

密级:

中国科学院研究生院

硕士学位论文

基于拓扑模型的行人检测方法研究

作者姓名: 高文

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 4 月

Part-based Pedestrian Detection Via Topology Model

By
Gao Wen

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

Graduate University of Chinese Academy of Sciences

April, 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：

中国科学院研究生院直属院系

研究生学位论文原创性声明

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

行人检测在智能交通、视频监控、多媒体检索等领域具有重要的应用价值。行人检测的相关方法与关键技术已经成为计算机视觉领域研究热点。然而，行人的多姿态、多尺度、局部遮挡等因素均会影响检测性能。近年来，基于部位模型的行人检测方法已经成为目前目标检测领域的热点，本文以此为研究内容，引入拓扑描述及建模方法。相关成果具有理论意义和应用价值。

论文提出了一种基于拓扑模型的行人检测方法，该方法将行人分为若干个部位，并对各个部位分别进行检测，保证每个待检测部位在最大响应值下输出检测结果。提出了一种行人拓扑描述子（Log-polar Topology Pattern）用以描述行人各个部位间的拓扑关系，并利用单类支持向量机（One-Class Support Vector Machine，OC-SVM）训练行人拓扑模型，对图片中的待检测样本进行二分类，实现图片与视频帧中的行人检测。

在行人部位检测的过程中，利用 HOG 特征和线性 SVM 分类器进行检测。在行人拓扑关系描述的过程中，考虑到部位检测过程中各部位的独立性，引入了整体拓扑描述。拓扑模型描述了有拓扑关联的部位间两两关系，同时能够描述整个行人的姿态特征。鉴于行人姿态描述的特殊性，采用 OC-SVM 作为分类器对拓扑关系进行建模。OC-SVM 应用于只包含一类样本的情况，在高维特征空间中构造一个尽可能多的包络正例样本且尽可能小的超球体，从而描述一类样本数据分布的二值模型。

本文提出的行人检测方法在两个公开数据集合上进行了实验验证，均取得了较好的效果，能够显著提高在多姿态、多尺度、局部遮挡等情况下的检测率。

关键词：行人检测，拓扑描述，部位检测，分类器, OC-SVM

Abstract

- **Gao Wen** (Computer Application)

Directed by: **Jiao Jian-bin** (Professor)

Pedestrian detection has a variety of applications in intelligent transportation systems, video surveillance, image compression, content-based multimedia retrieval. Therefore, it has become a most active area in computer vision research. Having a long research history, however, the variation of pedestrian appearances, crowd scenes, movements, occlusions and lighting conditions make it still an open problem. In recent years, part-based pedestrian detection has been one of the most representative methods. In this paper, we investigate the method of topology description and learning, which is highly desirable to research.

We present a pedestrian detection method based on parts detection and topological relation modeling. A pedestrian object is divided into parts and the intermediate outputs of parts detection is the maximum response values of each detection procedure. We present a pedestrian topological descriptor (Log-polar Topology Pattern) to describe the topological relations between all parts of the pedestrian and use an One-Class Support Vector Machine (OC-SVM) classifier to train pedestrian topology model. Finally, the sample image is classified to discriminate a pedestrian or non-pedestrian.

In parts detection procedure, we use HOG feature and linear SVM classifier to train independent classification part models. In pedestrian topology description procedure, considering the independence of each part detection, we introduce the overall topology description, which describes not only the pairwise relations among part pairs, but also the global topology relation. For pedestrian topology description, we adopt the OC-SVM classifier. OC-SVM is usually used in cases that samples are almost from one class, and OC-SVM can construct as small hyper-sphere as possible that contains as many samples as possible in high dimensions of feature space, which describes a binary model of probability distribution of the samples. The pedestrian topology sample set can be seen from one class, and we can obtain the

training model by OC-SVM to evaluate each training sample in sample set, and ultimately perform pedestrian detection.

We test our method in the public data sets of INRIA and TUD-Brussels, and experiments show that the topology descriptor is robust to occlusion and multi-postures. Our method obviously reduces the miss rate, which leads to less false detection compared with other methods.

KEY WORDS: Pedestrian Detection, Topology Description, Part detection, Classifier, One-Class SVM

目 录

摘 要	I
目 录	IV
图目录	VI
表目录	VII
第一章 绪论	1
1.1 课题背景和研究意义	1
1.2 国内外研究现状及发展趋势	2
1.3 行人检测中存在的问题	5
1.4 本文的研究内容	5
1.5 本文的组织结构	7
第二章 行人检测相关工作综述	9
2.1 目标检测的一般框架	9
2.2 特征表示	9
2.2.1 Haar-like特征	10
2.2.2 LBP特征	10
2.2.3 HOG特征	12
2.3 行人检测的代表性方法	13
2.3.1 基于整体模型的方法	13
2.3.2 基于部件模型的方法	16
2.4 行人检测数据集	19
2.5 本章小结	20
第三章 行人部位检测模型	22
3.1 HOG特征表示	22
3.2 行人部位分割与表示	24
3.2.1 行人部位分割	24
3.2.2 旋转部位检测	25
3.3 实验结果与分析	27
3.4 本章小结	28
第四章 行人拓扑描述与建模	29
4.1 基于拓扑模型的检测	29
4.2 拓扑关系描述	30
4.2.1 拓扑结构	31

4.2.2 拓扑描述子	32
4.3 OC-SVM方法	34
4.3.1 OC-SVM基本思想	34
4.3.2 基于OC-SVM的建模	36
4.4 实验结果与分析	41
4.4.1 实验数据	41
4.4.2 实验效果及对比分析	42
4.5 本章小结	46
总结与展望	47
参考文献	49
发表(录用)文章目录	55
致 谢	57

图目录

图 1-1 图像/视频目标检测的发展	1
图 1-2 基于整体模型机统计学习的目标检测一般框架	3
图 1-3 基于部位检测和关系模型的目标检测一般框架	3
图 1-4 行人检测研究框架	6
图 2-1 目标检测一般框架	9
图 2-2 Haar-Like特征	10
图 2-3 拓展的Haar-Like特征	10
图 2-4 LBP特征计算示意图	11
图 2-5 HOG特征块	12
图 2-6 最大间隔分类器	14
图 2-7 HOG-SVM 检测方法流程图	15
图 2-8 自适应的分类器组合算法示意图	17
图 2-9 基于贝叶斯推断的组合算法中行人部件定义示意图	18
图 2-10 基于判别训练部件的算法示意图	19
图 3-1 梯度向量直方图特征示意图	22
图 3-2 行人部位分割示意图	24
图 3-3 部位搜索示意图	25
图 3-4 部位旋转示意图	25
图 3-5 头部和腿部训练样本示例	26
图 3-6 腿部旋转检测结果	27
图 3-7 行人部位检测结果示例	28
图 4-1 基于拓扑的行人检测方法流程图	31
图 4-2 拓扑关系示意图	31
图 4-3 拓扑结构解释示例图	31
图 4-4 拓扑描述示意图	32
图 4-5 SVM最优分类超平面	34
图 4-6 OC-SVM的二维示意图	37
图 4-7 OC-SVM的超球面分类	38
图 4-8 训练样本示例	41
图 4-9 测试样本示例	41
图 4-10 拓扑模型的行人检测结果	44
图 4-11 实验曲线与对比	45

表目录

表 2-1 非线性迭代偏最小二乘算法	16
表 2-2 行人检测公开数据集	20
表 4-1 实验方法参数对比表	46

第一章 绪论

1.1 课题背景和研究意义

随着人类社会的发展，社会的不安全因素也随之增多。在众多不安全因素中，人为破坏是最为严重和最难于防范的。美国的“911”事件和伦敦地铁爆炸案更是给各国的安全防范部门敲响了警钟。人体目标检测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，自动行人检测技术能够帮助监控交通事故，并可以在十字路口、新建高速路端等进行安全预警；在汽车辅助驾驶系统中，通过车载视频实时分析路况和检测行人变化，能够对危险情况进行提示和报警；随着城市规模的扩大，住宅小区逐渐成为安全防范的重点，结合红外/近红外等视频设备，人体目标检测与自动报警技术可增强小区安全监控、减少犯罪，增强居民的安全感。在监控视频数据库的管理和检索系统中，人体目标及其他典型目标定位可以为语义检索提供支持。在国防领域，夜视探测设备中的人体目标检测和跟踪识别危险目标，能够帮助特种部队及空降兵探测周边区域情况。随着国家越来越重视采用视频系统对重要部门、敏感地点、公共场所等进行监控，高性能的人体目标检测技术呈现出越来越中重要的应用价值。



图 1-1 图像/视频目标检测的发展

在图像理解和计算机视觉研究领域，正面视角的汽车牌照、车辆、人脸等一些典型目标检测方法日趋成熟，检测性能已经能够支持很多应用系统的需求。如图 1-1 所示，这些目标一般近似为刚体，受到形变影响较少。同时，应用系统一般都要求摄像头和被检测目标之间相对位置固定，以保证拍摄的目标具有固定的视角。随着视频信号的采集质量越来越高，监控的视角、范围的日益扩大（如全景摄像机、摄像机网络技术的应用），捕获的人体目标视角更加多样，其中不仅包含了旋转视角，还包含了摄像机俯仰引起的目标角度变化。人体是一个非刚性的物体，在运动中可以呈现多种姿态。人体目标很容易被其他人体或者其他物体遮挡，使得拍摄到的目标丧失整体性，给检测带来很大困难。除此之外，很多应用系统（如：实时视频监控报警系统，车外行人告警系统）对目标检测具有实时性要求。这些因素使得人体目标检测成为计算机视觉领域内最具挑战的问题之一，也成为了图像理解和机器学习方法的理想研究平台。

所谓行人检测(Pedestrian Detection)就是根据人体的某些特征在图像或者图像序列中判断有无人体目标并定位其准确位置。其定位的精度直接影响到相关应用的效果。虽然行人检测方面的研究有了长足的进步，但是仍然没有一个适用于所有情况的通用方法。其主要原因是行人检测作为目标检测的一个特例有其特殊性，例如行人的视角多样性、背景的复杂性、人的身体以及衣着的差异、步态的不确定性、遮挡以及目标大小等因素。以上因素在人体目标模式上反映出的主要问题在于：表观的差异，目标模式的部分缺失，以及在目前的 PC 机上的计算高复杂性。这使得行人检测距离实际应用仍有较大距离。

本研究受到以下课题资助：

- 1) “天目 360 度智能监控系统产业化开发”，中科院研究生院—哈尔滨产业技术创新与育成专项资金项目，2011.01-2013.12，在研；
- 2) “视频中人体目标可靠检性检测方法与稳定性跟踪算法研究”，国家自然科学基金面上项目，60872143，2009.01-2011.12，已结题。

1.2 国内外研究现状及发展趋势

已有的行人检测研究通常可分为基于整体模型的方法和基于部位结构的方法。

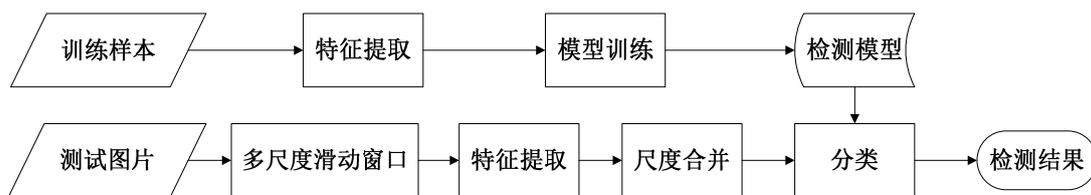


图 1-2 基于整体模型统计学习的目标检测一般框架

基于整体模型的方法将行人作为一个整体进行建模、描述与检测。其思路是利用固定大小的窗口对整幅图像或者感兴趣区域进行滑动窗口（sliding window）扫描，在滑动窗口内提取特征，再利用分类器进行模式分类判断该窗口是否包含人体。其关键问题是选择鲁棒、高效的特征（如轮廓、形状、颜色等）来表述行人，并使用模板匹配或者统计学习方法建模。其典型框图如图 1-2 所示。

Papageorgiou 等人^[1]首次使用了一组完备的 Haar 小波特征结合支持向量机（SVM）分类实现了行人的快速检测。Dalal 等人^[2,3]提出了一种梯度向量直方图（HOG）特征结合 SVM 分类器的行人检测方法。该特征在行人检测主要表征的是行人的轮廓信息，是通过计算局部区域上的梯度方向直方图构成人体特征，在 MIT 等数据库上取得了较好的效果，已经成为行人检测中比较主流的一种特征。在 HOG 的基础上，Zhu^[4]等人提出了变尺度 HOG（v-HOG）特征，相比原始 HOG 特征的固定尺度特性，其对于局部轮廓的变化更加敏感。Xu 在文献[5,6]中将 1 范数最小化作为分类准则应用到人体检测中，提出人体的稀疏表示与分类。Laptev^[7]同样使用了梯度方向直方图特征，通过加权 Fisher 线性判别分析法，将高维的直方图特征投影为 1 维的标量值，然后使用 AdaBoost 和级联分类器技术实现了快速的行人检测。

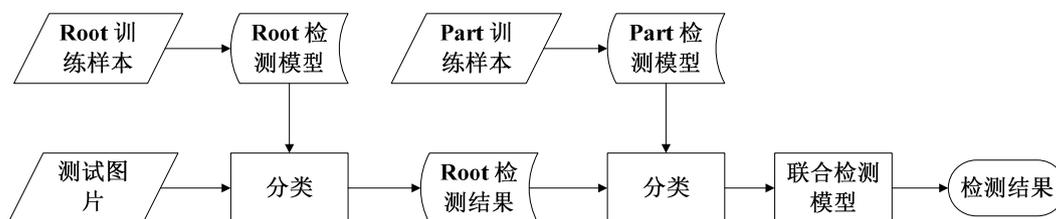


图 1-3 基于部位检测和关系模型的目标检测一般框架

在行人检测中主要的难点来自遮挡、多视角、大的类内差距以及光照变化，而基于部位结构的行人检测模型是可以较好解决以上问题的另一类目标检测解

决方案。基于部位的方法将一个目标建模分为其整体模型和各个部位模型的组合，通过对于不同部分的分别或者协同检测定位人体这样对于解决遮挡目标和多姿态目标等情况是比较有效的。其典型框图如图 1-3所示：

Mohan^[8]等提出了一个基于部位的检测系统。系统由 4 个部位检测器和一个装配检测器组成。4 个部位检测器分别为头，腿，左臂和右臂检测器。实验结果显示这个系统的检测性能比单个基于整体的方法要好，它能够检测出部分被遮挡的人和某些部位跟背景对比度低的人。Lin 在文献[9]中提出了一个将人体分为三个部位的分级的模板匹配检测框架，三个部位分别是，头和躯干，大腿，小腿。文章中扩展了层次匹配模板的思想，将整个人体形状模型分解成不规则部位，构建部位模板匹配树。模型中按照行人各个部位的变化程度来决定每个部位模板的数量，实验表明该方法对于遮挡问题有较好的效果。Wu^[10]提出了基于贝叶斯推断方法，求出最大后验概率解。他将人体分成头肩、躯干、腿部三个部位，分别训练分类器加以检测。在得到所有部位检测结果后，再用贝叶斯推断来分析哪里出现行人。

2009 年以来，Felzenszwal 等人^[11,12]提出了可形变的基于部分的模型，该模型对于一类物体建立多个长宽比不同的模型，首先利用 root 模型进行检测，检测到整体行人后再进行部分定位，最后各部分响应及形变的损失函数判别人体。该方法在 PASCAL2009 竞赛中也取得了最好的成绩。目前，对于人体姿态的研究也是一个热点问题。在文献[13]中，Park 等人将固定模板与基于部位的模型相结合，将检测的尺度扩大，较大的降低了漏检率。在文献[14]中，作者提出了一种基于部位检测和子结构的检测方法，主要是选取部位组合的最小覆盖集来达到检测的目的，可以解决一部分遮挡问题。文献[15,16]中，Andriluka 将人体划分成了十个部位，利用 shape context^[17]特征对不同部位进行检测，在此基础上利用后验概率最大化来判断人体的姿态，并在此基础上做了跟踪的相关实验。Yao 在文献[18]中用同样的特征对人体部位进行检测，提出了一个模型学习的方法对人体不同姿态进行判断，并将交互的物体与人体姿态互为上下文用于姿态识别和检测。

可以看出，人体检测在还有许多提升的空间。基于部位的检测越来越成为研究人员关注的重点。基于部位的人体检测中，对部位特征的选取和部位间关

系模型的构建是两个研究点，而这两方面各自的发展都将为此领域的研究带来提升使得复杂场景下的检测更加鲁棒。

1.3 行人检测中存在的问题

现有的人体检测算法，虽然在一定程度上能够完成对人体目标的正确检测，但是都或多或少的存在一些问题，列举如下：

- 1) 人体目标衣服的形态、颜色、样式的多变性。老人、小孩、男人、女人衣服饰物各不相同，多种多样，是人体检测的难点之一，在很大程度上限制了人体检测算法的误检率和检测精度。
- 2) 检测目标的部分及完全遮挡。遮挡会造成被检测目标特征的缺失，挑战了分类器的容忍度。比如，被部分遮盖了腿部信息的人其实同一个放置在盒子上的盆栽没有本质的区别（就计算机来看）。
- 3) 图片中行人的多尺度问题。由于行人在视频中的位置不同，以及其本身的大小造成了行人在监控场景下的尺度不同，但是现阶段所使用的检测窗口大小固定，这就给扫描时的定位造成了很大的困难。如今采用的主流方法是多尺度检测窗口的方法，虽然有效地解决了这个问题，但是却在无形之中增加了计算量。
- 4) 场景光线亮度变化。同一目标在不同的光线亮度条件下会产生有明显差异的目标图像，成为了是很多室外实际应用系统中阻碍性能提升的瓶颈。
- 5) 背景的影响。现实的背景中存在各式各样、五颜六色、形态各异的物体，例如：路边树木、电线干、栏杆、栅栏等它们都或多或少具有人体的部分特征信息，很容易引起人体检测算法失效。

1.4 本文的研究内容

本文的主要研究旨在提出一种基于拓扑模型的行人检测方法。针对在复杂场景下人体检测中遇到的遮挡问题，补偿单一特征检测的局限性。研究内容如下：

- 1) 人体各个部位的描述与检测：本文通过对行人进行语义分块，将行人分成若干个部位，对每个人体部位进行 HOG 特征提取，训练线性 SVM 分类器，对相应的人体部位样本进行二分类，从而得到对行人部位的检测结果。考虑到行人部位的多样性，尤其是行人呈现不同姿态时，部位产生的平面角度形变，尝试对于该问题提出解决方案，在图片中行人大小多尺度的基础上，对角度进行多尺度处理。
- 2) 人体各个部位之间关系模型的描述：空间关系模型建立是一个对人体的各个部位根据空间关系建模的过程。建立的模型将一个人体样本各个部位之间的拓扑关系表示为一个特征向量，一系列空间关系训练样本的特征向量输入到OC-SVM中进行训练，用于去除部分检测中误检测。研究内容如图 1-4。

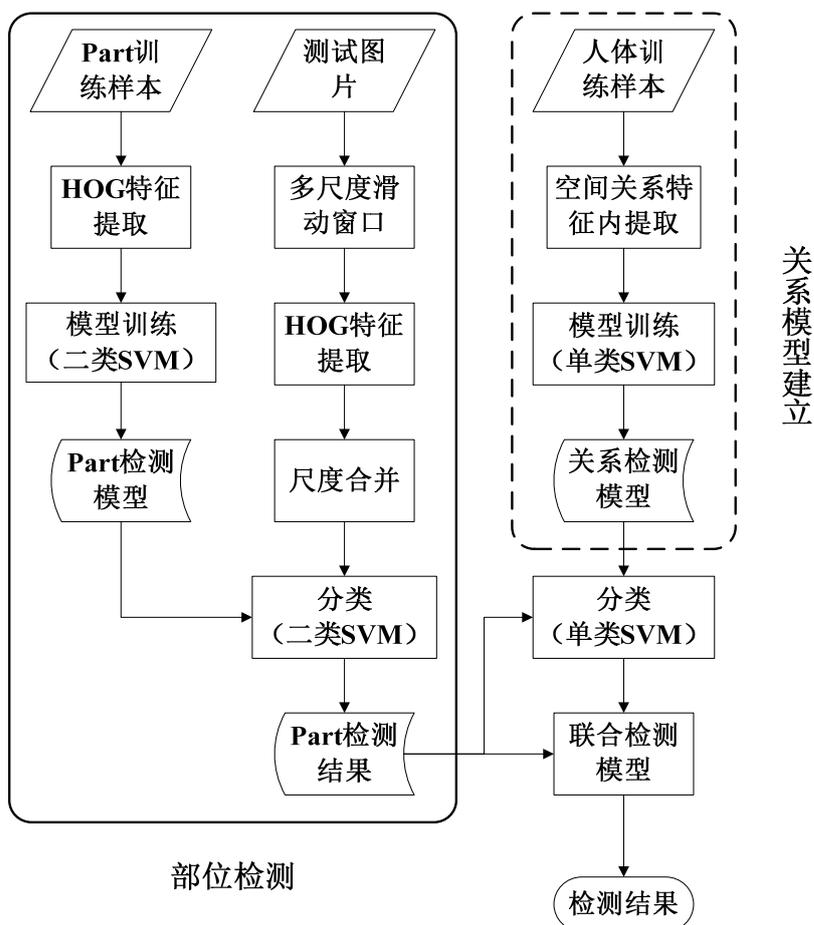


图 1-4 行人检测研究框架

1.5 本文的组织结构

第一章，绪论。论述了图片/视频帧中行人检测的研究背景和意义，阐述了国内外行人检测的研究现状以及发展趋势，分析了行人检测领域中存在的问题，并列出了本文的主要研究目标和研究内容。

第二章，行人检测相关工作综述。综述了行人检测框架、特征表示和代表性研究方法，并列举了行人检测常用的数据库。

第三章，行人部位检测模型。阐述了 HOG 特征的表示，以及行人部位分割与表示策略，同时给出了行人部位旋转检测的方法以及实验结果。

第四章，行人拓扑描述与建模。介绍了行人拓扑关系描述，阐述了行人部位间的空间关系，论述了 OC-SVM 的基本原理及其在行人检测上的应用。在不同数据集上进行了实验验证及对和分析。

结束语总结了本文的主要工作，展望了未来工作的方向。

第二章 行人检测相关工作综述

计算机视觉和模式识别相关会议每年都有人体检测的研究工作发表，提出一些新的研究视点或者改进一些经典算法。本章首先回顾了目标检测的一般框架，然后对一些代表性的工作进行综述。

2.1 目标检测的一般框架

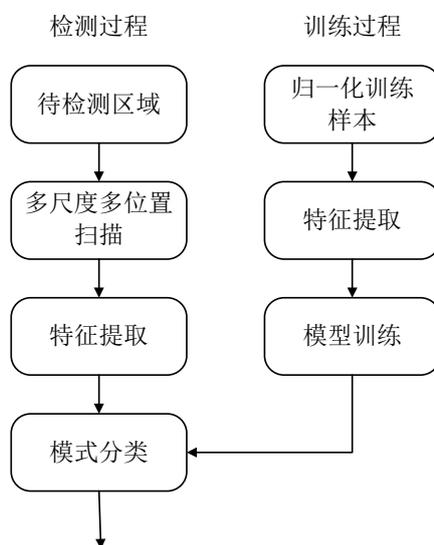


图 2-1 目标检测一般框架

如图 2-1所示，为目标检测的一般框架。一般来讲，目标检测主要分为两个部分：训练和检测。在训练过程中，训练样本需要通过归一化成为同一尺度，在统一的尺度下做特征提取和模型训练；检测过程中，需要对图片进行多尺度缩放，并在图片范围内做滑动窗口扫描，并对滑动经过的每个窗口做相应的特征提取，最后送入训练好的分类器进行分类，得到检测结果。

2.2 特征表示

基于整体模型和基于部件模型的检测方法的关键一环在于特征的选取，尤

其针对于人体的非刚性，特征描述显得尤为关键，对行人整体或部件的检测起着重要作用。下面介绍几种常见的特征描述子。

2.2.1 Haar-like特征

在 1997 年，M. Oren^[20]提出了用小波提取特征，再使用 SVM 来做分类的思想。M. Oren 的合作者 Papageorgiou 等人^[1]，在行人检测中使用两个尺度上的小波函数对图像进行 Haar 小波变换，选择能够表现类内紧致性的小波系数作为人体特征。后来，Viola 等人又对原始的 Haar-Like 特征进行了拓展，实现了对图像更精细的描述。该特征分别在人脸、人体等目标上进行了实验，在人脸检测上的效果好于人体检测的效果^[21,22]。

Haar-Like特征是由Haar小波演变而来的^[1,8,23]。如图 2-2，Harr-like特征的值等于灰色矩形框中所有像素的颜色值和减去白色矩形框中的所有像素的颜色值。(A)和(B)为两个矩形的特征，(C)为三个矩形的特征,(D)为四个矩形的特征。特征计算可通过积分图的计算进行加速^{[1,8][23]}。

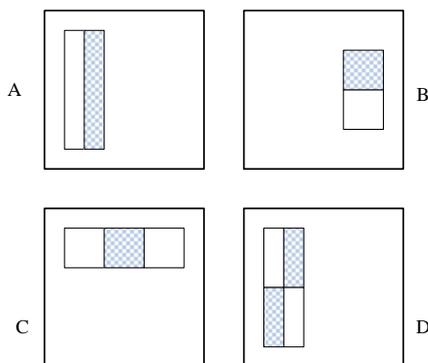


图 2-2 Haar-Like 特征

Viola等人又对上面基本的Haar特征进行了拓展，便有了现在的“拓展的 Haar-Like特征”^[23]，见图 2-3。



图 2-3 拓展的 Haar-Like 特征

2.2.2 LBP特征

局部二值模式LBP (Local Binary Pattern)^[19,25]虽然计算简单，但是由于其

在文理分析方面的强大能力而在目标检测中得到应用。LBP利用图像中某一像素点与周边像素点的灰度差构成了图像的细节纹理。LBP的计算过程如图 2-4 所示：以中心像素的 8 邻域为例，以中心像素的灰度值为阈值，邻域内的每个像素都与中心像素的进行比较，大于中心点的像素标记为“1”，小于的则标记为“0”，这样邻域内的二进制串起来组成了表征该点的LBP二进制串。再将该二进制串乘以他们各位的对应权值就可以得到一个十进制的LBP值。如果我们计算一个点的 P 邻域内的LBP值，计算公式如下所示：

$$LBP_{K,C} = \sum_{k=0}^K S(p_k - p_c) \cdot 2^k \quad (2-1)$$

$$\text{其中, } S(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2-2)$$

其中， p_k 为邻域像素值， p_c 为中心点像素值， 2^k 为对应点的权重。

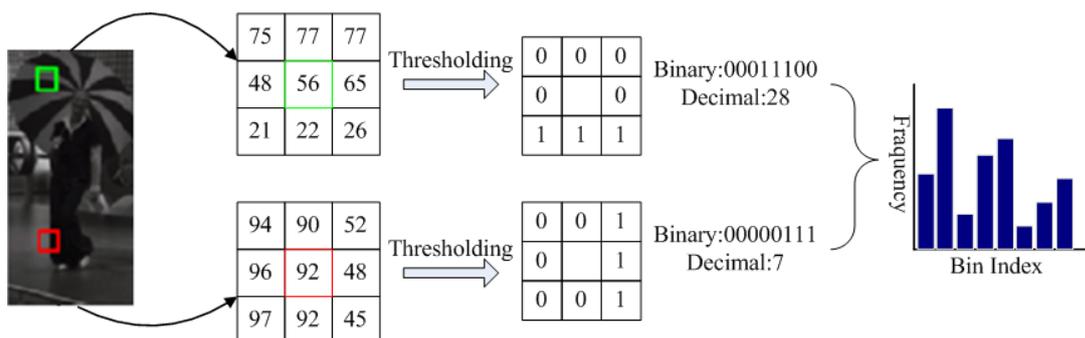


图 2-4 LBP 特征计算示意图

LBP的另外一种统计方式叫做 uniform LBP 特征。与原始的LBP特征相比，uniform 的 LBP 特征是将 K 邻域组成的二进制串首位相连成一个环，记录其中的 0/1 和 1/0 跳变次数。并且只计算跳变次数小于等于 2 次的环为 uniform 的，例如 00000000(0 次跳变)和 11100011(2 次跳变)是 uniform 的，但是 01010000 (4 次跳变) 和 01110101 (6 次跳变) 是非 uniform 的，计算 LBP 值时只计算 uniform 的值，其余的非 uniform 的 LBP 均设为统一的值。此时 LBP 的计算公式变为：

$$LBP_{K,C}^{uniform} = \begin{cases} \sum_{k=0}^K S(p_k - p_c) \cdot 2^k & \text{transformation} \leq 2 \\ \text{constant} & \text{transformation} > 2 \end{cases} \quad (2-3)$$

Uniform LBP 特征在降低特征存储量的同时没有降低检测识别的精度，说明 LBP 特征并不是越多越好，其主要信息保存在 Uniform 部分，而非 Uniform 部分代表噪声。

2.2.3 HOG特征

“Histograms of Oriented Gradients(HOG)由 Dalal 和 Triggs 于 2005 年提出。与 Linear SVM 结合用于目标检测，在 MIT 行人库上性能优异。为进一步测试其在复杂样本上的性能，Dalal 发布了 INRIA 行人样本库^[31]。在 INTRA 库上，HOG+SVM 算法在万分之一的误报率下取得了约 90%的检测率，成为了行人检测的一个里程碑^{[10][11]}。与 SIFT (Scale invariant feature transformation)^[24]特征类似，HOG 特征通过提取局部区域的边缘或梯度的分布，来表征局部区域内目标的形状。由于是在局部区域统计求得，HOG 特征对小的形变和配准误差有较强的鲁棒性。以下为 HOG 特征简介。

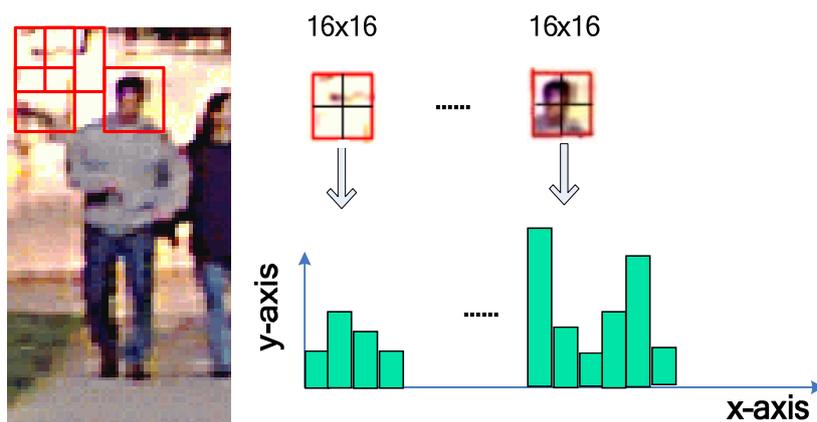


图 2- 5 HOG 特征块

而对于HOG特征而言，如图 2- 5，作者是将 64×128 的训练样本按照 8×8 个像素的小块 (cell) 进行划分，这样就形成了 $8 \times 16 = 128$ 个 cell。然后再将每相邻的 4 个 cell (田字形结构的 4 个 cell) 划分为 1 组 (block)，block 的形成是按照每个 cell 进行滑动生成的，由此原训练样本 (64×128) 便具有 $7 \times 15 = 105$ 个 block。对于每个 cell 都将其中的所有像素的梯度方向进行投影，形成每个 cell 各自的梯度方向直方图^{[2][3]}。这里的方向 bins 的数量被设定为 9，即每 20 度一个 bin，0-180 与 180-360 的方向采用对等角相等的方法进行归类划分。然后，再将每个 block 中的 4 个 cell 的梯度直方图的数据串联起来，由于每个 cell 的梯度直方图为一个

9 维的向量,所以每个block便是一个 36 维的向量。再将所有的block(对于 $64*128$ 像素的样本共 105 个block)依次串联起来,便形成了 $36*105=3780$ 维的特征^[3]。

2.3 行人检测的代表性方法

如 1.2 节所述,行人检测算法通常可以分为基于整体模型的方法和基于部件模型的方法。第一类:基于整体特征的方法。该方法通常是利用一个窗口在空间上和尺度上滑过所有行人可能出现的位置,在滑动窗口中提取如 Haar-like 特征^[20]、梯度向量直方图(HOG)^[3]和局部二值模式(LBP)^[19]等特征,然后利用如支持向量机(SVM)^[57,58], Adaboost^[59,60]、神经网络^[61]等分类器来判定每一个窗口与行人目标的相似度。最后利用非极大值抑制(Non Maximum Suppression, NMS)的方法获取目标在图像中的位置。该类方法将行人看作一个整体,而忽略了人的非刚性,因而在处理人体姿态,遮挡和视角的问题上显现的不够灵活。由于行人目标具有表现的多样性,因此该类方法的重点是寻找一个能够较好描述行人轮廓的特征。第二类:基于部件模型的方法。该方法首先检测出行人的各个部件,其次判别各个部件之间的关系是否满足行人的分布关系,从而判别该目标是否为行人。

2.3.1 基于整体模型的方法

基于整体特征的行人检测方法将行人看作一个整体来处理,忽略了人的非刚性^[42-47]。本小节以 HOG-SVM 方法和偏最小二乘算法(PLS)为代表,介绍基于整体模型的方法。

2.3.1.1 HOG-SVM检测方法

HOG 特征在 2.2.1 节已做了简单介绍,下面主要阐述一下支持向量机 SVM 的原理,以及 HOG-SVM 方法在行人检测中的应用。

SVM 由 Vapnik 首先提出 SVM 能被用作机器学习、模式识别、非线性回归等方面,是一种基于结构风险最小化准则的学习方法^[37,38,39,40],根据所采用核函数的不同,可分为“线性 SVM”和“非线性 SVM”两种^[19]。

线性 SVM 的核心思想是建立一个超平面作为决策曲面,使得正例和反例之间的分隔边界被最大化。SVM 是结构风险最小化的近似的实现方法。如图 2-6,

其中黑点表示+1，白点表示-1。

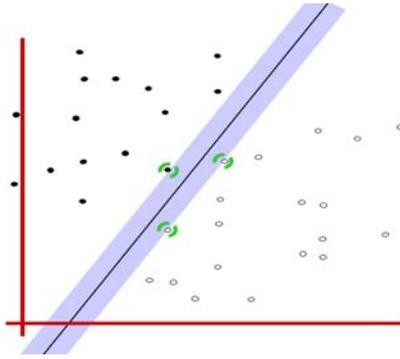


图 2-6 最大间隔分类器

对于线性可分的样本 $\{x_i, y_i\}, i=1, \dots, N$ 。其中 x_i 为样本特征向量, $y_i \in [+1, -1]$, 其分类面方程为 $|w^* x_i + b| = 0$ 。将分类面方程归一化, 使两类样本满足 $|w^* x_i + b| \geq 1$, 这样分类间隔为 $2/\|W\|$ 。要使分类间隔最大, 即要 $\|W\|$ 最小, 并要求样本满足:

$$y_i[(w^* x_i) + b] - 1 \geq 0 \quad (2-4)$$

那么, 求解最优超平面问题就可以表示成如下的约束优化问题: 即在式(2-4)的约束条件下, 求解最小化函数:

$$\phi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w^* w) \quad (2-5)$$

为了解决约束优化问题, 引入 Lagrange 函数:

$$L = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i y_i (x_i^* w + b) + \sum_{i=1}^n \alpha_i \quad (2-6)$$

其中 $\alpha_i > 0$ 为 Lagrange 系数, 现在的问题就是要求 L 的最小值。把式(2-6)分别对 W 、 α_i 和 b 求偏微分, 并令其等于 0, 就可以把上述问题转换为一个较简单的对偶问题。即在约束条件下, 对 α_i 求解下列函数的最大值:

$$\begin{cases} Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j \\ \sum_{i=1}^N \alpha_i y_i = 0, C \geq \alpha_i \geq 0 \end{cases} \quad (2-7)$$

式 (2-7) 中： C 为使用者所选的正参数。它控制机器的复杂度和不可分离点数之间的平衡，也可以被看作是一种“正则化”参数的形式。

对于样本本质上是非线性可分的情况，我们可以采取将特征空间映射到多维空间上去，使样本达到线性可分的状态，并且不改变样本的主体性质。因此引入了“核函数”的概念。

如图 2-7，在行人检测中，HOG特征针对行人的表征提取人体轮廓特征，采用对每个像素点求梯度的方法，利用统计学思路描述人体的整体轮廓。在具备合适的行人数据库后，SVM分类器对于具有 3780 维HOG特征的正反例样本进行训练，得到能够区分行人与非行人的超平面，从而对待检测的样本图片正确分类。

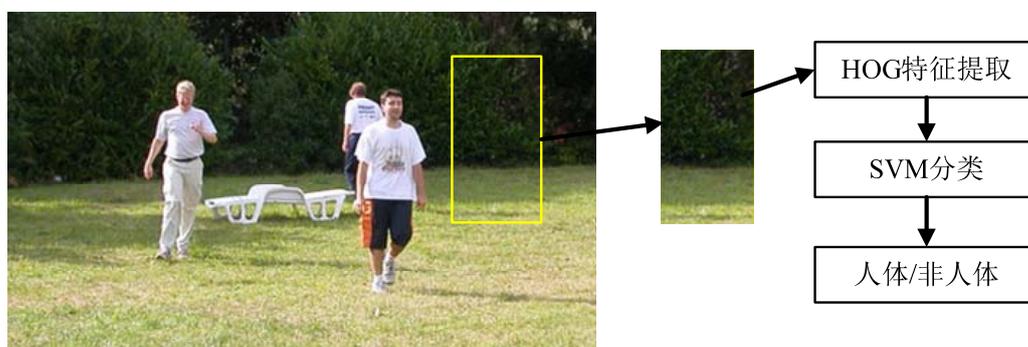


图 2-7 HOG-SVM 检测方法流程图

2.3.1.2 偏最小二乘算法 (PLS)

偏最小二乘法是一种新型的多元统计数据分析方法，它于 1983 年由 S.Wold 和 C.Albano 等人首次提出。其主要思想是通过隐变量之间的联系来构建多个数据集之间的关系。偏最小二乘算法存在一个基本假设，观测数据是由一个只受几个隐变量影响的系统或者过程产生。这个假设是偏最小二乘算法用来处理回归和分类问题，作为降维和建模工具的基础^[62]。

最小二乘回归算法的特点是：

- 1) 能够在自变量存在严重多重相关性的条件下进行回归建模；
- 2) 偏最小二乘回归更易于辨识系统信息与噪声；
- 3) 允许在样本点少于变量个数的条件下进行回归建模。

假设两个变量 X 和 Y ，偏最小二乘算法首先对变量 X 和 Y 进行零均值化，

然后对它们进行分解并从中同时提取成分，其公式如下所示：

$$\begin{aligned} X &= YP^T + E \\ Y &= UQ^T + F \end{aligned} \quad (2-8)$$

其中 T 和 U 是提取得到隐变量或者成分， P 和 Q 表示稀疏矩阵， E 和 F 表示残差。因此我们就可以通过计算 T 和 U 之间的关系得到 X 和 Y 之间的关系。偏最小二乘算法可以通过非线性迭代偏最小二乘算法（Nonlinear iterative partial leastsquares algorithm）构建一个权重向量 $W = \{w_1, w_1, \dots, w_p\}$ 和 c 去描述 T 和 U 之间的关系，其形式如下：

$$[cov(t, u)]^2 = [cov(Xw, Yc)]^2 = \max_{|r|=|s|=1} [cov(Xr, Ys)]^2 \quad (2-9)$$

其中， $cov(t, u) = t^T u / n$ 表示隐变量 t 和 u 的样本方差。在求解过程中，非线性迭代偏最小二乘算法首先随机初始化 u ，然后按照表 2-1 的算法进行迭代，直到收敛。

表 2-1 非线性迭代偏最小二乘算法

- 1) $w = X^T u / u^T u$
- 2) $\|w\| \rightarrow 1$
- 3) $t = Xw$
- 4) $c = Y^T t / t^T t$
- 5) $\|c\| \rightarrow 1$
- 6) $u = Yc$

2.3.2 基于部件模型的方法

在行人检测中主要的挑战大多来自于遮挡、多视角、大的类内差距以及光照变化，而基于部件结构的行人检测模型是可以较好解决以上问题方法^[48,49,50]。该类方法首先检测出行人的各个部件，然后将不同部件的信息综合起来得到最后的检测结果。在行人部件的定义上通常有两类分法：第一类将行人分为几个部件^[10,26,27]，如：头、左臂、右臂和腿等。另一类是按照图像本身的特性将行人分为很多图像块^[28,29,30]，而这些图像块可能没有明显的物理意义。以下部分综述了分类器组合、基于贝叶斯推断的组合和基于判别训练部件的算法。

2.3.2.1 自适应的分类器组合算法

Mohan等人^[8]在2001年提出了一个基于样本的行人检测框架，自适应的分类器组合算法（Adaptive Combination of Classifier）。如图2-8所示，该算法将行人分为头、腿、左肩和右肩四个部分，且由四个部件检测器和一个装配监测器组成。首先是训练各个部件的分类器，在输入图像中检测出各个部件。其次是训练组合分类器，该分类器以部件检测的分类器输出作为输入，得到最后的检测结果。实验结果显示该方法的性能由于部分基于整体的方法，能够检测出部分遮挡和人的某些部件和背景对比度低的人，同时也带来了检测耗时的问题，因此，在没有提速算法的前提下，需要在精度和效率之间做权衡。

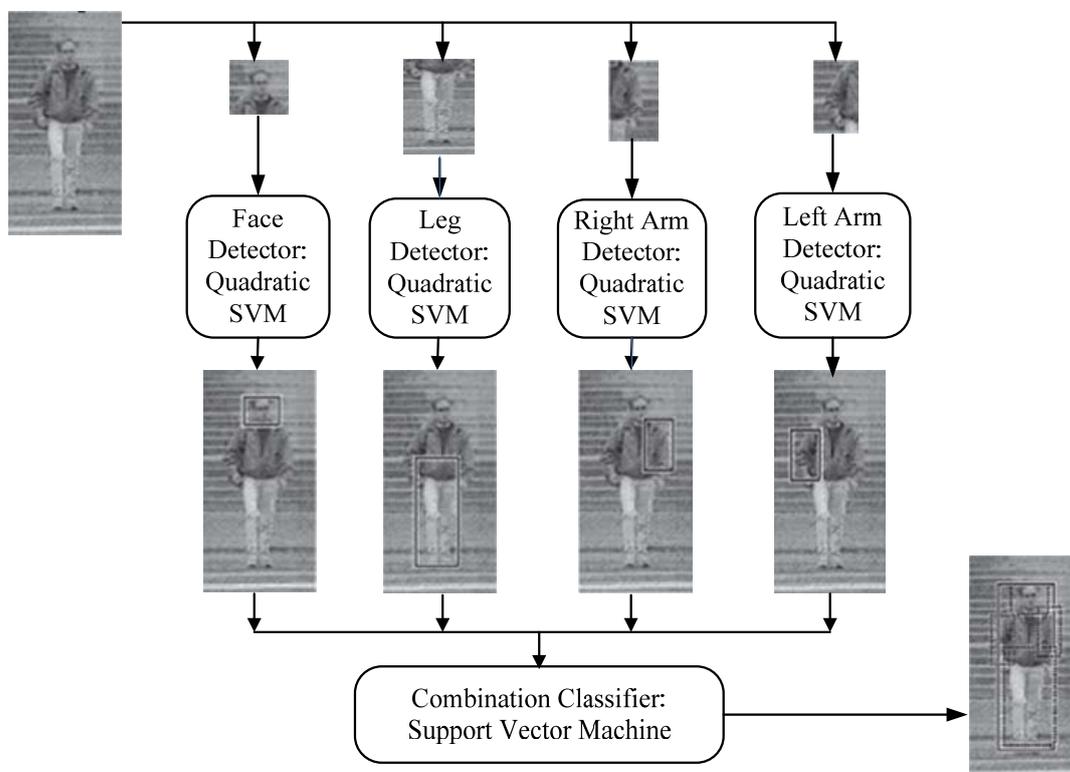


图 2-8 自适应的分类器组合算法示意图

2.3.2.2 基于贝叶斯推断的组合算法

Mohan的自适应的分类器组合算法中，利用分类器学习行人各部件之间的关系。Bo Wu^[10,25]在提出Edgelet特征的同时，提出了从贝叶斯推论的角度去寻找各部件之间的关系。如图2-9所示，Wu将行人分为了头肩、躯干和腿三个部分。利用事先训练好的分类器检测出各个部件，再通过贝叶斯推论来分析，哪

个位置为行人目标。基于贝叶斯推断的算法公式如下所示：

$$p(I|S) = p(Z^{FB} | P^{FB})p(Z^{HS} | S^{HS})p(Z^T | S^T)p(Z^L | S^L) \quad (2-10)$$

$$p(Z^{HS} | S^{HS}) \propto \prod_{Z_i^{HS} \in SD} p_{SD}^{HS} p(Z_i^{HS} | S_i^{HS}) \prod_{Z_i^{HS} \in FA} p_{FA}^{HS} \prod_{Z_i^{HS} \in FN} p_{FN}^{HS} \quad (2-11)$$

其中， $Z = \{Z_i; 1 \leq i \leq N\}$ 是实际检测结果， $S = \{S_i; 1 \leq i \leq K\}$ 表示我们假设存在的行人。FB、HS、T、L分别表示全身（Full Body）、头肩（Head Shoulder）、躯干（Torso）和腿（Leg）。人的后验概率是各部件及整体的似然概率的乘积。所以该算法的目标就是找到符合条件的S使 $P(Z|S)$ 最大。在图 2-9中，Wu假设各部件的结果是相互独立的，并且每个部件的检测结果可以有三种情况即检测成功SD（Successful Detection）、误检测FA（False Alarm）和漏检测FN（False Negative）。头部检测的似然概率与给定部件的条件概率、成功检测、误检测和漏检测概率相关。通过实验发现该联合检测方法不仅对单个人体的检测效果很好，而且能够处理人群拥挤的场景。但是，该方法只能检测近似正面或背面的人体，人体还必须是直立或行走的。

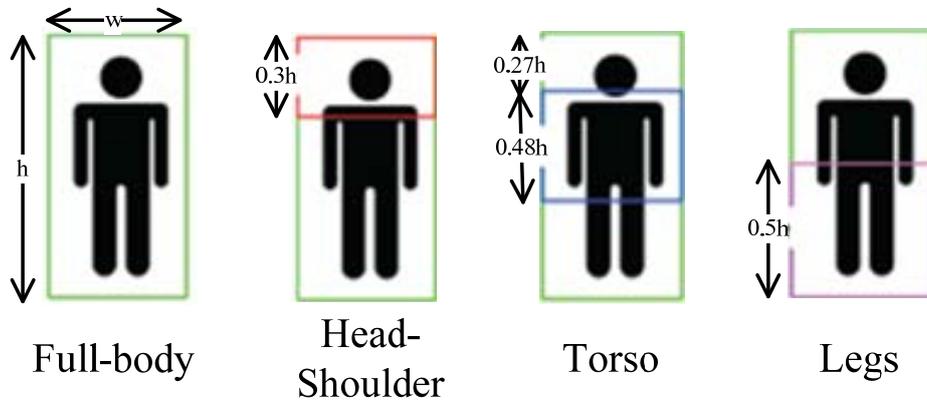


图 2-9 基于贝叶斯推断的组合算法中行人部件定义示意图

2.3.2.3 基于判别训练部件的模型算法

近几年，的基于判别训练部件的模型（Discriminatively Trained Part Based Models）取得了巨大的成功。其扩展算法赢得了 PASCAL VOC 目标检测竞赛的冠军，并且被成功运用到分割、基于静态图像的行为分类等领域。

基于判别训练部件的模型算法示意图如图 2- 10所示，该方法由一个包含

整个行人的根（root）滤波器和一些包含行人部件的局部（part）滤波器组成。其中局部滤波器所处的特征层的分辨率比根滤波器的分辨率高两倍。通过滤波以后得到整个行人和行人部位的反应，对局部特征相对于窗口所在的位置加入惩罚项，能够有效的描述行人部件在一定范围内的变化。通过结合整体以及局部部件结构，克服只有整体不能解决的形变、部分遮挡问题，表现出良好性能。

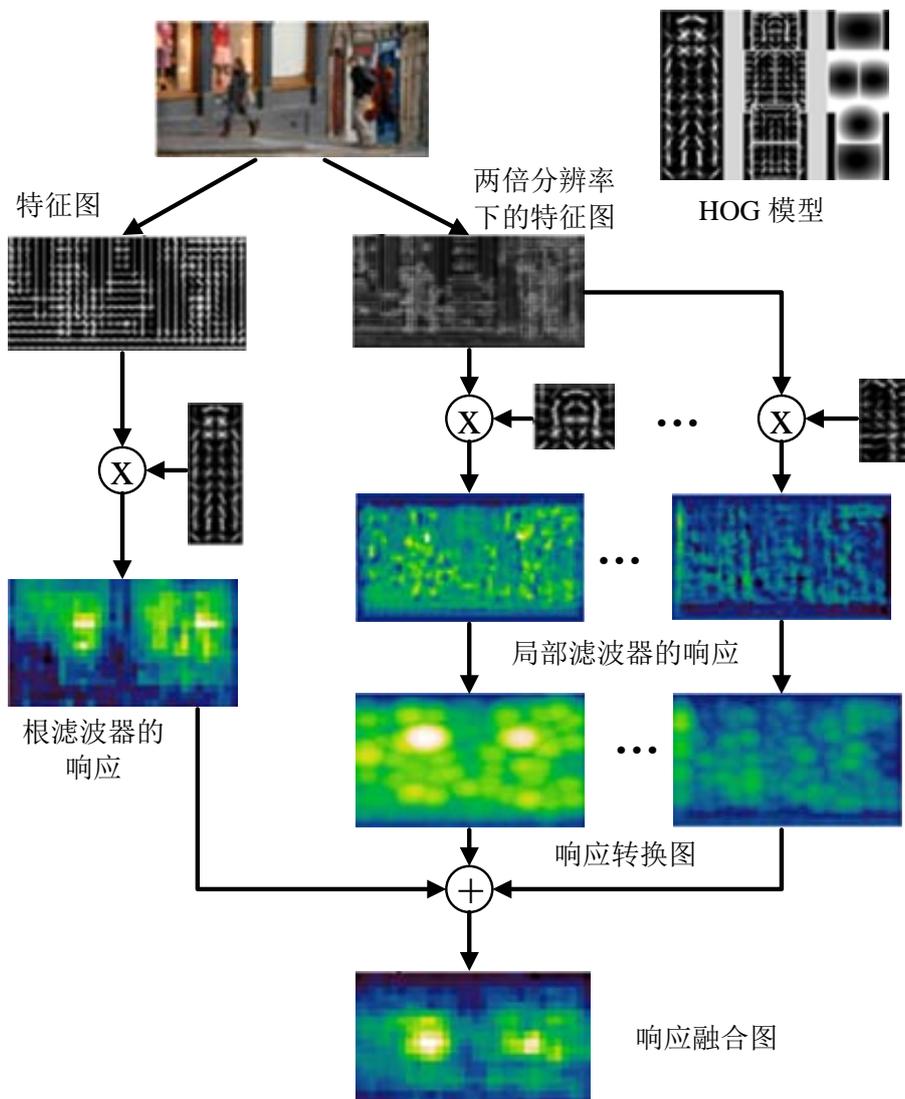


图 2-10 基于判别训练部件的算法示意图

2.4 行人检测数据集

表 2-2 列出了现有的一些主要行人检测数据库。可以看出根据图像获取装置的不同，所有的数据库大致可以分为两类：静态图像和视频。INRIA 数据库

自 2005 年发布以来，一直备受关注，推动了行人检测技术的快速发展。中国科学院研究生院模式识别与智能系统实验室 (Pri-SDL) 也发布了自己的行人数据库，供研究者训练和测试。PASCAL VOC 2007 数据库是近年来目标检测领域最具影响力和挑战性的数据库，最优秀的行人检测算法在这个数据库上的成绩也不理想，说明了行人这一类可形变的目标检测还有很大的研究空间。INRIA 数据库中行人目标的分辨率相对较高，而通过车辆上的移动摄像机拍摄的数据库目标相对较小，行人高度一般处在 50-100 个像素之间，从一个侧面反映了低分辨率情况下行人检测的迫切需求，尤其是针对移动行人检测平台的应用。

表 2-2 行人检测公开数据集

数据集	训练集	测试集	注解
INRIA People Dataset ^[31]	2416(64*128) / 1218(含标定)	1132(64*128) / 453(含标定)	1) 彩色图片 2) 背景复杂
Caltach Pedestrian Dataset ^[32]	250,000 帧, 350,000 bounding boxes 2300 行人 (标定)		1) 驾驶员视角 2) 视频帧序列
TUD-Brussels Pedestrian Dataset ^[33]	508 幅 640*480 的图片, 共 1326 个行人		多尺度、多视角的行人
VOC2007 dataset ^[34]	正样本 4096 / 负样本 3003	4528 行人	1) 彩色图像 2) 行人检测最具挑战和影响
TUD Multi-view Pedestrians Dataset ^[35]	4732 幅彩色图片 (含标定数据)	250 幅彩色图片	3) 图片尺寸不一 4) 含部位标定 5) 共 8 个视角
SDL Pedestrian Dataset ^[36]	正面 1000 幅 / 侧面 3050 幅 / 多视角 7550 幅	140 幅彩色图片 / 258 幅多视角图片	包含正面/侧面/多视角 64*128 像素 (训练)

2.5 本章小结

本章介绍了现有行人检测的相关工作，包括检测的一般框架，常用的特

征描述子、代表性方法与数据集合。可以看出，基于部件模型的算法相对于整体模型研究时间要短，但检测效果更加出色。基于整体特征的检测方法致力于寻找一个有效的整体特征来描述行人。而基于部件模型的方法对于遮挡，形变比较鲁棒，在各类数据库上取得了更好的效果。

第三章 行人部位检测模型

随着当代社会尤其是智能视觉监控对于行人检测与识别的迫切需求，每年的计算机视觉和模式识别的相关会议都有很多优秀的文章涌现，它们分别从不同的角度切入研究或者深入的改进一些经典算法，提高检测的精度。基于部位的检测方法已经成为行人检测领域研究的热点，并有了一定的突破。基于部位的检测关键在于部位检测模型以及部位间关系模型的构造，本章主要阐述行人部位的特征，部位分割与表示，以及旋转检测。

3.1 HOG特征表示

Dalal和Triggs^[3]在 2005 年首次提出了梯度向量直方图（Histogram of Oriented Gradient: HOG）特征，HOG是一种用于表征局部边缘梯度分布特性的描述符，其示意如图 3- 1 梯度向量直方图特征示意图 所示。其计算步骤如下：

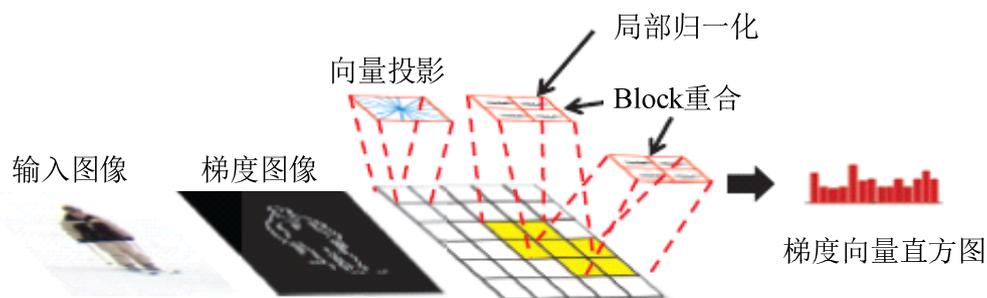


图 3- 1 梯度向量直方图特征示意图

第一步：梯度值计算。许多特征提取前都需要对图像进行预处理，例如颜色归一化和伽马校正，但是在 HOG 特征中不需要进行这一步操作，因为后续的归一化过程可以替代这些过程。在计算梯度图像时通常可以采取利用形如 $[-1; 0; 1]$ 和 $[-1; 0; 1]^T$ 的模板对图像进行滤波就可以获得图像上下和左右方向的梯度值。

第二步：构建块描述子。由于在 HOG 中将待检测窗口划分成一些相互重叠的块，每个块又被分成几个小单元。单元中的每一个像素的梯度值都被投影到直方图的某一个 Bin 上。在计算过程中单元可以是矩形也可以是圆形的，Bin 的角度范围根据梯度是否有方向性可以是 0 到 180° 或者 0 到 360°。统计每个单元内所有像素的梯度值在各个 Bin 上的分布就可以得到描述局部边缘梯度分布的单元梯度直方图向量。为了使 HOG 特征对于光照和对比度的变化鲁棒，就需要对梯度值进行局部归一化，所以将每个块内的单元串联在一起并进行归一化得到块描述子。

第三步：块归一化。Dalal 和 Triggs^[12]在文中介绍了四种：L1-norm、L1-sqrt、L2-norm 和 L2-hys。其公式如下所示：

L1-norm:

$$f = \frac{v}{\|v\|_1 + e} \quad (3-1)$$

L1-sqrt:

$$f = \sqrt{\frac{v}{\|v\|_1 + e}} \quad (3-2)$$

L2-norm:

$$f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \quad (3-3)$$

L2-hys: 在 L2-norm 的基础上进行截断（将 v 的最大值限制在 0.2）再归一化。其中 v 为待归一化向量， $\|v\|_k$ 是该向量的第 k 范式，e 是一个小常量。通过 Dalal 和 Triggs 的实验表明经过 L2-Hys、L2-norm 和 L1-sqrt 归一化后效果相差不多，而 L1-norm 的效果相对较差。但是经过归一化后 HOG 特征的检测效果都得到大大的提升。

通过以上三步的操作可以得到梯度向量直方图特征。该特征的优点是：首先由于 HOG 特征的操作是在局部的单元上进行的，并且图像的几何和光度形变通常只会出现在比单元更大的空间领域上，因此对图像的这两种形变可以保持良好的不变形。其次，在满足粗空域抽样、精细的方向抽样以及较强的局部光度归一化的条件下，只要行人大体保持直立姿态，一些细微的肢体变化都不

会影响检测的效果。通过比较可以发现，HOG 与 Lowe^[24]提出的 SIFT 的差异并不是很大，SIFT 特征的提取主要是在一张图像的兴趣点上进行，而 HOG 不再依赖于兴趣点的检测。在 INRIA^[26]数据库上，HOG 也在万分之一的误报率的情况下取得了近 90% 的检测率。

3.2 行人部位分割与表示

3.2.1 行人部位分割

基于部位的人体检测很重要的一个部分是对人体各个部位关系模型的构建，这里考虑从人体部位之间的空间关系入手。首要任务便是如何将行人分割为若干个部位。在图 3-2 中，我们把行人分为头①，左肩②，右肩③，左躯干④，右躯干⑤，大腿⑥，左脚⑦和右脚⑧共 8 个部位，寻找各个部位之间的空间关系。根据各个部位之间的拓扑关系（下一章将会详细阐述拓扑关系的构建策略），存在 1-2,1-3,2-4,3-5,4-6, 5-6,4-7,5-8,6-7,6-8，共 10 对节点间有拓扑联系。经验告诉我们，人体检测中，头部和躯干的遮挡概率比腿部的遮挡要小，为了模型构建的合理性，避免某个部位丢失或错检造成的严重影响，又加入 4-7,5-8 两组关系，具体描述见下一章的内容。因此，当构成人体的 8 个部位节点有任何一个被遮挡或者检测丢失时，其他部位节点都可以跟其他至少一个部位保持关联，这样可以在一定程度上解决某些遮挡问题，而且比起两两关系遍历的 28 对关联少了 9/14，运算上更简单。

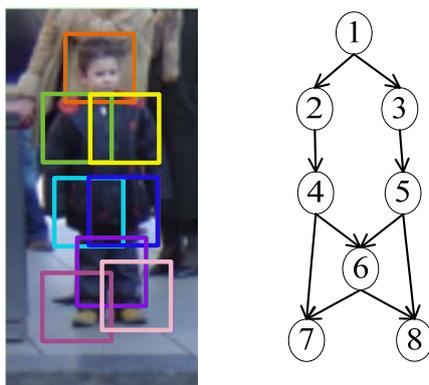


图 3-2 行人部位分割示意图

行人部位分割的策略为：在候选窗口中搜索 HOG 特征加权响应最大的部

位，其中，搜索框的大小固定，为预定义的部位块的大小。当搜索到候选窗口中响应最大的部位时，将该部位位置标记，并继续搜索响应次之，且与已标记的部位块重合不大于 50% 的块，依次迭代，直到找到 8 个部位块为止。部位块的搜索示意如下：

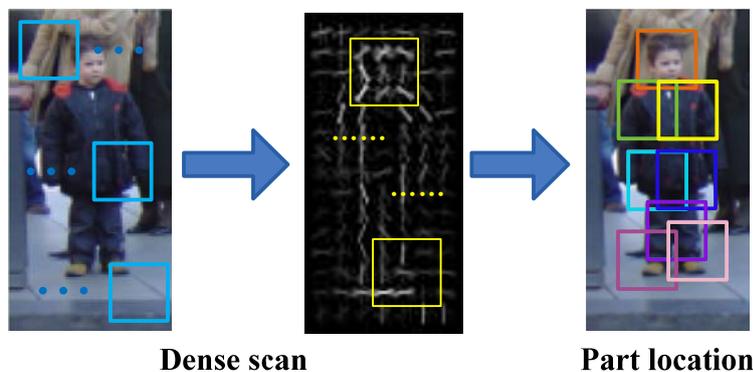


图 3-3 部位搜索示意图

3.2.2 旋转部位检测

在不同场景中，行人的姿态变化多种多样，由于人体的非刚体特性，行人的四肢在图像中的变化是最多的，尤其是四肢围绕关节的旋转。这一特性使得行人检测面临多姿态的问题，也是行人检测领域一直在研究解决的问题。这里，我们针对行人部位的形变旋转问题做一些初步的探讨。

我们知道针对行人检测中不同尺度行人检测的解决方案普遍是利用图像的多尺度缩放来处理。受该思想的启发，对于肢体的旋转，我们借用类似的思路，采用对图片的多尺度旋转策略处理部位的旋转问题。

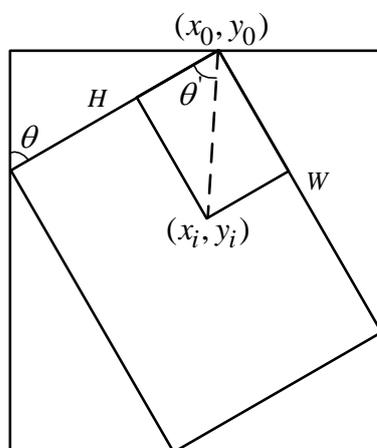


图 3-4 部位旋转示意图

如图 3- 4, 中间倾斜的矩形为行人某部位的样本框, H 为样本的高度, W 为样本的宽度, 其外接矩形是样本旋转 θ 角度后的矩形框, 所有计算单位都以像素计。 (x_i, y_i) 是旋转后的大图中的任意一点图像坐标, (x_i', y_i') 是旋转之前原图中对应 (x_i, y_i) 的点的坐标, 也即 (x_i', y_i') 是相对原始图像的坐标, 即所求。根据图像中的存在的几何关系, 有以下计算:

$$x_0 = H \cdot \sin \theta \tag{3-4}$$

$$y_0 = H \cdot \cos \theta \tag{3-5}$$

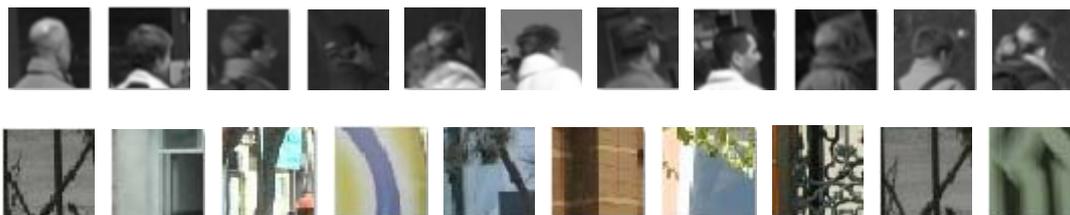
$$\theta' = 90 - \arctan \frac{y_i - y_0}{x_i - x_0} + \theta \tag{3-6}$$

$$x_i' = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \cdot \sin \theta' \tag{3-7}$$

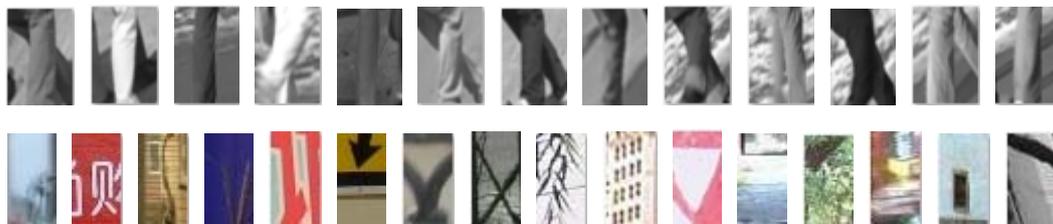
$$y_i' = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \cdot \cos \theta' \tag{3-8}$$

因此, 我们能够根据行人样本框的旋转角度, 确定样本框任意一点 (x_i, y_i) 在旋转了 θ 角度之后在原始图像上的新坐标 (x_i', y_i') 。这样, 就得到了多尺度旋转下的行人部位检测。

下面给出一些行人部位检测的训练样本示例图:



(a) 头部训练样本, 正反例样本示例



(b) 腿部训练样本, 正反例样本示例

图 3-5 头部和腿部训练样本示例

如图 3-5, 给出了头部和腿部的训练样本示例, 其中, 头部训练样本大小为 40×40 像素, 腿部大小为 32×48 像素。此外, 躯干训练样本大小为 72×72 像素。所有正例训练样本都是取自 INRIA 训练集; 反例训练样本是再 SDL 行人检测数据集中的反例样本上根据需要的像素随机抠出的反例图片。



图 3-6 腿部旋转检测结果

如图 3-6, 为以腿部为示例的旋转检测结果。我们可以看到, 在利用上述多尺度角度的匹配策略对行人部位在角度上按一定步长遍历, 可以得到较为理想的效果, 但也会存在一些误检和丢失的情况, 如图中雨伞上沿也被检测为腿部。实验中, 将多尺度旋转的参数设在 $10^\circ - 45^\circ$ 之间, 分别采用 10° 、 20° 、 30° 和 45° , 在精度上 10° 为步长的检测结果较好, 但是检测时间明显会比较慢, 由于部位旋转检测不仅仅有多尺度角度的遍历检测, 还包括多尺度缩放的遍历检测, 因此其时间复杂度会比单一的多尺度缩放要高。另外, 需要特别强调的是, 在进行多尺度检测之后, 窗口合并策略是值得重视的环节, 否则会有较多误检测产生。

3.3 实验结果与分析

这一节中, 我们给出了行人部位检测的方法在一些实验数据集上的检测效果, 以及定性的分析。

如图 3-7, 给出了行人部位检测的结果示例, 其中包括不同姿态和光照条件的行人。我们将行人分为 8 个部位, 对应检测结果中, 深绿色对应头部, 黄色和蓝色对应左右肩, 砖红色和粉色对应躯干的左右部分, 亮绿色对应大腿,

浅紫色和紫罗兰色对应左右脚。在实验结果中可以看出，行人部位的检测结果基本集中在行人边缘轮廓较为明显的区域中，如头肩和腿部。这样的结果与利用HOG特征描述人体部位及部位搜索策略的思想一致。同时，从结果中还能够发现，行人各个部位的检测不会多数存在较大的偏差，利用这个结论，我们可以在行人部位拓扑关系方面做进一步的研究，在下一章会详细阐述。



图 3-7 行人部位检测结果示例

3.4 本章小结

本章介绍了将行人划分为若干个部位，并利用部位的表示进行行人检测。首先介绍了用来描述行人部位的 HOG 特征及其提取规则和流程；接着对行人部位的分割与表示做了详细阐述，将行人分为 8 个部位，分别是头，左右肩，左右躯干，大腿和左右脚，另外还对部位的形变问题进行讨论，采用多尺度旋转的策略予以解决；最后，给出了行人部位检测的实验结果与分析。

第四章 行人拓扑描述与建模

视角和姿态是行人检测当中的难点,在本章中,通过将行人分为若干部位,着重对部位之间的拓扑关系进行描述和分类,以此达到对行人检测中多视角、多姿态问题的改进。对基于部位检测的方法而言,由于一般情况下各个部位检测器的鉴别性较差,检测器通常会产生大量的误检。另外,由于遮挡和/或部分检测器的不完备,可能造成部位的丢失。这些情况都对现有的基于部位的检测器提出了挑战。为了处理误检,通常需要考虑部位内部的几何约束。

本章首先介绍行人拓扑关系建模的框架,然后阐述如何对行人部位间的拓扑关系加以描述,以及采用的 OC-SVM 分类方法的基本理论及在行人检测中的应用。

4.1 基于拓扑模型的检测

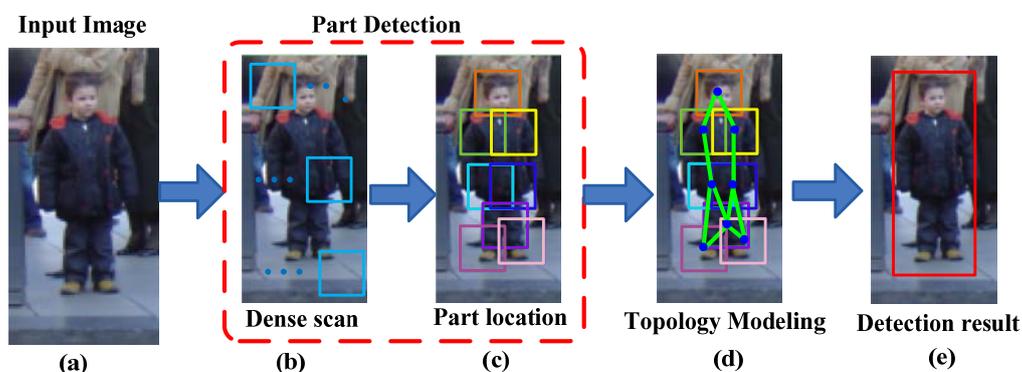


图 4-1 基于拓扑的行人检测方法流程图

如图 4-1,给出了基于拓扑的行人检测方法流程图。图中,(a)为输入的待检测样本,(e)为输出的样本检测结果,其中(e)中红色的矩形框为检测到的行人在图片中的具体方位,同时输出的还包括检测结果框的角点在图像坐标中的具体像素坐标,以及检测得到的分数。(b-c)为行人的部位检测,包括对每一个部位独立的检测,之后得到各个部位响应最大的位置作为部位检测的中

间结果，如 (c) 所示。(d) 中，在部位检测的基础上，对行人部位之间的关系进行了拓扑描述，图中的绿色边代表生成的拓扑结构，该拓扑结构在本章中采用提出了一个新的描述子Log-polar Topology Pattern来表示，拓扑特征提取后，再利用OC-SVM分类器对拓扑进行分类，之后结合部位检测的结果得到待检测样本在一定的阈值下是否被判别为行人并输出。

在公式 (4-1) 中， $S(x)$ 为行人检测结果的总分数，其中 $\sum_{i=1}^P \alpha_i \cdot F_{i,HOG}$ 表示部位检测的响应值，而 $\beta \cdot F_{LTP}$ 表示行人拓扑的响应值，参数 $F_{i,HOG}$ 为行人第 i 个标号部位的 HOG 特征描述子， α_i 为该部位 HOG 特征抽取后训练得到的权重； F_{LTP} 为行人拓扑描述子 LTP 特征， β 为拓扑权重。这里我们将每个行人定义为 P 个部位，在后面的章节中将会详细介绍。得到公式 (4-1) 中的检测分数 $S(x)$ 后，在公式 (4-2) 中，我们利用实验得到的阈值 τ 对行人进行判别，当 $S(x) < \tau$ 时， $f(x) = 0$ ，表示检测结果不是行人；否则，认为该待判别样本是行人，并将其在图片中标出，给出具体的图像坐标。

$$S(x) = \sum_{i=1}^P \alpha_i \cdot F_{i,HOG} + \beta \cdot F_{LTP} \quad (4-1)$$

$$f(x) = \begin{cases} 0, & \text{if } S(x) < \tau \\ 1, & \text{otherwise} \end{cases} \quad (4-2)$$

4.2 拓扑关系描述

本文的研究中，我们将行人分为 8 个部位，每一个部位都各自用HOG特征描述，在部位检测过程中得到行人 8 个部件相应最大的作为部件的检测输出。如图 4- 2所示，部位检测的结果（即标注的矩形框，不同颜色表示不同部位的检测结果），对应到拓扑关系为：拓扑中每个节点都与相应的部位检测结果一一对应。以部位检测结果的中心点的图像坐标作为拓扑对应的节点坐标，按照图 4- 2中的拓扑关系，将各个节点连接，构成包含 8 个节点，10 条有向边的拓扑关系图。下面介绍一下选择这样拓扑关系的理由。

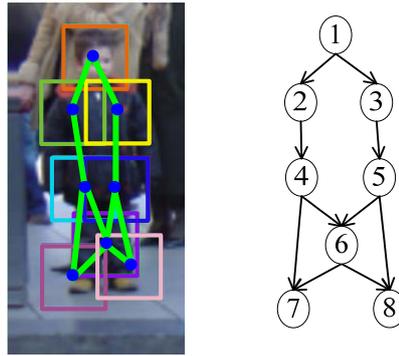


图 4-2 拓扑关系示意图

4.2.1 拓扑结构

由于在行人的部位检测阶段得到的结果不存在漏检，也即每一个部位在搜索的范围内最后输出的是响应最大，也是最相关的结果，这使得不论检测到的“部位”是否是完全正确的结果都会有一个局部最优解输出，不存在缺失。拓扑关系模型也正是基于这一点，对行人全局进行描述，根据行人本身的结构拓扑，以及根据既定规则选取的不可缺省的关系，一并作为行人的拓扑描述。

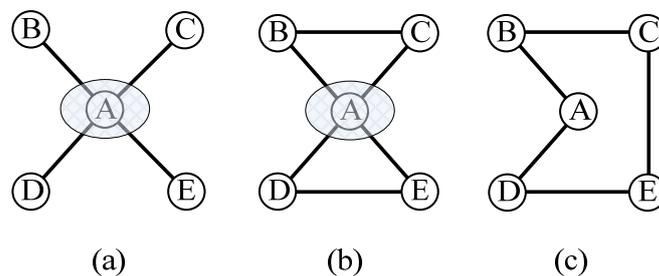


图 4-3 拓扑结构解释示例图

如图 4-3 所示，对一个包含 5 个部位的检测目标，(a) 包含 4 条无向边，采用星型结构。我们可以看出，当部位 A 丢失或者错检时，所有的 4 条边都没有了实际意义，无疑当出现这种情况的时候，这类拓扑是很失败的。(b) 中，相对于 (a) 而言，多出了 B-C 和 D-E 两条无向边。当部位 A 丢失或者错检时，仅仅有着两条边是有效的，边的关系的有效描述只有 1/3，显然也不是一个好的策略。为了避免以上两种拓扑关系带来的问题，也即，拓扑描述过分的依赖于某一个“关键”节点，而导致的拓扑描述的有效性大大降低的问题，进而我们提出了类似 (c) 中的方案。

针对拓扑模型过分依赖某个“关键”节点的问题，考虑选取 (c) 中的方法。

当一个包含 5 个部位的检测目标中，有任意一个存在较大的检测偏差，或者存在丢失误检现象时，(c) 中描述的拓扑关系只会缺少一对有意义的连接，对于某个节点不会存在过分依赖。这样的策略能较好的保证拓扑的稳定性和鲁棒性。

在对行人整体进行描述时，结合以上原理以及行人固有的拓扑结果，抽取了 10 对部位关系，如图 4- 3所示。在后面的实验结果与分析中，该拓扑结果体现了良好的效果，对于行人的多姿态以及遮挡问题都有较好的处理。

4.2.2 拓扑描述子

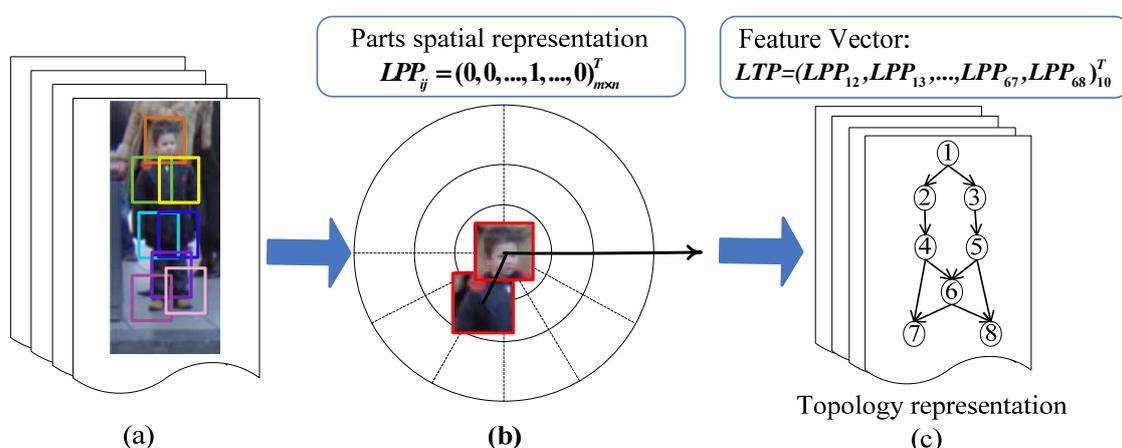


图 4- 4 拓扑描述示意图

(a) 训练样本集，(b) LPP 描述子示意图，(c) LTP 描述子示意图

在行人部位检测的基础上，我们需要对检测得到的各个部位进行拓扑关系描述，从而得到拓扑模型。如图 4- 4所示，(a) 为部位检测的中间结果，作为拓扑训练的样本，其中，行人的 8 个部位都被检测出来，并且在原始图片的图像坐标系下有具体位置数据及部位标签；(b) 是LPP描述子的定义，表示两个不同部位间的关系描述；(c) 为得到了有拓扑关系的所有部位的LPP描述之后，由LPP描述子构建LTP描述子的示意图，LTP描述子是针对行人整体拓扑的描述，刻画行人的全局拓扑。下面我们分别对LPP描述子和LTP描述子做更为详细的介绍。

LPP: Log-Polar Pattern描述子的简写。这里的Log-Polar是指我们的方法中是采用对数极坐标为参考系，以便于对行人部位间的空间拓扑关系进行更恰当

的描述。如图 4-5, LPP描述子是针对两个独立且有拓扑关联的部位的空间关系进行描述。这里, 我们选取每个有效部位检测结果的中心点, 作为该部位的唯一位置坐标(图像坐标)。将拓扑结构中单向边的起点作为对数极坐标的极点, 以该极点为起点指向另一个部位中心点的方向的单向连线作为极轴, 以此方式, 即可构建一个对数极坐标系。其中, 极角和极轴用来表示两个有拓扑关系的部位之间的空间关系。

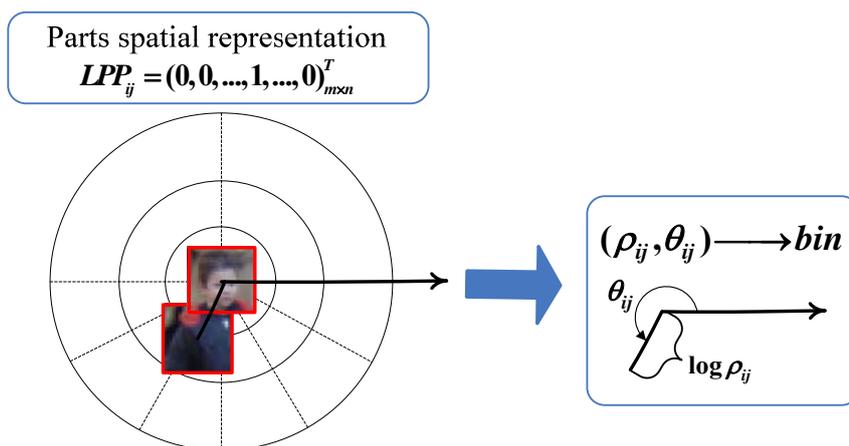


图 4-5 LPP 描述子

假设在对数极半径 $\log \rho$ 下有 m 个单元, 极角 θ 下有 n 个单元, 则这两个部位间的拓扑关系即可被表示为一个 $m \times n$ 维的二值向量。这里, 我们采用在图像坐标中计算欧氏距离的方法(见公式(4-3)), 计算部位 i 和部位 j 的空间位移 ρ_{ij} , 极角 θ_{ij} 可利用公式(4-4)计算出来。其中, X_i 和 Y_i 为部位 i 在图像坐标中的中心点在 X 和 Y 轴的坐标值, 参数 λ 取值 1,2, 是为了限制极角的范围在 $[0, 2\pi)$ 。 $\max(\rho_{ij})$ 表示各个部位之间空间距离最大的两个部位间的距离。

$$\rho_{ij} = \sqrt{|X_i - X_j|^2 + |Y_i - Y_j|^2} / \max(\rho_{ij}), i \neq j \quad (4-3)$$

$$\theta_{ij} = \arctan \frac{X_i - X_j}{Y_i - Y_j} + \lambda\pi, \lambda = 1, 2, \theta_{ij} \in [0, 2\pi) \quad (4-4)$$

当计算得到了 ρ_{ij} 和 θ_{ij} 后, 我们将 (ρ_{ij}, θ_{ij}) 投影到对数极坐标中预设的 $m \times n$ 个单元中, 置该单元为 1, 其余为 0。

LTP: Log-Polar Topology Pattern描述子的简写。在得到了行人有拓扑关系的各个部位间的LPP描述子后, 我们可以利用多个LPP描述子构建整个行人的拓

扑描述子，即LTP。我们结合行人固有的拓扑关系，如：头肩连接，左右腿关系等，以及 4.1.1 中详细阐述的拓扑选取方案，就被划分为 8 个部位的行人选取了 10 个单向连接，从而构成整个行人的拓扑关系，并利用上述的LPP描述子对每个单向连接进行描述。最后，按照图 4- 2中行人部位的标号将 10 组LPP连接，构成描述行人整体拓扑的LTP描述子。

4.3 OC-SVM方法

SVM（Support Vector Machine）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效果及对样本的宽松要求，广泛的应用于视频目标的检测与识别中。其中，OC-SVM方法于2001年由Scholkopf et al.^[27]在支持向量机(SVM)的基础上提出，但是它与 SVM 有很大的不同，主要体现在：其不要求大量的正例反例样本，且主要应用只包含一类样本信息中。OC-SVM 的基本思想是在特征空间中估计一块体积最小的部分，包含尽可能多的相关样本，同时保持较大的分类间隔。正是因为 OC-SVM 对样本的特殊要求，它可以应用到行人拓扑结构模型的分类中，通过对正样本集中的样本进行评价得出重构系数来重构目标，从而完成拓扑分类，并取得了较好的分类效果。

4.3.1 OC-SVM基本思想

基于统计学习的 VC 维理论和结构化风险最小化原理的支持向量机(SVM)是统计学系理论中较年轻的内容，它能够在有限的样本信息的前提下，在模型的复杂性和学习能力之间衡量，以获得最好的推广能力。

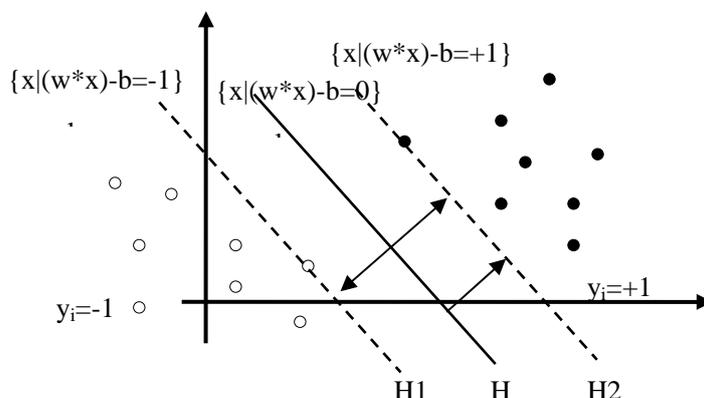


图 4- 1 SVM 最优分类超平面

SVM 的基本思想是:通过非线性变换 $\Phi(i)$ 将输入空间映射到一个高维特征空间中,将输入空间中线性不可分的数据集映射后变换为线性可分,然后在特征空间中求取最大的间隔超平面 $f(x) = w^T \Phi(x) + b$, 其中 w , b 分别是这个超平面的权值和和阈值。

如图 4-1, 空心和实心点分别代表正例、反例样本, H 为分隔超平面, H_1 和 H_2 为两个距离最远且平行于分隔超平面的超平面, H_1 和 H_2 之间的距离叫做最大分类间隔。

设训练样本集为 $\{(x_1, c_1), (x_2, c_2), \dots, (x_n, c_n)\}$, $i = 1, 2, \dots, n$, $x \in R^d$, $c \in \{+1, -1\}$, 其中 c_i 为 1 或 -1, 分别表示数据点属于正例、反例。 x_i 是表示一个样本的 n 维向量, 为了防止变量随机影响分类结果, 每个元素都被缩放到区间 $[0, 1]$ 或 $[-1, 1]$ 。设分隔超平面的数学表示为 $w \cdot x - b = 0$, w 是垂直于分隔超平面的向量, b 是间隔距离的位移。与最优分隔超平面平行, 且距离支持向量最近的超平面, 可由方程表示为:

$$\begin{cases} w \cdot x - b = +1 \\ w \cdot x - b = -1 \end{cases} \quad (4-5)$$

若训练数据样本线性可分, 则存在两个超平面, 其间没有样本点且距离最大, 通过几何计算可知两个超平面之间距离为 $2/\|w\|$, 最小化 $\|w\|^2$ 才可以得到最大的距离。此外, 为了保证超平面间隔之间没有样本数据点, 要确保所有的样本满足如下条件:

$$w \cdot x - b \geq 1 \text{ 或者 } w \cdot x - b \leq -1 \quad (4-6)$$

$$\text{即 } c_i(w \cdot x_i - b) - 1 \geq 0, 1 \leq i \leq n \quad (4-7)$$

经过转换后, 寻求最优分隔超平面的问题就变为在 $c_i(w \cdot x_i - b) - 1 \geq 0, 1 \leq i \leq n$ 的约束条件下最小化 $\|w\|$ 的问题, 也就是说转化为一个二次规划最优化的问题:

$$\min \|w\|^2 / 2, \text{ s.t. } c_i(w \cdot x_i - b) - 1 \geq 0, 1 \leq i \leq n \quad (4-8)$$

这样, 我们就可以得到最优分类面或最大间隔分类面, 其中, 在分类面 H_1 和 H_2 上的训练样本点称为支持向量(Support Vector, SV)。

若样本数据点是线性不可分的，则可以在公式 3-4 的条件 $c_i(w \cdot x_i - b) - 1 \geq 0, 1 \leq i \leq n$ 中增加一个松弛项 $\xi_i, i=1, 2, \dots, n$ ，同时在求解中加入松弛变量，要求 $\|w\|^2 / 2 + c \sum_{i=1}^n \xi_i$ 的最小值，这样公式 3-4 可以转化为：

$$\min \|w\|^2 / 2 + c \sum_{i=1}^n \xi_i, \text{ s.t. } c_i(w \cdot x_i - b) - 1 + \xi_i \geq 0, 1 \leq i \leq n \quad (4-9)$$

折衷考虑最小误差样本和最大分类间隔，则可以得到最优分类面，其中 $c > 0$ 是一个常数，对控制样本分类正确度惩罚的程度。

求解上面的二次规划，我们可以得到最优分类超平面函数为：

$$f(x) = \text{sgn}\{w^* \cdot x - b^*\} = \text{sgn}\left\{\sum_{i=1}^n \partial_i y_i (x_i \cdot x) - b^*\right\} \quad (4-10)$$

该公式中仅包含待分类样本与训练样本中的支持向量机的内积运算，因此，特征空间中的最优线性分类问题则转化为求解空间中的内积运算。

总之，SVM 专门针对在有限样本信息情况下，在模型的复杂性和学习能力之间寻求最优解，从而获得最好的学习能力。与 SVM 用于两类的分类问题不同，OC-SVM 主要是应用于密度估计和聚类。One-Class 概念在 1993 年 Moya 和 1997 年 Ritter 都有涉及过，而被引入到机器学习概念中则是在 1995 年由 Bishop 引入。OC-SVM 的基本思想是训练样本只有一类，标签为 1 或 -1，即用一种训练样本训练分类器，再用分类器测试标签为 1 和 -1 的测试样本。

Scholkopf^[27]等学者提出了两种办法来解决一类分类问题：

1) 通过核函数将数据映射到特征空间中，然后将这些数据同原点以最大的间隔分开。实际上，这就是一个 two-class 分类问题，负样本中的唯一元素就是原点，而所有的训练数据都是正样本；

2) 通过核函数将数据映射到特征空间中，寻找到一个超球面，尽可能包含所有的训练数据，并且球体越小越好；

这两种解决方法分别对应于 OC-SVM 的普通形式和球面形式，两种形式是完全统一的，但是针对于不同的应用场景和样本数据。在本文中，我们使用 OC-SVM 球面形式进行行人拓扑的学习和检测。

4.3.2 基于OC-SVM的建模

在行人拓扑建模中，我们采用OC-SVM方法建模。其中，正例样本为利用

4.2.1 中提到的方法提取的行人部位间拓扑描述。OC-SVM的基本思想就是假设某一类样本数据在特征空间中具有一定的概率分布，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。Schoelkopf等提出的OC-SVM巧妙的利用原点作为负类的代表，通过最大化原点和目标数据间的最小几何间隔 $\rho/\|w\|$ ，来寻找最优超平面 $w \cdot x - \rho = 0$ (其中 w 是超平面的法向量， ρ 是原点到超平面函数间隔)，使超平面尽量远离原点，从而最小化大部分目标数据所在的正半空间。如图 4-6 所示：

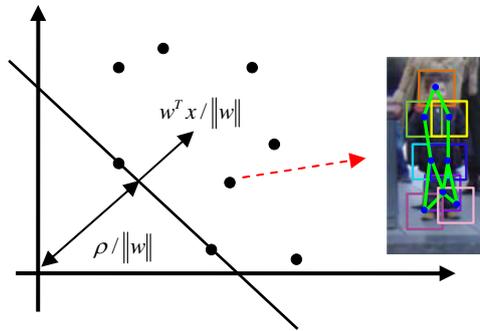


图 4-6 OC-SVM 的二维示意图

考虑数据集 $\{x_1, x_2, \dots, x_n\}, x_i \in \mathbb{R}^d, i=1, \dots, n$ ，记 Φ 是一个从输入空间到某个内积空间的特征映射，且有核函数 $K(x, y) = \Phi(x) \cdot \Phi(y)$ ，并引入松弛变量 $\xi_i, i=1, 2, \dots, n$ 。为了将数据与原点尽量分开，使 $\rho/\|w\|$ 最大化，OC-SVM 通过求解下面的优化问题实现：

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + \frac{1}{\nu n} \sum_i \xi_i - \rho, \\ \text{s.t.} \quad & (w \cdot \Phi(x_i)) \geq \rho - \xi_i, \xi_i \geq 0 \end{aligned} \quad (4-12)$$

其中， ν 是百分比估计，是边界支持向量的上界，是全部支持向量个数的下界。其相应的拉格朗日函数为：

$$L(w, \rho, \xi, \partial, \beta) = \frac{1}{2} \|w\|^2 + \frac{1}{\nu n} \sum_i \xi_i - \rho - \sum_i \partial_i (w \cdot \Phi(x_i) - \rho + \xi_i) - \sum_i \beta_i \xi_i \quad (4-13)$$

其中， $\partial, \beta \geq 0$ 是拉格朗日乘子。通过对拉格朗日函数 $L(w, \rho, \xi, \partial, \beta)$ 关于

w, ρ, ξ 求偏导，可以得到：

- 1) 系数 ∂_i 满足约束 $\sum_i \partial_i = 1, 0 \leq \partial_i = \frac{1}{vn} - \beta_i \leq \frac{1}{vn}$
- 2) 法向量 w 是训练样本集中各向量的线性组合： $w = \sum_i \partial_i \Phi(x_i)$

将上式代入到原始的拉格朗日函数（式）得到其对偶形式：

$$\begin{aligned}
 L(w, \rho, b) &= \frac{1}{2} \|\omega\|^2 + \frac{1}{vn} \sum_i \xi_i - \rho - \sum_i \partial_i (w \cdot \Phi(x_i) - \rho + \xi_i) - \sum_i \beta_i \xi_i \\
 &= \frac{1}{2} \sum_{i,j} \partial_i \partial_j k(x_i, x_j) - \sum_{i,j} \partial_i \partial_j k(x_i, x_j) \\
 &= \frac{1}{2} \sum_{i,j} \partial_i \partial_j k(x_i, x_j)
 \end{aligned} \tag{4-14}$$

于是得到 OC-SVM 原始优化问题（公式 3-8）的对偶问题：

$$\min \frac{1}{2} \sum_{i,j} \partial_i \partial_j k(x_i, x_j), \tag{4-15}$$

$$\text{s.t. } \sum_i \partial_i = 1, 0 \leq \partial_i \leq \frac{1}{vn}, i = 1, 2, \dots, n \tag{4-16}$$

通过经典的二次规划软件包或者序列最小优化算法，求解该二次拉格朗日乘子 ∂_i ， $\partial_i > 0$ 对应的样本点 x_i 为支持向量 SV，法向量 w 可以通过支持向量得到：

$$w = \sum_i \partial_i \Phi(x_i) = \sum_{i \in SV} \partial_i \Phi(x_i) \tag{4-16}$$

决策函数通过下式得到：

$$f(x) = \text{sgn}(w \cdot \Phi(x) - \rho) = \text{sgn}(\sum_i \partial_i k(x_i, x) - \rho) = \text{sgn}(\sum_{i \in SV} \partial_i k(x_i, x) - \rho) \tag{4-17}$$

其中 ρ 可以通过某个支持向量 x_i 求得：

$$\rho = w \cdot \Phi(x) = \sum_i \partial_i k(x_i, x) = \sum_{i \in SV} \partial_i k(x_i, x) \tag{4-18}$$

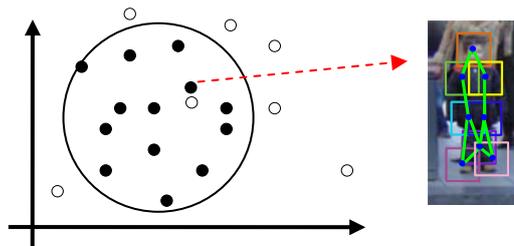


图 4-7 OC-SVM 的超球面分类

前面的理论是将one-class问题转化为two-class分类问题来解决,然后寻求一个能够将样本分开的由支持向量所支撑的最优超平面。下面我们寻找一个能够包含某类全部(或几乎全部)样本在内的由支持向量所支撑的最紧致超球面,如图4-7。

假设某单类样本集 $\{x_1, x_2, \dots, x_n\}$, 其中, $x_i \in \mathbb{R}^d, i=1, \dots, n$ 为样本特征向量, n 为样本数。希望找到一个几乎包含所有样本点 x_i 的体积最小的球。该球由球心 c 和最小半径 R 来描述。与SVM类似, 其最初的带有约束条件的优化问题为:

$$\begin{aligned} \min \quad & R^2 + \frac{1}{\nu n} \sum_i \xi_i, \\ \text{s.t.} \quad & \|\Phi(x_i) - c\|^2 \leq R^2 + \xi_i, \xi_i \geq 0 \end{aligned} \quad (4-19)$$

其中, ξ_i 是松弛变量, $\nu \in [0, 1]$ 控制着球的体积与被拒绝在该球之外的样本数目之间的平衡关系。 ν 越小意味着球的体积越大, 较少的样本被拒绝在该球之外; ν 越大意味着球的体积越小, 较多的样本被拒绝在该球之外。根据KKT条件, 引入拉格朗日系数 ∂, β , 其拉格朗日函数为:

$$L(R, c, \xi, \partial, \beta) = R^2 + \frac{1}{\nu n} \sum_i \xi_i - \sum_i \beta_i \xi_i - \sum_i \partial_i (R^2 + \xi_i - ((x_i \cdot x_i) - 2(c \cdot x_i) + (c \cdot c)))$$

其中 $\partial_i, \beta_i \geq 0$ 为拉格朗日乘子。由式对 R, c, ξ 的偏微分为0, 可导出如下约束条件:

$$\begin{aligned} \frac{\partial L}{\partial R} &= 2R(1 - \sum_i \partial_i) = 0 \Rightarrow \sum_i \partial_i = 1 \\ \frac{\partial L}{\partial c} &= -\sum_i 2\partial_i (x_i - c) = 0 \Rightarrow c = \sum_i \partial_i x_i \\ \frac{\partial L}{\partial \xi_i} &= \frac{1}{\nu n} - \partial_i - \beta_i = 0 \Rightarrow 0 \leq \partial_i \leq \frac{1}{\nu n} \end{aligned} \quad (4-20)$$

由此, 式3-14可转换成如下相对简单的对偶问题:

$$\begin{aligned} \min \quad & \sum_i \partial_i (x_i \cdot x_i) - \sum_{i,j} \partial_i \partial_j (x_i \cdot x_j), \\ \text{s.t.} \quad & \sum_i \partial_i = 1, 0 \leq \partial_i \leq \frac{1}{\nu n}, i=1, 2, \dots, n \end{aligned} \quad (4-21)$$

其中, ∂_i 是Lagrange乘积因子, 只有一部分是非零的, 每个非零的Lagrange乘积因子 ∂_i 对应的点就是通过训练支持向量机得到支持向量(SV)。决策函数

为:

$$f(x) = R^2 - \sum_{i,j \in SV} \partial_i \partial_j (x_i, x_j) + 2 \sum_{i \in SV} \partial_i (x_i, x) - (x, x) \quad (4-22)$$

同样，大多实际问题在原始空间中并不是球形分布的。引入一个映射 Φ 来解决非线性二次规划问题。 Φ 将原始空间的样本数据映射高维特征空间，通过在特征空间进行线性建模来实现原始空间的非线性建模问题。实际的算法是不需知道映射 Φ ，而是通过满足 Mercer 的核函数来实现。所以式可重写为：

$$\begin{aligned} \min \quad & \sum_i \partial_i K(x_i \cdot x_i) - \sum_{i,j} \partial_i \partial_j K(x_i \cdot x_j), \\ \text{s.t.} \quad & \sum_i \partial_i = 1, 0 \leq \partial_i \leq \frac{1}{vn}, i = 1, 2, \dots, n \end{aligned} \quad (4-23)$$

使用一类支持向量机对目标样本集进行训练，解决上述的二次规划问题。高维特征空间中的超球心为：

$$c = \sum_i \partial_i x_i \quad (4-24)$$

高维特征空间中的超球半径为：

$$R^2 = \sum_{i,j \in SV} \partial_i \partial_j K(x_i, x_j) + 2 \sum_{i \in SV} \partial_i K(x_i, x) - K(x, x) \quad (4-25)$$

高维特征空间中的超球面函数为：

$$f(x) = R^2 - \sum_{i,j \in SV} \partial_i \partial_j K(x_i, x_j) + 2 \sum_{i \in SV} \partial_i K(x_i, x) - K(x, x) \quad (4-26)$$

只要一种运算满足 Mercer 条件，它就可以作为这里的核函数使用，核函数 $K(x, x)$ 的选择是构造支持向量机的关键所在，常用的的核函数有

- 1) 线性核函数: $K(x, y) = x \cdot y$;
- 2) 多项式核函数: $K(x, y) = (x \cdot y + \theta)^q, q = 1, 2, \dots, \theta, q$ 为调整参数, q 为多项式次数;
- 3) 高斯径向基核函数: $K(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$;
- 4) Sigmoid 核函数: $K(x, y) = \tanh[b(x \cdot y) - c]$ 。

在 4.2 节中，介绍了 OC-SVM 的行人拓扑样本输入，我们就可以使用 OC-SVM 对样本进行分类，获得每个检测样本的特征向量。与线性二类 SVM 相比较，OC-SVM 使用一类的数据，从密度估计和聚类的角度出发，构建一个远离原点的超平面或者超球面。这个超球面对于数据分布的反应，是使用支撑

向量的近似估计，虽然只是一定程度上反应了数据的情况，但是给我们提供了一些重要的信息（如球心）可以用于行人拓扑的分类，对行人拓扑建模，这点完全不同于以往的 SVM 在该领域上的应用^[51-56]。

4.4 实验结果与分析

本节将提出的行人拓扑模型应用在行人检测中进行实验和对比分析。

4.4.1 实验数据

实验中，我们选择INRIA数据集[26]作为训练集，用来训练行人拓扑模型，供 2416 个正例样本，样本尺寸 64×128 像素。实验性能的测试上，我们采用了 2 个测试数据集，一个是INRIA测试集，共 288 幅图片，另一个是TUD-Brussels[30]测试集，共 509 幅图片。图 4-8 和图 4-9，分别给出了INRIA数据集中的训练样本和测试样本示例。

在所有的试验中，我们对 LPP 和 LTP 描述子的参数设置为：LPP 描述子的极轴按对数分为 5 个单位，角度分为 8 个单位，其中角度为非均匀划分，最大跨度为 90° ，最小跨度为 30° ；LTP 描述子最终抽取为一个 400 维的二值特征向量。最终的检测结果处理过程中，非极大值抑制合并窗口（NMS）阈值设为 0.5。所有的实验都是在普通计算机（Intel Core 2 2.50GHz CPU 和 4GB 内存）上运行的，MATLAB 实现拓扑检测算法和窗口合并及最终的图像输出。



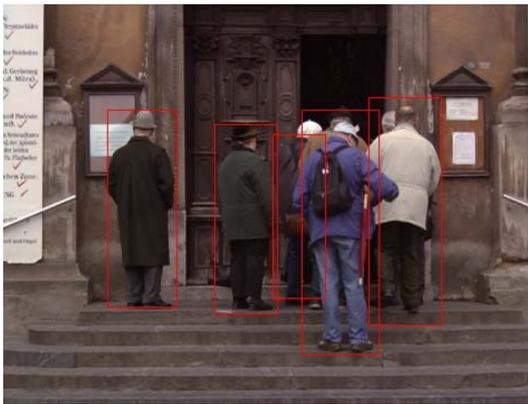
图 4-8 训练样本示例



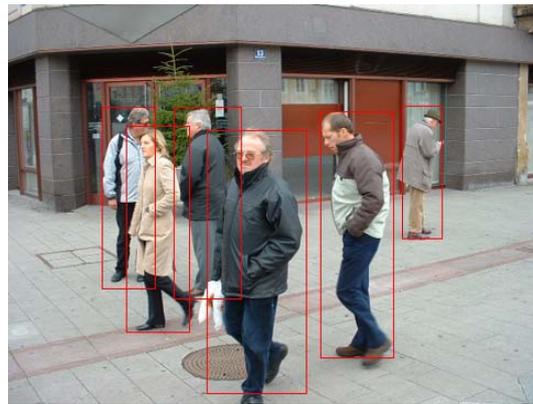
图 4-9 测试样本示例

4.4.2 实验效果及对比分析

如图 4- 10，列举了一些具有代表性图片，包含了背景复杂、姿态多变的人体样本，其中(a-f)是INRIA测试集中的图片示例，(g-o)是TUD-Brussels测试集中的图片示例。可以看到本文提出的基于拓扑模型的检测方法能够将上所有的人体目标正确分类，解决了直立人体的多视角和多姿态的检测问题，提升了检测率和精确率。例如，(a)(b)中的行人密集度较高，存在较大程度的局部遮挡，利用本文提出的方法能够将此类问题很好的解决，(c)和(d)中的实例说明对于姿态变化的行人，基于部位检测的策略发挥了很大的作用，行人拓扑关系的学习也让误检率进一步降低。(g-o)的图片中，行人在图片中的尺度与前面的图片相差较大，但运用本文的检测方法依然能够准确的定位行人。其中，左边图中的行人检测结果在其右侧放大，以便于分析。(g)(h)中骑车的人与路边行人都被准确检测到，并且对于(f)中的局部遮挡也很鲁棒。



(a)



(b)



(c)



(d)



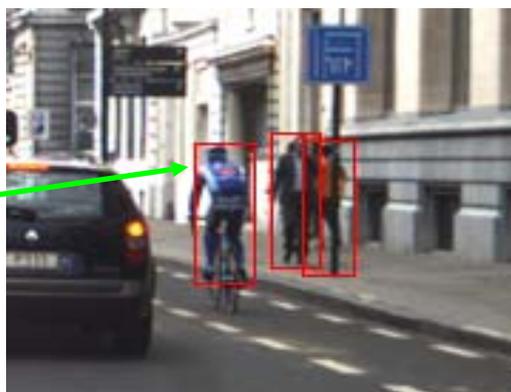
(e)



(f)



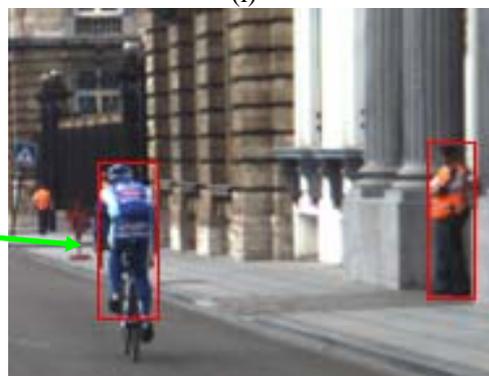
(f)



(i)



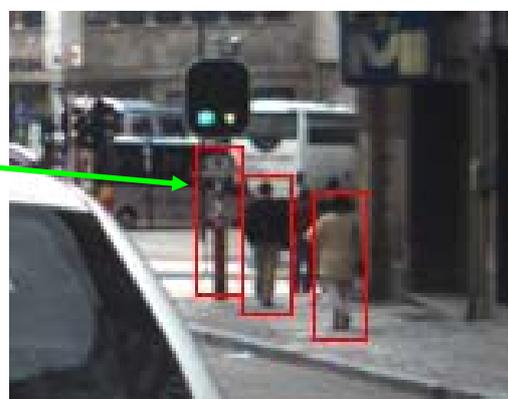
(g)



(k)



(l)



(m)

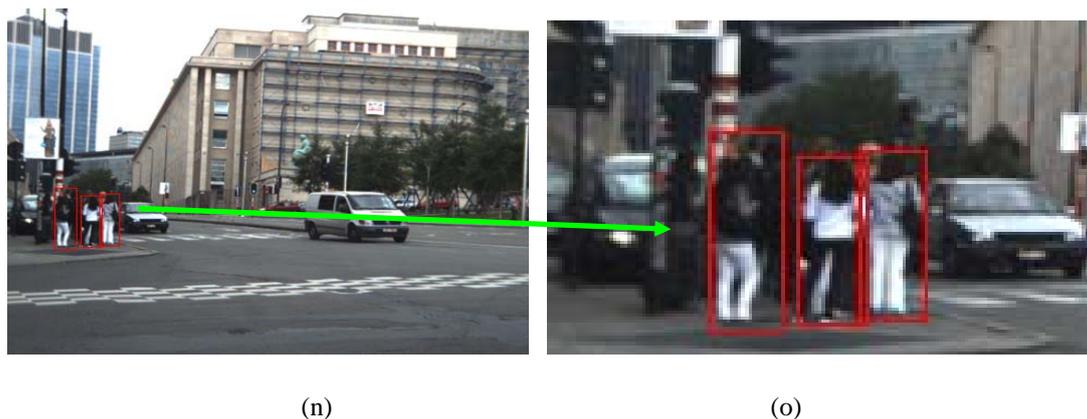


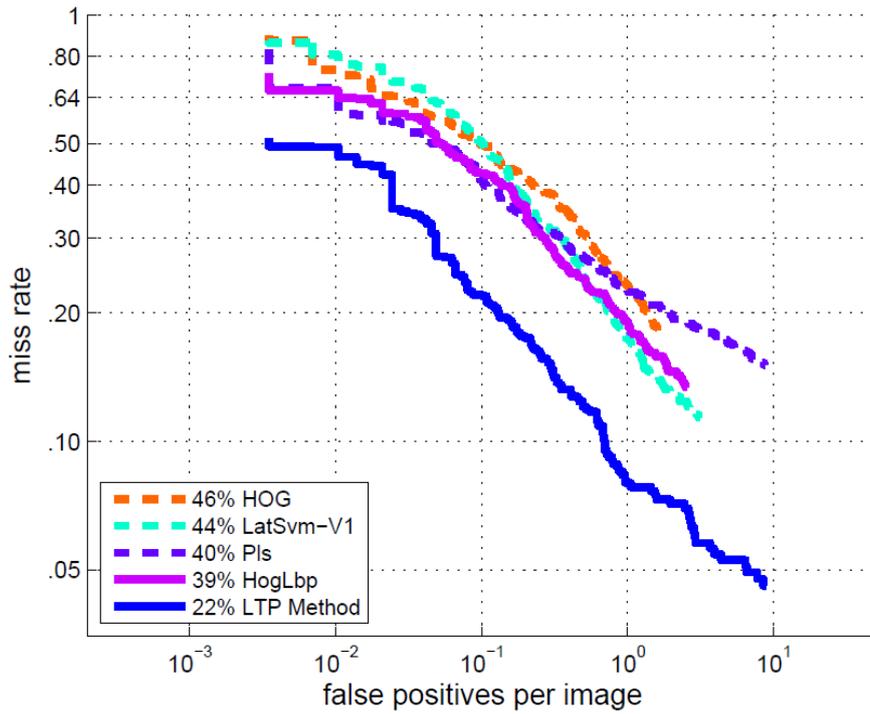
图 4-10 基于拓扑模型的行人检测结果
(a-f) INRIA 测试集, (g-o) TUD-Brussels 测试集

图 4-11展示了基于拓扑模型的方法（LTP）和现在一些经典算法比较的性能曲线。这些经典的算法主要包括Dalal的HOG+SVM[3]算法，Wang的HOG+LBP[28]算法，Schwartz的PLS算法[29]和Felzenszwalb的LatSVM算法[12]。性能曲线是根据丢失率（Miss Rate）和每幅图片的误报率(FPPI)来计算的，所有的点的绘制是在log尺度的坐标下。随着级数的增加，可以看出虚假率增加的同时，丢失率逐渐减少，丢失率和虚假率是相互矛盾的整体，所以在检测时需要平衡两者之间的关系。丢失率和虚假率的定义如下：

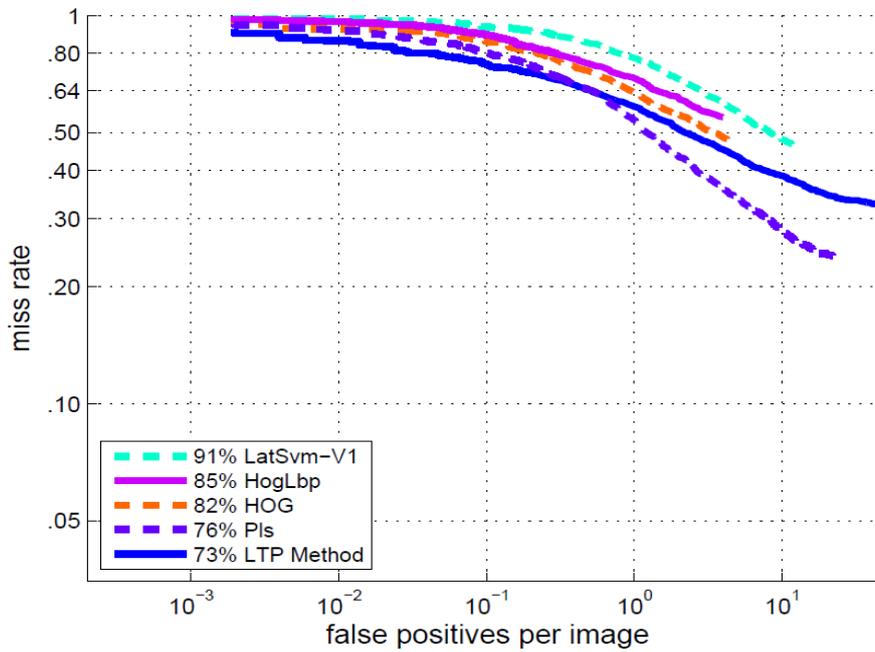
$$MissRate = \frac{\text{丢失的人体数}}{\text{总的实际人体数}} \quad (4-27)$$

$$FPPI = \frac{\text{虚假检测结果的个数}}{\text{图像数目}} \quad (4-28)$$

在对比曲线中可以看到，在INRIA测试集上，LTP算法的平均丢失率只有22%，而其他方法丢失率在39-46%之间，尤其，当FPPI=0.1时，丢失率明显低于20%，说明LTP算法对于行人检测的鲁棒性有较大的提高。在TUD-Brussels测试集中，LTP算法的平均丢失率明显低于其他算法，只有73%。为了方便对比，图4-11给出了曲线中所比较的各个算法的属性统计，包括是否是基于部位的算法，所采用的分类器，以及在INRIA和TUD-Brussels测试集上各自的平均丢失率的精确值。



(a)INRIA 测试集对比曲线



(b)TUD-Brussels 测试集对比曲线

图 4-11 实验曲线与对比

表 4-1给出了实验中方法的参数对比，包括该方法是否基于部位检测，分类器的选取，以及在INRIA和TUD-Brussels数据集上的结果。

表 4-1 实验方法参数对比表

Method	Part based	Classifier	Log-average miss rate	
			INRIA	TUD-Brussels
HOG ^[3]	--	Linear SVM	0.459788	0.818491
LatSVM-v1 ^[12]	√	Latent SVM	0.438304	0.913114
PLS ^[29]	--	Linear SVM	0.308773	0.761705
HOGLBP ^[28]	--	Linear SVM	0.390968	0.854415
LTP Method	√	Linear SVM+ OC-SVM	0.216689	0.732474

4.5 本章小结

本章提出了一种基于部位的行人拓扑模型，并在此基础上给出了 LTP 行人检测方法。介绍了拓扑关系建模，包括拓扑关系描述，以及提出的两个拓扑描述子 LPP 和 LTP。论述了所采用的 OC-SVM 分类方法，详细阐述了 OC-SVM 的基本原理以及在行人检测中的应用。最后，给出了基于拓扑模型的行人检测结果，在高遮挡、多姿态、多尺度条件下验证其有效性和鲁棒性，并与现有的一些经典算法的结果进行了对比。

总结与展望

本论文通过对部位的检测和部位间拓扑关系的学习实现静态图片或视频帧中的行人检测，所提出的方法能够较好的解决行人目标的局部遮挡，姿态多样性等问题。

论文简述了行人检测的研究意义、国内外研究现状和目前已存在的研究方法，分析了国内外主流的行人检测研究方法以及该领域的主要难题；对行人检测相关工作做了综述，包括行人目标检测的一般框架，阐述和分析了目前基于整体模型和基于部件模型的思路和方法，介绍了行人检测研究领域公开的数据集合；分析了目前国内外行人检测方法的一些优劣势，提出了基于拓扑模型的方法应用于行人检测研究。基于整体模型的检测方法，较为依赖于行人整体轮廓的描述，对于行人形变不大，背景简单的场景效果较好，但背景图片复杂，行人姿态差异较大的问题是该方法的瓶颈。基于部位模型的检测能够更好的处理行人多姿态和局部遮挡等难点问题。同时，基于部位模型的检测也是目前行人检测领域中的热点研究方向。提出了行人部位间拓扑关系学习的方法，将行人的姿态利用一个新的特征来描述，并利用 **OC-SVM** 训练行人拓扑模型，从而得到行人与非行人的二分类数学模型。最后，该方法在 **INRIA** 和 **TUD-Brussels** 公共数据集中进行了实验验证，并且与一些有代表性的方法做了对比分析。实验结果表明，基于拓扑模型的方法在行人检测应用中能够较好的解决行人目标在多姿态、背景变化，以及局部遮挡情况下的检测问题，能够取得较为鲁棒的实验效果。

论文提出的基于拓扑关系学习的方法虽然取得了一定的成功，但是仍然存在不足。首先，所提出的方法依赖于部位检测的结果，当行人出现大部分被遮挡的情况时，部位检测结果会出现较大偏差，从而导致行人整体拓扑分类不准确，使得最终的检测结果精度降低。另外，对于姿态变化过大的人体检测，如：运动员肢体变化幅度较大的情况，行人整体拓扑的学习便会存在困难。针对此类问题，在未来的工作中，拟加入多姿态行人拓扑的学习，研究不同类姿态的拓扑规律，将看似差别较大的人体姿态进行聚类，区分不同姿态拓扑的人体目标。

参考文献

- [1] C. Papageorgiou and T. Poggio. A Trainable System for Object Detection[J]. *International Journal of Computer Vision*, Vol. 38, pp.15-33, 2000.
- [2] N. Dalal. Human Detection using Oriented Histograms of Flow and Appearance[C]. *Proc. of the IEEE Conf. on European Conference on Computer Vision*, pp.428-441, 2006.
- [3] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 2, pp. 886-893, 2005.
- [4] Q. Zhu, S. Avidan, M. C. Yeh, and K. T. Cheng. Fast Human Detection using A Cascade of Histograms of Oriented Gradients[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol 2, pp.1491-1498, 2006.
- [5] R. Xu, B. Zhang, Q. Ye and J. Jiao. Human Detection in Images via L1-Norm Minimization Learning[C]. *Proc. of the IEEE Conf. on International Conference on Acoustics, Speech, and Signal Processing*, pp. 3566-3569, 2010.
- [6] R. Xu, B. Zhang, Q. Ye and J. Jiao. Cascaded L1-norm Minimization Learning (CLML) Classifier for Human Detection[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Pp.89-96, 2010.
- [7] I. Laptev. Improvements of Object Detection using Boosted Histograms[C]. *Proc. of the IEEE Conf. on British Machine Vision Conference*, 2006.
- [8] C. Mohan, C. Papageorgiou and T. Poggio. Example-based Object Detection in Images by Components[C]. *IEEE Trans. PAMI*, Vol. 23, No. 4, pp. 349-361, 2001.
- [9] Z.Lin, L. Davis, D. Doermann and D. DeMenthon. Hierarchical Part-template Matching for Pedestrian Detection and Segmentation[C]. *Proc. of the IEEE Int'l Conference on Computer Vision*, 2007.
- [10] B. Wu and Nevatia R. Detection of Multiple, Partially Occluded Humans in A Single Image by Bayesian Combination of Edgelet Part Detectors[C]. *Proc. of the IEEE Int'l Conference on Computer Vision*, Vol.1, pp.90-97, 2005.

- [11] P. Felzenszwalb, R. Girshick, D. McAllester and D. Ramanan. Object Detection with Discriminatively Trained Part Based Models[C]. *IEEE Trans. PAMI*, Vol. 32, No. 9, pp. 1627-1645, 2010, .
- [12] P. Felzenszwalb, D. McAllester, and D. Ramanan. A Discriminatively Trained, Multiscale, Deformable Part Model[C]. *IEEE Trans. PAMI*, pp. 1-8, 2008.
- [13] D. Park, D. Ramanan and C. Fowlkes. Multiresolution Models for Object Detection[C]. *Proc. of the IEEE Conf. on European Conference on Computer Vision*, pp. 241-254, 2010.
- [14] S. Dai, M. Yang and Y. Wu. Detector Ensemble[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp.1-8, 2007.
- [15] M. Andriluka, S. Roth and B. Schiele. People-Tracking-by-Detection and People-Detection-by-Tracking[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.
- [16] M. Andriluka, S. Roth and B. Schiele. Pictorial Structures Revisited: People Detection and Articulated Pose Estimation. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1014-1021, 2009.
- [17] D. Ramanan. Learning to Parse Images of Articulated Objects[J]. *Advances in Neural Information Processing Systems*, Vol. 134, 2006.
- [18] B. Yao, L. Fei-Fei. Modeling Mutual Context of Object and Human Pose in Human-Object Interaction Activities[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 17-24, 2010.
- [19] Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou. Discriminative Local Binary Patterns for Human Detection in Personal Album[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [20] M. Oren, C. Papageoriou, P. Sinha, E. Osuna and T. Poggio. Pedestrian Detection Using Wavelet Templates[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 193-199, 1997.
- [21] P. Viola and M. Jones. Robust Real-time Object Detection[J]. *Proc of International Journal of Computer Vision*, Vol. 57, pp. 137-154, 2001.
- [22] B.E. Goldstein, Sensation and Perception[M], *sixth ed. Wadsworth*, 2002.
- [23] P. Viola and M. Jones. Rapid Object Detection using A Boosted Cascade of

- Simple Features[C]. *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp.511-518, 2001.
- [24] L.F. Teixeira and L. Corte-Real. Video Object Matching Across Multiple Independent Views using Local Descriptors and Adaptive Learning[J]. *Pattern Recognition Letters*, Vol.30(2):157-167, 2009.
- [25] R.M. Haralick, K. Shanmugam, and I.H. Dinstein. Textural Features for Image Classification[J]. *IEEE Transactions on Systems, Man and Cybernetics*, Vol.3, No.6, pp.610-621, 1973.
- [26] K. Mikolajczyk, C. Schmid, and A. Zisserman. Human Detection Based on A Probabilistic Assembly of Robust Part Detectors[C]. *Proc. Of the IEEE Conf. on European Conference on Computer Vision*, pp. 69 - 82, 2004.
- [27] P. Felzenszwalb and D. Huttenlocher. Pictorial Structures for Object Recognition[J]. *International Journal of Computer Vision*, Vol.61(1), pp. 55-79, 2005.
- [28] B. Leibe, E. Seemann, and B. Schiele. Pedestrian Detection in Crowded Scenes[C]. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Vol.1, pp. 878 - 885, 2005.
- [29] S. Ioffe and D.A. Forsyth. Probabilistic Methods for Finding People[J]. *International Journal of Computer Vision*, Vol.43(1), pp. 45-68, 2001.
- [30] F. Porikli. Integral histogram: A Fast Way to Extract Histograms in Cartesian Spaces[C]. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Vol.1, pp. 829 - 836, 2005.
- [31] <http://pascal.inrialpes.fr/data/human/>
- [32] P. Dollár, C. Wojek, B. Schiele, and P. Perona. Pedestrian Detection: A Benchmark[C]. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 304 - 311, 2009.
- [33] C. Wojek, S. Walk, and B. Schiele. Multi-cue Onboard Pedestrian Detection, [C]. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 794-801, 2009.
- [34] <http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2007/index.html>.
- [35] <http://www.mis.tu-darmstadt.de/tud-brussels>

- [36] <http://coe.gucas.ac.cn/SDL-HomePage/resource.asp>
- [37] C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition[J]. *Data Mining and Knowledge Discovery*, Vol. 2, No.2, pp.121-167, 1998.
- [38] 邓乃扬,田英杰, 数据挖掘中的新方法---支持向量机[M], 北京: 科学出版社, 2004.
- [39] C. Burges. Advances in Kernel Methods-Support Vector Learning[M]. *MIT Press*, 1999.
- [40] V. Vapnik. Stastical Learning Theory[M]. Willey, 1998.
- [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*, pp.1443-1471,2001.
- [42] M. Gavrilad, J. Giebel. Shape-Based Pedestrian Detection and Tracking[C]. *Proc. of IEEE Conf. on Intelligent Vehicle Symposium* , Vol. 1, 17-21, pp. 8-14 , 2002.
- [43] I. Haritaoglu, R. Cutler and D. Harwood. Detection of People Carrying Objects using Silhouettes[C]. *Proc. of IEEE Conference on European Conference on Computer Vision*, pp.102-107, 1999.
- [44] E. Rivlin, M. Tudzsky and R. Goldenberg. A Teal-time System for Classification of Moving Objects[J]. *Proc. 16th Int'l Conf. on Pattern Recognition*, Quebec, Vol.3, pp.688-691, 2002.
- [45] A. Senior, A. Hampapur and Y.L. Tian. Appearance Models for Occlusion Handling[J]. *Image and Vision Computing*, pp.1233-1243, 2006.
- [46] A. Lipton, H. Fujiyoshi and R. Patil. Moving Target Classification and Tracking from Real-time Video[C]. *In Proc IEEE Workshop on Applications of Computer Vision*, pp:8-14 1998,.
- [47] R. T. Collins, A.J Lipton and T. Kanade. A System for Video Surveillance and Monitoring: VSAM Final Report[R]. Carenegie Mellon University, Technical Report: XMU-RI-00-12, 2002.
- [48] K. Mikolajczyk, C. Schmid, and A. Zisserman. Human Detection Based on a Probabilistic Assembly of Robust Part Detectors[M]. *Lecture Notes in Computer Science*, Vol. 3021, Berlin, Springer, pp, 69-82, 2004.

- [49] B. Leibe, A. Leonardis, and B. Schiele. Robust Object Detection with Interleaved Categorization and Segmentation[J]. *International Journal of Computer Vision*, Vol.77, pp. 259-289, 2008.
- [50] L. Wang, J. Shi, G. Song and I. Fan. Object Detection Combining Recognition and Segmentation[C]. *Proc. in IEEE Conf. on Asian Conference on Computer Vision*, pp. 189-192, 2007.
- [51] S. Avidan. Support Vector Tracking[J]. *IEEE Trans. PAMI*, pp.1064-1072, 2004.
- [52] 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]. *Proc. IEEE International Conference on Image Processing*, pp.1501-1504, 2005.
- [53] H.D. Dai, M. Chen, and S.L. Zhou. Support Vector Machine Based Adaptive Kalman Filtering for Target Tracking[C]. *Proc. in IEEE Conf. on General System and Control System*, pp. 114-118, 2007.
- [54] F. Tang, S. Brennan, Q. Zhao, and H. Tao. Co-tracking Using Semi-Supervised Support Vector Machines[C]. *Proc. in IEEE Conf. on International Conference on Computer Vision*, pp.992-999, 2007.
- [55] J. Zhang, D.W. Chen, and M. Tang. Combining Discriminative and Descriptive Models for Tracking[C]. *Proc. IEEE International Conference on Asian Conference on Computer Vision*, pp. 113-122, 2009.
- [56] N. Kabaoglu. Target Tracking Using Particle Filters with Support Vector Regression[C]. *Proc. in IEEE Transactions on Vehicular Technology*, pp. 2569-2573, 2009.
- [57] C. Cortes and V. Vapnik. Support-Vector Networks[J]. *IEEE Trans. Machine learning*, Vol.20, No.3, pp. 273-297, 1995.
- [58] B.E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers[C]. *Proc. of the fifth annual workshop on Computational Learning Theory*, pp. 144-152, 1992.
- [59] Y. Freund and R. Schapire. A Decision-theoretic Generalization of On-line Learning and An Application to Boosting[C]. *Proc. in IEEE Conf. on Computational Learning Theory*, pp. 23-37, 1995.
- [60] R. Schapire. The Strength of weak LSeamability[J]. *IEEE Trans. Machine*

- learning*, Vol.5, No.2, pp.197-227, 1990.
- [61] C. Diehl. Toward Efficient Collaborative Classification for Distributed Video Surveillance[D]. PhD thesis, Citeseer, 2000.
- [62] O. Tuzel, F. Porikli, and P. Meer. Learning On Lie Groups For Invariant Detection and Tracking[C]. *Proc. in IEEE Conf. on Computer Vision and Pattern Recognition*, pp.1-8, 2008.

发表(录用)文章目录

- Wen Gao, Xiaogang Chen, Qixiang Ye, Jianbin Jiao, “Pedestrian Detection Via Part-based Topology Model,” in Proc. International Conference on Image Processing (ICIP), 2012.

致 谢

在中国科学院研究生院攻读硕士学位三年的学习生活中，我付出了艰辛的努力，也经历了诸多坎坷，同时也得到了很大的收获。在此期间，模式识别与智能系统实验室的老师、师兄、师姐和同学们给予我极大的帮助和鼓励，让我领会到团队协作的精神，严谨的治学态度，和多角度的完备的思维方式。

首先，感谢我的指导老师焦建彬教授，对我的论文撰写给予悉心的指导和严格把关，并为我提供了良好的学习和实验环境。我的硕士论文从选题、开题、研究内容直到论文的撰写，都离不开焦老师的悉心指导。焦老师治学严谨，学识渊博，为我营造了良好的学术氛围。他从治学态度，专业知识，团队合作以及自我要求上都为我树立了良好的榜样，并在学习、生活上给予极大的关心和帮助，让我受益匪浅。这一切都将是今后学习、工作、生活的宝贵财富。

其次，感谢叶齐祥副教授，他在我的科研学习的各个方面都给予了极大的帮助和鼓励。感谢叶老师在我攻读硕士学位期间对我的每一点指导和教诲，在我的论文撰写和修改中倾注的心血。老师们渊博的专业知识、自强不息的学习精神和扎实的实践能力，在理论学习和项目开发过程中给我提供了耐心的引导和帮助。感谢韩振军老师，在硕士期间对我的指导和鼓励，他在科研和学习生活中认真负责的态度使我受益良多，也是我今后学习的好榜样。

感谢我的学长徐冉、陈孝罡，在科研中给我的指导和帮助。感谢我的小组同学武博、李策、纪颖夏、刘一飞、祝耀华，在这三年中，大家一起讨论，一起发现问题、解决问题，一起拼搏，我们不仅结下了深厚的友谊，更共同见证了研究生最后时光。感谢我的同学吴金晨、李潇宇、宁悦、李小君，我们互相鼓励，互相帮助，让我度过了人生中非常重要和开心的三年。

特别感谢我的父母，作为普通的工薪阶层，他们养育了我，并且让我受到了良好的教育，他们不懂我的论文写的是什么，也不会对我的论文有任何指手划脚，但他们默默的关心和关爱是对我最大的支持！

最后，感谢参加论文评审和答辩的各位老师。

高文

2012年4月

