

中国科学技术大学学位论文相关声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

本人授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅，可以将学位论文编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

保密的学位论文在解密后也遵守此规定。

作者签名： 陈达

2007 年 5 月 26 日

摘要

每年在汽车交通事故中死亡的行人数目庞大，基于保护行人、避免交通事故的目的，行人检测系统已成为目前国内外智能交通领域公认的一个研究热点。

行人检测系统是一种车载辅助驾驶系统，它具有自动识别处于危险位置的行人、为驾驶者提供警报并在危急时刻接管驾驶等功能。近年来，已经有一些汽车厂商推出了面向实用的系统；然而这些系统一般采用昂贵的传感设备，并且只装备在部分高档车型上，这不利于行人检测系统的推广，特别在我国更不符合建设节约型社会的需要和社会需求。因此，基于简单传感设备的行人检测技术研究具有重要的理论研究意义和很高的实际应用价值。

本文以采用单光学摄像头的行人检测系统为研究背景，重点研究其中的分类检测机制关键技术。论文的主要工作与特色有：

首先，设计并实现了一个只采用单光学摄像头的低成本行人检测原型系统。该系统先利用统计学习分类器过滤掉大量不含行人的图像，再利用支持向量机进行精确的分类检测；为了实现行人意图判别，该系统还加入了行人行进状态检测功能；

然后，针对支持向量机分类器的设计，我们提出了递进的两种基于进化的支持向量机分类检测方法。两种方法都同时采用行人外形和运动特征，但第一种方法使用进化算法来优化设计分类器的训练模型；而第二种方法进一步实现了特征集合和支持向量机分类器训练模型的联合优化。

实验表明，我们优化得到的分类器具有较高的检测率和很低的误报率，并具有可以调节的检测速度；初步实现的单光学行人检测原型系统在实际城市交通环境中具有较好的检测能力，经过进一步改进后具有较高的推广价值。

关键词：行人检测系统 支持向量机 进化算法

Abstract

On our road, traffic accidents involving pedestrian-vehicle collisions cause significant injury and fatality every year. In a goal to reduce such alarming statistics and protect pedestrians, many teams began to research pedestrian detection systems.

A pedestrian detection system is designed to detect pedestrians, alarm to driver and take over driving in some emergent situation. In recent yeas, some automobile manufactures worked out a few practical systems. However, only these luxury cars are equipped with these systems which usually use top grade sensor. This is not proper to popularize the intelligent vehicle system. To make the system more useful, it is important to improve the detection ability under the popular sensors equipped vehicles.

In this paper, all of our works are under the platform of single camera based pedestrian detection system. We focus on the classification method which is the key of pedestrian detection. The main advantages of the paper are as follows:

Firstly, a single camera based prototype is designed to detect pedestrian. The detection process includes two steps. In the first step, most images which don't contain a pedestrian are filtered by the static learning classification. In the second step, SVM classifier is used to determine the ultimate type of a candidate image.

Secondly, an evolutionary SVM classifier is designed for the pedestrian detection system. Considering that the training model and feature set have tremendous and associated effect to the performance of a SVM classifier, we use parameters of kernel and penalty parameter C to represent the training model and then apply evolutionary method to optimize them. Therefore a most suitable SVM classifier can be obtained for pedestrian detection.

Experiments show that our system based on this classifier has high detection rate and acceptable detection speed in the real city traffic environment. It will be a competitive candidate after some slight system progress.

Keyword: Pedestrian Detection System, Support Vector Machine, Evolutionary Algorithm

第一章 引言

1.1 行人检测系统的目标

随着汽车的普及，汽车交通事故发生的频率也与日俱增。据统计，2006 年全国共发生道路交通事故 378781 起，共造成 89455 人死亡；其中汽车交通事故占了绝大多数。作为事故中相对缺乏安装保障设施的行人，其生命安全所遭受的威胁尤为突出。

为了降低交通事故发生率，保护行人的生命安全，研究界和汽车产业界都开始了行人检测系统（PDS: Pedestrian Detection System）的研究。这种系统依靠汽车上的摄像和传感装置来识别汽车前方处于危险地带的行人，并提供驾驶员警报和紧急时刻接管驾驶等服务。就目前来说，绝大部分商用系统都依赖于高精度的视频采集和传感设备，比如红外摄像头，雷达等。这些昂贵的设备极大的增加了汽车的制造成本，一般也只配备在比较高档的汽车身上，不利于行人检测系统的推广。因此研究如何改进检测技术，用相对廉价一点的设备来识别行人是非常有意义的。

一个完整的行人检测系统要实现的功能包括行人的检测、跟踪、意图判别、危险评估、信号预警和真实预警等。其中最重要的目标是快速准确地检测出处于危险区域的行人。一般来说，该检测系统的工作方式是利用装置于汽车中的摄像头来实时捕捉前方信息，同时采用模式识别技术在采集到的连续帧中检测汽车前方处于危险地带的行人。这些识别技术大部分都采用机器学习算法从大规模训练样本中训练以得到一个分类器，然后使用这个分类器对每一帧所对应的整个危险区域进行穷举扫描以发现可能存在的行人。归根到底，行人识别系统的目标就是判断一幅图像中是否包含行人。如何选择合适的图像识别算法来训练分类器是决定一个行人识别系统性能的关键。

1.2 行人检测系统的特殊要求

模式识别技术已经在图像检测和机器人视觉研究中获得了广泛的成功，其基本思想就是利用机器学习算法对大规模样本进行训练得到所需要的模式分类器，已经在实验中验证了有效性的算法包括 Adaboost[1][2]，支持向量机（SVM：

Support Vector Machine) [3-10], 模拟退火[11], 神经网络[12-15]等等。后面章节会详细介绍这些算法如何应用到行人识别系统的分类器设计中。

大量的实验结果验证了模式识别技术在人脸检测这样一个典型问题中的有效性;然而,当把这些方法应用在行人检测中却遇到很多问题。其中最主要的问题是:在城市交通环境中,复杂多变的背景与路面情况带来复杂纹理,而我们的目标是要准确快速地将姿态各异的行人作为一类从背景剩余的所有物体中识别出来,并且不能以降低行人的识别率和增加误报率为代价。

当然情况也有好的一面,不同于传统图像识别,行人检测并不只是在一幅图像中、而是在一段连续帧中来完成。相对于背景,行人有其自身的运动信息:如果能想办法将行人的运动信息提取出来,并设计相应的分类算法,那么和传统图像识别方法结合起来后,将会在检测速度和检测率上得到很大提高。第一个这样设计 PDS 系统的人是 viola 等人,他于 2003 年首先同时采用形状和运动信息来设计分类器[1],之后有不少人对他们的工作进行了重复和局部改进。从这些工作可以看出,合理地利用多种信息源来设计分类算法能够有效提升检测效果,但是多种信息的融合本身也是一个比较有挑战性的课题。下图是行人检测中信息处理的一般流程。

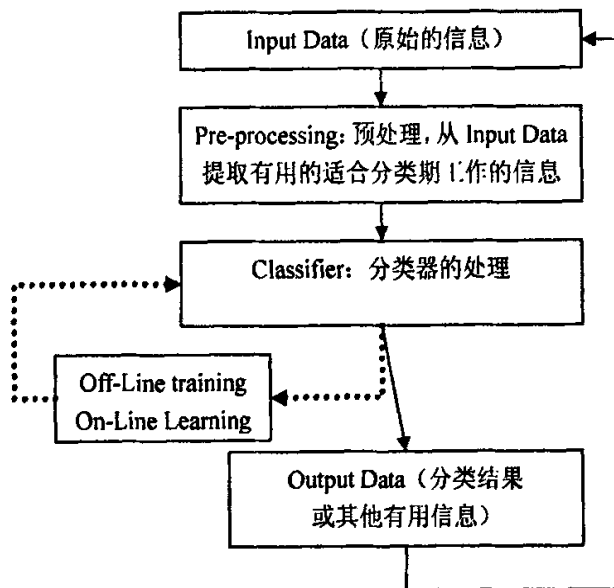


图 1-1 行人检测中信息处理的一般流程

1.3 行人检测系统研究现状

1.3.1 系统硬件配置与系统结构

本文的主要工作是采用自进化的 SVM 算法来设计 PDS 中的分类器。但由于分类工作所需要的信息源有很多,不同的信息源又会导致分类器设计方式有所区别,所以在正式介绍行人检测技术之前有必要对 PDS 所采用的不同传感器以及获取的信息源做简单的介绍。

当前主流的传感器有光学摄像头、红外摄像头、雷达、声纳/超声波、压力计。使用频率依次降低,前三者使用得较多,下表 1-1 给出了它们的基本功能和比较:

表 1-1 主流的传感器性能

传感器类型	主要功能	缺陷	使用多个的结果
光学摄像头	获取可见光环境中视频图像	图像较为复杂、不便于处理;受光照、天气(例如雨雪天气)影响严重;不适用于夜间以及行人与背景色调相似的情况	可以获得立体视觉图像,建立3维模型,强化纹理信息,可以利用更多的行人外部特征,
红外摄像头	根据热源获得视频图像,适合夜间,雨雪天气工作	设备昂贵,容易受热源干扰,例如汽车,路灯	和光学摄像头类似,可以获得立体图像,获得更多的纹理信息
雷达	精确测量测试一定范围内物体距离	无法获得物体纹理信息,无法区分障碍物体是什么	精确测量物体的距离,可以方便地实现跟踪

通过上表的分析可以发现,要想让设计的 PDS 变得更加精确,增加一些高精度的传感器是一种比较直观的做法。在我们后面的介绍中可以看到确实有人走这条路,然而这样做的代价就是汽车制造成本大大上升,对 PDS 的实际普及与推广有较大的限制性影响。因此,在相对廉价的光学摄像头下,想办法改进检测效果显得更有推广意义。

一般来说, PDS 的设计思路如图 1-2 所示。当前几乎所有的方法都遵循以下工作模式:

- (1) 确定系统要实现的功能,选取合适的硬件设备;
- (2) 根据自己选取的信息源,设计多信息源的不同过滤器。选择和设计的标准是:系统的功能、信息源必须足够,系统的检测性能必须达到一定的效果、满足系统实时要求,实时性目标在某种程度上是最重要的;

(3) 根据系统要实现的功能，决定是否需要设计这个模块，需要的话又要考虑计算代价。

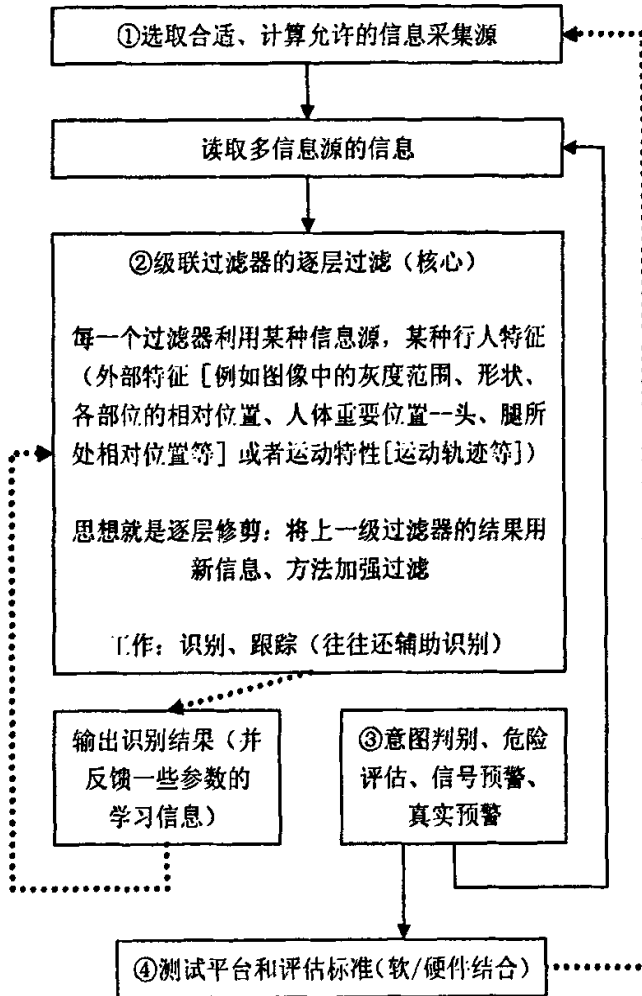


图 1-2 PDS 中的模块设计

从上图可以看出，模块①是决定信息来源的，而②是用来做图像检测的，也是本文的主要工作，③是我们将来需要考虑的。其中虚线路径代表一个自适应学习调整的过程，是模块优化的重要手段。

1.3.2 现有行人检测系统

智能交通系统是个新兴的热门研究领域，PDS 是其中一个重要的组成部分，应用前景主要是辅助驾驶、自动/半自动导航（军用 / 民用）。

PDS 实用系统的研究开发才刚刚起步，很少有能达到真正的理想的效果的实

用系统。2000 年之前的工作基本停留在监控应用背景下的行人识别。目前，代表世界最先进的 PDS 系统有两个：PROTECTOR[16]和 SAVE-U[17]这两个系统。这两个系统是一脉相承的，都是 DaimlerChrysler 公司下的机构来设计完成的。SAVE-U 比 PROTECTOR 要更先进成熟，硬件上主要提升是引入了新的信息源/硬件—5 个新型雷达，而 PROTECTOR 只使用了 2 个或者 3 个摄像头。由于硬件上的变化，SAVE-U 的体系结构上也多出了一些处理雷达信息的过滤器模块，关于分类算法的改进尚没有详细的说明文档，然而从参与设计系统的人员所发表的一些论文中可以看出一些信息，这在后面的方法介绍中会有提及。下面是这两个系统发展历程：

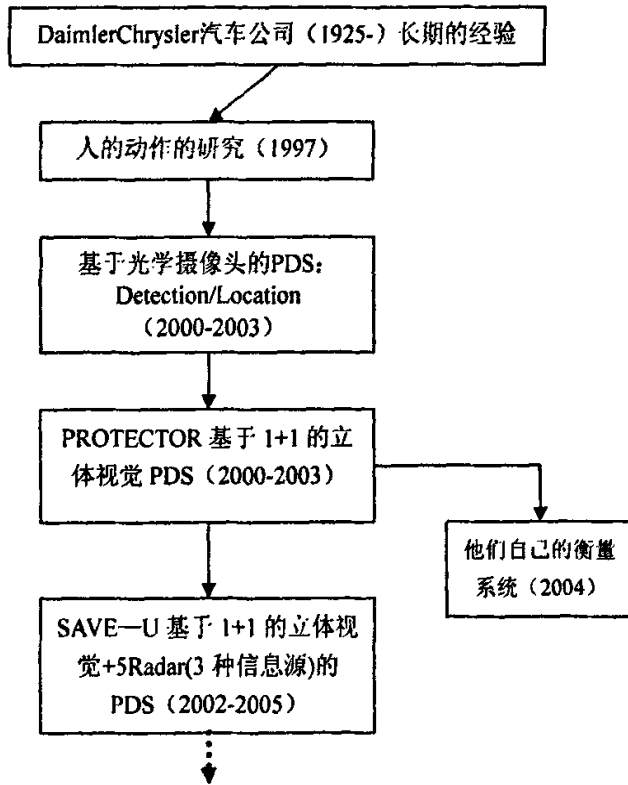


图 1-3 SAVE-U 与 PROTECTOR 系统的发展历程

从图 1-3 中可以看出，该系统的传感设备趋于多元化，技术上朝着多种信息源融合方向发展。

另外一个实用系统是由 US-Army 和 Italy 合作开发的 ARG0[18]，其发展路线见图 1-4，同样可以看出 ARG0 采用的传感设备也向着高精度与多元化发展。

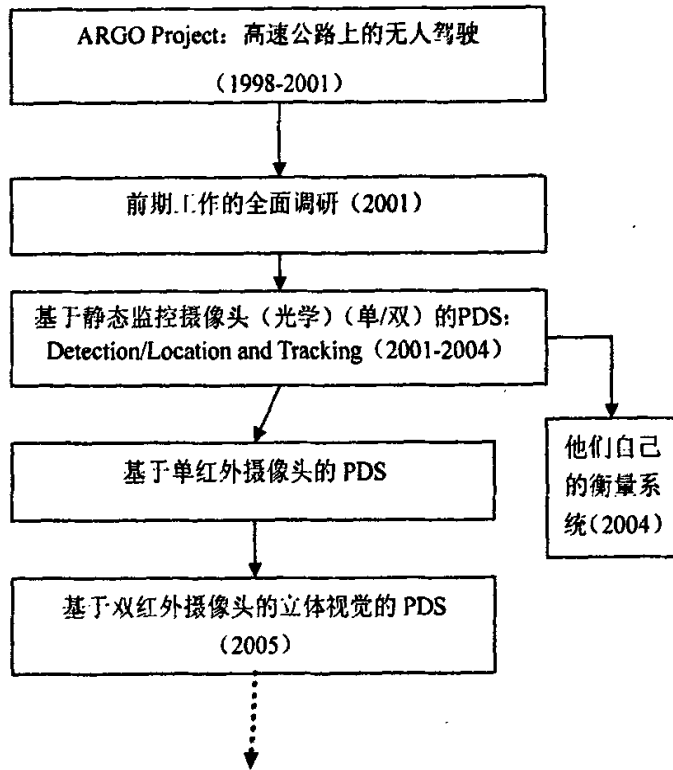


图 1-4 US-Army PDS 系统的发展历程

这两支队伍研究的目标并不一致（前者民用，后者军用），采用的硬件和方法也差别很大，但是都是注重持续研究、背景雄厚的队伍，是目前 PDS 研究队伍中成果最多的。前者有自主开发硬件的能力，所以使用了一些别的机构无法使用的硬件，获得更多的信息，利用多信息的融合得到了实用的效果；后者基于军事上常用的、成熟的红外技术治理采用最简单的方法解决这一复杂问题，也取得了不错的进展，虽然不及前者用复杂的系统解决复杂的问题的效果好，但是对 PDS 研究起了很大推广作用。

近两年国外公司和大学研究小组中也实现了一些其他的准实用系统，比如：以色列 MobilEye 公司开发的出芯片级别 PDS[2]；日本本田公司的基于红外摄像头的 PDS[5]；CMU[12]，MIT[9][10]都有比较重要的研究工作。

在国内，我们智能交通研究小组从 2004 年开始对行人检测系统进行研究，已经实现了一个基于单光学摄像头的原型 PDS[19-23]、清华大学[24][25]，西安交通大学[7]也在该领域做了许多研究工作。

从近几年发展来看，PDS 的趋势将是制造实时的可靠的应用软硬件产品。该

系统一方面有很大的硬件依赖性，很多工作中都提到了性能、方法的改进依赖于新型传感器的诞生和改进。另一方面，PDS 的性能还依赖于不断发展的图像检测技术，这是实用系统设计必须面对和解决好的问题。然而后期的人的运动意向识别、危险评估、预警决策更多的涉及机器智能，也是在解决好检测之后需要逐步改进和完善的。总的来说，一个实用的 PDS 要先做好第一步检测工作，才有实现后续功能的可能。

1.4 现有行人检测技术

当前可行的行人检测系统中的检测技术主要分为两类：基于外形的方法和基于运动的方法。然而越来越多些实用系统开始结合这两种方法来检测行人。

1.4.1 基于运动特征的方法

基于运动的检测方法一般是指通过观察行人走路的周期性步态特征来检测行人。它的主要优点是能够消除光线和背景纹理条件给传统图像识别带来的影响。不足之处在于需要连续帧的信息，可能是需要分析一个周期较长的特征，那样会影响整个检测系统得实时性。这种周期特征的检测主要采用傅立叶变换和人工神经网络两种方法。Cutler[26]等人方法是先计算不同时间间隔图像之间的相关性，然后对得到相关信号利用短时傅立叶变换进行分析，进而得到检测结果。Kressler[15]等人的方法是通过一个自适应时延神经网络对图像序列进行分析，进而判断该序列是否是行人的运动序列。除了以上两个对全身进行分析的工作，还有对行人腿部运动周期进行分析的，如 Curio[27]的工作。

1.4.2 基于外形特征的方法

基于外形特征的方法是指通过分析待检测图像的边缘和纹理特征来进行检测。与基于运动的方法相比，它可以使用一些传统的图像检测技术来进行识别，特别是可以识别静止的行人。基于外形的方法主要分为三类：

(1) 基于人体建模的方法

基于人体建模的方法主要是指根据人体结构的特点，构造一些行人的参数模型，包括 2D 或 3D 的。在检测行人时，提取待检测图像的原始特征来求解模型。该类方法的优点是可以处理遮挡问题，重建行人姿态，不足之处在于人体建模比

较困难,求解也相当复杂。采用 2D 建模的代表有:Sun[25],Zhao[28]和 Yasuno[29]。其中, Sun 将人体分为头,躯干和左右腿四部分建立模型,不仅有各部分的形状信息,还包含各部分的空间信息,这些限制条件有助于问题的求解,Zhao 针对人体的正面和侧面分别建立 2D 模型,Yasuno 等人的做法是将人体分为头部和躯干两部分来建模。跟 2D 模型相比,3D 模型在处理遮挡问题与视角变化上具有较大优势,然而在信息获取上要依赖更多的传感器。代表工作有 Bertozzi[30]等人针对不同行人姿态建立一个 3D 模型系列来对红外图像进行识别。

(2) 基于模版匹配的检测方法

基于模版匹配的方法是通过实现制作一些行人的轮廓模版,然后在检测时提取待检测图像的轮廓与模版库中的模版比较得到检测结果。该方法的优点是计算上简单,不足之处在于一个较完备模版库的建立比较困难。这方面的代表工作是 Gavril[11]实现的一个 PDS。其中他制作了 2500 个不同姿态的模版,并采用模拟退火算法将模版库组织成树形结构来匹配待检测图像,这样一种有粗到精的分层搜索策略有效地解决了因匹配量过大带来的检测速度问题。具体的匹配方法是先将模版与待检测图像的轮廓提取图像转化为距离变换(DT: Distance Transform)图像,然后计算两个 DT 图像的 Chamfer 距离[31]来判别它们相似度。由于其不错的检测效果,Gavril 实现的这个系统曾应用于 PROTECTOR 项目。不仅有对人的整体采用模版匹配的方法,还有对人体局部采用模版匹配的工作,Broggi[18]等人利用不同大小的头和肩部模版来进行匹配,该方法已被应用于 Param 大学开发的 ARGO 项目中。

(3) 基于统计学习的检测方法

基于统计学习的方法主要是从大量训练样本中通过机器学习寻找其内在的模式,从而得到相应的分类器。基于分类的优点是算法鲁棒性强,对于不同光线和场景的行人均具有较好检测能力。不足之处在于训练样本库的建立以及针对行人检测这一特定问题需要对机器学习算法作特殊定制。一般来说,机器学习分两步:第一是特征提取,第二是设计分类器的训练。特征提取是为了得到图像最关键的特征,设计分类器训练是为了提高分类器的泛化能力,也就是提高我们检测算法的鲁棒性。在行人检测方面,机器学习方法主要有三种:人工神经网络,SVM 以及 Adaboost。

人工神经网络的思想是通过学习的方法将模式的特征隐含在一系列的网络参数之中,可以描述极为复杂的模式。它已经成功地应用在人脸检测等模式识别领域中,目前在行人检测领域也有一些积极的尝试[12][13][14]。Szarv[13]等人

利用卷积神经网络 (CNN: Convolutional neural network) 来检测行人, 它的特点是直接使用待检测图像灰度值作为 CNN 的输入, 特征提取的细节隐藏在 CNN 内部网络结构中, 学习算法会自动构造最优特征。类似的工作还有 Franke[14]等人的成果。另外一种做法是 Zhao[12]等人采用图像梯度的幅度作为特征, 构造三层的前馈神经网络作为分类器来检测行人, 作者认为梯度的引入有助于消除光线干扰以及二值化时产生的阈值选择问题。

SVM 方法是 Vapnik[32]等人建立的一种统计学习方法, 它遵循结构风险最小化原则, 具有良好的泛化性, 可以避免以前神经网络训练中出现的过学习问题, 是一种具有广泛应用前景的小样本机器学习方法。在行人检测领域与哈尔小波特征结合使用最为广泛。代表性工作有: 在单光学摄像头 PDS 中, C. Papapeourgio[3]采用过完备的哈尔小波提取外形特征, 然后用 SVM 方法构造分类器。作者在文中详细分析了特征集合对 SVM 类器检测性能的影响, 是构造 SVM 分类器检测行人最早的一批研究人员。他们设计的 PDS 也成为当时一些汽车公司的候选。与其他特征结合使用的工作包括: 在双光学摄像头 PDS 中, Grant Grubb[4]等人的主要工作是采用 3D 建模技术过滤掉大部分复杂背景, 并设计了一种 3×3 的 Sobel 算子来提取图像边缘信息, 利用这种外形特征构造了两个并联的 SVM 分类器来同时检测两种不同行进状态的行人。作为尝试, 作者还实现了检测后的跟踪来提高检测率。在单红外摄像头 PDS 中, Fengliang Xu[5]等人提出了一个先检测再跟踪的实现方案, 第一步首先利用图像灰度信息构造热点图像外形特征, 然后训练 SVM 分类器来检测行人。其中有特色的地方是将行人分为三类, 并比较了用一个单分类器来检测这三类和三个串联的分类分别检测出这三类的实现方案。第二步用 Kalman filter prediction[33]和 mean shift tracking[34]来实现行人对象的跟踪。在他们实验中显示检测和跟踪的交替使用带来了更高的识别速度以及准确率。除此之外, 形状直方图特征[6]、描述符特征[8]、Gabor[7] 特征与 SVM 的结合也广泛地应用于现有的行人检测系统中。

Adaboost 是一种分类器组合的策略, 它的目的是将一些弱分类器组合成一个强分类器。Adaboost 得到的分类器具有较好的推广性能, 现在广泛地应用于模式识别和计算机视觉领域。代表性的工作是 Viola 的行人检测系统, 他采用 Adaboost 算法训练了一个统计学习分类器, 并第一次同时采用了外形和运动特征来训练分类器。其中矩形特征的构造和计算方法源于他以前做人脸识别的工作, 经过作者改进后将运动信息成功的引入特征集合中, 非常有效的支持了行人的检测。除了 Viola 的矩形特征, SIFT (Scale invariant feature transform) [35]特征因为其尺度不变的良好特性, 也在基于 Adaboost 的行人检测中得到广泛应用。代表性工作包括: Shashua[2]等人的成果, 他们设计的系统将识别阶段分成两个部

分，在单帧识别阶段利用 Adaboost 训练得到的分类器对行人进行识别，在多帧识别阶段通过分析行人的运动信息来进行综合的判决。在使用 Adaboost 训练分类器时，利用了类似 SIFT 特征的梯度方向的直方图特征。

总结起来，现在主流的 PDS 设计方法主要基于外形的方法来构造检测分类器，然而很多实验结果表明在复杂背景纹理中的检测结果还不是特别理想，新出现的一些方法采用一些短周期运动特征作为外形特征的补充，或者使用长周期运动检测的方法得到辅助检测结果来参与最后决策，大量的实验结果表明同时采用外形和运动信息的系统具有更好的检测能力。

1.5 本文研究的技术路线

经过对以往工作的一些分析，我们决定设计一个以 SVM 分类器为核心检测机制的级联分类器来检测行人。级联分类器的工作流程如下图所示：

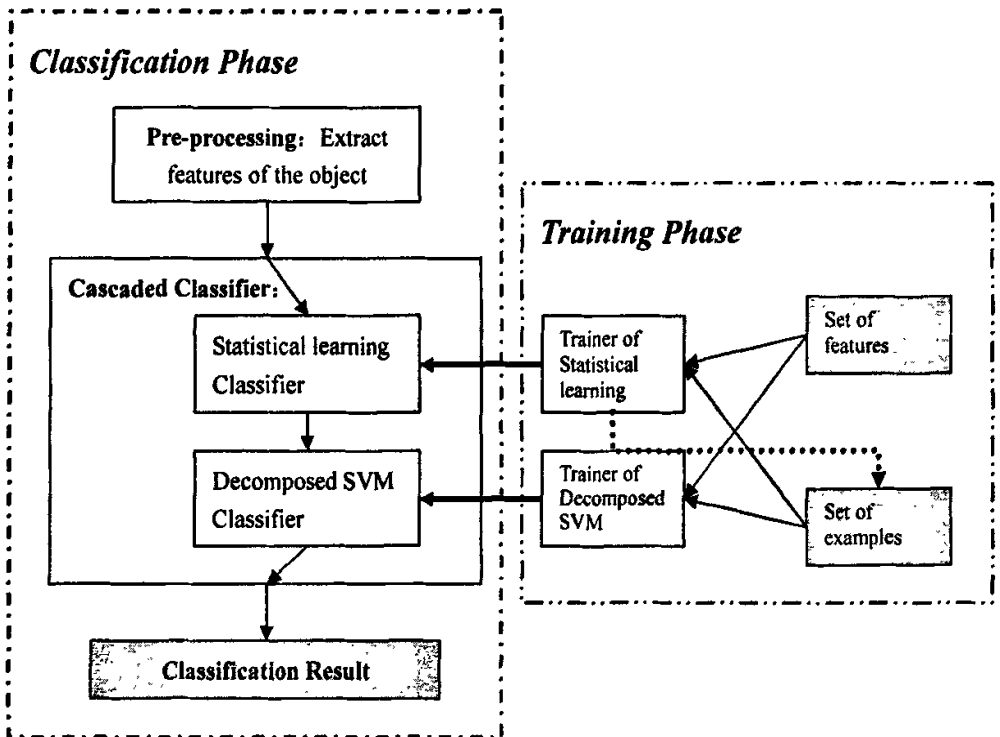


图 1-5 级联分类器的结构

级联分类器分类流程：在分类阶段，如图 1-5 所示，整个检测流程可以描述成以下几步：

- (1) **Pre-processing:** 预处理阶段主要工作是提取并计算待识别对象的特征向量。
- (2) **Statistical Learning Classification:** 这一阶段的工作是用统计学习分类器进行初选, 得到最有可能包含行人的子图像, 相当于为后面的 SVM 分类阶段做兴趣域 (ROS: Regions of Interest) 选取。这一层分类器的训练方法为 Adaboost 算法, 检测机制是如果分类器认为是正样本, 检测继续进行, 否则停止检测。
- (3) **SVM Classification:** 这一阶段的工作是用 SVM 分类器对上一层没被过滤掉的图像进行检测。这一层分类器训练算法为 SVM 方法, 检测机制是如果认为是正样本, 则输出结果为正样本, 否则输出结果为负样本。虽然 SVM 检测的速度是整个级联分类器的瓶颈, 当时由于上一层已经过滤掉绝大部分不含行人的图像, 故而能够最终得到一个比较满意的速度。

为了从实验中验证算法的有效性, 我们模拟了一个基于单光学摄像头的 PDS, 所采用的训练和检测样本均来自于车上实拍的视频, 在以后的每章中都对手实验结果进行了详细分析。

1.6 本文工作与内容安排

1.6.1 本文主要工作与特色

PDS 是因应行人安全保护而提出的一个具有创新性意义和实用价值的研究方向。现有的采用昂贵传感设备的系统难以大规模推广。因而, 只采用几个光学摄像头的 PDS 更具研究和开发价值。目前, 国内外都对此给以高度关注。

本文以一个只用单光学摄像头 PDS 的研制为背景, 重点研究了其中的分类检测技术。结合这一系统的特殊要求, 我们利用进化的支持向量机设计出合适的分类器, 在上述原型系统的检验中取得了很好的效果。

论文的主要工作与特色有:

- (1) 完成了一个面向城市交通环境的 PDS 原型系统的设计。其中同时采用了外形和运动两种特征, 提出了 Adaboost 与 SVM 结合的分类器设计方法, 实现了图像由粗到精的检测, 并实现了行人本身及其状态的检测;
- (2) 提出了一种基于进化 SVM 的 PDS 分类算法。该算法可以自进化得到合适的 SVM 训练模型, 从而得到合适的 SVM 分类器, 解决了以前通过实验方法需要人工干预以及无法得到全局最优训练模型的问题;
- (3) 提出了一种特征和 SVM 训练模型的联合进化算法。该算法可以同时获得适

当的关键特征和 SVM 训练模型；既提高了检测速度，也可以使检测率保持在较高水平。

1.6.2 本文的内容安排

本文的章节安排如下：

第一章 引言。主要介绍行人检测技术的现状，包括 PDS 中主流的传感器，检测方法以及一些具有实用价值的 PDS，并在最后提出了我们的技术路线；

第二章 单光学摄像 PDS。主要介绍我们设计的基于单光学摄像头的 PDS，包括级联分类器中特征的计算方法和不同层次的分类器所采用的机器学习算法；

第三章 基于进化 SVM 的行人检测。主要介绍我们如何利用进化算法来优化 SVM 分类器的训练模型，包括训练模型的编码解码，适应度函数的选取，进化方式等等；

第四章 特征集合与 SVM 训练模型联合优化。主要介绍特征选择与 SVM 分类器性能的关系，采用进化算法来实现特征集合与 SVM 分类器训练模型的联合优化，进一步优化了检测速度与检测率；

第五章 总结与展望。对全文进行总结并对下阶段工作进行了展望。

第二章 单光学摄像头 PDS 原型系统

本章介绍我们设计完成的一个基于单光学摄像头的行人检测系统。该系统的核心是一个级联检测器模块，可以检测静止或者运动的行人，以及行人可能的行进状态。按照我们的技术路线，首先介绍分类器所采用的特征计算方法，再介绍不同层次分类器的设计目标以及相应的训练和检测流程。

2.1 特征的选取

特征选取是 PDS 设计的第一步。一般来说，原始的灰度特征难以直接利用，需要经过特征提取以达到降低特征维数，得到反映模式本质属性的目的。在我们的系统设计中，特征选取采用了类似于 Viola 的矩形特征方法[1]，特征计算方法采用积分图像算法[36]。下面具体介绍。

2.1.1 特征表示与计算

1. 特征表示

初始特征集合的生成过程分为以下两步：

- (1) 在 32×16 的矩形框内部生成大量不同矩形（不同位置或不同大小）。每个矩形的长和宽都大于 2。
- (2) 每个矩形内部可以有下图的五种分布(二矩形，三矩形，四矩形)，这样每个矩形可以得到五个新的矩形，每个特征就是由一个新的矩形表示。



图 2-1. 五种基本外形特征

2. 特征计算

特征值的计算采取积分图像的方法,先分别求黑色区域和白色区域内部所有像素的灰度值之和,再求得到的两个和的差的绝对值。例如,对于第一种特征的特征值:设图片中位于黑色区域的灰度值和为 x ,位于白色区域的灰度值和为 y ,则该特征的特征值为 $abs(x-y)$ 。

运动特征的特征值是通过计算相邻两帧图片之间的差异得出的。我们把这些差异分为五种: $\{\Delta, U, D, L, R\}$,表示后一帧相对前一帧的五种运动形式:静止、向上、向下、向左和向右。其中每种差异的计算方法如下:

$$\Delta = abs(I_t - I_{t+1}) \quad (1)$$

$$U = abs(I_t - I_{t+1} \uparrow) \quad (2)$$

$$D = abs(I_t - I_{t+1} \downarrow) \quad (3)$$

$$L = abs(I_t - I_{t+1} \leftarrow) \quad (4)$$

$$R = abs(I_t - I_{t+1} \rightarrow) \quad (5)$$

I_t 和 I_{t+1} 分别表示连续两帧图像, $\{\uparrow \downarrow \leftarrow \rightarrow\}$ 是运动算子(例如 $I_{t+1} \uparrow$ 表示将第二帧图像 I_{t+1} 上移一个像素),这样可以得到包括第一帧以及 Δ, U, D, L, R 总共六种 PC 的分类器使用,其中后面五种都包含了运动信息。如图 2-2 所示

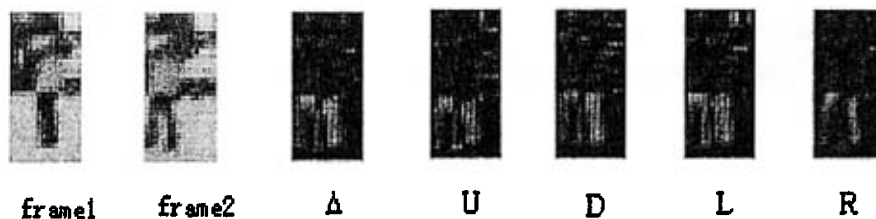


图 2-2 分类需要的六个图形

我们采用和 Viola 相同的方法计算运动特征的特征值。外部特征只是基于图像的灰度而不需要考虑前后两帧的差别,因此它的计算对象就是 I_t ,而运动特征的计算对象是 Δ, U, D, L, R 。根据 Viola 的介绍,运动特征的计算方法又可以分为以下三种方法:

(1) 第一种运动特征的特征值计算方法选择的特征是单矩形框特征。它计算 Δ 与

$\{U, D, L, R\}$ 之间的差别。对于一指定特征 a 的特征值 f_a 计算公式如下：

$$f_a = r_a(\Delta) - r_a(S) \quad (6)$$

其中 $S \in \{U, D, L, R\}$, $r_a(\)$ 为特征 a 的矩形区域的所有像素灰度值之和。

- (2) 第二种运动特征的特征值计算方法选择的特征是如图 1.1 所示的五种矩形框特征，它是通过对 U, D, L, R 计算特征的特征值得到的。对于一指定特征 a 的特征值的计算公式如下：

$$f_b = \phi_b(S) \quad (7)$$

其中 $\phi()$ 是按照计算外部特征特征值的方法对 $\{U, D, L, R\}$ 求出特征值。

- (3) 第三种运动特征的特征值计算方法选择的特征也是单框矩形特征。它的特征值是对 $\{U, D, L, R\}$ 计算一个矩形区域内所有像素的灰度值和。计算公式如下：

$$f_c = r_c(S) \quad (8)$$

其中 $S \in \{U, D, L, R\}$, $r_c(\)$ 为特征 c 的矩形框内所有像素的灰度值和。

2.2 统计学习分类器

我们的方法借鉴了 Viola 统计学习方法[1]的技术路线，但在具体实现上有所差别。注意到实际情况中负样本比正样本的数目要多的多，因此我们在 Adaboost 算法训练过程中对权值进行调整，加大对负样本识别错误时的惩罚力度，提高识别的正确率。Viola 设置了四层级联，用第一层训练外部特征，后三层训练运动特征。而我们的方法中为了提高负样本识别的正确率，采用了 10 层级联，前七层采用外部特征，并且串联连接。后三层使用运动特征，采用并联连接。Viola 采用的样本规模是 20×15 像素，而我们采用的规模是 32×16 像素。

2.2.1 训练统计学习分类器

统计学习分类器的训练流程如图 2-3 所示：

- (1) 输入包含连续两帧图像的正负样本，取连续两帧是为了能够计算短周期的运动特征；
- (2) 用上一节的方法生成从初始特征集合，并从中随机提取 60000 个特征；
- (3) 采用积分图像算法计算这些特征值；
- (4) 计算统计学习分类器的识别率和误报率，如果达到要求就完成训练，否则就进入下一步；

- (5) 采用 Adaboost 算法分类，得到一个弱分类器；
 (6) 将弱分类器加入级联结构，转到步骤(4)。

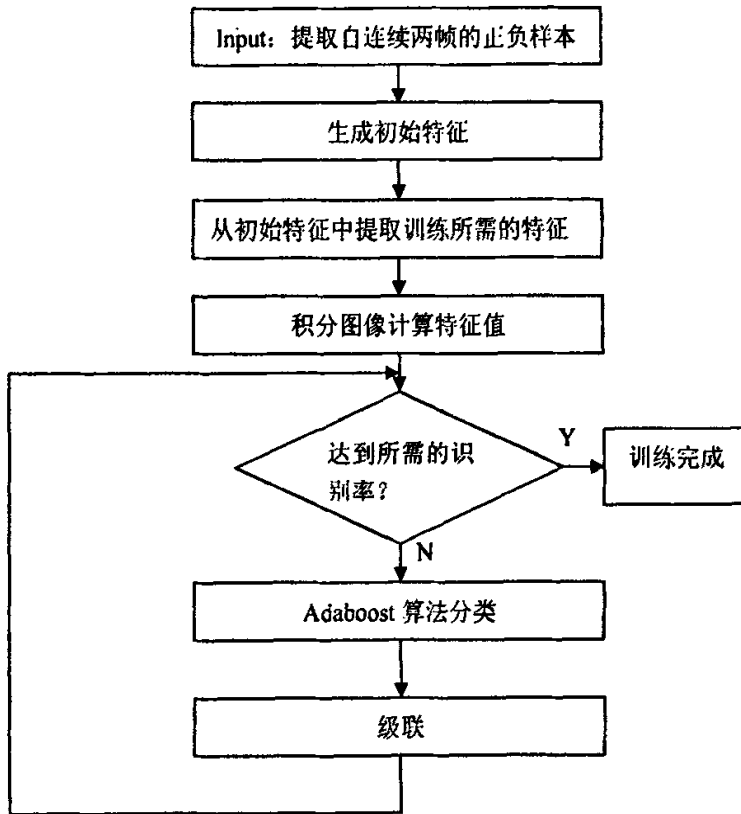


图 2-3 统计学习分类器训练流程图

2.2.2 统计学习分类器的级联结构

为了降低误报率，统计学习分类器本身也是采用级联的结构，只有通过前一级的样本才会被下一级的分类器判断。如图 2-4 所示，我们采用了十层级联。前七层使用的特征是外部特征，后三层使用的是运动特征，运动特征的加入使其更接近实际情况，可以进一步提高识别精度，这三层采取并联结构。一个样本只有在被前七层分类器和后三类分类器中的至少两个都判断是正样本时，才确认它为真样本。

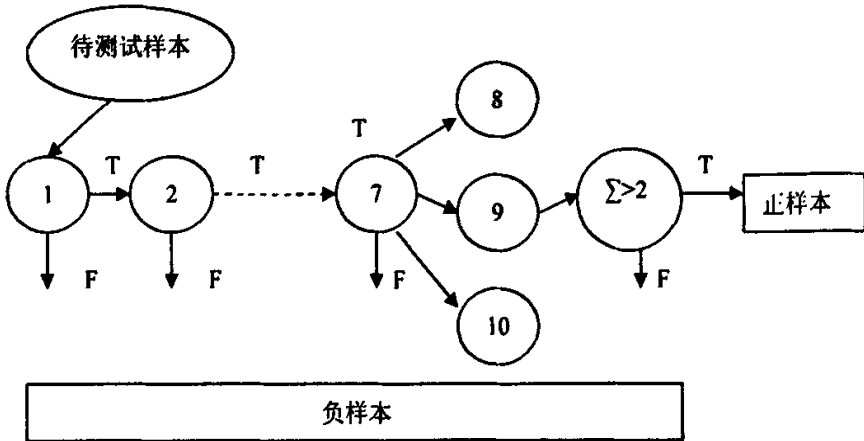


图 2-4 统计分类器的级联结构

在实现的过程中，前两级的负样本来源于手工框图得到。后八级的负样本是通过使用程序得到。我们把那些被前 $i-1$ 层识别为“正样本”的负样本作为第 i 级的负样本集合。每级分类器采用相同的正样本。

实现伪代码如下：

// 训练 n 级的分类器

P = 正样本集合。

N = 负样本集合

n_i 为第 i 级的分类器的特征个数

D_i 为第 i 级分类器的识别率

F_i 为第 i 级分类器的误报率

d 为分类器所允许的最小识别率

f 为分类器所允许的最大误报率

for $i=1$ to 10 do

$n_i=0$;

$f=0.1+i \times 0.05$

$d=0.95$

选择该层所用的特征；

do

$n_i \leftarrow n_i + 1$

使用 Adaboost 算法，对正样本集合 P 和负样本集合 N 训练，得到一个分类器。

使用该分类器对测试集合进行测试，得到该分类器的识别率 D_i 和误

```

    报率  $F_i$ 。
    if  $D_i > d$  then 重新设置该分类器的阈值, 使得  $D_i \geq d_i$ , 并重新对
        测试集合进行测试, 计算误报率  $F_i$  ;
    Until  $F_i \leq f$ ;
     $N \leftarrow \Phi$ 
    if  $i=1$  or  $i=2$ 
        then 读入手工框制的负样本到集合  $N$  中;
        else 把被当前级联分类器识别为“正样本”的负样本放入集合  $N$  中
    endif
endfor

```

2.3 SVM 简介

SVM 是一种基于统计学习理论的机器学习方法, 主要理论由 Vapnik[32] 等人建立。机器学习的基本思想是从观测数据出发寻找规律, 利用这些规律对未知数据进行预测。其理论基础传统统计学研究的是样本趋于无穷大时的渐进理论, 而 SVM 的理论基础是一种基于小样本情况下的机器学习理论。

2.3.1 统计学习理论的建立

机器学习的一般问题描述如下[37]:

已知 n 个观测样本 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

求: 最优函数 $y^* = f(x, w)$

满足条件: 期望风险 $R(w)$ 最小

$$R(w) = \int L(y, f(x, w)) dF(x, y) \quad (9)$$

其中 $L(y, f(x, w))$ 为损失函数

然而期望风险 $R(w)$ 要依赖联合概率 $F(x, y)$ 的信息, 实际问题中无法计算。因此在实际使用中一般用经验风险 $R_{\text{exp}}(w)$ 代替 $R(w)$,

$$R_{\text{exp}}(w) = \frac{1}{n} \sum_{i=1}^n L(y_i, f(x_i, w)) \quad (10)$$

可以看出传统方法有两点不足之处:

- (1) 经验风险最小不等于期望风险最小, 不能保证分类器的推广能力。
- (2) 经验风险只有在样本数无穷大趋近于期望风险, 需要非常多的样本才能保证

分类器的性能。

为了得到最佳的预测函数，需要找到经验风险最小和推广能力最大的平衡点，正是在这个背景下 Vapnik 等人创立了统计学习理论。其核心内容是经验风险 $R_{\text{emp}}(w)$ 和实际风险 $R(w)$ 之间以至少 $(1-\eta)$ 的概率满足下面的不等式：

$$R(w) \leq R_{\text{emp}}(w) + \Phi(h/\eta) \quad (11)$$

其中， h 是函数集的 VC 维[32]。关于统计学习理论更详细的介绍参见 Vapnik 的著作。

2.3.2 SVM 数学表示与问题求解

SVM 原始的数学表示如下[37]：

已知 n 个观测样本 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

求解：

$$\min \frac{1}{2} \|w\|^2$$

$$y_i(w \cdot x_i - b) \geq 1 (i = 1, 2, \dots, n)$$

即使得间隔 $2\|w\|^{-1}$ 最大的分类面：

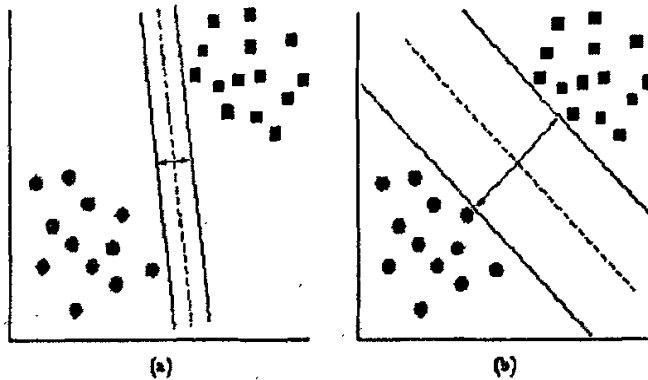


图 2-5 最大间隔分类面

求解该问题的对偶表示是一个二次优化问题[37]：

$$\min(L_D) = \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j x_i \cdot x_j - \sum_{i=1}^l \alpha_i = \frac{1}{2} \alpha^T Q \alpha - e^T \quad (12)$$

$$\begin{aligned} \text{约束条件: } 0 \leq \alpha_i \leq C \\ y^T \alpha = 0 \end{aligned} \quad (13)$$

其中 l 为样本数, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)^T$, $i = 1, \dots, l$; $Q = (Q_{ij})_{i,j=1,\dots,l}$, $Q_{ij} = y_i y_j x_i \cdot x_j$ 。

由于矩阵 Q 非负定, 该问题是个凸的最优化问题。Kuhn-Tucker 条件是其最优化应该满足的充要条件。

对于非线性可分的数据样本, 使用函数 $\Theta(x)$, 将所有样本点映射到高维空间, 则新的样本集为 $\{(\Theta(x_1), y_1), \dots, (\Theta(x_n), y_n)\}$ 。此时新样本集在高维空间可能转化为线性可分, 而且在训练问题中, 涉及到训练样本的数据计算只有两个样本向量点乘的形式。通过构造满足 Mercer 定理的核函数 $k(x_i, x_j) = \Theta(x_i) \cdot \Theta(x_j)$ 来避免高维空间的向量点乘, 这时对偶问题变成另一个二次优化问题:

$$\min(L_D) = \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^l \alpha_i = \frac{1}{2} \alpha^T Q \alpha - e^T \quad (14)$$

$$\begin{aligned} \text{约束条件: } 0 \leq \alpha_i \leq C \\ y^T \alpha = 0 \end{aligned} \quad (15)$$

其中 l 为样本数, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)^T$, $i = 1, \dots, l$; $Q = (Q_{ij})_{i,j=1,\dots,l}$, $Q_{ij} = y_i y_j k(x_i, x_j)$ 。由于矩阵 Q 非负定, 该问题仍然是个凸的最优化问题。Kuhn-Tucker 条件是其最优化应该满足的充要条件。

2.3.3 SVM 常用核函数介绍

线性内核:

$$k(x_i, x_j) = x_i \cdot x_j \quad (16)$$

多项式内核:

$$k(x_i, x_j) = ((x_i \cdot x_j) + 1)^q \quad (17)$$

径向基函数内核:

$$k(x_i, x_j) = \exp\left\{-\frac{|x_i - x_j|^2}{\sigma^2}\right\} \quad (18)$$

Sigmoid 内核:

$$k(x_i, x_j) = \tanh(\nu(x_i \cdot x_j) + c) \quad (19)$$

2.4 基于 SVM 的行人分类检测

SVM 由于其出色的学习性能和泛化推广能力，在图像识别和文本分类领域 [15] 得到了广泛应用。在第二章的检测技术介绍中已经提到很多人都用它来实现 PDS 中的分类器设计。不同的特征计算方法所得到的分类器差别也很大，下面分别介绍在我们设计中 SVM 分类器的训练准备和检测流程。

2.4.1 训练阶段

分类器的训练流程可以描述如下：

(1) 获取训练样本

构造 SVM 分类器所需要的样本来自于我们在汽车上用一 SONY 光学摄像头采集的城市交通视频。为了得到特征集中的运动信息，每个样本由两幅图像组成，这两幅图像分别剪切自连续帧中的相同位置，规格为 32×16 。正样本数量和负样本数量的比例为 1:3。

(2) 选取训练所需要的特征集合

在我们的级联分类器前端，已经用 Adaboost 方法初选了一个特征集合，这里我们可以根据要实现的 SVM 分类器检测速度从其中选取特征子集合。采用适当的运动特征可以在一定程度上消除复杂背景的影响，所以我们最终选取的特征集合为外形特征和运动特征的混合集。

假设特征集合的规模为 n ，那么每个特征 c_k 对应有一个特征计算函数 $f_k(A, B)$ ， $k=1, \dots, n$ ，其中 A, B 代表一个训练样本所包含的两个灰度矩阵。

(3) 将每个训练样本转化为 SVM 训练所需的特征向量

所有特征的计算方法 $f_k(A, B)$ ， $k=1, \dots, n$ ，在介绍特征的那一节已经提到。我们可以用这些函数将样本映射为特征向量：

具体来说，样本 S_i 由灰度矩阵 A_i 和 B_i 组成，其中 $i=1, \dots, l$ ， l 是训练样本的个数。

向量 $x_i = (f_1(A_i, B_i), f_2(A_i, B_i), \dots, f_n(A_i, B_i))$ 可以被计算出来，并被标记为训练样本 S_i 的特征向量

(4) 训练 SVM 分类器

分类器训练的目标是得到下面的分类函数 $f(x)$ ：

$$f(x) = \text{sgn}\left(\sum_{\text{Support Vector}} y_i \alpha_i k(x_i, x) - b\right) \quad (20)$$

其中 (x_i, y_i) 对应一个训练样本, x 表示待分类输入, $k(x, y)$ 为核函数, α_i 为待求拉格朗日乘子, b 为分类超平面偏移量。

训练分类器就是要得到最优的 α_i 。这等价于求解上一小节中提到的二次优化问题。该优化问题是个凸的最优化问题, Kuhn-Tucker 条件是其最优化应该满足的充要条件。本文采用 Joachims 的 SVM^{light} 算法来训练分类器[38], 其基本思想是如果存在不满足 Kuhn-Tucker 条件的样本, 则以某种方式选择 q 个样本作为当前训练集, 其他样本保持不变, 在此样本集上解决 QP 问题, 重复这一过程直到所有样本都满足 Kuhn-Tucker 条件, 另外 Joachims 实现了一种高效缓存机制来保存核函数的计算结果, 可以大大缩短训练时间。

2.4.2 检测阶段

检测阶段流程如下图:

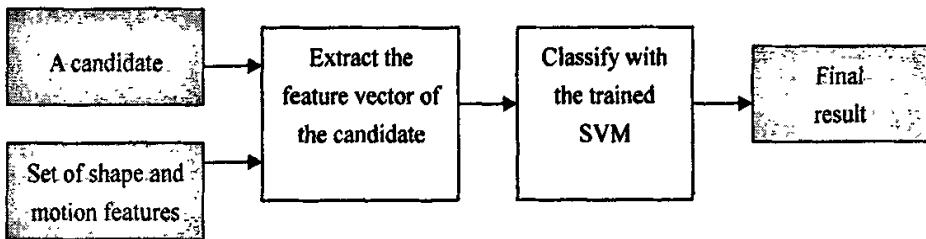


图 2-6 SVM 分类器的检测流程

如图2-6所示, 完成SVM分类器的训练之后, 可以按照下面的步骤来检测行人:

- (1) 依据先前选定的特征集为待检测图像 (包含灰度矩阵为 A' 和 B') 计算特征向量: $x' = (f_1(A', B'), f_2(A', B'), \dots, f_n(A', B'))$
- (2) 将特征向量 x' 作为决策函数 $f(x) = \text{sgn}\left(\sum_{\text{Support Vector}} y_i \alpha_i k(x_i, x) - b\right)$ 的输入, 可以得到检测结果: 1 表示该图像包含行人, -1 表示不包含行人。

2.5 实验与分析

本次实验中，PDS 运行在 PC 机上，CPU 型号为 Barton2500+，随机访问存储器为 1G DDR。所有的训练样本来自十段视频，这些视频拍摄于白天真实的城市交通环境中，视频画面大小为 320×240 ，汽车平均车速约为 25 公里/小时，这足以模拟一个真实 PDS 的实时采集，我们还准备了另外五段这样的视频用来做测试。目前，共制作了 2000 对含行人的图像作为正样本和 3000 对不含行人的图像作为负样本，其中图像的规模统一为 32×16 。

本次实验从样本中同时提取外形和运动信息来组成特征向量，然后用 Joachims 的分解算法来训练分类器。

如图2-7所示，每一个训练样本都由两幅图像组成，这两幅图像是从连续两帧中的同一位置剪切出来。

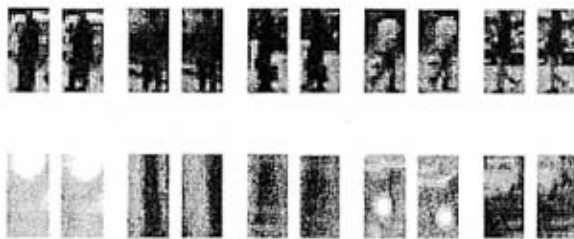


图 2-7. 训练样本：第一行表示包含行人的正样本，
第二行表示不包含行人的负样本

到目前为止，还没有一种指导性的策略能够确定 SVM 分类器所需要的特征集合中外形特征与运动特征的比例。在本系统的设计中，为了分析运动特征对检测效果的影响，我们手工构造了几个特征集合，所有集合的规模保持相同，但不同集合具有不同的外形与运动两种特征的比例。通过对 SVM 分类器的训练以及针对具体测试样本的性能测试，我们选取了其中比较有代表性的三个特征集合，它们的具体构成见表 2-1。为了能够直观地比较上述三种特征集合对 SVM 分类器训练的影响，我们还为每组特征集合对应的分类器画出了 ROC 曲线。

表 2-1. 三种不同的特征集合

特征集合 A	300 个外形特征
特征集合 B	240 个外形特征和 60 个运动特征
特征集合 C	120 个外形特征和 180 个运动特征

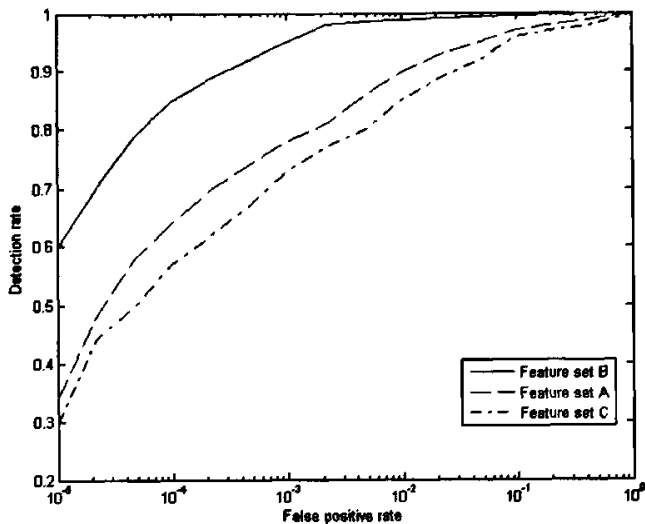


图 2-8 基于三个不同特征集合分类器的 ROC 曲线

图 2-8 描绘了不同分类器在测试视频中 ROC 曲线。其中，最上面的红色曲线反映了特征集合 B 对应 SVM 分类器的性能表现，中间的浅蓝色曲线反映了特征集合 A 对应 SVM 分类器的性能表现，最下面的深蓝色曲线特征集合 C 对应 SVM 分类器的性能表现。可以看出特征集合 B 训练得到的 SVM 分类器在相同误报率下具有最好的性能，特征集合 C 训练得到的 SVM 分类器性能最差。

在本次实验中，我们发现刚开始向以往采用的外形特征集合中加入运动信息时，训练出来的 SVM 分类器检测性能会提升，当运动特征的比例增加到约 20% 时，SVM 分类器的检测率达到一个最高点。进一步增加运动特征的比例反而会导致检测性能的下降，当比例达到约 60% 时，分类器的检测性能已经不如完全采用外形特征训练和检测的性能。这同上面 ROC 曲线反映的规律是一致的，训练集合 B 中运动特征比例为 20%，此时训练得到的分类器检测能力最强。

这一结果显示，我们采用运动特征来设计分类器是合理有效的。在复杂的场景中，完全采用外形特征将会得不到好的检测表现，因为这时行人的外形纹理特征将会和背景纹理特征混淆在一起而难以区分。这时，作为另一种信息的携带者，运动特征对行人的检测能起到非常重要的作用。然而，运动特征的比例并不是越高越好，在我们的实验中，运动特征的最佳比例约为 20%，更高的比例反而会带来性能的下降。这说明在我们选取的检测视频中，PDS 主要还是依靠外形特征来识别行人，运动特征只是在非常复杂场景中有着更好的表现。因此，只有当特征集合选好外形和运动特征合适的比例时，才能使训练出来的 SVM 分类器达到最佳性能。鉴于我们的实验也是手工选取特征比例来进行，下一步的工作应该更深

入的研究运动特征对检测性能的影响，以及研究如何自适应的选取特征比例。

2.6 行人行进状态检测

在一个功能完备的 PDS 系统中，仅仅识别出行人是不够的，更多的行人信息需要获取，比如行人的行进状态。对行人的识别与行进状态分类可以设计专门的多分类器模块来实现。目前的大部分工作中，为 PDS 所设计的分类器只是简单做了对行人的识别，更多的信息获取依赖于硬件或由别的模块来实现。

本章中第一部分的内容是使用 SVM 方法来为 2D 场景中的 PDS 系统构造多分类器模块。和以往工作不同的是，我们在识别行人的同时，还实现了对行人行进状态进行分类，将行人分为横穿马路和沿着马路两类。实验结果表明，这种多分类器模块在高效识别行人的同时，还能准确得到行人的行进信息，有助于实现 PDS 系统中关于行人的意图判别和预警。

2.6.1 多类 SVM 分类器

我们设计的多类分类器由三个两类分类器组成，并组织成一个二叉树的结构。第一个用来将可能包含横穿马路行人的图像和可能包含沿着马路行人的图像区分开，第二个用来将包含横穿马路行人的图像和不包含行人的图像区分开，第三个用来将包含沿着马路行人的图像和不包含行人的图像区分开。每一个分类器都采用 SVM 方法训练，其中第一个 SVM 分类器的训练样本由包含横穿马路行人的图像和沿着马路行人的图像区分开的图像组成，第二个 SVM 分类器的训练样本由包含横穿马路行人的图像和不包含行人的图像组成，第三个 SVM 分类器的训练样本由沿着马路行人的图像和不包含行人的图像组成。

实验结果显示，基于这样一个多类分类器的 PDS 可以很好的识别出行人，并得到他们的行进信息，这为后面实现意图判别提供了便利。

2.6.2 多类 SVM 分类器的结构

为了获取行人的行进信息，需要将一个待检测图像分为三类，包含沿着马路行人的图像，包含横穿马路行人的图像，以及根本不包含行人的图像。设计一个分类器来解决这样一个特殊的多类划分问题将会引出一些非常复杂的优化问题，

所以我们选择用三个两类分类器组合起来解决这个问题，即在检测出行人的同时得到他们的行进信息。这样一个组合多类分类器的结构如下图所示：

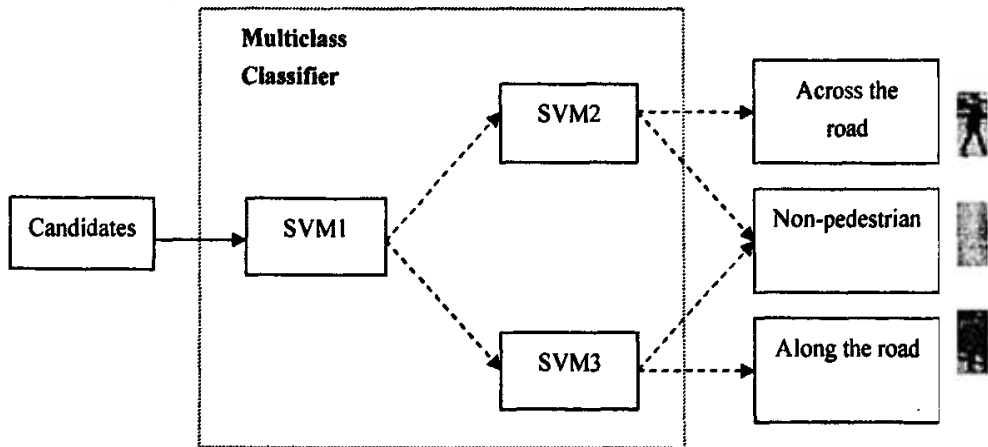


图2-9 多类分类器的结构

如图2-9所示，这个多类分类器的结构可以被认为具有3个节点的二叉树。其中，每个节点代表一个SVM分类器。分类器SVM1用来将可能包含横穿马路行人的图像与可能包含沿着马路行人的图像分开，事实上这里并没有将包含行人图像与不包含行人的图像分开，该分类器只是得到一种倾向，即得到的图像可能包含某一类行人或者不包含行人，剩下的分类工作由它后面的节点来完成。比如分类器SVM2就是用来进一步将这种倾向分类为包含沿马路行人的图像和不含行人的图像，而分类器SVM3用来将另一个分支的倾向细分为横穿马路行人的图像和不含行人的图像。

候选图像的分流流程可以这样描述：首先，它会被送往SVM1做分类。如果它被分类为包含横穿马路行人的图像，那么它会接着被送往SVM2做分类，经过这个分类器后就能确定它到底包含横穿马路行人的图像还是不含行人的图像。类似的，如果它在第一步被分类为包含沿马路行人的图像，那么它会被接着送往SVM3做分类，经过该分类器后才能确定它到底包含沿马路行人的图像还是不含行人的图像。所以每幅待分类图像会被分类两次才能最终确定它的类别。

上面描述的多类分类器结构有如下两个优点：

- (1) 同以往类似的工作相比，我们设计的多类分类器具有更强的分类能力。如果我们使用[3]中的结构，某幅包含行人的图像更有可能会被两个分类器都识别出来，从而得不到准确的行进状态。这是因为[3]中的分类器结构相当于我们的多类分类器结构去掉了分类器SVM1，某幅待分类图像会被两个分支同时

分类，试想在我们的结构中去掉SVM1后，一幅包含横穿马路行人图像同时被SVM2与SVM3分类，毫无疑问，它会在SVM2中得到自己准确的分类信息，然而在SVM3中，它也会有很大的可能性也被识别为正样本。如果我们使用[4]中的串联结构，那么某幅包含行人的图像可能会被前面的分类器识别为正样本，得到错误的行进状态。所以要想用多个两类分类器组合使用得到行人的行进状态，必须在第一步就将两种行进状态行人分开，至于那些本来就不包含行人的图像则无须在意它们被分到哪一类，因为它们会在接下来的分类当中被过滤掉。

- (2) 本章介绍的多类分类器只增加了些许的计算时间代价就可以获得行人行进状态。事实上，本章讨论的多类分类问题会使得两种行进状态的行人集合更加紧致，这样会使得用来识别它们的特征集合规模变小，从而每个两类SVM分类器的分类速度都会上升，最后的结果就是总的分类时间没有明显增加。在实验中还显示了一个现象：每个两类分类器的支持向量数目都明显小于以前系统中的单分类器的支持向量数目。

2.6.3 实验与分析

本次实验的平台同上次一样，但是训练样本集合有不同，手工分为三类样本：包含沿着马路行人的图像，包含横穿马路行人的图像，不含行人的图像。

表2-2中是我们以前采用单SVM分类器和本章采用多个两类SVM分类器的支持向量数目与特征数目。其中可以看出，多类分类器结构中每个两类分类器的支持向量数目都明显少于以前将行人全归为一类的SVM分类器。所用来训练的特征集合规模也有极大缩水，这样总的检测时间不会有明显增加。在保证检测速度的同时，检测率也有不错表现，这一点在表2-3中可以看到。

表2-2 不同SVM分类器的支持向量和特征数目

Type of classifiers	Positive SV	Negative SV	Feature (shape/motion)
SVM1	116	431	100/30
SVM2	125	416	110/20
SVM3	142	165	60/60
Single SVM	420	983	200/50

如表2-3所示，在三段不同的测试视频中，多类分类器结构都表现出了良好的检测性能，平均来说PDS的检测率可以达到89.6%，而误报率仅为0.13%。当行人被检测到时，他的行进状态被正确识别的几率为96.4%。这充分说明了将两类

SVM分类器组成多类分类器方法的有效性,改进后的PDS可以在无损检测速度的同时得到额外的行人行进状态。

表 2-3 多类分类器的检测性能

Test video	Detection rate (%)	False positive rate (%)	Moving style detection rate (%)
1	88.9	0.12	93.3
2	90.4	0.14	97.8
3	89.5	0.13	98.1

图2-10展示的是第一段测试视频中某一帧的检测结果。其中共有三位行人被检测出来,并被标记为不同的行进状态。黄色表示可能沿马路的行人,红色表示可能横穿马路的行人,汽车正前方的那位行人被准确识别出行进状态。

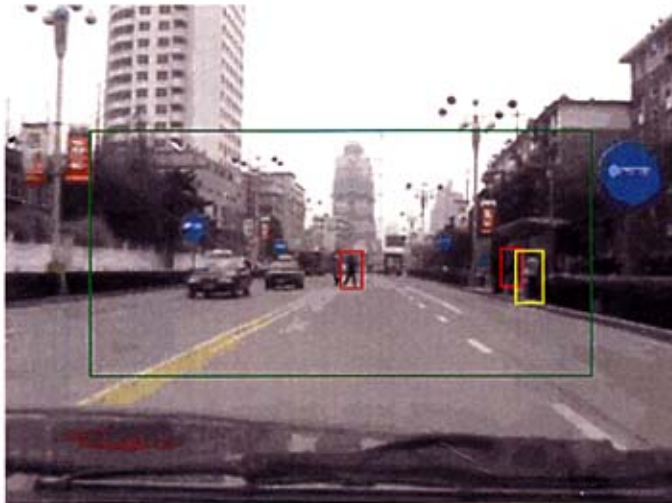


图 2-10 检测行人的同时得到他们的行进状态

通过对实验结果的分析,我们可以发现通过实现这种多类分类器结构,每个两类SVM分类器可以用更加紧密的训练样本集合和特征集合来训练,这能够同时减少训练和分类的时间。然而,当该多类分类器结构中拥有过多的两类分类器时,检测速度可能会迅速下降,因为一个候选图像会被检测很多次。在我们前面的工作中采用一个SVM分类器来检测一幅图像是否包含行人,由于行人这一类内部不同的行进状态,同本章中的任何一个两类分类器相比,它具有大得多的支持向量数目以及特征集合,更长的训练时间以及更慢的检测速度,同本章中两类分类器组成的多类分类器相比它仅在总的检测时间上占有一点优势,检测率在不

同的测试视频中互有高低。因此采用本章介绍的多类分类器替换原来PDS中的单分类器可以在稍微牺牲检测速度的情况下拥有更强的检测能力。

2.7 本章小结

本章的主要内容是为 PDS 设计一个级联分类器。它的分类模块主要由两部分组成：第一部分是一个统计学习分类器，采用 Adaboost 算法训练得到，主要作用是用来对每一帧中所有的待检测图像进行初选，并且能得到一个初选的特征集合，集合中的每个特征都具有在总排名中较靠前的分类能力；第二部分是 SVM 分类器，采用 Joachims 的 SVM^{light} 算法训练得到，它是决定 PDS 性能的核心分类器，此分类器的作用是将前面统计学习分类器初选出来的那些图像再次进行过滤，以决定它们是否真正包含有行人。此级联分类器在测试中显示出可以接受的速度以及不低于采用类似方法实现检测工作的检测率。

同以往的工作相比，我们的级联分类器有如下优点：

- (1) 分类器的级联有着更高的灵活性，可以取得正确率和速度的平衡，就是说在保证正确率的情况下可以调节检测速度。
- (2) 在训练统计学习分类器和 SVM 分类器时利用了包含运动信息的特征，这些运动特征易于计算，成为以往广泛使用的外形特征的合理补充。实验结果显示，运动特征的合理使用能够在一定程度上抵消城市交通环境复杂背景给行人检测带来的困难，明显提高正样本检测率。

在下一步工作当中，我们需要更深入的研究运动特征对行人识别的影响。同时，为了进一步提高检测速度需要研究更快的特征计算方法，以及如何用尽可能少的特征达到同样高的正确率，如何增加新的单个分类器完善当前的级联结构。

第三章 基于进化 SVM 的行人检测算法

PDS 最重要的目标是快速准确的检测出行人，提高 PDS 的检测性能是 PDS 研究中的核心课题。作为我们设计的 PDS 中最核心的部分，SVM 分类器的性能成为影响 PDS 检测性能的关键。

本章第一部分首先讨论了 SVM 分类器的训练模型优化的常用方法，然后给出了用 GA 去进化 SVM 的训练模型的具体方案。

3.1 现有 SVM 训练模型的获取方法

一个基于 SVM 分类器的 PDS 的检测性能很大程度上由 SVM 分类器的分类能力所决定，而 SVM 分类器的检测能力又在很大程度上受训练模型的影响（比如：核函数类型，核函数参数以及错分样本的惩罚参数等等），所以选择合适的训练模型对整个 PDS 的性能提升是非常关键的。然而从大部分文献中得到的信息来看，绝大部分 PDS 采用 SVM 做分类的时候，选择训练模型依靠的是经验或者几次简单的实验尝试。因此大部分 SVM 分类器都没能得到最优的训练模型，从而 PDS 的性能也没有达到最佳。

一般来说有三种方法常用来选择 SVM 的训练模型。第一种方法是最小化 radius-margin (RM)界[39]，常用的寻优算法为梯度下降法。当最小的 RM 界找到时，也就找到了近似最优的训练模型。一般做法是将 RM 界用训练模型的参数表示，然后用梯度下降法求解。

第二种方法是实验寻优法，如留一法(LOO: Leave One Out)，交叉验证法(cross validation)[40]。常用的做法是将样本等分为 n 份，然后依次将每一份作为测试样本，其余作为训练样本，这样得到 n 个分类器，并且可以测出其对应的检测率，然后求平均得到按经验估计的训练模型所对应分类器的检测率，当 $n \rightarrow \infty$ 时，这个均值就是该分类器的期望检测率。可以多次估计训练模型并通过交叉验证法得到近似检测率，然后选取较好的训练模型。

第三种方法是进化算法，其中又以遗传算法(GA: Genetic Algorithm)最为常用。GA 的基本思想是模拟自然界物种进化的方式来不断寻找更优的个体，已经形成了一套比较完整的理论体系。GA 提供了一种求解系统优化问题的通用框架，它不依赖于问题的具体领域，具有较强的搜索全局最优解的能力，已经广

泛应用到很多学科。训练一个性能良好的 SVM 分类器需要选择合适的训练模型。这个问题涉及到在一个较大的范围内搜索全局最优解，而 GA 的特点正好适合来解决这两个问题。

第一种方法从数学上考虑比较自然，然而实际求解过程非常复杂，而且不一定能得到全局最优解。第二种 LOO 方法需要多次实验，实际计算量非常大，也不一定得到全局最优解。本文采用 GA 来进化 SVM 的训练模型，从实验过程及结果来看，整个自进化的训练算法具有较强的鲁棒性，是一种合适的优化方法。

为了更简单有效的寻找最优训练模型，我们将训练模型参数编码，主要包括核函数的参数以及惩罚参数的编码，并用 GA 来寻找最优的个体。适应度评估函数采用 Joachims 的 $\xi\alpha$ -estimate 方法[18]。接下来内容是关于自进化 SVM 分类器训练的详细介绍。

3.2 进化 SVM 分类器的训练模型

为了使用进化算法得到一个性能良好的 SVM 分类器，我们首先需要将 SVM 的训练模型参数化，然后需要找出一个能合理评估 SVM 分类器性能的适应度函数，接着就是考虑进化算法的编码方案，以及遗传算子的选择。

3.2.1 采用的 SVM 训练模型

前面关于 SVM 的介绍中已经提到，求解一个 SVM 分类器的问题对应求解一个二次优化问题，其中一个约束条件是 $\forall i: 0 \leq \alpha_i \leq C$ ， C 限制了一个样本对最后分类超平面的影响，因此需要将它作为训练模型参数的一部分。同时，核函数决定了分类超平面的复杂程度，对分类性能有着极大的影响，因此核函数的类型和参数也应该作为训练模型参数的一部分。我们常用的核函数为多项式核函数 $k_{poly}(x, y) = (xy + 1)^d$ 与 RBF 核函数 $k_{RBF}(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$ ，它们对应的训练模型分别为 $M_1 = \{d, C\}$ 与 $M_2 = \{\sigma, C\}$ ，其中 d 为正整数， σ 为正实数。接下来我们会分别进化这两种模型。

3.3.2 采用的 SVM 性能评估方法

当我们使用 GA 来进化 SVM 训练模型时，需要对 SVM 分类器进行多次训

练和性能评估,所以需要一种高效的方法来评估分类器的性能。通常的做法是将样本划分为训练样本和测试样本。留一法和交叉验证法就是这类方法的代表,它们采用多种划分方式,而每种划分都对应一次 SVM 分类器的训练和测试。这种做法有一个最大的弊端就是计算代价太大。因为用 GA 做进化时,它的每一代包含有很多个体,每个个体对应一个训练模型,如果对每个训练模型都按照留一法和交叉验证法来进行评估,计算量是相当大的。

为了克服这个困难,我们选取 Joachims 的 $\xi\alpha$ -estimate 方法[41]并结合 PDS 的特点做了适当修正来进行性能评估。该方法的主要特点是不需要多次划分训练集和测试集,可以在一次训练后得到近似留一法的估计结果。

一共有三个指标可以被估计:第一种是错误率 $Err(f(x))$, 定义为一个样本被错误分类的概率;第二种是查全率 $Rec(f(x))$, 定义为一个正样本被正确识别也就是 $f(x)=1$ 的概率;第三种是查准率 $Prec(f(x))$, 定义为一个分类为正样本的样本 ($f(x)=1$) 实际上为正样本的概率,其中

$$f(x) = \text{sgn}\left(\sum_{\text{Support Vector}} y_i \alpha_i k(x_i, x) - b\right), \text{代表决策函数。}$$

PDS 最重要的目标是检测出行人,但是这又不能以增加误报率为代价。幸而在我们的级联分类器结构中,前面的统计学习分类器已经过滤掉了大部分不包含行人的图像窗口,所以位于后端 SVM 分类器可以适当牺牲一部分检测率来提高行人的识别率,这一思想反映在指标上就是用错误率和查全率来综合评价 SVM 分类器的性能。

为此我们定义了如下的适应度评估函数:

$$G(f(x)) = (1 - \lambda) \cdot (1 - Err(f(x))) + \lambda \cdot Rec(f(x)) \quad (21)$$

其中,参数 $\lambda \in (0,1)$ 是错误率 $Err(f(x))$ 和查全率 $Rec(f(x))$ 的折中,在我们的实验中, λ 的值一般不超过 0.5。

从上述公式可以看出,在行人检测这一特定问题中,为了提高行人的检测率,需要在分类器性能评估中额外增加查全率所占的权重,但由于仍然需要保持较低的误报率,所以最后取它们的一个折中来评估分类器性能。

3.3.3 训练模型的优化流程

1. 编码和解码

在我们的优化问题中要求的精度不是太高,所以我们采取二进制编码方案。对与训练模型 $M_1 = \{d, C\}$ 来说,我们设定 $d \in [1,5]$, $C \in [0.01,10.24]$, 多项式核函数次数 d 的编码长度为 2。对训练模型 $M_2 = \{\sigma, C\}$ 来说,我们设定 $\sigma \in [0.01,1.28]$, $C \in [0.01,10.24]$, RBF 核函数的参数 σ 为 7, 惩罚参数 C 在两种模型中的编码长度

一样，都为 10。在两种模型中，我们都将两种参数连接在一起来表示种群中的一个个体。下面是分别代表两种训练模型的编码个体：

$$P: \underbrace{101001100101}_{p_1} \quad \underbrace{00110010001011101}_{q_1} \quad \underbrace{}_{p_2} \quad \underbrace{}_{q_2}$$

图 3-1 P 和 Q 分别代表 M_1 和 M_2 的编码个体

如上图所示，对 P 来说，最左边的两位代表多项式次数 d ，其他的十位代表 C 。对应的解码方案为 $d = 1 + p_1$ 和 $C = 0.01 \cdot p_2$ ，其中 p_1, p_2 是与图中它们上面的二进制串对应的整数。对 Q 来说，最左边的七位代表 σ ，其余七位代表 C 。对应的解码方案为 $\sigma = 0.01 \cdot q_1$ 和 $C = 0.01 \cdot q_2$ ，其中 q_1, q_2 是与图中它们上面的二进制串对应的整数。

2. 适应度函数

适应度函数在 GA 中用来评价个体的好坏，适应度高的个体得到更多的保留机会，可以说适应度函数设计的好坏在很大程度上决定了最后得到的进化结果。在本文中，我们按照上述的解码方案得到个体对应的训练模型，然后用 Joachims 的分解方法进行训练得到分类器，同时也得到了上面提到的 $\xi\alpha$ -estimate 方法的估计指标，这样就可以将个体的适应度与分类器的性能对应起来。

简单的说就是对个体 m ，它有一个分类函数 $f(x)$ ，我们定义它们之间的映射关系为 $h: m \rightarrow f(x)$ ，这样就可以计算到 m 的适应度：

$$Fitness(m) = G(h(m)) = (1 - \lambda) \cdot (1 - Err(h(m))) + \lambda \cdot Rec(h(m)) \quad (22)$$

3. 算法流程

有了以上的准备，我们就可以用遗传算法来进化 SVM 训练模型了，这里采用基本遗传算子，其中交叉算子为单点交叉，变异算子为固定变异率的二进制变异。大致的进化流程如下所示：

- (1) 按先前制定的二进制编码方案残生初始种群。初始化种群代数 $gen = 0$
- (2) 将种群中个体按照先前指定的解码方案解码得到对应的训练模型。
- (3) 为每个个体 m 计算适应度 $Fitness(m)$ 。
- (4) 使用基本遗传算子来产生下一代个体，交叉率 $P_c = 0.5$ ，变异率 $P_m = 0.006$ ，交叉方式为单点交叉。
- (5) 将种群代数加 1， $gen = gen + 1$ ，当 $gen = 100$ 时，终止算法，否则转到步骤 (2)。

3.4 实验与分析

本节主要是对 GA 进化 SVM 分类器训练模型的一些实验和分析，在简单的介绍实验背景之后，我们重点关注两个对比实验，它们分别得出了下面的两条结论：

- (1) GA 进化的 SVM 分类器具有更好的检测性能。
- (2) RBF 核函数比多项式核函数更适合用在行人检测的分类器当中。

3.4.1 实验准备

本次实验中，我们的 PDS 运行在 PC 机上，CPU 型号为 Barton2500+，随机访问存储器为 1G DDR。所有的训练样本来自十段视频，这些视频拍摄于白天真实的城市交通环境中，视频画面大小为 320×240，汽车平均车速约为 25 公里/小时。

本次实验采用原来的方案从样本中提取外形和运动信息来组成特征向量，然后用 Joachims 的分解算法和性能估计方法来训练分类器以及测试 GA 优化训练模型算法有效性。

为验证进化 SVM 分类器的检测性能，我们完成了两个个大的对比实验。其中，第一个实验是进化训练模型的 SVM 分类器与一般人工选择训练模型的 SVM 分类器的比较，通过对比它们的性能评估 G 值与实际测试中的检测率，可以看出进化训练模型的 SVM 分类器具有明显的检测性能优势；第二个实验是采用 RBF 核函数的 SVM 分类器与采用多项式核函数的 SVM 分类器的比较，通过比较它们在进化过程中的性能提升，以及实际测试中的 ROC 曲线，可以看出采用 RBF 核函数的 SVM 分类器具有更强的检测能力与可优化潜力。

3.4.2 进化 SVM 分类器与一般分类器的比较

在实验中，我们分别采用性能估计公式中的 G 值和实际的检测率来对比进化的 SVM 分类器和一般分类器。

首先，我们随机选取了 6 个训练模型，并通过训练得到它们对应的 SVM 分类器，这时可以用性能估计公式计算各自的 G 值，然后用 GA 对训练模型进化得到对应的进化 SVM 分类器并计算出 G 值。如表 3-1 所示，无论是采用多项式核函数还是 RBF 核函数，进化的 SVM 分类器总是拥有比一般 SVM 分类器更高的 G 值。

表 3-1 不同分类的 G 值

	Evolutionary classifier	Normal classifier 1	Normal classifier 2	Normal classifier 3
G (Poly)	0.841	0.725	0.697	0.806
G RBF)	0.931	0.624	0.861	0.813

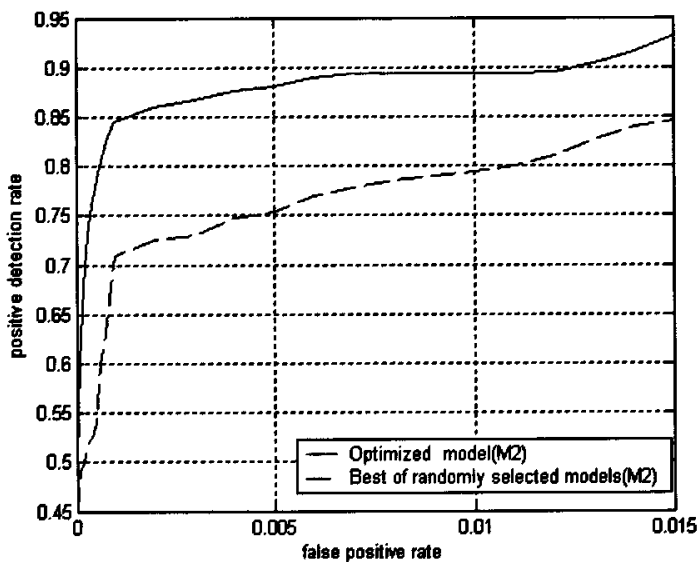


图 3-2 进化 RBF 核函数的分类器与一般 RBF 核函数分类器的 ROC 曲线

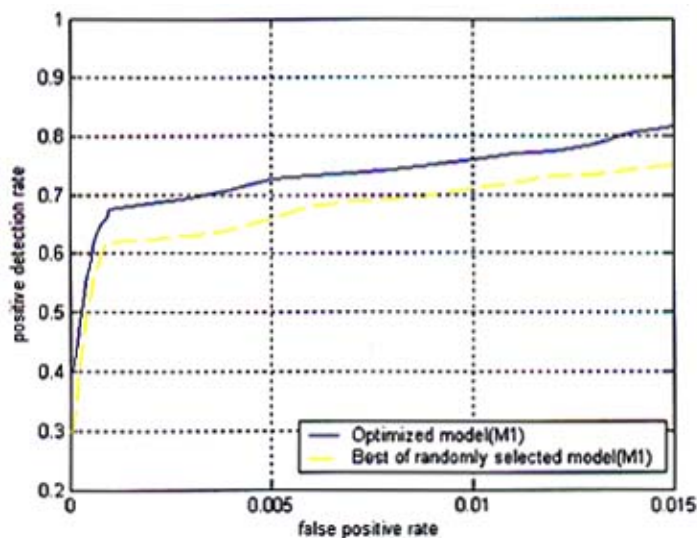


图 3-3 ROC 进化多项式核函数的分类器与一般多项式核函数分类器的 ROC 曲线

在用训练时得到的指标对分类器性能进行估计以后，我们又用测试样本对 SVM 分类器性能进行了实际意义上的测试。为了更加直观的检视分类器的性能，非常有必要分析反映正样本识别率和误报率折中关系的 ROC 曲线。如图 3-2 所示，ROC 曲线显示，在同样的误报率下，进化的 SVM 分类器（基于 RBF）比我们随机选择训练模型得到一般的 SVM 分类器（基于 RBF）的正样本识别率要高。类似的结论可以在进化的多项式核 SVM 分类器和一般多项式核 SVM 分类器的比较中得到，如图 3-3 所示。

在图 3-2 中，绿色的曲线表示随机选取 RBF 核函数训练模型的分类器的检测率和误报率的对应关系，红色的曲线表示进化选取 RBF 核函数训练模型的分类器的检测率和误报率的对应关系。

在图 3-3 中，黄色的曲线表示随机选取多项式核函数训练模型的分类器的检测率和误报率的对应关系，蓝色的曲线表示进化选取多项式核函数训练模型的分类器的检测率和误报率的对应关系。

另外，我们还测试了性能估计公式中参数 λ 对自进化 SVM 分类器的影响。如图 3-4 和图 3-5 所示， λ 的取值确实影响到训练模型的选择，并对最终进化得到的 SVM 分类器有着明显的影响。在图 3-4 中检测出了正前方的行人，在图 3-5 中不仅检测出了那个正前方的行人，还检测出了路边的一位行人，但是也出现了一个误检测。从理论上分析，当 λ 变大时，自进化的 SVM 分类器应该有更高的正样本识别率，但是相应的误报率也会有所上升，这同我们得到的实验结果是十分吻合的。因此，不同的 λ 取值将导致最后自进化 SVM 分类器有不同的检测倾向，我们需要在识别率和误报率之间达到平衡以适应 PDS 的要求。在本次实验中，在较低的精度下，我们试验了一些 λ 取值，最后发现当 $\lambda = 0.3$ 时效果较好，关于 λ 值能够接受的区间为 $[0.1, 0.5]$ 。



图 3-4 选取 $\lambda = 0$ 训练的分类器的检测结果

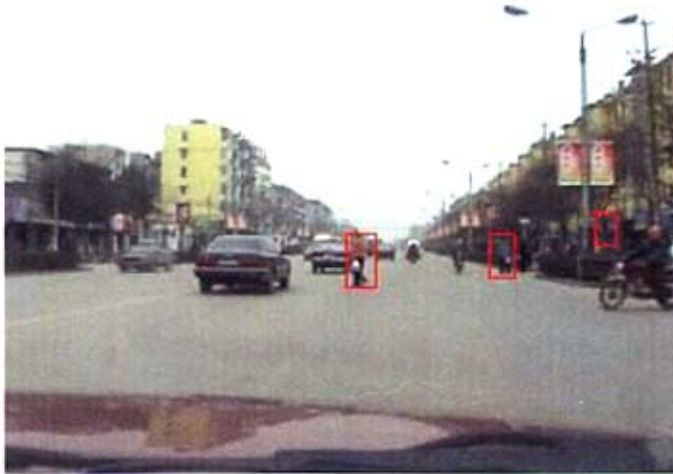


图 3-5 选取 $\lambda = 0.4$ 训练的分类器的检测结果

3.4.3 选择不同核函数的 SVM 分类器性能分析

首先, 让我们来比较分别采用多项式核函数和 RBF 核函数的 SVM 分类器优化结果。如图 3-6 所示, 横轴表示 GA 中进化的代数, 纵轴表示当前代数中最佳训练模型对应分类器的估计检测能力 (G 值)。红色曲线代表采用 RBF 核函数的训练模型 M_2 的优化过程, 蓝色代表采用多项式核函数的训练模型 M_1 的优化过

程。从图中可以清楚的看到，两种优化过后的训练模型都能够增强它们对应分类器的检测能力，对 M_1 的优化只是些许增加了代表分类器检测能力的 G 值，而对 M_2 的优化则十分明显的增加了其对应 G 值。训练模型 M_1 最终的优化结果为 $\{2,1.13\}$ ，对应的 G 值为 0.841，训练模型 M_2 的最终优化结果为 $\{0.47,0.09\}$ ，对应的 G 值为 0.931。这些实验结果显示，基于 RBF 核函数的 SVM 分类器在行人检测应用中具有更好的性能，而且能够通过进化方法来得到更多的性能提升。

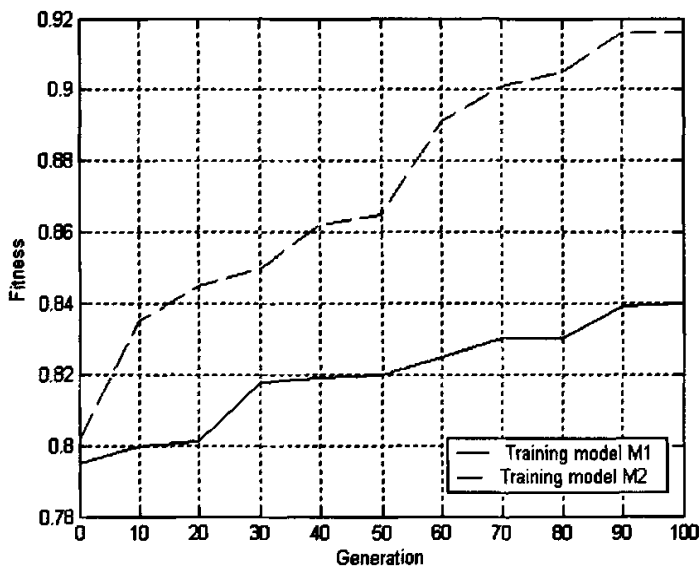


图 3-6 两种训练模型 M_1 和 M_2 的进化过程曲线

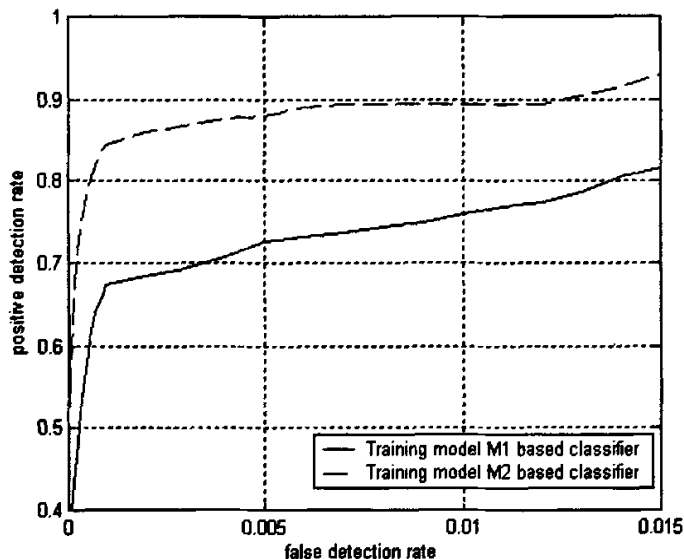


图 3-7 两种优化过的分类器的 ROC 曲线

接下来，我们来分析采用两种不同核函数的 SVM 分类器的实际检测能力。为了对它们进行更加直观的比较，我们画出了两个分类器对应的 ROC 曲线。从图 3-7 中可以看出，在相同误报率下，采用 RBF 核函数的 SVM 分类器具有更高的正样本识别率。这一结果同图 36 和表 3-1 中显示的性能估计结果非常吻合。它们都显示了采用 RBF 核函数的 SVM 分类器的优势。

3.5 本章小结

回顾本章，我们为单光学摄像头 PDS 设计了一种自进化的 SVM 分类器。主要思想使用 GA 来优化代表这一模型的参数，从而得到最优的 SVM 训练模型。其中一个有特色的地方是针对行人检测这一特殊问题，利用 Joachims 设计的分类器评测指标求加权和来估计分类器性能，这样不仅可以节省训练时间，还可以得到具有更好识别效果的 SVM 分类器。实验结果显示了这一方法的有效性，另外还显示了采用 RBF 核函数的 SVM 分类器具有更好的可优化性。

同以往的工作相比，自进化的 SVM 分类器有以下几个优点：

- (1) GA 能够无需人工干预的找到最优训练模型，使用该训练模型的 SVM 分类器可以达到更高的检测性能。
- (2) 前面的工作中，基于 SVM 分类器的 PDS 一般使用多项式核函数训练模型来训练分类器，然而在我们的实验中，使用 RBF 核函数的 SVM 分类器经过进化后可以取得明显高于多项式核函数 SVM 分类器的检测性能。

第四章 特征集合与 SVM 训练模型联合优化

为了提高 SVM 分类器的性能，需要解决的问题主要有两个。第一个问题是选择特征集，为了将候选图像中的行人快速准确的识别需要合适的特征集，特征集过于庞大影响分类速度，太少又失去分类的准确性。另一个问题是选择合适的训练模型，包括核函数及其主要参数和惩罚错分样本的折中参数的选择。

在现有的大多数 PDS 中，特征集和 SVM 训练模型都是通过经验或简单实验来给出，这样并不能给出最优的特征集和 SVM 训练模型。更为重要的是，这两个问题并不是孤立的，这是因为特征选择后也就在一定程度上限制了核函数和特征空间，所以特征集和 SVM 训练模型需要放在一起考虑。

本章采用 GA 来同时解决特征集和 SVM 训练模型的优选问题。特征集包括一个由 Adaboost 算法粗选的同时具有形状特征和运动特征的集合；训练模型用核函数参数和惩罚错分样本的折中参数来表示。每个进化的个体由这两类参数共同组成，采用 $\xi\alpha$ -estimate 方法来构造一个 SVM 分类器的个体的性能评估函数。

4.1 联合优化的原因

SVM 分类器性能的关键在于核函数的构造，核函数的类型，参数如何选择。比如说常用的多项式核函数，多项式的次数如何选择，还有 RBF 函数的宽度参数如何选择等等。另外惩罚错分样本的折中参数也有部分影响。目前国际上还没有形成一个统一的模式，也就是说最优 SVM 算法参数选择还只能是凭借经验、实验对比、大范围的搜寻或者利用软件包提供的交互检验功能进行寻优。而另一方面，对 SVM 分类器性能的优化大都将特征选择和训练模型选择分开来进行，这样做的不足之处在于没有考虑特征集和 SVM 分类器之间的内在联系。因为特征集合选定后，也就潜在的限定了特征空间，也就限定了最合适的核函数以及折中参数，所以对特征集和训练模型的优化必须同时进行。

到目前为止，评价 SVM 的性能有很多方法，比如理论上对 SVM 错分率上界的评估[3]，实验方法如留一法对 SVM 分类器检测率的评估[4]，Joachims 的 $\xi\alpha$ -estimate 方法[7]等等。虽然理论方法和留一法的评估都是无偏估计，但是计算上的复杂以及时间上的代价使得它们不如有偏估计 $\xi\alpha$ -estimate 方法更有吸引力，特别是当应用到下面一节所要描述的进化算法中时，Joachims 的方法效率会更高，这一点在我们的实验中得到了验证。

4.2 联合优化特征集合与 SVM 训练模型

训练一个性能良好的 SVM 分类器需要选择合适的特征集合和合适的训练模型。这两个问题的解决都涉及到在一个较大的范围内搜索全局最优解，一般的实验方法计算代价大而且结果不能取得全局最优，而 GA 全局搜索能力较强以及无须人工干涉的特点正好适合来解决这两个问题。

使用 GA 来联合优化 SVM 特征集和训练模型需要解决两个问题：

- (1) SVM 性能如何表述，由于使用 GA 需要多次进行 SVM 训练，所以需要选择一个高效的描述方法。
- (2) 如何用 GA 进化的个体描述 SVM 训练模型。

对于第一个问题，我们仍然采用 $\xi\alpha$ -estimate 来描述 SVM 分类器的性能。 $\xi\alpha$ -estimate 最大的优势在于它可以在一次 SVM 训练过后就能够得到识别率，误报率，比起留一法估计要节省大量的时间。对于第二个问题，本章只讨论使用 RBF 核函数的 SVM 分类器训练，具体做法将核函数距离参数 σ 和折中参数 C 组成 SVM 的训练模型，它们和代表特征选择的向量一起组成在 GA 中进化的个体。

基本遗传算法容易产生早熟和局部收敛的情况，本文采用自适应的遗传算法来进行特征选择和训练模型选择。本文所采用遗传算法的各方面特性如下：

(1) 个体编码与解码

一个个体由若干个基因组成，每个基因代表求解问题的一个参数，本优化问题有三个待优化参数：特征选择参数、核函数参数 σ 以及惩罚错分样本的折中参数 C 。

$$Q: \underbrace{11\dots 01001110100001011101}_{q_1} \underbrace{}_{q_2} \underbrace{}_{q_3}$$

图 4-1 联合优化的个体编码

解码时将各字段的二进制分别转化为特征编号，核函数参数 σ 以及折中参数 C 。

(2) 适应度函数选择：

适应度函数求值是一个将 SVM 训练和 SVM 分类器性能评估结合在一起的过程。对每个个体先采用 SVM^{light} 算法训练，然后采用 $\xi\alpha$ -estimate 方法来计算每个个体所训练出分类器的 Error 和 Recall，然后用它们的加权和来充当个体的适应度。计算公式仍然采用上一章中的性能估计公式。

(3) 遗传算子：

由于采用 SGA 易产生早熟和局部收敛现象, 本文采用一种自适应遗传算法来对 SVM 模型参数进行选取, 最后获得最佳训练模型。具体实施其来就是采用自适应的交叉和变异概率。交叉概率 p_c 和变异概率 p_m 越大, 算法产生新个体能力越强, 个体适应度波动越大。若 p_c 和 p_m 越小, 算法产生新个体能力越弱, 个体适应度波动越小。

定义 f_{\max} 和 f_{avg} 为种群最大适应度和平均适应度, f_c 为两个交叉个体适应度较大者, f_m 为变异个体适应度。 $f_{\max} - f_{\text{avg}}$ 在一定程度上代表种群的稳定性, 为了避免早熟, p_c 和 p_m 应跟 $f_{\max} - f_{\text{avg}}$ 成反比。而当种群接近全局最优时, 这种反比关系又会破坏个体的稳定性, 所以 p_c 和 p_m 应分别与 $f_{\max} - f_c$ 和 $f_{\max} - f_m$ 成正比, 这样可以降低最优个体不被破坏的概率。

这种关系可以用以下公式来描述:

$$p_c = k_c(f_{\max} - f_c)/(f_{\max} - f_{\text{avg}}) \quad (23)$$

$$p_m = k_m(f_{\max} - f_m)/(f_{\max} - f_{\text{avg}}) \quad (24)$$

其中, 参数 k_c 和 k_m 为小于 1 的常数。采用上述策略时, 对同一代种群中的个体选取了不同的 p_c 和 p_m , 适应度高的个体变化概率减小, 适应度低的个体变化概率增加, 从而实现了参数的自适应选取。

4.3 实验与分析

本次实验的系统是基于我们以前的基于单光学摄像头的 PDS[19], PC 配置为 Yanoh T2300, 1G DDR2 RAM, 实验视频共 10 段, 使用 sony 光学摄像机在真实城市交通环境中采集。

本次实验的目的是找出最优的特征集和 SVM 训练模型, 对 GA 来说就是找出具有最大适应度的个体。其中, SVM 分类器采用 RBF 核函数。下面是三个比较有说明意义的对比实验。

(1) 联合优化前后的检测结果对比

本次实验将特征选择和 SVM 训练模型优化分开进行的做法与联合优化的做法进行了详细的比较。我们发现联合优化的做法在提高分类器性能的同时可以有效压缩特征集合, 大概用原来一半数量的特征可以得到更好的性能, 具体实验结果如下表所示:

表 4-1 联合优化前后的 SVM 分类器特征集合与性能

Classifier	Feature set	Error	Recall	precision
SVM1	300	0.144	0.913	0.831
SVM2	138	0.137	0.927	0.895

在前面工作中，我们使用 Adaboost 算法选择出所需要的特征集合，然后利用这个特征集合构造后面 SVM 所需要的训练样本。事实上，若用 Adaboost 算法选出的特征集合不同，那么用相同的图像样本训练出的 SVM 分类器是不同的。所以先确定特征集合，再用 GA 对 SVM 训练模型进行参数选择得到的模型仅仅是对当前特征集合近似最优，是一个局部的近似最优解。而特征集合和训练模型联合优化将特征和 SVM 分类器联系起来，能够得到一个相对全局的近似最优解。另外一个好处则是经过联合优化后，特征集合的规模往往会变小。在我们的实验结果中，特征集合规模缩小了大约一半，但是 SVM 分类器的性能却得到了提高。这意味着在缩短检测时间的同时，提高了分类器的检测率。所以我们得出的结论是特征集合和 SVM 训练模型的联合优化对 PDS 性能的提升具有十分重要的意义。

(2) 适应度函数合理性考察

通过对测试样本的检验，我们发现采用 $\xi\alpha$ -estimate 方法得到的 Error 略高于实际检测中的 Error，Recall 略低于实际检测中的 Recall，总体来说能够比较真实反映个体的好坏。由于设计 PDS 最重要的目的是保护行人，准确的识别出行人最为关键，因此对 Error 和 Recall 采用比例加权得到适应度 G 值来评价个体更为合理。而这个权值就是 G 值估计公式中的 λ ，本次实验中取值为 0.4。

下表是采用进化模型和实验估值模型训练得到的分类器性能评估 G 值和实际测试中检测率的对比：

表 4-2 个体适应度与分类器检测率的比较

	Evolutionary Classifier(RBF)	Normal classifier(RBF)	Normal classifier (Poly)
G (RBF)	0.938	0.861	0.813
Detection rate	0.915	0.843	0.807

从实际的检测结果和个体适应度的关联可以看出，我们选取的适应度评估函数能够较真实的反映个体的好坏，适应度越高的个体对应训练得到的分类器检测率越高

(3) 联合优化的检测能力与单纯估计方法的对比

我们进化得到的 SVM 模型与传统的估计方法的结果比较。从下图中可以看出传统的实验估计方法所得到的训练模型不容易达到最优, 而利用 GA 搜索全局最优解的能力可以得到一个较好的训练模型, 从而使训练出来的 SVM 分类器具有最好性能。本文中采用特征集合和训练模型的联合优化进一步提升了这种方法的有效性。

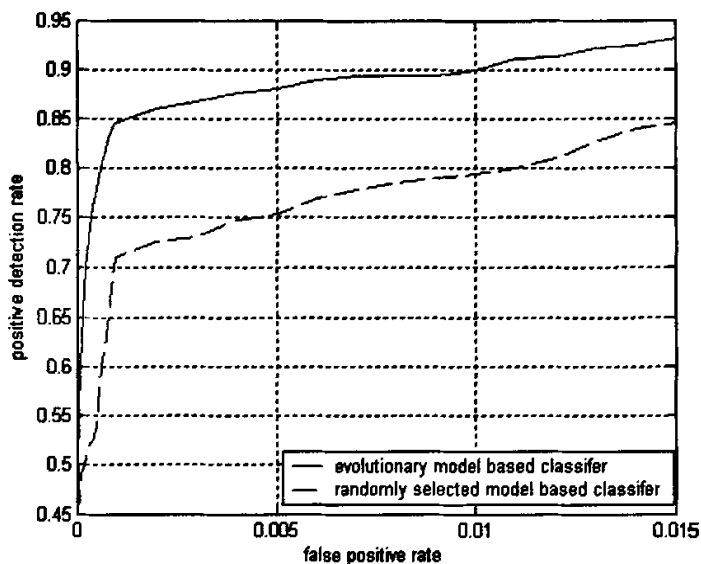


图 4-2 进化模型和估值模型训练出的分类器的 ROC 对比

4.4 本章小结

为了进一步提高 SVM 分类器的检测能力和检测速度, 我们分析了特征集合与 SVM 分类器的关系, 决定采取另外一种自进化的训练方式。具体做法是在训练中不直接使用前面工作中用 Adaboost 算法得到的特征集合, 先将代表特征选择的向量与 SVM 训练模型联合编码, 然后使用 GA 对其进行联合进化, 最后得到一个精简的特征集合 (大概能提升 80% 的检测速度), 并且分类器的检测性能也有提升。

和以往工作相比, 本章主要贡献有:

- (1) 实现了特征集合和 SVM 训练模型的联合优化, 比单独进化训练模型得到的分类器性能更好, 具有更高的识别率, 更低的误报率, 同时由于特征集合规

模的减小，使得分类器识别速度大大提高。

- (2) 用 GA 得到 SVM 训练模型不需要人工干预，所采用的性能评估方法比起 LOO 大大减少了计算时间，也没有梯度下降法计算上那么复杂，减小了获取最佳性能分类器的代价。

第五章 总结与展望

5.1 本文的主要工作

随着社会文明程度的提高,对行人的安全保障关注度也会越来越高,具备行人检测与辅助驾驶能力的行人检测系统已成为目前国内外智能交通领域的一个重要研究课题。

行人检测系统是一个利用安装在运动车辆上的摄像头、雷达或其他传感设备获取车辆前方的图像或者距离信息,然后根据获得信息从视频序列中检测出行人的系统。然而,近年来出现的实用系统一般采用昂贵的传感设备,并且只装备在部分高档车型上,这不利于行人检测系统的推广。因此,基于简单传感设备的行人检测技术研究具有重要的理论研究意义和很高的实际应用价值。

本文就行人检测系统中的检测技术以及分类器优化展开了讨论,提出并实现了一个以 SVM 分类技术为核心的级联分类器。在级联分类器的第一层采用 Adaboost 方法训练一个弱分类器的集合来过滤待检测图像,同时第二层采用 SVM 方法训练一个分类器来最终确定待检测图像的类别。其中,用来分类的特征集合包括外形和运动两种特征,外形特征中包含的是一种短周期的运动信息,这种特征的加入在一定程度上可以消除城市交通复杂背景纹理给传统外形特征检测带来的困难。

同时,为了改进 SVM 分类器的性能,本文采用 GA 来优化 SVM 分类器训练模型。不同于传统方法中采用 LOO 计算分类器检测率的技术,本文利用 Joachims 的 $\xi\alpha$ -estimate 方法来估计 SVM 分类器的检测率,并针对行人检测这一特殊问题设计了相应的适应度来评估 SVM 分类器的训练模型。

在进一步的研究中,我们发现特征集合的规模以及外形和运动两种特征比例对 SVM 分类器性能有着很大影响。虽然在前面实验中使用的特征集合已经过了 Adaboost 方法初选,但是与 SVM 结合使用并不一定能达到最好的效果。因此,我们采用同样的 SVM 性能估计方法和 GA 完成了特征集合与 SVM 训练模型的联合优化。经过优化后的特征集合规模大约减少了一半,大大加快了检测速度,并且训练出来的 SVM 分类器性能还略有提升。

实验结果表明,本文实现的以 SVM 为核心的级联分类器具有较高的检测率和可以接受的误报率,基于该分类器的单光学行人检测原型系统在实际城市交通环境中具有较强的检测能力,经过进一步改进后应该具有较高的推广价值。

5.2 展望

在接下来的工作中,可以考虑更加深入地分析不同检测方法的特点,进一步细化级联分类器的模块结构,同时需要研究更加有效的特征表示与计算方法,实现多种信息的融合,增强检测算法的鲁棒性。

另外为了将本文实现的原型系统实用化,还需要在保证检测能力的基础上,为该系统实现行人跟踪、意图判别,以及分级预警等功能。

行人检测系统作为智能交通系统中的一个研究热点具有十分重要的社会意义。随着对主流传感器下检测技术的不断改进,行人检测系统在未来将具有广阔的发展前景。

参考文献

1. Paul Viola, Michael Jones, Daniel Snow., Detecting pedestrians using patterns of motion and appearance, *Mitsubishi electric research laboratories*. TR2003-90 2003, 734-741
2. Shashua A., Gdalyahu Y., Hayun G., Pedestrian detection for driving assistance systems: single-frame classification and system level performance, *IEEE Intelligent Vehicles Symposium*, Parma, Italy. 2004,1-6
3. T. Evgeniou C. Papageorgiou and T. Poggio., A trainable pedestrian detection system, *IEEE Intelligent Vehicle Symposium*, 1998, 10: 241-246
4. Grant Grubb, Alexander Zelinsky Lars Nilsson, Magnus Rilb, 3D Vision Sensing for Improved Pedestrian Safety, *IEEE Intelligent Vehicle Symposium*, 2004, 19-24
5. Fengliang Xu, Xia Liu, and Kikuo Fujimura, Pedestrian Detection and Tracking With Night Vision, *IEEE Transactions on Intelligent Transportation System*, 2005, 6: 63-71
6. Dai Cong-Xia, Zheng Yun-Fei, Li Xin., Layered representation for pedestrian detection and tracking in infrared imagery, *IEEE Conference on Computer Vision and Pattern Recognition*, 2005,13-13
7. Cheng Hong, Zheng Nan-Ning, Qin Jun-Jie., Pedestrian detection using sparse Gabor filter and support vector machine, *IEEE Intelligent Vehicles Symposium*, Vienna, Austria, 2005, 58-587
8. Dalai N, Triggs B., Histograms of oriented gradients for human detection, *IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, 2005, 886-893
9. M. Oren *et al.*, Pedestrian detection using wavelet templates, *IEEE Computer Vision and Pattern Recognition Conference*, 1997, 193-199
10. A. Mohan and T. Poggio, Example-based object detection in images by components, *IEEE Trans. Pattern Anal. Machine Intell.*, 2001, 23: 349-361
11. D. Gavrilu, Pedestrian detection from a moving vehicle, *European conference on Computer Vision*, 2000, 37-49
12. L. Zhao and C. Thorpe, Stereo and neural network-based pedestrian detection, *Intelligent Transportation Systems*, Tokyo, Japan, Oct. 1999, 1(3): 148 -154
13. Szaras M., Yoshizawa A., Yamamoto M., Ogata J., Pedestrian detection with convolutional neural networks, *IEEE Intelligent Vehicles Symposium*, Las Vegas, Nevada, 2005, 224-229

14. Franke U., Joos A., Real-time stereo vision for urban traffic scene understanding, *IEEE Intelligent Vehicles Symposium*, Dearborn, 2000, 273-278
15. Wöhler C., Kressler U., Anlauf J. K., Pedestrian recognition by classification of image sequences: global approaches vs. local spatio-temporal processing, *15th International Conference on Pattern Recognition*, Barcelona, Spain, 2000, 2: 540-544
16. Gavrila D. M., Giebel J., Munder S., Vision-based pedestrian detection: the protector system, *IEEE Intelligent Vehicles Symposium*, Parma, Italy, 2004, 13-18
17. Tons M., Doerfler R., Meinecke M. M., Obojski M. A., Radar sensors and sensor platform used for pedestrian protection in the EC-funded project SAVE-U, *IEEE Intelligent Vehicles Symposium*, Parma, Italy, 2004, 813-818
18. Broggi A., Bertozzi M., Fascioli A., Sechi M., Shape-based pedestrian detection. *IEEE Intelligent Vehicles Symposium*, Dearborn, USA, 2000, 215-220
19. YW Xu, XB Cao, H Qiao, A low cost pedestrian detection system with a single optical camera, *the 6th World Congress on Control and Automation*, 2006, 8759-8763
20. D Chen, XB Cao, H Qiao, A SVM-based classifier with shape and motion features for pedestrian detection system, *IEEE Intelligent Vehicle Symposium 2006*, 331- 335
21. Chen D, Cao XB, Qiao H, Wang FY, "A Multiclass Classifier to Detect Pedestrians and Acquire Their Moving Style," *IEEE International Conference on Intelligence and Security Informatics*, 2006, 3975: 758 – 759
22. D. Chen, X.B. Cao, Y.W. Xu, H. Qiao, An Evolutionary Support Vector Machines Classifier for Pedestrian Detection, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, 4223-4227
23. Xianbin Cao, Hong Qiao, Fei-Yue Wang, Application of Cooperative Co-evolution in Pedestrian Detection Systems, *Lecture Notes in Computer Science*, 2005, 664-665
24. Jia Hui-Xing, Zhang Yu-Jin, A Survey of Computer Vision Based Pedestrian Detection for Driver Assistance Systems, *Acta Automatica Sinica*, 2007, 33(1): 84-90
25. Sun Hui, Hua Cheng-Ying, Luo Yu-Pin, A multi-stage classifier based algorithm of pedestrian detection in night with a near infrared camera in a moving car, *IEEE International Conference on Image and Graphics*, Hong Kong, China. 2004, 120-123
26. Cutler R., Davis L S., Robust real-time periodic motion detection, analysis and applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 781-796

27. Curio C., Edelbrunner J., Kalinke T., Tzomakas C., Werner von Seelen, Walking pedestrian recognition, *IEEE Transactions on Intelligent Transportation Systems*, 2000, 1(3): 155-163
28. Zhao Liang, Dressed Human Modeling, Detection, and Parts Localization, *CMU-RI-TR-01-19, Robotics Institute, Carnegie Mellon University*, 2001
29. Yasuno M., Yasuda N., Aoki M., Pedestrian detection and tracking in far infrared images, *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, Washington, DC, USA, 2004, 125-125
30. Bertozzi M., Broggi A., Fascioli A., Graf T., Meinecke M. M., Pedestrian detection for driver assistance using multi-resolution infrared vision. *IEEE Transactions on Vehicular Technology*, 2004, 53(6): 1666-1678
31. Leibe B., Seemann E., Schiele B., Pedestrian detection in crowded scenes, *IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, CA, 2005, 1: 878-885
32. V. Vapnik, *The Nature of Statistical Learning Theory*. Springer Verlag, 1995.
33. K. Rohr, Incremental recognition of pedestrians from image sequences, *Computer Vision and Pattern Recognition*, 1993, 8-13
34. Y. Cheng, Mode seeking and clustering, *IEEE Trans. Pattern Anal. Machine Intell.*, 1995, 17: 790-799
35. Lowe D. G., Distinctive image features from scale-invariant key points, *International Journal of Computer Vision*, 2004, 60(2): 91-110
36. P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *IEEE Conference on Computer Vision and Pattern Recognition*, 2001
37. C.J.C. Burges, A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, 1998, 2: 121-167
38. T. Joachims, Making large-scale SVM learning practical, in *Advances in Kernel Methods—Support Vector Learning*, B. Schölkopf, C. J. C. Burges, and A. J. Smola, Eds. Cambridge, MA: MIT Press, 1998, 1:511-518
39. O. Chapelle, V. Vapnik, O. Bousquet, et al, Choosing multiple parameters for support vector machines, *Machine Learning*, 2002, 131-159
40. K. Tsuda, S. Mika, et al, Learning to predict the leave-one-out error of kernel based classifiers, *Int. Conf. Artificial Neural Networks —ICANN 2001* Berlin: Springer-Verlag, 2001, 331-338
41. Joachims T., Estimating the generalization performance of a SVM efficiently, *the 17th International Conference on Machine Learning*, San Francisco, Morgan, 2000, 431-438

致 谢

随着这篇论文的完成，我在中国科技大学的学习生活也即将结束，之后我将奔赴工作岗位开始一段新的生活。在科大求学的这段日子里，学习和生活的点点滴滴都将成为我生命中最美好的回忆，值得我永远珍藏和品味。在这段珍贵的青春岁月里，许多的老师和同学都给了我莫大的关怀和帮助，这将使我终身受益，这里我衷心的感谢他们。

首先，衷心感谢我的导师曹先彬教授，感谢他三年来对我孜孜不倦的教诲。这三年来，曹老师渊博的知识、敏锐的思维、高深的学术造诣和严谨的治学态度使我受益匪浅。尤其重要的是，曹老师在教育我们做学问的同时，也教育了我们做人的道理。曹老师自身积极奋斗、永无止境的探索精神将永远激励我在今后的工作中刻苦努力，力求上进。

感谢本研究组的同学，他们在我平时的研究过程中和论文的完成过程中提供了很多的帮助，营造了良好的研究氛围。感谢郭圆平、尹洪章、许言午、张卫、高鹏、魏闯先、马静、何其真、贺涛、闫瑞、吴培、嘉晓岚等同学们对我学习和研究中的帮助，使我受益良多。

最后我要感谢我的亲人们，父母给予我的支持和教诲使我不断地努力和成长，他们的关爱与期望是我今后更加努力的动力，感谢我的女友，她的信任和支持给了我坦然面对困难和挫折的勇气。

感谢中国科学技术大学来对我的培养！今日我以我是科大的学生而自豪，希望明日科大以我是科大的学生而骄傲。

陈达

2007年5月

攻读硕士学位期间参加的科研和发表的学术论文

参与课题：

1. 973 子课题：大规模文本演化计算模型研究
2. 中科院复杂系统与智能科学重点实验室开放课题：演化计算及在行人检测中的应用

已发表学术论文：

1. Chen. D., Cao. X.B., Xu Y.W., Qiao H., A Multiclass Classifier to Detect Pedestrians and Acquire Their Moving Style, *IEEE International Conference on Intelligence and Security Informatics*, 2006, 758-759 (SCI:BEM80, EI:062910010953, ISTP:BEM80)
2. Chen. D., Cao. X.B., Xu Y.W., Qiao H., A SVM-based classifier with shape and motion features for a pedestrian detection system, *IEEE Intelligent Vehicles Symposium*, 2006, 331- 335(ISTP: BFI75)
3. Chen. D., Cao. X.B., Xu Y.W., Qiao H., An Evolutionary Support Vector Machines Classifier for Pedestrian Detection, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006 , 4223 - 4227