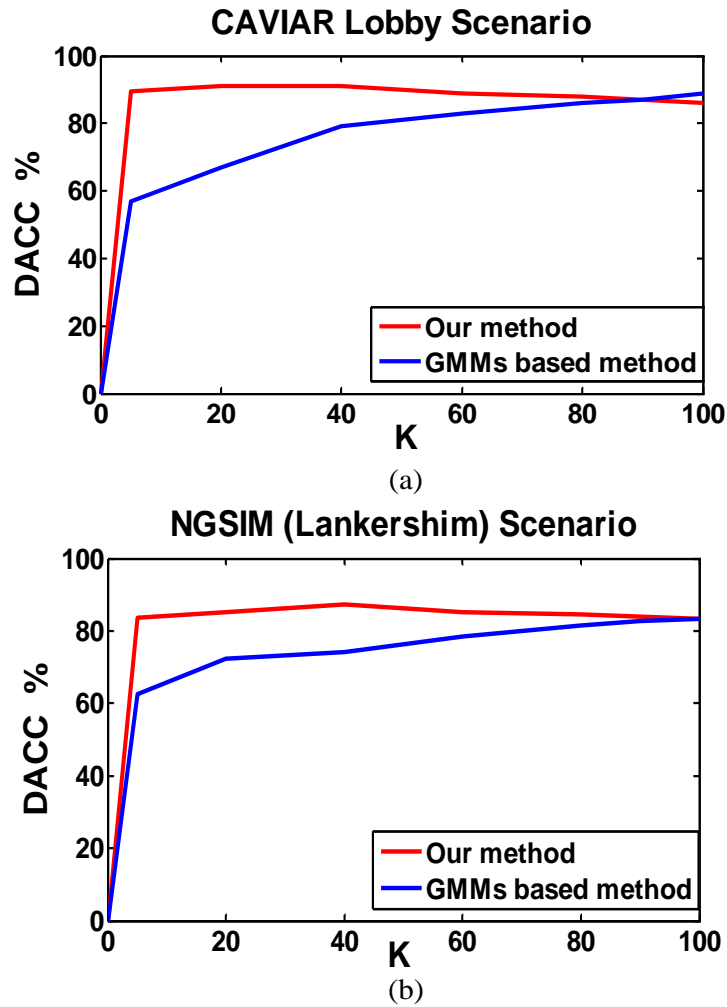


图 3-14 CAVIAR 和 NGSIM 数据集上检测性能对比图

3.6 本章小结

考虑到轨迹的多样性，本章对提出最小二乘三次 B 样条近似 (LCSCA) 的表示算法进行了论述和研究。首先，根据 B 样条基函数、B 样条曲线、Deboor-Cox 算法和函数近似的理论，概括出 LCSCA 目标轨迹表示的基本概念和用于轨迹表示的原理。接着，详细介绍了基于 LCSCA 目标轨迹表示算法，提出控制点

个数的参数选择问题需要进一步研究。最后，比较分析了不同的控制点个数的参数时控制点序列与原始轨迹数据的拟合情况，给出了具体的基于 **B** 样条曲线拟合误差的控制点个数的选择算法。

针对轨迹分类和异常检测问题，本章对所提出的基于稀疏重构分析的异常行为检测算法进行研究。首先，介绍了稀疏重构分析的基本理论和在相关领域的稀疏优势，以及 L_1 范数的求解方法。然后，论述了目标轨迹的样本集构造、目标轨迹的稀疏重构分析，详细说明了如何利用重构分析得到的稀疏重构系数和重构误差来检测异常行为。最后，给出了该算法的检测结果实验，从而对算法在目标轨迹长度不一、运动目标的运动方向多样等情况下的鲁棒性进行了验证，同时将该方法对于训练样本集较小时的有效性与其他算法进行了比较和分析。

第四章 基于局部约束下稀疏重构的目标轨迹分析

在实际应用中，由于监控场景中遮挡、背景变化等的影响，大部分由跟踪获取的轨迹数据是不完整的，具体表现为这些轨迹往往带有噪声和局部形变。为了解决这一问题，本章提出在轨迹学习模型中用轨迹划分的方法来引入局部性约束，设计出每段样本在其相似局部样本子集进行稀疏重构的模型，并依此构建一个加权的局部判别式字典，最后根据所定义的判别编码和损失加权解码算法，实现整个样本空间的轨迹分类和异常检测^[114]。这就是本章提出的基于局部约束下稀疏重构（Locality-constrained Sparse Reconstruction, LSR）的目标轨迹分析算法，图 4-1 显示了局部约束下稀疏重构的流程步骤图。

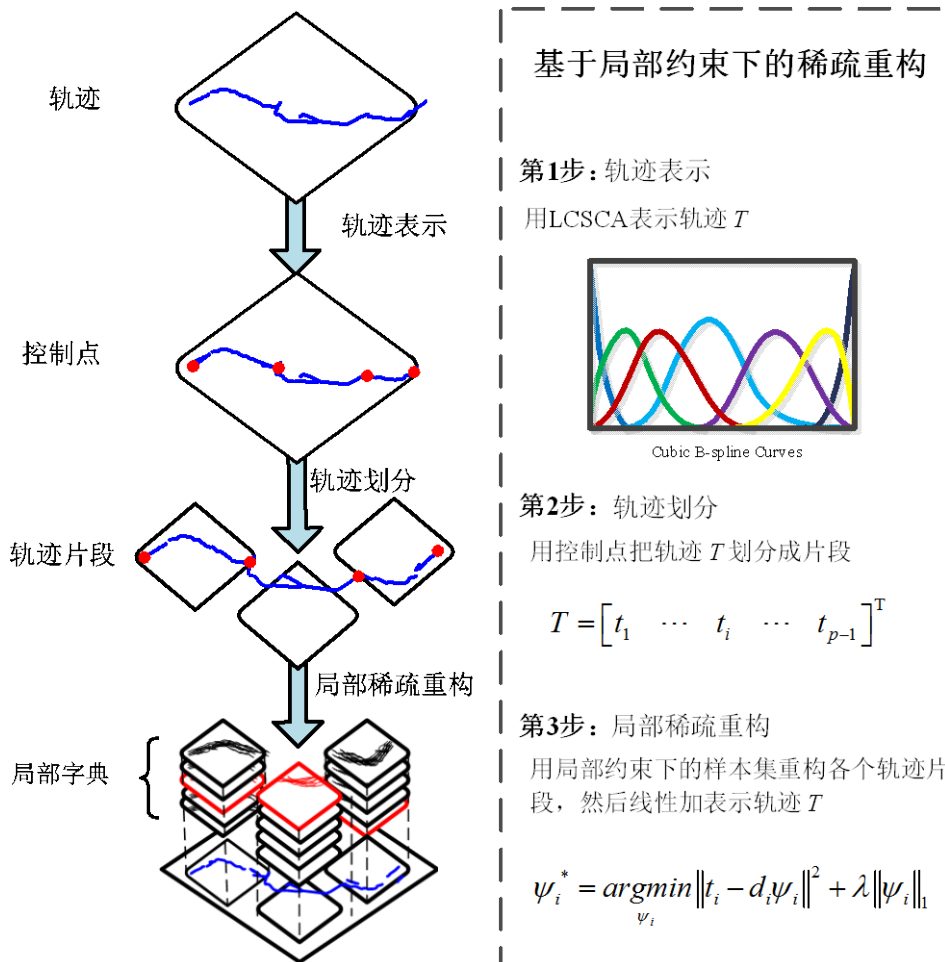


图 4-1 基于局部约束下稀疏重构的流程图

4.1 局部约束下稀疏重构分析的思路

局部约束下稀疏重构分析旨在借鉴文献^[92,116]中局部模型的思想，利用局部样本子集来重构测试轨迹，提出分段编码和基于损失加权解码算法，以提高轨迹在不完整、噪声和局部有形变情况下的分类和异常检测性能。

4.1.1 局部约束的定义

为了实现上述思路，首当其冲需要定义“什么是局部约束”。在轨迹学习中，局部约束包含两层含义：一是每个轨迹点仅和同一条轨迹上的若干个相邻的轨迹点有关联，即轨迹间的比较单位是若干个相邻的轨迹点组成；二是每个轨迹点的异常程度和类别计算仅涉及到其邻域内的轨迹点构成的基本比较单元，即基本比较单元仅和其邻域范围内的基本比较单元相比较，而不是整个数据空间。图 4-2 示意了轨迹的局部约束。

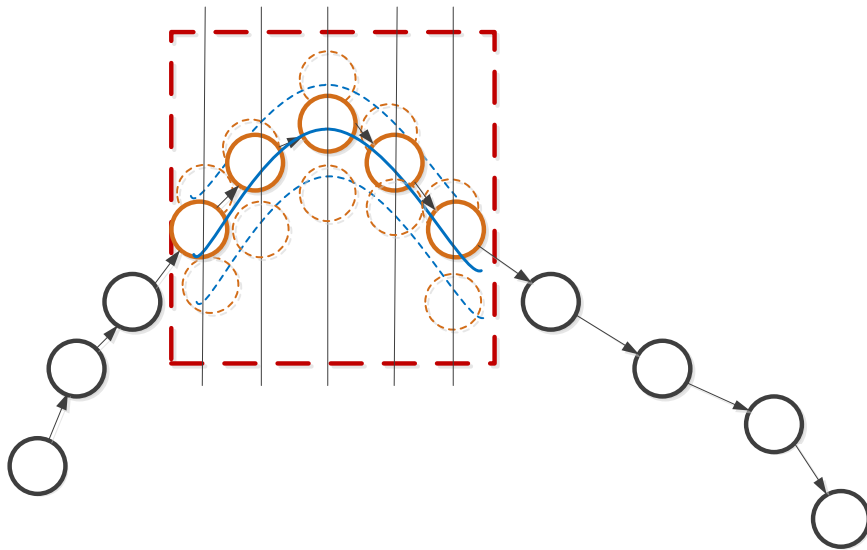


图 4-2 轨迹的局部约束示意图

有了上述的“局部约束”定义，我们可以循序渐进形成整个算法的思路。首先，按其定义来划分轨迹，本文采用的基本比较单元并不是固定长度的轨迹片段，而是按与轨迹形状相关的点来截取的轨迹片段，相比之下，物理描述意义更为合理。然后，利用局部约束，按照轨迹特征，计算基本比较单元与其邻域内的基本比较单元之间的稀疏重构误差，这可以理解为分段计算其局部相似

度。最后，在整个数据空间内，通过对所有分段结果的判别编码和加权解码，实现整条轨迹的判别。

4.1.2 局部字典与全局字典

本文前面第 3.2 节讲过，经典的稀疏重构问题是：

$$\min_x \|Dx - y\|_2, \text{ s.t. } \|x\|_1 \leq \varepsilon. \quad (4-1)$$

相应地，稀疏编码就是指在字典 D 已知的情况下，求解信号 y 的稀疏表示系数的过程，通常也被称作原子分解^[93]。在基于局部约束下稀疏重构的学习模型中，已知的训练样本集合作为给定字典，求解测试集在该字典上的稀疏表示系数。这里缺少字典更新过程，因此，该模型只是稀疏重构而有别于字典学习，具体的字典学习基本概念下一章将会详细解释。

在计算机视觉领域，当我们试图用一个字典匹配图像中各种各样的结构信息时，往往为了达到更好的表示结果，而要求字典兼顾各种纹理、边缘信息。然而，现实中即使同一幅图像也存在结构迥异的图像模式，例如一幅图像中可能存在行人、宠物和车辆，一段视频中可能存在不同形态类别的片段轨迹等等。因此，用一个字典去表示其中丰富的图像内容，结果往往难以匹配，甚至可能丢失许多图像的细节信息^[93,94]。于是，为了使字典具有局部自适应性，能够充分匹配局部区域的模式，我们修改为局部字典。局部字典和全局字典是一组相对的概念，局部字典是由基于某种相似性准则在一个区域内找到与之相似的区块集合，而全局字典是由整幅图像或全体样本构成的集合。换言之，全局字典对应一个矩阵 D ，那么局部字典就是该矩阵的 D 分块形式。上述概念的对比很容易得出一个结论，就是局部字典期望通过建立多个小的与信号 y 相似的集合，来换取对信号更好的稀疏表示 x 。

4.2 轨迹划分

从如何设计一种既能充分表现视频监控场景中各种各样的局部模式，又能充分利用轨迹本身固有的局部自相似性的局部字典出发，本节在相当大的程度上充分发掘轨迹局部区域的结构特点和轨迹的局部相似性，由轨迹划分的局部相似集合构成局部字典，为接下来展开局部字典上的稀疏重构研究做好铺垫。

4.2.1 样本集的划分

在最小二乘三次 B 样条近似表示算法^[95]中, 由于 B 样条基函数和 B 样条曲线都具有非常优良的局部性质, 所以该表示算法求出的控制点也对轨迹形状具有局部支撑性。这意味着, 以与轨迹形状密切相关的控制点来截取轨迹片段, 可以良好地保证了轨迹的局部性, 还具有非常合理的物理意义。于是, 本节采用控制点来划分样本集合 D 和测试轨迹 T , 公式如下:

$$T = \begin{bmatrix} t_1 \\ \vdots \\ t_i \\ \vdots \\ t_{p-1} \end{bmatrix}, \quad i = 1, \dots, p-1, \quad (4-2)$$

$$D = \begin{bmatrix} d_1 \\ \vdots \\ d_i \\ \vdots \\ d_{p-1} \end{bmatrix}, \quad i = 1, \dots, p-1, \quad (4-3)$$

$$d_i = \{a_1^1(i), a_1^2(i), \dots, a_1^K(i), \dots, a_j^K(i)\}. \quad (4-4)$$

其中, p 是和本文 3.1 节定义相同的控制点的个数, t_i 表示第 i 段轨迹片段, d_i 表示第 i 段轨迹片段的集合, 其中划分后截取的每个轨迹片段集合 d_i 分别由 DTW 方法^[45]归一化对齐。

4.2.2 局部性分析

由于上述轨迹划分严格地与轨迹的局部约束相对应, 所以我们可以由划分后的样本集合 d_i 来线性地稀疏重构估计轨迹片段 t_i , 公式如下:

$$t_i \approx d_i \psi_i, \quad (4-5)$$

其中, ψ_i 表示局部字典的的线性叠加的系数, 优化求解公式如下:

$$\psi_i^* = \underset{\psi_i}{\operatorname{argmin}} \|t_i - d_i \psi_i\|^2 + \lambda \|\psi_i\|_1. \quad (4-6)$$

该问题的求解方法与本文 3.3 节相似, 因此不再多做赘述。公式 4-6 中, λ 表示

正则项，可以控制模型的稀疏度。由于每个轨迹片段 t_i 只与局部相似的样本集合 d_i 相关，所以该方法所涉及的稀疏模型是在局部字典上的稀疏重构，称之为局部稀疏重构。

图 4-3 显示了四种行为轨迹的控制点位置，可以看出红色的控制点簇往往集中在轨迹形状改变的位置，体现了控制点对轨迹的局部支撑性。图 4-4 中最右侧示例的同一类行为模式的轨迹，经过红色的控制点簇划分成左侧 6 段。可以看出，图示中每一段由若干个临近位置的轨迹点组成，每一簇内的各个轨迹段基本形状相似、位置临近。

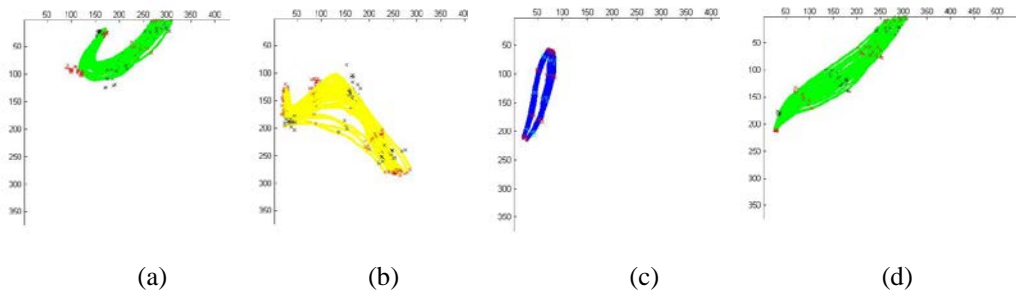


图 4-3 控制点的局部支撑示意图

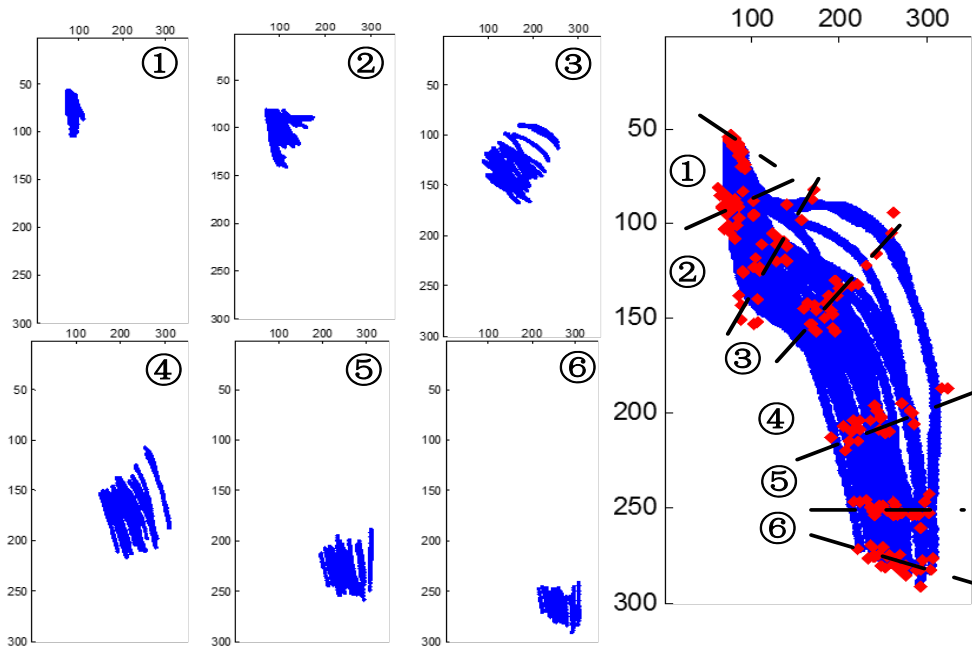


图 4-4 轨迹划分的局部约束示意图

4.3 判别编码矩阵

上一节中，轨迹划分已经利用局部约束，按照轨迹特征，计算基本比较单元与其邻域内的基本比较单元之间的稀疏重构误差，这可以理解为分段计算其局部相似度。接下来，需要在整个数据空间内设计一种多类分解框架，有效利用所有分段结果，实现整条轨迹的分类判别和检测模型。判别编码是一种多分类器融合技术，可以利用分类器的差异性来得出合适的分类结果，从而减少识别的错误率并提高分类器的鲁棒性。

4.3.1 判别编码的来源

纠错输出编码 (Error Correcting Output Codes, ECOC) 是一种常见的多类分解框架，它可以有效地把多类问题分解为多个二类问题。利用纠错输出编码解决多类分类问题，关键在于构建基于数据的编码矩阵。针对如何构建基于数据的编码矩阵问题，Alpaydin 和 Mayoraz^[96] 率先提出基于数据的反向传播编码确定方法。随后，Utschick 和 Weichselberger^[97] 利用期望最大化法则，通过对最大似然目标函数进行优化找出最适合样本空间的编码矩阵。2006 年，Pujol 和 Alpaydin 等^[98] 提出一种判别式编码方法，利用决策树结构逐步地构造类间间隔最大的子类，从而得出编码矩阵。2008 年，Escalera 和 Tax 等^[99] 针对样本集现行不可分问题，提出最基类子集再分割的编码方法。Escalera 还指出基于数据编码矩阵能最大可能获得数据类别子类划分，而类别子类划分往往为相关性较小且最易于分类的二类划分，因此基于此类划分构造的二分类器就能达到较高准确率，从而实现分类效果的整体提高。

ECOC 框架是用一种二元或三元的编码矩阵实现多类类别分解和基分类器集成的方法。常用的编码矩阵中，二数码用 $\{-1,+1\}$ 表示，三数码用 $\{-1,0,+1\}$ 表示，“-1”表示一类，“+1”表示另一类，“0”表示该码字对应类不参与此列的基分类器训练。图 4-5 展示了一个常见的 ECOC 框架，图中编码矩阵的每一行代表某一类 C_i 的码字，每一列代表样本的一种二类划分，码元 $\{-1,0,+1\}$ 分别以黑色、灰色和白色示意。在训练阶段，每个基分类器 f_i 的训练样本按编码矩阵的列划分后训练得出相应二类分类器。

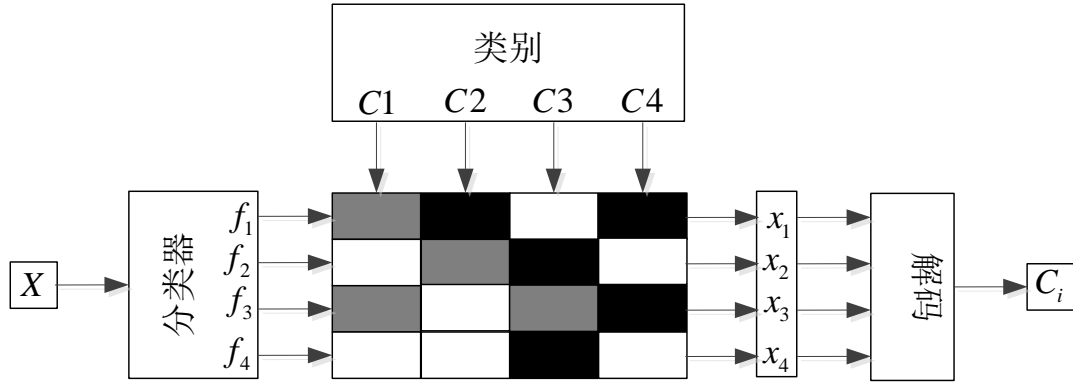


图 4-5 纠错输出编码框架示意图

4.3.2 判别编码矩阵的定义

表 4-1 一个判别编码矩阵的样例

编码 轨迹段号	类别					
	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>Class 4</i>	...	<i>Class N_c</i>
Tracklet 1	0	0	1	0	...	0
Tracklet 2	1	0	0	0	...	0
Tracklet 3	0	0	0	1	...	0
Tracklet 4	0	0	0	0	...	1
...
Tracklet N_p	0	0	0	1	...	0

借鉴上述思想，我们定义判别编码矩阵 (Discriminate Code Matrix, DCM) 用于轨迹的分段编码。基于上述轨迹划分和局部约束下稀疏重构的结果，计算每个轨迹片段计算重构误差：

$$\varepsilon_{ij}(t_i) = \|t_i - d_i \delta_j(\psi_i^*)\|_2, \quad j = 1, \dots, N_c, \quad (4-7)$$

然后，定义判别编码矩阵 $M \in \{0,1\}^{N_p \times N_c}$ ，元素 M_{ij} 为：

$$M_{ij} = \begin{cases} 0, & \text{当 } j \neq \underset{j}{\operatorname{argmin}}(\varepsilon_{ij}) \text{ 时} \\ 1, & \text{当 } j = \underset{j}{\operatorname{argmin}}(\varepsilon_{ij}) \text{ 时} \end{cases}, \quad j = 1, \dots, N_c, \quad (4-8)$$

其中， N_p 表示轨迹片段的个数， N_c 表示码字的类别个数。公式 4-7 中特征函数 δ_j 和上文 3.4 节定义相同，用于保证字典集中同一类的稀疏系数相同。表 4-1 所示是某条轨迹所求得的一个判别编码矩阵样例示意，可以看出，每行的分类中仅有一个非零，这意味着每个轨迹片段只能属于局部字典中的某一类行为模式。

4.4 损失加权解码和最小化能量函数解码

4.4.1 损失加权解码算法

表 4-2 基于损失加权解码的分类算法

算法 3 基于损失加权解码的分类算法

输入： 编码矩阵 M 。

输出： 测试轨迹 T 的分类标签 j^* 。

1. 初始化：权值 W ，其中 $w_{ij} = \text{MAX}$ ，
 2. for $j=1$ to N_c do
 3. for $i=1$ to N_p do
 4. if $M_{ij} \neq 0$ and $\sum_{j=1}^{N_c} M_{ij} \neq 0$ then
 5. 计算权值 $w_{ij} \leftarrow M_{ij} / (\sum_{i=1}^{N_p} M_{ij})$ 。
 6. end if
 7. end for
 8. 计算损失误差 $e_j \leftarrow \sum_{i=1}^{N_p} w_{ij} L(\varepsilon_{ij})$ 。
 9. end for
 10. 返回 $j^* \leftarrow \text{argmin}(e_j)$ 。
-

为了对轨迹 T 实现分类判别和检测模型，我们提出了基于损失加权解码 (Loss Weighted Decode, LWD) 的分类算法，算法流程详见表 4-2。损失加权

解码算法是将判别编码矩阵中各个轨迹字段的类别信息综合归一化为权值，然后定义加权后的重构误差为解码函数中的损失，当该损失值取最小时给出整条轨迹综合后的类别信息，即返回轨迹分类的标签值。该算法中，第 5 步中 W 表示权值矩阵，第 8 步中损失函数 $L(\varepsilon_{ij}) = \varepsilon_{ij}$ 直接定义为重构误差，第 10 步会由最小化损失误差 e_j 来返回该测试轨迹的标签 j^* 。

4.4.2 最小化能量函数解码

考虑到一段轨迹的在监控场景中的出口和入口信息对轨迹分类影响很大时，我们在损失加权解码中分别引入轨迹起始点和终止点的距离约束，提出了基于最小化能量函数解码（Energy Minimization Decode, EMD）的分类算法，该算法流程详见表 4-3。EMD 算法与 LWD 算法相似，都是以解码函数作为优化目标，区别在于 EMD 算法在 LWD 算法所定义的损失上，还加入了轨迹片段的两个端点距离约束（见算法步骤 3），归一化后的端点距离约束和加权后的重构损失一起作为解码函数中的优化目标。在算法第 10 步中，能量函数的第一项与 LWD 算法的加权损失相同，第二项是与轨迹片段的端点距离相关的函数，可以看出当轨迹端点距离越大时，轨迹相似信息越小。换言之，分类函数优化目标是不仅要使轨迹片段的重构误差损失越小，还要保证该段轨迹与稀疏重构的局部轨迹集的不相似度尽可能的小。与算法 3 类似，本算法最终返回测试轨迹 T 的标签 j^* 。

此外，我们还设计了分段的局部约束下的异常轨迹检测模型。与 3.4 节中检测模型相似，首先，计算每个轨迹片段 t_i 的重构误差比：

$$H_i = \min_j \frac{1/\varepsilon_{ij}}{\sum_{j=1}^{N_c} (1/\varepsilon_{ij})}, \quad (4-9)$$

接着，轨迹 T 的异常检测模型为：

$$F = \begin{cases} 0, & \text{当 } \sum_{i=1}^{N_p} \text{sign}(H_i - \theta)H_i < 0 \\ 1, & \text{其他} \end{cases}, \quad (4-10)$$

其中， θ 表示判断阈值，实验中经验阈值 $\theta=0.03$ 。公式 4-10 中，当且仅当所有的重构比值都大于阈值 θ 时，目标轨迹才被判断不属于异常行为。当目标轨迹被判断为不属于异常行为时，我们还可以使用算法 3 和算法 4 为目标轨迹 F 作

分类。

表 4-3 基于最小化能量函数解码的分类算法

算法 4 基于最小化能量函数解码的分类算法

输入： 编码矩阵 M 。

输出： 测试轨迹 T 的分类标号 j^* 。

1. 初始化权值 W ，其中 $w_{ij} = \text{MAX}$ 和 $w_i = 1/N_p$ ，
 2. 计算相似 $S_i \leftarrow \exp(-D_i^2 / \sigma^2)$ ，
 3. **for** $i=1$ to N_p **do**
 4. **for** $i=1$ to N_p **do**
 5. **if** $M_{ij} \neq 0$ and $\sum_{j=1}^{N_p} M_{ij} \neq 0$ **then**
 6. $w_{ij} \leftarrow M_{ij} / (\sum_{j=1}^{N_p} M_{ij})$ 。
 7. **end if**
 8. **end for**
 9. $S_i \leftarrow \text{normalize}_{(0,1)}(S_i)$ 。
 10. $e_j \leftarrow \sum_{i=1}^{N_p} (w_{ij}L(\varepsilon_{ij}) + w_i(1 - S_i))$ 。
 11. **end for**
 12. 返回 $j^* \leftarrow \text{argmin}(e_j)$ 。
-

4.5 实验结果分析

4.5.1 本章实验的数据描述

在这一节，我们为了分析基于局部约束下稀疏重构模型的算法性能，采用室内大厅 CAVIAR 和停车场 Carpark 两个数据集加以验证。CAVIAR 和 Carpark

数据集中分别包括了一些来自室内大厅的人和停车场的车的不完整、有噪声和局部有变化等轨迹。由于我们使用的轨迹数据是视频跟踪的标注结果，为了更好地验证该算法对不完整轨迹的有效性，我们仿真模拟了视频跟踪的轨迹，形成一个约有 10% 缺失的轨迹数据集合，如表 4-4。

表 4-4 基于局部约束下稀疏重构的目标轨迹分析的实验数据集列表

数据集		介绍
CAVIAR	Annotated	室内大厅的正常/异常行为的标注轨迹
	Incomplete	室内大厅的正常/异常行为的不完整轨迹
Carpark	Annotated	停车场的正常/异常行为的标注轨迹
	Incomplete	停车场的正常/异常行为的不完整轨迹

4.5.2 基于局部约束下稀疏重构模型的轨迹分析实验

我们在室内大厅 CAVIAR 和停车场 Carpark 的数据集上分别验证了基于局部约束下稀疏重构的轨迹分类和异常检测方法。



图 4-6 CAVIAR 和 Carpark 两个场景中待分类的轨迹示例

在轨迹分类的验证实验中，我们沿用第 3.5 节中的 LCSCA 表示方法，首先提取轨迹特征，将数据集中的所有轨迹序列表示为最小二乘的三次 B 样条近似的控制点序列，采用正确分类率 CCR 作为算法分类性能评价指标。在每个监控场景的分类性能的对比实验中，各采用了三种方法和两个数据集。所采用的三种方法分别是基于稀疏重构分析 (Sparse Reconstruction Analysis, SRA) 的轨迹分类方法、局部约束下的稀疏重构 (Locality-constrained SRA, LSR) 采用损失加权解码 (Loss Weighted Decode, LWD) 的方法、局部约束下的稀疏重构

(Locality-constrained SRA, LSR) 采用最小化能量函数解码 (Energy Minimization Decode, EMD) 的方法。CAVIAR 场景下的两个数据集是 CAVIAR Lobby Scenario 的 annotated 数据集和 incomplete 数据集, 同样地, Carpark 场景实验也用到了两个数据集是 Carpark Scenario 的 annotated 数据集和 incomplete 数据集。图 4-6 是 CAVIAR 和 Carpark 两个场景中待分类的轨迹示例, 图 4-7 和图 4-8 分别是两个场景中不同类别的一条轨迹示例。



图 4-7 CAVIAR 数据集来自不同类别各一条轨迹的示例

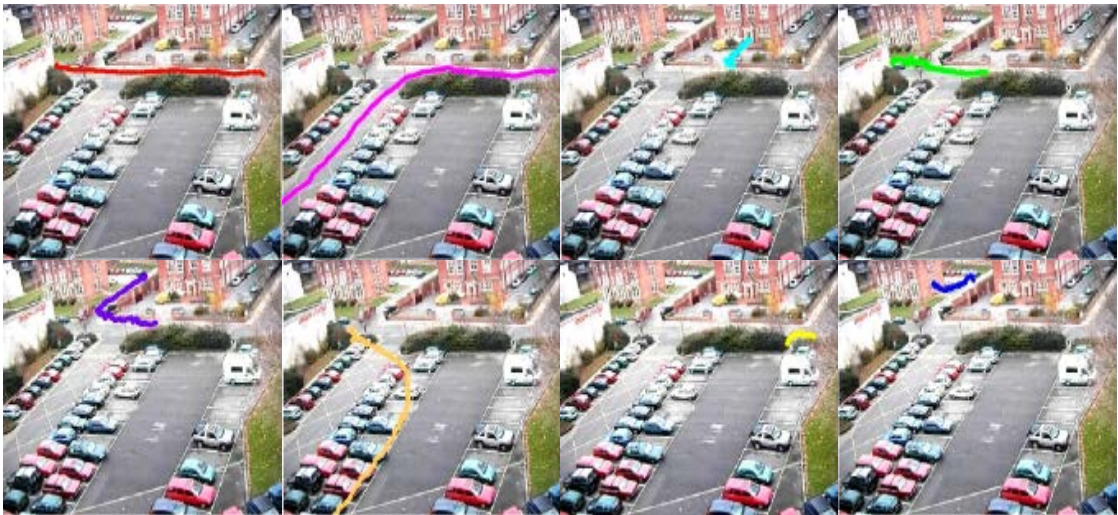


图 4-8 Carpark 数据集中不同类别的各一条轨迹的示例

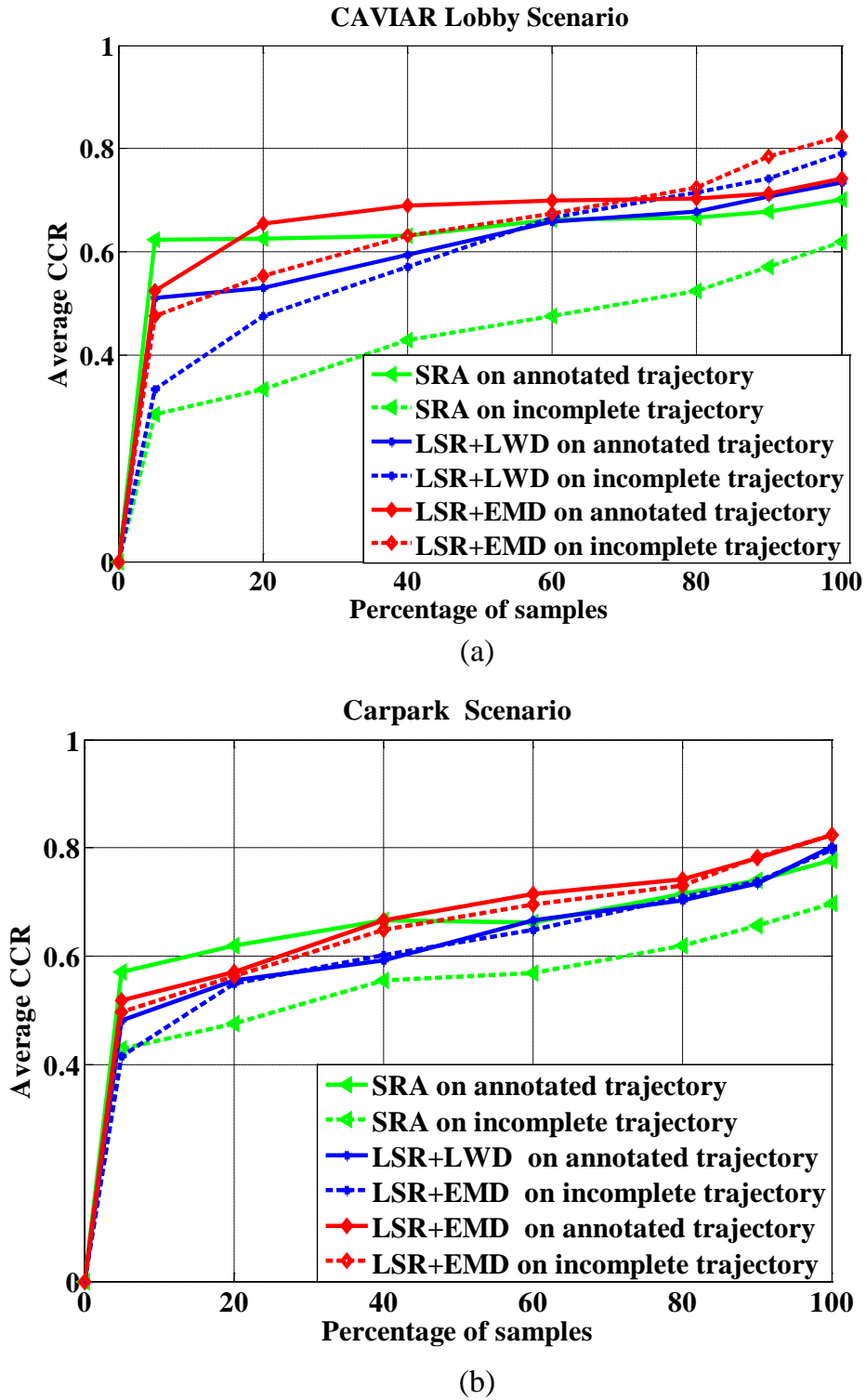


图 4-9 局部约束下稀疏重构模型的轨迹分类性能对比图

图 4-9 给出了两个监控场景在三种方法和两个数据集上的分类性能的结果对比,展示了正确分类率 CCR 随着重构样本集的样本数 K 增加的变化曲线对比。可以看出,当训练样本数达到 $K=10$ 以后,局部约束下的稀疏重构方法 LSR+LWD (蓝色曲线) 和 LSR+EMD (红色曲线) 的分类结果明显优于稀疏重构方法 (绿色曲线)。尤其是,图 4-9(a)和图 4-9(b)中给出的 LSR+EMD (红色曲线)方法性能无论在 annotated 数据集上的结果(红色实线),还是在 incomplete 数据集上的结果(红色虚线),都比其他两种方法表现更优。图 4-9(a)中,CAVIAR 的不完整轨迹 incomplete 数据集上 LSR+EMD 方法 (红色虚线) 的分类准确度达到 $CCR=72.77\%(\pm 9.62\%)$, LSR+LWD 方法 (蓝色虚线) 可以达到 $CCR=69.70\%(\pm 9.29\%)$ 。图 4-9(b)中,CAVIAR 的不完整轨迹 incomplete 数据集上 LSR+EMD 方法 (红色虚线) 的分类准确度达到 $CCR=67.60\%(\pm 7.41\%)$, LSR+LWD 方法 (蓝色虚线) 可以达到 $CCR=63.70\%(\pm 6.65\%)$ 。经过对比分析,所提出算法的优良性能是来源于稀疏重构模型中引入的局部约束。这是因为一条轨迹中关键的形状结构信息往往集中在一小段轨迹上,而能决定这条轨迹分类的恰恰就是这些关键片段。因此,这个实验可以证明,局部约束下稀疏重构模型做轨迹分类时能够有效利用轨迹片段的局部信息,保证稀疏重构系数的有效稀疏,从而得到良好的分类性能。此外,图 4-10 给出了待分类轨迹与 LSR 方法中所选取的稀疏重构样本集对比示意,图 4-10(a)来自于 CAVIAR 数据集,图 4-10(b)来自于 Carpark 数据集。

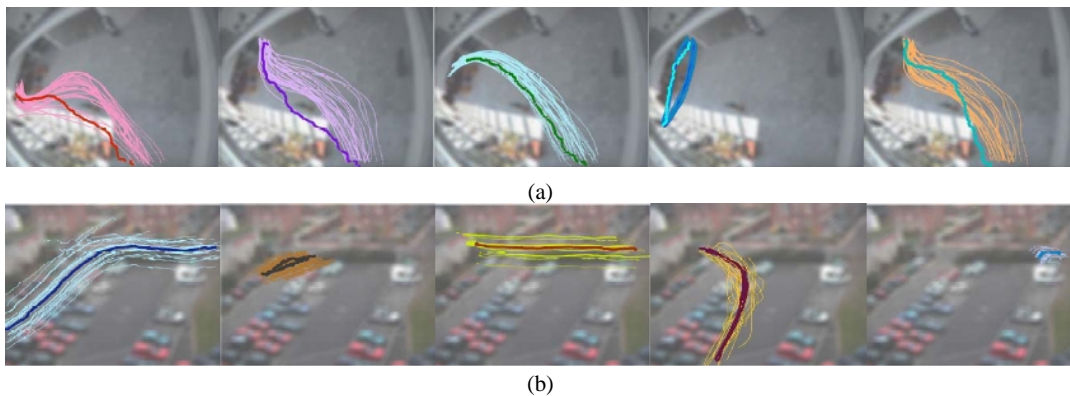


图 4-10 CAVIAR 和 Carpark 数据集中待分类轨迹与其重构样本示意图

异常轨迹检测中,我们还将基于局部约束下稀疏重构 (LSR) 模型的性能检测精确度 DACC 与两种方法做了对比分析。这两种方法分别是第 2.3.3

节基于高斯混合模型(Gauss Mixture Models, GMMs)的异常检测方法和第 3.4.2 节基于稀疏重构(Sparse Reconstruction Analysis, SRA)的异常检测方法。CAVIAR 和 Carpark 监控场景中分别在两个数据集(annotated 数据集和 incomplete 数据集)上的检测精确度结果如表 4-5, 该表格列出了训练样本数达

表 4-5 局部约束下稀疏重构模型的异常轨迹检测性能对比

数据集		方法	DACC (%)
CAVIAR	Annotated	GMMs	84.75
		SRA	87.42
		LSR	93.21
	Incomplete	GMMs	74.00
		SRA	83.72
		LSR	89.50
Carpark	Annotated	GMMs	84.60
		SRA	87.32
		LSR	89.75
	Incomplete	GMMs	79.41
		SRA	85.29
		LSR	88.24

到 $K=100$ 以后的检测结果。其中, CAVIAR 的 annotated 数据集上 LSR 的检测精确度高达 $DACC=93.11\%$, 同时在 incomplete 数据集检测性能也很好, 可以达到 $DACC=89.50\%$ 。Carpark 数据集上性能也明显高于其他两种方法, 分别在 annotated 数据集达到 $DACC=89.75\%$, 在 incomplete 数据集上达到 $DACC=88.24\%$ 。

4.6 本章小结

本章提出基于局部约束下稀疏重构(Locality-constrained Sparse Reconstruction, LSR)的目标轨迹分析算法。在监控场景的样本空间中, 以 LCSCA 表示的轨迹控制点做样本集划分, 得到的局部字典既能充分利用轨迹本身固有的局部相似性, 又可以表现监控场景中样式繁多的局部模式。本章设计了每段样本在其相似局部样本子集稀疏重构的模型, 并依此构建一个加权的局部判别式字典, 最后根据所定义的判别编码和损失加权解码算法, 实现整个样本空间的轨迹分类和异常检测。本章给出了在不同程度的不完整轨迹情况下轨

迹分类和异常检测的鲁棒性验证，将该算法与其他现有算法进行了结果比较和定量分析。实验结果表明，在整体性能上，本章提出的基于局部约束下稀疏重构的目标轨迹分析方法鲁棒有效。

第五章 基于聚类的局部稀疏重构的目标轨迹分析

运动目标轨迹分析大多是采用监督学习的算法，一般需要获取更多的标记样本，这样做成本通常很高，而且实际中要想找到直接有利于分类的数据比较困难。针对这一问题，本章提出了一种新的基于聚类的局部稀疏重构（Cluster-based Locality-constrained Sparse Reconstruction, CLSR）目标轨迹分析方法，如图 5-1 所示。

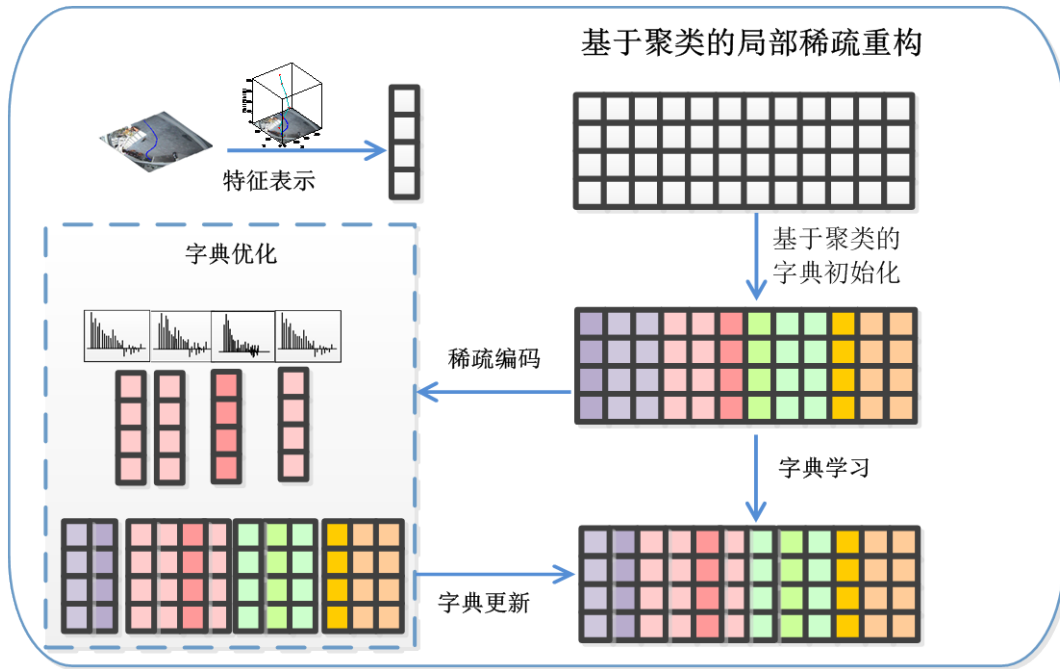


图 5-1 基于聚类的局部稀疏重构的目标轨迹分析流程图

CLSR 方法构建了一个增量的字典学习框架：首先，通过 K-Means 算法得到待学习轨迹 K 类的聚类结果。然后，将聚类结果中每类的一部分作为训练样本，利用稀疏表示的系数更新选择进行增量学习的训练样本集，迭代训练得到 K 个字典。最后，再对每条轨迹作局部字典的稀疏重构，利用稀疏重构误差做轨迹分类。该方法的非监督增量学习的机制融合了同类轨迹的标签一致选择与分类器优化的过程，主要解决了现有技术在实际应用中训练样本的类别标注非常难以获取的问题，相比现有非监督方法显著地提高了轨迹学习性能，可以用于

目标轨迹分析。

在本章，聚类初始化阶段使用 K-Means 算法，字典学习阶段使用标签一致的 K-奇异值分解 (Label Consistent K-Singular Value Decomposition, LC-KSVD) 算法，局部稀疏重构阶段使用最小化能量函数解码算法。最后，我们还设计出增量的字典学习算法，用于在线学习轨迹字典。

5.1 基于聚类的局部稀疏重构的思路

前文中提出的轨迹分析模型都属于有监督的学习方法，总是需要获取大量的标记样本，但在实际应用中，轨迹的训练样本的类别标注却非常困难，往往缺少已学习模型中需要的类别先验。运动目标轨迹分析大多是采用监督学习的算法，然而实际中要想找到这种有利于分类的比较合理的表示方法的数据是较为困难的，即往往缺少已学习模型中的类别先验。因此，我们试图提出一种基于聚类的轨迹学习方法。由于在上一章节我们采用局部模型的思想，所提出的局部约束下的样本子集稀疏重构模型、分段编码和损失加权解码算法经过验证初步有效。因而，本章我们将继续沿用局部稀疏重构的思路，形成基于聚类的局部稀疏重构目标轨迹分析。

然而，正如本文第 4.1.2 节中所讲，稀疏重构与字典学习比较，虽然两者都有稀疏编码，但是稀疏重构缺少字典更新的过程。因此，这里引入字典学习的理论，求解出更能够通过其中少量原子的线性组合表示训练集中的各个信号的字典基向量，从而使稀疏重构时的系数稀疏度更高，这将意味着能够更好地恢复出原始信号^[100]。总之，基于聚类的局部稀疏重构的方法旨在利用结构聚类法来解决轨迹样本字典的类别先验问题，同时，使得字典中的数据在类别数目尽可能少的前提下，保证数据类内相似度高，类间差别很大，从而提高分类准确度。

此外，目前基于监督学习的目标轨迹分析已经在众多的研究中获得发展，但监督学习需要人工标定数据样本的缺点也是其结合实际应用发展的最大阻碍。在非监督的基础下进行监控视频中的运动目标轨迹分析，探寻减少掉庞大的人工干预代价的方法，具有科学意义。非监督学习在减少标注时间与代价，提高机器学习性能、提高算法适应能力方面具有重要的实用价值。

5.2 基于 K-Means 的轨迹聚类

5.2.1 K-Means 聚类

K-Means 聚类是数据挖掘领域中的一种非监督聚类分析算法，主要目的是通过不断地取距离种子点最近的均值，将数据按照一定的相似属性分成不同的子类^[43,101]。

假设训练样本为 $\{y_1, y_2, \dots, y_N\}$ ，算法流程如下：

(1) 随机选取 K 个聚类初始中心点 $\mu_1, \mu_2, \dots, \mu_K$ ；

(2) 计算待聚类样例与每类中心点的距离值，将待聚类样例与各个中心点距离比较，取其距离最近的类：

$$c_i = \arg \min_j \|y_i - \mu_j\|^2, \quad (5-1)$$

(3) 计算更新该类中心点：

$$\mu_j = \frac{\sum_{i=1}^n y_i | c_i = j}{\sum_{i=1}^n 1 | c_i = j}, \quad (5-2)$$

(4) 根据步骤 2 和步骤 3 将 n 个样例聚类；

(5) 重复步骤 2、3、4，直至达到收敛条件，即每类中心点不再变化或变化很小。按照每个样例 x_i 与其类中心点距离平方和，收敛条件可以定义为：

$$J(c, \mu) = \sum_{i=1}^n \|y_i - \mu_{c_i}\|^2, \quad (5-3)$$

而 K-Means 聚类算法的目的就是调整收敛条件函数值直到最小。经过聚类后的样本空间每类中心点相对较远，同类样本距离中心点相对集中。

5.2.2 学习样本的特征选择

对于视频监控中运动目标轨迹分析研究，学习样本是运动目标的轨迹。基于聚类的局部稀疏重构方法首要工作就是学习出聚类的样本集。构建样本字典时采用第四章的局部稀疏重构思路，利用局部样本子集来重构测试轨迹，采用分段编码和基于最小化能量函数的解码算法。这种轨迹分段的方法能够降低了样本重构集合的特征维数，使学习样本个数远大于特征维数，保证了字典学习

中对样本数远大于特征维数的要求^[119,120]。这里与第四章的稀疏重构方法不同之处在于所用到的字典是经过学习更新得到的优化结果，下一节将详细介绍字典学习的模型和求解过程。

此外，考虑到下一步字典学习方法对样本同构性的要求，该方法需要对学习样本加以特征选择。这里依然考虑轨迹的形状结构信息，我们并不是直接计算轨迹样本间的轨迹点相似距离，而是计算与之对应的 LCSCA 表示控制点的相似距离，此处的距离采用 DTW 距离（具体介绍详见第 2.3.3 节）。

5.3 增量字典学习模型

5.3.1 字典初始化

样本集 $\{y_1, y_2, \dots, y_N\}$ 中的各个样本 y_i 聚类后得到 $y_i^{(j)}$ 的集合，其中 j 是聚类中心 c_j 的标签。于是，经过 K-Means 聚类，样本集变成了 $\{y_1^{(1)}, \dots, y_M^{(l)}, \dots, y_N^{(K)}\}$ 。

字典学习问题也就变成了从样本集 $Y \in \mathbb{R}^{p \times N}$ 中学习字典 $D \in \mathbb{R}^{p \times M}$ ，求解字典的优化模型为：

$$\langle D, X \rangle = \underset{D, X}{\operatorname{argmin}} \|Y - DX\|_2^2, \quad \text{s.t. } \forall i, \|x_i\|_0 \leq \varepsilon, \quad (5-4)$$

其中， $\|Y - DX\|_2^2$ 表示重构误差， $D = [d_1, \dots, d_M] \in \mathbb{R}^{p \times M}$ 表示字典基向量， $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$ 表示稀疏系数， $\|x_i\|_0 \leq \varepsilon$ 表示模型的稀疏度约束。该模型可以通过迭代算法求解，首先用 N 个聚类后训练样本来初始化字典 $D^{(0)}$ ，结果如下

$$D^{(0)} = [C^{(1)}, \dots, C^{(k)}, \dots, C^{(K)}], \quad C^{(k)} = [y_1^{(1)}, \dots, y_{N_k}^{(k)}], \quad (5-5)$$

其中， $C^{(k)}$ 是第 5.2.1 节的聚类结果，满足 K 个 N_k 的总和为 N 。将 $D^{(0)}$ 代入公式 5-4 采用 OMP 算法^[83]计算稀疏系数为：

$$x_i = x^*(y_i, D) = \underset{x}{\operatorname{argmin}} \|y_i - Dx\|_2^2, \quad \text{s.t. } \|x\|_0 \leq \varepsilon. \quad (5-6)$$

5.3.2 字典学习

稀疏系数 x_i 可以直接用于轨迹特征向量 y_i 的分类器 $f(x_i)$ 的输入, 构建该多类分类器的模型如下:

$$W = \operatorname{argmin}_W \sum_{i=1}^N L\{h_i, f(x_i, W)\} + \lambda_1 \|W\|_2^2, \quad (5-7)$$

其中, $W \in \mathbb{R}^{K \times M}$ 表示模型参数, K 表示类别个数, $L(\cdot)$ 是分类损失函数, h_i 是 y_i 的标签, λ_1 是正则项的参数, 用于防止模型过拟合。

按照文献[102-104]的证明, 可以将字典和上述分类模型联合学习优化。于是, 得到字典和分类器的联合优化的目标函数, 如下:

$$\langle D, W, X \rangle = \operatorname{argmin}_{D, W, X} \|Y - DX\|_2^2 + \sum_{i=1}^N L\{h_i, f(x_i, W)\} + \lambda_1 \|W\|_2^2, \quad s.t. \forall i, \|x_i\|_0 \leq \varepsilon. \quad (5-8)$$

5.3.2.1 过完备字典训练方法

过完备字典可以是一种预先选取的基底, 也可以从训练样本训练所得。通常, 为了学习更适用于样本集中信号的稀疏表示, 研究者会选择通过训练得到过完备字典。各种获取字典的方法通常都可以分为两个步骤: (1) 稀疏编码阶段, 即给定字典, 求解训练集在该字典上的稀疏表示系数; (2) 字典更新阶段, 即假设稀疏表示系数已知且固定, 更新字典。因此, 各种字典学习算法都集中于解决这两个问题: 如何求解稀疏表示稀疏和如何更新字典。现有非常具有代表性的方法有: 最大似然法^[105]、最优方向法^[106]、最大后验概率^[107]、正交字典级联^[108]、K-SVD^[109]等。

5.3.2.2 K-SVD 方法

Elad 提出的 K-SVD 算法^[107]在字典学习领域得到了广泛关注, 该算法是目前应用最多的字典学习方法之一。具体地, K-SVD 求解问题如下:

$$\min_{D, X} \left\{ \|Y - DX\|_F^2 \right\}, \quad s.t. \forall i, \|x_i\| \leq T_0, \quad (5-9)$$

在稀疏编码阶段, 假设字典 D 固定, 因为:

$$\|Y - DX\|_F^2 = \sum_{i=1}^N \|y_i - Dx_i\|_2^2, \quad (5-10)$$

公式 5-9 可以分解为 D 个独立子问题:

$$\min_{x_i} \{\|y_i - Dx_i\|\} \quad s.t. \|x_i\|_0 \leq T_0, \quad i = 1, 2, \dots, N, \quad (5-11)$$

这些子问题可以通过任意一种追踪算法求解。

在字典更新阶段，K-SVD 算法依次更新字典中的各个原子，然后更新该原子对应的稀疏表示系数。如果用 x_T^k 表示系数矩阵 X 的第 k 行，则公式 5-10 可以写成形式：

$$\begin{aligned} \|Y - DX\|_F^2 &= \left\| Y - \sum_{i=1}^K d_i x_T^i \right\|_F^2 \\ &= \left\| \left(Y - \sum_{i \neq k} d_i x_T^i \right) - d_k x_T^k \right\|_F^2, \\ &= \|E_k - d_k x_T^k\|_F^2 \end{aligned} \quad (5-12)$$

这里将 DX 分解成 K 个秩为 1 的矩阵的和。其中， E_k 为去掉第 k 个原子后的表示残差，找出所有使用了原子 d_k 的信号，计算信号在去掉 d_k 后的表示残差，然后对残差矩阵奇异值分解（Singular Value Decomposition, SVD），得到新原子和对应的表示系数，这样可以保证所有信号表示系数的稀疏度不会变大。

5.3.2.3 LC-KSVD 方法

一个线性分类器 $f(x_i)$ 的性能取决于输入的稀疏系数 x_i 的判别性。稀疏表示系数在任意字典下计算求得的稀疏重构误差具有很强的判别性，即来自同一类的样本具有相似的稀疏表示，只需要一个简单的线性分类器就能达到良好的多类分类器学习。

为了在字典 D 上得到具有判别性的稀疏系数，Jiang 等^[110]提出在公式 5-4 中加入标签一致正则项和分类误差项，字典学习模型修改为：

$$\langle D, W, A, X \rangle = \operatorname{argmin}_{D, W, A, X} \|Y - DX\|_2^2 + \alpha \|Q - AX\|_2^2 + \beta \|H - WX\|_2^2, \quad s.t. \forall i, \|x_i\|_0 \leq \varepsilon. \quad (5-13)$$

其中， α 和 β 是正则化参数， $Q = [q_1, q_2, \dots, q_N] \in \mathbb{R}^{M \times N}$ 是输入样本 Y 的判别式稀疏系数矩阵， A 是线性变换矩阵， $\|Q - AX\|_2^2$ 表示判别式稀疏编码误差，它使分类器更具有判别性。 H 是输入样本的类别矩阵， W 是线性分类器参数， $\|H - WX\|_2^2$ 表示分类误差，它使分类器更利于分类。举个例子，如果给出字典 $D = [d_1, \dots, d_4]$ 和样本 $Y = [y_1, \dots, y_4]$ ，当 y_1 、 y_2 、 d_1 和 d_2 的标签为 1，而 y_3 、 y_4 、

d_3 和 d_4 的标签为 2 时, 则 Q 定义为:

$$Q = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \quad (5-14)$$

5.3.3 字典优化

5.3.3.1 字典求解模型

下面介绍字典模型的最优化求解方法。基于 $A \in \mathbb{R}^{M \times M}$ 是可逆矩阵的假设, 公式 5-13 的求解模型^[108]为:

$$\langle D, W, A, X \rangle = \operatorname{argmin}_{D, W, A, X} \left\| \begin{pmatrix} Y \\ \sqrt{\alpha} Q \end{pmatrix} - \begin{pmatrix} D \\ \sqrt{\alpha} A \end{pmatrix} X \right\|_2, \quad s.t. \forall i, \|x_i\|_0 \leq \varepsilon, \quad (5-15)$$

在新的求解模型中, 令 $Y_{new} = (Y^T \quad \sqrt{\alpha} Q^T \quad \sqrt{\beta} H^T)^T$ 和 $D_{new} = (D^T \quad \sqrt{\alpha} A^T \quad \sqrt{\beta} W^T)^T$, 其中 D_{new} 的列示归一化的, 则上式等价于:

$$\langle D, W, A, X \rangle = \operatorname{argmin}_{D, W, A, X} \|Y_{new} - D_{new} X\|_2^2, \quad s.t. \forall i, \|x_i\|_0 \leq \varepsilon. \quad (5-16)$$

这就可以通过第 5.3.2.2 节的 K-SVD 算法求得 D_{new} 和 X , 从而同时得出 D 、 W 和 A , 解决了局部最优的问题, 适用于字典小、类别多的情况。

5.3.3.2 基于聚类的字典学习算法

针对上述字典模型, 我们用第 5.3.1 节、第 5.3.2.3 节的定义分别初始化 $D^{(0)}$ 、 $Q^{(0)}$ 和 $H^{(0)}$, 而 $A^{(0)}$ 和 $W^{(0)}$ 由多类岭回归方法^[111]初始化如下:

$$A^{(0)} = Q^{(0)} X^T (X X^T + \lambda_1 I)^{-1}, \quad (5-17)$$

$$W^{(0)} = H^{(0)} X^T (X X^T + \lambda_1 I)^{-1}, \quad (5-18)$$

稀疏系数 X 由公式 5-4 求得。在此基础上, 我们提出基于聚类的字典学习算法如表 5-1。最后, 对求得的 D_{new} , 根据公式 5-19 求得 \hat{D} 、 \hat{A} 和 \hat{W} , 使其满足 L_2

范数的归一化条件 $\forall k, \left\| \left(d_k^T, \sqrt{\alpha} a_k^T, \sqrt{\beta} w_k^T \right)^T x_i \right\|_2 = 1$ 。

$$\hat{D} = \left[\frac{d_1}{\|d_1\|_2}, \dots, \frac{d_K}{\|d_K\|_2} \right], \hat{A} = \left[\frac{a_1}{\|d_1\|_2}, \dots, \frac{a_K}{\|d_K\|_2} \right], \hat{W} = \left[\frac{w_1}{\|d_1\|_2}, \dots, \frac{w_K}{\|d_K\|_2} \right]. \quad (5-19)$$

表 5-1 基于聚类的字典学习算法

算法 5. 基于聚类的字典的学习算法

1. **输入:** $Y, \alpha, \beta, \varepsilon$ 。
2. **输出:** D, A, W 。
3. 计算 $D^{(0)}, Q^{(0)}, H^{(0)}, A^{(0)}, W^{(0)}$,
4. 初始化 $D^{(0)}$,
5. 根据第 5.2.1 节的聚类结果, 定义 $Q^{(0)}$ 和 $H^{(0)}$,
6. 将 Y 和 $D^{(0)}$ 代入公式 5-4 求出 $X^{(0)}$,
7. 由公式 5-17 和公式 5-18 计算 $A^{(0)}, W^{(0)}$ 。
8. 计算 $Y_{new} = \begin{pmatrix} Y \\ \sqrt{\alpha} Q^{(0)} \\ \sqrt{\beta} H^{(0)} \end{pmatrix}$ and $D_{new} = \begin{pmatrix} D \\ \sqrt{\alpha} A^{(0)} \\ \sqrt{\beta} W^{(0)} \end{pmatrix}$,
9. 由公式 5-16 用 K-SVD 算法更新 D_{new} ,
10. 对 Y 计算稀疏系数 X ,
11. 根据稀疏表示系数的聚类结果更新 Q 和 H ,
12. 最后由 D_{new} 代入公式 5-19 计算出 $\hat{D}, \hat{A}, \hat{W}$ 。

5.3.3.3 增量的字典学习算法

在此基础上, 当训练样本集分批进入学习系统时, 就需要把训练集划分成几个独立的子集, 依次在各个子集上做增量学习。本文将基于聚类的字典放入增量学习的框架^[108], 提出了增量的字典学习算法如表 5-2。这种增量学习的技术旨在将旧样本集上学习得到的字典基加入到新增样本中进行训练学习, 从而得出新的分类模型。该方法可以充分利用历史学习的结果, 从而显著地节省后续训练的时间。

表 5-2 增量的字典学习算法

算法 6. 增量的字典学习算法

1. **输入:** $Y, Q^{(0)}, H^{(0)}, D^{(0)}, A^{(0)}, W^{(0)}, \mu, \gamma, \rho, \nu_1, \nu_2, \hat{T}$ 。
2. **输出:** D, A, W 。
3. **for** $t=1, \dots, \hat{T}$ **do**
4. 交换排列训练样本集合 Y ,
5. **for** $i=1, \dots, N$ **do**
6. y_i 代入公式 5-4 计算 x_i ,
7. 寻找稀疏系数 x_i 中非零项所对应的样本组成有效集合 Λ_i ,
8. 计算隐微分函数中所需要的辅助变量 ϕ_1 和 ϕ_2 ,
9. 选择学习率 $\rho_t = \min(\rho, \rho_{i_0}/i)$,
10. 由梯度公式 5-21、公式 5-22 和公式 5-23 更新 D 、 A 、 W ,
11. D 、 A 、 W 代入公式 5-4 得出稀疏系数, 按系数聚类后同时更新 Q 和 H ,
12. 计算 $D^{(t)} = D^{(t)} - \rho_t \frac{\partial L^i}{\partial D^{(t)}}$, $A^{(t)} = A^{(t)} - \rho_t \frac{\partial L^i}{\partial A^{(t)}}$, $W^{(t)} = W^{(t)} - \rho_t \frac{\partial L^i}{\partial W^{(t)}}$,
13. 将 $D^{(t)}$ 的每一列投影到一个单位圆, 即公式 5-19 的归一化处理。
14. **end for**
15. 更新 $D^{(t+1)} = D^{(t)}$, $A^{(t+1)} = A^{(t)}$, $W^{(t+1)} = W^{(t)}$ 。
16. **end for**

按照增量学习理论, 我们首先读入一小部分训练样本, 然后增量地更新模型。首先, 定义 D 、 A 、 W 在增量学习框架中的目标函数如下:

$$\begin{aligned} \min_{D, W, A} \sum_{i=1}^N L^i \{D, y_i, W, h_i, A, q_i\} + \frac{\nu_1}{2} \|W\|_2^2 + \frac{\nu_2}{2} \|A\|_2^2 \\ \text{s.t. } x_i = \arg \min_x \|y_i - Dx\|_2^2 + \gamma \|x\|_1, \quad \|d_j\|_2^2 \leq 1 \end{aligned} \quad (5-20)$$

其中, $L^i \{D, y_i, W, h_i, A, q_i\} = \mu \|q_i - Ax_i\|_2^2 + (1-\mu) \|h_i - Wx_i\|_2^2$, 分别有 $i \in \{1, \dots, N\}$ 和 $j \in \{1, \dots, M\}$ 。然后, 令 $L_1^i \{D, y_i, A, q_i\} = \|q_i - Ax_i\|_2^2$ 和 $L_2^i \{D, y_i, W, h_i\} = \|h_i - Wx_i\|_2^2$, 根据随机梯度下降算法^[112]有函数 L^i 对字典 D 的梯度为:

$$\frac{\partial L^i}{\partial D} = \mu \frac{\partial L_1^i}{\partial D} + (1 - \mu) \frac{\partial L_2^i}{\partial D}. \quad (5-21)$$

同理，函数 L^i 对 A 和 W 梯度为：

$$\frac{\partial L^i}{\partial W} = (1 - \mu)(Wx_i - h_i)x_i^T + v_1 W, \quad (5-22)$$

$$\frac{\partial L^i}{\partial A} = \mu(Ax_i - q_i)x_i^T + v_2 A. \quad (5-23)$$

最后，我们利用梯度下降方法增量地更新字典模型，学习率设为 $\min(\rho, \rho i_0 / i)$ ，初始学习率 ρ 为 10， $i_0 = \hat{T}/10$ 的 i_0 与迭代次数 \hat{T} 相关。此外，一些正则化参数如 v_1 、 v_2 和 γ 在实验时会通过学习获得。至此，字典 \hat{D} 已经完成了学习过程，对于待分类的测试轨迹 y_i ，我们可以由字典学习模型的公式 5-4 计算其稀疏表示系数 \hat{x}_i ，然后采用一个线性的分类器：

$$l = \hat{W} \cdot \hat{x}_i, \quad (5-24)$$

得出轨迹分类结果。 l 是一个标签向量，该向量元素中最大值的索引就是轨迹样本 y_i 的分类标签。

5.4 实验结果分析

5.4.1 算法的分类性能验证

为了验证基于聚类的局部稀疏重构模型，本节对 CAVIAR 数据集的 22 类轨迹和 Carpark 数据集的 16 类轨迹进行实验。CAVIAR 数据集中，每个样本数据有不同长度的轨迹，最小长度达 150 维，每种类别分别包含有 100 个训练样本数据，实验采用 1100 个训练样本数据和 1121 个测试样本数据。Carpark 数据集中，每个样本数据有最小长度的轨迹维度是 60 维，每种类别分别包含有 15 个训练样本数据，实验采用 124 个训练样本数据和 145 个测试样本数据。

由于基于聚类的局部稀疏重构模型（Cluster-based Locality-constrained Sparse Reconstruction, CLSR）是从非监督的轨迹聚类思想出发，所以本节在轨迹分类的验证实验中选用的方法有：分层 K-Means 聚类（Hierarchical K-Means Clustering, HKMC）^[43]、并行谱聚类（Parallel Spectral Clustering, PSC）^[113]

和高斯混合模型（Gauss Mixture Models, GMMs）^[20]。此外，CAVIAR 数据集上，CLSR 方法字典学习的正则化参数分别为 $\alpha = 0.01$ 和 $\beta = 1$ ，CAVIAR 数据集上，CLSR 方法字典学习的正则化参数分别为 $\alpha = 1$ 和 $\beta = 1$ 。

表 5-3 列出了分层 K-Means 聚类、并行谱聚类、GMMs 和 CLSR 的轨迹分类在 CAVIAR 数据集上的实验结果。本节分别使用每种类别随机选用 $M = \{5, 15, 25, 35, 45, 50\}$ 个样本来做对比，从表格中可以看出，本文提出的基于聚类的局部稀疏重构（CLSR）的轨迹分类方法显然优于其他方法，正确分类率 CCR 比分层 K-Means 聚类方法平均高出 27.78%，比并行谱聚类方法平均高出 22.57%，比 GMMs 方法平均高出 23.45%。由此可见，CLSR 方法的非监督增量学习机制融合了同类轨迹的标签一致选择与分类器优化的过程，即使在小的训练样本集上也能保证同一类别的轨迹在字典上的重构系数非常相似。

表 5-3 CAVIAR 数据集上四种方法的分类性能对比

方法	分类正确率 (%)					
	每个类别训练样本个数					
	5	15	25	35	45	50
分层 K-Means 聚类	19.8	28.4	35.9	36.7	38.9	45.9
并行谱聚类	21.9	23.6	37.1	38.6	47.4	68.3
GMMs	-	-	-	-	-	38.6
CLSR	38.7	58.4	65.5	67.3	70.1	72.3

图 5-2 从上到下依次展示了四种方法在 CAVIAR 数据集上几种类别的轨迹的分类情况，从左到右依次为所有类别、第 1 类、第 4 类、第 9 类和第 19 类轨迹。例如，图 5-2(a) 一行显示的是分层 K-Means 聚类方法在所有类别、第 1 类、第 4 类、第 9 类和第 19 类轨迹样例上的分类结果，同理，第二行图 5-2(b) 显示并行谱聚类方法结果，第三行图 5-2(c) 对应 GMMs 方法，第四行图 5-2(d) 是 CLSR 的结果。可以看出，第 9 类（绿色）和第 19 类（红色）轨迹有着位置相近、形状相似的共同点，所以二者不易区分，右上角的分层 K-Means 聚类方法结果中就存在两种方法交叉分错的情况。然而，CLSR 方法由于采用分段局部重构，不仅更准确地描述了轨迹的弯曲形状，还能用到最小化能量函数中的轨迹端点信息约束，所以局部约束和端点先验保证了该方法更好的分类性能。

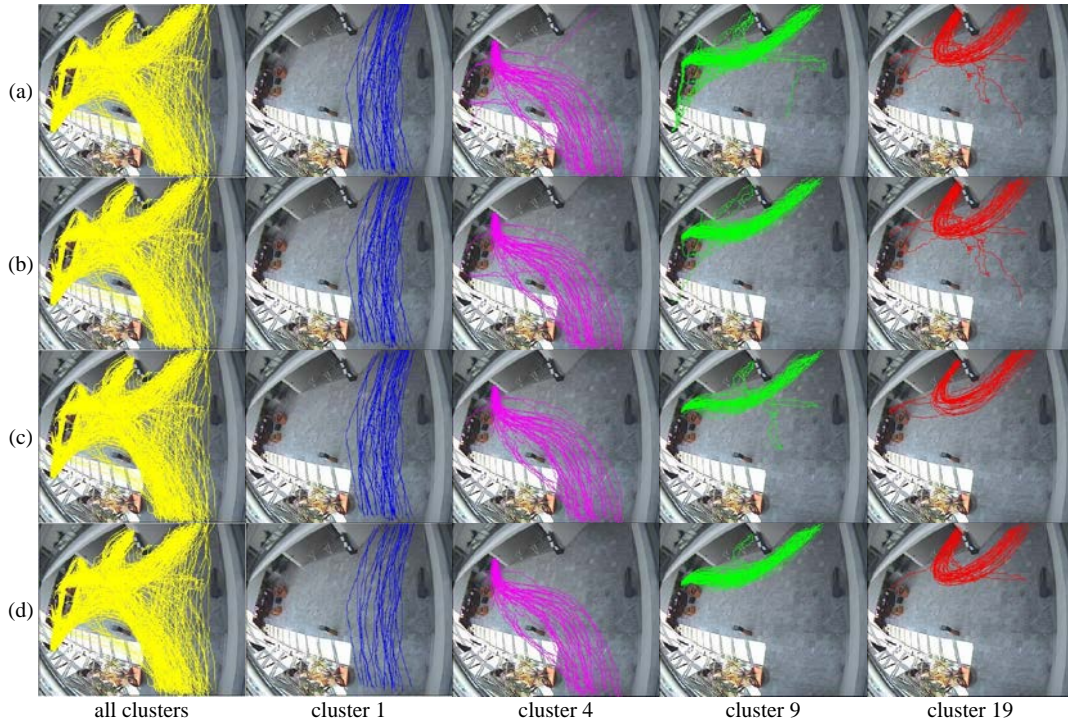


图 5-2 CAVIAR 数据集上四种方法的轨迹分类结果对比图

表 5-4 列出了上述四种方法在 Carpark 数据集上轨迹分类的实验结果。本节分别使用每种类别随机选用 $M = \{5, 10, 15\}$ 个样本来做对比，从表格中可以看出，本文提出的 CLSR 方法的正确分类率 CCR 也明显高于其他方法，验证了 CLSR 方法的非监督增量学习机制在轨迹分类问题上的有效性。图 5-3 从上到下依次展示了四种方法在 Carpark 数据集上几种类别的轨迹的分类情况，从左到右依次为所有类别、第 1 类、第 3 类、第 6 类和第 7 类轨迹。从图中可以看出，第 3 类（紫色）、第 6 类（绿色）和第 7 类（红色）三类轨迹存在错误分类情况，这是因为都经过了同一个十字路口，轨迹中存在一段方向相同、形状相似的区域，导致这些轨迹难以区分。尤其对于紫色的第 3 类从路口东侧向西南方向行驶的轨迹和第 4 类从路口西侧向西南方向行驶的轨迹，二者在路口向西南转弯后有大段相同位置相似形状的轨迹，这两类轨迹非常难分类正确。然而，在其他类别上，CLSR 方法还是表现出了良好鲁棒的分类性能。

通过这两组实验结果可知，CLSR 方法学出的基于聚类的稀疏重构模型能够相比显著地提高了非监督的轨迹学习的性能，可以用于目标轨迹分析的轨迹分类问题，解决现有技术在实际应用中训练样本的类别标注非常难以获取的问

题。

表 5-4 Carpark 数据集上四种方法的分类性能对比

方法	分类正确率 (%)	每个类别训练样本个数		
		5	10	15
分层 K-Means 聚类		17.9	33.7	37.5
并行谱聚类		18.5	36.9	45.8
GMMs		-	-	37.8
CLSR		38.1	66.7	69.2

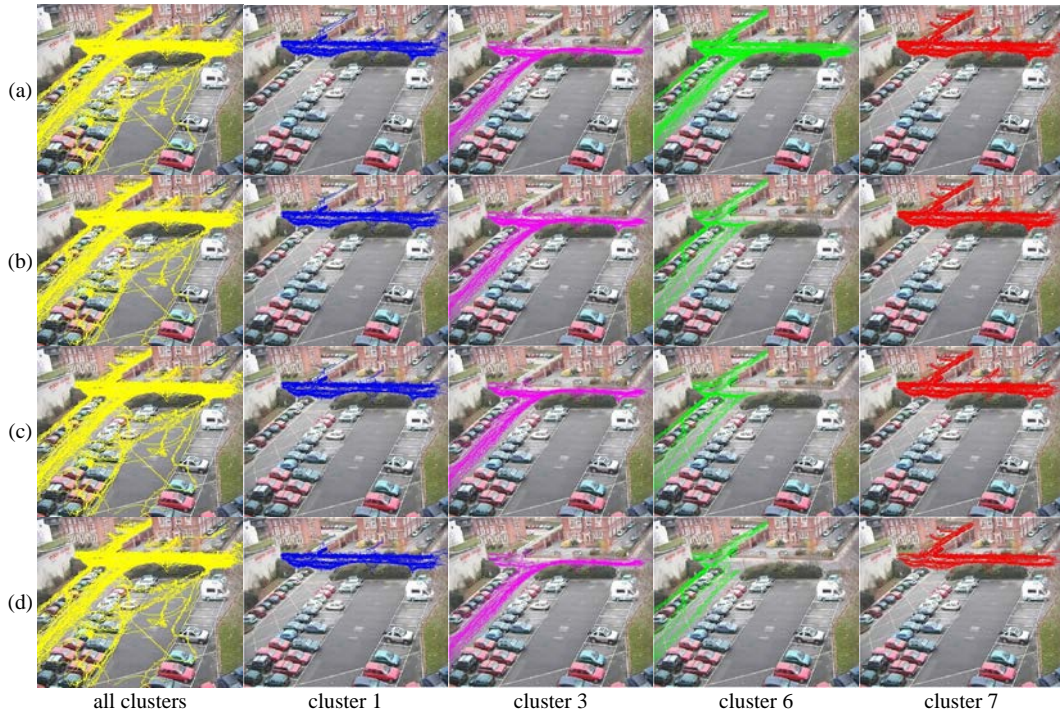


图 5-3 Carpark 数据集上四种方法的轨迹分类结果对比图

5.4.2 增量学习的有效性验证

为了进一步对所提出的增量字典学习方法的有效性进行验证，本节对两个数据集上的增量学习框架进行实验。本节的增量学习模型中需要涉及到正则化

因子、学习率、迭代次数等参数,通过优化学习,得到的参数值分别为: $\nu_1 = 0.001$ 、

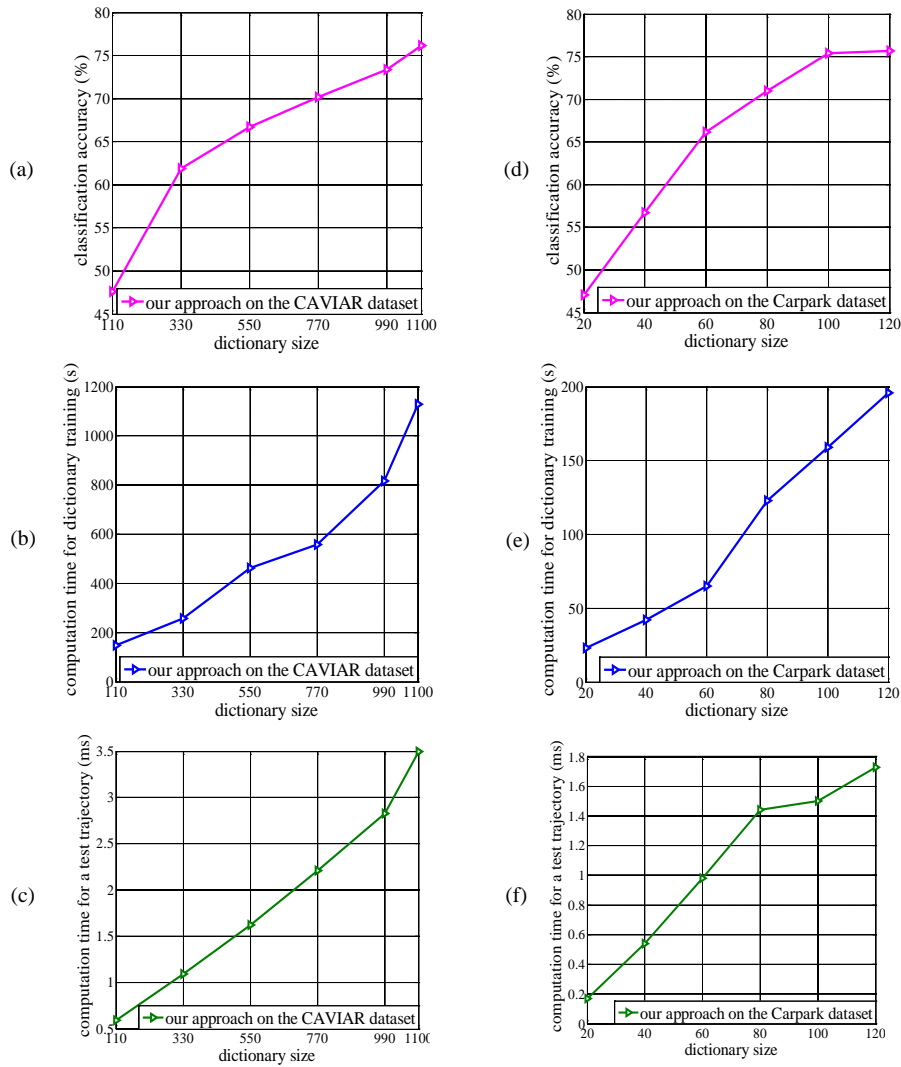


图 5-4 CAVIAR 和 Carpark 数据集上增量学习模型的分类和耗时性能对比

$\nu_2 = 0.001$ 和 $\gamma = 1$ 。CAVIAR 数据集上迭代次数取 $\hat{T} = 20$ 、稀疏系数判别误差的参设为 $\mu = 0.5$,学习字典的样本数目依次取 $M = \{110, 330, 550, 770, 990, 1100\}$;而 Carpark 数据集上迭代次数取 $\hat{T} = 10$ 、稀疏系数判别误差的参设为 $\mu = 0.3$,学习字典的样本数分别取 $M = \{20, 40, 60, 80, 120\}$ 。图 5-4 对比了 CAVIAR 数据集上不同字典大小时的正确分类率、训练样本耗时和测试样本平均耗时的变化情况。图 5-4(a)显示采用一些初始样本来增量地更新模型方法可行,图 5-4(b)表明随着字典学习样本数目的增加,训练耗时也线性地增加,而且每个样本训

练的平均耗时不超过 1s，在图 5-4(c)中可以看到当字典样本数达到 1100 时，每个测试样本平均耗时 3.5ms。与此类似，我们在图 5-4(d)、图 5-4 (e)、图 5-4 (f)中也分别对比了 Carpark 数据集上增量字典学习模型的正确分类率、训练样本耗时和测试样本平均耗时的变化情况，验证了增量学习模型的有效性。

5.5 本章小结

本章阐述了一种基于聚类的局部稀疏重构的目标轨迹分析方法，首先采用 K-Means 聚类做轨迹初始化和学习样本的特征选择，然后在增量学习的框架中，逐步实现字典初始化、字典学习和字典优化，最后将学习得到的字典经过局部稀疏重构后用于轨迹分类。在两个数据集上的实验结果表明，基于聚类的局部稀疏重构的目标轨迹分析方法与其他几种有代表性的非监督方法相比，轨迹分类的准确率有了明显提高，有助于解决现有技术在实际应用中训练样本的类别标注非常难以获取的问题。

第六章 总结与展望

目标轨迹分析是计算机视觉的一个重要问题,涉及到模式识别与智能系统、图像处理、统计学、机器学习等多个方面。经过近二十多年的深入研究和发
展,目标轨迹分析技术已在安防监控、交通分析以及视频检索等方面广泛应用。然
而,到目前为止,一般意义上的目标轨迹分析还有很大发展空间。本文面向视
频监控中基于轨迹的运动目标轨迹分析的基础理论,针对轨迹的形状、方向和
长度多样、局部噪声、学习样本标记少等典型难点问题,逐步提出了非监督的
目标轨迹分析方法。

6.1 本文工作总结

本文针对目前视频监控中运动目标轨迹分析存在的问题,提出一些新方法,
旨在能够及时对目标的轨迹行为判断分类并检测目标的异常行为。本文的主要
研究成果如下:

(1) 传统的基于轨迹分量的特征表示方法,如轨迹的目标位置、速度分量
信息、流向量等,存在表示序列长度不一致的问题。因此,基于目标轨迹的曲
线近似表示方法,比如 Haar 小波变换、DFT 变换、Chebyshev 多项式函数近
似等,被广泛应用于目标运动轨迹的表示研究。然而,这些方法只是实现了对
轨迹数据统一长度的表示而忽略了对轨迹数据行为信息中重要特征(如形状和
方向等)的提取。因此,本文提出了一种基于最小二乘三次 B 样条近似(LCSCA)
的轨迹表示方法,该方法通过对目标轨迹的三维时变序列数据的变换求解来解
析各种轨迹曲线,可以有效地提取轨迹的形状和方向特征,给出统一长度的轨
迹表示。由于监控场景中的目标行为多样、各类轨迹学习样本数目相对较少,
本文进一步提出了一种基于稀疏重构(SRA)的目标轨迹分析算法,设计出基
于稀疏重构的目标轨迹分析模型,并构建了轨迹分类和异常检测的分类器。实
验结果表明,与经典的 GMMs 方法相比,本文提出的基于 SRA 的目标轨迹分
析方法在小样本数据集的轨迹学习上鲁棒性有了明显的提高。

(2) 在视频监控中,由于场景的复杂,目标运动轨迹的获取往往受到遮挡、
背景变化等因素的影响,大部分的轨迹数据带有噪声和局部形变,导致分类精

度的降低。本文提出了一种基于局部约束下稀疏重构 (LSR) 的目标轨迹分析算法, 在轨迹学习模型中用轨迹划分的方法来引入局部性约束, 设计出每段样本在其相似局部样本子集稀疏重构的模型; 另外, 还构建了一个加权的局部判别式字典, 根据所定义的判别编码和损失加权解码算法, 实现整个样本空间的轨迹分类和异常检测。该算法不仅可以很好地解释视频监控中目标行为的分类和异常检测很大可能仅由轨迹的几个控制点变化所导致, 而且能够通过局部稀疏理论很好地选择出最具有代表性的相似分段来稀疏重构测试样本, 降低了错分率, 提高了轨迹分类和异常检测的效率。在室内大厅 CAVIAR 和停车场 Carpark 两个视频监控数据集上的实验证明了所提出方法的鲁棒性和有效性。

(3) 当前视频分析的大多有监督的学习方法, 一般需要获取更多的标记样本, 需要耗费的人力物力成本很高, 而且实际中要想找到直接有利于分类的数据比较困难。本文提出了一种基于聚类的局部稀疏重构 (CLSR) 目标轨迹分析方法。该算法利用增量字典学习的特性, 首先提出了基于 K-Means 聚类的增量学习轨迹特征局部字典方法, 然后使用局部字典的最小化能量函数解码算法, 不仅解决了学习模型中缺少类别先验的问题, 而且由于选择片段内相似度高的局部样本集合来稀疏表示轨迹, 所以提高了目标行为分类的准确度, 解决了现有技术在实际应用中训练样本的类别标注非常难以获取的问题。与其他几种有代表性的非监督轨迹学习方法相比, 基于 CLSR 的目标轨迹分析方法具有较高的准确率。

6.2 未来工作展望

目标轨迹分析是一项充满了挑战的研究课题。尽管本文针对其中的几个关键问题进行了一些探索和尝试, 取得了一些研究成果。但是, 必须指出的是, 目标轨迹分析仍然是一个开放问题, 将现有技术直接应用于视频监控分析, 在理论和应用上都还存在一些明显不足有待于研究和完善。结合本文研究中所遇到的问题, 关于未来的研究工作提出以下展望:

(1) 本文提出的 LCSCA 轨迹表示虽然解决了轨迹表示的形状、方向、长度等问题, 但带来了算法中的控制点个数 p 的优化选择问题。实际上, 离线的参数优化显然并不能满足工业应用中的需求。所以, 我们正在尝试研究轨迹表示参数的在线选择问题, 目前在可扩展子空间学习上做了初步尝试。如何更好地

定义轨迹字典学习模型，从而实现轨迹特征表示和字典基向量相结合的优化选择，这是下一步工作的首要任务。

(2) 本文在轨迹分类和异常检测中做了深入的研究，然而，关于轨迹分析结果的语义表达，目前本文只是用简单的类别表示或异常预警，这对于智能化的视频监控显然是不够的。因此，下一步研究工作可以尝试在目标运动中给出轨迹的类别概率，能够更形象地表达分析的语义，这也将涉及轨迹预测的问题，目前同一课题组的硕士生正对此问题展开研究。

(3) 目前在轨迹分析，非监督特征学习目前研究还比较少，如何自动学习获取鲁棒的轨迹的行为模式表达对于轨迹分析应用具有至关重要的意义。本文将非监督的聚类和字典学习结合在一起，不过只是从方法上找到了 **K-Means** 和 **LC-KSVD** 的契合点，还需要进一步从理论分析上深入研究。

最后需要指出的是，视频监控中运动目标轨迹分析算法研究是受到实际需求驱动的应用型研究，需要针对应用系统中的实际问题，不断研究新技术、新算法，解决新问题。

参考文献

- [1] Sumpter N., Bulpitt A.. Learning spatio-temporal patterns for predicting object behavior[J]. *Image and Vision Computing*, 2000, 18(9): 697-704.
- [2] 肖雪娟. 车辆运动行为的视觉分析[D]. 北京: 中国科学院自动化研究所, 2005.
- [3] 谷军霞, 丁晓青, 王生进. 行为分析算法综述[J]. *中国图像图形学报*, 2009: 377-387.
- [4] Tjondronegoro D., Chen Y.P.P., Pham B.. Content-based video indexing for sports applications using integrated multi-modal approach[C]. *In: Proceedings of ACM International Conference on Multimedia*, 2005: 1035-1036.
- [5] Makris D., Ellis T.. Automatic learning of an activity-based semantic scene model[C]. *In: Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2003: 183-183.
- [6] Remagnino P., Tan T., Baker K.. Multi-agent visual surveillance of dynamic scenes[J]. *Image and Vision Computing*, 1998, 16(8): 529-532.
- [7] Hu W., Xie D., Tan T.. A hierarchical self-organizing approach for learning the patterns of motion trajectories[J]. *IEEE Transactions on Neural Networks*, 2004, 15(1): 135-144.
- [8] 夏凡, 王宏. 基于局部异常行为检测的欺骗识别研究[J]. *智能系统学报*, 2007: 12-19.
- [9] Zhu G., Huang Q., Xu C., et al. Human behavior analysis for highlight ranking in broadcast racket sports video[J]. *IEEE Transactions on Multimedia*, 2007, 9(6): 1167-1182.
- [10] 李力, 王飞跃, 郑南宁等. 驾驶行为智能分析的研究与发展[J]. *自动化学报*, 2007, 33(10): 1014-1022.
- [11] 毛治力, 张宝昌, 徐彬. 基于 Q 学习的无人机航迹规划[J]. 2012.
- [12] 韩磊. 图像序列中人的行为分析和识别方法 [D]. 北京: 北京理工大学, 2009.
- [13] Yamato J., Ohya J., Ishii K.. Recognizing human action in time-sequential images using hidden markov model[C]. *In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1992: 379-385.
- [14] Davis J.W., Bobick A.F.. The representation and recognition of human movement using temporal templates[C]. *In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1997: 928-934.
- [15] Ben-Arie J., Wang Z., Pandit P., et al. Human activity recognition using multidimensional indexing[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(8): 1091-1104.
- [16] Luo Y., Wu T.D., Hwang J.N.. Object-based analysis and interpretation of human motion in sports video sequences by dynamic Bayesian networks[J]. *Computer Vision and Image Understanding*, 2003, 92(2): 196-216.
- [17] Shimozaki M., Kuniyoshi Y.. Integration of spatial and temporal contexts for action recognition by self organizing neural networks[C]. *In: Proceedings of IEEE Conference on*

- Intelligent Robots and Systems*, 2003, 3: 2385-2391.
- [18] Owens J., Hunter A.. Application of the self-organising map to trajectory classification[C]. *In: Proceedings of IEEE Conference on Visual Surveillance*, 2000: 77-83.
- [19] Sillito R.R., Fisher R.B.. Semi-supervised Learning for Anomalous Trajectory Detection[C]. *In: Proceedings of British Machine Vision Conference*, 2008:1-10.
- [20] Sillito R.R., Fisher R.B.. Incremental one-class learning with bounded computational complexity[C]. *In: Proceedings of IEEE Conference on Artificial Neural Networks*, 2007: 58-67.
- [21] Hu W., Tan T., Wang L., et al. A survey on visual surveillance of object motion and behaviors[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 2004, 34(3): 334-352.
- [22] Wang X., Ma X., Grimson W.E.L.. Unsupervised activity perception in crowded and complicated scenes using hierarchical bayesian models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(3): 539-555.
- [23] Wang X., Ma K.T., Ng G.W., et al. Trajectory analysis and semantic region modeling using nonparametric hierarchical bayesian models [J]. *International Journal of Computer Vision*, 2011, 95(3): 287-312.
- [24] Liu H.C., Wu D.B., Yih J.M., et al. Fuzzy possibility C-Mean based on mahalanobis distance and separable criterion[J]. *Transactions on Biology and Biomedicine*, 2007, 4(7): 93-98.
- [25] Kohonen T.. *Self-Organizing Maps* [M]. Heidelberg: Springer Series in Information Sciences, 1997.
- [26] Lee J.G., Han J., Whang K.Y.. Trajectory clustering: a partition-and-group framework[C]. *In: Proceedings of International Conference on ACM Special Interest Group on Management of Data*, 2007: 593-604.
- [27] Cuevas E.V.. *Kalman filter for vision tracking*[R]. Berlin: Freie University, 2005.
- [28] Brown R.G., Hwang P.Y.C.. *Introduction to random signals and applied Kalman filtering*[M]. New York: Wiley, 1992.
- [29] Collins R.T.. Mean-shift blob tracking through scale space[C]. *In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2003, 2: II, 234-40.
- [30] Comaniciu D., Ramesh V., Meer P.. Real-time tracking of non-rigid objects using mean shift[C]. *In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2000, 2: 142-149.
- [31] 蔡荣太, 吴元昊. 视频目标跟踪算法综述[J]. *视频应用与工程*, 2010, 34 (12): 135-149.
- [32] Han Z., Jiao J., Ye Q.. A fast object tracking approach based on sparse representation[C]. *In: Proceedings of IEEE International Conference on Image Processing*, 2011: 1865-1868.
- [33] Han Z., Jiao J., Zhang B., et al. Visual object tracking via sample-based adaptive sparse representation (AdaSR)[J]. *Pattern Recognition*, 2011, 44(9): 2170-2183.
- [34] Han Z., Ye Q., Jiao J.. Online feature evaluation for object tracking using kalman filter[C]. *In: Proceedings of IEEE International Conference on Pattern Recognition*, 2008: 1-4.
- [35] Li L., Han Z., Ye Q., et al. Visual object tracking via one-class SVM[C]. *In: Proceedings*

- of *IEEE International Conference on Asian Conference on Computer Vision Workshops*, 2011: 216-225.
- [36] Li L., Han Z., Jiao J., et al. Visual object tracking via sparse reconstruction[C]. *In: Proceedings of International Conference on Internet Multimedia Computing and Service*, 2010: 151-154.
- [37] Kalal Z., Matas J., Mikolajczyk K.. Online learning of robust object detectors during unstable tracking[C]. *In: Proceedings of IEEE International Conference on Computer Vision Workshops*, 2009: 1417-1424.
- [38] Fisher R.B., Santos-Victor J. and Crowley J.. CAVIAR test case scenarios [DB/OL]. URL <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>.
- [39] Cambridge Systematics Inc. NGSIM vehicle trajectory datasets [DB/OL]. URL <http://ngsim.camsys.com/>.
- [40] Dee H.M., Hogg D.. Detecting inexplicable behavior[C]. *In: Proceedings of British Machine Vision Conference*, 2004: 1-10.
- [41] Johnson N., Hogg D.. Learning the distribution of object trajectories for event recognition[J]. *Image and Vision Computing*, 1996, 14(8): 609-615.
- [42] Stauffer C., Grimson W.E.L.. Learning patterns of activity using real-time tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 747-757.
- [43] Hu W., Xiao X., Fu Z., et al. A system for learning statistical motion patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(9): 1450-1464.
- [44] Zhang Z., Huang K., Tan T.. Comparison of similarity measures for trajectory clustering in outdoor surveillance scenes[C]. *In: Proceedings of IEEE International Conference on Pattern Recognition*, 2006, 3: 1135-1138.
- [45] Zhang Z., Huang K., Tan T., et al. Trajectory series analysis based event rule induction for visual surveillance[C]. *In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2007: 1-8.
- [46] Piciarelli C., Foresti G.L.. On-line trajectory clustering for anomalous events detection[J]. *Pattern Recognition Letters*, 2006, 27(15): 1835-1842.
- [47] Piciarelli C., Micheloni C., Foresti G.L.. Trajectory-based anomalous event detection[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2008, 18(11): 1544-1554.
- [48] Vasquez D., Fraichard T., Laugier C.. Growing hidden markov models: an incremental tool for learning and predicting human and vehicle motion[J]. *The International Journal of Robotics Research*, 2009.
- [49] Naftel A., Khalid S.. Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space[J]. *Multimedia Systems*, 2006, 12(3): 227-238.
- [50] Sillito R.R., Fisher R.B.. Parametric trajectory representations for behaviour classification[J]. *Multimedia Systems*, 2006, 12: 227-238.
- [51] Barber S.. All of statistics: a concise course in statistical inference[J]. *Journal of the Royal Statistical Society: Series A*, 2005, 168(1): 261-261.

- [52] Press W.H.. *Numerical recipes 3rd edition: The art of scientific computing*[M]. Cambridge university press, 2007.
- [53] 刘洋. 基于视觉的车辆路面行为分析关键技术研究[D]. 北京: 中国科学院研究生院, 2010.
- [54] Oliver N.M., Rosario B., Pentland A.P.. A bayesian computer vision system for modeling human interactions[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 831-843.
- [55] Schölkopf B., Platt J.C., Shawe-Taylor J., et al. Estimating the support of a high-dimensional distribution[J]. *Neural computation*, 2001, 13(7): 1443-1471.
- [56] Wang X., Ma K.T., Ng G.W., et al. Trajectory analysis and semantic region modeling using nonparametric hierarchical bayesian models[J]. *International Journal of Computer Vision*, 2011, 95(3): 287-312.
- [57] Li X., Han J., Lee J.G., et al. *Traffic density-based discovery of hot routes in road networks*[M]. Advances in Spatial and Temporal Databases, 2007: 441-459.
- [58] 郝久月, 李超, 高磊, 熊璋. 智能监控场景中运动目标轨迹聚类算法[J]. 北京航空航天大学学报, 2009: 1083-1087.
- [59] Knorr E.M., Ng R.T., Tucakov V.. Distance-based outliers: algorithms and applications[J]. *International Journal on Very Large Data Bases*, 2000, 8(3-4): 237-253.
- [60] Li X., Han J., Kim S., et al. ROAM: rule-and motif-based anomaly detection in massive moving object data sets[C]. In: *Proceedings of SIAM International Conference on Data Mining*, 2007, 7: 273-284.
- [61] Lee J.G., Han J., Li X.. Trajectory outlier detection: A partition-and-detect framework[C]. In: *Proceedings of International Conference on Data Engineering*, 2008: 140-149.
- [62] 刘良旭, 乐嘉锦, 乔少杰, 宋加涛. 基于轨迹点局部异常度的异常点检测算法[J]. 计算机学报, 2011, 34(10): 1966-1975.
- [63] Hermes C., Wohler C., Schenk K., et al. Long-term vehicle motion prediction[C]. In: *Proceedings of International Conference on Intelligent Vehicles Symposium*, 2009: 652-657.
- [64] Bashir F.I., Khokhar A.A., Schonfeld D.. Object trajectory-based activity classification and recognition using hidden markov models[J]. *IEEE Transactions on Image Processing*, 2007, 16(7): 1912-1919.
- [65] Schumaker L.. *Spline functions: basic theory*[M]. Cambridge University Press, 2007.
- [66] Unser M., Aldroubi A., Eden M.. B-spline signal processing II Efficiency design and applications[J]. *IEEE Transactions on Signal Processing*, 1993, 41(2): 834-848.
- [67] Gordon W.J., Riesenfeld R.F.. Bernstein-Bézier methods for the computer-aided design of free-form curves and surfaces[J]. *Journal of the Association for Computing Machinery*, 1974, 21(2): 293-310.
- [68] De Boor C.. A practical guide to splines[J]. *Mathematics of Computation*, 1978.
- [69] 崔明根, 吴勃英, 王德明, 丁效华编著. 数值分析原理[M]. 科学出版社, 2003.
- [70] 阎庆旭, 陈兆斗, 刘慧芳. Weierstrass 逼近定理的应用[M]. 数学的实践与认识, 2004:

174-176.

- [71] Hastie T., Tibshirani R., Friedman J., et al. *The elements of statistical learning*[M]. New York: springer, 2009.
- [72] Candès E.J., Romberg J., Tao T.. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information[J]. *IEEE Transactions on Information Theory*, 2006, 52(2): 489-509.
- [73] Donoho D.L.. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2006, 52(4): 1289-1306.
- [74] Baraniuk R.. A lecture on compressive sensing[J]. *IEEE Signal Processing Magazine*, 2007: 118-121.
- [75] DeVore R.A.. Deterministic constructions of compressed sensing matrices[J]. *Journal of Complexity*, 2007, 23(4): 918-925.
- [76] Tsaig Y., Donoho D.L.. Extensions of compressed sensing[J]. *Signal processing*, 2006, 86(3): 549-571.
- [77] Bradley P.S., Mangasarian O.L.. Feature selection via concave minimization and support vector machines[C]. In: *Proceedings of IEEE International Conference on Machine Learning*, 1998, 98: 82-90.
- [78] Chen S., Donoho D.. Basis pursuit[C]. In: *Proceedings of IEEE International Conference on Signals, Systems and Computers*, 1994, 1: 41-44.
- [79] Tibshirani R.. Regression shrinkage and selection via the Lasso[J]. *Journal of the Royal Statistical Society. Series B*, 1996: 267-288.
- [80] Kashin B.S.. Diameters of some finite-dimensional sets and classes of smooth functions[J]. *Mathematics of the USSR-Izvestiya*, 1977, 41(2): 334-351.
- [81] Elad M., Bruckstein A.M.. A generalized uncertainty principle and sparse representation in pairs of bases[J]. *IEEE Transactions on Information Theory*, 2002, 48(9): 2558-2567.
- [82] Donoho D.L., Huo X.. Uncertainty principles and ideal atomic decomposition[J]. *IEEE Transactions on Information Theory*, 2001, 47(7): 2845-2862.
- [83] Candès E.J., Romberg J.K., Tao T.. Stable signal recovery from incomplete and inaccurate measurements[J]. *Communications on Pure and Applied Mathematics*, 2006, 59(8): 1207-1223.
- [84] Mallat S.G., Zhang Z.. Matching pursuits with time-frequency dictionaries[J]. *IEEE Transactions on Signal Processing*, 1993, 41(12): 3397-3415.
- [85] Tropp J.A., Gilbert A.C.. Signal recovery from random measurements via orthogonal matching pursuit[J]. *IEEE Transactions on Information Theory*, 2007, 53(12): 4655-4666.
- [86] Chen S.S., Donoho D.L., Saunders M.A.. Atomic decomposition by basis pursuit[J]. *SIAM Journal on Scientific Computing*, 1998, 20(1): 33-61.
- [87] Figueiredo M.A.T., Nowak R.D., Wright S.J.. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2007, 1(4): 586-597.
- [88] Yang J., Peng Y., Xu W., et al. Ways to sparse representation: a comparative study[J].

- Tsinghua Science & Technology*, 2009, 14(4): 434-443.
- [89] Mairal J., Elad M., Sapiro G.. Sparse representation for color image restoration[J]. *IEEE Transactions on Image Processing*, 2008, 17(1): 53-69.
- [90] Wright J., Yang A.Y., Ganesh A., et al. Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [91] Yang J., Wright J., Huang T., et al. Image super-resolution as sparse representation of raw image patches[C]. In: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2008: 1-8.
- [92] Thiagarajan J.J., Spanias A.. Learning dictionaries for local sparse coding in image classification[C]. In: *Proceedings of IEEE International Conference on Signals, Systems and Computers*, 2011: 2014-2018.
- [93] Jia X., Lu H., Yang M.H.. Visual tracking via adaptive structural local sparse appearance model[C]. In: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2012: 1822-1829.
- [94] Feng J., Zhou J., Jain A.K.. Orientation field estimation for latent fingerprint enhancement[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(4): 925-940.
- [95] Li C., Han Z., Ye Q., et al. Visual abnormal behavior detection based on trajectory sparse reconstruction analysis[J]. *Neurocomputing*, 2013, 119: 94-100.
- [96] Alpaydin E., Mayoraz E.. Learning error-correcting output codes from data[C]. In: *Proceedings of IEEE International Conference on Artificial Neural Networks*, 1999, 2: 743-748.
- [97] Utschick W., Weichselberger W.. Stochastic organization of output codes in multiclass learning problems[J]. *Neural Computation*, 2001, 13(5): 1065-1102.
- [98] Pujol O., Radeva P., Vitria J.. Discriminant ecoc: A heuristic method for application dependent design of error correcting output codes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(6): 1007-1012.
- [99] Escalera S., Tax D.M.J., Pujol O., et al. Subclass problem-dependent design for error-correcting output codes[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(6): 1041-1054.
- [100] Dong W., Li X., Zhang D., et al. Sparsity-based image denoising via dictionary learning and structural clustering[C]. In: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2011: 457-464.
- [101] Hartigan J.A., Wong M.A.. Algorithm AS 136: A k-means clustering algorithm[J]. *Applied Statistics*, 1979: 100-108.
- [102] Mairal J., Ponce J., Sapiro G., et al. Supervised dictionary learning[C]. In: *Proceedings of IEEE International Conference on Advances in Neural Information Processing Systems*, 2009: 1033-1040.
- [103] Yang J., Yu K., Huang T.. Supervised translation-invariant sparse coding[C]. In: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*,

2010: 3517-3524.

[104] Pham D.S., Venkatesh S.. Joint learning and dictionary construction for pattern recognition[C]. *In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2008: 1-8.

[105] Lewicki M., Sejnowski T.. Learning overcomplete representations[J]. *Neural computation*, 2000, 12(2): 337-365.

[106] Engan K., Aase S.O., Hakon Husoy J.. Method of optimal directions for frame design[C]. *In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1999, 5: 2443-2446.

[107] Murray J.F., Kreutz-Delgado K.. An improved FOCUSS-based learning algorithm for solving sparse linear inverse problems[C]. *In: Proceedings of IEEE International Conference on Signals, Systems and Computers*, 2001, 1: 347-351.

[108] Lesage S., Gribonval R., Bimbot F., et al. Learning unions of orthonormal bases with thresholded singular value decomposition[C]. *In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, 5: 293-296.

[109] Aharon M., Elad M., Bruckstein A.. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.

[110] Jiang Z.L., Lin Z., Davis L.S.. Label consistent k-svd: learning a discriminative dictionary for recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(11): 2651-2664.

[111] Golub G.H., Hansen P.C., O'Leary D.P.. Tikhonov regularization and total least squares[J]. *SIAM Journal on Matrix Analysis and Applications*, 1999, 21(1): 185-194.

[112] Mairal J., Bach F., Ponce J.. Task-driven dictionary learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(4): 791-804.

[113] Chen W.Y., Song Y., Bai H., et al. Parallel spectral clustering in distributed systems [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(3): 568-586.

[114] Li C., Han Z.J., Ye Q.X., et al. Locality-constrained sparse reconstruction for trajectory classification[C]. *In: Proceedings of IEEE International Conference on Pattern Recognition*, 2014: 2602-2606.

[115] Chen X.G., Ye Q.X., Zou J.L., et al. Trajectory analysis via replicated softmax based model [J]. *Signal, Image and Video Processing*, 2014, 8(1): 183-190.

[116] Jiang Z.L., Zhang G.X., Davis L.S.. Submodular Dictionary Learning for Sparse Coding[C]. *In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2012: 3418-3425.

[117] 魏龙翔, 何小海, 滕奇志. 结合 Hausdorff 距离和最长公共子序列的轨迹分类[J]. *电子与信息学报*, 2013, 35(4): 784-790.

[118] Zhu Y., Lucey S.. Convolutional sparse coding for trajectory reconstruction[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 529-540.

参考文献

- [119] Ye J., Zhao Z., Wu M.. Discriminative k-means for clustering[C]. *In: Proceedings of IEEE International Conference on Advances in Neural Information Processing System*, 2008: 1649-1656.
- [120] Zografos V., Ellis L., Mester R.. Discriminative Subspace Clustering[C]. *In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, 2013:2107-2114.

附录：主要符号对照表

BP	基追踪 (Basis Pursuit)
CAGD	计算机辅助几何设计 (Computer Aided Geometric Design)
CHMM	耦合隐马尔可夫模型 (Coupled Hidden Markov Model)
CLSR	聚类的局部稀疏重构 (Cluster-based Locality-constrained Sparse Reconstruction)
DCM	判别编码矩阵 (Discriminate Code Matrix)
DFT	离散傅里叶变换 (Discrete Fourier Transform)
DTW	动态时间规整 (Dynamic Time Warping)
ECOC	纠错输出编码 (Error Correcting Output Codes)
EMD	最小化能量函数解码 (Energy Minimization Decode)
FCM	模糊 C-均值 (Fuzzy C-Means)
FPR	误报率 (False Positive Rate)
FSM	有限状态机 (Finite State Machine)
FSOM	模糊自组织映射 (Fuzzy Self-Organizing Map)
GHMM	增长的隐马尔可夫模型 (Growing Hidden Markov Model)
GMMs	高斯混合模型 (Gauss Mixture Models)
HKMC	分层 K-Means 聚类 (Hierarchical K-Means Clustering)
HMM	隐马尔可夫模型 (Hidden Markov Model)
LCKSVD	标签一致的 K-奇异值分解 (Label Consistent K-Singular Value Decomposition)
LCSCA	最小二乘三次 B 样条的曲线逼近 (Least-squares Cubic Spline Curves Approximation)

附录：主要符号对照表

LCSS	最长公共子序列 (Longest Common Subsequence)
LSR	局部约束下稀疏重构 (Locality-constrained Sparse Reconstruction)
LWD	损失加权解码 (Loss Weighted Decode)
MP	匹配追踪 (Matching Pursuit)
One-class SVM	一类支持向量机 (One Class Support Vector Machine)
OMP	正交匹配追踪算法 (Orthogonal Matching Pursuit)
OPSR	梯度投影稀疏重构 (Gradient Projection for sparse Reconstruction)
PCA	主成分分析 (Principle Components Analysis)
PSC	并行谱聚类 (Parallel Spectral Clustering)
SOM	自组织映射 (Self-Organizing Map)
SRA	稀疏表示分类 (Sparse Representation Classification)
SRC	稀疏重构分析 (Sparse Reconstruction Analysis)
SVD	奇异值分解 (Singular Value Decomposition)
TLD	检测学习跟踪 (Tracking-Learning-Detection)
TPR	命中率 (True Positive Rate)
TRAOD	轨迹异常检测 (Trajectory Outlier Detection)
UKR	非监督核回归 (Unsupervised Kernel Regression)

作者简介与论文发表

2009.09-2012.07	中国科学院研究生院	物流工程	硕士
2012.09-2015.06	中国科学院大学	计算机应用技术	博士

已发表文章目录

1. **Ce Li**, Zhenjun Han, Qixiang Ye, Shan Gao, Lijin Pang, and Jianbin Jiao. Locality-constrained sparse reconstruction for trajectory classification[C]. *In: Proceedings of IEEE International Conference on Pattern Recognition (ICPR)*, 2014: 2602-2606. (EI)
2. **Ce Li**, Zhenjun Han, Qixiang Ye, and Jianbin Jiao. Visual abnormal behavior detection based on trajectory sparse reconstruction analysis [J]. *Neurocomputing*, 2013,119: 94-100.(SCI)
3. **Ce Li**, Zhenjun Han, Qixiang Ye, and Jianbin Jiao. Abnormal behavior detection via sparse reconstruction analysis of trajectory [C]. *In: Proceedings of IEEE International Conference on Image and Graphics (ICIG)*, 2011:807-810. (EI)
4. Shan Gao, Zhenjun Han, **Ce Li**, Qixiang Ye, and Jianbin Jiao. Real-time Multi-pedestrian Tracking in Traffic Scenes via an RGB-D based Layered Graph Model [J]. *IEEE Transaction on Intelligent Transportation Systems (TITS)*, 2014. (SCI)
5. Xiaogang Chen, Qixiang Ye, Jialing Zou, **Ce Li**, Yanting Cui, and Jianbin Jiao. Trajectory analysis via replicated softmax based model [J]. *Signal, Image and Video Processing (SIVP)*,2014,8(1):183-190.(SCI)
6. Wei Yang, Liguozhang, Wei Ke, **Ce Li**, and Jianbin Jiao. Minimum entropy models for laser line extraction[C]. *In: Proceedings of International Conference on Computer Analysis of Images and Patterns (CAIP)*, 2013:499-506. (EI)

7. Shan Gao, Zhenjun Han, **Ce Li**, and Jianbin Jiao. Real-time multi-pedestrian tracking based on vision and depth information fusion [C]. *In: Proceedings of Pacific-Rim Conference on Multimedia (PCM)*, 2013:708-719. (EI)
8. Yifei Liu, Zhenjun Han, Qixiang Ye, Jianbin Jiao, and **Ce Li**. Sparse feature representation for visual tracking [C]. *In: Proceedings of IEEE International Conference on Systems and Informatics (ICSI)*, 2012:2050-2054. (EI)

已修改专著目录

1. 焦建彬, 叶齐祥, 韩振军, 李策. 《视觉目标检测与跟踪》[M]. 北京: 科学出版社, 2015.

致 谢

光阴荏苒，岁月如梭，转眼间博士学习生涯却已接近尾声。回首过去，思绪万千，我曾幸运地踏入了这座优秀的学府，也曾在这里经历了诸多坎坷，付出了艰辛的努力，同时也得到了很大的收获，这段重要而又美好的时光值得永生珍藏。在此毕业论文完成之际，我想由衷地向曾经给予我无数帮助的老师、同学、朋友和家人表示深深的谢意！

首先，要特别感谢我的恩师焦建彬教授！焦老师治学严谨求实，一丝不苟，我的论文从选题、研究方法以及论文撰写上，每一环节无一不倾注着他的心血和指导。感谢焦老师在工作中对我的谆谆教导，恩师的国际化的视野、前沿而精髓的学术造诣、严谨求实的治学风格使我受益匪浅。此外，他还在在生活和思想上关心我、照顾我，焦老师从容、乐观、以身立行的为人风格使我学到了很多珍贵的道理，深刻影响着我今后的工作和生活。再次表示万分的感谢！

其次，衷心感谢实验室叶齐祥副教授和韩振军副教授！两位老师在我的理论学习和实验研究中给予了耐心的指导和帮助，他对于研究领域精深的理解以及严谨的治学态度，开阔了我的研究视野，启发了我在研究中发现新问题。此外，两位老师还在生活上给予了极大的关心和帮助，再次表示感谢！

感谢师兄师姐、师弟师妹们在科研上的帮助与合作，感谢一起学习生活中结下的深厚友谊。感谢小分队、4ever 和熊伙伴们在我遇到困难时的陪伴和鼓励，感谢一直以来朋友带给我精神上的莫大支持。

再次，深深地感谢我的父母多年来的支持与鼓励，养育之恩无以为报。感谢我的姐姐在生活上的热心帮助，手足之情尽在心中。家人们对我工作和学业的热情鼓励和殷切期望，都给我注入了巨大的动力，他们无私的爱和无条件的支持一直指引我奋发拼搏、积极向上，使我能够顺利完成博士学业。

最后，感谢在百忙之中抽出宝贵时间参加论文评审和答辩的各位老师！

李策

2015 年 4 月