

密级: _____

中国科学院研究生院

硕士学位论文

基于稀疏表示的快速目标跟踪算法研究

作者姓名: _____ 刘一飞 _____

指导教师: _____ 焦建彬 教授 中国科学院研究生院 _____

学位类别: _____ 工学硕士 _____

学科专业: _____ 计算机应用技术 _____

培养单位: _____ 中国科学院研究生院 _____

2012年5月

Fast Object Tracking based on Sparse Representation

By

Yifei Liu

A Paper Submitted to

Graduate University of Chinese Academy of Sciences

In partial fulfillment of the requirement

For the degree of

Master of Computer Application Technology

Graduate University of Chinese Academy of Sciences

May, 2012

中国科学院研究生院直属院系 研究生学位论文原创性声明

本人郑重声明：所提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：

中国科学院研究生院直属院系 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

视频目标跟踪是计算机视觉领域的重要研究内容之一，也是高层次视频处理和应用理解的基础。目标跟踪融合了图像处理、模式识别、人工智能、自动控制等相关领域的研究成果。跟踪技术通过在连续的视频序列中找到图像中目标的确切位置，获取完整的目标运动轨迹，因此，被广泛地应用于视频监控、机器人导航、医学图像分析、气象分析以及国民经济的各个领域。本文以此为研究内容，具有重要的理论意义和广阔的应用价值。

目标表示和目标匹配搜索是目标跟踪过程中重要的两个步骤。本文提出了基于特征稀疏表示的快速目标跟踪算法，首先提取目标的颜色和梯度向量统计特征对目标进行描述，并且通过稀疏表示对目标特征进行评估和选择，然后基于此在搜索区域内对目标进行快速的匹配搜索。本文的主要贡献和研究内容如下：

1、对目标综合特征提取的改进。改进的综合直方图基于局部颜色直方图（HC, Histogram of Color）和局部梯度方向直方图（HOG, Histogram of Oriented Gradient），能够更好的融合目标的局部颜色、局部轮廓等特征，可以有效地对目标进行描述以适应目标跟踪的需要。

2、提出一种面向跟踪的目标特征稀疏表示。利用稀疏表示（Sparse Representation）的方法对目标的综合特征进行评估和选择，使得较好的特征描述能力以及前景-背景区分能力的特征得以保留，从而得到目标的鲁棒的优化表示模型，明显地提高了跟踪的准确性。

3、基于稀疏表示的均值漂移（Mean-shift）快速跟踪算法。继承了传统 Mean-shift 算法的迭代简单与实时性高的优点，融合了特征稀疏表示以增强对目标与背景的分力，同时将模板更新算法和目标尺度变化加入到传统的均值漂移跟踪算法中来，进一步显著地提高了目标跟踪实时性，扩大了算法的应用范围。

关键词：目标跟踪，均值漂移，稀疏表示

Abstract

Yifei Liu(Computer Application)

Directed by: **Jianbin Jiao**(Professor)

Visual object tracking is one of the most important branches in the computer vision, and it is also the foundation of video processing and understanding in high level. To find the precise location and then the complete motion trajectory in continuous image sequence, it combines advanced technologies and research achievements in image processing, pattern recognition, artificial intelligence, automatic control and other relative fields. It has been widely applied in video surveillance, robot navigation, medical image analysis, meteorological analysis and other field. Therefore, this subject has important theoretical significance and wide practical value.

In this paper, we propose a fast object tracking method based on sparse feature representation in Mean-shift tracking framework. We firstly extract the color and oriented gradient features to describe the object, and then evaluate and select the subset of features by sparse representation, and finally the object will be located in the searching region and the scale of the object will be calculated based on the sparse representation. The contributions of this paper are as follows.

1) **Improved combined feature representation.** The improved Histogram of Oriented Gradient and Color (HOGC) can better precisely describe the local color, local contour of the object.

2) **Sparse representation for adaptive object tracking.** Sparse representation is employed to evaluate and select the subset of the combined features, which makes the dimensions much lower and reduces the complexity to calculate the feature set of the object and its candidates. It optimizes the object model, and increases the accuracy and speed of the tracking procedure markedly.

3) **Fast object tracking based on sparse representation in Mean-shift Framework.** The new proposal tracking method effectively combines the idea of discriminative property of Sparse Representation with the advantages of high speed

and simpleness in traditional Mean-shift tracking algorithm. In addition, scale adaptation of the object during the tracking procedure is also considered and a new template updating algorithm is proposed in our tracking method, which improves the accuracy, flexibility and robustness, and expands the field of application obviously.

KEY WORDS: Object Tracking, Mean-shift, Sparse Representation

目 录

摘 要	I
目 录	V
图目录	VII
表目录	VIII
第一章 绪论	1
1.1 课题背景和研究意义	1
1.2 视频跟踪的国内外研究现状	2
1.3 视频跟踪的主要分类和难点	3
1.3.1 视频跟踪分类	3
1.3.2 视频跟踪研究难点	4
1.4 本文的研究内容	5
1.5 本文的组织结构	6
第二章 目标跟踪算法基础	9
2.1 目标表示模型	9
2.1.1 目标表示	9
2.1.2 目标特征提取	10
2.2 目标匹配搜索	11
2.2.1 基于目标表达和相似性度量的目标匹配算法	13
2.2.2 基于均值漂移的目标快速搜索算法	15
2.3 基于稀疏表示的目标跟踪基础	17
第三章 基于综合特征稀疏表示的目标跟踪算法	19
3.1 稀疏表示的基本理论及其应用	19
3.1.1 稀疏表示的基本理论	19
3.1.2 稀疏表示在图像处理中的主要应用	21
3.2 基于综合特征稀疏表示的目标跟踪算法	22
3.2.1 改进 HOGC 特征的选取	22
3.2.2 目标特征的稀疏选择	24
3.2.3 目标的匹配搜索	25
3.3 基于综合特征稀疏表示的目标跟踪实验	28
第四章 基于稀疏表示与 Mean-shift 框架的快速目标跟踪算法	33
4.1 快速目标跟踪算法基础	33
4.2 基于稀疏表示与 Mean-shift 的快速目标跟踪算法	36

4.3 基于特征稀疏表示的快速目标跟踪实验	37
4.3.1 实验步骤	38
4.3.2 实验结果	38
4.3.3 算法性能分析	40
总结与展望	41
参考文献	43
发表文章目录	51
附录：主要符号对照表	52
致 谢	53

图目录

图 2-1 目标表示模型	9
图 2-2 目标跟踪示意图	11
图 2-3 Mean-shift 示意图.....	16
图 3-1 基于综合特征稀疏表示的目标跟踪示意图	19
图 3-2 视频帧中的运动目标	22
图 3-3 HOGC 特征提取图.....	23
图 3-4 HOGC 特征稀疏前后对比.....	25
图 3-5 综合特征与模板匹配的结合	26
图 3-6 综合特征与经典 Mean-shift 的目标跟踪结果对比图.....	29
图 3-7 基于综合特征稀疏表示的人体跟踪结果	30
图 3-8 四种跟踪方法平均 DER	31
图 4-1 搜索区域积分示意图	33
图 4-2 特征稀疏前后得到的权值图对比	37
图 4-3 目标尺寸计算模板	37
图 4-4 基于稀疏表示与 Mean-shift 框架的跟踪结果.....	39

表目录

表 4-1	基于稀疏表示的快速目标跟踪算法	38
表 4-2	四种跟踪方法平均迭代次数	40

第一章 绪论

1.1 课题背景和研究意义

计算机视觉是研究应用计算机系统实现类似人类视觉系统理解客观世界的新兴学科，它的中心任务就是对视频图像进行理解。目前它是人工智能和生命科学研究的热点领域之一，近年来关于计算机视觉的研究取得了长足进步，研究成果已经开始逐步应用于社会生活的方方面面。

目标跟踪是计算机视觉的重要研究内容之一，也是当前计算机视觉研究中尚未根本解决的难点问题。目标跟踪使用摄像头等传感器捕捉图像序列，通过一系列的计算和分析，得出运动目标在每一帧图像中的二维坐标位置，然后，将图像序列中同一运动目标关联起来，得到完整的目标运动轨迹。简单的说，目标跟踪技术就是在连续的视频序列中，找到下一帧图像中目标的确切位置，为视频序列分析提供基础运动信息和数据，是后续的各种高级处理，如目标识别与分类、目标特征提取等高层次的视频处理和应用理解的基础，也是视频监控、导航等实时应用的关键组成部分，将日益广泛地应用于科学技术、航空宇航、国防建设、医药卫生以及国民经济的各个领域。

目标跟踪在以下领域将发挥重要作用，在国防领域，目标跟踪可以应用于导弹发射以及自动跟踪危险目标，为战场监控起到相当大的作用；在民用领域，该技术可用于重点建筑（核电站、水电站、水库、银行、重要桥梁等）及重要地点（国家重点保护建筑、博物馆、体育场馆、地铁站、监狱等）的安全保卫工作；在智能交通领域，可以及时监控交通事故并可以在十字路口、新建高速公路端等进行自动行人跟踪和安全预警；也可在车载的智能系统中融入行人车辆跟踪技术，对驾驶员进行驾驶状况进行安全预警，有效避免车祸的发生。在民众日常生活保障上，可用于检测可疑人物可疑行为，用于小区监控，减少犯罪率，为人民正常生活提供保障。

然而到目前为止，现有的目标跟踪算法虽然能够在一定程度上完成对运动

目标的跟踪，但这些算法大多是基于目标某一种特征集合的跟踪，对目标特征描述不够完备，并且无法自动适应目标和环境的大变化，当背景发生较大变化时，往往会失效；同时对于运动目标的跟踪需要计算机进行大量的运算，很难达到实际情况中实时性的要求。因为现实世界中的物体、特别是场景是非常丰富的，这就使得用以进行目标检测与跟踪的图像也是千变万化、甚至可以说是无穷的，同时目标在成像过程中从三维空间转到二维投影平面，会丢失一部分信息，而且目标在运动过程中可能存在旋转、遮挡等情况，所以基于计算机视觉的运动目标跟踪技术是一项需要完善的工作，解决这些问题也一直是跟踪领域的前沿和热点。

作者在论文工作期间，参与了以下课题研究：

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）；
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）；
3. “复杂环境下动态目标跟踪技术”中科院“百人计划”择优支持项目（项目编号：99T3009EA2）。

1.2 视频跟踪的国内外研究现状

近年来，对基于计算机视觉的运动目标跟踪问题正处于研究高潮中。国外关于视频跟踪的研究起步比较早，理论研究和系统设计方面都受到很多政府机构和学者的密切关注。在视频目标跟踪的理论研究方面，美国自然科学基金委员会及美国军方对复杂环境下多目标的动态检测、识别及跟踪算法的研究及其应用非常重视，多次资助有关公司进行相关算法研究；美国国防部、美国陆军 SBIR 项目；美国海军水下战事中心等部门也先后投巨资资助该类算法的开发及其相关应用研究并取得了可观的成果。近几年，国外的众多研究机构对视频目标跟踪也展开了深入的研究。如法国国家实验室（INRIA）、牛津大学动态视觉组、MIT 媒体实验室、CMU 机器人研究所、马里兰大学帕克学院、DCS of Maryland, CSD of U. Southern California, NEC Research, CSC of U. Reading, CS of UC Berkeley 等；自适应跟踪和智能跟踪的思想被相继提出，DARPA 成立了

自动目标识别工作组 ATRWG 进行目标智能识别和跟踪等相关研究。国外设计视频目标跟踪的原型系统也很多,早在 20 世纪 50 年代初期, GAC 公司就为美国海军研制开发了自动地形识别跟踪系统。其他出名的还有,美国 DARPA 设立的 VSAM 项目; Maryland 大学的实时视频监控系统 W4^[1]; 欧洲委员会自主的 AVS-PV (Advanced Video-based Surveillance Prevention of Vandalism in metro stations) 项目^[2]; 日本开展的 CDVP 计划等。

在国内,许多高等院校和研究机构在视频目标跟踪领域也投入了相当的研究精力,例如,中国科学院自动化研究所、中国科学院研究生院、清华大学,哈尔滨工业大学,西安交通大学,浙江大学,上海交通大学等一些高校和科研院所也逐渐开展视频目标跟踪方面的研究,并取得了一定研究成果。其中,中国科学院研究生院工程应用研究中心与清华大学合作将目标跟踪技术应用飞行器进近导航,对飞行器安全降落起到一定的作用;中科院计算所 JDL 实验室对于体育视频中的目标分割、检测、跟踪、和精彩片断分析进行了深入的研究^{[3][4]};中科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开发了人体运动视觉分析系统、交通行为事件分析系统、交通场景监控系统和智能轮椅视觉导航系统^[5];清华大学开发了自然环境中运动检测、跟踪和全景图生成等功能的视觉侦查系统^{[6][7]};西安交通大学的人工智能与机器人研究所使用光流和帧差两种算法深入研究了自适应巡航控制中的车辆跟踪;上海交通大学与上海市交警总队合作开发了 SUATS 系统用于交通信号控制。

1.3 视频跟踪的主要分类和难点

1.3.1 视频跟踪分类

根据不同的分类标准,可以对视频跟踪进行不同的分类。

1、根据被跟踪目标的数目可以将跟踪算法分为单目标跟踪算法和多目标跟踪算法两类。单目标跟踪算法可以通过辅以数据关联等算法来拓展到多目标跟踪中来。

2、根据被跟踪目标的属性可以将跟踪算法分为刚性目标跟踪和非刚性目标跟踪。刚性目标在跟踪过程中形体状态不发生改变,可以通过简单的集合图形

进行模拟；非刚性目标的形体在跟踪过程中会发生改变，需要提取其不变性特征或者将其分解为若干子块后关联建模。

3、根据目标背景是否变化可以将跟踪算法分为静止背景下的跟踪算法和运动背景下的跟踪算法。在背景不变或者很少变化的情况下，可以通过对背景建模后采用背景减除和数据关联的方法对目标进行检测与跟踪；在动态背景下，背景建模变得难以实现，因此需要事先训练得到的检测器来检测特定目标，然后对其进行跟踪。

4、根据跟踪速度的要求可以将跟踪算法分为实时跟踪和非实时跟踪。实时跟踪对跟踪算法的速度要求很高，在监控等领域有着广泛的应用，需要系统做出快速反应；非实时跟踪大多对算法的精度要求很好，如在体育视频编辑中需要准确的得到运动员的运动轨迹。

5、根据传感器的类型可以将跟踪算法分为可见光图像跟踪和可见光谱以外的图像跟踪。非可见光可以为目标跟踪提供更多的信息，因此在多与前者结合以增强跟踪的准确性与鲁棒性，如红外图像跟踪^[8]等。

现在大部分的算法是基于动态背景下单目标或者多目标、刚体与非刚体一体的研究，同时有些又结合了红外光等信息进行跟踪。这些跟踪算法中，部分算法是针对缩小目标搜索范围提出的，这类算法主要是通过滤波方法预测目标下一时刻可能出现的位置，从而缩小目标的搜索范围来缩短目标搜索时间以满足系统实时性的要求；部分是针对提高跟踪匹配的准确性而提出的，这类算法主要是通过选取稳定的、可区分性好的特征，并可以与红外光、雷达等信息融合，从而提高目标的搜索速度和匹配的正确度。

1.3.2 视频跟踪研究难点

在视频运动目标跟踪中，我们通过使用摄像头 CCD 捕捉视频图像序列，将现实目标数字化，接着对跟踪目标进行初始化，初始化时可以采用目标检测的方法检测目标的初始位置，或者通过手动标注标明我们需要跟踪的目标，最后才开始我们真正的跟踪过程。在目标跟踪的采集、标注、匹配搜索等整个过程中，主要有以下难点。

1、采集设备的成像影响。在采集设备 CCD 采集视频图像的过程中，三维的现实目标被转化为二维图像目标，损失了深度信息，给目标跟踪带来很大困

难。这个问题可以采用双目摄像机或者加上激光设备获取图像的深度信息等加以解决。

2、目标的颜色、形状和尺度变化。目标在运动过程中存在的旋转、偏移以及采集设备的焦距变化都可能会引起视频中目标的颜色、形状和尺度的变化。提取更为全面综合的目标特征可以适应目标运动中的变化，较好地解决这个问题。

3、目标的部分或全部遮挡。目标运动过程中出现的被背景或者其他目标遮挡会导致目标在图像中的表达发生剧烈变化，很容易导致跟踪失败。对目标的不同部分赋以不同的特征提取权值，采用稀疏表示的方法来实时地表示目标，并且对目标的运动轨迹进行预测，可以较好地解决这个问题。

4、复杂环境下的目标跟踪。在场景光线发生变化或者目标与背景相似度较高的情况下，目标跟踪变得较为困难。根据目标与背景的图像特征，选取区分力较强的特征对目标进行表达可以有效地区分目标与背景。

5、初始化对跟踪带来的影响。在跟踪之前，目标需要通过检测或者手动标注的方法来初始化，因此，目标跟踪的结果与目标检测算法或者手动标注的准确性密切相关。选取更为精确的目标检测算法可以保证目标跟踪的有效性。

6、目标跟踪算法的实时性。现实需求往往对目标跟踪的效率有较高要求，以方便为下一步分析提供基础，如果目标跟踪的效率过低可能会不能适应现实需求。采用更为快速、准确的跟踪算法可以满足目标跟踪的实时性要求。

1.4 本文的研究内容

本文主要改进了基于综合颜色直方图和梯度方向直方图的目标跟踪特征，同时提出了对目标特征集合在线的进行稀疏表示，使目标及背景发生变化时对目标能够进行紧致有效表示；然后在 Mean-shift 跟踪框架下，基于稀疏表示对权值图像的获取进行改进，并对目标尺寸进行自适应更新，使得目标跟踪过程在复杂环境中仍然具有较好的准确性、实时性和鲁棒性。本文的主要研究内容如下：

1、改进了目标综合特征的提取方法。在目标特征提取过程中，结合目标的局部颜色与局部轮廓信息，得到改进的综合直方图 HOGC (Histogram of

Oriented Gradient and Color), 适应目标跟踪的需要。

2、提出了稀疏自适应的综合特征表示。基于改进的综合直方图 HOGC, 在跟踪的初始化过程中建立在线训练样本集合, 用以对跟踪特征的实时稀疏选择, 从而得到跟踪目标具有前景-背景区分性的稀疏特征表示。同时, 在跟踪过程中对训练样本进行实时更新, 在保证特征描述具有区分性的基础上, 进一步保证特征的环境自适应性与连续性, 从而能够更好的解决前景-背景在跟踪过程中的变化对跟踪效果的影响。

3、稀疏表示与 Mean-shift 框架相结合的快速目标跟踪方法。由于目标匹配搜索过程中迭代递归方法比穷举搜索的效率要高很多, 因此采用 Mean-shift 方法对目标进行搜索。同时, 将该算法存在的目标表达较为单一、缺乏模板更新以及尺寸固定等缺点加以解决与改进, 使得目标跟踪更为高效。

1.5 本文的组织结构

第一章, 绪论。主要论述了视频跟踪算法的研究背景和意义, 分析了国内外的研究现状以及目标跟踪算法中存在的难题, 列出了本文的主要研究目的和研究内容。

第二章, 目标跟踪算法基础。主要综述了目标跟踪中的主要研究内容和方法, 对比了各个方法的特点与应用范围; 同时介绍了基于稀疏表示的目标跟踪算法基础。

第三章, 基于综合特征稀疏表示的目标跟踪算法。主要论述了稀疏表示在处理信号、图像等领域的应用及求解方法。详细介绍了基于稀疏表示的目标描述, 并结合局部穷举目标搜索算法给出了跟踪算法在复杂场景下的运动目标跟踪实验, 同时将该算法与其他经典算法进行了比较并对比较结果进行了分析。

第四章, 基于稀疏表示与 Mean-shift 的快速目标跟踪算法。基于 Mean-shift 跟踪框架, 结合目标的稀疏特征表达, 提出了基于稀疏表示以及优化匹配搜索的目标跟踪算法, 并针对性地对传统 Mean-shift 算法进行了改进。最后, 对跟踪算法在运动目标尺度伸缩变化、部分及完全遮挡和部分形变等情况下进行实验对比, 同时将该算法与传统的跟踪算法的性能进行了定量比较与分析, 验证了算法的有效性和高效性。

最后是对现有工作的总结和对未来工作的展望，以及对如何进一步提高跟踪算法的鲁棒性、准确性和实时性等研究热点和难点的探讨。

第二章 目标跟踪算法基础

作为计算机视觉领域的热点和核心问题之一，近年来研究人员对目标跟踪进行了大量的研究，并且取得了一定成果。本章将对目标的表示进行介绍与讨论，并主要介绍目标跟踪的一些传统算法。在接下来的第三章和第四章，我们将详细介绍基于综合特征稀疏表示和 Mean-shift 框架下的快速目标跟踪方法。

2.1 目标表示模型

目标跟踪的第一步就是要为目标建立模型进行表示。根据不同的应用需要，目标可以采用中心点表示，几何图形表示等，同时对于表示目标的模板需要提取需要表达的特征来区分目标与背景，为目标的预测匹配提供良好的基础。

2.1.1 目标表示

目标表示模型主要可以分为点表示和区域图形表示。点表示又可以分为中心点表示和多点表示；区域图形表示主要有：矩形框表示、椭圆表示、分块图形表示以及轮廓表示等。具体可见图 2-1。

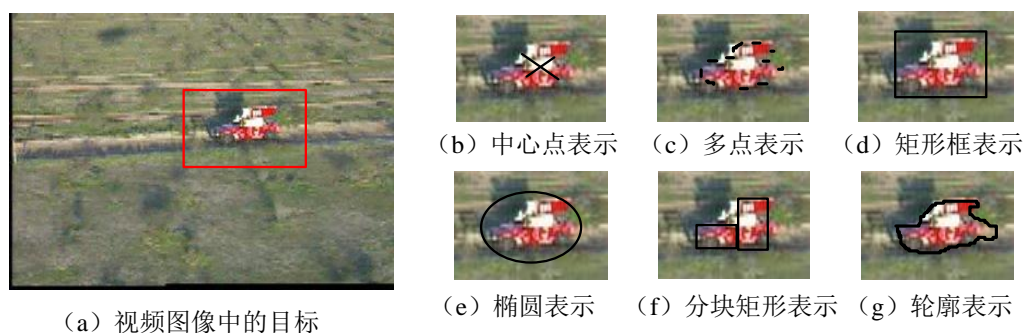


图 2-1 目标表示模型

点表示分为中心点表示和多点表示。当目标区域较小并且目标为刚体时，点表示可以被用来建立目标模型，但是当目标在运动过程中存在形变或者颜色等特征变化时，点表示将不适用于进行目标跟踪。点表示在滤波跟踪中有着广泛地应用。

在大多数的目标跟踪过程中，目标很难用一个点来表示，所以目标区域常

被用一个或多个几何图形来表示。当目标为非刚体时，我们可以用分块表示或者轮廓表示的方法来对目标进行表示，这些都得到了广泛地应用。

目标表示模型与接下来的目标匹配搜索算法密切相关，目标模型的选取也往往由目标本身的特点和搜索算法的特点所决定。

例如，点模型常常用于表示图像中较小的目标：Veenman 等人用它来表示碟子中的种子，Shafique 和 Shah 用它来表示较远距离的鸟；图形模型常常用于表示图像中类似形状的目标：Han 等人用矩形框来表示视频序列中的汽车，S. Avidan 用矩形框来表示视频序列中的人脸；对于具有复杂形状及形变的目标，分块表示可以被用来描述目标。

2.1.2 目标特征提取

目标描述和特征的选择密切相关。选择适当的特征在目标跟踪中具有重要的作用。通常，好的特征应该具有唯一性，因此可以很容易地将目标从特征空间中区分出来。例如对于基于轮廓的目标表示，边缘通常被用来作为相应的特征。目标跟踪过程中常用的特征如下^[9]：

1、颜色 目标的颜色主要由两个物理因素决定，一个是光源的功率谱分布，一个是目标的表面反射性质。在图像处理领域，RGB 空间是最常用的颜色空间，但是 RGB 在颜色感知上是不均匀的。 $L^*u^*v^*$ 和 $L^*a^*b^*$ 是感知均匀的颜色空间，而 HSV 是近似均匀的颜色空间，但是这些空间对噪声很敏感。

2、边缘 与颜色特征相比，边缘的一个重要性质是对光照变化不敏感。目前最为流行的边缘检测方法是 Canny 算子。

3、光流 光流定义了像素的密集位移向量场，其通过假设相邻图像帧的对应像素之间的亮度恒定不变而计算得到。光流常被用来表示基于运动的分割与跟踪的特征。

4、纹理 纹理是局部区域亮度变化性质（如，平滑性、规则性等）的一种描述。和颜色相比，纹理需要一个预处理过程来产生纹理描述子。已有的纹理描述子有很多种，如灰度共生矩阵、Laws 纹理描述子、小波、易控的金字塔等。与边缘特征相似，纹理特征对于光照的变化也不敏感。

在各种特征中，颜色在跟踪领域是使用得最为广泛的特征。但是，颜色对于光照的变化很敏感，因此，如果应用环境下光照变化较大，应该采用其他特

征（如，边缘、纹理等）进行跟踪。同时，基于各种特征的组合特征来更加准确完备的描述目标还有待进一步的研究。

2.2 目标匹配搜索

近年来，科研人员在改进传统的跟踪方法^[10-32]的同时，也开发研究了大量的新算法^[33-41]。目标统计特征^[10-13]常常被用来描述目标；背景和前景的差值及其相应的改进^[14-16]被用来搜索和匹配目标；其他传感器和双目摄像头可以有效获得目标的深度信息^[17-19]；Huang 等人^[21-25]将核函数纳入目标跟踪中来；卡尔曼滤波器^[26-27]和粒子滤波^[28-32]在目标跟踪也得到了广泛的应用；Han 等人^{[33-34][38]}将跟踪作为分类问题进行研究；Avidan 等人^[35-37]发展了支持向量机在目标跟踪中的应用；Kabaoglu 等人^[38-41]进一步发展了支持向量机与粒子滤波进行目标跟踪的应用。这些算法可以分为两大类：确定性方法和随机方法。在确定性跟踪算法中，目标的外观、颜色分布、轮廓信息等先验知识用于建立目标模板，其中基于表观的方法可以统计估计目标在特征空间或核空间内的分布规律^[42-45]。确定性方法通常将跟踪问题简化为一个最优化问题，如何定义好这个最优化函数是确定性方法的关键。随机跟踪方法的基本思想是，已知当前时刻以及之前时刻的所有参数度量，对目标当前状态进行估计。它利用状态空间对当前跟踪系统的运动进行模型化，该方法引入了概率统计的思想，将不确定性观察（如概率密度函数）与不同的状态相联系，从而不再假设运动系统输入与输出的完全确定性。

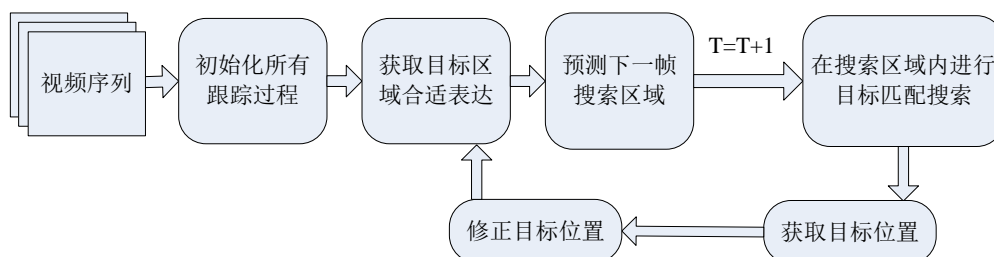


图 2-2 目标跟踪示意图

在确定性跟踪方法中，误差平方和（Sum of Squared Difference, SSD）方法是最广泛使用的与目标模板匹配相似性度量函数。SSD 方法虽然能够很好的反映目标的细节，但是当目标的外观发生较大变化时，SSD 方法会失败；同时由

于 SSD 方法计算量较大, 较难达到实时性跟踪的要求, 因此 SSD 方法被不断改进以适应实时跟踪系统的要求。基于光流场的目标跟踪是一种典型的确定性跟踪方法。对目标进行跟踪时, 根据预先设定好的相似性度量函数, 在当前视频帧的某个区域内, 通过搜索或者迭代算法找到与目标模板或者目标表观最为相似的区域。Mean-shift 目标跟踪方法, 以先前目标位置为初始位置, 利用无参数概率密度梯度, 搜索与目标模板颜色直方图最类似的窗口作为当前目标位置。Mean-shift 最早是由 Fukunaga 等人于 1975 年提出来的, 近年来被广泛用于寻找局部最优区域^{[46][47]}。Mean-shift 算法的形式简洁, 收敛条件宽松, 只要确定了跟踪目标, 整个跟踪过程就不需额外的参数输入。由于无需对概率密度分布进行估计, 就可以直接沿着梯度方向搜索局部最大值, 大大减少了运算量, 所以这种算法具有良好的实时性。Mean-shift 算法用于视觉跟踪有以下几个优势: 首先, 算法计算量不大, 在目标区域已知的情况下完全可以做到实时跟踪; 其次, 作为一个无参数概率密度估计算法, 很容易和其他算法结合使用; 然后, 采用加权直方图模型, 对部分遮挡、目标旋转、形变和背景运动不敏感。但是它缺少模板更新算法, 并且对目标的尺度变化比较敏感, 同时多采用颜色统计直方图对目标特征进行描述带来局限性, 因此近年来科研人员对初始的方法进行了大量改进以及与其他方法相结合, 取得良好的目标跟踪效果, 如 Collins 等人通过在线选择的方法来选取最具区分力的颜色特征^[48], Bouttefroy 等人使用投影 Kalman 滤波的 Mean-shift, Avidan 将分类器与 Mean-Shift 跟踪框架相结合^[49]等, 这些都改进了传统 Mean-shift 的不足。基于目标与背景进行分类的目标跟踪方法避开了目标状态估计的复杂过程, 通过分类达到很高的跟踪精度。然而这种方法需要大量的正例反例样本的学习构造分类器, 因为跟踪环境的不同样本的选择具有特殊性, 这个成为了使用的一个难点, 同时需要设计分类器和对样本进行训练, 较难满足跟踪实时的需要。

随机跟踪方法可以很好地弥补确定性跟踪方法中存在以及较难解决的如鲁棒性较差、出现遮挡后丢失等问题, 因此被广泛应用并多以一种与确定性跟踪方法相结合的方式出现。它采用预测滤波方法, 模型化跟踪目标的运动状态, 估计目标运动的状态方程。当运动模型为线性或高斯分布并且观察模型也为线性时, 采用 Kalman 滤波方法可以获得最优。当状态空间为非线性、非高斯时,

目前没有找到能够准确解析描述的模型,经常采用马尔科夫滤波进行目标跟踪,其中最常用的是时序蒙特卡罗滤波^{[50][51]}的方法,即粒子滤波。通过使用状态空间的加权随机采样表示目标状态的概率分布,计算最大后验概率获得最优目标状态,它是目前应用最为广泛的算法^[52]。随机跟踪方法不再假设跟踪系统输入与输出的对应确定性,融合了概率统计的思想,把如概率密度函数的不确定性观察和不同状态进行联系。由于在跟踪过程中,目标以及环境相互之间存在遮挡、不断变化等很多不确定性的因素,通过使用随机性的跟踪方法达到了更加鲁棒的效果,因此随机跟踪方法更适合于大多数的视觉目标跟踪。

本节主要介绍常用的目标匹配搜索算法:基于相似性度量的目标匹配算法以及基于均值漂移的目标搜索算法。两者是确定性的目标匹配跟踪方法,需要对目标本身进行建模,并找出具备区分力的目标特征进行目标表达。

2.2.1 基于目标表达和相似性度量的目标匹配算法

根据跟踪目标的表达和相似性度量,目标跟踪算法可以分为:基于区域统计特征的跟踪算法、基于模型匹配的跟踪算法、基于轮廓的跟踪算法、基于特征的跟踪算法。

基于区域统计特征的跟踪方法是把图像中运动目标连通区域的共有特征信息作为跟踪特征的一种方法。在图像中常用的共有特征信息如颜色特征、纹理特征等。这种方法不需要在视频序列中找到完全相同的特征信息,而是通过计算候选区域与原始目标之间的区域统计特征的相关性来确定跟踪目标的最佳位置。例如,利用小区域特征进行室内单目标人体的跟踪。该方法将人体看作由头、躯干、四肢等部分所对应的小区域块的联合体,通过分别跟踪各个小区域块最终完成对整个人体的跟踪。同时,基于区域统计特征的跟踪方法常常利用滤波技术来估计和预测区域的几何形状和运动速度。

基于模型匹配的跟踪方法通过建立模型来表示目标,也就是对运动目标进行建模,然后在图像序列中匹配这个建立好的模型来实现目标的跟踪。然而,在实际应用中跟踪的目标大部分都是非刚体的,其形状在不断发生变化。即使是刚体目标,由于拍摄视角及拍摄距离的不断变化也会导致其形态不停地发生变化,很难得到准确的几何模型。因此,一些学者提出通过变形轮廓模板来进行目标跟踪。目前有两种较为流行的可变形模型:自由式的可变形模型和参数

可变的模型。前者主要通过满足一些简单的约束条件(连续性、平滑性等)来跟踪任意形状的运动目标;后者通过使用参数公式或变形公式来描述目标的形状。

基于轮廓的跟踪算法通过对目标的分割和边缘提取实现对运动目标的跟踪。然而,当目标被部分及全部遮挡时,跟踪算法将受到很大的限制。基于模型的跟踪算法采用目标的物理知识而不仅仅是目标的图像信息建立模型,因此基于 3D 模型跟踪算法是比较完善的。该方法的局限在于无先验知识的实际跟踪系统中很难得到目标的模型。基于特征表达的跟踪算法通过提取具有尺度伸缩、旋转等不变性的特征点(如 SIFT、KLT、Harris Corners、SUSAN Corners 等),解决尺度伸缩和部分遮挡等问题。SIFT 特征点主要是通过图像金字塔的多尺度极值点检测得到的,用梯度方向进行描述,因此这种特征点具有尺度和旋转不变性等优点,在图像配准和跟踪方面具有很高的匹配精度和鲁棒性。然而,特征自身的检测和提取往往计算复杂难以满足跟踪系统实时处理的要求。基于区域统计特性的跟踪算法利用的是目标图像的全局描述信息,很难精确定位目标的位置。同时,当背景或者其他目标的区域统计特性相似的时候,跟踪算法也会失效(如基于颜色或者灰度直方图的跟踪算法)。

基于特征匹配的跟踪方法不考虑运动目标的整体特征,即不关心具体的运动目标,而只通过其特征来进行跟踪。由于图像采样时间间隔通常很小,可以认为这些特征在运动形式上是平滑的,因此可以利用其完成目标的整个跟踪过程。无论是刚体运动目标还是非刚体运动目标,利用基于特征匹配的方法进行目标跟踪时主要包括特征提取和特征匹配两个过程。在特征提取中要选择适当的跟踪特征,并且在后续的视频帧提取相应的特征;在特征匹配中将提取的当前帧图像中目标的特征与特征模板相比较,根据比较的结果来确定目标,从而实现目标的跟踪。例如使用特征点对人体进行运动跟踪时,首先把需要跟踪的每一个人用一个矩形框封闭起来,在封闭框内进行跟踪特征的提取并建立特征模板,然后在后续的视频帧中提取相应的特征并与特征模板进行匹配,从而通过找到最佳的匹配位置完成跟踪。除了用单一的特征来实现跟踪外,还可以将多个特征信息进行综合作为跟踪特征,来提高跟踪的稳定性。另外,如果运动目标简单,可以将整个目标作为特征模板来进行跟踪,这种方法也被称为模板匹配。

在目标跟踪中，特征搜索匹配就是在所得到的目标和跟踪目标之间进行特征相似度度量。一般使用向量的形式来表示特征，相似度度量就是计算两个向量之间的距离。在图像处理技术中常用的相似性度量方法有欧氏距离、街区距离、棋盘距离、加权距离、巴特查理亚系数、Hausdorff 距离等。其中，Hausdorff 距离通常用于衡量特征集之间的相似性；直方图之间的相似性通常采用巴特查理亚系数；欧氏距离由于其简便性，是使用最广泛的相似度度量。

2.2.2 基于均值漂移的目标快速搜索算法

Mean-shift 最早是由 Fukunaga 等人于 1975 年提出来的，其最初含义是偏移的均值向量。但随着 Mean-shift 理论的发展，Mean-shift 的含义也逐步地发生着变化，目前所说的 Mean-shift 算法一般是指一个迭代的过程，即先算出当前点的偏移均值，移动该点到其偏移均值，然后以此点为新的起始继续移动，直到满足一定的条件结束。

1、基本 Mean-shift

给定 d 维空间 R^d 中的 n 个样本点 $x_i, i=1, \dots, n$ ，在 x 点的 Mean-shift 向量的基本形式定义为：

$$M_h(x) \equiv \frac{1}{k} \sum_{x_i \in S_h} (x_i - x) \quad (2-13)$$

其中， S_h 是一个半径为 h 的高维球形区域，满足以下关系的 y 点的集合，

$$S_h(x) \equiv \{y : (y-x)^T (y-x) \leq h^2\} \quad (2-14)$$

其中， k 表示在这 n 个样本点 x_i 中，有 k 个点落入区域 S_h 中。

我们可以看到 $(x_i - x)$ 是样本点 x_i 相对于点 x 的偏移向量，公式 2-13 定义的 Mean-shift 向量 $M_h(x)$ 就是对落入区域 S_h 中的 k 个样本点相对于点 x 的偏移向量求和然后再平均。直观上讲，如果样本点 x_i 从一个概率密度函数中采样得到，由于非零的概率密度梯度指向概率密度增加最大的方向，因此从平均上来说， S_h 区域内的样本点更多的落在沿着概率密度梯度的方向。因此，对应的 Mean-shift 向量 $M_h(x)$ 应该指向概率密度梯度的方向。

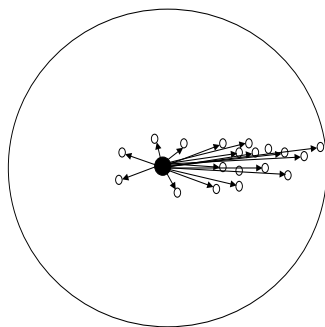


图 2-3 Mean-shift 示意图

如图 2-3 所示，大圆圈所圈定的范围就是 S_h ，小圆圈代表落入 S_h 区域内的样本点 $x_i \in S_h$ ，黑点就是 Mean-shift 的基准点 x ，箭头表示样本点相对于基准点 x 的偏移向量，因此我们可以得到，平均的偏移向量 $M_h(x)$ 会指向样本分布最多的区域，也就是概率密度函数的梯度方向。

2、扩展 Mean-shift

从公式 2-13 我们可以看出，只要落入 S_h 区域内的采样点，无论其离 x 远近，对最终的 $M_h(x)$ 计算的贡献是一样的。然而，一般来说离 x 越近的采样点对估计 x 周围的统计特性越有效，因此我们引进核函数的概念，即在计算 $M_h(x)$ 时可以考虑距离的影响，认为在所有的样本点 x_i 中重要性并不一样，因此我们对每个样本都引入一个权重系数。

因此，基本的 Mean-shift 形式扩展如下：

$$M(x) \equiv \frac{\sum_{i=1}^n G_H(x_i - x) w(x_i) (x_i - x)}{\sum_{i=1}^n G_H(x_i - x) w(x_i)} \quad (2-15)$$

其中： $G_H(x_i - x) = |H|^{-1/2} G(H^{-1/2}(x_i - x))$ ； $G(x)$ 是一个单位核函数； H 是一个正定的对称 $d \times d$ 矩阵，我们一般称之为带宽矩阵； $w(x_i) \geq 0$ 是一个赋给采样点 x_i 的权重。

在实际应用的过程中，带宽矩阵 H 一般被限定为一个对角矩阵 $H = \text{diag}[h_1^2, \dots, h_d^2]$ ，甚至可以更简单的被定义为正比于单位矩阵，即 $H = h^2 I$ 。

由于后一形式只需要确定一个系数 h ，因此常常被采用，因此公式 2-22 又可以被改写为：

$$M_h(x) \equiv \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i) (x_i - x)}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)} \quad (2-16)$$

可以看到，如果对所有的采样点 x_i 满足：

- (1) $w(x_i) = 1$;
- (2) $G(x) = \begin{cases} 1 & \text{if } \|x\| < 1 \\ 0 & \text{if } \|x\| \geq 1 \end{cases}$ 。

则公式 2-16 完全退化为公式 2-13，也就是说扩展的 Mean-shift 在某些情况下会退化为最基本的 Mean-shift。

2.3 基于稀疏表示的目标跟踪基础

稀疏表示起源于压缩感知理论，即当信号集合中大部分元素为 0 时，可以将信号投影到低维空间，并可以从投影信号中完全地恢复初始信号。在图像处理和模式识别领域，稀疏表示在特征评估和特征选择中的应用也越来越受到研究人员的重视，并且 ICCV'09 与 CVPR'10 中分别开设了一系列的关于稀疏表示专题报告。

Mei 等人^[55]将跟踪看作为粒子滤波框架下的稀疏估计，在搜索区域的候选目标被目标模板和小模板稀疏表示之后，他们选取其中最小投射错误的候选项作为跟踪的结果。已知目标模板 $T = [t_1, \dots, t_n]$ ，则候选目标项 y 可以用式 2-17 表示。

$$y \approx Ta = a_1 t_1 + a_2 t_2 + \dots + a_n t_n \quad (2-17)$$

其中 $a = [a_1, \dots, a_n]$ 是目标模板系数向量。加入误差量 e 后，式 2-17 可改为

$$y \approx [T, I, -I] \begin{bmatrix} a \\ e^+ \\ e^- \end{bmatrix} \square Bc, \text{ s.t. } c \geq 0 \quad (2-18)$$

其中 e 为误差量, I 为误差量对应的小系数模板, B 的定义为 $B = [T, I, -I]$, c 的定义为 $c^T = [a, e^+, e^-]$ 。这样, 当目标出现遮挡或者形变等情况时, 式 2-18 可以通过调整 e 和 I 来适应变化, 目标的跟踪问题也就转化为下面的最优化问题。

$$\min \|Bc - y\|_2^2 + \lambda \|c\|_1 \quad (2-19)$$

其中 $\|\bullet\|_1$ 和 $\|\bullet\|_2$ 分别代表 1 范式和 2 范式。

Han 等人^[33]将 L1-norm 最小化的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。

在他们研究的基础上, 我们采用稀疏表示进行目标特征稀疏选择, 评估和选取更为有效的特征进行目标跟踪。在跟踪过程中, 我们可以将目标的跟踪看作为目标与背景的分类问题, 因此, 在目标出现形变或者被遮挡的情况下, 选取不同的特征会带来不同的结果, 或者对不同的特征赋予不同的权值也会影响不同的结果。由于颜色信息对目标的表达往往是不完整的, 我们选取目标的颜色和梯度方向信息对目标进行表达, 同时又将稀疏表示应用到目标的特征评估和特征选择中来, 对不同区分力的特征赋以不同的权值, 更好地将目标与背景进行区分, 也因此可以得到更为准确的跟踪结果, 尤其在目标与背景相似度高的情况下效果更好。

第三章 基于综合特征稀疏表示的目标跟踪算法

人类视觉系统（Human Vision System, HVS）的最新研究^[53]表明，一小部分外界环境的因素，如颜色、纹理、形状以及尺度变化等，可以对人类的视觉神经元细胞产生大部分的刺激，大部分的外界环境因素对视觉细胞的刺激是几乎可以忽略的。同样，由目标检测，如人脸检测^[54]等相关的实验中用到的目标稀疏表示以及相关的研究得到启发，我们可以推断，在目标跟踪中，对于目标和背景的分所产生目标特征是稀疏的，而稀疏的样本子集比样本全集带来的目标背景区分力可能会更大。

因此，本章提出了基于综合特征稀疏表示的目标跟踪算法。首先，在初始化过程中建立目标与背景的训练样本集合；其次，基于样本集合对跟踪目标进行实时的特征稀疏选择，并自适应地更新训练样本集合；最后，基于目标的综合特征稀疏表示，对目标进行跟踪。

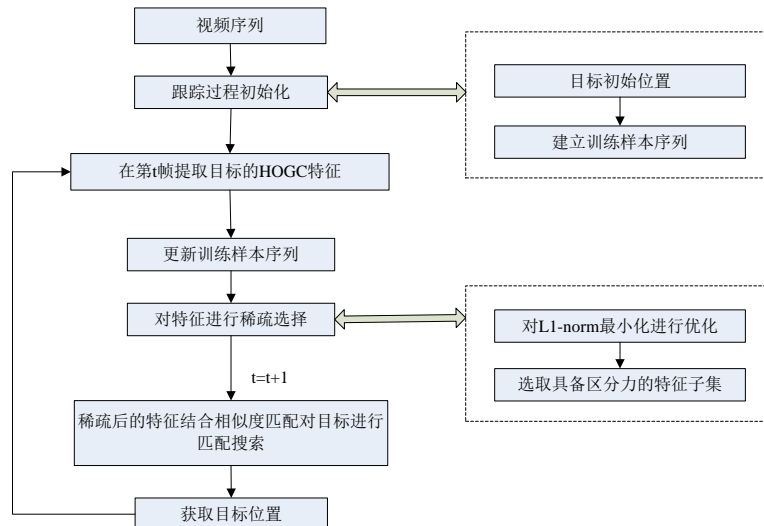


图 3-1 基于综合特征稀疏表示的目标跟踪示意图

3.1 稀疏表示的基本理论及其应用

3.1.1 稀疏表示的基本理论

稀疏表示（Sparse Representation）起源于由斯坦福大学的 David Donoho 和加州理工学院的 Emmanuel Candes 等人提出的压缩感知理论（Compressed

Sensing)^[56-68]，理论证明了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，是对奈奎斯特采样定理的突破。近年来，稀疏表示吸引了信号处理、模式识别以及图像处理等领域研究人员的关注，在很多相关的问题上得到了广泛的研究和应用^[69-72]。

广义上，稀疏是指信号集合，如向量集合或者矩阵集合的大部分元素为0，那么可以将变换系数线性投影为低维观测向量，同时又保持了信号重建时所需的信息，可以通过进一步求解稀疏优化问题就能够精确地或高概率精确地重建原始信号。由于过完备字典中的各信号元素的表示向量之间并不是线性无关的，而是不同程度上信号元素存在冗余，因此新的信号元素可以有多种由字典库中的信号元素构成的线性组合。稀疏表示就是找到一种由少数信号元素构成的线性组合，使得信号能量只集中于这些较少的元素上，从而揭示了信号的主要特征和内在结构。

上述的变换域中的过完备字典、满足非相干性的测量矩阵和快速鲁棒的信号重建等三个问题的关键在于信号是否可以进行稀疏表示，如果能进行稀疏表示，应该使用何种模型去表示，和模型的解法等一系列问题。在压缩感知的最初稀疏表示模型中 L0-norm 被用来解决这个问题，如式(3-1)所示。

$$\begin{aligned} \min_x \quad & \|x\|_0 \\ \text{s.t.} \quad & y = Ax \end{aligned} \quad (3-1)$$

上述的 L0-norm 优化模型中， $\|x\|_0$ 指向量 x 中不为0的元素的个数， 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 是接受端对信号进行重构之后的结果。于是，信号的稀疏表示问题最终转化为和式的最优化问题。Donoho 和 Elad^[73]等学者在 2003 年证明，如果矩阵 A 满足： $\sigma(A) \geq 2\|x\|_0$ ，那么 L0-norm 最优化问题具有唯一解。然而，虽然证明出模型的解是存在唯一的，但是上述优化模型的求解过程是 NP 难问题。

随后，Tao 和 Candes 在 2006 年合作证明了在限制等距性质(Restricted Isometry Property, RIP)条件下，L0-norm 优化问题与模型(3-2)中 L1-norm 优化问题具有相同的解，其中 RIP 条件指稀疏信号在观测矩阵作用下必须保持几何性质相一致。

$$\begin{aligned} \min_x \quad & \|x\|_1 \\ \text{s.t.} \quad & y = Ax \end{aligned} \quad (3-2)$$

在 RIP 性质提出之后，稀疏表示终于在理论上和实践中得到统一。后来很多研究学者也注意到稀疏表示的优点，将其用于模式识别和计算机视觉等领域中。

目前稀疏表示的实现主要有三种方式：直接优化法，迭代收缩方法和 L_p 范数正则化方法。

1、直接优化法：这种方法是指直接使用 0-范数最小化作为稀疏性追求的目标函数，优化模型与公式(3-1)的模型类似。这类优化问题的求解一般都使用贪婪算法去求解，典型的代表算法有匹配追踪法和正交匹配追踪法等。

2、迭代收缩优化法：这种方法是著名 LASSO 方法。该方法是针对回归问题，建立最小二乘误差和 1 范数正则化的约束模型，并提出使用贪婪算法去求解。它可以通过不断松弛模型(3-3)中的 ε 因子，来实现迭代收缩求解。

$$\begin{aligned} \min_x \quad & \|Ax - y\|_2 \\ \text{s.t.} \quad & \|x\|_1 \leq \varepsilon \end{aligned} \quad (3-3)$$

3、范数正则化优化法：这个方法目前研究比较热门，常见的有基追踪方法，软阈值方法(如梯度投影稀疏重构方法和 Bregman 迭代法)。

3.1.2 稀疏表示在图像处理中的主要应用

近年来，随着稀疏表示理论研究的不断突破和发展，稀疏表示越来越受到信号处理、图像处理、模式识别等领域研究人员的关注，在很多相关的问题上得到了广泛的研究和应用。

2008 年，Julien Mairal 等提出了基于稀疏表示的图像重建，他们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。John Wright 和 Yi Ma^[54]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 L1-norm 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。在人脸有部分遮挡的时候，该算法有很大优势，可以自动找出被遮挡的部分，并且对该部分进行重建。Jianchao Yang^[74]等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，然后根据高低分辨率之间的对应性，从而他们实现了超分辨率重建。而且，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。Zhenjun Han^[34]将线性稀疏重构应用到多目标跟踪中，很大程度上解决了目标遮挡情况下的目标跟踪。在图像处理的其他很多方面，如图像融合、图像目标检测、图像恢复等领域也不断发展，提出了令人瞩目的基于稀疏表示的算法。

3.2 基于综合特征稀疏表示的目标跟踪算法

3.2.1 改进 HOGC 特征的选取



图 3-2 视频帧中的运动目标

由于点模型不适应于大部分的目标跟踪，而目标轮廓模型需要对目标进行分割，因此我们采用矩形框对目标进行标示，将矩形框内的目标特征提取后近似认为是目标本身的特征，如图 3-2 所示。同样地，颜色直方图（Histogram of Color, HC）在过去的几年中被大量的应用于目标跟踪，并且取得了不错的效果。但是大量的实验表明，当目标和背景颜色比较相似的时候，基于颜色的跟踪算法往往不能取得很好的效果；而目标检测中被经常采用的梯度方向直方图（Histogram of Oriented Gradient, HOG）提取目标的边缘和轮廓特征，并且对目标的颜色变化不敏感。虽然 HOG 特征相比 HC 计算量大，并且在目标出现大面积平滑区域时性能较差，但是 Dalal 等人已经通过实验验证了固定尺度的 HOG 和 SIFT 对目标的描述能力几乎一样强。因此我们将 HC 和 HOG 两种特征融合在一起，称之为综合直方图（Histograms of Orientated Gradient and Color, HOGC）。

我们将颜色空间中每个颜色分量（R、G 和 B）均匀量化为一个 16 维的颜

色直方图，三个颜色分量总计形成了一个 48 维的颜色直方图。同样，将梯度向量量化为一个 8 维的梯度向量直方图。当给定特定的图像大小和特征尺度，梯度直方图可以很好地表达出跟踪目标的局部轮廓。此特征因为只保留方向特征，不保留强度特征，因此对于亮度、光照等变化具有一定的适应性。它与颜色直方图可以形成目标表示上的互补性。

HOGC 特征提取的详细过程：将跟踪目标转化为灰度图像；将跟踪目标窗口归一化到固定大小的图像窗口：32×32 像素；然后把图像划分为 16 个 8×8 的小区域（cell），并且每个 2×2 的相邻小区域组成一个图像块（block，标号为从①到⑨），因此每个相邻的图像块有 2 个小区域彼此重合；本文将每个图像块的梯度方向统计投影到一个 8 维的梯度方向直方图中，同时计算小图像块中的 48 维颜色直方图，将颜色直方图与梯度直方图组合到一起，这样每个图像块中的特征为 56 维（图 3-3c 所示）。从图 3-3a 中可以看到，由于每个图像窗口包含 9 个图像块，因此每个目标得到 504 维的梯度方向直方图来描述其局部轮廓。每个像素的梯度方向计算如下（如图 3-3b 所示）：

$$\begin{aligned}
 I &= G(0, \sigma) * I_0 \\
 dy &= I(h+1, w) - I(h-1, w) \\
 dx &= I(h, w+1) - I(h, w-1) \\
 ori(h, w) &= a \tan 2(dy, dx) \quad ori \in [0, 2\pi]
 \end{aligned}
 \tag{3-4}$$

其中，上式是计算一个像素在特定尺度下一个像素点方向的过程。 $G(0, \sigma)$ 表示一个高斯函数， I_0 表示原图像， I 表示经过平滑滤波后的图像。 $ori(h, w)$ 表示像素梯度方向（目标轮廓方向）。对目标区域/背景区域的所有像素进行累计，可以获得 504（9 个图像块，每个块 8 个方向向量，48 个颜色向量）维 HOGC 特征。

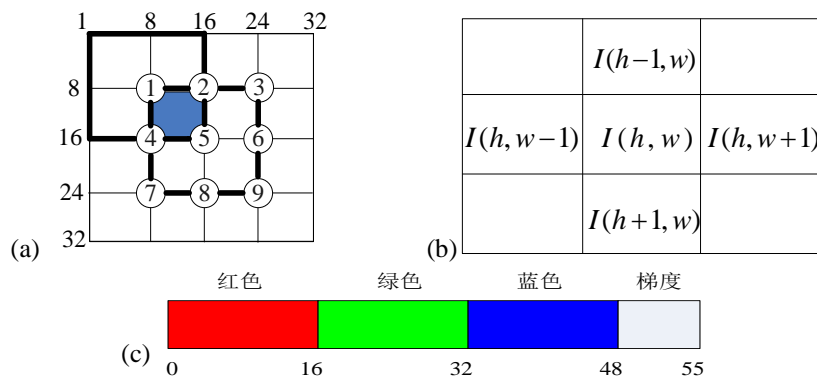


图 3-3 HOGC 特征提取图 (a) 目标窗口划分 (b) 计算梯度向量模板 (c) 图像特征向量

3.2.2 目标特征的稀疏选择

在目标跟踪的初始化过程中, 利用前 M 帧的跟踪结果及其背景构分别构造样本集合中的正例 ℓ_j^+ 与反例 ℓ_j^- , 获得构建 $2M$ 个训练样本 $training_set = \{(\ell_j^+, \ell_j^-)\}, j=1\dots M$, 提取了 HOGC 特征; 同样, 在跟踪过程中, 采用类似堆的先进先出形式对样本集中的样本进行实时替换, 即将最新的跟踪正例和反例样本替换样本集中样本。这样, 由于每次只更新 $1/M$ 的样本序列, 即使是一次较差甚至是错误的样本更新过程对整个样本集合的影响也是非常小的, 这样就能有效避免了跟踪过程中经常产生的模板漂移问题, 保证了跟踪的稳定性。

基于在线训练样本集, 通过 L1-norm 最小化的方法对目标特征进行稀疏选择, 这样既保证了对特征稀疏选择, 又保证了选择的结果具有很好的前景/背景区分力。这个过程可以优化建模如下:

$$\begin{aligned} \min \quad & \|S\|_1 \\ \text{s.t.} \quad & y_i \cdot (w^T x_i) \geq \alpha \end{aligned} \quad (3-5)$$

其中, $\|\bullet\|_1$ 代表 L1-norm; 约束条件保证训练样本能够被正确分类; $S \in R^N$ 是特征的前景背景区分能力; α 保证了正例样本和反例样本之间的距离最少为 2α 。由于这个优化模型较难直接解决, 需要引入向量来有效地解决这个问题。定义 $\eta \in R^N$, $\mu \in R^N$ 且 $\eta \geq 0, \mu \geq 0$ 。同时令 $\eta^i = (S^i)_+, \mu^i = (-S^i)_+, i=1, 2, \dots, N$, i 表示特征集合的维数, $(\bullet)_+$ 被定义为取正运算符 $(S^i)_+ = \max\{0, S^i\}$ 。然后便得到 $S = \eta - \mu$ 以及 $\|S\|_1 = I_N^T \eta + I_N^T \mu$, $I_N = [1, 1, \dots, 1]^T$ 。公式 (3-5) 便可被重新写为下面的凸优化模型:

$$\begin{aligned} \min \quad & I_N^T \eta + I_N^T \mu \\ \text{s.t.} \quad & \begin{cases} (\eta - \mu)^T \ell_j^+ \geq \alpha \\ (\eta - \mu)^T \ell_j^- \leq -\alpha \\ \eta \geq 0 \\ \mu \geq 0 \end{cases} \end{aligned} \quad (3-6)$$

经过稀疏特征选择之后的特征如图 3-4 所示:

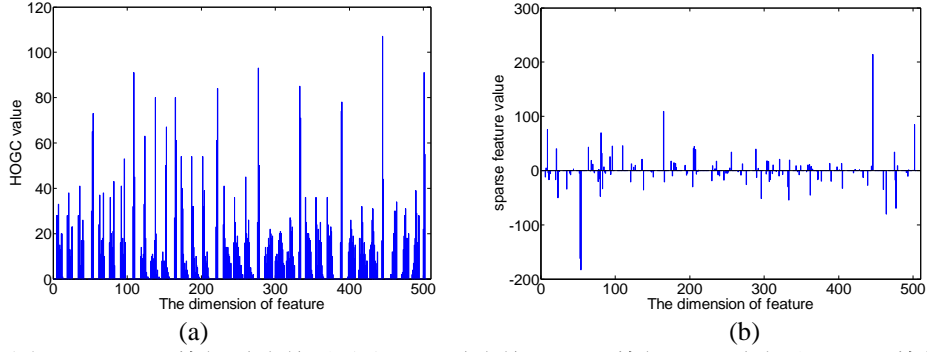


图 3-4 HOGC 特征稀疏前后对比 (a) 稀疏前 HOGC 特征 (b) 稀疏后 HOGC 特征

3.2.3 目标的匹配搜索

目标跟踪是对跟踪目标在搜索区域 Ω_t 内进行搜索的过程。相似性度量是在 Ω_t 内找到与跟踪目标最匹配的跟踪结果 $(x, y)_t$ 通过最大化 $p((x, y)_t | \Omega_t, (x, y)_{t-1})$ 。我们通过贝叶斯推理过程来求 $p((x, y)_t | \Omega_t, (x, y)_{t-1})$ 的最优解，如下所示：

$$\begin{aligned}
 & \max_{(x, y)_t} p((x, y)_t | \Omega_t, (x, y)_{t-1}) \\
 &= \max_{(x, y)_t} \int p(F_t((x, y)_t^c) | F_{t-1}^i, p(x, y) | \Omega_t) d\Omega_t \\
 &= \max_{(x, y)_t} \left[\sum_{i=1}^N (p(F_t^i | F_{t-1}^i, p(x, y) | \Omega_t)) \right] p(x, y) | \Omega_t d\Omega_t
 \end{aligned} \tag{3-7}$$

其中， $(x, y)_t^c$ 定义为 Ω_t 中第 c 个候选目标的位置， $p(F_t^i((x, y)_t^c), F_{t-1}^i)$ 表示为 $F_t^i((x, y)_t^c), F_{t-1}^i$ 这两个特征的相似度， $p(F_t^i)$ 表示为 t 时刻第 i 维特征的置信度，此处将其视为 t 时刻第 i 维特征的稀疏系数 $w_t(i)$ ，因此公式(3-7)可以如下描述：

$$\begin{aligned}
 & \max_{(x, y)_t} p((x, y)_t | \Omega_t, (x, y)_{t-1}) \\
 &= \max_{(x, y)_t} \int \left(\sum_{i=1}^N (w_t(i) p(F_t^i((x, y)_t^c), F_{t-1}^i)) \right) p((x, y)_t^c | \Omega_t) d\Omega_t
 \end{aligned} \tag{3-8}$$

特征搜索匹配就是在所得到的目标和跟踪目标之间进行特征相似度度量，而综合特征很包含了目标区域的颜色、边缘等信息，因此可以与目标匹配搜索算法很好地结合起来实现目标跟踪，如图 3-5 所示。

$$\min_{pos \in \Omega_t} (\| F(t_0) - F(C_t(pos)) \|_1) = \min_{pos \in \Omega_t} (\| \sum_{i=1}^n (F(t_0)_i - F(C_t(pos))_i) \|_1) \tag{3-9}$$

其中， $F(t_0)$ 是初始目标的 HOGC 特征， $C_t(pos)$ 是搜索区域 Ω_t 中在位置 pos 的候选项， $F(C_t(pos))$ 代表候选项 $C_t(pos)$ 的 HOGC 特征， F_i 是第 i 维的 HOGC 特

征。相似度量常就是计算两个向量之间的距离，是最基本的目标匹配搜索方法之一。在图像处理技术中，常用的相似性度量方法有欧氏距离、街区距离、棋盘距离、加权距离、巴特查理亚系数、Hausdorff 距离等。其中，欧氏距离由于其简便性，是使用最广泛的相似度量；Hausdorff 距离通常用于衡量特征集之间的相似性；直方图之间的相似性通常采用巴特查理亚系数。

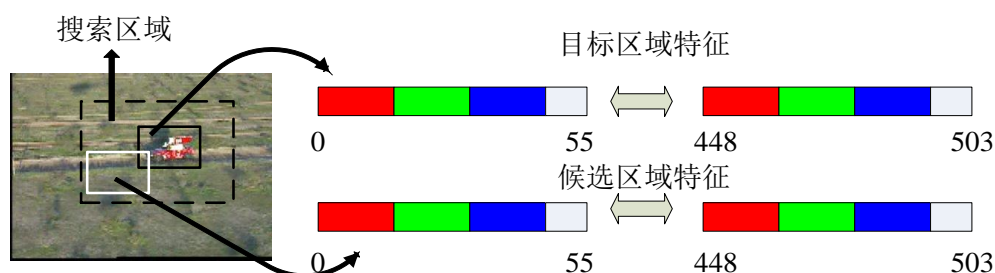


图 3-5 综合特征与模板匹配的结合

1) 欧氏距离。已知二维平面上的两个点 P_1, P_2 ，坐标分别为 $P(x_1, y_1), P(x_2, y_2)$ ，它们之间的欧式距离为坐标向量的平方差之和的开方，公式如下所示：

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3-10)$$

2) 加权距离。二维平面上的两个点 P_1, P_2 ，坐标分别为 $P(x_1, y_1), P(x_2, y_2)$ ，它们之间的加权距离为坐标向量的每一项的绝对距离的加权和，公式如下所示：

$$d = \sum_{i=1}^n w_i |x_i - x'_i| + \sum_{i=1}^n w_{i+1} |y_i - y'_i| \quad (3-11)$$

其中， $w_i = 1$ 。街区距离、棋盘距离是加权距离特殊情况，当 $w_i = 1, w_{i+1} = 1$ 时，加权距离就是街区距离，即坐标向量每一项的绝对距离之和；当 $w_i = 1, w_{i+1} = 1$ 时，加权距离就是棋盘距离，即坐标向量第一项的绝对距离；当 $w_i = 1, w_{i+1} = \sqrt{2}$ 时，加权距离就变成类欧式距离。

3) 巴特查理亚系数，简称 Bhattacharyya 系数或巴氏系数。已知两个向量点 P_1, P_2 ，坐标分别为 $P(x_1, y_1), P(x_2, y_2)$ ， P_1, P_2 之间的巴氏系数为坐标

向量的每一项的乘积的开方，几何意义是两个 m 维向量 \vec{a} 和 \vec{b} 夹角的余弦。公式如下所示：

$$d_{\vec{a}, \vec{b}} = \sqrt{\sum_{i=1}^m x_i y_i} \quad (3-12)$$

4) Hausdorff 距离 (Hausdorff Distance, HD)。HD 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 HD 距离定义如下：

$$d_{HD}(A, B) = \max \left\{ \max_{a_i \in A} \min_{b_j \in B} \|a_i - b_j\|, \max_{b_j \in B} \min_{a_i \in A} \|b_j - a_i\| \right\} \quad (3-13)$$

其中， $d_{HD}(A, B) = \max_{a_i \in A} \min_{b_j \in B} \|a_i - b_j\|$ ，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 HD，表示点集 A 中所有点到点集 B 的距离 $d_B(a_i)$ 的最大值。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，HD 使用不同的范式形式，灰度特征的 HD 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；彩色特征的 HD 使用 Bhattacharyya 系数等。

5) 改进的 Hausdorff 距离。HD 的改进算法主要是对有向 HD 的定义进行修正，如部分 HD 距离和基于平均距离值的 MHD。

部分 HD 是指先求出点集 A 中所有点到点集 B 的距离，然后将这些距离由小到大排序，其中序号 k 的距离即为 $h_k(A, B)$ 。部分 HD 对目标存在严重遮挡或图像退化时可以获得较好效果。部分 HD 的有向 HD 定义如下：

$$d_{HD}^k(A, B) = \min_{a_i \in A} \max_{b_j \in B} \|a_i - b_j\| \quad (3-14)$$

基于平均距离值的 MHD，是指先求出点集 A 中所有点到点集 B 的距离，然后选择这些距离的平均值，MHD 的公式表示如下：

$$d_{MHD}(A, B) = \frac{1}{N_A} \sum_{a_i \in A} d_B(a_i) \quad (3-15)$$

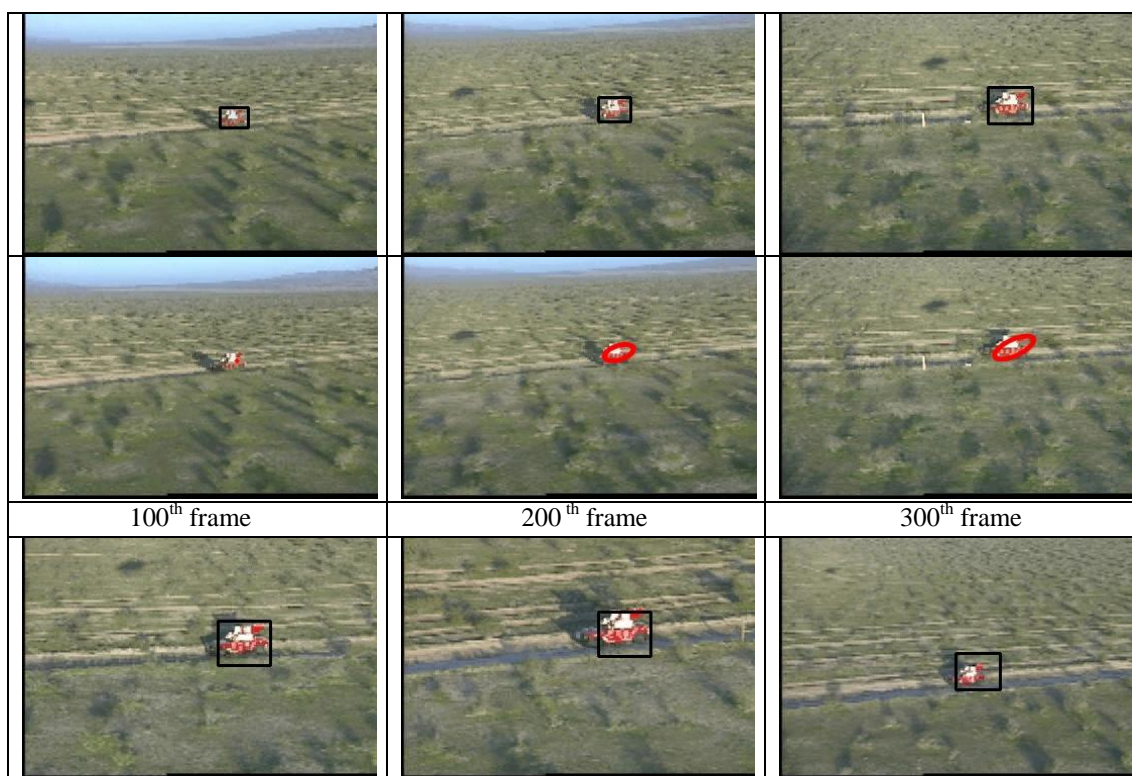
部分 HD 和平均距离值的 MHD 的融合距离，将上面两种方法融合在一起，它的公式表示如下：

$$d_{HB}(A, B) = \frac{1}{k+A} \left(d_{HD}^k + d_{MHD} \right) \quad (3-16)$$

3.3 基于综合特征稀疏表示的目标跟踪实验

本文所使用的测试视频集合，主要来自于目前比较流行的公共测试集和实验室我们自己的跟踪测试集，主要包括：CAVIAR data set^[75]、VIVID data set^[76]和SDL data set^[77]。这些测试视频针对运动背景和前景的变化，其中背景和目标的相对变化主要包括了跟踪目标之间的相互遮挡、目标的尺度变化、目标的旋转、背景光照变化以及相似背景等。在测试视频中不仅有对人体的跟踪，也有对运动车辆的跟踪。

如图 3-6 所示，上面一行矩形表示是使用我们的方法的跟踪结果，下面一行椭圆表示是使用经典 Mean-shift 方法的跟踪结果。图 3-6 中在视频数据来自于 VIVID 数据集，在图 3-6a 中摄像机从高空拍摄红车在草地上行驶，而行驶过程中会出现尺度与形状的变化，图 3-5a 显示了两种跟踪方法都可以在这段视频中成功跟踪红车。在图 3-6b 中摄像机同样从高空拍摄车辆进行 360 度转弯，在这段视频中车辆的表面发生了很大的变化（200th、500th），同时也会出现超车时两车距离很近的情况（1500th），图 3-6b 中显示了，在车辆外观发生变化时，两种方法都可以成功跟踪车辆；但是在 1500th 帧时，经典 mean-shift 方法跟踪失败，因此我们的方法对于车辆与背景的变化具有更稳定的跟踪效果。



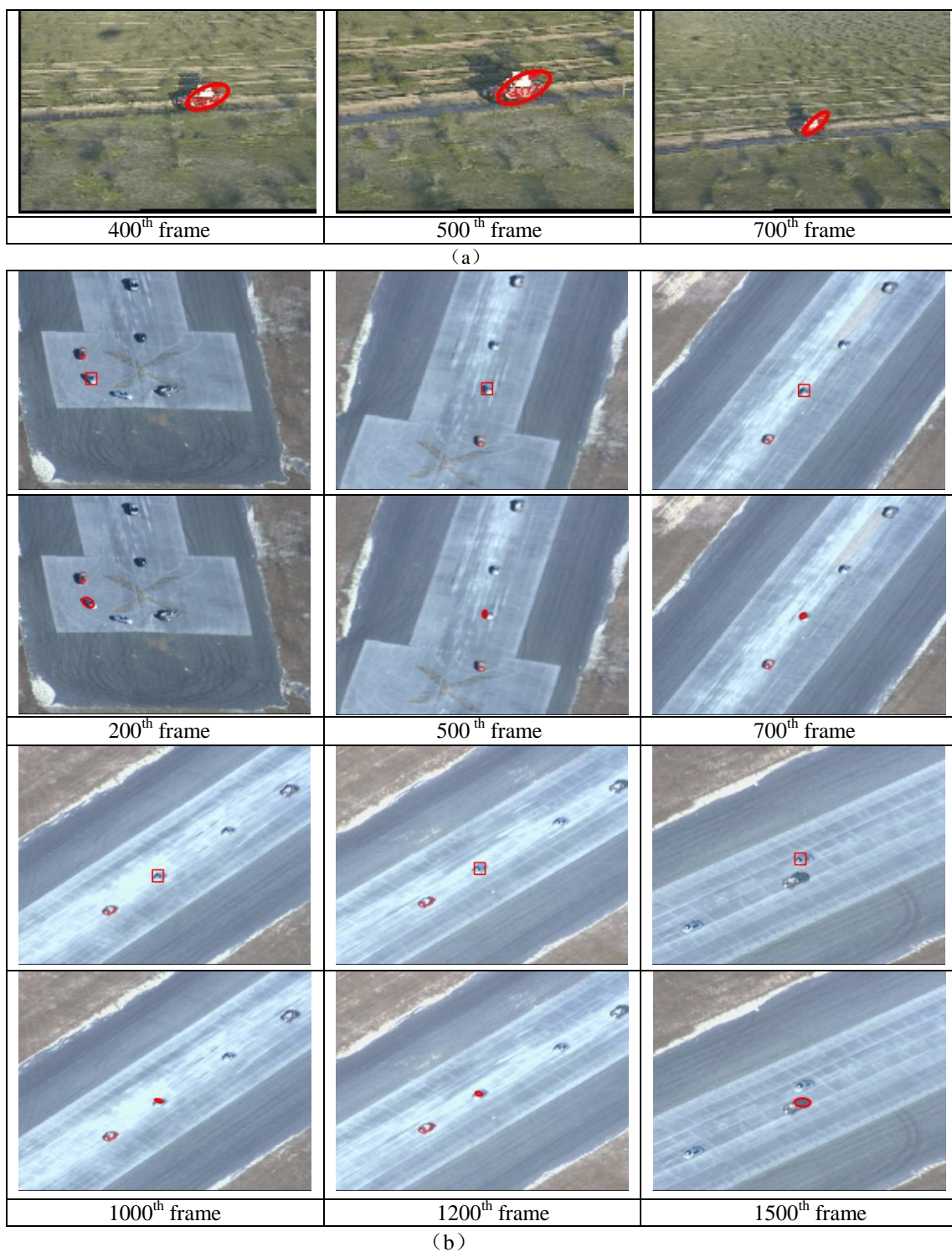
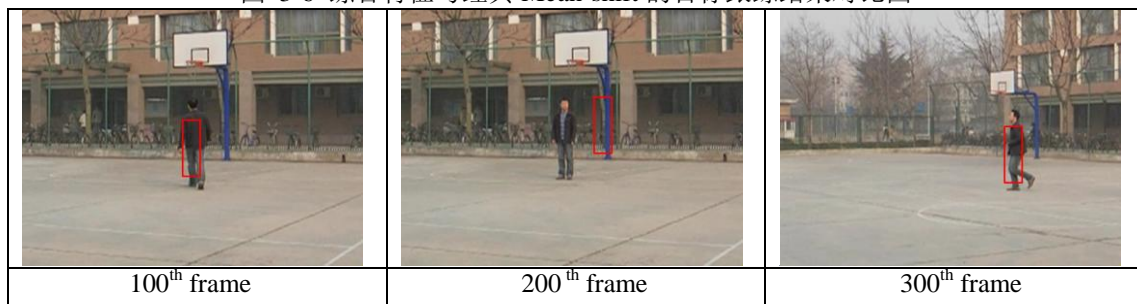
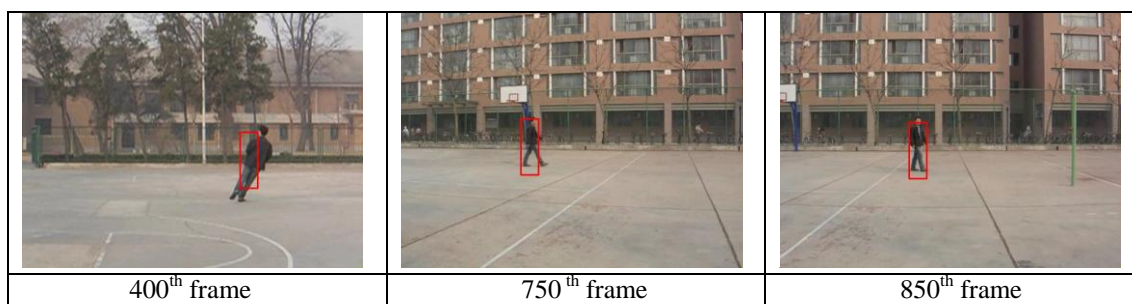


图 3-6 综合特征与经典 Mean-shift 的目标跟踪结果对比图





(c)

图 3-7 基于综合特征稀疏表示的人体跟踪结果

在上面的图 3-7 中，视频数据来自于 SDL 实验室，摄像机拍摄行人在篮球场散步，行人在运动过程中会发生剧烈的形变和尺度变化（100th、400th、750th），同时又与周围环境具有相似的形状和颜色，因此非常具有挑战性。经典 Mean-shift 跟踪方法^[46]和 Collins method^[48]基于目标的颜色特征对目标进行跟踪，在本段视频序列中很快就跟踪失败，而整体跟踪(Ensemble tracking method)^[49]的跟踪结果则与分类器选择的特征及其他因素有关，我们展示基于综合特征稀疏选择跟踪方法的跟踪结果。由图可知，在大部分情况下可以成功跟踪，并且当中间会出现错误跟踪结果时，可以重新找到目标并进行跟踪，展现了很好的鲁棒性。

为了量化跟踪算法跟踪目标的准确率，我们首先对跟踪的视频进行了人工标注，然后采用 Han 等人^[33]定义的跟踪错误率（displacement error rates, DER）来将跟踪算法的跟踪结果与手工标注的结果进行比较。其定义如下：

$$\text{DER} = \frac{\sum_{t=1}^T \text{displacement}(t)}{\sum_{t=1}^T \text{length}(t)} \quad (3-17)$$

从 DER 的定义可知：跟踪算法的性能越好，其 DER 便会越小；跟踪算法越稳定，其 DER 的在图像上表现的越平稳，这样便可以更加直观地比较跟踪算法的性能。

我们使用来自不同数据集^[75-77]的 6 段视频数据，将本章所提到的基于综合特征稀疏选择的跟踪算法与目前比较流行的其他三种跟踪算法进行比较，其中包括：经典的 Mean-shift 跟踪方法^[46]、基于特征选择的跟踪算法(Collins method)^[48]和通过分类器进行整体跟踪的跟踪算法（Ensemble tracking method）^[49]。并且使用前面提到的跟踪错误率 DER 评价指标，由于部分视频序列会存在部分跟踪算法跟踪失败的情况，我们选取了视频序列中前面视频帧的跟踪结

果展示，三个方法的结果曲线图如图 3-8 所示。从这个图中可以看出，相对于其他三种方法，我们的方法获得了较好的效果。

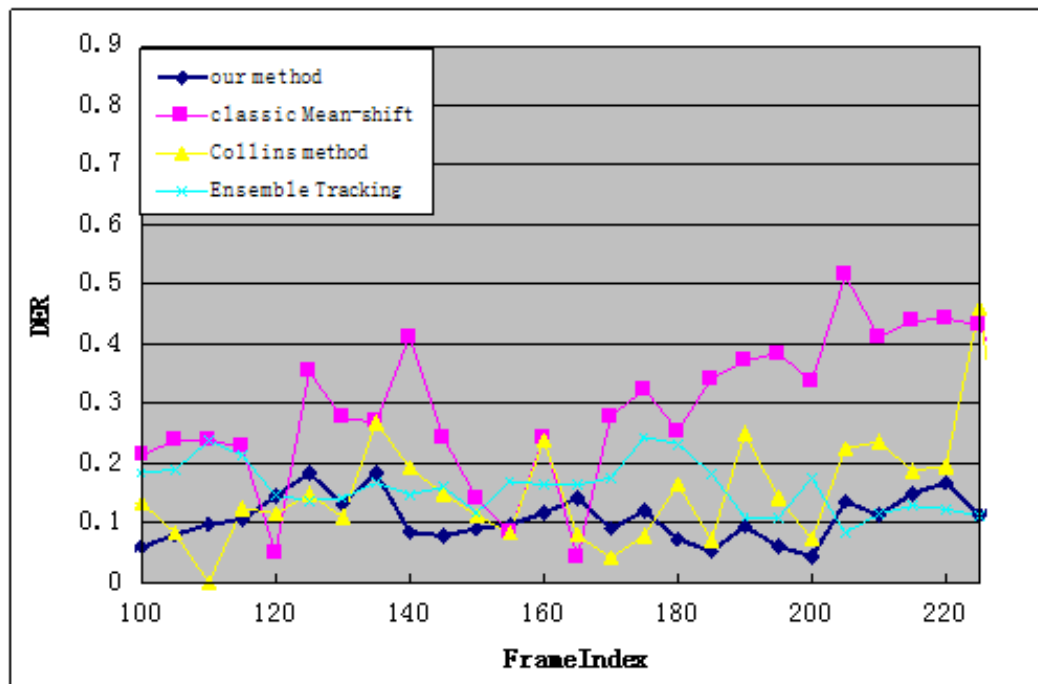


图 3-8 四种跟踪方法平均 DER

第四章 基于稀疏表示与 Mean-shift 框架的快速目标跟踪算法

由于现实情况的需求，目标跟踪往往要求具备实时性好的特点。然而，为了准确地获取目标的信息，目标表达的特征越来越复杂，Haar 小波、轮廓特征、纹理特征、颜色特征及其组合被越来越多地应用在了目标表达中来，因此目标匹配搜索的时效性对目标跟踪的最终算法性能有很大的影响。本章我们主要对目标的快速跟踪进行介绍与讨论，并通过实验验证效果。

4.1 快速目标跟踪算法基础

现有的视频目标跟踪算法，在一定程度上完成了目标的跟踪，但是大多存在跟踪精度不高、特征提取匹配的计算复杂度太大、鲁棒性不够等问题，难以满足现实对跟踪算法的要求。如何提高目标跟踪算法的实时性、鲁棒性和精确性，是目前视频跟踪算法研究的热点和难点。

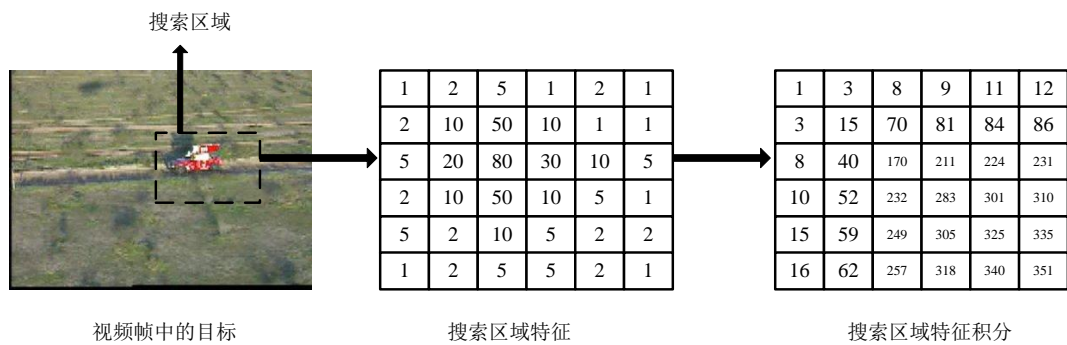


图 4-1 搜索区域积分示意图

模板匹配作为一种常用的相似度度量方法，在目标变化较少的情况下，具有很高的准确性，因此在目标跟踪中有着广泛地应用。但是在匹配过程中候选项的特征提取需要进行大量的计算，当目标较大或者搜索区域较大时，往往不能达到实时性的要求。在已有的算法中，有的通过缩小搜索区域的方法来加快匹配的速度，有的通过改进特征提取的方法减少提取的时间，还有一种被广泛使用的方式是特征积分图的方法来减少候选项特征集的计算时间。

在目标搜索区域的每个候选项特征提取前，我们可以通过将整个搜索区域的特征积分图提取，然后通过特征积分图的加减运算得到每个候选项的特征向量。这样，搜索区域中任何位置的特征向量可以有积分图中该位置以及前面左上、左、上等四个位置的特征向量积分制通过加减运算得到，大大减少了目标特征提取的时间。如图 4-1 所示。

Mean-shift 算法的形式简洁，收敛条件宽松。只要确定了跟踪目标，整个跟踪过程就不需额外的参数输入。由于无需对概率密度分布进行估计，就可以直接沿着梯度方向搜索局部最大值，大大减少了运算量，所以这种算法具有良好的实时性。

在传统的算法中^[46]，我们采用一个目标的灰度或颜色分布来描述这个目标，假设目标中心位于 x_0 ，则该目标可以表示为：

$$\hat{q}_u = C \sum_{i=1}^n k \left(\left\| \frac{x_i^s - x_0}{h} \right\|^2 \right) \delta [b(x_i^s) - u] \quad (4-1)$$

加权的灰度或颜色直方图，使距离目标中心的像素点具有较大的权值，具有抵抗部分遮挡和边缘噪声影响的优点。

候选的位于 y 的目标可以描述为：

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{x_i^s - y}{h} \right\|^2 \right) \delta [b(x_i^s) - u] \quad (4-2)$$

因此目标跟踪可以简化为寻找最优的 y ，使得 $\hat{p}_u(y)$ 与 \hat{q}_u 最相似。

$\hat{p}_u(y)$ 与 \hat{q}_u 的相似性用 Bhattacharrya 系数 $\hat{\rho}(y)$ 来度量，即

$$\hat{\rho}(y) \equiv \rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y) \hat{q}_u} \quad (4-3)$$

公式(4-3)在 $\hat{p}_u(\hat{y}_0)$ 点泰勒展开可得，

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(y_0) q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (4-4)$$

把公式(4-2)带入公式(4-4)，整理可得：

$$\rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p(y_0) q_u} + \frac{C_h}{2} \sum_{i=1}^n w_i k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (4-5)$$

其中，

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q_u}{p_u(y_0)}} \quad (4-6)$$

对公式(4-6)右边的第二项，可以利用 Mean-shift 算法对其进行最优化。

Mean-shift 算法用于视频跟踪有以下几个优势：首先，算法计算量不大，在目标区域已知的情况下完全可以做到实时跟踪；其次，作为一个无参数概率密度估计算法，很容易和其他算法结合使用；然后，采用加权直方图模型，对部分遮挡、目标旋转、形变和背景运动不敏感。但是，其也有不足之处，比如：缺乏必要的模板更新算法，整个跟踪过程中窗宽的大小保持不变，因此当目标有尺度变化时，可能跟踪失败；颜色直方图是一种比较通用的目标特征描述子，当背景和目标的颜色直方图相似时，跟踪效果往往不好；另外，Mean-shift 算法是局部最优的优化算法，当出现多个局部峰值时，算法可能会收敛于局部最优值而造成跟踪算法失效。

针对 Mean-shift 算法存在的缺陷，许多改进的算法展现出了很好的实验效果。卡内基-梅隆大学的 Collins 等人^[48]根据颜色特征对目标的视角、目标之间的部分遮挡以及目标形状的不敏感性提出了一种基于颜色的目标特征描述的新方法。他们所提出的跟踪特征集合是颜色特征 R、G 和 B 的线性组合。当提取得到前景和背景的跟踪特征之后，Collins 等人首先计算了每种特征在前景与背景之间的相似性，定义如下：

$$L(i) = \log \frac{\max\{p(i), \delta\}}{\max\{q(i), \delta\}} \quad (4-7)$$

其中， $p(i)$ 定义为目标中的第 i 个特征，而 $q(i)$ 为背景中的第 i 个特征。 $L(i)$ 为目标与背景基于第 i 个特征的 \log 似然比，包含了前景与背景在某一个特征之间的相似度信息。然后通过统计第 i 个特征在视频帧中得到的每一个 $L(i)$ ，并计算其方差便得到了目标与背景在第 i 个特征之间的判别力，其中 var 定义为某一特征相似度的方差，公式表达如下：

$$VR(L; p, q) = \frac{\text{var}(L; (p+q)/2)}{[\text{var}(L; p) + \text{var}(L; q)]} \quad (4-8)$$

Liang 等人^[3]提出了一种类似的方法，但是特征的判别力是通过贝叶斯误差率定义的，即通过目标和前景颜色特征直方图的交集进行计算得到的。Wang

等人^[4]提出了一种通过 Adaboost 学习选取 Harr 特征子集作为跟踪特征集的方法。Avidan^[49]考虑到分类器在目标检测中的良好性能，提出了与分类器相结合的综合跟踪方法 (Ensemble Tracking)，事先对目标和背景的样本通过训练得到强分类器，并且在跟踪过程中实时更新分类器，取得了更高的准确度与适应度，但是却一定程度上牺牲了目标跟踪的实时性。

在他们研究的基础上，我们提取目标的综合直方图，将目标的局部颜色和轮廓特征有效地融合起来，使得描述目标更加准确，所包含关于目标的信息量更多，便于计算机处理和进行有效的跟踪；然后通过 1 范数最小化的方法来得得到目标特征的稀疏表示，获取搜索区域的概率权值图，最后在该图中运行 Mean-shift 算法，找到目标在下一帧图像中的位置。

4.2 基于稀疏表示与 Mean-shift 的快速目标跟踪算法

由式(4-6)可知，Mean-shift 在跟踪中的主要应用是寻找最优候选目标 y ，使得 $\hat{p}_u(y)$ 与 \hat{q}_u 最相似，具有计算量小，同时目标旋转等形变不敏感等优点，而综合特征弥补了颜色直方图的单一性，稀疏表示和选择又可以选取具备区分目标与背景能力的特征子集，因此综合特征稀疏选择与 Mean-shift 算法相结合具有很强的互补性。之前的许多改进算法在于对搜索区域的权值计算方法进行了改进，我们首先采用特征稀疏选择得到系数向量 S ，如式(4-11)所示。

$$\begin{aligned} & \min \|S\|_1 \\ & s.t. \begin{cases} S^T M_j^+ \geq \alpha \\ S^T M_j^- \leq -\alpha \end{cases} \end{aligned} \quad (4-11)$$

其中， M^+ 为目标区域的正例训练样本， M^- 为背景区域的反例训练样本， α 为正例与反例的区分度。这样，权值 w_i 可由式(4-12)得到。

$$\begin{cases} w_i = S_i & , S_i \geq \varepsilon \\ 0 & , else \end{cases} \quad (4-12)$$

采用稀疏选择的搜索区域目标概率权值图像较稀疏前有了更明显的区分力，权值图中的像素点亮度呈现出向两级（0 和 255）分化的趋势，减少了灰色（中间）像素点的数量，如图 4-2 所示。

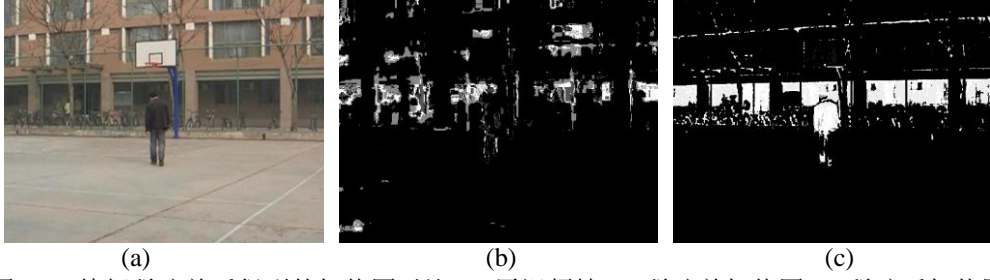


图 4-2 特征稀疏前后得到的权值图对比 (a)原视频帧 (b)稀疏前权值图 (c)稀疏后权值图

同样地，传统的 Mean-shift 算法不能对目标的大小进行自适应改变，本节借鉴 Liang 等人^[3]的思路，将目标边缘部分进行模板计算。但是由于一阶梯度在目标形状不规则时容易出现误差，我们采用二阶梯度来计算目标的尺寸，即假设在目标的边缘部分，目标内部出现的特征变化要远小于目标与特征之间的特征变化。如图 4-3 所示，目标的中心点 (x_0, y_0) 确定后，目标的上下左右边界可以分别由式(4-13),(4-14),(4-15)以及(4-16)确定。

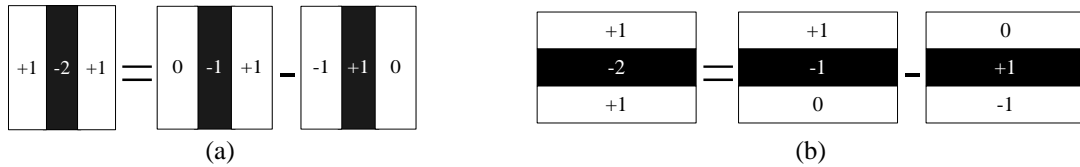


图 4-3 目标尺寸计算模板 (a) 目标左右侧计算模板 (b)目标上下侧计算模板

$$x_L = \arg \max_{|x-x_L| \leq \lambda w, y=y_0} C_W \otimes \omega_t(x, y) \quad (4-13)$$

$$x_R = \arg \max_{|x-x_R| \leq \lambda w, y=y_0} C_W \otimes \omega_t(x, y) \quad (4-14)$$

$$y_T = \arg \max_{x=x_0, |y-y_T| \leq \lambda h} C_H \otimes \omega_t(x, y) \quad (4-15)$$

$$y_B = \arg \max_{x=x_0, |y-y_B| \leq \lambda h} C_H \otimes \omega_t(x, y) \quad (4-16)$$

$$C_I \otimes \omega_t(x, y) = \sum_{i=-w/2}^{w/2} \sum_{j=-h/2}^{h/2} C_I(i, j) \times \omega_t(x+i, y+j) \quad (4-17)$$

其中， λ 决定每一帧的变化范围，运算符 \otimes 由式(4-17)决定， w 和 h 分别为目标在 $t-1$ 时刻的宽度和高度，这样当目标出现尺寸变化时，跟踪算法仍然可以成功跟踪目标并适应目标的尺寸变化。

4.3 基于特征稀疏表示的快速目标跟踪实验

本节将对基于综合特征稀疏表示的快速目标跟踪算法进行总结，并通过在

数据集上进行实验来验证算法的性能。通过与其他经典的跟踪算法进行实验对比，我们可以发现，我们的算法在目标发生颜色与形状变化以及复杂背景（如部分遮挡）情况下具有较好的效果，并且与其他方法相比本算法展现出了良好的效率。

4.3.1 实验步骤

基于稀疏表示的快速目标跟踪算法的基本实验步骤，如表 4-1 所示。

表 4-1 基于稀疏表示的快速目标跟踪算法

1. 初始化.

初始化目标位置，计算目标与对应背景的特征；

2. 构造训练样本集.

对目标特征向量与背景特征向量分别构建训练集合 M^+ 与 M^- ；

3. 获取目标搜索区域概率权值图.

3.1 使用 M^+ 与 M^- 计算目标样本的稀疏系数 S ；

3.2 由稀疏系数 S 得到搜索区域所有位置的概率权值；

4. 目标跟踪.

4.1 通过梯度向量递归的方法搜索目标中心点；

4.2 计算目标的大小变化并得到目标的位置与大小；

5. 训练样本集更新.

5.1 将新得到的目标样本替换掉训练样本集中的最初始样本；

5.2 跳转到步骤 3，直到完成跟踪过程。

由上述的实验步骤可以看出，本文的方法与传统 Mean-shift 方法相比，主要在于权值的计算方法不同，我们在得到目标特征向量和背景特征向量的基础上采用稀疏表示的方式来获取目标特征向量的稀疏系数，这样得到的权值可以有效地区分目标与背景；通过对训练样本集的实时更新，稀疏表示更可以有效应对目标部分遮挡的情况。

4.3.2 实验结果

图 4-4 中的 (a) 和 (b) 两个视频段分别来自于 CAVIAR 数据集与来 SDL

数据集。在 (a) 中, 目标跟踪的难点在于当行人经过中间立柱时(80th 到 120th 帧), 行人与立柱的颜色几乎相同, 仅仅依靠颜色特征在跟踪过程中非常容易跟踪失败, 我们采用了综合特征的稀疏表示在 Mean-shift 框架下来跟踪行人, 虽然跟踪过程中出现了抖动现象, 但是仍然能够成功完成跟踪过程。在 (b) 中, 视频中出现了严重的目标遮挡 (300th 到 600th 帧), 同时目标有很大的外观变化 (700th 帧), 并且遮挡物与目标之间非常相似, 属于比较难的跟踪视频, 我们的方法可以正确的追踪目标。这两段视频的跟踪结果表明, 我们的方法可以有效地解决视频中复杂背景下的目标部分遮挡与外观变化等情况。

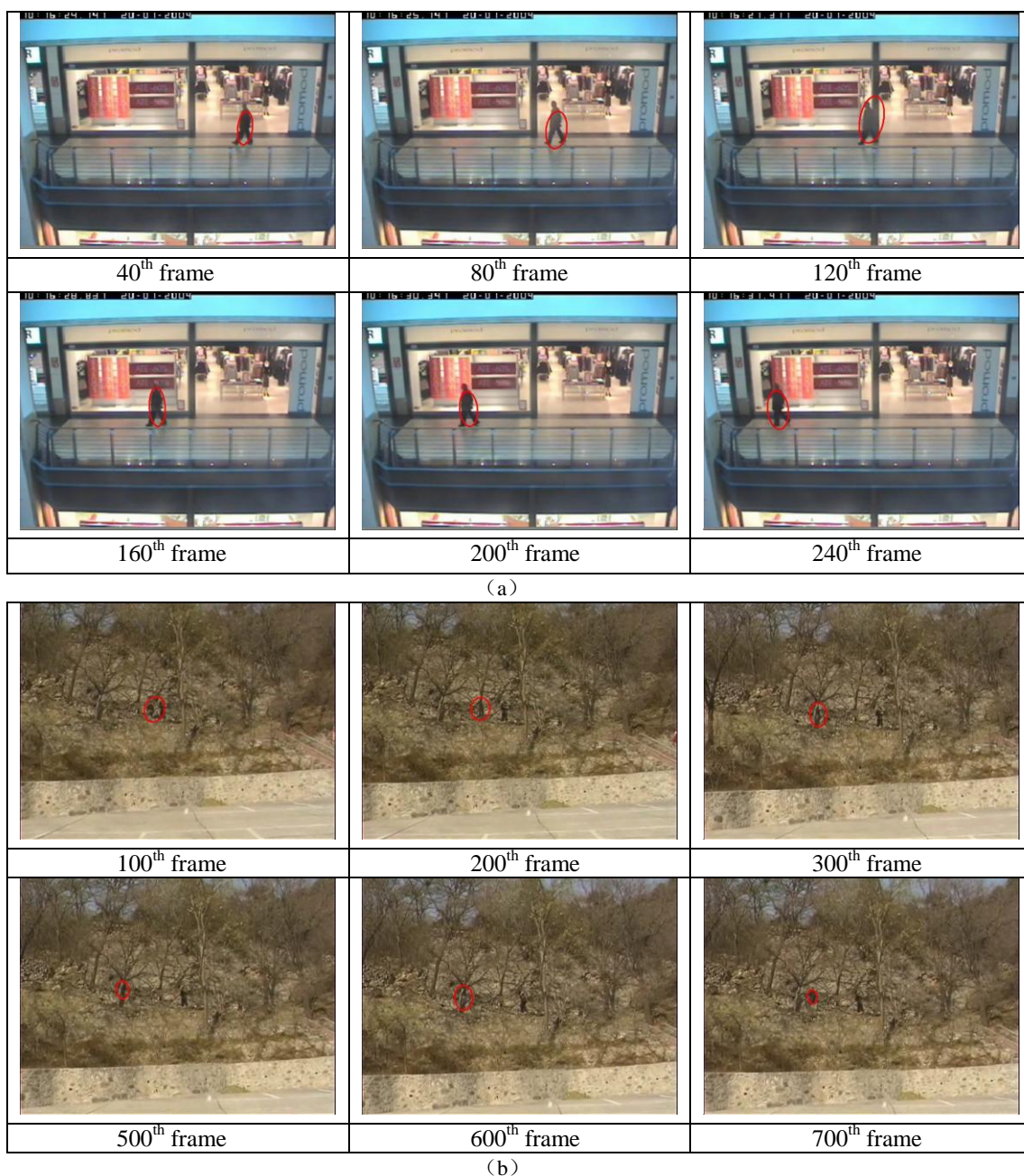


图 4-4 基于稀疏表示与 Mean-shift 框架的跟踪结果

4.3.3 算法性能分析

为了量化跟踪算法跟踪目标的速度和效率，我们使用平均迭代次数来对几种 Mean-shift 框架算法的效率进行对比，结果如表 4-2 所示。

通过表 4-2 中的结果我们可以看到，相对于经典 Mean-shift 算法，其他三种算法的平均迭代次数都有了较大程度的改进，这是由于对目标和背景进行区分/分类选择特征的结果，我们的算法迭代次数略低于基于特征选择的跟踪算法和整体跟踪，提高了跟踪的效率。

表 4-2 四种跟踪方法平均迭代次数

算法	Classic Mean-shift	Online selection	Ensemble tracking	Our method
迭代次数	5.72	3.10	3.83	3.02

总结与展望

复杂背景下的目标跟踪是目标跟踪中的难点，本文通过提取目标的综合特征，结合改进 Mean-shift 跟踪算法来进行实验，实现对目标的快速跟踪。针对目标跟踪中的目标变化、复杂背景以及算法性能等难点，本文针对性的对这些问题加以研究并取得了较好的效果。

我们在目标表达和目标匹配搜索两个目标跟踪的步骤中分别加以改进，提出了一种面向跟踪的目标特征稀疏表示对进行目标表达，融合了目标的颜色、局部轮廓、梯度方向等特征，并且通过对目标样本和背景样本进行 L1-norm 最小化的方式，利用稀疏表示对目标特征评估选择；同时将模板更新算法加入到传统的均值漂移跟踪算法中来，在假设目标边缘变化显著，即周围的目标点与背景点区别较大的情况下对目标尺度进行计算，提高了跟踪算法的性能。我们分别在三个数据集进行实验验证，并且将实验结果与经典算法进行了对比。结果表明我们的算法在部分遮挡、目标与背景变化等情况下仍然能够取得较为准确和鲁棒的结果。

本文提出的算法虽然取得了一定成功，但是仍然有改进的空间。由于没有将滤波器的运动预测融入到算法中来，在目标跟踪过程中，我们的跟踪结果不够稳定，会出现抖动现象。另外，我们提出了解决目标尺度变化的模板计算方法，在大多数目标形状规则的情况下能够取得较好的效果，但是在目标形状不规则或者剧烈变化时往往不能适应目标的尺寸变化。我们希望在将来的工作中把滤波器算法与本文算法相融合，并对目标形变时的尺寸变化进行深入研究。

参考文献

- [1] A. Sumpter, and A. Bulpitt. Learning spatio-temporal patterns for predicting object behaviour [J]. *Image and Vision Computing*, 2000: 649-658.
- [2] D. Makris, and T. Ellis. Automatic learning of activity-based semantic scene model [J]. *Proceedings of IEEE Advanced Video and Signal Based Surveillance*, 2003: 183-188.
- [3] D.W. Liang, Q.M. Huang, W. Gao, and H.X. Yao. Online selection of discriminative features using Bayes error rate for visual tracking [C]. *7th Pacific-Rim Conference on Multimedia*, 2006: 547-555.
- [4] J. Wang, X. Chen, and W. Gao. Online selecting discriminative tracking features using Particle Filter [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 1037-1042.
- [5] Weiming Hu, Dan Xie and Tieniu Tan. A hierarchical self-organizing approach for learning the patterns of motion trajectories [J]. *IEEE Transactions on Neural Networks*, 2004: 135-144.
- [6] 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.
- [7] 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.
- [8] A. Yilmaz, K. Shafique, and M. Shah, Target tracking in airborne forward looking infrared imagery [J], *Image and Vision Computing*, 2003:623-635.
- [9] A. Yilmaz, O. Javed, and M. Shah, Object Tracking: A Survey [J], *ACM Computing Surveys*, 2006.
- [10] G. Bradski. Real time face and object tracking as a component of a perceptual user interface [C]. *Proceedings of IEEE Workshop on Applications of Computer Vision*, 1998: 214-219.
- [11] N. Papanikolopoulos, P. Khosla, and T. Kanade. Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision [J]. *IEEE Transactions on Robotics and Automation*, 1993: 14-35.

- [12] J. Shi, and C. Tomasi. Good features to track [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994: 593-600.
- [13] M. Isard, and A. Blake. CONDENSATION-conditional density propagation for visual tracking [J]. *International Journal on Computer Vision*, 1998:5-28.
- [14] H. Nguyen, and A. Smeulders. Tracking aspects of the foreground against the background [C]. *Proceedings of European Conference on Computer Vision*, 2004: 446-456.
- [15] S. Baker, and I. Matthews. Lucas-Kanade 20 years on: A unifying framework [J]. *International Journal on Computer Vision*, 2004: 221-255.
- [16] Weiming Hu, Dan Xie and Tieniu Tan. A hierarchical self-organizing approach for learning the patterns of motion trajectories [J]. *IEEE Transactions on Neural Networks*, 2004: 135-144.
- [17] H. Bohyung, and L. Davis. Object tracking by adaptive feature extraction [C]. *IEEE International Conference on Image Processing*, 2004: 1501-1504.
- [18] 李美花. 三维人体跟踪 [D]. 大连理工大学硕士学位论文, 2004.
- [19] T. Yang, Q. Pan, J. Li, and S.Z. Li. Real-time multiple objects tracking with occlusion handling in dynamic scenes [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 970-975.
- [20] 彭宁嵩, 杨杰, 刘志, 张凤超. Mean-Shift 跟踪算法中核函数窗宽的自动选取 [J]. *软件学报*, 2005: 1542-1550.
- [21] Y. Wu. Robust visual tracking by integrating multiple cues based on co-inference learning [J]. *International Journal of Computer Vision*, 2004: 55-71.
- [22] R. Collins. Mean-shift blob tracking through scale space [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2003: 234-240.
- [23] M. Isard, and A. Blake. Condensation - conditional density propagation for visual tracking [J]. *International Journal of Computer Vision*, 1998: 5-28.
- [24] T.X. Han, V. Ramesh, Y. Zhu, and T.S. Huang. On optimizing template matching via performance characterization [C]. *International Conference on Computer Vision*, 2005: 182-189.
- [25] B.Han, Y. Zhu, D. Comaniciu, and L. Davis. Kernel-based Bayesian filtering

- for object tracking [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 227-234.
- [26] Y. Huang, and I Essa. Tracking multiple objects through occlusions [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 1051-1058.
- [27] A. Elgammal, R. Duraiswami, and L. Davis. Probabilistic tracking in joint feature-spatial spaces [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2003: 781-788.
- [28] Z. Khan, T. Balch, and F. Dellaert. A rao-blackwellized particle filter for eigentracking [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2004: 980-986.
- [29] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. Color-based probabilistic tracking [C]. *European Conference on Computer Vision*, 2002: 661-675.
- [30] David G. Lowe. Distinctive image features from scale-invariant key points [J]. *International Journal of Computer Vision*, 2004: 91-110.
- [31] K. Okuma, A. Taleghani, N. de Freitas etc. A boosted particle filter: multi target detection and tracking [C]. *European Conference on Computer Vision*, 2004: 28-39.
- [32] C. Stauffer, and W. Grimson. Adaptive background mixture models for real-time tracking [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1999: 246-252.
- [33] Z.J. Han, Q.X. Ye, and J.B. Jiao. Online Feature Evaluation for Object Tracking using Kalman Filter [C]. *International Conference on Pattern Recognition*, 2008.
- [34] Z. Han, Q. Ye, Y. Liu, et al. Feature evaluation by particle filter for adaptive object tracking [J]. *Proceedings of International Society for Optical Engineering*, 2009.
- [35] S. Avidan. Support Vector Tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004: 1064-1072.
- [36] G.Y. Zhu, D.W. Liang, Y. Liu, Q.M. Huang, and W. Gao. Improving Particle Filter with Support Vector Regression for Efficient Visual Tracking [C]. *IEEE International Conference on Image Processing*, 2005: 1501-1504.

- [37] H.D. Dai, M. Chen, and S.L. Zhou. Support Vector Machine Based Adaptive Kalman Filtering for Target Tracking [J]. *General System and Control System*, 2007: 114-118.
- [38] F. Tang, S. Brennan, Q. Zhao, and H. Tao. Co-tracking Using Semi-Supervised Support Vector Machines [C]. *IEEE International Conference on Computer Vision*, 2007: 992-999.
- [39] J. Zhang, D.W. Chen, and M. Tang. Combining Discriminative and Descriptive Models for Tracking [C]. *Asian Conference on Computer Vision*, 2009.
- [40] N. Kabaoglu. Target Tracking Using Particle Filters with Support Vector Regression [J]. *IEEE Transactions on Vehicular Technology*, 2009: 2569-2573.
- [41] B. Scholkopf, J.C. Platt, J. Shawe-Taylor, A.J. Smola, and R.C. Williamson. Estimating the support of a high-dimensional distribution [J]. *Neural computation*, 2001:1443-1471.
- [42] A. Yilmaz, K. Shafique, and M. Shah, Target tracking in airborne forward looking infrared imagery [J], *Image and Vision Computing*, 2003:623-635.
- [43] D. Chen, and J. Yang. Robust object tracking via online spatial bias appearance model learning [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007: 2157-2169.
- [44] J.Q. Wang, and Y.S. Yagi. Integrating color and shape-texture features for adaptive real-time object Tracking [J]. *IEEE Transactions on Image Processing*, 2008: 235-240.
- [45] N. Dalal, and B. Triggs, Histograms of oriented gradients for human detection [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp:886-893, 2005.
- [46] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003: 564-577.
- [47] J. Wang, X. Chen, and W. Gao. Online selecting discriminative tracking features using Particle Filter [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 1037-1042.
- [48] R. Collins, and Y. Liu. Online selection of discriminative tracking features [C]. *Proceedings of 9th IEEE International Conference on Computer Vision*, 2003:

- 346-352.
- [49] S. Avidan. Ensemble tracking [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2005: 494-501.
- [50] M. Isard, and A. Blake. CONDENSATION-conditional density propagation for visual tracking [J], *International Journal on Computer Vision*, 1998: 5-28. .
- [51] E. Cuevas, D. Zaldivar, and R. Rojas. Particle filter for vision tracking [C], *Technical Report B, Fachbereich Mathematikund Informatik, Freie Universit ät Berlin*, 2005.
- [52] 王建宇. 基于序列蒙特卡洛滤波算法的视觉目标跟踪 [D], *哈尔滨工业大学博士论文*, 2006.
- [53] T. Serre. Learning a dictionary of shape-components in visual cortex: Comparison with neurons, humans and machines [D], *Ph.D dissertation, MIT*, 2006.
- [54] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma. Robust face recognition via sparse representation [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2009: 210-227.
- [55] Xue Mei, Haibing Ling [C]. Robust visual tracking using l1 Minilization, *International Conference on Computer Vision*, 2009: 1436-1443.
- [56] R. DeVore. Deterministic constructions of compressed sensing matrices [J]. *Journal of Complexity*, 2007: 918-925.
- [57] D. Donoho, and Y. Tsaig. Extensions of compressed sensing [J]. *Signal Processing*, 2006: 533-548.
- [58] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk. An architecture for compressive imaging [C]. *IEEE International Conference on Image Processing*, 2006: 1273-1276.
- [59] J. Ma. Single-pixel remote sensing [J]. *IEEE Geoscience and Remote Sensing Letters*, 2009: 199-203.
- [60] S.F. Cotter, and B.D. Rao. Sparse channel estimation via matching pursuit with application to equalization [J]. *IEEE Transactions on Communications*, 2002: 374-377.
- [61] G. Taubock, and F. Hlawatsch. A compressed sensing technique for OFDM

- channel estimation in mobile environments: exploiting channel sparsity for reducing pilots [C]. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008: 2885-2888.
- [62] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk. Single-pixel imaging via compressive sampling [J], *IEEE Signal Processing Magazine*, 2008: 83-91.
- [63] B.A. Olshausen, and D.J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images [J]. *Nature*, 1996: 607-609.
- [64] D.L. Donoho. For most large underdetermined systems of equations, the minimal l_1 -norm near-solution approximates the sparsest near-solution [J]. *Communications on Pure and Applied Mathematics*, 2006: 907-934.
- [65] S. Chen, and D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit [J]. *Society for Industry and Applied Mathematics Journal on Scientific Computing*, 1999: 33-61.
- [66] D. Donoho, and X. Huo. Uncertainty principles and ideal atomic decompositions [J]. *IEEE Transactions on Information Theory*, 2001: 2845-2862.
- [67] E. Candes, J. Romberg and T. Tao. Robust uncertainty principles: exact signals reconstruction from highly incomplete frequency information [J]. *IEEE Transactions on Information Theory*, 2006: 489-509.
- [68] D. Donoho. Compressed sensing [J]. *IEEE Transactions on Information Theory*, 2006: 1289-1306.
- [69] J. Yang, J. Wright, T. Huang, and Y. Ma. Image super-resolution as sparse representation of raw image patches [C], *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [70] T. Wan, N. Canagarajah, and A. Achim. Compressive image fusion [C]. *IEEE International Conference on Image Processing*, 2008.
- [71] K. Egiazarian, A. Foi, and V. Katkovnik. Compressed sensing image reconstruction via recursive spatially adaptive filtering [C]. *IEEE International Conference on Image Processing*, 2007.

- [72] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration [J]. *IEEE Transaction on Image Processing*, 2008: 53-69.
- [73] D.L. Donoho, and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via L1-norm minimization [C], *Proc. Nat. Aca. Sci*, 2003: 2197-2202.
- [74] J. Yang, J. Wright, T. Huang, and Y. Ma. Image super-resolution as sparse representation of raw image patches [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [75] CAVIAR Test Case Scenarios at: <http://homepages.inf.ed.ac.uk/rbf/CAVIAR>.
- [76] VIVID data set at: <http://www.vividevaluation.ri.cmu.edu/datasets>.
- [77] SDL Data Set at: <http://coe.gucas.ac.cn/SDL-HomePage>.

作者简历

刘一飞 男 汉族 共青团员

教育经历

- 2005 年 9 月至 2009 年 7 月 同济大学 电子信息工程专业 工学学士
- 2009 年 9 月至 2012 年 7 月 中国科学院研究生院 计算机应用技术专业 工学硕士

已发表（录用）文章目录

- Yifei Liu, Zhenjun Han, Qixiang Ye, Jianbin Jiao, Ce Li, “Sparse Feature Representation for Visual Tracking”, International Conference on Systems and Informatics, 2012. (accepted)

附录：主要符号对照表

HC	颜色直方图 (Histogram of Color)
HOG	梯度方向直方图 (Histogram of Oriented Gradient)
HOGC	梯度方向与颜色综合直方图 (Histogram of Oriented Gradient and Color)
CAM-Shift	连续自适应均值漂移基 (Continuously Adaptive Mean-shift)
SR	稀疏表示 (Sparse Representation)
SRC	稀疏表示分类 (Sparse Representation Classification)
OPSR	梯度投影稀疏重构 (Gradient Projection for sparse Reconstruction)
DER	位置错误率 (Displacement Error Rates)

致 谢

在攻读硕士学位三年的学习生活中，我经历了诸多坎坷，也付出了艰辛的努力，同时也得到了很大的收获。在毕业论文完成之际，由衷地感谢这三年来曾经给予我无数帮助的老师、同学、朋友和家人。

首先，我要感谢我的导师焦建彬教授将我列入门墙，使我有机会来到中科院研究生院，得以与实验室的师生一起度过三年美好的时光。我要感谢焦老师在我取得进步和遇到挫折时给予的鼓励和支持，感谢他在我论文的撰写和修改中倾注的心血，感谢他对我攻读硕士期间顽劣之处的宽容和包涵。三年来，恩师对我的学习和工作给予了悉心的指导，给我的研究提供了良好的环境和机会，让我受益匪浅。他严谨的治学态度和平易近人的学者风范深深地影响着我，他诲人不倦的精神和对我的谆谆教导让我感动，他的这些言传身教将使我受益终生。

其次，感谢叶齐祥老师和韩振军师兄。他们在我的科研学习中的理论和实验方面给予了耐心的指导，并在论文的撰写和审稿中给予了很多帮助。他们渊博的专业知识，自强不息的学习精神、扎实的动手能力和对我的谆谆教导，使我从他们身上学到了不少可贵的东西，让我受益匪浅。

此外，张立国师兄和陈孝罡师兄在学习和生活上也给予了很多帮助；在我论文撰写期间，实验室的高文、李策、武博、纪颖夏、杨威、高山等同学也给予了我很多支持和帮助。

特别要感谢我的亲人们，尤其是我的父母和外公外婆，感谢他们多年来一直给我最无私的关怀，无微不至的关心和爱护，感谢他们为我所做的一切。亲人们对我的爱和期望一直指引我奋发拼搏、积极向上，我将对此终生铭记。

最后，感谢参加论文评审和答辩的各位老师。

刘一飞

2012年4月