

分类号 TP391.41

密级 _____

UDC _____

编号 _____

中国科学院研究生院 博士学位论文

多视角步态分析与识别

于仕琪

指导教师 谭铁牛 研究员

中国科学院自动化研究所

申请学位级别 工学博士 学科专业名称 模式识别与智能系统

论文提交日期 2007年10月24日 论文答辩日期 2007年11月25日

培养单位 中国科学院自动化研究所

学位授予单位 中国科学院研究生院

答辩委员会主席 袁保宗教授

Multi-view Gait Analysis and Recognition

Dissertation submitted to
Institute of Automation, Chinese Academy of Sciences
in partial fulfillment of the requirements
for the degree of
Doctor of Engineering

by

Yu Shiqi

(Pattern Recognition and Intelligent Systems)

Dissertation Supervisor: Prof. Tan Tieniu

独创性声明

本人声明所递交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其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任何贡献均已在论文中作了明确地说明并表示了谢意。

签名：_____日期：_____

关于论文使用授权的说明

本人完全了解中国科学院自动化研究所有关保留、使用学位论文的规定，即：中国科学院自动化研究所有权保留送交论文的复印件，允许论文被查阅和借阅；可以公布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论文。

（保密的论文在解密后应遵守此规定）

签名：_____导师签名：_____日期：_____

摘 要

由于严峻的公共安全形势，智能视觉监控越来越受到重视。步态识别能够从远距离识别出人的身份，对提高监控系统的智能性至关重要。步态识别的难点之一是步态特征作为一种行为特征，具有很强的不稳定性。特别是在视角变化、衣着变化、携带物品、光照变化、时间变化等情况下，提取稳定的具有区分能力的特征变得非常困难。本文针对步态识别中的视角变化这一常见变化，进行了一系列研究，并对衣着和携带物品变化等因素也有所涉及。此外，还对基于步态的性别识别这一新方向进行了探索。本文的主要工作有：

1. 目前步态识别研究尚不成熟，步态识别缺乏统一评价标准，而且有很多问题尚没有研究清楚。针对这一现状，创建了一个步态识别算法评价框架。这个评价框架包括一个大规模的步态数据库（CASIA 步态数据库数据集 B），三组实验和一系列评价指标。该框架可以评价某个步态识别算法对视角变化，衣着变化和携带物品变化的稳定性，指导步态算法的设计。
2. 哪个视角最适合用于步态识别？视角变化是如何影响识别率的？这两个问题是步态识别研究中的两个重要问题。本文对视角变化与步态识别性能的关系进行了建模和分析，分析得出侧面视角是步态识别中的最佳视角的结论；另外，还建立起视角与识别率之间的关系模型。这些对步态识别研究具有一定的指导意义。
3. 目前虽然有大量的步态识别算法被提出，但是这些算法大部分都是基于特定视角的，对视角变化不鲁棒，这极大的限制了步态识别的应用。文中分别提出了一种线性和一种非线性的模型，将一个视角的步态数据合成为该人在另一个视角的数据，来解决测试数据的视角跟注册数据的视角不一致问题。大量的实验验证了该算法对于解决视角变化问题非常有效。
4. 已有的研究证明可以根据步态区分性别。性别可以用于步态特征的粗分类，用于提高步态数据库的检索速度和提高步态识别的准确率。另外，性别识别还可以提高监控系统的感知能力，对进入场景中的行人进行信息搜集。在文中使用了多种特征和多种分类器从多个视角进行性别识别研究，

并比较了各种特征以及分类器性能, 在实验中发现了一些有用的结论, 如身体的哪些部分最能表现人的性别, 哪些特征最有区分能力等。针对步态特征常见的视角变化问题, 提出了解决方法, 获得了令人鼓舞的结果。除此之外, 还进行了跨人种的性别识别, 获得了比较高的识别率, 证明了不同人种的相同性别的人的步态是相似的。

关键词: 生物特征识别, 步态识别, 步态分析, 多视角, 视觉监控, 性别识别

Abstract

With the increasing requirements for security, intelligent visual surveillance gained more and more attentions. Gait can be used as a kind of biometric to identify human at a distance. Gait recognition is a key element for improving the intelligence of the surveillance systems. Compared with other biometric features, the shortcoming of gait is that it is not robust to variations, such as view, clothing, carrying condition, illumination. View variation is very common in visual surveillance. In the dissertation we investigate view variation in gait analysis and recognition. Some other variations, clothing and carrying condition, are also considered. Besides, some research work is also done on gait based gender classification. This dissertation mainly includes the following issues:

1. Gait recognition is in its immaturity, and there is still no standard to evaluate different algorithms. We propose a evaluation framework to do this work and advance gait recognition technology. The framework contains a large gait database (Data Set B in CASIA Gait Database), 3 sets of experiments and some metrics. The framework can evaluate an algorithm's robustness to view, clothing and carrying condition variations.
2. There are two remaining open problems in gait recognition. One is which view is the most suitable one for gait recognition and why it is. Another is how view angle variation affects the performance of gait recognition. We proposes two models, a geometrical one and a mathematical one, in an attempt to address these two questions, and investigate and analyze the effect of view angle on the performance of appearance-based gait recognition.
3. Currently most gait recognition algorithms are view dependent and not robust to view variation. We proposed a linear model and a non-linear model to synthesize the gait feature from one view to another view. The models can be used when the probe angle is not equal to the gallery an-

gle. Experimental results show that the models can solve view variation problem.

4. Some previous research shows that gait feature can be used for gender classification. Gait feature can be divided into two groups by gender, male and female, so gender can help to speed up gait database retrieval, improve gait recognition accuracy and surveillance systems' perception. We give a comprehensive study on gait based gender classification, which includes experiments by human observers and computer algorithms, comparison of different gait features and classifiers, experiments from multi-view and view invariant gender classification. We also do cross-race gender classification experiments and gained inspiring results.

Keywords: Biometrics, gait recognition, gait analysis, multi-view gait feature, visual surveillance, gender classification

目 录

摘要	iii
Abstract	v
目录	vii
第一章 引言	1
1.1 步态识别的应用背景	1
1.2 相关研究领域	2
1.2.1 人体检测与跟踪	2
1.2.2 人体行为识别	3
1.2.3 医学步态分析	3
1.3 步态识别历史与现状	3
1.4 步态识别研究的难点	4
1.5 研究内容和论文结构	5
第二章 步态识别研究现状	7
2.1 步态识别系统框架	7
2.1.1 轮廓分割	8
2.1.2 步态周期计算	9
2.1.3 特征提取和表达	11
2.1.4 分类器	15
2.2 多视角步态识别	18
2.2.1 虚拟体模型	18
2.2.2 身体参数	19
2.2.3 特征合成	20
2.2.4 研究现状小结	21

第三章 多视角步态数据库	23
3.1 公开的步态数据库	23
3.2 多视角步态数据库的设计	29
3.3 多视角步态数据库的采集	30
3.3.1 采集环境设置	30
3.3.2 数据采集	32
3.4 数据管理与共享	33
第四章 步态识别算法性能分析与评价	37
4.1 步态识别算法评价	37
4.1.1 实验设计	38
4.1.2 算法评价标准	41
4.2 视角与识别率关系模型	43
4.2.1 识别率函数	44
4.2.2 实验结果	45
4.2.3 几何模型	48
4.2.4 代数模型	49
4.3 本章小结	51
第五章 视角无关的步态识别	53
5.1 视角无关的步态识别框架	53
5.2 特征合成方法	55
5.3 线性模型	56
5.3.1 原理	56
5.3.2 步态特征	57
5.3.3 主成分分析法降维	58
5.3.4 实验与分析	59
5.4 非线性模型	63
5.4.1 神经网络	64

5.4.2	实验和分析	65
5.5	结论与展望	68
第六章	基于步态的性别识别	69
6.1	相关研究	69
6.2	由人识别性别	70
6.3	步态特征与分类器	72
6.4	数据库	72
6.5	实验和分析	73
6.5.1	各种不同特征和分类器组合的实验	74
6.5.2	最佳分类特征	76
6.5.3	轮廓与轮廓线	77
6.5.4	空间域特征分析	79
6.5.5	PCA 域特征分析	83
6.6	跨人种的性别识别	83
6.7	多视角下的性别识别	85
6.8	视角无关的性别识别	85
6.9	本章小结	87
第七章	结论与展望	91
7.1	工作总结	91
7.2	未来发展趋势	92
7.2.1	行人检测与分割	93
7.2.2	关键影响因素的评价	93
7.2.3	融合	94
7.2.4	基于步态的视频检索	94
7.2.5	基于步态的个人信息获取	94
	参考文献	97

发表文章目录	107
致谢	109

表 格

3.1	目前已经公开的步态数据库	28
4.1	实验集 A 实验结果 (GEI 算法)	38
4.2	实验集 B 实验结果 (GEI 算法)	39
4.3	实验集 C 实验结果 (GEI 算法)	39
4.4	实验集 A 实验结果 (KFD 算法)	40
4.5	实验集 B 实验结果 (KFD 算法)	40
4.6	实验集 C 实验结果 (KFD 算法)	41
4.7	视角无变化时的评价标准列表	42
4.8	GEI 和 KFD 算法在视角无变化时的比较	42
4.9	GEI 和 KFD 算法在视角变化时的比较	43
5.1	几种视角无关的步态识别方案	53
5.2	不使用合成特征的实验结果 (实验集 A)	62
5.3	线性投影合成特征的实验结果 (实验集 A)	62
5.4	三种不同算法对视角鲁棒性的评价参数	63
5.5	二层神经网络合成特征的实验结果 (实验集 A)	67
5.6	线性模型与二层神经网络模型在视角变化时性能评价标准	67
6.1	由人观察轮廓识别性别的识别率	71
6.2	使用 CASIA 步态数据的性别识别实验结果	74
6.3	使用 Soton 大步态数据的性别识别实验结果	75
6.4	使用混合步态数据的性别识别实验结果	75
6.5	使用最佳分类特征时的实验结果	77
6.6	噪声对轮廓和轮廓线特征的影响	79
6.7	使用不同的 PCA 系数获得的识别率 (混合库)	83

6.8 跨人种性别识别结果(训练数据为亚洲人,测试数据为欧洲人)	84
6.9 跨人种性别识别结果(训练数据为欧洲人,测试数据为亚洲人)	85
6.10 视角变化时的性别识别的识别率(SVM分类器)	87
6.11 使用线性模型合成的特征进行性别识别的识别率(SVM分类器)	88
6.12 视角变化时性别识别性能评价标准	89

插 图

2.1	基于视觉的步态识别系统框架	7
2.2	基于背景减除的人体分割示意图	9
2.3	一个步态周期 [1]	10
2.4	一个序列中所有轮廓的宽度	10
2.5	轮廓宽度向量的自相关系数	11
2.6	轮廓投影特征 (水平投影和垂直投影)	12
2.7	步态能量图计算示意图	13
2.8	轮廓线 (contour) 表达方式	13
2.9	人体主要部分的跟踪结果 (左侧 3 图) 以及 4 个人的大腿角度变化曲线 (右图) [2]	15
2.10	k 近邻分类器 (k=5)	16
2.11	支持向量机	17
2.12	论文 [3] 中的虚拟体模型	18
2.13	论文 [4] 中使用的特征	19
2.14	论文 [5] 中的投影模型	20
3.1	Georgia Tech 数据库例图	24
3.2	CMU Mobo 数据库例图	24
3.3	HID—UMD 数据库 (数据集 1) 例图	25
3.4	HID—UMD 数据库 (数据集 2) 例图	25
3.5	CASIA 数据库 (数据集 A) 例图	26
3.6	Soton 数据库例图	26
3.7	USF Gait Challenge 数据库例图	27
3.8	CASIA 数据库 (数据集 C) 例图	27
3.9	数据库采集方案示意图	31

3.10	步态数据采集软件界面	31
3.11	采集软件的 Capture 部分	32
3.12	数据库采集场景	33
3.13	CASIA 数据库 (数据集 B) 例图	34
4.1	多视角步态识别示意图 (注册集角度为 θ_g , 测试集角度为 θ_p)	44
4.2	GEI 算法的正确识别率 (横坐标是 θ_p , 纵坐标是 CCR)	46
4.3	KFD 算法的正确识别率 (横坐标是 θ_p , 纵坐标是 CCR)	47
4.4	几何模型示意图	48
4.5	GEI 法的正确识别率曲面	49
4.6	KFD 法的正确识别率曲面	50
4.7	通过模型计算出的 GEI 法的正确识别率曲面	51
4.8	通过模型计算出的 KFD 法的正确识别率曲面	52
5.1	监控系统中常见的行走情况	54
5.2	视角无关的步态识别框架	55
5.3	多个视角的特征在不同特征空间表示	56
5.4	将 54° 视角数据映射到 90° 的矩阵	59
5.5	两人的第 1 个和第 2 个 PCA 系数 (* 为 54° 特征; \circ 为 90° 特征; + 为利用 54° 特征合成的 90° 特征)	60
5.6	注册集角度为 90° , 测试集角度分别为 $0^\circ, 18^\circ, \dots, 180^\circ$ 时的识别率	61
5.7	多层前馈神经网络模型	64
5.8	sigmoid 激活函数	65
5.9	注册集角度为 90° , 测试集角度分别为 $0^\circ, 18^\circ, \dots, 180^\circ$ 时的识别率	66
6.1	播放人体轮廓的软件	71
6.2	CASIA 步态数据库数据集 B 中不同性别的视频	73
6.3	Soton 大步态数据库中不同性别的视频	74
6.4	受噪声影响的轮廓	78
6.5	真实轮廓	78

6.6	人体轮廓各个部位的 F 统计量	80
6.7	混合数据库中不同性别的特征的均值	81
6.8	步态能量图的前 30 个 PCA 系数的 F 统计量 (混合库)	81
6.9	步态能量图的前 4 个 PCA 系数 (混合库, 蓝 * 代表男, 红 o 代表女)	82
6.10	步态能量图的前 10 个 PCA 特征向量 (混合库, 黑色代表负值, 白色代表正值, 灰色代表零值)	84
6.11	在不同视角下性别识别结果 (o 为 kNN 结果, * 为 SVM 结果)	86

第一章 引言

1.1 步态识别的应用背景

在智能监控中，远距离的确定监控场景中的人的身份是一个充满挑战而且又很有应用前景的方向。随着公共场所的安全形势越来越严峻，大量的监控摄像头已经被应用于银行、商店、办公楼、停车场、地铁站等场所。但是目前的监控系统一般只能用来记录发生的事情，很难起到预警和报警作用。一个真正的智能监控系统能够自动分析摄像机采集的图像数据，对数据进行理解。当盗窃发生或发现具有异常行为的人时，系统能准确及时地发出警报，从而避免犯罪的发生，或者在第一时间赶到现场控制犯罪行为。监控系统不仅需要判断人是否存在、人的位置和行行为，而且还需要识别这个人是谁。因此非接触式远距离的身份识别研究已经逐渐成为一个热门的研究方向。

在监控场景中对行人进行身份识时，距离往往比较远，所以难以获取特征。因为这个原因，常用的生物特征如虹膜、指纹、手形、声音、签名等无法用于监控场景进行人的身份识别。人脸虽然在一定程度上可以从远距离获取，但是这个距离不能太远。步态特征是目前发现的一种能满足要求的生物特征。

步态 (Gait: A particular way or manner of moving on foot [6].) 的意思是足部带动身体移动的方式。我们使用“步态识别” (gait recognition) 来指根据人的行走模式来识别个体身份 [7]。本文中的“步态”指广义的步态，既包含动态的行走信息，又包含静态信息，如身高、腿长、肩宽等。广义的“步态识别”也可以被称作“远距离身份识别”，即无论采取何种特征，只要能够从远距离识别出人的身份，则达到目的。与其他特征相比，步态有着如下优点：

1. 可远距离获取。人脸、虹膜、手形和指纹等生物特征通常需要近距离或接触式的采集。如人脸和虹膜通常需要距离比较近以保证采集的图像足够清晰，而指纹和手形往往需要与传感器进行接触来完成采集。步态可以从几十米甚至上百米远的距离远处采集到。这一优点是其他生物特征难以比拟的。
2. 非接触性和非侵犯性。步态可以非接触的非常隐蔽的采集到。大部分生物

特征往往需要用户的配合来完成采集，而步态不需要。这一点在智能监控中十分重要。

3. 难于隐藏。在公共场合中的每个人都要走路(当然腿部残疾者例外)，如果一个人要刻意改变自己的步态，那他的行为往往变得可疑而更容易引起注意。

在文献 [8] 中，描述了这样的一个系统：一座大楼入口处装有摄像头，当有人走到门口时，步态识别系统能够识别出人的身份，并确定他是否具有进出大楼的权限，如果有权限，则自动将门打开。这样的步态识别系统是所有步态识别研究者的目标。但是在开放的环境下，仅仅根据步态对人进行身份识别，具有一定的挑战性。步态识别的另外一个应用是监控系统中的视频检索，当在监控系统中发现一个形迹可疑的人后，可以根据该人的步态特征，在监控系统储存的视频中检索该人还出现在哪些场景中，可以快速确定该人的历史行动路线。除了身份识别，步态特征还可以用来判断人的性别和年龄，进行人的粗分类，可以从监控场景中搜集信息，用以提高监控系统的感知能力。

1.2 相关研究领域

步态分析在医学和心理学领域已经有着长期的研究，但基于视觉的自动步态分析与识别在最近十年才引起了研究者的注意。与步态分析与识别相关的领域有视频中人的检测、跟踪，人的行为识别，医学上的步态分析等。

1.2.1 人体检测与跟踪

利用附加在人体上的标记物来提取特征可以用在实验室中的步态研究，但是如果将步态研究应用到智能视觉监控中，这样的办法是无法实现的，因为无法为所有人的身上都贴满标记，所以需要利用计算机视觉的方法来自动的检测视频中的人，跟踪人体的各部分，以获取人的运动数据。

目前有很多种行人跟踪和检测方法 [9, 10, 11, 12]。这些方法中，有些是基于模型的方法，有些采用非模型的方法。如 Ning 等人 [12] 用梯形模型来跟踪人体的各个部位，而 Dai 等人 [11] 则是利用人体的形状和外观特征来进行行人的检测和跟踪。人体的检测和跟踪是步态分析与识别系统的前端，检测和跟踪质量的提高，对于提高步态识别算法的性能起到决定性作用。

1.2.2 人体行为识别

人体行为识别的目的是识别出人的各种行为，如弯腰、挥手、蹲、跳、行走等动作。比较早的方法是采用附在身体上的发光点（moving light displays, MLD）[13]来提取运动特征。这种方法的实验环境设置过于复杂，目前大部分方法改用从视频中用计算机视觉的方法提取行为特征。如 Bobick 和 Davis 使用运动能量图像（motion energy images, MEI）和运动历史图像（motion history images, MHI）来描述图像序列中人的运动 [14]。Wang 等人采用人体轮廓的 R 变换来提取运动特征 [15]。

行为识别跟步态识别相同之处是它们都提取人的行为特征进行分类。它们的不同之处是：行为识别是提取不同人运动特征的共性特征，以达到对不同的人都能识别出其行为的目的；而步态识别是提取不同人的个体特性，以达到进行身份识别的目的。

1.2.3 医学步态分析

医学步态分析 [16, 17, 18] 的主要目的是分析并比较正常的步态与非正常的步态，检测出非正常的步态，推测某种类型的非正常步态是由何种原因引起等。医学步态分析还可以为非正常的步态提供矫正方案。

在早期的医学步态分析中，步态特征的获取一般采用跟踪附加在身体主要关节上的标记物来实现 [16, 17]。通过跟踪，可以获得标记物的运动轨迹，如头部的运动轨迹，胳膊和手的摆动轨迹，以及膝盖和脚踝的移动轨迹等。然后用这些标志物的轨迹来分析步态。在 [17] 中，Murray 根据实验，发现人能够只根据标记物的运动轨迹来识别出熟悉的人，但是并没有在生物识别领域对这个问题进行深入研究。

1.3 步态识别历史与现状

针对步态的研究最早出现在三四十年前，当时的研究主要集中在医学步态分析方面 [17]。在上世纪九十年代，很多研究者开始把步态应用于人的身份识别，开辟了一个全新的方向。在步态识别研究中，2000年美国国防部高级研究局（Defense Advanced Research Projection Agency, DARPA）赞助的 Human ID 计划（Human Identification at a Distance）起了很大的促进作用。在这一计划的资助下，南佛罗里达大学、南安普敦大学、马里兰大学、卡内基梅隆大学、

麻省理工学院等很多研究单位开展了大量有意义的工作。中科院自动化所、西安交通大学、中国科学技术大学等国内研究单位和高校也对步态识别进行了相关研究，取得了丰硕的研究成果。目前在小规模步态数据库上获得了 100% 的识别率，在大规模数据库（约100人）上获得了大约 80% 的识别率。这些研究结果令人鼓舞，让人看到步态识别的可行性，激励了更多的研究者对步态识别的兴趣，并投入到步态识别的研究中。

虽然步态识别已经取得了很大的进展，但是由于步态特征容易受到其他因素影响，鲁棒性差，目前仍有大量问题需要进行解决，如识别性能与视角和衣着等因素的关系，步态特征中的哪些部分最有区分能力，如何提取鲁棒的步态特征等。

1.4 步态识别研究的难点

生物特征可以大致分为两类：生理特征和行为特征 [7]。生理特征源自人体部分的直接度量，如人脸、虹膜、指纹、掌纹、手形、视网膜等。行为特征是当人体执行某种行为的时候的特征，如步态，签名，声纹，击键等。与生理特征相比，行为特征在特征鲁棒性方面往往有所欠缺。步态是一种典型的行为特征，因此也具有行为特征的鲁棒性差这一些特点。步态特征的鲁棒性容易受到下面几种因素的影响：

1. 视角的变化。在步态识别中，视角是一个很难控制的因素。步态识别的一个很大优点是隐蔽性，这也就是被识别对象往往不知道被识别，这就无法要求被识别对象进行相应的配合，他们一般不会按照步态识别系统设计者的意图去行走。视角变化后，摄像机获取的人的外形将会发生很大变化，目前大部分算法无法处理这种变化。
2. 衣服的变化。衣服容易引起体形的变化，比如夏天的穿短裤和 T 恤衫的体形，与冬天穿的长外套的体形相比有很大的不同。在裙子和长外套这类衣服的掩盖下，人的步态特征往往会被严重遮挡，难以提取有区分能力的特征。同时，衣服的变化也会引起行走方式的变化。
3. 携带物品的变化。当行人携带物品（包，箱，书等）时，目前的视觉技术很难将携带的物品跟人体分开，而且携带物品将也会影响人的行走方式。

4. 身体的变化。人在受伤、怀孕以及体重变化的情况下，步态将会有比较大的变化。
5. 情绪的变化。人在情绪高涨的情况下，步态轻快；在情绪低落的情况下，步态沉重；在醉酒的情况下，步态往往变得不规律。
6. 光照和复杂背景的影响。在步态识别中，首先需要将人跟背景分割开。在背景比较复杂的情况下，前景往往难以从背景中被分割出来。另外，单方向的光照容易引起影子，也给人体分割造成很大的难度。

另外，由于步态识别所使用的数据一般都是视频数据，所以具有存储量大、处理复杂耗时等缺点。

1.5 研究内容和论文结构

本论文以智能视觉监控为背景，针对步态识别中的视角变化问题，进行了一系列分析和研究。首先建立了一个多视角数据库，提出了一个用于评价算法各种变化因素鲁棒性的框架，并设计了提取视角无关的步态特征的算法。另外还利用步态进行粗分类，即区分男性步态还女性步态，取得了令人鼓舞的结果。具体各章的安排如下。

第二章介绍了步态识别系统的基本框架，并详细介绍了其中的基本模块和一些常用算法，以及目前的研究现状。除此之外，在该章中还特别的介绍了多视角步态识别的现状。

第三章首先介绍了国际上已经公开的步态数据库，并分析了它们的创建目的和优缺点。随后分析视角无关的步态识别算法的需求，并为了满足这种需要，设计了 CASIA 步态数据库数据集 B (多视角数据集)。CASIA 步态数据库数据集 B 目前是国际上规模最大的公开数据库，它的设计、采集过程、以及后期管理都在这一章中做了详细介绍。

第四章针对目前的步态识别研究缺少通用的评价标准这一现状，提出了一个步态识别评价框架来尝试解决这个问题。这个评测框架可以评价一个算法对视角、衣着和携带物品的鲁棒性。同时，并对视角变化与步态识别性能的关系进行了建模和分析，得出了一些有指导意义的结论。

第五章提出了解决视角变化问题的方法。使用了线性模型和非线性的模型，将一个视角的步态数据合成为该人在另一个视角的数据，这也可以被视为一种

提取视角无关特征的方法。大量的实验验证了算法对于解决视角变化是有效的。

第六章介绍步态分析在性别识别中的应用。本章对基于步态的性别识别进行了较全面的研究。首先，由人观察运动的人体轮廓来判断性别，这个实验验证了一个结论，即可以由步态判断出性别。在随后的研究中，使用了多种特征和多种分类器进行性别识别研究，并比较各种特征以及分类器性能，在实验中发现了一些有用的现象和结论。除此之外，还进行了跨人种的性别识别。性别识别同样也有视角变化问题，对于视角变化，以及视角无关的性别识别，在这一章中也进行了分析和研究。

第七章对全文进行了总结，并展望了步态分析与识别未来的方向。

第二章 步态识别研究现状

因为步态独特的优点，步态识别受到越来越多的重视，很多研究者在尝试开发稳定性好并且识别率高的步态识别系统。一个典型的步态识别系统包括数据采集、从背景中分割出人体、提取步态特征和使用分类器进行识别等步骤，在本章随后的内容中将对这些步骤进行具体描述。

步态特征提取的难点是如何应对各种变化因素（如视角、衣着、携带物品等）。视角变化是影响步态特征提取的一个重要变化因素，一些研究者进行了探索性的研究，这些相关研究内容将在本章第二部分中介绍。

2.1 步态识别系统框架

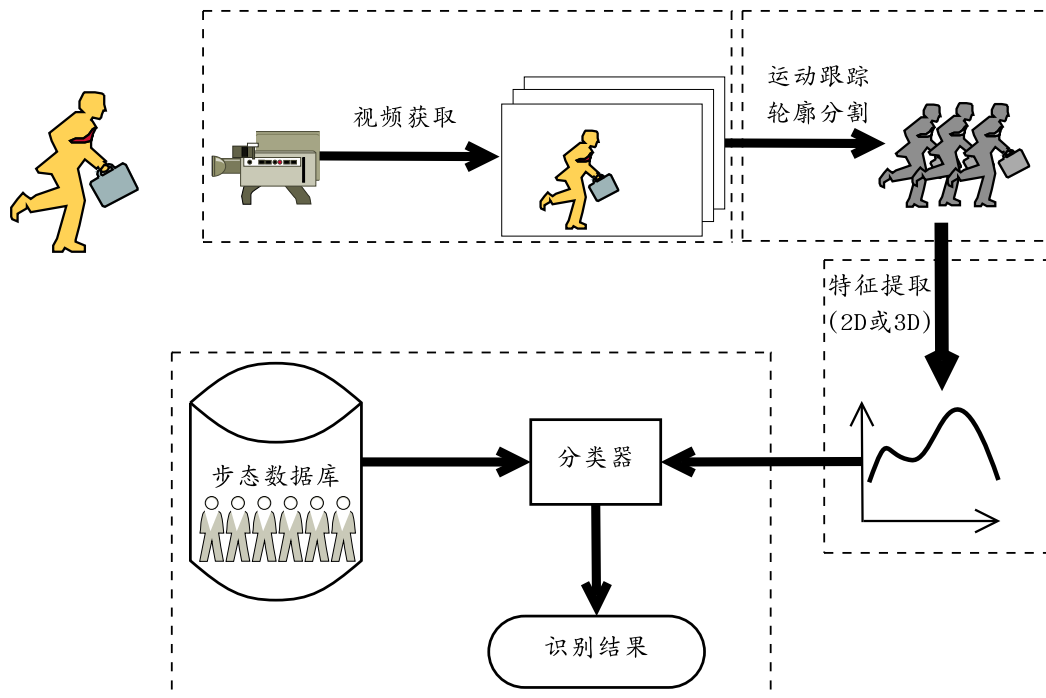


图 2.1: 基于视觉的步态识别系统框架

一个典型的步态识别系统首先使用摄像头采集视频数据，通过运动检测，检测到里面的行人，将人体轮廓分割出来；然后把人体轮廓送入特征提取模块，

通过建立人体结构模型或者非结构的方法提取步态特征；提取到的特征通过识别模块跟数据库中的数据进行比对，判断出人的身份。具体流程如图 2.1。下面详细介绍各个模块中的常用算法。

2.1.1 轮廓分割

无论对于基于模型的步态识别方法，还是对于基于人体整体的步态识别方法，大都需要将人体跟背景图像区分开，分割 (segment) 出人体轮廓是必不可少的一步。这一步对于后面的特征提取与分类非常重要，轮廓的质量决定着最终的识别效果。然而，由于复杂的背景、变化的光照、阴影等原因，分割出好的轮廓是非常有挑战性的工作。下面介绍两种常用的方法。

2.1.1.1 背景减除法

在简单背景情况下，背景减除法可以获得很好的分割效果。这种方法简单高效，目前被广泛采用。其基本思路是使用当前图像跟背景图像相减来获得当前图像上的运动区域。背景图像一般采用当前时刻以前的图像进行加权平均运算获得，计算公式如下式所示：

$$B_t = \alpha I_{t-1} + (1 - \alpha) B_{t-1} \quad (2.1)$$

其中 B_t 是当前的背景图像， I_{t-1} 是上一时刻的图像， B_{t-1} 是上一时刻的背景图像， α 是更新系数，决定着背景图像的更新速度。通过背景的自动更新，可以处理光照变化和背景变化等因素。

计算出背景图像后，将当前图像跟背景图像相减 (公式 2.2)，可以得到运动区域，为了得到二值化的前景图像 F_t^b ，需要对图像进行二值化操作 (公式 2.3)。

$$F_t = |I_t - B_t| \quad (2.2)$$

$$F_t^b(x, y) = \begin{cases} 1, & \text{if } F_t(x, y) > T; \\ 0, & \text{else.} \end{cases} \quad (2.3)$$

公式 2.3 中的 T 是阈值，可以根据具体情况选择。一般来说，图像 F_t^b 中往往会存在一些噪点和一些小的空洞，可以通过形态学操作 (腐蚀和膨胀) 去除。当前图像，背景图像以及从当前图像计算出的前景图像如图 2.2 所示。

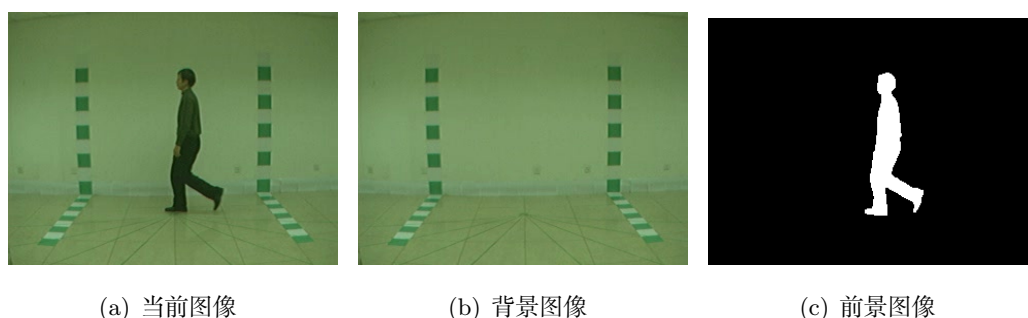


图 2.2: 基于背景减除的人体分割示意图

2.1.1.2 高斯混合模型

在比较复杂的情况下，例如存在光线的变化，树叶的摇动，这个时候简单的背景减除法很难将前景提取出来。高斯混合模型（Gaussian Mixture Model, GMM）是目前流行的一种方法，可以在一定程度上解决复杂背景问题。高斯混合模型方法并不是为整副图像建立一个统一的模型，而是为每个像素建立一个多高斯的混合。高斯混合模型的数学表达式如下：

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \cdot \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2.4)$$

其中 K 是高斯模型数目， $\omega_{i,t}$ 是第 i 个模型在时刻 t 的权重， $\mu_{i,t}$ 是第 i 个模型在时刻 t 的均值， $\Sigma_{i,t}$ 是第 i 个模型在时刻 t 的协方差矩阵，函数 η 是一个高斯概率密度函数，函数 $P(X_t)$ 描述的是该点是否是背景的概率。

高斯混合模型是一个自学习的模型，它能够自动的调整模型里的参数（均值和方差）来适应背景的变化。具有较好的效果，目前被广泛用于固定场景的背景建模。具体的高斯混合模型的自学习过程在论文 [19] 中有详细描述。

2.1.2 步态周期计算

通过前面提到的算法得到前景图像后，由于人与摄像机的距离可能是变化的，离摄像机越近，轮廓图像越大，离摄像机越远，轮廓图像越小。为了消除这种影响，需要首先对轮廓的高度进行归一化，获得具有相同高度的轮廓图像。

同一只脚两次相邻触地时刻之间行走动作作为一个步态周期（图 2.3）。目前很多步态识别算法需要正确计算出步态的周期，以步态周期为单位进行特征提

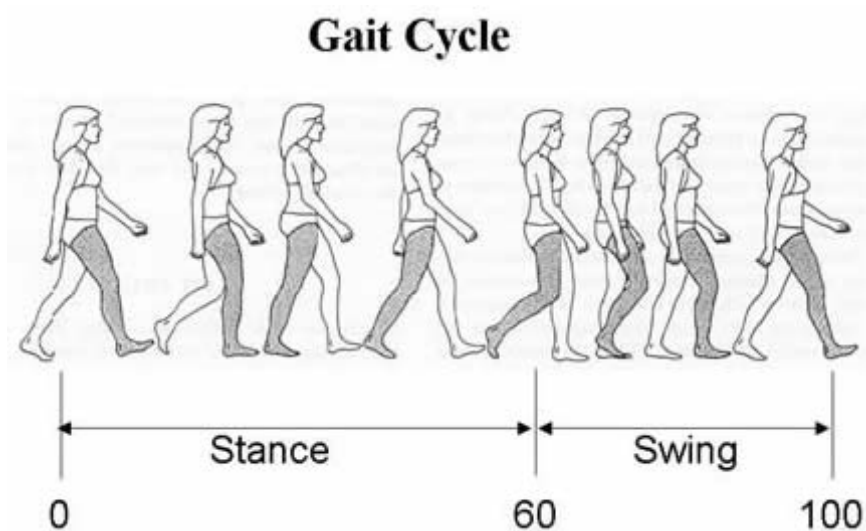


图 2.3: 一个步态周期 [1]

取，所以正确的计算出步态的周期非常重要。当人行走的时候，轮廓的宽度呈现周期性变化，如图 2.4 所示。假如轮廓的宽度向量为 $\mathbf{v} = [v_1, v_2, \dots, v_n]$ ，那么 \mathbf{v} 与 \mathbf{v} 的周期平移会非常相似，因此可以从轮廓的宽度变化计算出步态的周期。但是由于轮廓噪声等因素的影响，宽度变化曲线往往并不平滑，难以直接从宽度变化曲线上自动的获得步态周期。

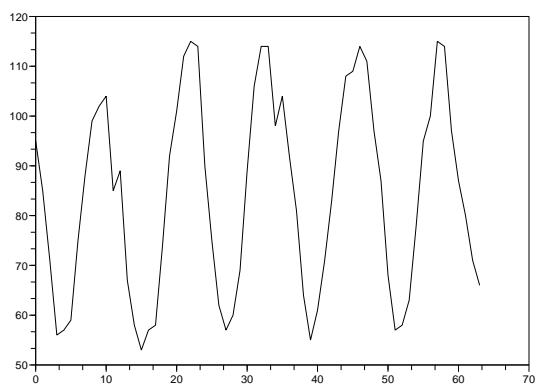


图 2.4: 一个序列中所有轮廓的宽度

相关系数可以用来表示两个向量之间的相关程度（相似程度），轮廓的

宽度向量 \mathbf{v} 与宽度向量 \mathbf{v} 的周期平移向量非常相似，所以宽度向量 \mathbf{v} 的自相关系数会在周期的倍数处达到局部最大值。计算宽度向量 \mathbf{v} 的自相关系数 $\mathbf{c} = [c_1, c_2, \dots, c_m]$ 的公式如式 2.5 所示：

$$c_m = \frac{1}{n} \sum_{k=1}^{n-m} (v_k - \mu)(v_{m+k} - \mu) \quad (2.5)$$

其中 μ 是向量 \mathbf{v} 的均值。

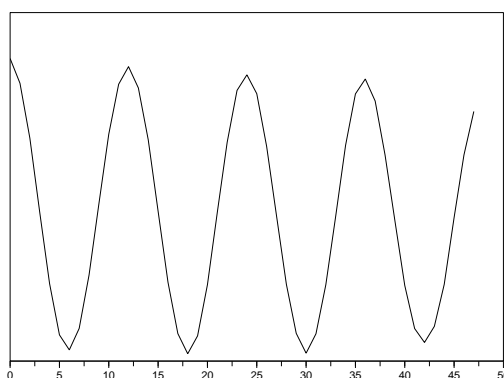


图 2.5: 轮廓宽度向量的自相关系数

图 2.4 中的宽度向量的自相关系数绘出来如图 2.5 所示。从图 2.5 中可以发现，自相关曲线比宽度向量曲线更加平滑。由于自相关系数曲线比较平滑，所以很容易的计算出极值点的位置，然后计算这些极值点间隔的平均值的两倍，则可以得到步态周期。例如图 2.5 中曲线分别在 12, 24, 36 处达到极大值。因为步态信号与其半周期平移信号也具有很高相似性，所以极值间的间隔（12 帧）为半周期，步态的周期为 24 帧。将本方法应用到 CASIA 多视角步态数据库，除了个别轮廓分割完全失败的序列，本方法都可以很准确的检测出序列中步态的周期。

2.1.3 特征提取和表达

目前大部分用于步态识别的特征可以分为基于模型的特征 [4, 20, 21, 22, 2, 23, 24]，和基于整体轮廓（非模型）的特征 [25, 26, 27, 28, 29, 30, 31, 32] 两大类。基于模型的特征一般采用人体关节跟踪方式提取人行走时的关节角度变化，



图 2.6: 轮廓投影特征 (水平投影和垂直投影)

肢体运动轨迹, 以及身高、腿长等物理信息。为了获取高精度的特征, 基于模型的方法一般要求高清晰的输入图像, 而且具有较高的计算复杂度, 但算法的稳定性却不一定很好。基于整体轮廓的方法具有比较低的计算复杂度, 对图像质量要求不高, 目前更多的步态识别算法是采用基于整体轮廓的方法。

2.1.3.1 原始轮廓

在 [25] 中 Sarkar 等人提出了一个基准算法, 这个算法使用原始的人体轮廓 (silhouette) 作为特征进行识别。每个轮廓用一个图像来表示, 一个序列中的所有轮廓可以用一个 3 维矩阵表示。当两个序列进行相似度比较时, 一个序列在时间轴进行平移, 寻找与另一个序列的最佳匹配。得到最佳匹配位置后, 然后计算两个序列的相似度。直接用原始轮廓进行识别虽然简单, 但是缺点是显而易见的: 计算复杂度很高, 且无法处理速度变化情况。

2.1.3.2 轮廓投影

轮廓投影是人体轮廓图像沿着某个方向投影得到, 投影值为投影方向上非零元素的个数。常用的投影方向为水平投影, 垂直投影和对角线投影 (图 2.6) [33, 34]。在论文 [33] 中, Tan 等人通过实验分析, 发现水平投影特征的区分能力优于垂直投影的区分能力。当然, 也可以将多个方向的投影特征融合起来进行识别以提高识别率。投影特征跟原始轮廓特征相比, 大大的降低了特征的维数, 所以可以降低计算复杂度。



图 2.7: 步态能量图计算示意图

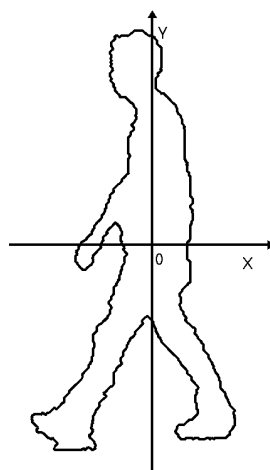


图 2.8: 轮廓线 (contour) 表达方式

2.1.3.3 步态能量图

步态能量图 (gait energy image, GEI) 法, 也被称作平均轮廓法, 最早在 Bobick 和 Davis 的论文 [14] 中被提出, 用于行为识别。后来 Han 等人 [29] 将这个方法用于步态识别, 取得了很大的成功。步态能量图的定义比较简单, 为一个步态周期内的所有轮廓的平均:

$$F(i, j) = \frac{1}{T} \sum_{t=1}^T I(i, j, t) \quad (2.6)$$

其中 I 是轮廓图像, (i, j) 是图像上的坐标, t 是序列中帧的编号。步态能量图上某一点的值的大小代表了这一位置出现人体的可能性。步态能量图的优点是 不容易受到轮廓噪声干扰, 具有比较好的鲁棒性, 但是特征维数较高, 可以用 PCA 等方法对步态能量图进行降维来降低特征维数。

2.1.3.4 轮廓线

前面提到的轮廓 (silhouette) 特征不仅记录了轮廓边缘的点, 还记录了轮

廓内部的点, 具有比较多的数据冗余。如果采用轮廓线 (contour) 来表示, 只需考虑轮廓边缘的点, 数据表达会更加紧凑。轮廓线表达一般将轮廓边缘表示为复平面空间的一个复数点:

$$z_i = x_i + j \cdot y_i, (i = 1, 2, \dots, N) \quad (2.7)$$

其中 x_i 和 y_i 分别是第 i 个点的横坐标和纵坐标。使用轮廓线特征, Wang 等人曾提出了基于统计形状分析的步态识别算法 [35] 和基于时空轮廓分析的算法 [7, 27]。

轮廓线特征虽然表达紧凑, 但是也带来了一个缺点, 就是轮廓线容易受到噪声的影响。针对这个问题, 我们提出了一个解决方案 [36], 其基本思想是在计算轮廓相似度之前, 先对轮廓进行匹配。这可以极大的提高算法对噪声的鲁棒性。

2.1.3.5 关键傅立叶描述子

关键傅立叶描述子 (key Fourier descriptors, KFD) 是我们提出的另一种步态特征 [28]。关键傅立叶描述子通过对人体轮廓线进行傅立叶变换后, 选择关键的描述子组成。一个步态周期中的所有轮廓线上的点 (复数点) 连成一个长的向量:

$$\mathbf{z} = [z_1, z_2, \dots, z_N]^T$$

全部傅立叶描述子可以通过对向量 \mathbf{z} 进行离散傅立叶变换得到。这些傅立叶描述子的关键部分可以通过文章 [28] 中的方法获得。

关键傅立叶描述子的一个明显的优点是不需要将所有的步态周期对齐。很多其他的步态特征提取方法要求将所有步态周期对齐, 比如将两脚距离最远的姿势作为周期的开始。由于傅立叶系数的幅度是时移不变的, 关键傅立叶描述子利用的是幅度信息, 去除了相位信息, 所以这种特征与步态周期的人起始姿势无关, 避免了对齐步态周期时引入的误差, 也降低了计算复杂度。

2.1.3.6 基于模型的特征

目前有多种基于模型的步态特征, 如关节角度变化、身高、腿长、步幅等。因为行走运动的关节角度变化包含丰富的识别信息 [37], 所以人体的关节变化是一种经常被使用的基于模型的步态特征。要提取出关节变化, 关键的一

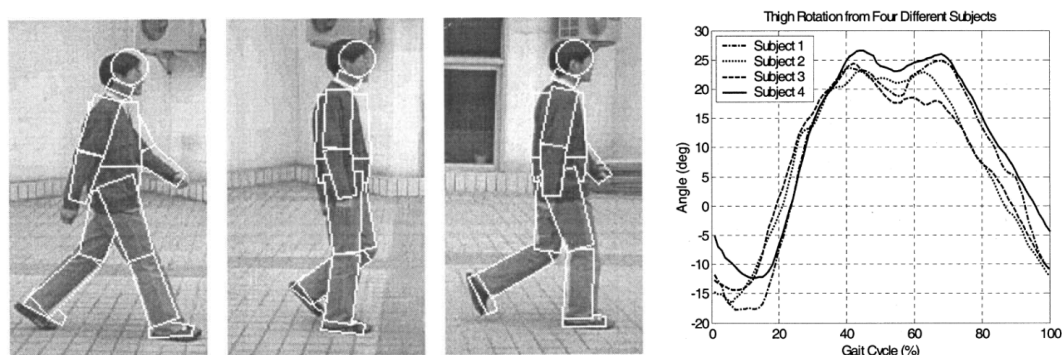


图 2.9: 人体主要部分的跟踪结果 (左侧 3 图) 以及 4 个人的大腿角度变化曲线 (右图) [2]

步是要准确跟踪出身体的每一个部分, 如头、躯干、上臂、下臂、大腿、小腿等。Wang 等人使用了粒子滤波的方法来跟踪身体的关键部分 [2], 跟踪结果如图 2.9 中左侧 3 幅图所示。数据库中某 4 个人的大腿的角度变化曲线如图 2.9 中最右侧的图所示, 从图中可以明显看出, 不同人的大腿角度变化曲线相差较大, 因此跟踪出的各个关节角度变化特征可以用来进行步态识别。另外, 还有一些其他的方法进行人体跟踪, 如 Zhang 等人所使用 Metropolis-Hasting 方法 [24]。

2.1.4 分类器

分类器的主要作用是确定一个未知类别的样本的类别。步态识别的一个特点是每一类的样本少, 特征维数高; 而如果使用步态特征进行性别识别、或者年龄段估计, 则每一类的样本数会相对多一些。对于不同的分类问题, 应该选择不同的分类器。在后面几章提到的工作中, 使用的分类器主要有如下几种: k 近邻分类器 (包括最近邻分类器), 贝叶斯分类器和支持向量机分类器。

2.1.4.1 k 近邻分类器

k 近邻分类器是 [38] 一种简单但是被经常使用的分类器。它根据特征空间内, 距离待分类的样本最近的 k 个已知类别样本来确定待分类样本的类别。如图 2.10 所示, 待分类样本 ‘*’ 的 5 近邻中有三个 ‘+’, 两个 ‘-’, 所以 ‘*’ 的类别应该为 ‘+’。 k 近邻分类器虽然是一个简单的分类器, 但是它具有比较强的适应能力, 目前被广泛使用。

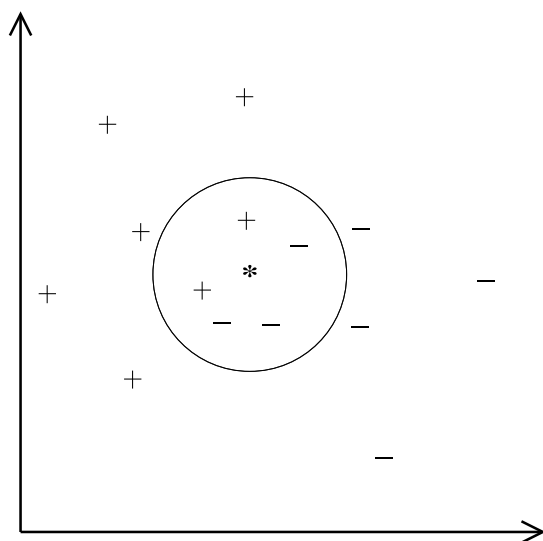


图 2.10: k 近邻分类器 (k=5)

2.1.4.2 贝叶斯分类器

贝叶斯分类器假定某一类别的样本在特征空间里呈正态分布，未知类别样本 \mathbf{x} 属于类别 i 的可能性是：

$$f_i(\mathbf{x}) = \frac{1}{(2\pi)^{N/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_i)^T \Sigma_i^{-1} (\mathbf{x} - \mu_i)\right) \quad (2.8)$$

其中 μ_i 和 Σ_i 分别是第 i 类的均值和方差， $|\Sigma_i|$ 是 Σ_i 的范数。 x 属于可能性最大的那个类别 i^* ：

$$i^* = \arg \max_i f_i(\mathbf{x}) \quad (2.9)$$

其中 i^* 是分类结果，是类别标签。

2.1.4.3 支持向量机

支持向量机 [39] (support vector machines, SVMs) 是一种最大化分类间隔的分类器。支持向量机可以同时最小化经验分类误差和最大化分类间隔。它将特征映射到更高维空间内，然后在高维空间内寻找一个线性的分类超平面。靠近分类面的特征向量被称作支持向量，意思是分类面由这些向量支持（影响），其他的特征向量不会影响分类面。

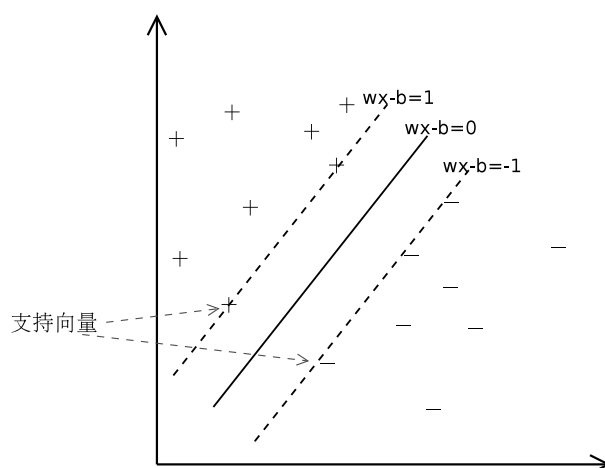


图 2.11: 支持向量机

给定一组标记好类别的样本 (\mathbf{x}_i, l_i) , $i = 1, 2, \dots, n$, 其中 $\mathbf{x}_i \in R^N$ 而且 $l_i \in \{-1, +1\}$, 支持向量机需要求解下列方程:

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & l_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \\ & \xi_i \geq 0 \end{aligned} \quad (2.10)$$

其中 C 是错误的惩罚系数, ϕ 是映射特征向量到高维空间的映射函数。一般来说不需要使用映射函数 ϕ 的显式表达, 而可以转化为核函数的表达。核函数 $K(\mathbf{x}_i, \mathbf{x}_j)$ 被定义为 $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ 。下面列出了四个常用的核函数。

- 线性核函数: $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$;
- 多项式核函数: $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d, \gamma > 0$;
- 径向基函数 (radial basis function, RBF):
 $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \gamma > 0$;
- sigmoid 核函数: $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r)$ 。

论文 [40] 指出, 线性核函数是径向基函数的一个特例, 且在某些参数下 sigmoid 核的作用跟径向基函数的类似, 因此可以将径向基函数做为首选核函数。



图 2.12: 论文 [3] 中的虚拟体模型

2.2 多视角步态识别

目前大多数步态识别算法都是针对特定视角的，而且一般采用侧面视角（行走方向跟图像平面平行）进行识别，在第四章中也证明了侧面视角是最适合进行识别的视角。但在实际应用中的问题是，无法要求被监控的行人按照固定的路线行走，期望总是能够得到侧面视角的图像是不切合实际的，因此需要分析多个视角的步态特征之间的关系，设计能从任意视角根据步态识别人的身份的算法来解决这一问题。目前，已经有一些研究者开始尝试对此问题进行研究。

2.2.1 虚拟体模型

MIT 的 Shakhnarovich 等人于 2001 年提出了一个与视角无关的步态识别算法 [3]。这个算法的输入为从几个视角拍摄的视频，首先将视频中的人的轮廓提取出来，然后用这些轮廓创建一个虚拟的体模型（visual hull），如图 2.12 所示。这个虚拟体模型不仅包含形状信息，而且还包含纹理信息。该方法不是直接从输入视频中提取步态特征，而是在虚拟的三维环境中，设置虚拟摄像机，从虚拟人体的侧面拍摄步态视频，从虚拟的步态视频中提取特征。无论人的行走方向如何，多个相机都可以将人的虚拟体建起来，用虚拟摄像机从侧面拍摄，来提取步态特征，因此这个算法是一个与视角无关的算法。

Shakhnarovich 的虚拟体算法为视角无关的步态识别提供了一个很新的思

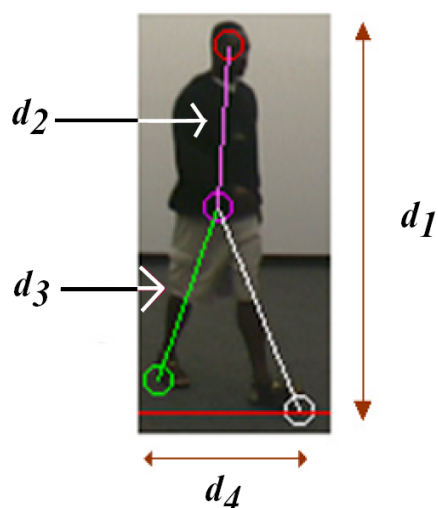


图 2.13: 论文 [4] 中使用的特征

路,但是这个算法在实际应用中会受到一定限制。首先,多个摄像头需要同步,这在实现上增加了复杂度;其次,监控区域为多个摄像头重叠的拍摄区域,比起单摄像机监控的区域要小;再者,为了提高模型精度,需要增加摄像机数目,这会大大增加算法复杂度。

2.2.2 身体参数

使用多摄像机的步态识别在应用中会受到很大限制,另外一个思路就是建立起不同视角的特征之间的关系模型,将从各个视角采集来的特征通过这个关系模型转换到一个基准角度(例如侧面视角)。Johnson 和 Bobick 的算法 [4] 就采用了这种思路。他们使用的特征是身体的静态参数,这些参数是身高 (d_1),躯干长度 (d_2)、腿的长度 (d_3) 和步幅 (d_4) (图 2.13)。

$$\mathbf{x} = [d_1, d_2, d_3, d_4] \quad (2.11)$$

不同角度的特征之间可以通过一个系数向量 $\mathbf{f} = [f_1, f_2, f_3, f_4]$ 建立联系。这个向量可以由两个角度的特征训练得到。使用某个视角的特征 \mathbf{x} 合成侧面视角特征 \mathbf{x}^* 可以通过下面的公式实现:

$$\mathbf{x}^* = \mathbf{f} \cdot \mathbf{x} = [f_1x_1, f_2x_2, f_3x_3, f_4x_4] \quad (2.12)$$

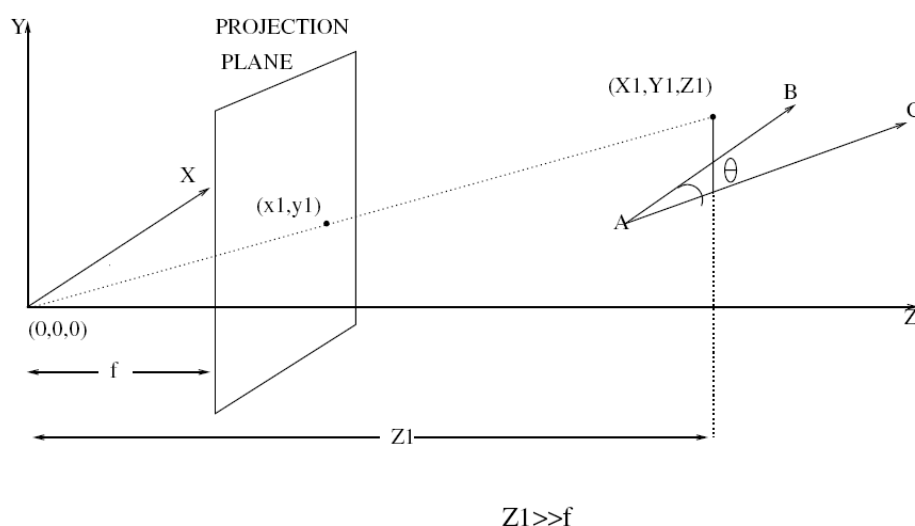


图 2.14: 论文 [5] 中的投影模型

BenAbdelkader 等人也提出了一个基于身体参数的步态识别方法 [41]。在识别中,使用的特征为步长、步频和身高。该算法在 41 人的数据库上获得了 65% 的最高识别率,说明身体参数具有一定的区分能力。

与前面的虚拟体的方法相比,这种方法只使用一个摄像机,且实现简单。但是本方法采用的特征—身体参数—的区分能力一般来说不是很高。在小规模的测试数据上,人的身高与身材比例这样的静态参数能够区分不同的人,但是当数据规模变大以后,会有很多人身高和身材比较相似,这个时候识别率将会明显下降。

2.2.3 特征合成

Kale 等人在 2003 年提出了一个用投影模型合成侧面视角轮廓的方法 [5],与前面两个算法相比,这个算法更加成熟,实用性更强。这个算法先计算出行走路线与图像平面的夹角 θ (图 2.14 中的 θ),夹角可以使用透视投影的方法或者用光流的方法计算得到。计算出夹角之后,假设人体是一个平面模型,将人体轮廓平面旋转 θ 角度,投影到侧面轮廓。Kale 文中的使用的步态特征是合成的轮廓的水平投影(第 2.1.3.2 节)。实验证明这个算法可以提高识别率。

但是由于这个算法假设人体是一个平面,而人体实际上是三维的体而不是平面,这种假设跟实际情况有比较大的差别。在行走角度跟图像平面夹角比较

大的时候，合成的侧面轮廓的身体躯干部分变得比真实轮廓宽，跟真正的侧面视角轮廓差别比较大。这种方法能提高腿部区域的区分能力，但是却不能提高躯干部分的区分能力。

2.2.4 研究现状小结

从前面的分析可以看出，已有的多视角工作对视角变化问题进行了探索性的研究，但是尚没有对视角与性能的关系进行深入分析，也没有有效的算法能够提取视角无关特征。视角变化对步态分析与识别仍是一个非常大的挑战，还需要更多更深入的研究。

第三章 多视角步态数据库

俗话说“工欲善其事，必先利其器”，要进行步态识别研究，首先需要一个好的步态数据库，用来评价设计的算法。随着步态识别研究的深入，很多研究单位都创建了步态数据库，这些数据库的创建极大的促进了步态识别研究的发展。但是随着步态识别研究的逐渐深入，原有的一些规模比较小、考虑因素比较少的数据库逐渐不能满足步态研究的需求。目前需要具有更多人数、更多变化因素以及更有挑战性的数据库。

3.1 公开的步态数据库

早期的步态识别研究是探索性的研究，这个时候的数据库规模一般比较小。文献中发现的最早的步态数据库是 1998 年加州大学圣地牙哥分校 (University of California, San Diego, UCSD) 的研究者创建的，这个数据库在论文 [42] 和 [43] 中被提及并被使用。UCSD 步态数据库包含 6 个人，每个人 7 个视频序列，共有 42 个序列。

2000 年，美国国防部高级研究计划局 (DARPA) 资助了 HID (Human Identification at a Distance) 项目，极大的促进了步态识别研究的发展，很多研究者开始进行远距离行人识别的研究，并创建了多个的步态数据库。麻省理工学院的人工智能实验室在 2001 年创建了一个 24 人的步态数据库 [20]。这个数据库中包含 14 个男性，10 个女性，每个人行走的时候，摄像机从侧面拍摄。每个人拍摄 4-22 个视频序列，总共拍摄了 194 个序列。

佐治亚理工学院 (Georgia Institute of Technology) 创建了一个包含 20 人的数据库。这个数据库采集过程中考虑了多种因素，如环境变化 (室内，室外)，相机到行人距离变化，视角变化等。而且除了拍摄视频数据外，还采集了人体关节的运动信息 (图 3.1)。这是一个相对比较全面的步态数据库。

卡耐基-梅隆大学 (Carnegie Mellon University) 用利用 3D 房间拍摄了一个多视角的数据库: Mobo 数据库。这个数据库的独特之处在于当人行走时，6 个摄像头从不同的角度进行拍摄 (图 3.2)，跟其他数据库相比具有更多的视角，这对于研究视角无关的步态识别算法比较有用。同时，除了视角外，在采集数据

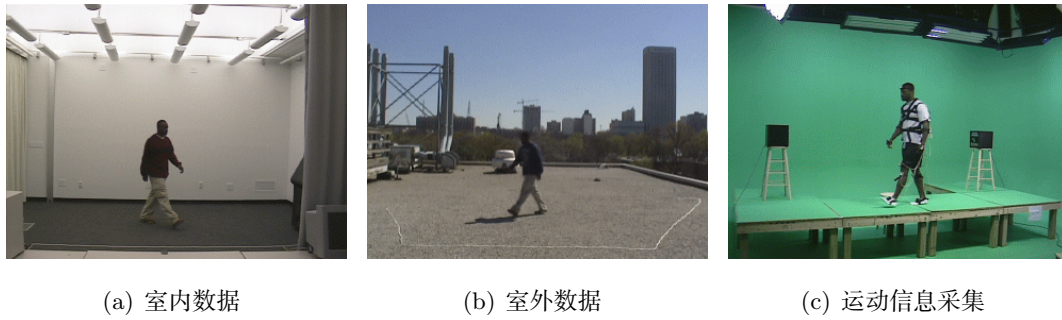


图 3.1: Georgia Tech 数据库例图

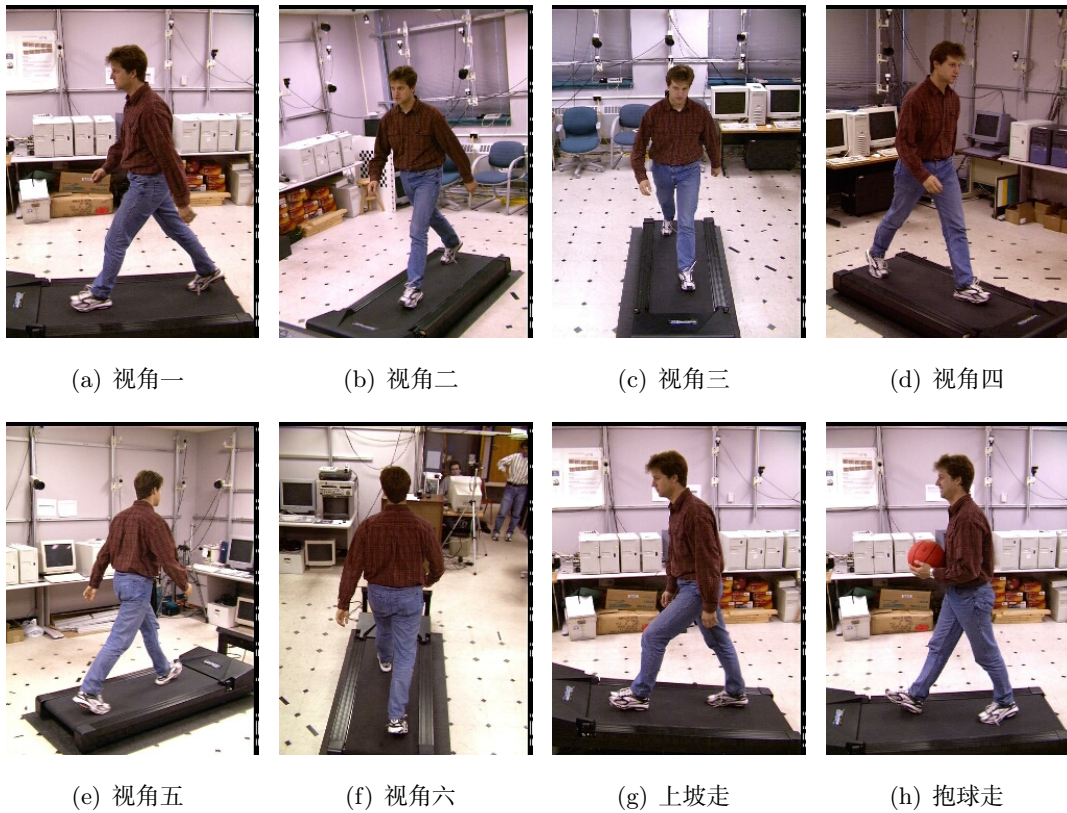


图 3.2: CMU Mobo 数据库例图

时其他因素也被考虑进来。被采集的人以四种方式行走：慢走，快走，上坡走和抱球走。但是采集数据时，人在跑步机上行走，而不是在路面上自然行走，这可能在一定程度上会影响到人的步态，使得在采集数据时的步态跟正常行走有一定差别。



图 3.3: HID-UMD 数据库 (数据集 1) 例图



图 3.4: HID-UMD 数据库 (数据集 2) 例图

马里兰大学 (The University of Maryland) 创建了包含人数更多的 HID-UMD 数据库 [44, 45]。这个数据库包含两个数据集，数据集 1 包含 25 人，从正面、右侧面、左侧面和背面四个视角拍摄 (图 3.3)。数据集 2 跟数据集 1 相比，包含更多人数，55 人，但是只有正面和一个侧面两个视角 (图 3.3)。HID-UMD 数据库虽然包含了比较多的人，但是图像分辨率很低，特别是数据集 1，行人在图像中很小，对算法很有挑战性。目前步态识别算法尚不成熟，所以创建能够容易提取清晰轮廓的数据库更有意义。

中国科学院自动化研究所最早在 2001 年创建了创建的 CASIA 步态数据库数据集 A [46] (即以前的 NLPR 步态数据库, 图 3.5)。这个数据库包含 20 人，在室外环境中从三个视角拍摄：正面视角，与正面成 45 度视角和侧面视角。每个人拍摄 12 个序列，共有 240 个序列。这是一个小规模的数据库。



图 3.5: CASIA 数据库 (数据集 A) 例图

除了前面提到的小规模的数据库外,目前有四个人数超过 100 人的大数据库,分别是由英国的南安普顿大学 (University of Southampton) 创建的 Soton 大步态数据库,美国南佛罗里达大学 (University of South Florida) 创建的 Gait Challenge 步态数据库,中国科学院自动化研究所创建的多视角步态数据库 (CASIA 步态数据库数据集 B, 本文将重点描述) 和红外步态数据库 (CASIA 步态数据库数据集 C)。目前这三个研究单位拥有的步态数据库规模最大,且都多次更新。



(a) 室内正常行走

(b) 室内跑步机上行走

(c) 室外行走

图 3.6: Soton 数据库例图

南安普顿大学创建了两个数据库,一个是人数较少(12人)小数据库,另一个是大数据库(115人) [47, 48]。小数据主要是用来评价衣着,鞋子,携带物品等因素对步态识别的影响。大数据分别在室内行走,室内跑步机和室外拍摄(图 3.6),主要是用来进行步态特征的唯一性研究。另外, Soton 数据库中的室内数据在拍摄时悬挂了绿色的幕布作为背景,背景非常简单,所以可以提取高质量的人体轮廓,这对分析研究步态特征是否具有识别能力非常有利。

与 Soton 数据库不同,南佛罗里达大学创建的 Gait Challenge 数据库 [25, 49] (图 3.7) 充分考虑了实际应用中的各种干扰因素。2001 年 5 月,首先创



图 3.7: USF Gait Challenge 数据库例图

建了一个 74 人规模的数据库，在拍摄中，人沿一个椭圆路线行走，两台摄像机分别从两个视角拍摄。除了视角，变化因素还有：两种鞋子类型，不带包和带包，草地和水泥地面。同年 11 月，又进行了第二次数据采集，使得数据库内的总人数达到 122 人，且有 33 人参加了两次数据采集，这样就可以分析时间变化的影响。这个数据库中的数据由于在室外采集，背景相对复杂，且人离摄像机比较远，所以提取出的人体轮廓质量一般不是很高，而且由于人沿着椭圆路线行走，视角一直保持变化，这对步态识别算法提出了很高的挑战。这个数据库可以用来评价一个算法在实际应用中的性能，但是并不适合分析单个因素（比如视角）的影响，因为这个数据库涉及的多个因素同时存在，且图像噪声也会带来一些干扰。

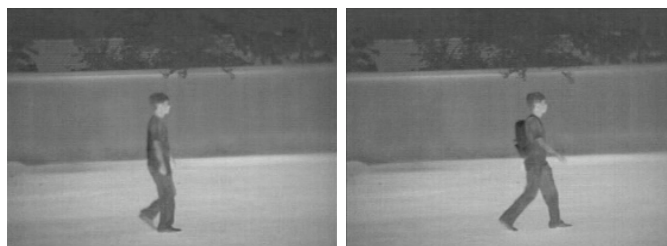


图 3.8: CASIA 数据库 (数据集 C) 例图

为了研究夜间的步态识别，中国科学院自动化所在 2005 年 7—8 月还创建了一个红外步态数据库 (CASIA 步态数据库数据集 C)。这个数据库中的数据是用红外 (热感) 摄像机在夜间拍摄 (图 3.5)，包含 153 人，每个人在 4 种条件下行走：正常行走，快走，慢走，带包走。该数据库中的每个人共有 10 个序列，总共有 1,530 个序列。这个数据库是目前唯一公开的大规模红外步态数据库。

前面介绍的各个数据库的详细情况列于表 3.1 中。

表 3.1: 目前已经公开的步态数据库

数据库	人数	序列数	场景	创建时间	变化因素
UCSD 数据库[42]	6	42	室外	1998 年	无
MIT AI 数据库 [20]	24	194	室内	2001 年	视角, 时间
Georgia Tech 数据库[50]	20	188	室内, 室外	2001 年	视角, 时间, 距离
CMU Mobo 数据库[44, 45]	25	600	在室内跑步机上	2001 年 3 月	6 个视角, 速度, 携带物品
HID-UMD 数据库 (数据集 1)	25	100	室外	2001 年 2-5 月	4 个视角
HID-UMD 数据库 (数据集 2)	55	220	室外	2001 年 7 月	2 个视角
Soton 小数据库[47]	12	-	绿色背景的室内	-	衣着, 携带物品, 鞋子, 视角
Soton 大型数据库[48, 47]	115	2,128	绿色背景的室内, 跑步机, 室外	2001 年 夏季	视角, 场景
Gait Challenge 数据库[25, 49]	122	1,870	室外	2001 年 5 月, 11 月	2 个拍摄点, 路面, 鞋子, 携带物品, 时间
CASIA 步态数据库 (数据集 A) [46]	20	240	室外, 水泥路面	2001 年 12 月	3 个视角
CASIA 步态数据库 (数据集 B) [51, 46]	124	13,640	室内	2005 年 1 月	视角, 衣着, 携带物品
CASIA 步态数据库 (数据集 C) [46]	153	1,530	室外, 水泥路面, 红外	2005 年 7-8 月	速度, 携带物品

3.2 多视角步态数据库的设计

最近十年，步态识别研究取得了比较大的进展，研究者们普遍认为步态可以作为一种特征进行身份识别。但是步态特征是一种行为特征，与虹膜、指纹、人脸等生理特征相比，特征的鲁棒性有所欠缺。步态特征容易受到各种因素的影响。

进行步态识别的研究，需要对影响步态识别的各种变化因素进行分析。分析出这些因素如何影响识别，对于提取鲁棒的步态特征以及设计有效的分类器都有很重要的意义。在变化因素中，比较重要的有视角、衣着、携带物品、速度、鞋子、路面等因素，其中前三种又是最常见的三种变化，因为在视频监控中，步态的优势是隐蔽性，因此我们无法要求行人按照规定的直线方向行走、不改变衣着或者不能携带物品。

研究这些变化因素产生的影响需要一个包含这些因素的步态数据库，在上一节提到的数据库中有多个数据库采集了各种变化因素的数据。但是这些数据库大多数规模太小，大都不超过 50 人，这样的小规模数据目前已经难以满足步态分析与识别的要求。在创建 CASIA 数据库数据集 B 和 C 之前，超过 100 人的大规模数据库有两个：Soton 大数据库和 Gait Challenge 数据库。Soton 大数据库有 6 个视角，在室内拍摄，可以容易的分割出人的轮廓。但这个数据库的缺点是变化较少，对每个场景，只有两个视角，而且公开的数据大都只有一个视角（侧面视角），所以无法使用这个数据库对视角等变化因素进行研究。Gait Challenge 数据库在数据采集过程中包含了大量变化因素，数据在室外采集，可以用来评价一个算法在实际应用中的效果。但是这个数据库并不适合研究单个变化因素的影响，因为采集数据时，行人沿着椭圆路线行走，视角一直在变化，而且背景相对复杂，分辨率不高，很难提取高质量轮廓。这样，轮廓噪声和视角变化无法跟其他变化因素分开，所以无法对视角、衣着以及携带物品等因素进行单独研究。

目前需要这样的一个步态数据库：数据在室内简单背景下拍摄，易于提取人体轮廓；从多个视角拍摄数据，两个相邻视角之间的夹角相对比较小，这样的数据可以用来分析视角变化与识别性能的关系；尽量包含比较多的变化因素，如衣着变化和携带包裹与否；数据规模要大，人数应该超过 100 人，每个人的视频序列要尽量多。

CASIA 多视角数据库数据集 B 充分考虑了视角、衣着和携带包裹这三种最

重要的变化因素。对于视角变化，由于人体是左右对称的，所以可以只拍摄人体的一侧则可。从人体正面开始，以 18° 为角度间隔从身体左侧进行数据拍摄，直到人体背面，共有 11 个视角。数据库的具体设计指标如下：

- 包含 100 人以上；
- 行走状态有正常行走，穿大衣行走和带包行走三种，每种行走状态拍摄 2-6 次；
- 数据拍摄角度分别与人的行走方向成 $0^\circ, 18^\circ, 36^\circ, \dots, 180^\circ$ (间隔为 18°)，共 11 个拍摄方向；
- 11 个角度的数据同时采集；
- 场景中放置定标块；
- 视频尺寸： 320×240 像素；
- 室内采集。

3.3 多视角步态数据库的采集

3.3.1 采集环境设置

为了采集 CASIA 多视角步态数据库，专门建立了一间数据采集实验室，采集实验室内的采集设备如图 3.9 所示。11 只摄像头 (型号：Fametech 318SC) 呈半圆形分布，相邻摄像头的拍摄方向夹角为 18° 。这 11 只摄像头分别连到 3 台计算机上，计算机 1 和计算机 3 分别连 4 只摄像头，计算机 2 连接 3 只摄像头，网络控制程序在计算机 2 上运行。三台计算机通过网络相连。

与其他数据库不同，该数据库在采集时，需要 11 个摄像头同时采集数据。如果完全采用人工操作，需要为每个摄像头配备一个操作人员，这不仅需要很多人力，而且难以协调和控制。为了方便控制这 11 个摄像头，我们专门设计了一个多视角采集软件。采集软件分两部分，分别是负责从摄像头抓数据的 Capture 和负责控制 Capture 的 Controller (图 3.10)。Capture 的内部有两个模块：网络通讯模块和视频采集模块 (图 3.11)，网络通讯模块负责从网络接收命令和发送信息，视频采集模块负责根据网络通讯模块发来的命令从摄像头采集

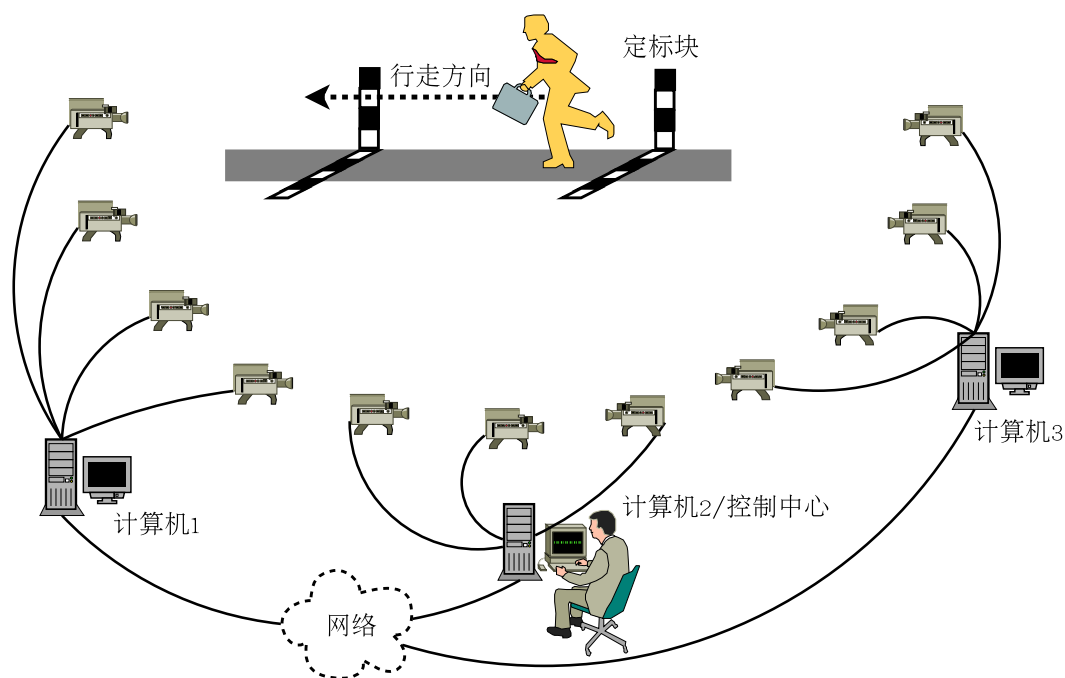
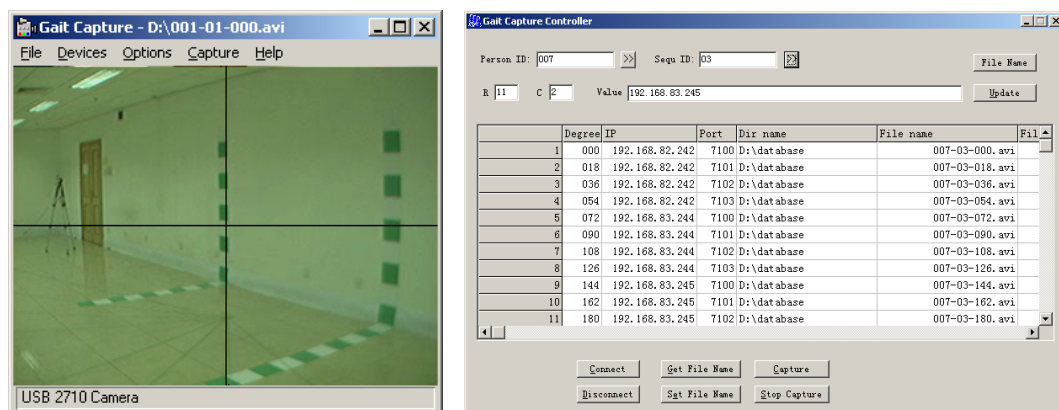


图 3.9: 数据库采集方案示意图



(a) 采集端界面 (Capture)

(b) 控制端界面 (Controller)

图 3.10: 步态数据采集软件界面

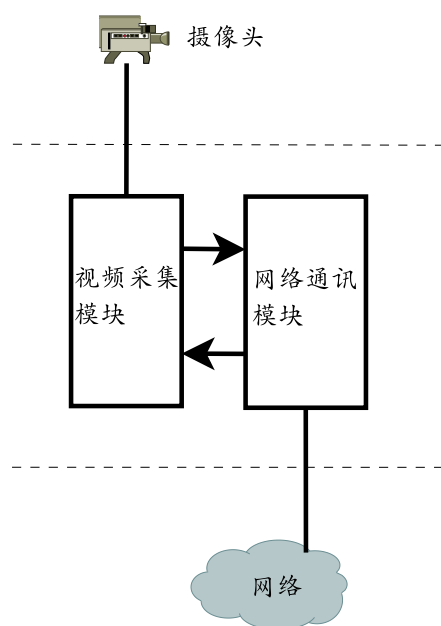


图 3.11: 采集软件的 Capture 部分

视频数据并保存。Controller 负责发送命令到各个 Capture。在软件辅助下，控制操作 11 个摄像头就变得比较简单方便。

既然人行走的时候，从多个视角同时拍摄图像，那么场景的一些几何信息，人体的一些参数，比如身高等，可以用三维重建计算取到。为了能够获得场景中的三维信息，4 条定标带被分别贴在垂直的墙上和水平的地面上。定标带上是绿色和白色相间的方格子，格子的大小为 $20\text{cm} \times 20\text{cm}$ ，定标带之间的距离是 240cm 。设置好的采集环境如图 3.12 所示。

3.3.2 数据采集

在录制视频数据之前，先记录采集对象的个人信息，个人信息包括姓名，性别和身高。身高为包含鞋子和头发的高度。在采集视频数据时，要求采集对象以自然的状态沿规定的路线行走。当采集对象走入采集区域，由 Controller 发出采集指令，11 只摄像头同时开始采集；当采集对象离开采集区域时，Controller 发出停止采集的指令，摄像头停止采集工作，并将采集到的视频文件被存到计算机硬盘上。

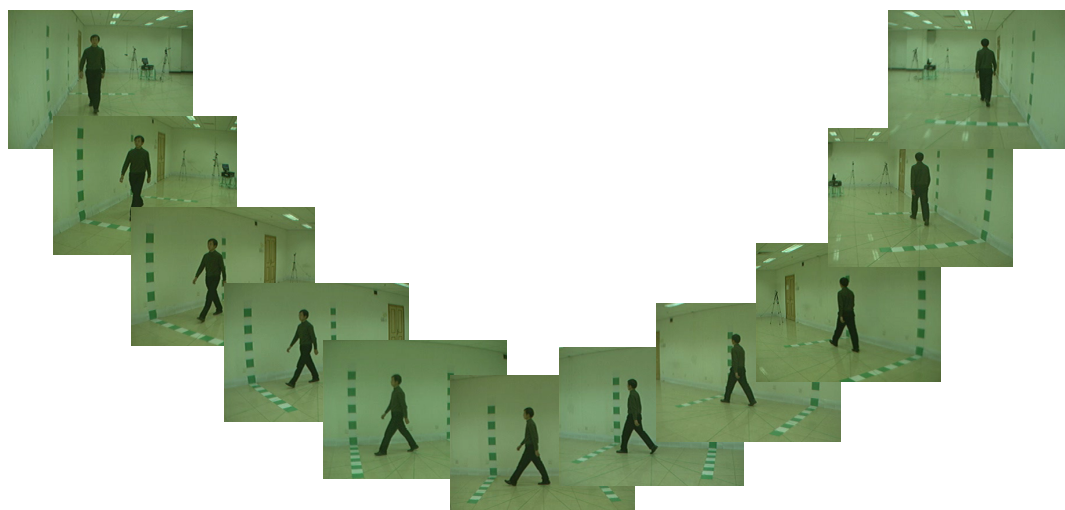


图 3.12: 数据库采集场景

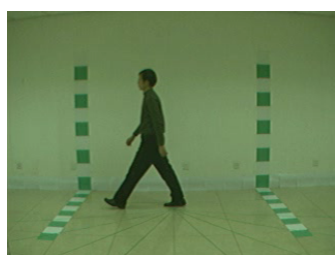
对于每一个采集对象，都进行了三种行走状态的数据采集。首先是正常行走状态，采集 6 个序列；然后是改变衣着，即穿上大衣，再采集 2 个序列；最后是不穿大衣，但携带包裹状态下行走，也采集 2 个序列。这样，每个人每个视角共采集 10 个序列。在采集过程中，改变衣着时所穿的大衣为采集对象自己的衣服，包裹也为采集对象自己的包裹，类型有背包、斜挎包、女式包等，采集对象根据自己的习惯选择携带包的方式（拎、挎、背等）。最终共采集了 124 人的步态数据，其中男性有 93 人，女性 31 人，中国人 123 人，欧洲人 1 人。采集到的数据如图 3.13 所示。

3.4 数据管理与共享

CASIA 步态数据库数据集 B 是一个大规模的步态数据库，共有 $124 \times 11 \times 10 = 13,640$ 个视频序列，如果以未压缩的视频格式进行存储，大约需要 400GB 的存储空间，这不便于视频的处理和传输。为了减少存储数据库所需空间，XviD



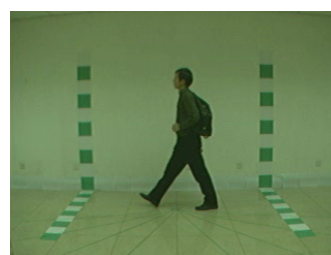
(a) 同时从11个视角拍摄的图像



(b) 正常行走



(c) 穿大衣行走



(d) 带包行走

图 3.13: CASIA 数据库 (数据集 B) 例图

编码器 (XviD codec)¹ 被用来对视频文件进行压缩, XviD 是一种非常优秀的 MPEG-4 压缩算法, 它能在保证视频质量的同时进行高压缩比的压缩。压缩后的数据库大小为 6GB。

为了促进步态识别研究的发展, 使步态作为一种生物特征迅速在智能视觉监控以及其他领域被广泛应用, 中国科学院自动化研究所向国内外同行免费提供 CASIA 步态数据库。数据库申请者需要首先签署一份使用协议, 协议中声明 CASIA 步态数据库为中国科学院自动化研究所所有, 得到使用授权的申请者可以在科学研究中使用本数据库, 但未经许可不可以在商业中使用此数据库。使用者未经许可不可以修改或向第三方发放本数据库。另外, 如果使用者在本数据库帮助下取得了科研成果, 有将成果向中国科学院自动化研究所告知的义务。

¹<http://www.xvid.org>

该协议为 CASIA 步态数据库知识产权的保护提供了法律保障。

为了方便广大数据库申请者方便的获得数据库,目前提供了两种获取数据库方法:如果申请者具有速度比较快的互联网连接,申请者可以通过因特网从中国科学院自动化研究所的数据库专用 FTP 上下载数据库;另外,申请者也可以通过快递公司或其他途径到中国科学院自动化研究所取以 DVD 光盘为介质的步态数据库。同时,为了让更多的科研工作者了解 CASIA 步态数据库,还为本数据库建立了专门的网页²,该网页详细介绍了数据库的各种参数,以及获得数据库的方法。截至 2007 年 10 月 23 日,已经有 27 个国家和地区的 138 个研究单位申请了 CASIA 步态数据库。

²<http://www.cbsr.ia.ac.cn/Gait%20Database.htm>

第四章 步态识别算法性能分析与评价

步态识别是一个新兴的研究方向,近几年引起了研究者的巨大兴趣,提出了很多种步态识别算法。虽然大量的算法被提出,每个作者都声称自己的算法与其他的算法相比,在某个方面具有优越性。但是这些比较往往不是在同样的标准下进行,比如数据库的选择、注册样本的数目、算法实现中的一些参数的设置等都会影响算法的比较。因此目前亟需一个通用的评价标准来对各种不同的步态识别算法进行评价。

美国南佛罗里达大学的研究者 Sarkar 等人提出了一个步态识别的评价标准 HGCP (The HumanID Gait Challenge Problem) [25, 49], 来尝试解决前面提到的问题。HGCP 包括一个 122 人的大数据库, 一个基准 (baseline) 算法和 12 个实验。这 12 个实验考虑了视角、鞋子、行走路面、衣着、时间和携带包裹等因素的影响。如果要评价一个步态识别算法的性能, 可以用这个算法的 12 个实验结果跟基准算法进行对比。不过 HGCP 中所用的数据库为室外拍摄, 受各种噪声影响, 比较难获取高质量的轮廓。同时鞋子、行走路面、衣着、时间变化和携带包裹等影响因素跟噪声混在一起, 难以准确分辨出性能降低是由于噪声, 还是由于其他因素引起。另外, 数据库中的人的行走路线为一个椭圆路线, 因此在人行走的过程中, 行走方向跟拍摄方向的夹角一直在变化, 而不是一个固定的角度, 所以难以评价不同角度的影响。综合来说, Sarkar 等人提出的评价标准, 为步态识别算法提供了一个很有挑战性的环境, 但是却无法具体细致的分别分析各种因素产生的影响。

为了详细的分析各种因素分别引起的影响, 本文提出了一个步态识别算法评价框架, 这个评价框架包含三部分: 一个大数据库, 三组实验和一系列评价标准。除此之外, 为了深入分析视角变化对识别率的影响, 提出了一个用于估计任意角度的识别率的数学模型。

4.1 步态识别算法评价

本评价框架所包含的数据库 CASIA 步态数据库数据集 B 已经在第三章中介绍, 下面将详细介绍实验和评价标准的设计。

4.1.1 实验设计

CASIA 步态数据库数据集 B 包含 124 个人, 11 个视角, 3 种变化条件。针对这个数据库, 设计了 3 组实验 (实验集 A, 实验集 B 和实验集 C) 来评价步态识别算法。

实验集 A 只使用了正常行走状态 (无衣着变化和携带包裹) 的数据。实验集 A 的每一个实验, 首先选择所有人的某个角度的正常行走状态的 6 个序列中的前 4 个序列, 放入注册集, 然后再任选一个角度的正常行走状态的后 2 个序列放入测试集。这样两个视角任意组合, 实验集 A 共有 121 (11×11) 个实验。实验集 A 可以用来评价视角变化对识别性能的影响, 或者算法对视角变化的鲁棒性。采用步态能量图法 (gait energy image, GEI, 第 2.1.3.3 节) 得到的实验集 A 的实验结果列在表 4.1 中。

实验集 B 可以用来评价衣着和视角对识别性能的影响。跟实验集 A 一样, 实验集 B 中的任意一个实验, 注册数据是正常行走状态的 6 个序列中的前 4 个序列。但是测试集是改变衣着 (穿大衣) 后的 2 个序列。实验集 C 可以用来评价携带包裹和视角对识别性能的影响。注册集跟实验集 A 和 B 一样, 但是测试集是携带包裹的 2 个序列。采用步态能量图法得到的实验集 B 和 C 的实验结果分别列在表 4.2 和 4.3 中。这样就设计了 3 组共 363 个实验。

表 4.1: 实验集 A 实验结果 (GEI 算法)

CCR (%)	测试集角度 θ_p (正常行走 #5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_r (正常行走 #1-4)	0°	99.2	31.9	9.3	4.0	3.2	3.2	2.0	2.0	4.8	12.9	37.9
	18°	23.8	99.6	39.9	8.9	4.4	3.6	3.6	5.2	13.7	33.5	10.9
	36°	4.4	37.9	97.6	29.8	11.7	6.9	8.1	13.3	23.4	13.3	2.0
	54°	2.4	3.6	29.0	97.2	23.0	16.5	21.4	29.0	21.4	4.8	1.2
	72°	0.8	4.4	7.3	21.8	97.2	81.5	68.1	21.0	5.6	3.6	1.6
	90°	0.4	2.4	4.8	17.7	82.3	97.6	82.3	15.3	5.2	3.6	1.2
	108°	1.6	1.6	2.0	16.9	71.4	87.9	95.6	37.1	6.0	2.0	2.0
	126°	1.2	2.8	6.0	37.5	33.5	22.2	48.0	96.8	26.6	4.4	2.0
	144°	3.6	5.2	28.2	18.5	4.4	1.6	3.2	43.1	96.4	5.6	2.8
	162°	12.1	39.1	15.7	2.4	1.6	0.8	0.8	2.4	5.2	98.4	28.6
	180°	41.1	19.8	8.1	3.2	2.0	0.8	1.6	3.6	12.5	51.2	99.6

表 4.2: 实验集 B 实验结果 (GEI 算法)

CCR (%)	测试集角度 θ_p (穿大衣#1-2)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	24.6	6.9	3.2	1.6	0.8	0.8	0.8	0.8	1.6	5.6	8.1
	18°	4.4	27.0	18.5	6.9	0.8	0.8	0.8	2.4	5.6	11.7	2.8
	36°	1.6	8.5	30.2	16.5	1.2	1.2	1.6	6.9	9.3	3.6	0.8
	54°	0.8	2.4	10.1	30.6	5.6	4.4	7.7	14.1	5.6	2.4	0.8
	72°	0.0	2.4	5.6	7.7	31.0	21.8	14.9	8.9	2.8	2.4	0.4
	90°	1.2	2.4	4.0	6.0	20.6	32.7	16.5	6.0	3.6	3.2	0.8
	108°	1.6	2.0	2.4	4.8	17.7	27.8	30.2	9.3	4.8	2.0	1.6
	126°	1.6	1.6	1.6	4.4	10.1	10.1	18.5	26.2	8.9	1.6	1.6
	144°	2.4	2.8	4.0	12.5	4.4	2.4	4.4	18.1	30.6	1.2	2.0
	162°	2.8	7.7	9.7	2.0	0.4	0.8	0.8	1.6	4.0	27.0	6.5
	180°	9.3	6.0	3.2	0.8	1.2	0.0	0.0	1.6	5.6	12.5	27.4

表 4.3: 实验集 C 实验结果 (GEI 算法)

CCR (%)	测试集角度 θ_p (带包行走#1-2)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	80.2	20.2	5.2	2.4	2.4	2.0	2.0	2.0	4.4	14.5	25.8
	18°	16.9	76.2	36.7	7.3	3.6	2.8	2.8	4.8	10.9	18.1	8.1
	36°	3.6	23.0	74.6	24.2	9.3	8.1	7.3	10.5	15.3	6.5	1.6
	54°	0.8	2.8	19.4	66.5	19.0	13.3	14.5	18.5	7.3	4.0	1.6
	72°	0.4	4.8	6.5	8.9	60.5	31.0	22.2	11.7	4.0	3.6	1.2
	90°	0.4	2.8	5.2	8.5	42.3	52.0	31.9	9.7	6.0	3.2	2.0
	108°	1.6	1.2	3.6	7.3	39.9	44.0	57.3	23.4	6.5	3.6	1.6
	126°	0.8	2.8	4.0	17.3	25.4	14.9	27.8	65.7	14.5	2.0	1.6
	144°	2.0	4.8	15.3	11.7	6.9	2.0	4.4	31.9	64.1	2.4	1.2
	162°	7.7	23.8	13.7	3.6	2.4	2.0	2.4	4.4	6.0	68.1	19.0
	180°	30.2	12.9	6.5	2.0	2.0	2.0	1.2	2.8	7.7	31.5	80.2

表 4.4: 实验集 A 实验结果 (KFD 算法)

CCR (%)	测试集角度 θ_p (正常行走#5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	71.8	5.2	5.2	2.4	0.8	1.2	1.2	0.8	1.6	3.2	33.1
	18°	3.6	49.2	14.5	4.4	2.8	3.2	4.0	3.6	4.0	8.9	4.0
	36°	2.8	12.1	72.6	11.7	3.6	2.8	2.0	3.2	14.1	10.9	2.4
	54°	2.0	3.2	10.5	69.4	7.7	2.4	4.4	14.9	10.9	3.2	0.8
	72°	0.4	0.8	2.8	12.9	77.8	16.9	25.0	8.9	2.4	1.2	0.0
	90°	0.4	0.8	3.6	4.4	23.0	75.0	20.6	4.0	2.0	0.8	1.2
	108°	0.4	2.4	3.2	5.2	20.6	21.4	69.8	10.5	2.8	1.2	0.8
	126°	0.8	3.6	4.8	14.9	11.7	5.6	14.1	71.4	10.5	3.6	1.6
	144°	2.0	6.9	16.1	12.1	4.0	2.4	2.8	12.5	71.0	11.7	3.2
	162°	2.8	10.9	10.9	1.6	2.0	2.4	2.8	6.0	11.3	72.2	3.6
	180°	30.6	3.2	4.8	1.6	2.0	2.4	1.2	2.8	3.6	7.3	67.7

表 4.5: 实验集 B 实验结果 (KFD 算法)

CCR (%)	测试集角度 θ_p (正常行走#5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	10.5	1.6	2.0	1.2	0.8	0.8	1.2	1.6	2.0	1.6	6.0
	18°	0.4	7.7	5.6	2.4	1.2	0.0	1.2	0.8	1.2	4.0	0.8
	36°	1.2	3.2	7.3	4.0	0.8	1.2	0.8	0.8	6.0	1.6	1.6
	54°	1.2	1.2	2.0	12.1	1.6	1.2	0.8	2.0	2.8	2.4	1.6
	72°	1.6	2.0	2.0	3.6	10.5	6.9	5.6	4.8	1.2	2.0	0.4
	90°	1.6	1.2	0.8	2.4	5.2	9.3	5.6	2.8	1.6	2.0	0.8
	108°	2.0	1.2	0.4	2.0	3.6	5.2	6.5	3.2	0.8	1.2	0.8
	126°	2.4	1.2	0.8	3.2	2.8	2.4	2.8	6.9	0.8	1.6	2.4
	144°	0.8	1.6	2.0	2.8	2.0	1.2	1.6	0.8	8.5	2.8	0.8
	162°	1.6	2.4	4.0	1.2	1.2	1.2	0.8	0.4	2.8	10.5	0.8
	180°	6.0	0.8	0.8	0.8	1.6	1.6	2.8	1.6	3.6	2.8	6.0

表 4.6: 实验集 C 实验结果 (KFD 算法)

CCR (%)	测试集角度 θ_p (正常行走#5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	27.8	4.8	2.4	2.0	0.4	2.4	2.0	2.4	2.4	2.0	12.9
	18°	3.6	19.0	6.9	3.2	2.0	1.6	2.4	2.4	2.0	8.9	3.2
	36°	1.2	6.0	22.2	6.9	2.4	4.4	1.2	2.0	8.5	7.7	3.6
	54°	1.2	2.8	4.4	19.4	6.9	2.0	2.0	7.7	5.6	2.8	1.6
	72°	0.8	3.2	2.0	5.2	14.1	11.3	10.5	6.5	2.0	2.8	1.2
	90°	0.0	1.6	0.8	2.4	8.5	14.1	5.2	3.2	0.4	1.6	0.8
	108°	2.0	2.8	3.6	2.8	10.9	12.1	15.7	5.2	3.2	1.2	2.0
	126°	1.6	2.8	3.6	8.9	6.0	5.2	9.7	19.4	5.2	1.6	1.6
	144°	3.2	6.5	8.1	5.6	2.8	2.4	6.5	9.7	25.8	8.5	3.6
	162°	2.8	8.9	5.2	2.4	2.0	1.6	2.0	4.4	7.3	25.8	3.6
	180°	14.5	1.2	2.4	2.0	3.2	2.0	2.8	2.4	2.8	4.4	26.6

4.1.2 算法评价标准

上一节提到的 3 个实验集中的 363 个实验结果可以准确的刻画一个步态识别算法对视角、衣着和携带包裹的鲁棒性, 以及这些因素对算法的影响。但是我们希望得出一些定性的结论, 比如“视角变化对算法 A 的影响比对算法 B 的影响小”, “衣着变化对算法 C 的影响很小”等, 所以需要设计一些简单直观的评价标准。

首先设计了三个评价参数 Δ^A , Δ^B 和 Δ^C , 其定义如方程 4.1 所示。

$$\begin{aligned}
 \Delta^A &= \frac{1}{11} \sum_{i=1}^{11} C_{ii}^A \\
 \Delta^B &= \frac{1}{11} \sum_{i=1}^{11} C_{ii}^B \\
 \Delta^C &= \frac{1}{11} \sum_{i=1}^{11} C_{ii}^C
 \end{aligned} \tag{4.1}$$

其中 C^A , C^B 和 C^C 分别是实验集 A, B 和 C 中的正确识别率, 如表 4.1 - 4.3 所示, 是一个 11×11 的矩阵。 C_{ii}^X 是 C^X ($X \in \{A, B, C\}$) 对角线上的元素。

Δ^A , Δ^B 和 Δ^C 是在视角没有变化的情况下的平均识别率。而且在实验集 A (表 4.1) 中, 对角线上的实验是在各种条件没有明显变化的情况下进行的,

所以 Δ^A 可以评价视角以及其他因素没有明显变化的情况下算法的识别性能。同样的, Δ^B 可以用来评价算法对于衣着改变的鲁棒性; Δ^C 可以用来评价算法对于携带包裹的鲁棒性。这三个评价参数的作用如表 4.7 所示。

表 4.7: 视角无变化时的评价标准列表

评价参数	被评价因素
Δ^A	条件无明显变化
Δ^B	改变衣着
Δ^C	改变携带条件

采用步态能量图 (gait energy image, GEI, 第 2.1.3.3 节) 和关键傅立叶描述子 (key Fourier descriptors, KFD, 第 2.1.3.5 节) 方法的性能评价参数 Δ 如表 4.8 所示。从这个表中的数据我们可以看出在无视角变化的情况下, 步态能量图特征明显优于关键傅立叶描述子特征。

前面提出的三个评价标准是用来衡量在视角没有明显改变的情况下的识别性能, 而评价算法对视角的鲁棒性, 可以用下列评价标准:

$$\begin{aligned}
 \sigma_i^A &= \frac{1}{N \cdot \Delta^A} \sum_{|j-k|=i} C_{jk}^A \\
 \sigma_i^B &= \frac{1}{N \cdot \Delta^B} \sum_{|j-k|=i} C_{jk}^B \\
 \sigma_i^C &= \frac{1}{N \cdot \Delta^C} \sum_{|j-k|=i} C_{jk}^C
 \end{aligned} \tag{4.2}$$

$$i \in \{1, 2, \dots, 10\} \quad j, k \in \{1, 2, \dots, 11\}$$

其中 N 是公式中参与求和的 C_{jk}^X 的数目, $X \in \{A, B, C\}$ 。上式中 σ_i^X 是当测试数据的角度跟注册数据角度相差 $18 \cdot i^\circ$ 时, 正确识别率与无角度差别情况下识别率的比值, 可以用来评价识别率随角度变化的关系。对角度变化不敏感的算

表 4.8: GEI 和 KFD 算法在视角无变化时的比较

算法	Δ^A	Δ^B	Δ^C
步态能量图法 (GEI)	97.7%	28.9%	67.8%
关键傅立叶描述子法 (KFD)	69.8%	8.7%	20.9%

表 4.9: GEI 和 KFD 算法在视角变化时的比较

i	σ_i^A (GEI)	σ_i^B (GEI)	σ_i^C (GEI)	σ_i^A (KFD)	σ_i^B (KFD)	σ_i^C (KFD)
1	0.418	0.423	0.350	0.173	0.354	0.324
2	0.164	0.192	0.139	0.084	0.232	0.187
3	0.086	0.118	0.094	0.051	0.192	0.122
4	0.076	0.097	0.078	0.053	0.164	0.132
5	0.064	0.101	0.061	0.055	0.168	0.128
6	0.076	0.089	0.078	0.066	0.216	0.169
7	0.071	0.111	0.078	0.066	0.212	0.173
8	0.155	0.158	0.139	0.073	0.222	0.234
9	0.142	0.149	0.159	0.047	0.138	0.110
10	0.404	0.301	0.413	0.456	0.689	0.656

法应该有比较大的 σ_i^X (σ_i^X 越接近 1 越好), 这样的算法是我们需要设计的算法。

同前面比较 Δ 一样, 我们也对采用步态能量图和关键傅立叶描述子方法的性能评价参数 σ_i 进行比较, σ_i 的值被列在表 4.8 中。从这个表中的数据我们可以看出只有视角变化的情况下, 步态能量图特征优于关键傅立叶描述子, 但是如果既有视角变化, 又有衣着变化或者携带物品变化, 关键傅立叶描述子的大部分 σ_i 值要高于步态能量图的值。也就是步态能量图特征对视角变化更鲁棒些, 即当几种变化因素同时存在的时候, 关键傅立叶描述子更优一些。

综上所述, 本文提出的评估框架可以有效的评价步态识别算法的性能, 并能衡量算法对各种变化因素的鲁棒性。

4.2 视角与识别率关系模型

在步态识别中, 有很多因素影响识别的性能, 比如视角、衣着、携带物品、路面、光照、精神状态等。在这些影响因素中, 视角是非常重要的一个因素。目前虽然有大量的步态识别算法被提出, 但是这些算法大部分都是基于特定视角的, 对视角变化不鲁棒, 这大大的限制了步态识别的应用。除此之外, 在步态识别中还有两个未解决的问题: ① 哪个视角最适合用于步态识别? Kale 等人[5]认为侧面视角最好, 但是目前为止并没有在理论上或者实验上证明侧面视角优于

其他视角；② 视角变化是如何影响识别率的？直观上来讲，注册集的角度跟测试集的角度差别越大，识别率越低。但是它们之间的精确关系尚需研究和发掘。上述两个问题的答案会对步态识别系统的设计有着重要的指导意义。

在本节后面的部分，将提出两个模型来尝试回答上面提到的两个问题。一个模型是几何模型，用来进行定性的分析；另一个模型是代数模型，用来进行定量的分析。

4.2.1 识别率函数

在此处，将正确识别率 (correct classification rate, CCR) 作为评价步态识别算法的指标。如图 4.1 所示，假设注册集的拍摄视角跟人行走方向的夹角为 θ_g ，测试集的拍摄视角跟人行走方向的夹角为 θ_p ，那么，正确识别率可以视作 θ_g 和 θ_p 的函数，如下所示：

$$CCR = f(\theta_g, \theta_p) \quad \theta_g \in [0^\circ, 360^\circ), \theta_p \in [0^\circ, 360^\circ) \quad (4.3)$$

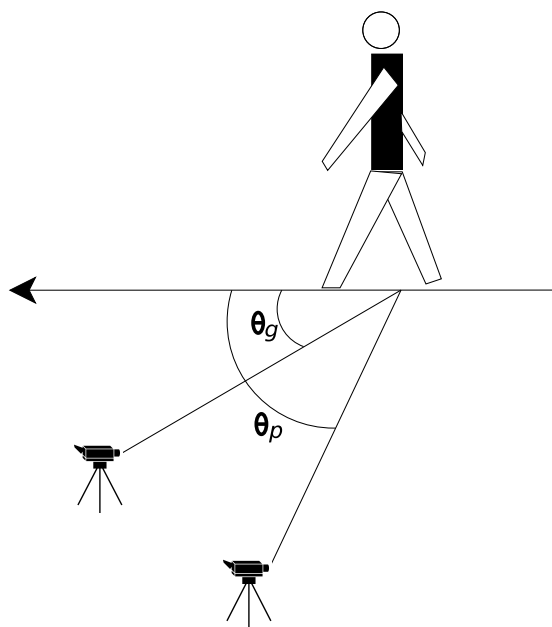


图 4.1: 多视角步态识别示意图 (注册集角度为 θ_g ，测试集角度为 θ_p)

如果能够得到函数 $f(\theta_g, \theta_p)$ 的数学表达式，那么 θ_g 和 θ_p 与识别性能的关系将会很容易得到。但是，获得 $f(\theta_g, \theta_p)$ 的数学表达式，并不是一件很

容易的事情。可以说通过实验的方法，在空间 $\mathbb{P} = [0^\circ, 360^\circ) \times [0^\circ, 360^\circ)$ 中任意的一点 (θ_g, θ_p) 获取 f 的值几乎是不可能的，因为空间 \mathbb{P} 是一个连续空间。一个变通的方法是先计算 f 在 \mathbb{P} 的一个离散子集上的取值，然后通过数据拟合的方式，估算 f 在 \mathbb{P} 上的取值。假设 \mathbb{P} 的一个离散子集上是 $P = \{\Delta\theta, 2\Delta\theta, \dots, 360^\circ\} \times \{\Delta\theta, 2\Delta\theta, \dots, 360^\circ\}$ ，其中 $\Delta\theta$ 是一个很小的角度。

如前面描述，步态数据应该从视角 $\Delta\theta$ 到 360° 以 $\Delta\theta$ 的增幅采集。如果摄像机离人较远，从人体左侧拍摄到的轮廓跟从右边拍到的轮廓会非常相似，因此步态数据只需要从人的一侧拍摄则可。在随后实验使用的数据是 CASIA 步态数据库数据集 B，这个数据集中的数据是从角度 $\{0^\circ, \Delta\theta, 2\Delta\theta, \dots, 180^\circ\}$ 拍摄的，其中 $\Delta\theta = 18^\circ$ 。

在离散集 $\{0, 1, 2, \dots, \lfloor \frac{180}{\Delta\theta} \rfloor\} \times \{0, 1, 2, \dots, \lfloor \frac{180}{\Delta\theta} \rfloor\}$ 上的正确识别率可以通过实验的方式获得，离散式的正确率表达方式如公式 4.4 所示：

$$\begin{aligned} CCR = F(n, k) &= f(n \cdot \Delta\theta, k \cdot \Delta\theta) \\ n &= 0, 1, 2, \dots, \lfloor \frac{180}{\Delta\theta} \rfloor \\ k &= 0, 1, 2, \dots, \lfloor \frac{180}{\Delta\theta} \rfloor \end{aligned} \quad (4.4)$$

连续空间上的函数 $f(\theta_g, \theta_p)$ 的估计值 $\tilde{f}(\theta_g, \theta_p)$ 可以通过对 $F(n, k)$ 拟合和插值计算得到，令得到的 $\tilde{f}(\theta_g, \theta_p)$ 满足 $\tilde{f}(\theta_g, \theta_p) \approx f(\theta_g, \theta_p)$ 。这样就可以得到在任意角度 θ_g 和 θ_p 的正确识别率的估计值。在后面的小节中，将以步态能量图法 (GEI) 和关键傅立叶描述子法 (KFD) 为例，建立步态识别算法的评估模型。

4.2.2 实验结果

在实验中共采用了两种算法 (步态能量图法和关键傅立叶描述子法) 来验证步态识别评价模型。实验跟前面第 4.1.1 小节中所述的实验协议一致，且只采用来实验集 A 的数据来建立模型 (其他实验集有相似规律)，正常行走的前 4 个序列放入注册集，后 2 个序列放入测试集。这两种算法的实验结果列在表 4.1 和 4.4 中。

表格中的数据不是很直观，因此将表 4.1 和 4.4 中的数据表示在图 4.2 和 4.3 中。为了使实验更加充分，并计算了 rank=3 和 rank=5 的数据，绘制在图 4.2 和 4.3 中。图 4.2 和 4.3 中的横坐标为测试集的角度，纵坐标为正确识别率。

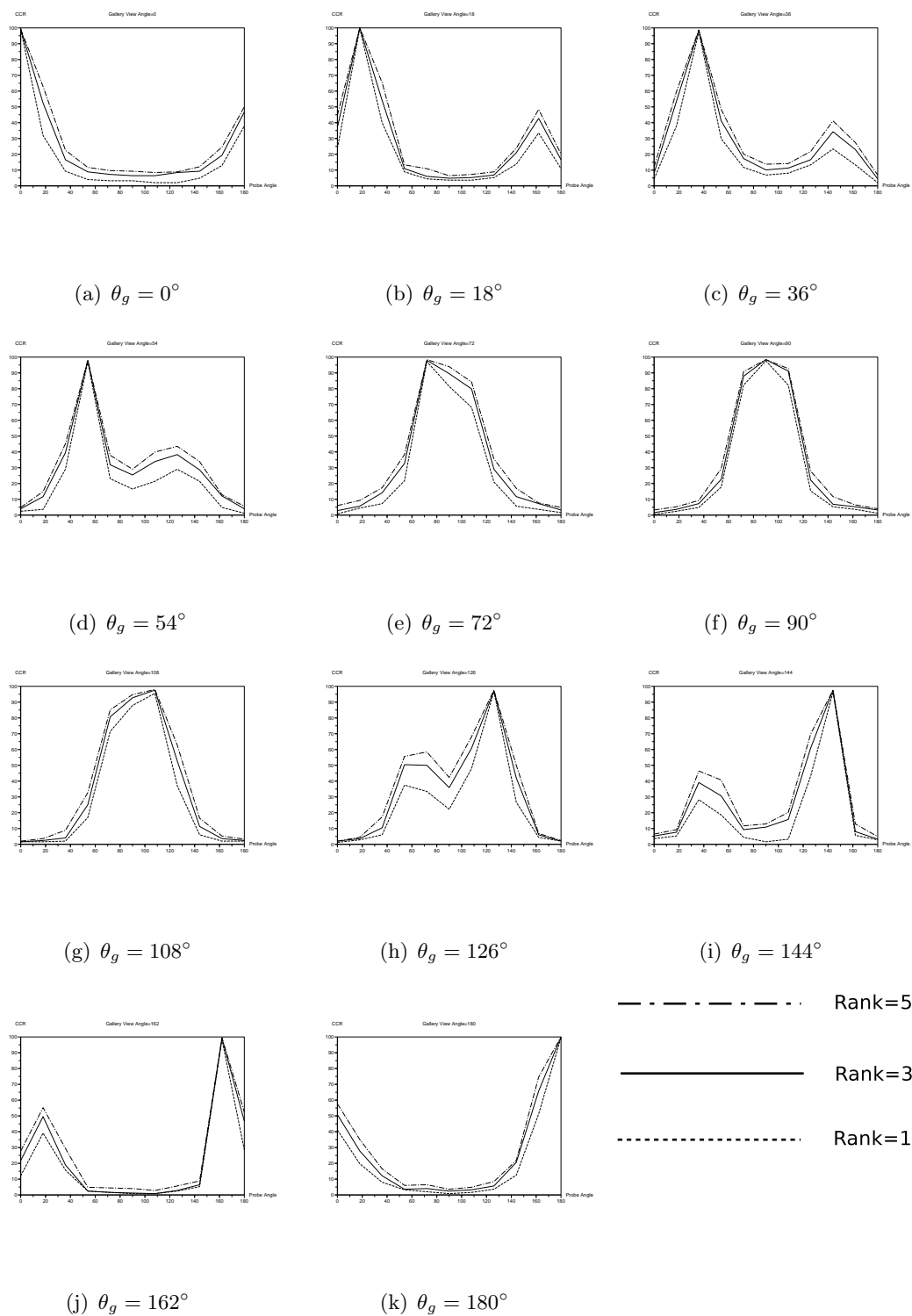


图 4.2: GEI 算法的正确识别率 (横坐标是 θ_p , 纵坐标是 CCR)

第四章 步态识别算法性能分析与评价

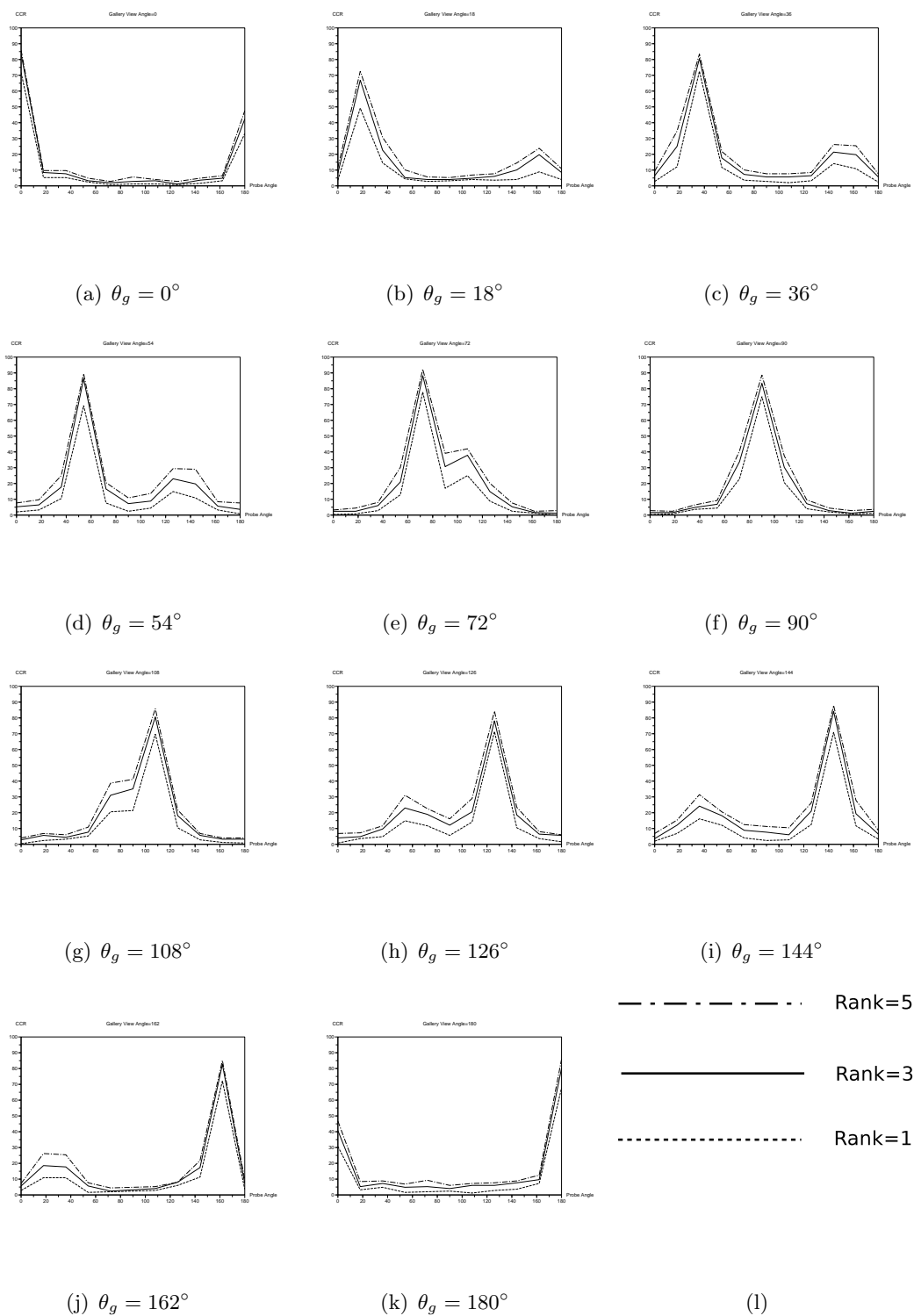


图 4.3: KFD 算法的正确识别率 (横坐标是 θ_p , 纵坐标是 CCR)

从图 4.2 和 4.3 中我们可以发现, 曲线存在两个峰值, 一个峰值出现在 $\theta_p = \theta_g$ 处, 另一个略低的峰值出现在 $\theta_p = 180^\circ - \theta_g$ 处。当 θ_p 取值在 90° 附近时 (图 4.2(f) 和 4.3(f)), 两个峰叠加在一起, 合二为一, 获得了一个顶端相对比较平坦的峰。这说明当注册集角度是 90° 时, 对视角变化具有最好的鲁棒性。另外, 还可以发现, 当 $\theta_g = \theta_p$ 的时候, 正确识别率基本保持一个恒定的数值, 可以将 $\theta_g = \theta_p$ 时的正确识别率记为 C_M ; 另外, 当 $\theta_p = 180^\circ - \theta_g$ 也基本保持恒定 (这些情况除外: $(\theta_g, \theta_p) \in \{(108^\circ, 72^\circ), (90^\circ, 90^\circ), (72^\circ, 108^\circ)\}$), 可以将 $\theta_p = 180^\circ - \theta_g$ 时的正确识别率记为 C_m 。

4.2.3 几何模型

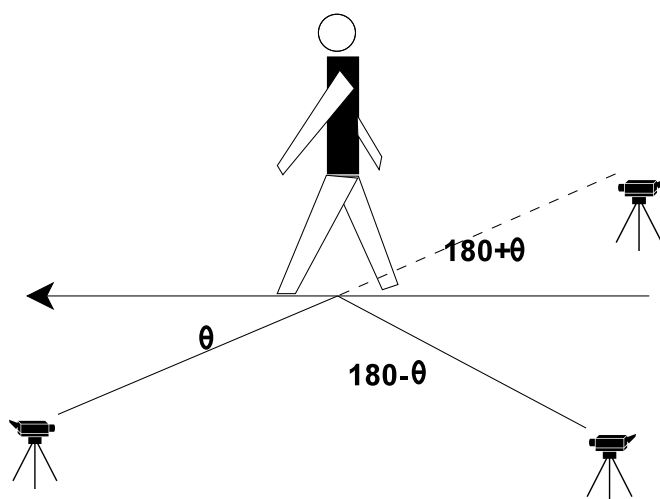


图 4.4: 几何模型示意图

上一小节中提到图 4.2 和 4.3 中的曲线存在两个峰, 这两个峰的出现应该与人体的左右对称性有关。假设从角度 θ , $180^\circ - \theta$ 和 $180^\circ + \theta$ 分别拍摄 3 幅图像, 如图 4.4 所示, 并用 $S(\theta)$, $S(180^\circ - \theta)$ 和 $S(180^\circ + \theta)$ 分别表示从这 3 个角度拍摄到的人体轮廓, 那么当摄像机离人体相对比较远的时候可以认为:

$$S(\theta) \approx \overrightarrow{S}(180^\circ + \theta) \quad (4.5)$$

其中符号 \rightarrow 表示图像沿水平方向翻转。另外, 由于人体是左右对称的, 所以有:

$$\overrightarrow{S}(180^\circ + \theta) \approx S(180^\circ - \theta) \quad (4.6)$$

当然，由于人体不可能严格对称，而且相机也不可能在无限远处，所以上面公式中使用了约等于符号 \approx 。

从前面两个公式 (4.5 和 4.6)，可以很容易的得到：

$$S(\theta) \approx S(180^\circ - \theta) \quad (4.7)$$

这说明了在角度 θ 拍摄到的轮廓与在角度 $180^\circ - \theta$ 拍摄到的轮廓非常相似，所以，当注册集的角度是 θ 的时候，用角度为 $180^\circ - \theta$ 的数据进行测试，可以得到相对较高的正确识别率，这也就是图 4.2 和 4.3 中曲线上两个峰中较小的峰出现的原因。

4.2.4 代数模型

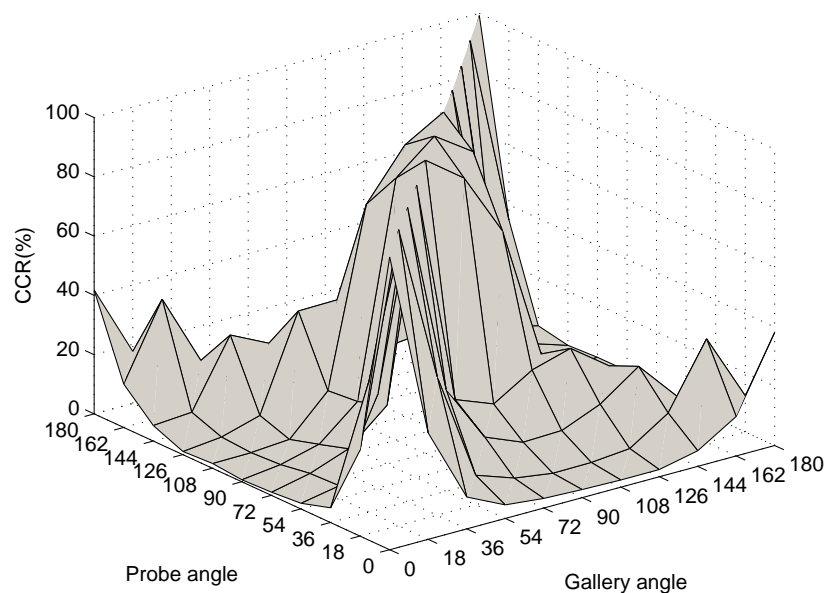


图 4.5: GEI 法的正确识别率曲面

将表 4.1 和 4.4 中的数据以三维曲面的方式绘制出来，如图 4.5 和 4.6 所示。希望能够得到一个简单的（参数少的）数学模型，通过这个模型，可以估计在任意角度 (θ_g, θ_p) 的正确识别率。

在图 4.5 和 4.6 存在两个岭，分别对应于图 4.2 和 4.3 中曲线上的两个峰，通过观察这两个岭的形状，会很直接的想到用一个混合高斯模型来描述图 4.5 和

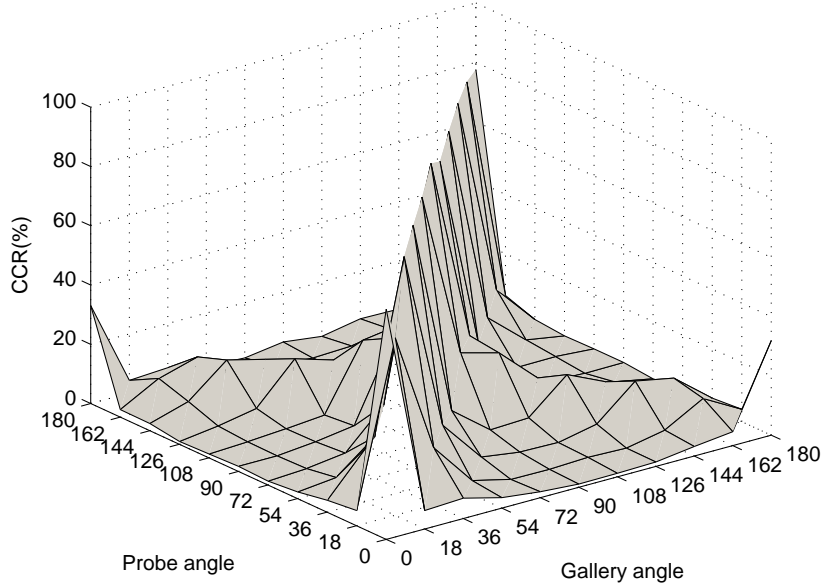


图 4.6: KFD 法的正确识别率曲面

4.6 中的正确识别率曲面。由于正确识别率 $CCR = f(\theta_g, \theta_p)$ 是一个 2 维函数，所以设计了下面的函数来描述：

$$\tilde{f}(\theta_g, \theta_p) = C_M e^{-\frac{(\theta_g - \theta_p)^2}{2\sigma^2}} + C_m e^{-\frac{(180^\circ - \theta_g - \theta_p)^2}{2\sigma^2}} \left[1 - e^{-\frac{(\theta_g - \theta_p)^2}{2\sigma^2}} \right] \quad (4.8)$$

其中 C_M 和 C_m 跟在第 4.2.2 小节中定义的一样， C_M 是无角度变化时的正确识别率，可以通过计算正确识别率表格中对角线上元素的平均值得到； C_m 是当注册集角度与测试集角度之和为 180° 时的正确识别率，可以通过计算正确识别率表格中反对角线上元素（ 72° ， 90° 和 108° 处的值除外）的平均值得到。 σ 是一个常数，可以表示当测试集的角度与注册集的角度不同时，步态识别性能降低的速度， σ 的值越大说明性能降低的越慢。针对前面的实验结果，使用 Matlab 曲线拟合工具箱（Curve Fitting Toolbox）[52] 可以计算得 σ 的取值是 15° 的时候可以有很小拟合误差。 $C_M e^{-\frac{(\theta_g - \theta_p)^2}{2\sigma^2}}$ 和 $C_m e^{-\frac{(180^\circ - \theta_g - \theta_p)^2}{2\sigma^2}}$ 是两个高斯函数，它们可以用来描述图 4.5 和 4.6 中的两个岭。 $1 - e^{-\frac{(\theta_g - \theta_p)^2}{2\sigma^2}}$ 是一个权重系数，用来保证 $\tilde{f}(\theta_g, \theta_p)$ 的值不超过 1。

通过公式 4.8 计算出的两种算法的正确识别率曲面如图 4.7 和 4.8 所示，从图中可以看出，公式 4.8 可以很好的估计正确识别率。所以这个代数模型可以很

好的刻画当角度变化时步态识别正确识别率的变化规律。公式 4.8 的另一个优点是只需要两个参数 C_M 和 C_m 则可估计在任意角度的正确识别率，也就是说只要计算出当无角度变化的时候的正确识别率 C_M 和注册集角度和测试集角度之和为 180° 时的正确识别率 C_m ，则可估计在任意角度的正确识别率。

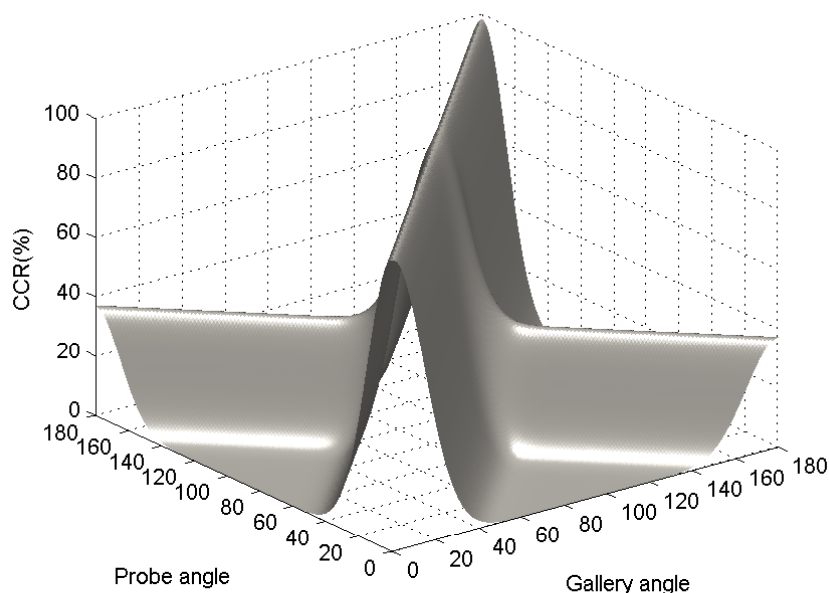


图 4.7: 通过模型计算出的 GEI 法的正确识别率曲面

4.3 本章小结

本章共分两部分，分别对步态识别算法的评价以及视角变化问题进行了分析。首先针对现在步态识别算法繁多但是评价标准欠缺的现状，提出了一个评价框架，这个框架包含一个大数据库、三组实验和一系列评价标准。这个框架可以评价一个步态算法对视角变化、衣着变化以及携带物品的鲁棒性，为步态识别算法提供评价标准。

本章的第二部分回答了步态识别中与视角相关的两个重要问题：① 哪个视角最适合用于步态识别？② 视角变化是如何影响识别率的？对于第一个问题，当视角没有变化的时候，各个角度的正确识别率无明显变化，但是侧面视角（ 90° ）特征对角度变化的鲁棒性最好所以侧面视角优于其他视角。对于视角

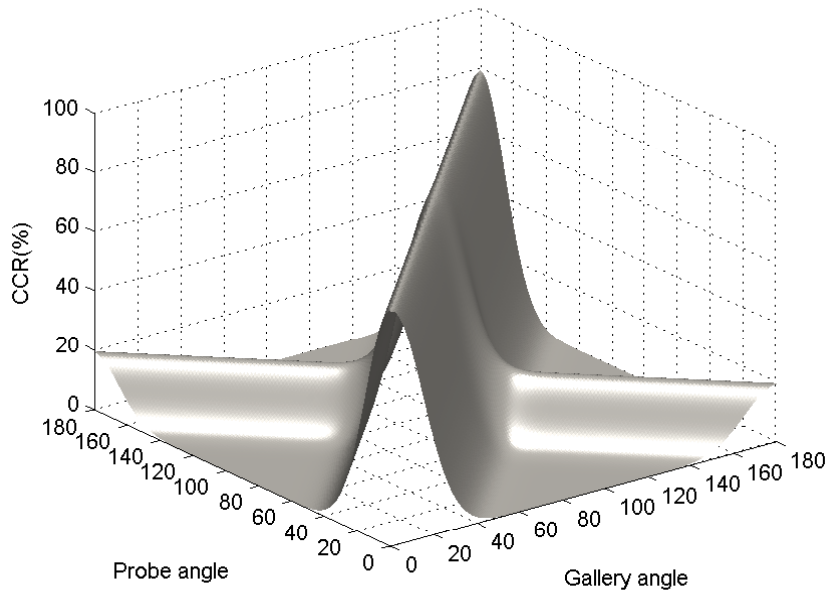


图 4.8: 通过模型计算出的 KFD 法的正确识别率曲面

跟识别率的关系，第 4.2.3 小节提出的几何模型给出了定性的分析；第 4.2.4 小节提出的代数模型给出了定量的分析，详细的描述了视角变化跟识别率的关系。

第五章 视角无关的步态识别

步态识别的一个重要应用是在智能视觉监控方面，在监控场景里面，我们不能期望场景里面的行人都按照一定的方向行走，所以无法获取固定视角的步态特征。如在图 5.1.(a) 中，人行走的方向跟拍摄方向不是垂直方向，无法获取最有效的侧面视角，而且对同一个人，在不同的场景里的摄像头对此人的拍摄方向也可能不相同。在图 5.1.(b) 中的情况则更有挑战性，里面人的行走路线为一圆弧曲线，而不是直线，在行走过程中视角不停的变化。

解决这个问题，需要设计视角无关的步态识别算法。设计视角无关的步态识别算法的核心问题是当测试集 (probe set) 中的数据拍摄角度跟注册集 (gallery set) 中的拍摄角度不一样的时候，如何衡量两个步态序列的相似性。

5.1 视角无关的步态识别框架

解决步态识别中的视角变化问题，根据注册集和测试集所用的视角数目，可以将解决方案分为三种 (表 5.1)。第一种解决方案为采集注册数据和测试数据的时候，都采用多个摄像头从多个视角采集。论文 [3] 就是采用这种方法。在拍摄数据的时候，采用视角越多，获得的步态数据就越完善，人的行走方向的变化引起的影响也就越小。随着摄像机数目的增加，步态的识别率会增高，但是系统配置以及计算的复杂度也会随之增加，实用性会变差。

第二种方案简化了测试集的采集，只采用一个视角。在这种情况下，采集注册数据的时候，采用环绕人体的多摄像机采集。在识别阶段，先估计出测试步态序列的行走方向，然后跟注册集中行走方向最相近的步态序列进行比较。这种方案比方案一实用性更强，但是在采集注册数据时仍需要一个专门的多摄

表 5.1: 几种视角无关的步态识别方案

	注册数据视角数	测试数据视角数
方案一	N	N
方案二	N	1
方案三	1	1



图 5.1: 监控系统中常见的行走情况

像机环境。

目前监控摄像头被广泛的安装在各种场合，比如办公楼、交通路口等地方，这些场景的摄像头拍摄区域大部分都没有重合区域，无法同时从多个视角对某一个人进行拍摄。所以一种更实用的方案（方案三）是无论注册数据，还是测试数据，都只从一个视角采集。这种方案的优点是采集数据无需特定的采集环境，可以采用现在已被广泛安装的摄像头。论文 [4, 5] 采用的都是这种方案。这种方案对硬件环境的要求比较低，实用性强，易于推广使用，

在前面提到的三种方案中，方案三具有最强的实用价值，但是对步态识别算法也提出了比较大的挑战。方案三需要解决的问题是，当注册数据的拍摄视角跟测试数据的视角不一致的时候，如何计算这两个数据的相似性。本文提出了一个针对方案三的视角无关的步态识别框架（图 5.2）。这个框架分三个阶段：训练阶段，注册阶段和识别阶段。在训练阶段，使用的数据为多个人多个视角的特征，然后根据这些特征，训练出一个不针对特定人的通用的视角无关模型。注册阶段需要采集行人的某一个角度的特征，放入注册集中。虽然可以注册任一个角度，但是选择适合识别的角度将增加算法的鲁棒性和提高算法的识别率。实验分析已经证明侧面视角（行走方向跟图像平面平行）是最优选择（详见第 4.2 节），所以注册时建议采用侧面视角。在识别阶段，可以使用在训练阶段获得的视角无关模型，将任一角度的步态测试特征合成为注册角度的特征，然后利用合成的特征跟注册集中的数据进行相似性度量，确定出待识别人的身份。

本框架的优点是训练数据跟注册数据无关，训练阶段获得的模型是对所有

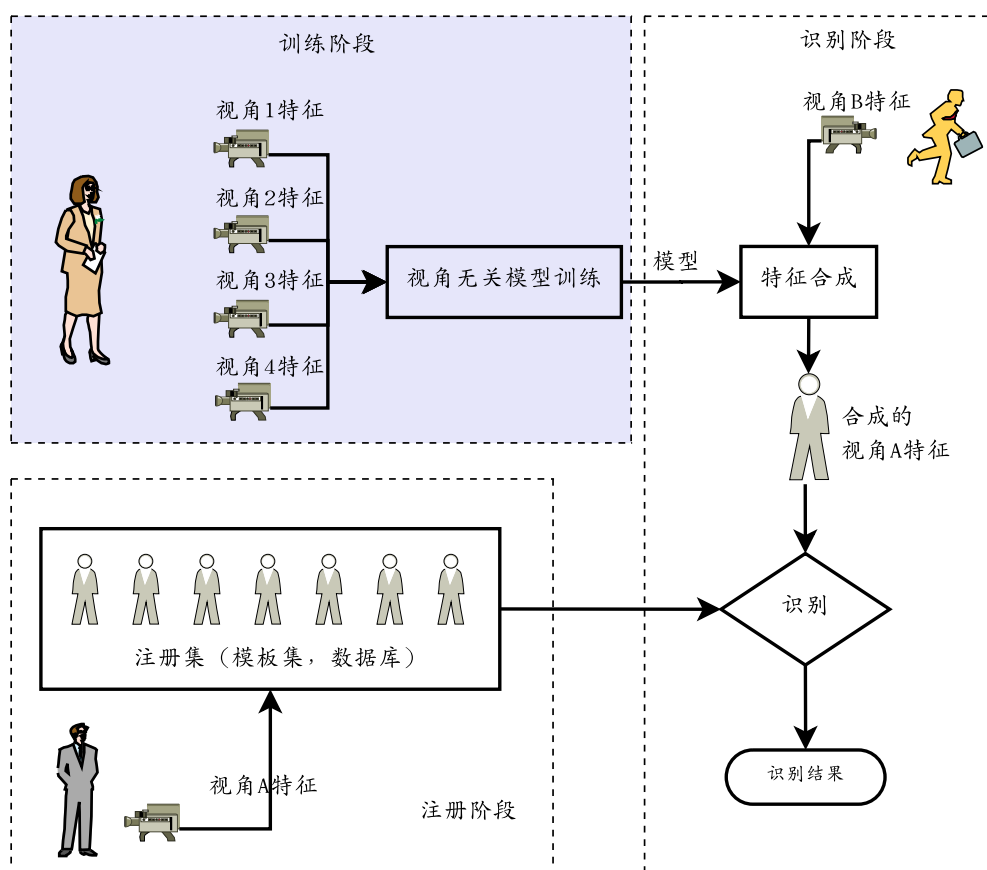


图 5.2: 视角无关的步态识别框架

人都适用的模型，模型可以推广到任何人；而且在注册和识别阶段，只需要一个视角，便于在监控系统中应用。

5.2 特征合成方法

步态特征作为一种行为特征，容易受到视角和衣着等因素的影响，跟人的生理特征相比，步态特征具有较大的类内距离与类间距离比值。一个人的不同视角的步态特征在特征空间内往往不会聚集在一起，而是不同的视角特征分布在空间中的不同位置。如图 5.3 所示，在特征空间 A 中，Person A 的两个视角的数据（view 1, view 2）并不聚集在一起。另一人 Person B 的数据也是如此。这样，类内数据的分散分布将会造成分类困难。

根据摄像机透视成像原理，在获取视角 1 的步态数据的时候，不同的人的成像参数（摄像机内参和外参）是一样的，同样在获取视角 2 的步态数据的时候

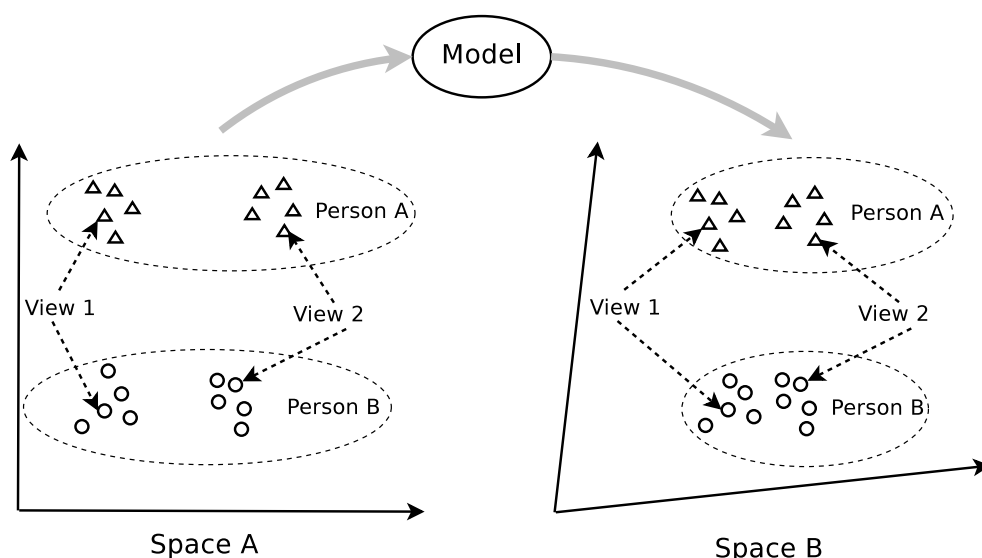


图 5.3: 多个视角的特征在不同特征空间表示

候, 不同的人的成像参数也是一样的。因此, 不同人的视角 1 数据跟视角 2 数据之间应该有相似的规律。如果能够发掘出这个规律, 那么就可以依据这个规律, 将同一人的不同视角的特征, 聚集在一起。即如图 5.3 所示, 通过这个规律, 建立起一个模型, 将特征空间 A 中的数据投影到另外的一个空间 B, 使得 Person A 的不同视角的数据聚集在一起, Person B 的不同视角的数据也聚集在一起。这样可以大大降低类内散度, 提高特征的区分能力。

5.3 线性模型

5.3.1 原理

假如对某个行人从角度 α 获取的特征向量为 \mathbf{f}_α , 从角度 β 获取的特征向量为 \mathbf{f}_β , 那么可以采用线性映射的方法, 将特征 \mathbf{f}_α 投影到另一个空间, 使得投影后的特征跟 \mathbf{f}_β 非常接近。线性投影模型 $\mathbf{X}_{\alpha,\beta}$ 的求解可以描述为:

$$\mathbf{X}_{\alpha,\beta}^* = \arg \min_{\mathbf{X}_{\alpha,\beta}} \|\mathbf{f}_\beta - \mathbf{X}_{\alpha,\beta} \cdot \mathbf{f}_\alpha\| \quad (5.1)$$

假如有 n 对分别从角度 α 和角度 β 的步态特征向量 $(\mathbf{f}_{\alpha,1}, \mathbf{f}_{\beta,1}), (\mathbf{f}_{\alpha,2}, \mathbf{f}_{\beta,2}),$

..., $(\mathbf{f}_{\alpha,n}, \mathbf{f}_{\beta,n})$, 那么可以从下面方程组中求解 $\mathbf{X}_{\alpha,\beta}$:

$$\begin{cases} \mathbf{f}_{\beta,1} = \mathbf{X}_{\alpha,\beta} \cdot \mathbf{f}_{\alpha,1} \\ \mathbf{f}_{\beta,2} = \mathbf{X}_{\alpha,\beta} \cdot \mathbf{f}_{\alpha,2} \\ \dots \\ \mathbf{f}_{\beta,n} = \mathbf{X}_{\alpha,\beta} \cdot \mathbf{f}_{\alpha,n} \end{cases} \quad (5.2)$$

这样模型 $\mathbf{X}_{\alpha,\beta}$ 的求解就变成了一个线性方程组的求解, 可以使用最小二乘法容易的求得 $\mathbf{X}_{\alpha,\beta}$ 。

求得 $\mathbf{X}_{\alpha,\beta}$ 后, 可以应用于识别阶段。假如注册的时候, 拍摄角度是 β , 在识别阶段, 需要从角度 α 拍摄到的步态视频中识别出视频中人的身份。那么先对从角度 α 拍摄到的视频中提取步态特征 \mathbf{f}_{α} , 然后使用投影模型 $\mathbf{X}_{\alpha,\beta}$ 将 \mathbf{f}_{α} 投影到另外一个空间, 得到 \mathbf{f}'_{β} 。其中 \mathbf{f}'_{β} 可以看作是从 \mathbf{f}_{α} 合成的在 β 角度的特征。

5.3.2 步态特征

目前已经有多种步态特征被用于身份识别, 有基于模型的特征, 也有基于整体轮廓的特征。由于基于模型的特征获取相对困难, 且稳定性较差, 所以在实验中采用了基于整体轮廓的特征。步态能量图特征 (第 2.1.3.3 节) 是一种基于轮廓整体的特征, 这种特征对噪声影响具有较好的鲁棒性, 且具有比较好的区分能力, 目前被广泛使用。

假如步态特征的特征向量维数为 m , 那么在方程 5.2 中需要求解的变量个数 ($\mathbf{X}_{\alpha,\beta}$ 的维数) 为 m^2 , 约束 (方程) 的个数为 mn , 要使得该方程有解, 至少需要满足:

$$m^2 \leq mn$$

也就是需要满足:

$$m \leq n$$

也就是特征的维数不能大于样本的数目, 否则方程 5.2 无解。

在实验中使用的数据库为 CASIA 步态数据库中的数据集 B (详见第三章)。该数据库中共有 124 人, 每个人每个视角正常行走的步态序列数为 6, 所以特征的维数不能大于 $124 \times 6 = 744$ 。步态能量图是一个二维的图像, 如果维数不大于 744, 那么图像最大只能是 27×27 。如果使用这么小的步态能量图, 那

么它的区分能力会大大降低。为了解决特征维数太高的问题，主成分分析法被用来对特征进行降维。

5.3.3 主成分分析法降维

主成分分析法 [38] (principal component analysis, PCA) 是一种将高维空间的数据转换到低维空间的方法，目前在特征提取中被广泛使用。PCA 寻找一个线性投影，将数据投影到一个新的坐标系统中。在这个新的坐标系统中，第一个坐标表示最大的变化 (第一主成分)，第二个坐标表示次大的变化 (第二主成分)，以此类推。低阶的成分被认为是更能代表原始数据的部分，高阶的成分往往会被认为是噪声。通过保留低阶成分，抛弃高阶成分，可以达到降维的目的。

假设原始数据是 $(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)$ ，其中 n 是样本的数目， \mathbf{y}_i 是一个长度为 M 的向量。如果我们希望将特征向量的维数从 M 降低到 m ，将 \mathbf{y}_i 的降维为 \mathbf{f}_i ，可以按照下面的方法实现。

首先计算原始样本集的均值 \mathbf{u} 和协方差矩阵 \mathbf{C} ：

$$\mathbf{u} = \frac{1}{n} \sum_{i=1}^n \mathbf{y}_i \quad (5.3)$$

$$\mathbf{C} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{u})(\mathbf{y}_i - \mathbf{u})^T \quad (5.4)$$

如果矩阵 \mathbf{C} 的秩为 L ，那么可以通过奇异值分解 (singular value decomposition, SVD) 计算出 \mathbf{C} 的 L 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_L$ 以及与特征值对应的特征向量 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_L$ 。

通常来说，特征值 λ_i 越大，对应的原始数据的变化也越大。对特征值进行排序，使得 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L$ ，保留前 m 个最大的特征值对应的特征向量，那么原始数据 \mathbf{y}_i 可以通过下面的公式转换为 \mathbf{f}_i ：

$$\mathbf{f}_i = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m]^T \mathbf{y}_i \quad (5.5)$$

这样特征维数就从 M 维降低到 m 维。前 m 个成分更能够表示原始数据，它们所占的比重为：

$$\rho = \sum_{i=1}^m \lambda_i / \sum_{j=1}^L \lambda_j \quad (5.6)$$

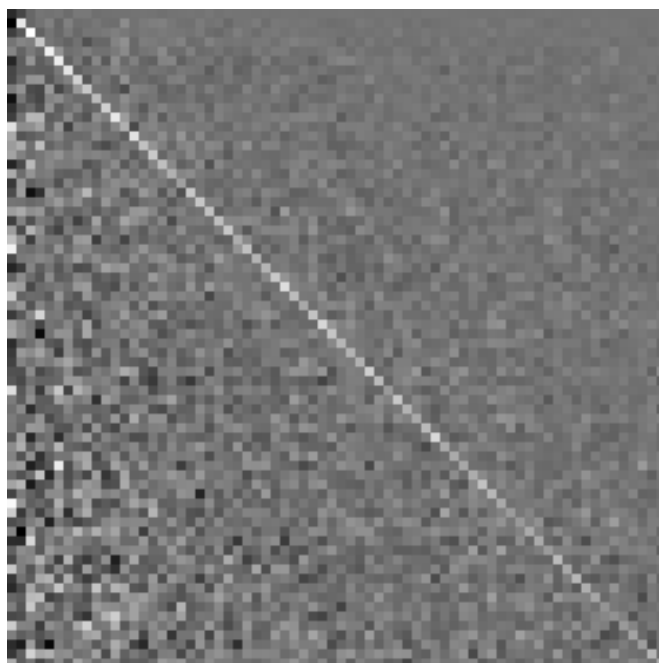


图 5.4: 将 54° 视角数据映射到 90° 的矩阵

5.3.4 实验与分析

在实验中, 采用了 CASIA 步态数据库中的数据集 B 对算法进行评价。首先将提取出的步态能量图特征用 PCA 进行降维, 将特征降低至 70 维。在本实验中, 前 70 个主成分可以表达大约 95% 的原始数据的变化。

因为侧面视角 (数据库中的 90° 视角) 跟其他视角相比, 最适合进行步态识别 (详见第 4.2 节)。首先将侧面视角作为注册视角, 即注册集 (模板集) 中的所有数据都是侧面视角采集的数据。在识别阶段, 当未知身份的角度为 θ ($\theta \neq 90^\circ$) 时, 使用将 θ 角度特征转为 90° 特征的投影矩阵 $\mathbf{X}_{\theta,90}$, 合成 90° 特征:

$$\mathbf{f}'_{90} = \mathbf{X}_{\theta,90} \cdot \mathbf{f}_\theta \quad (5.7)$$

然后使用合成的特征 \mathbf{f}'_{90} 进行识别。由于此处的主要目的是比较合成特征的有效性, 所以采用了简单的最近邻分类器来识别身份。

在实验中, 为了保证训练过程不引入测试数据的信息, 在计算投影矩阵 $\mathbf{X}_{\theta,90}$ 过程中不使用被测试人的数据, 即采用留一法。假若要识别第 i 人, 那么将除第 i 人之外的 123 人的数据放入训练集, 训练得到投影矩阵, 然后使用投影矩

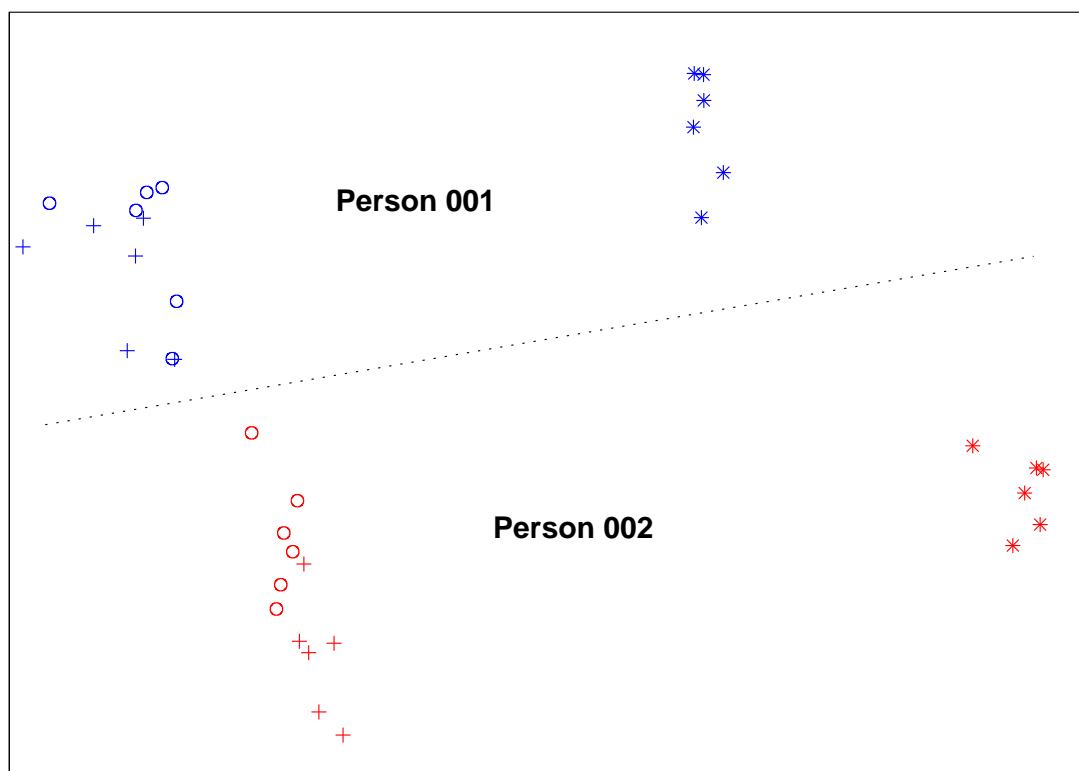


图 5.5: 两人的第 1 个和第 2 个 PCA 系数 (* 为 54° 特征; \circ 为 90° 特征; + 为利用 54° 特征合成的 90° 特征)

阵合成第 i 人的 90° 视角特征, 进行识别。由于投影矩阵的计算是一个线性方程组的求解, 可以比较容易的计算出。图 5.4 以图像的方式显示了将 54° 视角数据映射到 90° 的矩阵, 这个矩阵的大小为 70×70 , 图中越亮的点表示该位置的值越大, 越暗的点表示该位置的值越小。用这个投影矩阵可以将 54° 特征合成的 90° 特征。图 5.5 中显示了数据库中 ID 为 001 和 002 的两个人的 54° 特征, 90° 特征以及利用 54° 特征合成为 90° 特征。从图中我们可以很清楚的看出同一个人的合成的 90° 特征跟原始的 90° 特征混合在一起, 这样非常有利于分类。

当注册集角度为 90° 时, 测试集角度分别为 $0^\circ, 18^\circ, \dots, 180^\circ$ 时 (横坐标) 的正确识别率如图 5.6 所示。图中 Non 表示不使用合成的特征, 只使用从步态能量图中提取的 70 个主成分系数的识别结果; Linear 表示使用线性投影模型合成的 90° 特征的识别结果。从图中我们可以看出, 当测试角度跟注册角度差别 36° 时, 也就是测试角度为 54° 或者 129° 时, 如果不使用合成特征, 识别率小于 50%; 如果使用合成的特征, 识别率可以提高到 75% 左右。除了这两个角度, 在

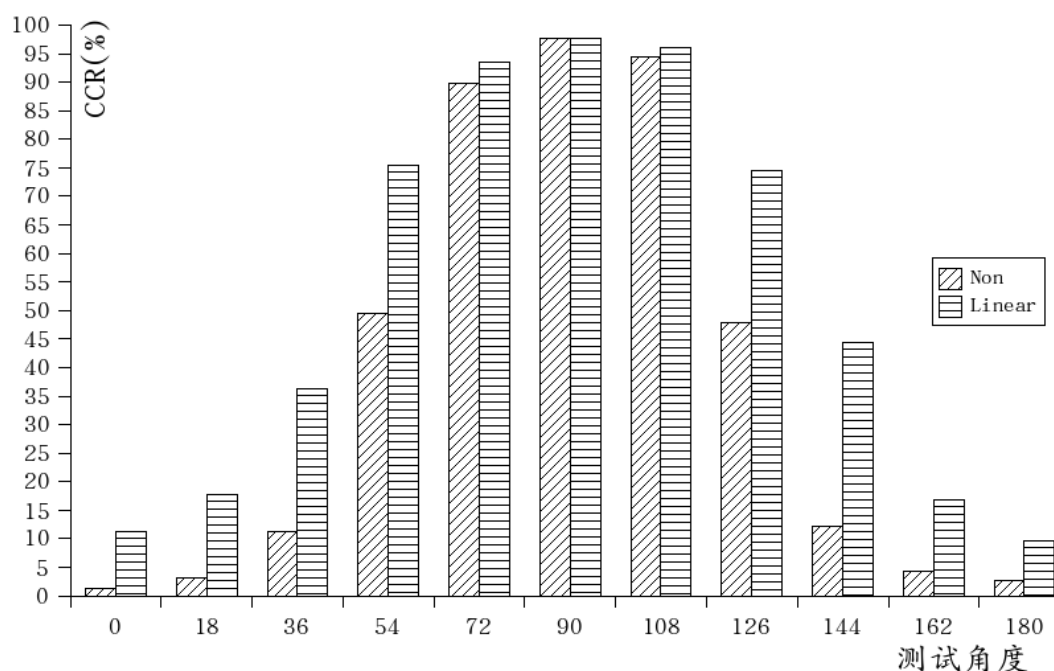


图 5.6: 注册集角度为 90° , 测试集角度分别为 $0^\circ, 18^\circ, \dots, 180^\circ$ 时的识别率

其他角度, 使用合成特征也可以大幅度的提高识别率。

在本文第四章中, 设计了一系列实验, 并提出了评价标准来评价一个算法对视角变化的鲁棒性。前面提出的用线性模型合成步态特征的方法, 对视角变化的鲁棒如何, 应该也可以用这个标准进行评价。要评价算法对视角的鲁棒性, 首先要先进行实验集 A 中的 121 个实验。使用步态能量图主成分分析系数 (GEI+PCA) 算法的实验结果列在表 5.2 中, 对 GEI+PCA 特征进行线性合成方法的实验的实验结果列在表 5.3 中。

在第四章的公式 4.2 定义的 σ_i^A 参数可以用来评价视角变化时, 步态识别算法的鲁棒性。 σ_i^A 中的 i 表示测试数据角度与注册数据角度夹角为 $18 \cdot i^\circ$ 时的性能。 σ_i^A 的值越高表示视角变化对步态识别率的影响越小。表 5.4 中列出了几种算法的 σ_i^A 值。表中 GEI 表示所用步态特征是步态能量图, GEI+PCA 表示特征是步态能量图的 PCA 系数, GEI+PCA+Linear 表示训练特征是原始的步态能量图的 PCA 系数, 但测试特征是用线性投影模型合成的特征。

从表 5.4 中可以看出, GEI+PCA+Linear 一系列中的所有 σ_i^A 值均大于其他方法的 σ_i^A 值。因此使用线性投影模型将测试角度的特征合成为注册角度的特征,

表 5.2: 不使用合成特征的实验结果 (实验集 A)

CCR (%)	测试集角度 θ_p (正常行走#5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	99.2	77.0	23.0	4.4	3.6	1.2	2.4	1.2	7.3	23.4	43.1
	18°	81.5	99.6	87.5	31.0	15.3	9.7	6.9	8.5	18.5	43.1	20.6
	36°	35.1	84.7	97.6	82.7	41.9	24.6	20.2	32.3	36.7	23.8	9.7
	54°	11.3	33.5	81.9	97.2	77.4	56.5	56.5	53.6	40.3	13.3	3.2
	72°	5.6	11.7	35.5	73.0	97.2	93.5	78.6	54.0	21.8	6.0	3.2
	90°	1.2	3.2	11.3	49.6	89.9	97.6	94.4	48.0	12.1	4.4	2.8
	108°	2.8	2.8	14.9	50.0	71.8	92.7	95.6	81.9	46.4	8.9	2.8
	126°	4.4	10.5	31.5	52.0	64.9	57.7	83.1	96.8	87.1	29.0	9.7
	144°	10.5	19.4	39.9	37.5	25.8	12.5	33.9	89.1	96.4	40.3	20.2
	162°	26.2	43.1	32.3	10.5	6.5	5.6	8.9	14.5	27.4	98.4	71.4
	180°	45.6	24.2	8.1	2.0	1.2	2.0	2.0	4.4	13.7	73.0	99.6

表 5.3: 线性投影合成特征的实验结果 (实验集 A)

CCR (%)	测试集角度 θ_p (正常行走#5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	99.2	78.6	42.3	24.2	8.5	7.3	8.5	18.1	35.9	59.7	86.7
	18°	88.7	99.6	92.3	60.9	25.8	16.1	16.5	20.2	29.0	37.9	60.1
	36°	49.2	94.0	97.6	89.9	57.7	31.5	25.4	23.8	24.2	23.0	25.8
	54°	17.3	67.3	88.3	97.2	89.5	67.7	52.4	38.7	26.6	11.7	12.5
	72°	9.3	29.8	61.3	92.7	97.2	93.5	81.5	61.3	34.3	12.9	8.1
	90°	11.3	17.7	36.3	75.4	93.5	97.6	96.0	74.6	44.4	16.9	9.7
	108°	12.9	10.9	24.2	54.4	80.6	91.1	95.6	94.8	69.4	25.8	9.3
	126°	13.7	11.7	23.8	44.8	53.6	70.2	92.3	96.8	93.5	60.5	17.7
	144°	24.2	21.4	21.0	28.2	29.4	41.1	68.5	91.9	96.4	90.3	33.5
	162°	57.3	39.1	21.8	15.7	10.5	19.4	27.8	64.5	89.1	98.4	74.6
	180°	85.5	56.9	33.1	14.9	8.9	8.1	9.3	20.6	35.1	77.0	99.6

表 5.4: 三种不同算法对视角鲁棒性的评价参数

i	σ_i^A		
	GEI	GEI+PCA	GEI+PCA+Linear
1	0.418	0.803	0.917
2	0.164	0.409	0.637
3	0.086	0.230	0.361
4	0.076	0.165	0.222
5	0.064	0.146	0.162
6	0.076	0.132	0.146
7	0.071	0.134	0.197
8	0.155	0.208	0.334
9	0.142	0.241	0.598
10	0.404	0.454	0.881

可以提高算法对视角变化的鲁棒性。另外还可以从表 5.4 中发现, GEI+PCA 列中的数值均高于 GEI 列中的数值, 所以对步态能量图特征进行 PCA 压缩, 也可以提高算法对视角变化的鲁棒性。

5.4 非线性模型

人体是一个非常复杂的形体, 人体可以分成几个关键的部分, 如头、躯干、上臂、小臂、大腿、小腿等, 这些关键部分通过关节连接起来, 肢体的这些主要部分在人运动时都有很高的自由度, 而且某些部分, 比如躯干, 还可以弯曲或者扭曲。如此高度复杂的形体, 从理论上讲, 无法用一个线性模型将一个视角拍摄到的特征合成为另一个视角的特征。使用非线性模型对步态特征进行合成, 应该更能发掘出不同视角的步态特征之间的联系。

有很多种非线性的数据拟合算法可以用于合成步态特征, 例如多项式法 [53], 支持向量机 [54], 神经网络 [55, 56] 等。对于多项式法, 虽然可以转化为线性问题进行求解, 但是由于输入向量是高维向量, 而步态识别中训练样本却是有限的, 往往无法进行求解。支持向量机的一个缺点是输出只能是一个数值, 不能是一个向量, 如果需要拟合数据到一个向量, 需要多个支持向量机并行运作, 计算复杂度很高。神经网络可以重复使用输入数据, 可以部分的解决样本数量

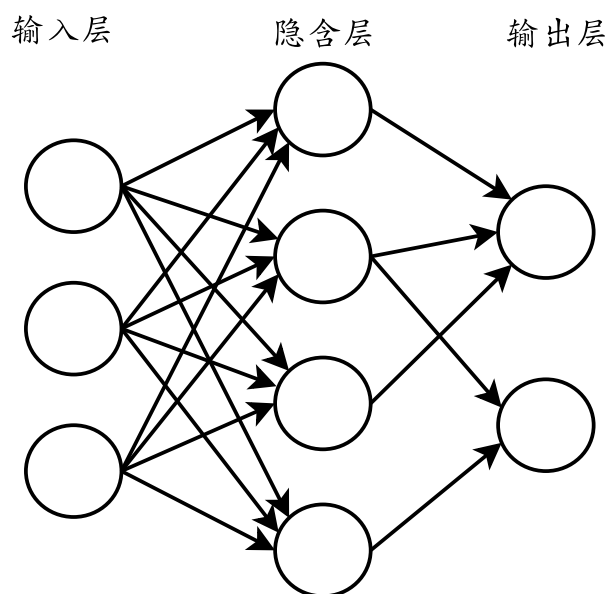


图 5.7: 多层前馈神经网络模型

不足问题，所以选择了神经网络来将一个视角的特征映射到另一个视角。

5.4.1 神经网络

多层前馈神经网络 (multi-layer perceptrons, MLP) 目前是最常用的一种神经网络类型。多层前馈神经网络包含一个输入层，一个输出层，以及几个隐含层。每一层都包含一个或者几个神经元，这些神经元跟上一层和下一层的神经元连接。图 5.7 表示了一个三层模型，这个神经网络模型包含 3 个输入神经元，2 个输出神经元，以及 4 个隐含神经元。

多层前馈神经网络中的所有神经元都是相似的，每个神经元都有几个输入和一个输出。从上层传来的值，经过加权求和输入神经元，神经元使用一个激活函数将输入神经元的数值转化为另外一个数值。也就是说，假如第 n 层的输出为 (x_1, x_2, \dots, x_J) ，第 $n+1$ 层的输出可以按照下面的公式计算：

$$u_i = \sum_{j=1}^J (w_{i,j}^{n+1} x_j) + w_{i,bias}^{n+1} \quad (5.8)$$

$$y_i = f(u_i)$$

式中 $w_{i,j}$ 为权重， f 为激活函数。激活函数有多种，在本实验中使用了 sigmoid

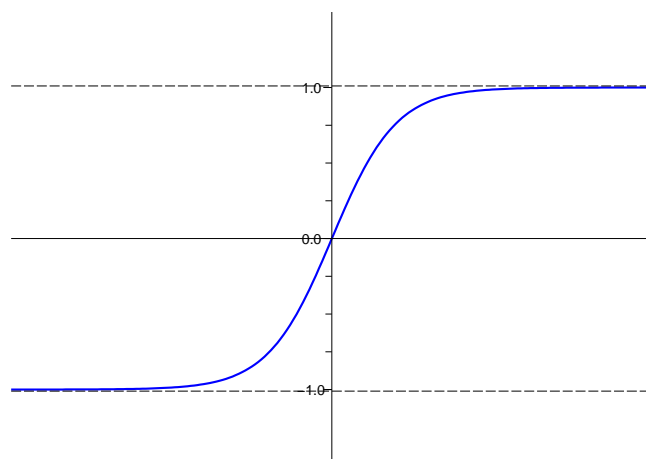


图 5.8: sigmoid 激活函数

函数，sigmoid 激活函数定义如下：

$$f(x) = \beta \frac{1 - e^{-\alpha x}}{1 + e^{-\alpha x}} \quad (5.9)$$

当 $\alpha = 1$, $\beta = 1$ 时，sigmoid 函数曲线如图 5.8 所示。

神经网络的工作原理如下面所述。首先神经网络将特征向量作为输入，输入层里的神经元个数等于特征向量的维数。然后这些输入传递到第一个隐含层，隐含层的输出值通过式 5.8 计算得到，然后传递到下一层，直到到达输出层。因此可以看出要使用神经网络，关键是先得到所有的权重 $w_{i,j}$ 的值。权重 $w_{i,j}$ 的值计算可以通过使用训练数据，不停的迭代来调整 $w_{i,j}$ 的值，直到输出误差足够小时停止迭代。多层前馈神经网络有很多种算法，本文中使用了 [56] 中的 RPROP 算法来实现。

5.4.2 实验和分析

实验采用的特征跟前面线性模型中的一样，为步态能量图的前 70 个 PCA 系数。当训练从 α 角度到 β 角度的合成模型时，神经网络的输入为 α 角度的特征，输出为 β 角度的特征，输入层和输出层的神经元个数跟特征的维数相等，都

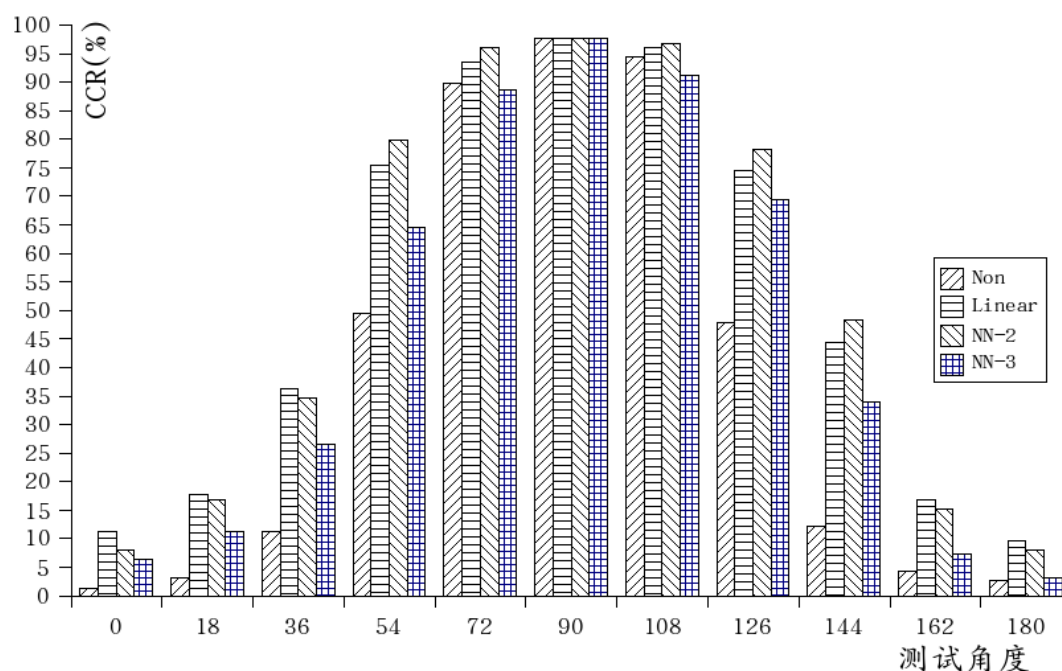


图 5.9: 注册集角度为 90°, 测试集角度分别为 0°, 18°, ..., 180° 时的识别率

为 70。实验中共采用了两种神经网络模型，一种是含有一个隐含层（70 个神经元）的三层模型，另一种是没有隐含层的二层模型。

当注册步态数据的角度为 90°, 测试数据的角度分别为 0°, 18°, ..., 180° 时, 各种算法的识别率结果如图 5.9 所示。图中 Non 表示不使用合成的特征, 只使用从步态能量图中提取的 70 个 PCA 系数的识别结果; Linear 表示使用线性投影模型的识别结果; NN-2 表示使用二层神经网络模型的识别结果; NN-3 表示使用三层神经网络模型的识别结果。从图中可以看出, 三层神经网络可以提高识别率, 但是提高的效果没有线性模型和二层神经网络明显。在视角变化比较小的时候, 也就是测试角度为 54°, 72°, 108°, 126°, 144° 时, 二层神经网络优于线性模型, 但是在角度变化比较大的时候, 也就是测试角度为 0°, 18°, 36°, 162°, 180° 时, 线性模型对于性能的提高要比二层神经网络模型高。在理论上, 三层神经网络比二层神经网络有着更强的表达能力, 但是在实验中并没有获得更好的结果。三层神经网络虽然有着更强的表达能力, 但是需要有更多的样本来训练这个网络, 以使之避免受噪声影响以及避免产生过学习现象。但在本实验中, 样本的数量是相对有限的, 这些有限的样本难以训练出一个复杂的三层神经网络模型。

表 5.5: 二层神经网络合成特征的实验结果 (实验集 A)

CCR (%)	测试集角度 θ_p (正常行走#5-6)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_g (正常行走#1-4)	0°	99.2	76.6	42.7	24.2	8.9	4.8	6.5	19.4	38.7	58.1	85.5
	18°	91.9	99.6	93.5	69.3	24.2	15.3	18.6	19.4	26.6	35.5	56.5
	36°	47.6	94.3	97.6	95.2	59.7	31.4	25.8	21.0	23.4	18.6	26.6
	54°	15.3	71.8	91.9	97.2	91.9	72.6	51.6	36.3	26.6	8.9	15.3
	72°	7.3	30.6	65.3	96.0	97.2	97.6	82.3	62.9	34.7	13.7	8.9
	90°	8.1	16.9	34.7	79.8	96.0	97.6	96.8	78.2	48.4	15.3	8.1
	108°	12.9	9.7	25.8	54.8	79.8	94.3	95.6	95.2	74.2	27.4	8.1
	126°	16.1	12.1	22.6	46.0	50.8	71.0	96.0	96.8	93.5	66.1	16.1
	144°	27.4	21.8	18.6	29.0	29.0	41.9	71.8	91.9	96.4	90.3	33.9
	162°	55.6	46.0	21.0	16.1	12.1	17.7	27.4	63.7	88.7	98.4	74.2
	180°	87.1	56.5	34.7	18.6	8.9	9.7	9.7	16.1	30.6	75.8	99.6

表 5.6: 线性模型与二层神经网络模型在视角变化时性能评价标准

i	σ_i^A	
	GEI+PCA+Linear	GEI+PCA+NN-2
1	0.917	0.932
2	0.637	0.660
3	0.361	0.357
4	0.222	0.217
5	0.162	0.157
6	0.146	0.139
7	0.197	0.201
8	0.334	0.356
9	0.598	0.580
10	0.881	0.883

为了更完整的评价算法对视角变化的鲁棒性,选择了两种神经网络中效果好的二层神经网络,进行了第四章中设计的 121 个实验,实验结果列在表 5.5 中。由这个实验得到的评价标准 σ_i^A 列在表 5.6 中。从表中可以看出,当测试特征跟注册特征差别较小时,也就是测试角度跟注册角度差别在不超过 36° ,或者是 180° 时,二层神经网络的评价标准 $\sigma_i^A, i = 1, 2, 10$ 的值都大于线性模型的评价标准值,此时二层神经网络有更好的效果。但是当视角差别比较大的时候 ($2 < i < 10$ 时),线性模型有更好的效果。

因为线性模型可以视作非线性模型的一种特例,所以非线性的神经网络模型与线性模型相比,有更强的表达能力,因此当角度变化比较小的时候,使用线性神经网络模型合成的特征跟真实的特征更相似,也就是算法对视角变化的鲁棒性也更高。但是当角度差别比较大时,来自两个视角的训练数据之间的关系变弱,神经网络对于这种微弱关系的挖掘能力不如线性模型高。

5.5 结论与展望

本章中提出的视角无关的步态识别框架,在注册和识别阶段只需要一个视角的数据,具有比较高的实际应用价值。另外,基于线性模型和非线性的神经网络模型,可以利用一个视角的特征有效的合成另一视角的特征,合成的特征可以获得令人满意的识别率。而且该模型可以用来表示不在训练库中的人的步态,对于新进的样本具有比较好的适应能力。

另外,虽然目前视角无关的步态识别研究取得一定的进展,但是尚不够深入,将来可以从如下几个方面进行视角无关的步态识别方法研究:

- 视角无关特征的提取。例如可以融合三维和二维来提取特征,采用三维的方法建立模型,与二维特征进行互补,进行视角无关特征提取。
- 研究视角变化时步态特征在特征空间里的变化规律。当视角变化时,人的步态特征会在特征空间内以某种规律变化,可以使用流形等方法对这种规律进行分析。发掘出这个规律对于步态识别的研究将有很大的指导意义。

第六章 基于步态的性别识别

智能视觉监控是步态识别的重要应用领域，当场景中出现行人时，步态识别系统便识别出该人的身份。但是由于监控场景一般为开放的场景，即经常会有一些未在识别系统中注册的行人出现，比如进入某大楼的访客。当识别系统不能识别出这些人的身份时，系统应该提供一些行人的有用信息，比如系统可以提供“一个身高 175 厘米体形微胖的中年男性进入大楼”。在这个例子里，如果系统不能识别行人身份，则会提供行人的身高（175 厘米）、体形（微胖）、年龄段（中年）和性别（男性）信息。该技术可以大大提高系统对场景的理解能力，有助于推动视觉监控向更加智能化的方向发展。除此之外，还可以使用身高、体形、年龄段和性别信息对步态特征进行粗分类，不仅可以提高步态数据库检索的速度，而且还可以提高步态识别的精度。

一些研究者已经进行了利用视觉的方法获取人的身高以及体形的工作 [4, 41]。对于年龄的估计，目前大部分工作都以人脸 [57, 58] 和声音 [59, 60] 为特征，目前尚没有专为估算年龄而设计的步态数据库，进行基于步态的年龄估计尚有一定的困难。基于步态的性别识别，已有一些研究者进行了开创性研究，而且有一些记录了性别信息的数据库可以使用，在本章中，我们将对基于步态的性别识别进行全面的的研究，包括特征分析、跨人种实验、多视角实验分析等。

6.1 相关研究

目前文献中与性别识别相关的工作很多都是基于人脸 [61, 62, 63] 或者声音 [64, 65] 的。对于基于步态的性别识别，Kozlowski 等人在比较早的时候就进行了研究 [66]。在他们的实验中，将一些发光点固定在人体的主要关节上，由人观察活动的发光点来判断性别，获得了 63% 正确识别率。他们还发现运动的发光点比静止的发光点更易于用来区分性别。后来提出的很多方法 [67, 68] 也是使用附在关节上的发光点来进行性别识别。其中 Davis 等人提出的方法 [67] 是一种自动的性别识别方法，他们采用了三模态主成分分析法（three-mode PCA）来从运动的发光点中提取特征，这种算法在 40 人（20 男，20 女）的数据库上获得了 95.5% 的识别率。Davis 等人还让 15 个人通过观察运动的发光点来判断性别，只获得了 69% 的识别率，这跟论文 [66, 68] 中的结果大致在一个水平上。从

以前的工作中我们可以看出，由人观察从关节上采来的点数据，不能获得很高的识别率，相反的，计算机可以获得比较高的识别率，

人的肩髌比，腰髌比，以及人行走时的身体晃动也可以用来区分性别 [69]。一般来说，男性的肩膀较宽，腰较粗，髌部较窄，而女性则恰恰相反。同时男性走路时容易晃动肩膀，而女性更多的是臀部的扭动。不过，从视频中精确的提取这些特征是一项很有挑战性的工作。

一些研究者也提出了基于人体轮廓的性别识别方法。例如 Lee 等人从运动的人体轮廓中提取步态特征 [20]。人体轮廓首先被分成 7 部分，然后用一个椭圆来拟合每个部分。椭圆的中心，长轴跟短轴的比例，以及长轴的方向被提取出作为步态特征。Lee 等人的方法在一个 24 人（14 男，10 女）的数据库上获得了 84.5% 的正确识别率。跟 Lee 等人的方法不同，Yoo 等人的方法是用从轮廓中提取的二维棍棒模型来表示人体，这个棍棒模型包含头，颈，肩，腰，髌，膝盖和脚踝。使用支持向量机进行分类，在一个 100 人（84 男，16 女）的数据库上获得了 96% 的识别率。

一般来说，步态包含人体的形状信息和运动信息。在基于步态的性别识别中，哪种信息贡献更大，或者说哪种信息的区分能力更好？在论文 [70] 中，McDonnell 等人认为如果使用发光点特征，运动信息区分能力更好；如果使用轮廓，则人体形状的区分能力更好。在后面的实验中，本文也将对这一问题进行分析。

6.2 由人识别性别

在进行使用计算机识别人的性别实验之前，我们首先进行了一个由人观察步态来识别性别的实验。一共 30 个人（15 男，15 女）参加了这个实验。实验中使用的数据来自 CASIA 步态数据库数据集 B（详见第三章）。CASIA 步态数据库数据集 B 共包含 124 人，但是只有 31 位女性。性别识别是一个两类问题，所以最好数据集中男女的数目相等，故在本实验中数据来自这 31 个女性和随机选择出的 31 位男性。同时为了研究在视角不同的情况下，基于步态的性别识别能否受到影响，在实验中使用了前面选出的 62 人的 0° ， 36° 和 90° 三个角度的数据。

为了去除衣服颜色和减小衣服式样等先验知识对识别结果的影响，在实验中只是使用了二值的人体轮廓。人体轮廓以 25 帧每秒（跟原始视频速度一致）

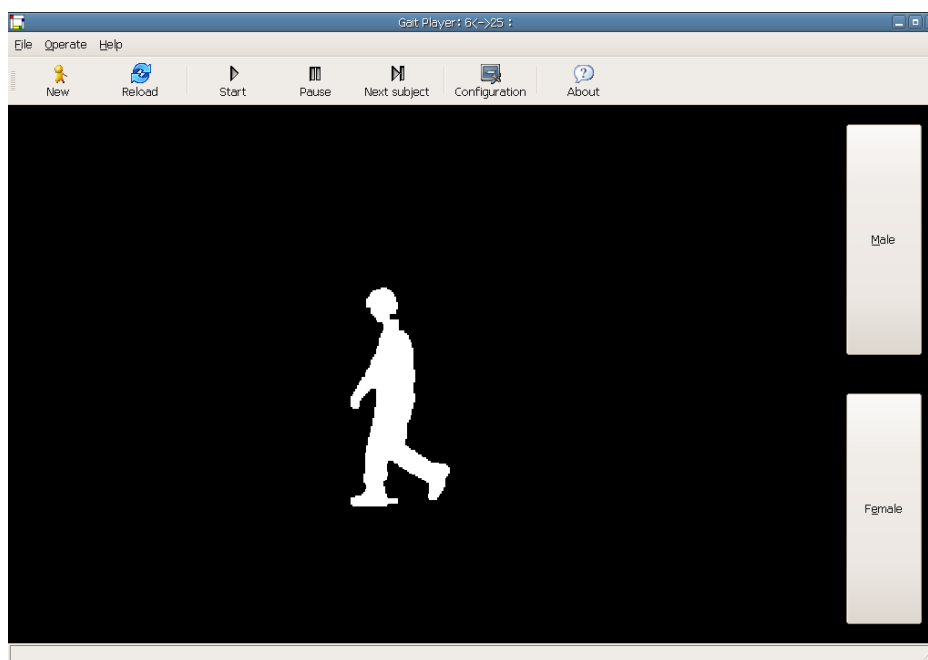


图 6.1: 播放人体轮廓的软件

表 6.1: 由人观察轮廓识别性别的识别率

角度	男观察者	女观察者	综合
0°	86.13%	78.92%	82.53%
36°	90.97%	86.13%	88.55%
90°	95.59%	95.38%	95.48%

的速度在屏幕上循环播放，如图 6.1 所示。实验者被要求观察计算机屏幕上运动的轮廓，并判断其性别。判断出性别后，用鼠标点击屏幕上的“男”或“女”按钮，由软件将识别结果记录。然后软件自动的随机选择出下一个人的步态轮廓在屏幕上播放，等待实验者判断性别。本实验的实验结果列在表 6.1 中。

从实验结果可以发现，侧面视角（90°）与正面视角（0°）相比，可以获得更高的识别率。这一点也是符合常理的，因为侧面视角可以更多的表示不同性别的特征，特别是胸部，臀部以及发型的信息可以被更好的表达。因此可以说人体的侧面轮廓在对性别的区分方面，比正面轮廓更有优势。在这一点上，性别识别跟身份识别有着相似性，因为身份识别的最佳角度也是侧面视角（参见第 4.2 节）。这些结论都是由观察者得来的，如果让计算机来识别性别，分类算法

能否根据步态区分不同的性别? 另外, 如果分类算法可以区分性别, 那么哪个视角的步态数据区分能力最强?

在本实验中, 侧面视角的性别识别率超过了 95%, 但是在文献 [68, 71] 中由观察者获得的性别识别率却只有 60% 左右。在他们的研究中, 用的特征是附加在人身上的发光点, 在使用这种特征的时候, 由识别算法获得的识别率远远超过观察者。为何识别算法能够从几个跟踪点判断出性别, 而观察者却很难做到? 这应该是认识领域的一个值得研究的方向。

除此之外, 还可以发现在实验中的 30 位观察者中, 男性观察者区分性别的能力要强于女性观察者。而且特别是正面视角, 男性观察者的识别率比女性观察者高 7.21%。由于目前的实验只有 30 位观察者, 而且这个问题涉及不同性别的人的认识能力, 需要进一步研究。

6.3 步态特征与分类器

在实验中, 由于拍摄视频的摄像机都是固定的, 所以可以使用背景减除法(第 2.1.1 节)来分割人体轮廓。另外, 由于人体的轮廓大小不一, 为了处理的方便, 需要将所有的轮廓都归一化到相同的大小。

轮廓归一化后, 可以提取出步态特征。在后面的实验中, 主要使用了如下几种步态特征: 步态能量图 [29], 步态能量图的 PCA 系数, 人体轮廓线 (contour), 人体轮廓线的 PCA 系数, 人体轮廓线的关键傅立叶描述子 [28]。这些特征的具体描述可以参考第 2.1.3 节。在实验中, 当使用 PCA 系数作为步态特征时, 都仅仅选取了前 10 个系数。

提取出步态特征以后, 需要将特征输入分类器进行分类。用到的分类器有 k 近邻 (knn) 分类器, 贝叶斯分类器和支持向量机 (svm)。关于分类器的详细描述可以参考第 2.1.4 节。

6.4 数据库

测试测试性别识别算法的有效性需要一个比较大的步态数据库。目前已经公开的规模比较大 (超过 100 人) 的数据库有 Gait Challenge 数据库 [72], Soton 大数据库 [48], 以及 CASIA 步态数据库数据集 B。在这三个数据库中, CASIA 步态数据库数据集 B 和 Soton 数据库记录了人的性别, 所以实验中采用了这两个数据和一个合成的更大的混合数据库来测试算法。

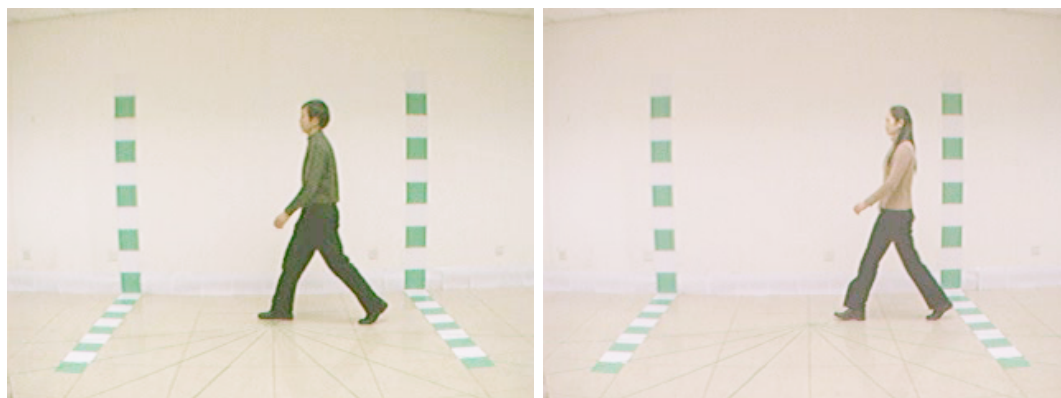


图 6.2: CASIA 步态数据库数据集 B 中不同性别的视频

跟第 6.2 节中一样，从 CASIA 步态数据库数据集 B 中选择了 31 位男性和 31 位女性的数据进行实验，所有被选出的人为中国人 / 亚洲人。该数据库中每个人的数据是从 11 个视角采集，每个视角有 6 个正常行走的步态序列。这样就构成了一个包含 62 人，11 个视角，一共有 4092 个步态序列。CASIA 步态数据库数据集 B 中的两个不同性别的样本如图 6.2 所示。

Soton 大数据库是由英国的南安普顿大学创建，包含 115 个人。所有的视频都在室内拍摄，拍摄的数据为侧面数据。在这个数据库中，女性只有 24 人，所以选择了 24 位女性的数据和随机选择了 24 位男性，总共 48 人的数据进行实验。跟 CASIA 数据库不同，Soton 大数据库包含多种人种，选出的这 48 人既有亚洲人，也有欧洲人。由于这个数据库只有一个侧面视角，每个人选择了侧面视角的 6 个序列。Soton 大数据库中的两个不同性别的步态序列如图 6.3 所示。

为了获得一个更大的步态数据库，将从 CASIA 步态数据库中选出的 62 人的侧面视角数据和 Soton 大数据库中选出的 48 人的数据合在一起，构成了一个更大的数据库。这个新数据库被称作混合库，包含 110 人。混合库中的数据由不同的设备，在不同的条件下采集，是一个包含多个人种和多种变化因素的具有挑战性的数据库。

6.5 实验和分析

为了发掘步态特征在性别识别中的潜力，特别是各种步态特征和分类器的潜力，我们设计了一系列实验来评价这些特征和分类器。在随后的部分也分析了人体的哪些部分具有比较好的性别区分能力问题。除此之外，还进行了跨人

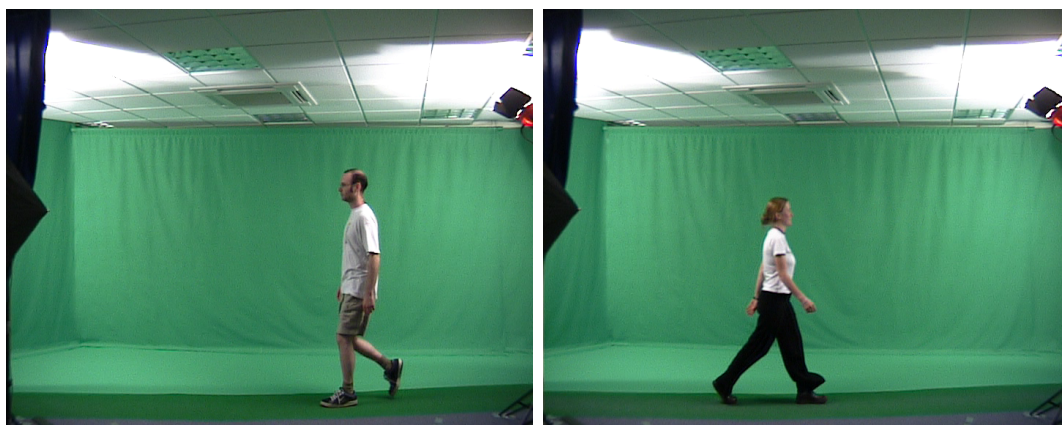


图 6.3: Soton 大步态数据库中不同性别的视频

表 6.2: 使用 CASIA 步态数据的性别识别实验结果

步态特征	kNN (k=21)	Normal Bayes	SVMs
步态能量图	90.86%	-	91.67%
步态能量图的 PCA 系数	90.86%	86.29%	93.55%
轮廓线	73.66%	-	85.75%
轮廓线的 PCA 系数	68.82%	65.05%	74.19%
关键傅立叶描述子	73.66%	76.08%	85.48%

种的实验, 实验充分证明了基于步态的性别识别的可行性。

6.5.1 各种不同特征和分类器组合的实验

在使用 CASIA 步态数据库的实验中, 首先是对侧面视角步态数据的实验。实验中使用了 m 重交叉验证 [38] (m -fold cross-validation) 来评价不同的步态特征和分类器。数据库中的每个人有 6 个侧面视角的步态序列。数据库中的 62 个人被分成 31 个无重合数据的组, 每组中有 1 个男性的 6 个序列和 1 个女性的 6 个序列, 一个组中共有 12 个序列。这样就可以构成 31 重的交叉验证 ($m=31$), 即每次取出一个分组作为测试集, 剩下的 30 个分组为训练集。这样保证了同一个人的步态数据不可能同时出现在测试集和训练集中。在这个 31 重的交叉验证中共有 31 种实验组合, 实验结果 (正确识别率) 为这 31 种组合的平均值。使用 CASIA 步态数据库的所有特征和分类器的组合实验结果列在表 6.2 中。

表 6.3: 使用 Soton 大步态数据的性别识别实验结果

步态特征	kNN (k=21)	Normal Bayes	SVMs
步态能量图	85.07%	-	85.07%
步态能量图的 PCA 系数	83.68%	79.17%	86.81%
轮廓线	72.57%	-	86.81%
轮廓线的 PCA 系数	60.42%	67.71%	61.81%
关键傅立叶描述子	65.97%	78.47%	84.38%

表 6.4: 使用混合步态数据的性别识别实验结果

步态特征	kNN (k=21)	Normal Bayes	SVMs
步态能量图	87.27%	-	91.36%
步态能量图的 PCA 系数	87.88%	80.76%	90.76%
轮廓线	73.18%	-	76.06%
轮廓线的 PCA 系数	66.06%	65.45%	62.73%
关键傅立叶描述子	70.30%	60.45%	81.21%

在使用贝叶斯分类器的时候,假设数据的分布是一个高斯分布。如果特征为高维空间的特征,比如步态能量图和步态轮廓线,由于维数过高,很难建立高斯模型,所以当步态能量图或步态轮廓线为步态特征的时候,没有使用贝叶斯分类器进行分类。

跟使用 CASIA 步态数据库的实验一样,使用 Soton 大数据库和混合数据库的实验也采用了 m 重交叉验证。在交叉验证中,每次将两个人(一男一女)的步态数据放入测试集,其他的数据放入训练集。使用 Soton 大数据库和混合数据库的实验结果列在表 6.3 和表 6.4 中。

从列在表 6.2-6.4 中的实验结果可以发现跟其他特征相比,步态能量图的 PCA 系数具有最好的区分能力。在分类器里面,支持向量机优于其他分类器。步态能量图的 PCA 系数特征和支持向量机这一组合基本上取得了各组实验的最高识别率。

虽然实验中所使用的轮廓线特征保留了时间信息,而步态能量图特征几乎完全抛弃了时间信息,只保留了体形特征,但是实验结果却显示步态能量图特征得到了更好的实验结果。这说明人体的体形特征是一种有效的区分性别的特

征, 这跟 McDonnell 的论文 [70] 中的结论是一致的。

6.5.2 最佳分类特征

利用 PCA 得到的特征为最佳描述特征 (the Most Expressive Features, MEF), 虽然在前面的实验中步态能量图的 PCA 系数获得了比较好的实验结果, 但是最佳描述特征未必是最佳分类特征。性别识别是一个分类问题, 所以应该尝试提取最佳分类特征来进行性别识别。一种最佳分类特征可以使用线性鉴别分析 [73] (linear discriminant analysis, LDA) 来获得。

LDA 提取特征的方法如下所述。假设一个步态特征向量为 \mathbf{x}_{ij} , 其中 $i \in \{1, 2\}$, $j \in \{1, 2, \dots, c_i\}$, c_i 是第 i 类的样本数。向量 $\mathbf{y}_{ij} = \mathbf{w}^T \mathbf{x}_{ij}$ 是 \mathbf{x}_{ij} 在另一个空间的表达。在鉴别分析中, 我们希望找到一个最佳投影 \mathbf{w} 来最大化类间距离与类内距离的比值 $J(\mathbf{w})$ 。

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}} \quad (6.1)$$

$$\mathbf{S}_b = \sum_{i=1}^2 \sum_{j=1}^{c_i} (\mathbf{x}_{ij} - \mathbf{m}_i)(\mathbf{x}_{ij} - \mathbf{m}_i)^T \quad (6.2)$$

$$\mathbf{S}_w = \sum_{i=1}^2 (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \quad (6.3)$$

其中 \mathbf{m}_i 是第 i 类的均值, \mathbf{m} 是所有样本的均值。在 [74] 中已经证明了对于两类问题, 最优的投影矩阵可以这样求解得到:

$$\mathbf{w}^* = \mathbf{S}_w^{-1}(\mathbf{m}_1 - \mathbf{m}_2) \quad (6.4)$$

MDA 特征可以通过下面的方程求得:

$$\mathbf{y}_{ij} = \mathbf{w}^{*T} \mathbf{x}_{ij} \quad (6.5)$$

因为对于二类问题 \mathbf{w}^* 是一个向量, 所以特征 $\{\mathbf{x}_{ij}\}$ 被投影到 1 维空间里。

一般来说, 步态的原始特征 $\{\mathbf{x}_{ij}\}$ 一般具有很高的维数, \mathbf{S}_w 很容易变成奇异矩阵而不可逆, 这样 \mathbf{w}^* 就不能通过公式 6.4 计算得到。要使得 \mathbf{S}_w 变成可逆矩阵, 一个办法是对原始特征先利用 PCA 进行降维 (PCA 降维详见第 5.3.3 节), 然后使用 LDA 提取特征。利用 PCA 和 LDA 提取特征的过程如下式所示:

表 6.5: 使用最佳分类特征时的实验结果

特征	CASIA 数据库	Soton 数据库	混合数据库
GEI+PCA+LDA	94.89%	86.46%	90.45%
Contour+PCA+LDA	74.19%	60.76%	65.00%

$$\mathbf{z}_{ij} = \mathbf{w}^{*T} \mathbf{y}_{ij} = \mathbf{w}^{*T} [\mathbf{e}_1, \mathbf{e}_1, \dots, \mathbf{e}_m]^T \mathbf{x}_{ij} \quad (6.6)$$

其中 $[\mathbf{e}_1, \mathbf{e}_1, \dots, \mathbf{e}_m]$ 是 PCA 中的投影矩阵, 在实验中, 使用了前 10 个主成分, 所以 $m = 10$ 。 \mathbf{w}^* 是从特征 $\mathbf{y}_{ij} = [\mathbf{e}_1, \mathbf{e}_1, \dots, \mathbf{e}_m]^T \mathbf{x}_{ij}$ 计算出的。

因为性别识别是一个两类问题, 所以 \mathbf{z}_{ij} 是一个标量, 而不是一个向量, 因此可以简单的设置一个阈值来实现分类。

$$L(\mathbf{z}_{ij}) = \begin{cases} male, & \text{if } \mathbf{z}_{ij} \geq T; \\ female, & \text{else.} \end{cases} \quad (6.7)$$

其中阈值 T 可以设置为 $\{z_{ij}\}$ 的均值:

$$T = \frac{1}{2} \sum_{i=1}^2 \left(\frac{1}{c_i} \sum_{j=1}^{c_i} \mathbf{z}_{ij} \right)$$

在实验中, 对步态能量图特征和轮廓线特征的 PCA 系数利用 LDA 进行了最佳分类特征提取, 实验结果列在表 6.5 中。在前面的实验中, 使用支持向量机对步态能量图的 PCA 系数进行分类获得最好的实验结果, 在此处, 最佳分类特征获得了跟前面最好结果不相上下的识别率。但是利用最佳分类特征的计算复杂度要比利用支持向量机低, 而且训练时的复杂度也低。从步态能量图上提取的最佳分类特征具有很强的区分能力, 它在保证识别率的同时还降低了计算复杂度。但是, 基于轮廓线的最佳分类特征并没有获得令人满意的结果。为何基于轮廓线的特征总是劣于基于步态能量图的特征, 这个问题将在下一节中分析。

6.5.3 轮廓与轮廓线

从理论上讲, 从同一物体提取的轮廓 (silhouette) 和轮廓线 (contour) 包含同样的信息。步态能量图是基于轮廓的特征, 是所有轮廓的平均, 所以它丢弃了时间信息。前面用到的轮廓线特征, 保留了时间信息。因此轮廓线特征应该

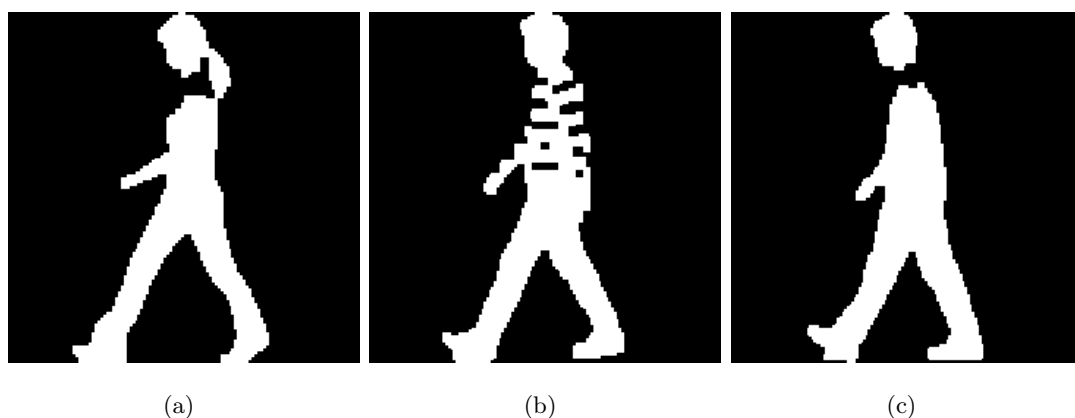


图 6.4: 受噪声影响的轮廓

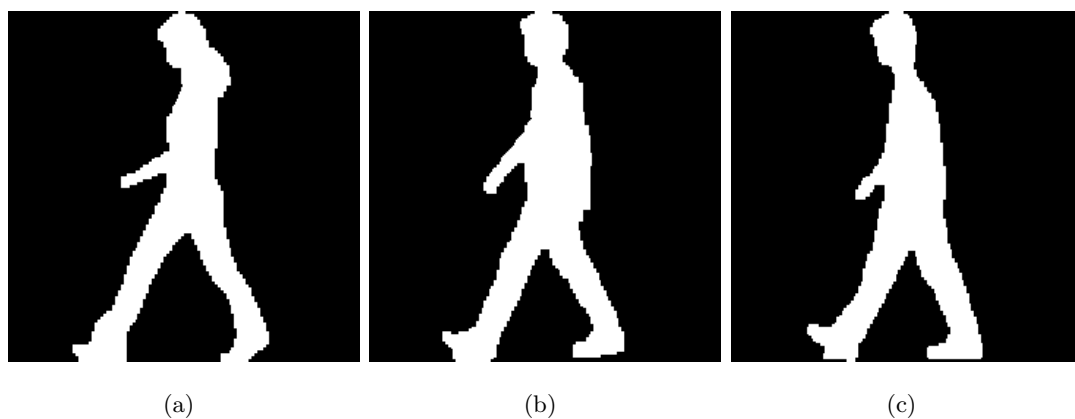


图 6.5: 真实轮廓

具有比步态能量图包含更多的信息，但是为何在前面实验中，步态能量图却又优于轮廓线特征呢？跟在我们的论文 [36] 中分析的一样，有噪声存在的情况下，轮廓线一般比轮廓更加不稳定。浏览数据库中的轮廓图像，可以发现一些受噪声影响，没有被很好的分割的轮廓（图 6.4）。

一般来说，轮廓的表达方式跟轮廓线是不一样的。轮廓一般用一幅二值图像来表示，而轮廓线一般是用一个向量来表示所有的轮廓外边缘上的点，所以轮廓特征的维数要比轮廓线特征的维数高很多。这样也就意味着轮廓的表达方式有很大的数据冗余。数据冗余虽然“浪费”了存储空间，但是一般来说数据冗余可以提高特征对噪声的鲁棒性。为了评价特征对噪声的鲁棒性，通过手工制作出了图 6.4 中的真实轮廓（ground truth silhouette），如图 6.5 中所示。

表 6.6: 噪声对轮廓和轮廓线特征的影响

特征	图 6.4(a)	图 6.4(b)	图 6.4(c)
轮廓	4.43%	8.45%	2.34%
轮廓线	35.08%	32.50%	28.45%

下面的公式可以用来评价当存在噪声时, 被“污染”的特征与真实特征的差距, 可以评价特征对噪声的鲁棒程度。

$$\rho = \frac{|\mathbf{x} - \mathbf{x}_{gt}|}{|\mathbf{x}_{gt}|} \quad (6.8)$$

其中 \mathbf{x} 是被噪声“污染”了的特征向量, \mathbf{x}_{gt} 是真实的特征。 ρ 的值越大, 表示噪声影响越大。图 6.4 提取的轮廓和轮廓线与它们的真实数据计算出的 ρ 值列在表 6.6 中。从表 6.6 中可以看出, 同样的噪声, 对轮廓线产生的影响要比对轮廓产生的影响大很多。这也是在前面的实验中, 基于轮廓线的特征的认识率要低于基于轮廓的特征的原因。

要提高基于轮廓线的步态特征对噪声的鲁棒性, 一种方法是提高人体检测和分割的准确性, 但是从背景中非常完美的分割出人体, 特别是在复杂背景下, 是一件非常困难的事情。另外一种解决方法是使用更高效的轮廓线相似性度量方法, 例如在度量两个轮廓线的相似程度时, 先将两个轮廓线各部分匹配, 例如头跟头匹配, 手跟手匹配, 然后在匹配基础上再衡量两个轮廓线的相似程度, 可以提高轮廓线特征对噪声的鲁棒程度。

6.5.4 空间域特征分析

很多人会问到的一个问题是: 人体的哪个部分在性别识别中起着最重要的作用呢? 为了回答这个问题, 我们使用方差分析 (Analysis of variance, ANOVA) 来计算人体不同部分的对性别的区分能力。ANOVA 的 F 统计量是一个用来衡量特征区分能力的度量。F 统计量可以用下式计算得到 [75]:

$$\mathcal{F} = \frac{\frac{1}{c-1} \sum_{i=1}^c n_i (\bar{x}_i - \bar{x})^2}{\frac{1}{n-c} \sum_{i=1}^c \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2} \quad (6.9)$$

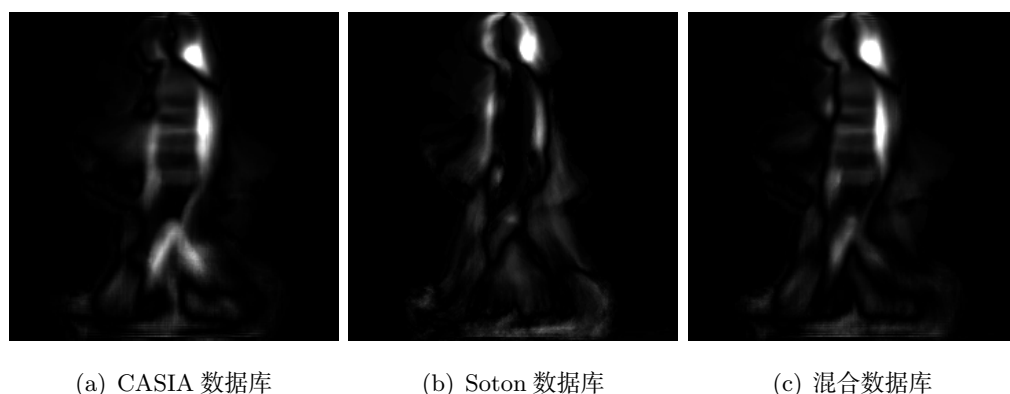
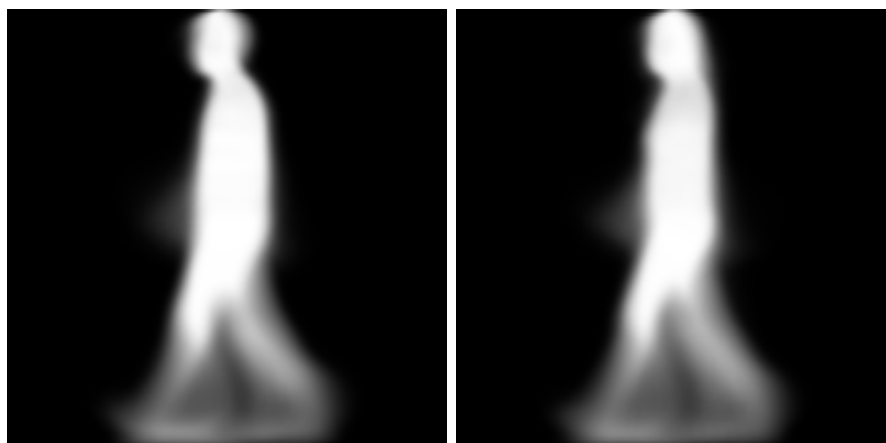


图 6.6: 人体轮廓各个部位的 F 统计量

其中 x_{ij} 是第 i 类中的第 j 个特征向量, c 是类别数目, n_i 是第 i 类的样本数目, 并且有 $n = \sum_{i=1}^c n_i$, \bar{x}_i 是第 i 类中样本的平均值, \bar{x} 是所有样本的平均值。

因为步态能量图在前面的实验中获得了比较高的识别率, 所以采用步态能量图的 F 统计量来分析身体的哪个部分对性别识别贡献大。在使用公式 6.9 计算 F 统计量过程中, 式中的 x 为步态能量图里的一个像素, 而不是整个步态能量图, 所以计算出的 \mathcal{F} 值也可以用一个图像来表示。从三个数据库 (CASIA, Soton 和混合数据库) 中计算出的 F 统计量分别显示图 6.6.(a), (b) 和 (c) 中。图 6.6 中某处的像素点越亮, 表示此处的区分能力越强; 像素点越暗, 表示此处的区分能力越弱。

从图 6.6 可以发现, 头发区域是一个区分能力比较强的区域, 这很容易理解, 因为一般男性为短发, 而女性多数为长发。另外一个区分能力比较强的区域是胸部和背部, 不同性别的人在背部的区别是因为女性躯干比男性细, 衣服更紧身。另外, Soton 数据库中的数据显示胸部也是一个能够明显区分性别的区域, 而 CASIA 数据库中的数据却没有显示这个结论。其原因应该是由衣着因素引起, Soton 数据库的采集时间是夏季, 穿衣服比较少, 而 CASIA 数据库的采集时间是冬季, 穿衣服比较多, 所以胸部特征不明显。另外, 对混合数据库中所有男性的步态能量图和女性的步态能量图分别求平均, 得到了两个平均图像 (图 6.7)。由这两个平均图像, 可以得到相似的结论, 即头颈后部、背部以及胸部是对性别识别贡献比较大的区域。



(a) 所有男性的步态能量图的均值 (b) 所有女性的步态能量图的均值

图 6.7: 混合数据库中不同性别的特征的均值

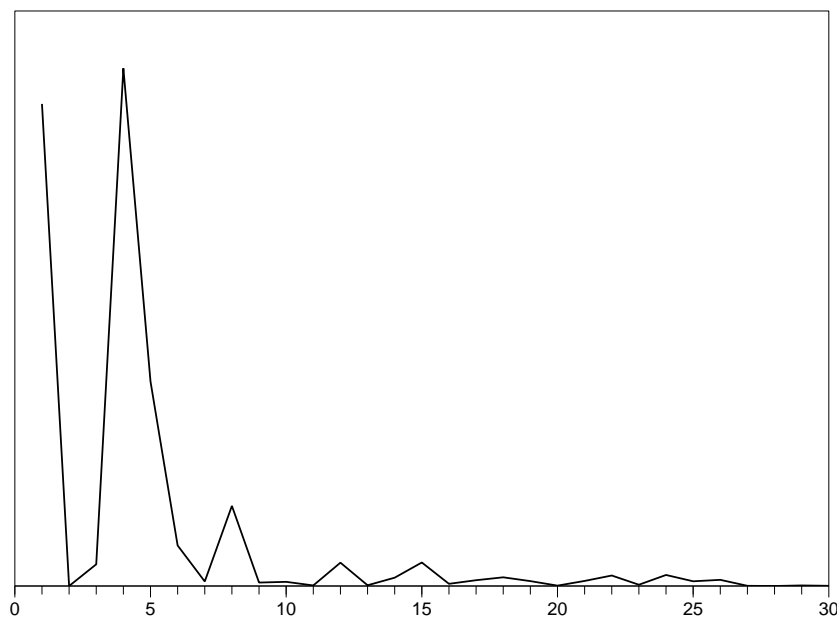
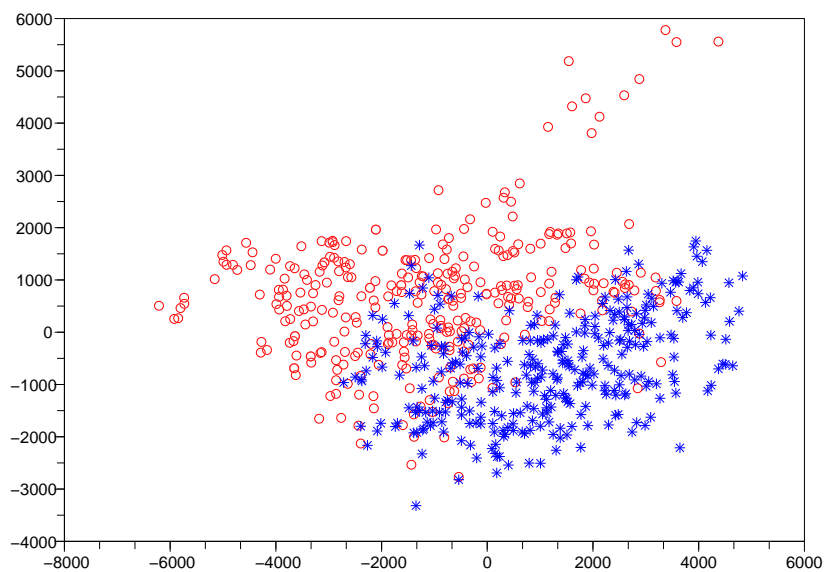
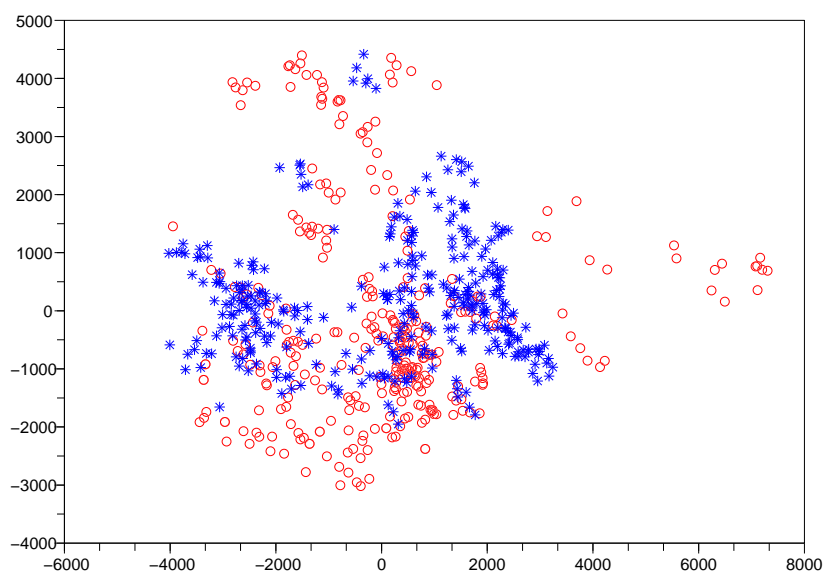


图 6.8: 步态能量图的前 30 个 PCA 系数的 F 统计量 (混合库)



(a) 第 1, 4 个系数



(b) 第 2, 3 个系数

图 6.9: 步态能量图的前 4 个 PCA 系数 (混合库, 蓝 * 代表男, 红 o 代表女)

表 6.7: 使用不同的 PCA 系数获得的识别率 (混合库)

PCA 系数	kNN 分类器	贝叶斯分类器	支持向量机
1,4,5,8	90.76%	76.82%	90.91%
2,3,6,7,9,10	69.39%	66.36%	71.97%

6.5.5 PCA 域特征分析

在前面的实验中,步态能量图的 PCA 系数取得比较高的识别率,所以进行了各个 PCA 系数的区分能力的分析。实验中依然使用了 ANOVA 中的 F 统计量来衡量各个维度上的 PCA 系数。从混合数据计算出的前 30 个 PCA 系数的 F 统计量如图 6.8 所示。从图中可以很明显的看出第 1、4、5 和 8 个 PCA 系数具有比较好的区分性能,其他的 PCA 系数则区分能力不强。将前 4 个 PCA 系数以 2 维的形式显示出来,如图 6.9 所示。从图 6.9(a) 可以看出,只使用第 1 和 4 PCA 系数,可以用一个简单的线性分类器,将大部分不同性别的数据分开。但是对于图 6.9(b),不同性别的数据混在一起,难以分开。

只使用 4 个最具有区分能力的 PCA 系数,以及使用前 10 个 PCA 中其他的 6 个系数进行性别识别,得到的实验结果列在表 6.7 中。只用 4 个最具有区分能力的系数,获得的识别率跟用所有的前 10 个系数差别不大,甚至在使用 kNN 分类器的时候,只用 4 个系数获得了比用前 10 个系数更高的识别率。而其他 6 个系数,却只获得了 70% 左右的识别率,对于两类问题来说,这是一个很低的识别率,因为抛硬币方式尚有 50% 的正确率。所以说,只有第 1、4、5 和 8 个 PCA 系数对性别识别做出了明显的贡献。前 10 个特征向量的图像表示形式显示在图 6.10 中,在最具有区分能力的 4 个特征向量中,可以由图看出第 1 和 5 系数分别对应头发部分,第 4 系数分别对应背部部分,因此也就是头发和背部区域具有比较强区分能力,这跟上一节的结论相同。

6.6 跨人种的性别识别

对于不同的人种,很多生物特征也会不同。比如欧洲人的人脸跟亚洲的和非洲的就不同。对于虹膜也是如此,不同人种的虹膜纹理和颜色是不一样的 [76]。对于步态,不同人种的步态是否也会不一样? 如果使用一种人种的数据进行训练,能否根据步态识别出另一人种的性别? 在后面,我们设计了几个实验来回答这两个问题。

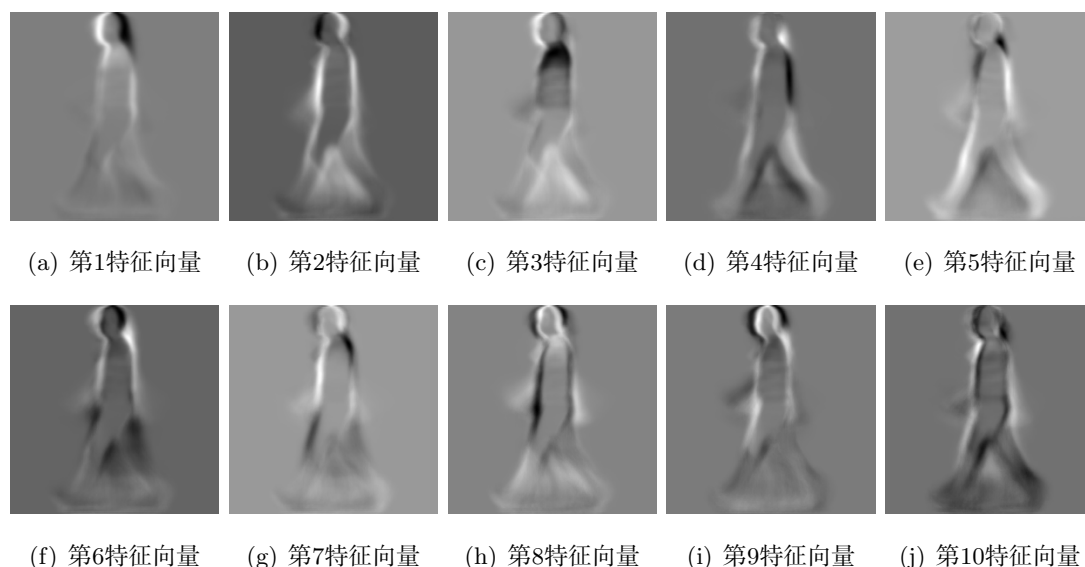


图 6.10: 步态能量图的前 10 个 PCA 特征向量 (混合库, 黑色代表负值, 白色代表正值, 灰色代表零值)

表 6.8: 跨人种性别识别结果 (训练数据为亚洲人, 测试数据为欧洲人)

kNN (k=21)	贝叶斯分类器	支持向量机	最佳分类特征
76.11%	74.44%	90.00%	90.00%

在前面的实验中使用的来自 CASIA 步态数据数据集 B 的数据都是亚洲人 (中国人), 但是 Soton 数据却包含很多人种, 有欧洲人, 也有亚洲人和非洲人。为了评价跨人种的性别识别, 设计了两组实验: 第一组实验将 CASIA 数据的 62 个亚洲人数据作为训练数据, 选自 Soton 数据库的 30 个欧洲人作为测试数据。因为 Soton 库中只有 15 个欧洲女性, 所以只选择了这 15 个欧洲女性和另外 15 个欧洲男性。第二组实验跟第一组实验类似, 不同的只是在第二组实验中, 将从 Soton 数据库中选出的 30 个欧洲人的数据作为训练数据, 从 CASIA 选出的 62 人的数据作为测试数据。

在每一组实验中有 4 个实验。这 4 个实验的前 3 个使用相同的特征, 为步态能量图的前 10 个 PCA 系数, 分类器分别为 kNN 分类器, 贝叶斯分类器和支持向量机; 第 4 个实验使用的特征为 LDA 提取的最佳分类特征, 使用一个阈值进行分类。这两组实验的实验结果分别列在表 6.8 和表 6.9 中。

首先从实验结果中可以发现, 表 6.9 中的识别率要低于表 6.8 的识别率。引

表 6.9: 跨人种性别识别结果 (训练数据为欧洲人, 测试数据为亚洲人)

kNN (k=21)	贝叶斯分类器	支持向量机	最佳分类特征
68.55%	53.23%	85.22%	81.18%

起这个问题的原因应该是在第二组实验中, 训练数据为欧洲人, 训练样本比较少 (只有 30 人) 导致, 而第一组中的训练数据为 62 人。除了人种差异外, 两个数据库的采集条件、采集设备、采集时间等都不同, 所以这两组实验应该是非常有挑战性的实验。但是从表 6.8 和表 6.9 中我们可以发现, 跨人种的基于步态的性别识别, 最高识别率可达 90.00%。这说明跨人种的基于步态的性别识别是可行的, 不同人种的步态没有明显差别。

6.7 多视角下的性别识别

在上一节的实验中, 只有侧面视角的步态特征被用来识别性别。在第 6.2 节中, 我们从实验中发现, 在由人观察轮廓判断性别的实验中, 侧面视角获得了最高的识别率, 正面视角的识别率最低。自动性别识别算法是否也会有同样的规律? 为了分析不同视角下性别识别的准确率, 我们使用 CASIA 数据库中的 11 个视角的步态数据来进行性别识别实验。

在实验中, 注册数据和测试数据都来自同一个视角, 所使用的特征为在前面实验中取得比较好的效果的步态能量图的前 10 个 PCA 系数, 实验依然使用 m 重交叉验证。使用 k 近邻分类器和支持向量机对 11 个视角的数据进行性别识别的实验结果显示在图 6.11 中。从实验结果中我们可以发现使用正面 (0°) 或背面 (180°) 的步态特征进行性别识别, 得到的识别率最低。

6.8 视角无关的性别识别

如果在实际应用中使用时基于步态的性别识别, 那么也会跟步态识别一样, 不可避免的遇到视角变化问题。也就是当测试数据的视角跟注册数据的视角不一致时, 如何保证依然能够正确的识别出性别。在本文第四章中, 提出了一个评价步态识别算法的框架。这个框架中提供了 3 组实验, 可以用来评价步态识别算法对视角变化、衣着变化和携带物品变化的鲁棒性。在此处, 也可以使用这个框架来评价基于步态的性别识别对视角变化的鲁棒性。

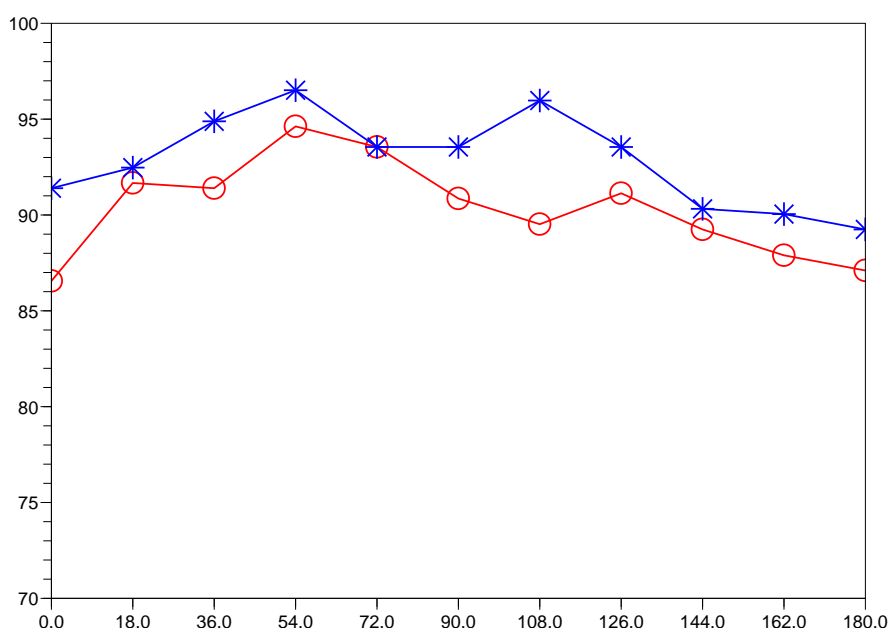


图 6.11: 在不同视角下性别识别结果 (○为 kNN 结果, * 为 SVM 结果)

实验集 A 中共有 121 个实验, 对于其中的每一个实验, 假设注册数据的视角为 α , 测试数据的视角为 β , 实验依然采用 m 重交叉验证, 在每一重中, 取出男女各一人的 β 角度数据作为测试数据, 将其余的 60 人的数据 α 作为注册数据。使用步态能量图的前 10 个 PCA 系数和支持向量机分类器的实验结果列在表 6.10 中。从中可以发现, 如果测视角跟注册角差别比较大, 性别识别的识别率降至 50% 左右。因为性别识别是一个两类问题, 所以即使采用抛硬币方式猜测, 都可以达到 50% 的识别率。因此识别率在 50% 左右意味着分类完全失败。

在第五章中, 提出了两种步态特征合成方法: 基于线性投影的线性模型和基于神经网络的非线性模型。这两种方法可以把从一个视角采集的步态特征合成为另一个视角的特征, 用于测试视角与注册视角不一致问题。这种特征合成方法应该也可以用于性别识别。由于基于神经网络的非线性模型复杂度比较高, 与线性模型相比, 在步态识别里面并没有明显的提高, 所以只使用第 5.3 节中的线性模型来合成特征。使用合成的步态特征和支持向量机分类器得到的实验结果如表 6.11 所示。从实验结果可以看出, 合成的步态特征大大的提高了性别识

表 6.10: 视角变化时的性别识别的识别率 (SVM 分类器)

CCR (%)	测试集角度 θ_p (正常行走)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_r (正常行走)	0°	91.40	85.75	86.29	50.00	50.00	50.00	50.00	50.00	54.84	85.75	81.99
	18°	84.14	92.47	83.60	50.81	50.00	50.00	50.00	50.00	56.99	78.49	82.26
	36°	84.68	84.95	94.89	95.70	80.91	63.44	60.22	90.86	88.44	80.65	79.03
	54°	50.00	71.77	92.47	96.51	89.78	81.45	83.87	93.28	90.32	51.88	50.00
	72°	50.00	50.27	75.81	93.55	93.55	95.70	93.01	89.52	69.35	50.00	50.00
	90°	50.00	50.00	48.92	90.32	94.89	93.55	91.40	81.99	49.46	50.00	50.00
	108°	50.00	50.00	62.37	93.28	92.20	93.82	95.97	93.82	62.90	51.61	50.00
	126°	50.00	68.01	79.84	93.01	87.37	77.96	85.48	93.55	86.29	68.82	50.00
	144°	82.80	75.27	76.61	79.03	73.66	59.14	76.08	90.05	90.32	92.20	80.11
	162°	80.38	81.45	66.40	62.63	48.39	48.12	50.81	66.40	83.60	90.05	88.71
180°	71.51	70.97	61.02	51.08	50.00	50.00	50.00	45.97	62.90	83.33	89.25	

别的识别率, 在表 6.11 中, 最低的识别率值也达到了 82.8%。

此外, 在步态识别评价框架中设计的用于评价算法对视角的鲁棒性的参数 σ_i^A , 也可以用于此处评价性别识别算法的鲁棒性。不使用合成特征 (GEI+PCA) 和使用合成特征 (GEI+PCA+Linear) 这两种情况下的评价参数 σ_i^A 的值如表 6.12 所示。 σ_i^A 的值越大, 表明算法的鲁棒性越好。从表中可以看出, 使用合成的特征, 算法的鲁棒性明显优于不使用合成特征的方法。

6.9 本章小结

随着智能视觉监控研究受到越来越多的关注, 基于步态的性别识别也逐渐引起了重视。基于步态的性别识别是一个很新而且没有被深入研究的方向。特别是基于整体轮廓特征的性别识别, 目前只有相对比较少的人进行这个方面的研究。在这一章中, 对步态识别进行了全面的实验性研究, 获得了鼓舞人心的实验结果, 并获得了一些有用的结论。

在实验中, 使用 CASIA 步态数据库数据集 B 中的 62 个人 (31 男, 31 女) 获得的最高识别率为 94.89%。另外还发现, 对于噪声的影响, 基于轮廓的特征要比基于轮廓线的特征更加鲁棒。在所有实验中使用的分类器中, 支持向量机要优于 kNN 和贝叶斯分类器, 其原因是支持向量机具有比较好的学习能力和推广能力。特征采用步态能量图的 PCA 系数, 分类器采用支持向量机, 这一组合可以获得比较高的识别率。如果从步态能量图的 PCA 系数中提取最佳分类特征, 可以获得比较高的识别率, 而且计算代价比较小。

表 6.11: 使用线性模型合成的特征进行性别识别的识别率 (SVM 分类器)

CCR (%)	测试集角度 θ_p (正常行走)											
	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	
注册集角度 θ_r (正常行走)	0°	91.40	88.98	91.13	93.55	91.67	89.25	91.40	90.32	91.13	90.32	92.47
	18°	85.48	92.47	92.47	93.82	92.47	91.13	90.32	90.86	92.74	91.13	89.25
	36°	90.32	88.17	94.89	92.74	95.97	96.51	94.62	91.94	92.20	93.82	94.09
	54°	93.01	89.52	90.59	96.51	93.82	93.01	94.09	95.97	95.97	95.43	97.04
	72°	95.16	95.70	92.74	93.55	93.55	93.55	91.67	93.55	96.51	96.24	94.89
	90°	94.62	93.28	95.70	93.01	93.28	93.55	92.20	90.86	92.20	94.89	96.24
	108°	93.55	94.62	93.01	94.62	93.01	93.01	95.97	91.13	86.29	94.09	93.01
	126°	89.78	93.55	93.82	94.35	94.09	88.71	92.74	93.55	82.80	84.95	88.17
	144°	91.94	88.44	88.71	90.59	93.82	93.28	89.25	91.13	90.32	86.02	84.95
	162°	90.32	88.98	90.32	91.94	89.78	90.59	90.59	89.78	89.52	90.05	86.02
	180°	88.44	90.86	90.86	90.86	90.59	90.59	91.13	88.71	88.17	89.52	89.25

除了对各种步态特征和特征进行分析外,还分析了身体的不同区域的区分性别的能力,得出颈后部,背部和胸部为三个区分能力比较强的区域。另外,本章还研究了跨人种的性别识别,使用亚洲人作为训练数据,欧洲人作为测试数据,获得了 90.00% 的识别率;使用欧洲人作为训练数据,亚洲人作为测试数据,获得了 85.22% 的识别率。这些实验说明即使在跨人种情况下,同一性别的步态还是更相似一些。

由人观察轮廓判断性别的实验,以及识别算法的多视角实验都得出了一个相同的结论,即与正面和背面步态特征相比,侧面的步态特征更有利于性别识别。除此之外,本章还进行了视角无关的性别识别研究,取得了令人满意的结果。

另外,在目前的实验中,观察者和识别算法都获得了相当的识别率,在未来,基于步态的性别识别的识别率是否还有进一步提高空间?在性别识别方面,识别算法是否能够远超越人的感知?这些问题值得进一步研究。

表 6.12: 视角变化时性别识别性能评价标准

i	σ_i^A	
	GEI+PCA	GEI+PCA+Linear
1	0.963	0.973
2	0.828	0.973
3	0.655	1.003
4	0.654	1.007
5	0.663	1.000
6	0.643	0.994
7	0.647	0.987
8	0.785	0.984
9	0.860	0.971
10	0.826	0.974

第七章 结论与展望

7.1 工作总结

步态识别是根据人走路的姿势以及人体形状来识别人的身份，是一个非常具有挑战性的问题。步态识别的难点之一是步态特征具有很强的不稳定性。特别是在视角变化、衣着变化、携带物品、光照变化、时间变化等情况下，鲁棒的提取具有区分能力的特征变得很困难。本文针对步态识别中的视角变化这一常见变化，进行了一系列研究，并对衣着和携带物品变化、以及性别识别也有所涉及。本文的主要贡献如下：

⊖步态识别算法评价框架

目前步态识别研究尚不成熟，步态识别中的很多问题没有研究清楚，且缺乏统一评价标准。针对这一现状，我们提出了一个步态识别算法评价框架。这个评价框架包括一个大规模的步态数据库（CASIA 步态数据库数据集 B），三组实验和一系列评价指标。CASIA 步态数据库数据集 B 包含 124 个人，11 个视角，3 种行走条件，共有 13,640 个视频序列，目前是国际上规模最大的公开步态数据库。该数据库已经对国内外科研同仁共享，促进了步态识别研究的发展。这个框架包含三个实验集（实验集 A，B 和 C），每个实验集包含 121 个实验。实验集 A 中包含视角变化，实验集 B 中包含视角变化和衣着变化，实验集 C 中包含视角变化和携带物品变化。为了定量的比较算法的鲁棒性，我们设计了一组评价标准 Δ 和 σ_i ($i = 1, 2, \dots, 10$) 来量化算法对各种变化的鲁棒性。该步态识别算法评价框架可以用于指导步态算法的设计。

⊖对视角变化与步态识别性能关系建模和分析

哪个视角最适合用于步态识别？视角变化是如何影响识别率的？这两个问题是在研究步态识别中两个需要回答的问题。本文对视角变化与步态识别性能的关系进行了建模和分析，发现各个角度的识别率无明显差别，但是侧面视角对视角变化具有比较好的鲁棒性，得出侧面视角是步态识别中的最佳视角的结论。另外，还建立起视角与识别率之间的关系模型，用我们提出量化的模型，只需要指定几个参数就可以预测任意注册角度和测试角度组合的识别率。这些工作对于发掘步态识别性能与视角变化的关系具有一定意义。

③ 视角无关的步态识别

目前虽然有大量的步态识别算法被提出，但是这些算法大部分都是基于特定视角的，对视角变化不鲁棒，这大大的限制了步态识别的应用。文中分别提出了线性和非线性的模型，将一个视角的步态数据合成为该人在另一个视角的数据，来解决测试数据的视角跟注册数据的视角不一致问题。线性模型为一个线性变换，将测试角度采集的特征从原来的特征空间映射到一个新的空间，使得特征在新空间的表达跟注册角度的特征相近。这个线性变换可以通过最小二乘法求解得到。对于非线性模型，我们使用的是神经网络。用神经网络将测试角度采集的特征拟合到注册角度的特征。大量的实验证明了这两种模型对于解决视角变化非常有效，特别是角度变化不大于 36° 时。

④ 性别识别

在视觉监控中，如果某个行人的数据没有被注册，那就不能根据步态识别出该人的身份。在这个时候，依然可以根据步态识别出行人的性别、年龄和体形等信息，作为步态识别的补充，提高监控系统的智能水平。另外，性别识别可以作为一种粗分类手段，提高步态数据库的检索速度和提高步态识别的准确度。本文中对基于步态的性别识别进行了较全面的研究。首先，由人观察运动的人体轮廓来判断性别，最高获得了 95.48% 的识别率。这个实验验证了一个结论，即可以由步态判断出性别。在随后的研究中，使用了多种特征和多种分类器进行性别识别研究，并比较各种特征以及分类器性能。在 110 人的数据库上取得了 91.36% 的最高识别率。另外，在特征分析中还发现，头颈后部、背部和胸部是最具有区分性别能力的区域；正面和背面特征区分性别的能力要弱于侧面特征；用于步态识别的视角无关特征提取模型同样可以用于性别识别，在视角变化情况下能够明显提高识别率。除此之外，还进行了跨人种的性别识别。跨人种的性别识别获得了比较高的识别率，证明了不同人种的相同性别的人的步态是相似的。

综上所述，本文提出了一套各种变化因素的评估框架，用于评估不同的步态识别算法。并对步态识别中常见的视角变化进行了分析和建模，提出了解决方法。此外，还对基于步态的性别这一新方向进行了探索，取得了一定的研究成果。

7.2 未来发展趋势

最近十年，在研究者们努力下，步态识别获得了很大的发展。但是步态

识别中还有许多尚未解决的问题和不成熟的地方。为了开发出成熟稳定的步态系统，尚有很多问题需要进一步的研究。

7.2.1 行人检测与分割

步态识别的第一步是从视频中检测到运动区域，然后判断是否是行人，如果是行人，则将人的轮廓从背景图像中分割出。这一步是非常关键的一步，因为轮廓质量将直接决定着后面每一步的难度，以及最终的识别率。目前的大部分步态识别都假设背景比较简单，且场景中只有一个人在行走。这样的假设在实际应用中是难以达到的。

运动物体的检测是图像处理中的经典难题，因为光照、阴影、晃动的树、天气等因素会使得鲁棒地检测运动区域变得很困难，至今尚没有一个通用的方法。在未来可以将行人检测的方法和分割的方法相结合，利用已知的人体运动知识，将行人与其他运动物体（如汽车等）区分开，然后根据人体的先验的形状信息，来提高人体的分割质量。

7.2.2 关键影响因素的评价

步态作为一种行为特征，容易受到各种因素的影响，如角度、衣着、携带物品、步速、拍摄距离、光照、时间等。在开发出实用的步态识别系统之前，这些常见的关键影响因素都需要用科学的方法进行深入分析和评价。

对于视角变化，在前面的章节中已经进行了分析。但是在实际应用中，视角变化可能更加复杂。如为了获得更大的监控范围，监控系统中的摄像机一般悬挂在高处，以俯视的角度进行拍摄。这样不仅有水平方向的视角变化，而且还会有垂直方向上的视角变化。这些都需要进行详细的研究，分析出视角变化对识别的影响。衣着也是一个很重要的因素。夏天的裙子，冬天的大衣，帽子等这些都无一不对步态特征的提取有巨大的挑战。携带物品也是常见现象，而且现有的方法难以将携带的物品跟人体区分开。检测出携带的物品，并将该物品从人身上去除是一个值得尝试的方向。另外，一些生物特征，如指纹和虹膜，可以保持一生不变。但是步态却会随着年龄的变化、体重的变化等而发生变化，如何应对这一变化，也是一个极具挑战性的问题。

7.2.3 融合

为了提高识别率,可以采用融合的方法。此处的融合有两个层面的融合,一个是各种步态特征的融合,另一个是步态特征跟其他生物特征的融合。

步态特征既包含动态特征,又包含静态特征。动态特征有人体关节的角度变化,人体部位的运动曲线等,静态特征有身高,身体各部位的形状等。如何合理的在各个层面(特征层,匹配分数层等)融合这些特征,是值得研究的方向。

步态是一种弱生物特征,在实际应用中,可以跟其他生物特征进行融合,以提高识别率,或者提高人机交互功能。在目前流行的各种生物特征中,人脸特征跟步态特征一样,可以从远距离获取,具有非接触和非侵犯的特点,比较适合跟步态进行融合。这两种特征可以融合在一起,来提高识别的准确性。

7.2.4 基于步态的视频检索

随着计算机技术的发展,出现了海量的视频数据。如何管理这些数据成为了一个越来越重要的问题,而基于内容的视频检索技术是解决这一问题的方法。基于内容的视频检索有着广泛的应用和广阔的发展前景,是当前多媒体研究领域的热点。视频检索需要对视频内容进行理解,而步态分析与识别可以用来帮助理解视频,如视频中出现了哪个人物,以及人物的行为等。

基于步态的视频检索也可以用于视觉监控系统。在实际中,经常有这样的事情发生:保安人员会遇到一些可疑的人,但由于保安人员无执法权力,只能简单的询问可疑人。在无证据的情况下,也不可能一旦怀疑某人,就立刻报警。基于视频的检索系统可以用来检查可疑人是否真的做了违法的事情,一旦发现可疑人,立刻将该人的步态输入系统,通过步态特征检索出该人去过哪些地方,做过什么事情,来确认该人是否真的可疑,这样可以缩短整个安全保卫系统的反应时间。这样的应用一般只需要检索当天的视频数据,衣着和体形等一些跟时间相关的影响因素不需要考虑,所以在技术上是可行的。

7.2.5 基于步态的个人信息获取

在目前的商业社会,可以说大量客户的个人信息意味着财富,因为可以从中间挖掘出大量有用信息。如一个大商场想了解哪些年龄段的人对某个商品感兴趣,不同年龄和性别的人的购物行为等信息,可是这样的信息是很难获取的。一方面有技术方面的原因,另外一个更大的障碍是隐私方面的问题,因为隐私保

护越来越被重视，很多国家都有各种各样的法规来保护个人隐私，所以这类信息获取起来非常困难。

步态特征可以用来获取行人的信息，如年龄和性别，而不会引起隐私问题。因为步态分析系统获取的年龄和性别只是计算机估计出的信息，而不是准确信息，同时这些信息并没有跟这个人的姓名等资料联系起来。这就好像一个专家站在那儿，快速的估计并记录周围每一个人的年龄和性别，甚至职业。

在技术上，根据步态来获取信息是可行的。本文前面的研究证明了可以根据步态来识别性别。而年龄，也可从步态识别出。因为老态龙钟的老人的步态跟精力旺盛的年轻人毕竟有些不同。同时，人脸可以跟步态融合，更准确的获取性别、年龄等信息。

总之，在未来，随着计算机视觉技术的发展，在各种需求的刺激下，步态分析与识别必定会有更大的发展。同时步态分析与识别也会反过来促进计算机视觉技术的发展。

参考文献

- [1] Web site of the Department of Mechanical and Materials Engineering, Queen's University, Canada. <http://me.queensu.ca/people/deluzio/GaitAnalysis.php>.
- [2] Liang Wang, Huazhong Ning, Tieniu Tan, and Weiming Hu. Fusion of static and dynamic body biometrics for gait recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(2):149–158, Feb. 2004.
- [3] G. Shakhnarovich, L. Lee, and T. Darrell. Integrated face and gait recognition from multiple views. In *Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages I439–I446. Hawaii, USA, Dec. 2001.
- [4] Amos Y. Johnson and Aaron F. Bobick. A multi-view method for gait recognition using static body parameters. In *Proc. of 3rd International Conference on Audio and Video Based Biometric Person Authentication*, pages 301–311. Halmstad, Sweden, Jun. 2001.
- [5] Amit Kale, Amit K. Roy Chowdhury, and Rama Chellappa. Towards a view invariant gait recognition algorithm. In *Proc. of IEEE Conference on Advanced Video and Signal Based Surveillance*, pages 143 – 150, July 2003.
- [6] Joseph P. et al. Pickett, editor. *The American Heritage Dictionary of the English Language*. Houghton Mifflin Compan, fourth edition, 2000.
- [7] 王亮. 步态分析与识别: [博士论文]. 北京: 中国科学院自动化研究所. 2003年.
- [8] Nikolaos V. Boulgouris, Dimitrios Hatzinakos, and Konstantinos N. Plataniotis. Gait recognition: a challenging signal processing technology for biometric identification. *IEEE Signal Processing Magazine*, 22(6):78 – 90, November 2005.

-
- [9] D. M. Gavrilu and L. S. Davis. 3-d model-based tracking of humans in action: a multi-view approach. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, page 73, Washington, DC, USA, 1996. IEEE Computer Society.
- [10] Dariu Gavrilu. Pedestrian detection from a moving vehicle. In *Proceedings of the 6th European Conference on Computer Vision-Part II*, pages 37–49, 2000.
- [11] Congxia Dai, Yunfei Zheng, and Xin Li. Pedestrian detection and tracking in infrared imagery using shape and appearance. *Computer Vision and Image Understanding*, 106(2-3):288–299, 2007.
- [12] Huazhong Ning, Tieniu Tan, Liang Wang, and Weiming Hu. Automatic extraction and description of human gait models for recognition purposes. *Image and Vision Computing*, 22(5):429–441, May 2004.
- [13] R. Rashid. Towards a system for the interpretation of moving light displays. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2(6):574–581, 1980.
- [14] Aaron F. Bobick and James W. Davis. The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):257–267, 2001.
- [15] Ying Wang, Kaiqi Huang, and Tieniu Tan. Abnormal activity recognition in office based on r transform. In *Proceedings of the IEEE International Conference on Image Processing*, pages I:341–344, San Antonio, Texas, USA, Sep. 2007. IEEE Computer Society.
- [16] M. P. Murray, A. B. Drought, and R. C. Kory. Walking patterns of normal men. *Journal of Bone and Joint Surgery*, 46-A(2):335–360, 1964.
- [17] M. P. Murray. Gait as a total pattern of movement. *American Journal of Physical Medicine*, 46(1):290–332, Feb. 1967.

- [18] H. M. Lakany, G. M. Haycs, M. Hazlewood, and S. J. Hillman. Human walking: tracking and analysis. In *Proc. of IEE Colloquium on Motion Analysis and Tracking*, pages 5/1–5/14, 1999.
- [19] Chris Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *Proceedings of the IEEE Computer Science Conference on Computer Vision and Pattern Recognition (CVPR-99)*, pages 246–252, Los Alamitos, Jun. 1999. IEEE.
- [20] L. Lee and W. E. L. Grimson. Gait analysis for recognition and classification. In *Proc. of 5th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 155–162. Washington D.C., USA, May 2002.
- [21] Raquel Urtasun and Pascal Fua. 3d tracking for gait characterization and recognition. In *Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 17–22. Seoul, Korea, May 2004.
- [22] David K. Wagg and Mark S. Nixon. On automated model-based extraction and analysis of gait. In *Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 11–16. Seoul, Korea, May 2004.
- [23] C. Y. Yam, M. S. Nixon, and J. N. Carter. On the relationship of human walking and running: automatic person identification by gait. In *Proc. of International Conference on Pattern Recognition*, pages 287–290. Quebec, Canada, 2002.
- [24] Rong Zhang, Christian Vogler, and Dimitris Metaxas. Human gait recognition at sagittal plane. *Image Vision Computing*, 25(3):321–330, 2007.
- [25] Sudeep Sarkar, P. Jonathon Phillips, Zongyi Liu, Isidro Robledo Vega, Patrick Grother, and Kevin W. Bowyer. The humanoid gait challenge problem: Data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):162–177, Feb. 2005.

-
- [26] Amit Kale, Aravind Sundaresan, A. N. Rajagopalan, Naresh P. Cuntoor, Amit K. roy Chowdhury, Volker Krüger, and Rama Chellappa. Identification of humans using gait. *IEEE Transactions on Image Processing*, 13(9):1163–1173, Sep. 2004.
- [27] Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1505–1518, December 2003.
- [28] Shiqi Yu, Liang Wang, Weiming Hu, and Tieniu Tan. Gait analysis for human identification in frequency domain. In *Proc. of the 3rd International Conference on Image and Graphics*, pages 282–285. Hong Kong, China, December 2004.
- [29] Ju Han and Bir Bhanu. Statistical feature fusion for gait-based human recognition. In *Proc. of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages II842–II847. Washington DC, USA, June 2004.
- [30] A. Kale, A. Rajagopalan, N. Cuntoor, and V. Krüger. Gait-based recognition of humans using continuous HMMs. In *Proc. of 5th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 336–341. Washington D.C.,USA, May 2002.
- [31] Guoying Zhao, Rui Chen, Guoyi Liu, and Hua Li. Amplitude spectrum-based gait recognition. In *Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 23–28. Seoul, Korea, May 2004.
- [32] Rocío Díaz de León and Luis Enrique Sucar. Human silhouette recognition with Fourier descriptors. In *Proc. of 15th International Conference on Pattern Recognition*, pages 709–712, Sep 2000.

- [33] Daoliang Tan, Kaiqi Huang, Shiqi Yu, and Tieniu Tan. Uniprojective feature for gait recognition. In *Proc. of The 2nd International Conference on Biometrics*, pages 673–682, Seoul, Korea, August 2007.
- [34] Daoliang Tan, Shiqi Yu, Kaiqi Huang, and Tieniu Tan. Walker recognition without gait cycle estimation. In *Proc. of The 2nd International Conference on Biometrics*, pages 222–231, Seoul, Korea, August 2007.
- [35] Liang Wang, Tieniu Tan, Weiming Hu, and Huazhong Ning. Automatic gait recognition based on statistical shape analysis. *IEEE Transactions on Image Processing*, 12(9):1120–1131, Sep. 2003.
- [36] Shiqi Yu, Daoliang Tan, Kaiqi Huang, and Tieniu Tan. Reducing the effect of noise on human contour in gait recognition. In *Proc. of The 2nd International Conference on Biometrics*, pages 338–346, Seoul, Korea, August 2007.
- [37] Rawesak Tanawongsuwan and Aaron Bobick. Gait recognition from time-normalized joint-angle trajectories in the walking plane. In *Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages Vol. 2, 726–731. Kauai Marriott, Hawaii, USA, December 2001.
- [38] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. ISBN: 0-471-05669-3. Wiley, second edition, 2001.
- [39] Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, New York, NY, USA, 1992. ACM Press.
- [40] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. A practical guide to support vector classification.
- [41] Chiraz BenAbdelkader, Ross Cutler, and Larry S. Davis. View-invariant estimation of height and stride for gait recognition. In *Proceedings of the*

-
- ECCV 2002 Workshop Copenhagen on Biometric Authentication*, pages 155–167. Copenhagen, Denmark, June 2002.
- [42] James J. Little and Jeffrey E. Boyd. Recognizing people by their gait: the shape of motion. *Videre: Journal of Computer Vision Research*, 1(2):2–32, 1998.
- [43] James B. Hayfron-Acquah, Mark S. Nixon, and John N. Carter. Automatic gait recognition by symmetry analysis. In *Proc. of 3rd International Conference on Audio and Video Based Biometric Person Authentication*, pages 272–277. Halmstad, Sweden, Jun. 2001.
- [44] Human ID at CMU. <http://www.hid.ri.cmu.edu>.
- [45] Ralph Gross and Jianbo Shi. The CMU Motion of Body (MoBo) Database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, June 2001.
- [46] Center for Biometrics and Chinese Academy of Sciences Security Research, Institute of Automation. <http://www.cbsr.ia.ac.cn>.
- [47] Automatic Gait Recognition for Human ID at a Distance at Soton. <http://www.gait.ecs.soton.ac.uk>.
- [48] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter. On a large sequence-based human gait database. In *Proc. of 4th International Conference on Recent Advances in Soft Computing*, pages 66–71. Nottingham, UK, 2002.
- [49] Human ID Challenge Problem at USF. <http://figment.csee.usf.edu/GaitBaseline>.
- [50] Human Identification at a Distance at Gatech. <http://www.cc.gatech.edu/cpl/projects/hid>.

- [51] Shiqi Yu, Daoliang Tan, and Tieniu Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *Proc. of the 18th International Conference on Pattern Recognition (ICPR06)*, pages 441–444. Hong Kong, China, August 2006.
- [52] The Mathworks Inc. *Curve Fitting Toolbox User's Guide, Version 1*. <http://www.mathworks.com/>.
- [53] John Shawe-Taylor and Nello Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [54] Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
- [55] John A. Hertz, Richard G. Palmer, and Anders Krogh. *Introduction to the Theory of Neural Computation*. Addison-Wesley, 1991.
- [56] Martin Riedmiller and Heinrich Braun. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *Proc. of the IEEE International Conference on Neural Networks*, pages 586–591, San Francisco, CA, 1993.
- [57] C. M. Scandrett, C. J. Solomon, and S.J. Gibson. Towards a semi-automatic method for the statistically rigorous ageing of the human face. *IEE Proceedings of Vision, Image and Signal Processing*, 153(5):639–649, Oct. 2006.
- [58] Yun Fu, Ye Xu, and Thomas S. Huang. Estimating human age by manifold analysis of face pictures and regression on aging features. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 1383–1386, Beijing, China, July 2007.
- [59] Markus Iseli, Yen-Liang Shue, and Abeer Alwan. Age-and gender-dependent analysis of voice source characteristics. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, pages I:389–392, Toulouse, France, May 2006.

-
- [60] M. Gerosaa, S. Lee, D. Giuliani, and S. Narayanan. Analyzing children's speech: An acoustic study of consonants and consonant-vowel transition. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, pages I:393–396, Toulouse, France, May 2006.
- [61] Amit Jain and Jeffrey Huang. Integrating independent components and linear discriminant analysis for gender classification. In *Proc. of 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 159–163. IEEE Computer Society, 2004.
- [62] Baback Moghaddam and Ming-Hsuan Yang. Learning gender with support faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):707–711, 2002.
- [63] Yunus Saatci and Christopher Town. Cascaded classification of gender and facial expression using active appearance models. In *Proc. of 7th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 393–400. IEEE Computer Society, 2006.
- [64] H. Harb and Liming Chen. Gender identification using a general audio classifier. In *Proceedings of the 2003 International Conference on Multimedia and Expo*, pages 733–736, Washington, DC, USA, 2003. IEEE Computer Society.
- [65] Stefan Slomka and Sridha Sridharan. Automatic gender identification optimised for language independence. In *1997 IEEE TENCON - Speech and Image Technologies for Computing and Telecommunications*, pages 154–148, 1997.
- [66] L. T. Kozlowski and J. E. Cutting. Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, 21:575–580, 1977.
- [67] James W. Davis and Hui Gao. Gender recognition from walking movements using adaptive three-mode PCA. In *Proceedings of the 2004 Conference on*

- Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 1*, page 9, Washington, DC, USA, 2004. IEEE Computer Society.
- [68] Nikolaus F. Troje. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2(5):371–387, 9 2002.
- [69] George Mather and Linda Murdoch. Gender discrimination in biological motion displays based on dynamic cues. In *Proc. of Biological Sciences*, volume 258 of *B*, pages 273–279. the Royal Society of London, 1994.
- [70] Rachel McDonnell, Sophie Joerg, Jessica K. Hodgins, Fiona Newell, and Carol O’Sullivan. Virtual shapers & movers: Form and motion affect sex perception. In *Proc. of Symposium on Applied Perception in Graphics and Visualization*, Tübingen, Germany, July 2007.
- [71] Jungwon Ryu and Sung-Bae Cho. Gender recognition of human behaviors using neural ensembles. In *Proceedings of International Joint Conference on Neural Networks*, volume 1, pages 571–576, Washington, DC, USA, 2001.
- [72] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. W. Bowyer. The gait identification challenge problem: Data sets and baseline algorithm. In *Proc. of International Conference on Pattern Recognition*, pages 385–388. Quebec, Canada, August 2002.
- [73] Daniel L. Swets and John (Juyang) Weng. Using discriminant eigenfeatures for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):831–836, 1996.
- [74] R. O. Duda and P. E. Hart. *Pattern Classification*. John Wiley and Sons, 2000.
- [75] Morton B. Brown and Alan B. Forsythe. The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16(1):129–132, Feb. 1974.

- [76] Xianchao Qiu, Zhenan Sun, and Tieniu Tan. Global texture analysis of iris images for ethnic classification. In *Proceedings of International Conference on Biometrics 2006*, pages 411–418, Hong Kong, China, January 2006.

发表文章目录

- [1] **Shiqi Yu**, Tieniu Tan, Daoliang Tan and Kaiqi Huang. A Study on Gait Based Gender Classification. *Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [2] **Shiqi Yu**, Daoliang Tan, Kaiqi Huang, and Tieniu Tan. Reducing the effect of noise on human contour in gait recognition. In Proc. of 2nd IAPR International Conference on Biometrics. Seoul, Korea, Aug 2007.
- [3] **Shiqi Yu**, Daoliang Tan, and Tieniu Tan. A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. In Proc. of the 18th International Conference on Pattern Recognition (ICPR06). Hong Kong, China. August 2006.
- [4] **Shiqi Yu**, Daoliang Tan, and Tieniu Tan. Modelling the Effect of View Variation on Appearance-based Gait Recognition. In Proc. of the 7th Asian Conference on Computer Vision (ACCV06). Hyderabad, India. Jan. 2006.
- [5] **Shiqi Yu**, Liang Wang, Weiming Hu, and Tieniu Tan. Gait Analysis for Human Identification in Frequency Domain. In Proc. of 3rd International Conference on Image and Graphics, pages 282-285. Hong Kong, China, Dec. 2004.
- [6] Daoliang Tan, **Shiqi Yu**, Kaiqi Huang, and Tieniu Tan. Walker recognition without gait cycle estimation. In Proc. of The 2nd International Conference on Biometrics, pages 222 - 231, Seoul, Korea, August 2007.
- [7] Yuan Wang, **Shiqi Yu**, Yunhong Wang and Tieniu Tan. Gait Recognition Based on Fusion of Multi-view Gait Sequences. In Proc. of the International Conference on Biometrics 2006, pages 605-611. Hong Kong, China. Jan. 2006.

-
- [8] Daoliang Tan, Kaiqi Huang, **Shiqi Yu**, and Tieniu Tan. Uniprojective feature for gait recognition. In Proc. of The 2nd International Conference on Biometrics, pages 673 – 682, Seoul, Korea, August 2007.
- [9] Daoliang Tan, Kaiqi Huang, **Shiqi Yu**, and Tieniu Tan. Orthogonal diagonal projections for gait recognition. In Proc. of 2007 IEEE International Conference on Image Processing, pages I/337-I/340, San Antonio, Texas, USA.
- [10] Daoliang Tan, Tieniu Tan, and **Shiqi Yu**. Efficient Night Gait Recognition Based on Template Matching. In Proc. of the 18th International Conference on Pattern Recognition (ICPR06). Hong Kong, China. August 2006.
- [11] 谭铁牛, 黄凯奇, 于仕琪. 一种基于步态的性别识别方法. 发明专利, 专利申请号: 200710098465.5, 公开(公告)号: IB071640, 公开(公告)日: 2007年4月18日。
- [12] 刘瑞祯, 于仕琪. OpenCV 教程—基础篇. 北京: 北京航空航天大学出版社, 2007年6月, 书号: 978-7-81124-035-1.

致 谢

值此论文完成之际，衷心感谢我的导师谭铁牛研究员在我攻读博士期间在科研、生活和其他各个方面给予的无微不至的关怀和孜孜不倦的教诲。在数点我这些年的收获和成熟之余，我非常庆幸自己选择中科院自动化所，师从谭铁牛研究员。是谭老师的引导，使得我对科学研究有所领悟，完成本文。论文从开始的选题到最后的修改都是在谭老师的悉心指导下完成的。谭老师渊博的学识和敏锐的学术眼光拓宽了我的视野；他对事业的执著追求和严谨的治学态度使我深受影响，必将使我受益终生。

感谢模式识别国家重点实验室的各位老师和同学。特别感谢胡卫明研究员，他的帮助使我很快地进入科研工作，少走了弯路。感谢王蕴红教授、黄凯奇副研究员和孙哲南助理研究员的指导和帮助。感谢现在已经在澳大利亚墨尔本大学的王亮，他给我提供了大量无私的帮助。感谢魏育成、楼建光、马力、柳崎峰、杨金锋、徐成华、薛文芳、黄俊洲、刘威、傅周宇、于昆、单彩峰、胡敏、王彦情、林翰青、谢丹、于小亿、崔家礼、李江伟、王谦、信伦、王莹、舒丽霞、陈文海、辛乐、何鹏、王恒奎、郭磊、邱显超、张彰、王媛、肖雪娟、范伟、覃道亮、王亮生、郝瑛、刘舟、蔡莹皓、韦卓识、钟诚、李邦昱、谭剑波、张兆翔、韩宇飞、董文博、刘亮、黄永刚、陈小川、李真、何召峰、董晶、黄永祯、王时全、何倩等同学，能和这些志同道合的朋友们共事，对我而言是一段美好的人生经历。

衷心感谢综合办公室的连国臻、赵薇、马丽霞三位老师以及研究生部的邸凌、李磊、卜树云、毛磊、曹娟等老师。我们的生活、学习和科研离不开他们的行政和后勤工作。

我要感谢我的父母，在我二十余年的求学生涯中，是他们一直在背后给予我支持和力量，希望我的成绩能够给他们带来快乐。我还要感谢我的女友，是她的爱给了我克服困难的勇气和积极进取的热情。

于仕琪

2007年10月