

学校编码: 10384

分类号 \_\_\_\_\_ 密级 \_\_\_\_\_

学号: 22120051302325

UDC \_\_\_\_\_

厦 门 大 学

硕 士 学 位 论 文

复杂背景下的行人检测技术研究

**Pedestrian Detection in Clutter Scene**

苏 松 志

指 导 教 师: 李 绍 滋 教 授

专 业 名 称: 计 算 机 应 用

论 文 提 交 日 期: 2008 年 4 月

论 文 答 辩 时 间: 2008 年 月

学 位 授 予 日 期: 2008 年 月

答 辩 委 员 会 主 席: \_\_\_\_\_

评 阅 人: \_\_\_\_\_

2008 年 月

# 厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：苏松志

2008年6月1日

# 厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1. 保密（ ），在      年解密后适用本授权书。
2. 不保密（  ）

（请在以上相应括号内打“√”）

作者签名：苏松志

日期：2008年6月1日

导师签名：

日期：      年    月    日

## 摘 要

复杂背景和遮挡问题是当前行人检测技术研究中的难点，对这两个问题的解决，有助于行人检测技术进一步实用化、市场化，创造更高的经济价值和社会价值，同时行人检测的相关技术可以被人脸检测、车辆检测等目标检测技术所借鉴，促进计算机视觉和模式识别等相关学科的发展，具有较大的理论意义和学术价值。本文针对这两个问题展开了研究。

针对复杂背景问题，本文全面分析了基于混合高斯模型运动目标检测方法的优缺点，针对其更新速度慢、收敛性差的缺点提出了相应的改进方法，通过实验验证了该改进方法的有效性；针对复杂背景中阴影对运动目标轮廓的影响，采用基于色度畸变的阴影消除方法；针对混和高斯模型在运动目标的运动方向与摄像机成像平面垂直时，只能检测到目标的边缘轮廓，引入了种子区域增长方法来解决该问题。

针对遮挡问题，详细介绍了隐式形状模型的行人检测方法。该方法首先通过感兴趣点检测算子提取表征行人的局部表观码本，然后通过训练获得码本模型的空间共生分布模型，在此模型基础上将行人识别和分割过程融入到一个概率框架中，最后通过最小描述长度消除重叠的候选人体区域之间的模糊性。该方法采取的是“先识别，后分割”的方法，基于局部特征的描述子，为遮挡问题提供了一条解决思路。

为了排除场景中非人体运动目标，本文提出了基于形状信息的行人分类方法总体框架，详细阐述了如何对所获得的运动目标轮廓解卷绕以进行特征提取，以及样本的收集和人体模型的建立过程。描述了基于 PCA 的维数约简方法和基于 EM 的人体建模具体实现过程，实验结果表明，我们提出的行人分类方法是有效的，在 4007 个测试样本中，只有 27 个误判。

**关键词：**行人检测；混合高斯模型；隐式形状模型；人体轮廓模型

厦门大学博硕士学位论文摘要库

## Abstract

Clutter background and partial occlusions are two challenging tasks of pedestrian detection in real world environments. It is important to solve these problems so that pedestrian detection can be put into practical applications. The technologies used in pedestrian detection can also be transferred to other object detection fields, such as face detection, and vehicle detection, and promote the development of computer vision and pattern recognition. So we focus on how to attack the problems of clutter background and partial occlusions in this thesis.

As to the problem of clutter background, we make a comprehensive study of the Gaussian Mixture Mode (GMM), and then improve it. The experimental results show that our improved method is effective. A shadow detection scheme based on chrominance distortion is also introduced. We also use seeded region grow method to solve a problem of GMM---when the direction of moving objects is vertical to the imaging plane of camera, it can only detect the parts of the objects.

With respect to the problem of partial occlusions, we describe a novel method based on Implicit Shape Mode (ISM). Following a common consensus in the field of object categorization, we do not assume that a figure-ground segmentation is available prior to recognition. The combination of recognition and segmentation into one process is made possible by our use of ISM, which integrates both into a common probabilistic framework. In addition to the recognition and segmentation result, it also generates a per-pixel confidence measure specifying the area that supports a hypothesis and how much it can be trusted. With this confidence, we derive an extension of the approach to handle multiple objects in a scene and resolve ambiguities between overlapping hypotheses with a novel Minimum Description Length (MDL) based criterion.

Finally, we proposed a shape-based classification method for pedestrian. It unwraps the contour of a pedestrian into a one dimensional distance, and then uses Expectation Maximum (EM) to model it. The experimental result shows that there are only 27 false-classified samples in 4007.

**Key Words:** pedestrian detection; Gaussian Mixture Model; Implicit Shape Model; Human Contour Model;

厦门大学博硕士学位论文摘要库

## 目 录

<b>第一章 绪论</b> .....	<b>1</b>
1.1 选题背景与研究意义 .....	1
1.2 国内外研究现状 .....	5
1.2.1 技术现状.....	5
1.2.2 典型系统的性能.....	14
1.3 本文的主要工作 .....	15
1.4 本文的组织结构 .....	16
<b>第二章 基于改进的高斯混合模型运动目标检测</b> .....	<b>18</b>
2.1 减背景方法在行人检测技术中的应用 .....	19
2.1.1 减背景方法的原理和过程.....	19
2.1.2 行人检测技术中的减背景方法.....	21
2.2 基于高斯混合模型的运动目标检测 .....	22
2.3 改进的高斯混合模型 .....	24
2.4 种子区域增长 .....	27
2.5 基于色度畸变的阴影消除 .....	28
2.5.1 阴影的光学属性.....	30
2.5.2 基于色度畸变的阴影消除方法.....	32
2.6 性能评估.....	35
2.7 本章小结.....	39
<b>第三章 基于形状信息的行人分类方法</b> .....	<b>40</b>
3.1 分类框架 .....	41
3.2 特征提取 .....	41
3.3 样本收集 .....	43
3.4 人体模型的建立 .....	44
3.4.1 基于 PCA 的维数约简.....	44
3.4.2 EM 算法.....	47
3.4.3 实验结果.....	48
3.5 本章小结.....	50
<b>第四章 基于隐式形状模型的行人检测方法</b> .....	<b>51</b>
4.1 局部特征检测算子和描述子 .....	52
4.1.1 检测算子.....	52
4.1.2 描述子.....	55
4.2 码本特征 .....	56
4.2.1 码本产生.....	56
4.2.2 码本聚类.....	57
4.3 行人识别方法 .....	58
4.4 自顶向下的分割 .....	60
4.5 基于 MDL 的验证 .....	61



4.6 本章小结.....	64
<b>第五章 总结和展望.....</b>	<b>65</b>
5.1 本文工作总结.....	65
5.2 展望.....	65
<b>参考文献.....</b>	<b>68</b>
中文.....	68
英文.....	68
A.....	68
B.....	68
C.....	69
D.....	70
E.....	70
F.....	70
G.....	70
H.....	71
J.....	71
K.....	71
L.....	71
M.....	72
N.....	73
O.....	73
P.....	73
Q.....	73
S.....	73
T.....	74
V.....	74
W.....	75
X.....	75
Y.....	75
Z.....	76
网络资源.....	76
攻读硕士期间发表的论文.....	77
<b>致 谢.....</b>	<b>78</b>

## Contents

<b>1</b>	<b>Introudction.....</b>	<b>1</b>
1.1	Background and Motivation .....	1
1.2	Related Works .....	5
1.2.1	State-Of-Art.....	5
1.2.2	Performance of Some Pedestrian Detection Systems.....	14
1.3	Contributions .....	15
1.4	Outline .....	16
<b>2</b>	<b>Moving Object Detection Based on Improved Gaussian Mixture Model .....</b>	<b>18</b>
2.1	Background Subtraction .....	19
2.1.1	Principles.....	19
2.1.2	Background Subtraction used in Pedestrian Detection.....	21
2.2	Moving Object Detection Based On Gaussian Mixture Model ...	22
2.3	An Improved Method for Gaussian Mixture Model .....	24
2.4	Seeded Region Grow .....	27
2.5	Cast Shadow Removing Based on Chrominance Distoration ....	28
2.5.1	Property of Cast Shadow.....	30
2.5.2	Cast Shadow Removing.....	32
2.6	Performance Evaluation .....	35
2.7	Conclusion .....	39
<b>3</b>	<b>A Shape-Based Method for Pedestrian Classification....</b>	<b>40</b>
3.1	System Overview .....	41
3.2	Feature Abstraction .....	41
3.3	Collect Samples .....	43
3.4	Modeling Human Body .....	44
3.4.1	Principle Component Analysis.....	44
3.4.2	Expertation Maximization.....	47
3.4.3	Experimental Results.....	48
3.5	Conclusion .....	50
<b>4</b>	<b>A Novel Method for Pedestrian Detection Based on Implicit Shape Model .....</b>	<b>51</b>
4.1	Local Detectors and Descriptors .....	52
4.1.1	Detectors.....	52
4.1.2	Descriptors.....	55
4.2	Codebook .....	56
4.2.1	Codebook Generation.....	56

4.2.2 Codebook Clustering.....	57
<b>4.3 Recognition Method .....</b>	<b>58</b>
<b>4.4 Up-Down Segmentation .....</b>	<b>60</b>
<b>4.5 Verification Based On MDL .....</b>	<b>61</b>
<b>4.6 Conclusion .....</b>	<b>64</b>
<b>5 Conclusion and Future Work.....</b>	<b>65</b>
5.1 Conclusion .....	65
5.2 Future Work .....	65
<b>References .....</b>	<b>68</b>
<b>Papers in Chinese .....</b>	<b>68</b>
<b>Papers in English .....</b>	<b>68</b>
A.....	68
B.....	68
C.....	69
D.....	70
E.....	70
F.....	70
G.....	70
H.....	71
J.....	71
K.....	71
L.....	71
M.....	72
N.....	73
O.....	73
P.....	73
Q.....	73
S.....	73
T.....	74
V.....	74
W.....	75
X.....	75
Y.....	75
Z.....	76
<b>Internet Resources .....</b>	<b>76</b>
<b>My Published Papers.....</b>	<b>77</b>
<b>Acknowledgement .....</b>	<b>78</b>

## 第一章 绪论

### 1.1 选题背景与研究意义

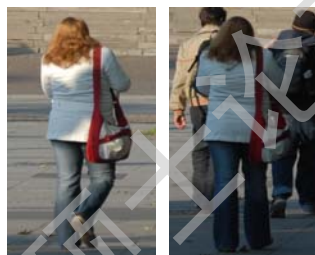
对人类感知信息（视觉、味觉、听觉等）的理解和描述是当前人工智能研究中的热点和难点问题之一。俗话说：“百闻不如一见”，人类从外界获得的信息中，80%是通过视觉得到的。让机器（如计算机）拥有视觉，这是人类多年以来的梦想。随着最近计算机视觉这一研究领域的发展，一些科研成果已经成功地应用到诸如智能视觉监控（Intelligent Video Surveillance[Wren, 1997] [Collins, 2000][Haritaoglu, 2000][Yilmaz, 2006][Shah, 2007]）、人体运动分析(Human Motion Analysis[王亮, 2002][Moeslund, 2006])、高级感知接口（Advanced Perceptual Interface）、虚拟现实（Virtual Reality）、基于模型的编码（Model-based Coding, 如 MPEG-4）等中，获得了极大的经济价值和社会价值。但目前计算机对场景中的目标识别、理解和描述远远不能达到像生物所具有的高效性和灵活性，对图像或者视频理解成为了当前计算机视觉研究领域中的前沿方向。

在对图像或者视频理解的研究中，第一步（也是最重要的一步）就是要鲁棒地、精确地获得感兴趣的研究目标在场景中的位置、大小、形状和姿态等信息，即目标的检测与分割。在早期的研究中一般预先测定所要研究的目标在场景中的信息，所以这方面的研究相对较少。随着社会应用需求的发展，如无人值守的智能监控、智能车辆、基于内容的图像检索及家庭服务机器人等，要求系统能够实时地、自动地完成目标的检测。因此研究一种能够在复杂的环境中自动检测感兴趣的目标（如车辆、行人等）成为了这一研究领域的科研人员的关注焦点。在目标检测技术的研究中，由于行人检测拥有巨大的市场前景和广泛的应用场合，如（1）银行、商店、矿井、机场、火车站和停车场等对安全较为敏感场合的异常情况的自动报警；（2）远距离的人的身份鉴别（如步态识别（Gait Recognition）[Dawson, 2002][Boyd, 2005][王科俊, 2007]）；（3）公共场所行人拥挤状态分析及商店中消费者流量的统计；（4）汽车辅助驾驶；（5）虚拟现实技术中人体动作的建模等，因此一些汽车生产商、大学和研究机构相继开始了行人检测技术的研究。欧盟从 2000 年到 2005 年连续资助 PROTECTOR[Gavrila, 2004] 和 SAVE-U[HTTP,1]项目，开发了两个以计算机视觉为核心的行人检测系统；意大

利 Parma 大学开发的 ARGO[Broggi, 2000]智能车辆也包括一个行人检测模块；以色列的 MobilEye[Shashua, 2004]公司开发了芯片级的行人检测系统；日本本田汽车公司[Xu, 2005]开发了基于红外摄像机的行人检测系统；国外的大学如 CMU、MIT 和国内的西安交通大学、清华大学和上海交通大学等也在该领域做了许多研究工作。



(a) 行人的外观 (appearance) 变化非常大，比如一个男的穿黑色西装和一个女的穿五颜六色的裙子，他们的表现相差甚远。



(b) 行人的外观因光照的变化而变化，上图为行人在阳光下和在树阴下的不同外观。



(c) 行走过程中胳膊和腿的位置使得行人的外观变化不一。



(d) 行人被观察的角度可能是正面、背面或者侧面等。



(e) 行人间的相互遮挡、背包等附属品对外观的影响



(f) 复杂的背景，如纹理较强的区域、垂直的边缘等，这些容易引起误检。

图 1-1 行人检测所面临的挑战 (a ~ f)

Figure.1-1 The challenges of Pedestrian Detection (a ~ f)

尽管存在如此巨大的市场前景和众多的研究机构参与其中，但是由于行人检测问题涉及到模式识别、图像处理、计算机视觉和机器学习等多学科交叉的知识，同时受到穿着、光照、身体姿态、尺度、视角和复杂背景等的影响[Seemann, 2007a] (如图 1-1 所示)，至今未能（也不可能）设计出一种通用的、实时的、鲁棒的检测算法。因此目前大部分行人检测技术的研究，都是基于一定的假设条件下(如表 1-1[Moeslund, 2001]) 进行的，如假设稳定的光照、静态的背景、行人之间或者行人与背景之间无遮挡、摄像机相对静止、特定形态的着装等，这样就避开了复杂背景、遮挡等诸多问题的处理。然而，在现实的场景中，背景是动态变化的，如树枝的摇动、日光灯的突然打开或关闭、喷泉和物体的移入或者移出等。而且场景中的行人并不总是单个出现，经常是并肩或者成群出现，行人之间相互遮挡，给行人检测带来了极大的困难。虽然有研究人员提出通过 3D 来解决遮挡问题，但是这种方案需要多摄像机协同工作，成本高，实时性能较差，不能在一些实时性要求较高的场合（如汽车辅助驾驶、智能监控等）中应用。因此有必要研究一种能在真实场景中工作的实时的、鲁棒的、低成本的行人检测方法。但在当前的技术研究中，不管是采用基于全局特征的描述子（如颜色、纹理、形状等），还是基于局部特征的描述子（如 Edgelet[Wu, 2005][Wu, 2007a]、Shapelet[Sabzmeydani, 2007]、HOG[Dalal, 2005] [Dalal, 2006]等），或者通过视频结合了运动信息[Viola, 2005]，研究者的主要思路是通过构造一个明确的人体模型，然后通过统计学习的方法（如 Boosting[Viola, 2005] [Wu, 2007b]、Bayesian[Zhao, 2003][Dimitrijevic, 2006][Lin, 2007]等）判定图像或者视频中是否存在行人。尽管这些方法在一定的假设条件下能取得较好的检测性能，但是由于背景复杂，再加上人体的外观千变万化，构造一个显式的人体形状模型适用于真实的场景，挑战性极大。



表 1-1 人体运动捕捉中的常用假设条件

Table 1-1 the typical assumptions made by motion capture systems

运动假设 (约束)	与人体外观相关的假设
	环境
1. 目标在工作区内	1. 稳定的光照
2. 摄像机静止或恒定运动	2. 静态背景
3. 在任何时刻工作区内只有一个人	3. 均匀背景
4. 目标始终朝向摄像机	4. 摄像机参数已知
5. 目标的运动方向与摄像机平面平行	5. 专用硬件
6. 无遮挡	目标 (即人体)
7. 缓慢的且连续的运动	1. 初始姿态 (角度) 已知
8. 单一的或少量的肢体运动	2. 目标类别已知
9. 目标运动模式已知	3. 目标上有标记
10. 物体在水平面上运动	4. 特定颜色的着装
	5. 紧身衣

综合上的分析我们知道,复杂背景和遮挡问题的研究是当前行人检测技术研究中的难点,对这两个问题的解决,有助于行人检测技术进一步实用化、市场化,创造更高的经济价值和社会价值,同时行人检测的相关技术可以被人脸检测、车辆检测等目标检测技术所借鉴,促进计算机视觉和模式识别等相关学科的发展,具有较大的理论意义和学术价值。正是在这样的背景下,本论文分别针对这几个问题展开了研究。针对复杂背景问题,我们设计了改进的高斯混合模型(Gaussian Mixture Model, GMM[Stauffer, 1999][KadewTraKuPong, 2001])算法,使之能在真实的复杂场景中工作,为动态的背景建模,准确提取运动目标的区域,同时我们设计了一种基于轮廓特征的行人分类方法,对检测到的目标区域进行分类,以获得感兴趣的区域(Regions Of Interest, ROIs)——行人;针对遮挡问题,由于显式形状模型的局限性(形状特征变化不一,且难于获得包含各式各样形状的训练样本),我们采用了隐式的形状模型(Implicit Shape Mode, ISM[Leibe, 2004][Seemann, 2007b][Leibe, 2007]),该方法能有效解决行人的外观因受关节位置、穿着和背包等的影响而呈现多样性的问题以及行人间相互遮挡的问题。

## 1.2 国内外研究现状

### 1.2.1 技术现状

目前关于行人检测技术的研究中,按摄像机是否移动可以分为:基于固定摄像机的和基于运动摄像机的。基于固定摄像机的行人检测技术主要用在智能视频监控和人体运动分析等领域中,其研究的主要思路是:首先利用减背景、光流法或者帧差法等检测出运动物体的区域;然后利用人体的形状特征、肤色特征或者运动特征等对检测出的运动物体区域进行分类。基于移动摄像机的行人检测技术(如车辆辅助/自动驾驶系统中的行人检测),由于摄像机的运动很难对动态的背景建模,无法利用帧间的变化信息,不能采用“运动目标检测—运动目标分类”这一思路,因此主要的研究方法是将行人检测看作是一类模式识别问题,首先从静态单幅图像中提取行人特征,然后进行模式分类。当然,也可以有其它的分类方法。如果按照图像来源来划分,可分为基于动态图像的(即序列图像或者视频)和基于静态图像的(包括数字化的照片、数码相机拍摄的图片或者是动态视频中的某一帧);如果按照颜色信息来划分,可分为基于彩色的和基于灰度的;如果按照图像背景的复杂程度来划分,可以分为基于简单背景的(指背景的颜色或纹理等属性已知,在该条件下只利用行人的轮廓、颜色或者运动等少量特征便可准确检测出行人)和基于复杂背景的(指背景的类型和特征不受约束,某些区域可能在色彩、纹理或者形状等特征与人体相似,必须将人体的多种特征融合起来才可能准确检测出行人);如果按照摄像头的感光光谱范围来划分,可分为基于可见光的和基于红外线的。

以上的划分并不是绝对了,由于真实场景中的行人受到光照、遮挡、穿着、人体姿态和复杂背景的诸多因素的影响,通常要将多种检测技术综合起来。尽管如此,现有的行人检测系统一般包括两个模块[贾慧星, 2007]:感兴趣区域(ROIs)分割和目标分类(即行人识别)。ROIs 分割的目的从图像中提取行人的候选区域作进一步验证,避免穷尽搜索,以提高系统的速度。目标分类是行人检测系统的核心,它对得到的 ROIs 进行验证,其本质是模式识别中“二分类问题”,即区分人体和非人体,它的性能决定了整个系统的检测率和鲁棒性。下面分别描述这两个模块的研究现状。



Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库